View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Northumbria Research Link

Northumbria Research Link

Citation: Aziz, Hayder, Gao, James, Maropoulos, Paul and Cheung, Wai Ming (2003) Open Standard, Open Source and Peer to Peer Methods for Collaborative Product Development and Knowledge Management. In: 12th Symposium of Product Data Technology EUROPE, 25th - 27th September 2003, Manchester.

URL:

This version was downloaded from Northumbria Research Link: http://nrl.northumbria.ac.uk/29275/

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: http://nrl.northumbria.ac.uk/policies.html

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

www.northumbria.ac.uk/nrl

northumbria UNIVERSITY NEWCASTLE

Open Standard, Open Source and Peer to Peer Methods for Collaborative Product Development and Knowledge Management

Hayder Aziz, James Gao, Enterprise Integration, Cranfield University, UK.
Paul G Maropoulos, Wai M. Cheung, School of Engineering, University of Durham, UK.

1. Introduction

Tools such as product data management (PDM) and its offspring product lifecycle management (PLM) enable collaboration within and between enterprises [1]. Large enterprises have invariably been the target of software vendors for development of such tools, resulting in large 1entralized applications. These are beyond the means of small to medium enterprises (SME). Even after these efforts had been made, large enterprises face numerous difficulties with PLM [2]. Firstly, enterprises evolve, and an evolving enterprise needs an evolving data management system. With large applications, such configuration changes have to be made at the server level by dedicated staff. The second problem arises when enterprises wish to collaborate with a large number of suppliers and original equipment manufacturer (OEM) customers. Current applications enable collaboration using business-to-business (B2B) protocols. However, these do not take into account that disparate enterprises do not have unitary data models or workflows. This is a strong factor in reducing the abilities of large enterprises to participate in collaborative projects.

The enterprises' prime asset, its knowledge, is not managed coherently thus perpetuating an 'invisible' limit on the company's knowledge, based on the impulse of knowledge workers' recollections of previous experience. In addition, the problems for inter-enterprise collaboration are also inherent when enterprises with different level of detail and different nomenclature collaborate. SMEs have up to now been left out of developments in PDM/PLM,and knowledge management in spite of forming the majority of the world's engineering community. This negatively affects the ability of SMEs to manage their knowledge, reuse existing expertise, collaborate with other SMEs and larger enterprises. This weakness affects not only the SMEs but also the large enterprises that SMEs work with. Smaller engineering enterprises lack the infrastructure and manpower for complex solutions.

This study has focused on the experiences of automotive and discrete machining companies in integrating their product 1entralized cycles with their supply chains and their customers in real-time, and also to enable the companies to make rapid appraisal and cost-estimation for their customers, from simple concept designs. The system has to be able to reuse the company's existing base of knowledge and to push the manufacturing knowledge higher up into the design chain to reduce the need for costly and time consuming reworks and engineering changes.

Based on the above set of problems facing industry, the authors have drawn up three core requirements to codify the needs of SMEs in managing their product and process information i.e.

(i) Enabling project managers and all knowledge workers to have access to the functionality and to create and manage knowledge within their domain according to the agreed nomenclature and ontological representation;
(ii) Information created has to be in a form that can be queried, reused and transformed into new representations through the use of rules and agents; and
(iii) Enabling the real-time collaboration between SMEs, and larger partners, by facilitating the fast and cost-less construction of Ves;

To test the requirements and try to meet the objectives the authors took an evolutionary approach by assessing the current technology and methods, and then constructing three new example applications. The as-is technology used was a web-based commercial PLM system. The three test scenarios constructed are: (i) Modification of the PLM system for flexible and 1entralized1 data model; (ii) Implementation of a functionally equivalent system based on open source tools and (iii) The implementation of a peer-to-peer (P2P) based system.

Lihui et al [3] wrote an extensive survey of collaborative design systems, and highlighted eight areas as having scope for development including: System architecture for web-based collaborative design, Collaborative Conceptual Design Modelling and Data Sharing, Conceptual Design Selection, Knowledge Management in Collaborative Environments. Beckett [4] discussed the topic of communication and

understanding in Ves between unfamiliar participants. He discussed the best tools and standards to apply in VE settings for collaboration and knowledge management and has an overview of various methodologies and applications. Camarinha-Matos [5] on a similar note, reviewed current trends in VE developments, and conclude that there is a need to develop a 2entralized framework for Ves, to enable 2entralized22, international collaboration and rapid deployment.

Kim [6] developed a 'Distributed open-intelligent PDM' system, which adopts ISO standard STEP, whilst offering standard PDM functions. A dynamic and flexible workflow model is implemented. This could greatly enhance the flexibility of the system. Goh [7] proposed the STEP workflow management facility as a rule based workflow for PDM that is compliant with the Workflow Management Coalition (WfMC) guidelines. As part of the proposal, an object oriented data model driven system to store STEP entities is described. Zha [8] proposed a STEP based application to manage the entire product lifecycle. The information that is not already defined in STEP is modelled in EXPRESS. The system is focused primarily on assembly mating features and does not consider the machining requirements of each component. Zhou[9] tried to solve the problems of functional design knowledge management within a platform neutral setting through the definition of STEP models with the addition of semantics and the use of Artificial Neural Networks to aid the selection process. The authors' implementation used express data models and is accessible through a web portal. Vasara [10] proposes ARACHNE, the adaptive network strategy to enable integration between 87 enterprises, the authors highlight the benefits of peer to peer networks to achieve synergy between collaborators. Their developed methodology called RosettaStone enabled many-to-many integration between enterprises using a three-tier architecture.

To overcome the above mentioned problems, a new methodology was devised. In order to achieve the management of knowledge and integration with downstream applications the STEP standard was adopted. Its advantages are that it is a mature and internationally agreed standard. However, it has the disadvantages of not being deployed widely. The main reason is the cost associated with compatible tools as well as the lack of heterogeneous collaboration tools on the web. For this reason the format used for sharing the data has been changed from STEP Part-21 to the semantic web formats to enable the seamless management and sharing of knowledge over the web and enable heterogeneous systems to query and infer knowledge from the system.

2. Knowledge Representation Methodology

The authors define knowledge as the semantically complete definition of a domain's information that is both machine readable and interpretable. Thus a knowledge base contains an ontology which defines the classes and their relationships, instances or objects that form the domain 'data', meta-data that constrains the data within a particular domain (transforming it into information), and Universal Resource Identifiers (URI) that allow the global identification and contextual interpretation of the information. In a document oriented PDM/PLM, word processor documents and some excel worksheets contain the company's knowledge assets and these are stored on the server with some 'meta-tags' to identify the basic purpose of the file. The same is true for the storage of the geometric models within the system, as each CAD file is stored with some tag information. Both documents A and B within the system contain unique knowledge (kn1, kn2 kn3 etc.). However, kn1 is stored in both documents. In this example, a user viewing or modifying document A or B will need to be instinctively aware of such data duplication and update both documents kn1 values. In most instances such 'simple' issues cause major problems for larger teams where the volume of documents contained in the vault is too large for individuals to be instinctively aware of it. Additionally the problem of finding knowledge, as opposed to documents, is intrinsic.

The solution provided for this problem is a user modifiable object-oriented ontology for managing all information in the product development process as distinct objects within the PLM systems and tied together with URIs. This enables the system to create reports that mimic the 'layout' of static documents but rely on a single source of up-to-date knowledge, eliminating duplication, and enabling the user to modify sub components and assemblies of geometric models without retrieving the complete model.

The resource description framework (RDF) has been used in this application as the format for storing the knowledge base, as opposed to eXtensible mark-up Language (XML). The reasons for this choice are the extra flexibility and 'machine-understandable' format of RDF graph triple model as opposed to the simple 'machine-readable' XML based mark-up vocabularies. In effect, any RDF-parser can derive the semantics and context from the URI and metadata attached to every instance. RDF has been modelled in this instance using the Protégé KBS and 2entralized22 was made with ontoviz and tgviz. The ontology was sub- classed from the base data model of the PLM system for flexibility, ease of deployment and to make use

of the lifecycle and workflow functions offered by the PLM systems. Integration with the p2p system was through a Protégé plugin.

The second component uses STEP AP-224 based automatic process planner by Sharma [11] to generate plans from concept designs. Aggregate process plans for assemblies are generated using Cheung's [12 – 16] process planner. The STEP AP-224 feature models are defined in Protégé and use the Java expert system shell (JESS) to define the rules and functions of the standard and the constraints these impose enable error-checking for the user during the definition of features. However, these constraints can only be applied to individual entities within the STEP model and cannot enforce any constraints between entities/features. Due to the size of the standard, only a subset of AP-224 has been translated. The definition of the Express (STEP) data types is also contained in the Clips interface. This is very flexible as it allows for the inclusion of STEP data types to other components within the ontology on a need basis without having to have any expertise in Express or any other programming language. The above enables the mixing of feature and meta-data information in the knowledge base, meaning that users can access the information stored in STEP models using queries and RDF parsers. This integration at low level between the geometric, feature and 'meta-data' within a single environment is intended to reduce repetition and errors, and also enables the reuse of all the data created during the conceptual design process.

3. Collaboration on Distributed Designs

One of the first problems of collaboration is trying to understand what the other says and means. The previous section showed the methods used to create the data models. This development is interrelated to the problem of collaboration and the authors overcame one of the perennial problems of collaborative design between enterprises of different sizes and complexity, that is, mapping between low and high content data models which results in irretrievable loss of information from the high data model. This problem cannot be overcome traditionally by creating a mapping from one data model to another. Instead the authors have sought to create a universal project oriented ontology that can be created, shared and used by all parties collaborating in an enterprise in real-time. This eliminates the problems faced when low-end suppliers collaborating with advanced enterprises face integration issues. Using the building-block ontologies and STEP standards, all enterprises, small and large, have access to enterprise features. With this problem out of the way, a number of mechanisms to establish the collaboration and manage the knowledge were investigated.

Business to Business (B2B) integration is the traditional method for companies to collaborate. For example, integrating 15 companies together on a one-to-many basis where a single repository manages the project knowledge and workflows would need 14 separate mappings. To empower the individual enterprises within the collaborative environment and 3entralized3 the system would require many-to-many i.e. 94 B2B integrations!

3.1 Application Service Providers

Application Service Providers (ASPs) can be set up in two ways, either by a large 'controlling' enterprise, or through independent third party hosting. These services intend to provide the same utilities as enterprise level systems but in a non-enterprise specific service. ASPs offer project and product data management vaults where the administrator can 3entraliz the third-party portal for their own use. Advantages include Reduced cost for the enterprise as maintenance and backup is delegated. Increased opportunities if customers and VE initiating enterprises seek out partners through the portal. This is a "Democratic" system where no one enterprise controls the server and data. It sets down de-facto standards for data exchange, to which other enterprises in the same domain will adhere to in order to join the network of enterprises.
There are of course some fundamental disadvantages to the use of ASPs for product development, and these include: The bandwidth and server bottleneck problem associated with 3entralized services. The security fears of intellectual property rights being compromised. The potential risks of downtime and data losses in an "uncontrollable" environment and the liability issues associated with it. The difficulty of creating direct interfaces from the enterprise system to the ASPs portal. The exact functionality required may not be available from the "generic" ASP.

There are already some ASPs operating in the automotive and aeronautical sector enabling supply companies to interact and bid openly for contracts with OEMs and then manage the project/product information on the portal. However due to the disadvantages highlighted above, the authors sought to find

a third way. Whilst traditional client-server systems can operate in a collaborative manner, for example over a LAN or Internet, they are not truly distributed as they are 4entralized.

3.2 Peer to Peer Systems

Peer to Peer (P2P) applications address the needs of de-centralised organisations to collaborate and share knowledge regardless of geographical location. The principle of P2P has been around for a long time, and is today implemented in a number of applications such as instant messaging and file-sharing (www.GNUtela.com). There are already a number of P2P PDM in existence. Primarily aimed at the lower end of the market. The two commercial applications are AutoManager workflow from Cyco (www.cyco.com) and Columbus from Oasys Software (http://www.oasys-software.com/). The latter is available for free, and aimed at AutoCAD users within the construction sector. However they are crude solutions relying on the underlying file system and adding some "meta tags" to files for version control. As an example of what can be achieved, Alibre is a P2P CAD/PDM and collaboration tool in one. It uses the STEP standard and combines low cost and fast configuration.

The advantages offered by P2P applications are (i) no single point of failure, the network is alive as long as one peer is on-line, (ii) distributed sharing of bandwidth storage and processing power, so the system becomes more powerful as more users attach, (iii) lower running cost due to the lack of servers or high bandwidth central nodes, as well as (iv) maintaining individual control of the shared knowledge. P2P groups can be used to create profiles of the peer, and also more importantly of the peer's list of contacts within different domains. These profiles can be used within the network to search for and assess people's competences, interests, and memberships of trusted groups, and can aid in the construction of new relationships based on commonalities and third party assessments. There have been a number of issues that reduce the performance of the system using pure P2P architecture. The lack of indexing and routing services in P2P degrades the peer discovery and query functions. In order to leverage the advantages of client/server systems with the independence and interoperability of P2P systems a hybrid system where “super peers” act as peers to the extended P2P network and as a server to the enterprise's internal peer network is used. In addition rendezvous peers can be assigned to manage some of the peer information assigned to particular peer nets or projects. This hybrid has been shown to have the best potential for high-performance de-centralised services.

3.3 Inter-enterprise Communication Architecture

Since no two enterprises are the same, the idea of using XML based messaging for inter-enterprise collaboration is not easy as the two company schemas have to be mapped to each other. However, a new methodology developed and applied in this instance has been to utilise the open standard and open source philosophies. In order to achieve the speedy interoperability a standard has to be set for basic messaging. Some like STEP PDM implement the complete information structure of the engineering enterprise. Others like JuxtaPose (JXTA) implement only the “messaging” components. The choice of standards to use depends on the level of “standardisation” that all enterprises can adhere to. In the view of the authors only a very low level subset of all enterprises can be “standardised”. This small subset should exist only as a medium to enable communication, identification and access control management. All other aspects of the collaborative data environment is enterprise specific (although constructed from a subset of ontological components).
This project created a two tier architecture for communications, speeding up of the integration process is further enhanced through the use of open standard semantic web enabled ontology formats for storing the information, plus the use of open-standards in addition to company ontologies to define those ontology components.

4. Implementations

4.1 Commercial PLM System

Windchill is a traditional document management tool, whilst it can store all manner of data and make revision controls. It does not however have an intelligent method of containing and persisting information in an object oriented format. To alleviate this, the functionality of the system has been extended to include the management of knowledge from Protégé. The integration with Protégé was made using the workflow

engine. Windchill is inherently centralised, overly complex to set up and requires a long period of time for customisation. The system's main strengths are in the workflow tools. However, even here there is the problem of lock-in. The workflow and any customisation carried out cannot be reused on another system. In addition windchill has an application layer for inter-enterprise communications, and inherent in its weakness is the B2B paradigm which has been elaborated previously. The RosettaNet standard implemented can decrease the amount of customisation needed. In addition, unlike the ontological 2-tier system, Windchill's data models are neither portable nor standardised.

The data model can be modified by an expert who has to model, program, compile, update the database and integrate the code into Windchill before it can be used. This does not provide for the flexibility and ease of use for modelling ontologies that a project manager needs. The system, implementation and running costs are very large and only practical if enterprises share it between each other over a long period of co-operation.

4.2 Open Standard and Open Source tools

Open source applications are widely deployed, and the engineering/manufacturing sector can benefit greatly from leveraging the available source code for easily customised applications. Two applications used in this project are the Protégé ontology editor and Sun's JXTA peer to peer protocol. They offer the best usability, portability and cost/performance capability within their niches, and when allied to a scalable open source database such as SAP-DB, can offer enterprise level performance for zero capital expenditure and low customisation cost. Moreover they free the customer from software vendors and the vagaries of obsolescence. Other open-source tools used in this work include OPEN-CASCADE (STEP modeller) and the openflow engine that forms the heart of the open source PLM.

In the application the bottom layer consists of the database MySQL. On top of this layer are the two open source gateways, Apache for serving static and PHP based web pages and the Tomcat servlet engine for Java server pages (JSP) based applications. These three layers (database, web server and application server) form the server side of the system. Figure 4.1 illustrates the three-tier architecture.

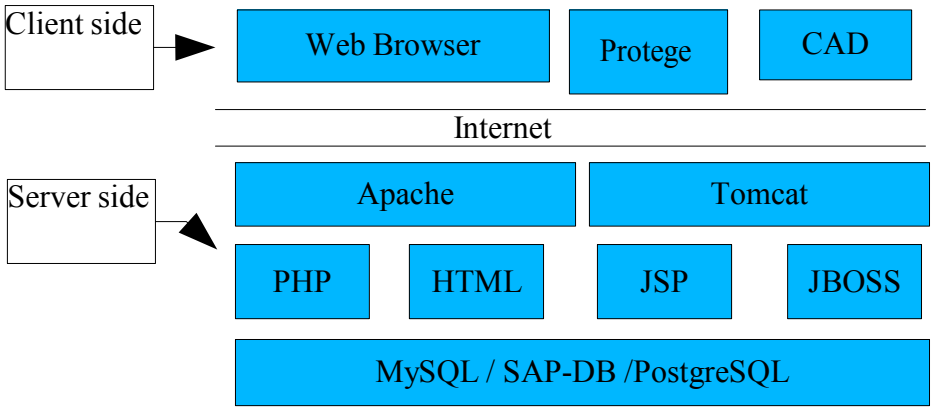

Figure 4.1, high-level view of open source server and user interface architecture

On the client side, the user has three main applications i.e. (1) the web browser through which interactions with the PLM system are carried out, (2) the Protégé Java applet that allows the user to query and manage the knowledge base and (3) a CAD system to enable the user to create and manipulate the STEP based models held in the PLM system.

The above systems are the evolution of product development management systems from document centric PLM systems to knowledge centric, intelligent systems of the future. However, it had been discovered during the course of the project that centralisation, by its very nature, is an inhibiting factor for inter-enterprise collaboration as the only methods available for collaboration in such an environment are B2B custom integration and the centralised web-enabled or ASP models. Portals in this instance would simply be classified as B2B integrations as each data source has to be separately mapped into the portal.

4.3 Peer to Peer Implementation

The back end consists of the open source SAP-DB database with the Java database connectivity (JDBC) connector to Protégé. Connectivity is achieved using an open source implementation of JXTA open standard P2P network protocol (www.jxta.org). The choice was made because JXTA implements a unique but anonymous identification mechanism for peers and for rendezvous peers. As well as “advertisement” implemented for all peers that give information about the peer to other peers. Rendezvous peers can act as managers for peer groups and store the peer advertisements for the group for distribution to other P2P networks. An extension to enable RDF queries and ontologies to be shared over P2P is used to share the knowledge base. Queries, project management of collaborative groups, group chat and instant messaging are readily implemented by the JXTA protocol. The systems’ settings enable enterprises and users without static addresses to collaborate using dynamic addressing, and this flexibility as well as the users’ ability to work off-line (that cannot be done with web based systems) empowers users in all possible network scenarios.

5. Example

The enterprise’s processes were split into lifecycle states and defined in workflow processes. Due to the complexity of the overall processes, only the conceptual design phase was modelled into the ontology and workflows. The sequence of processes for this phase are:

- Customer request for quotation is submitted including requirements and associated verification methods;
- Generation of specifications through analysis of earlier customer requests and testing against the new verification methods;
- Creating concept design options from the specifications by pattern matching against earlier concepts developed with similar specifications;and
- Selection of concept and utilising the verification methods;

User interaction through the interface as shown in Figure 5.1 was straightforward. Workflow sequences acted as widgets guiding the user through different configurations.

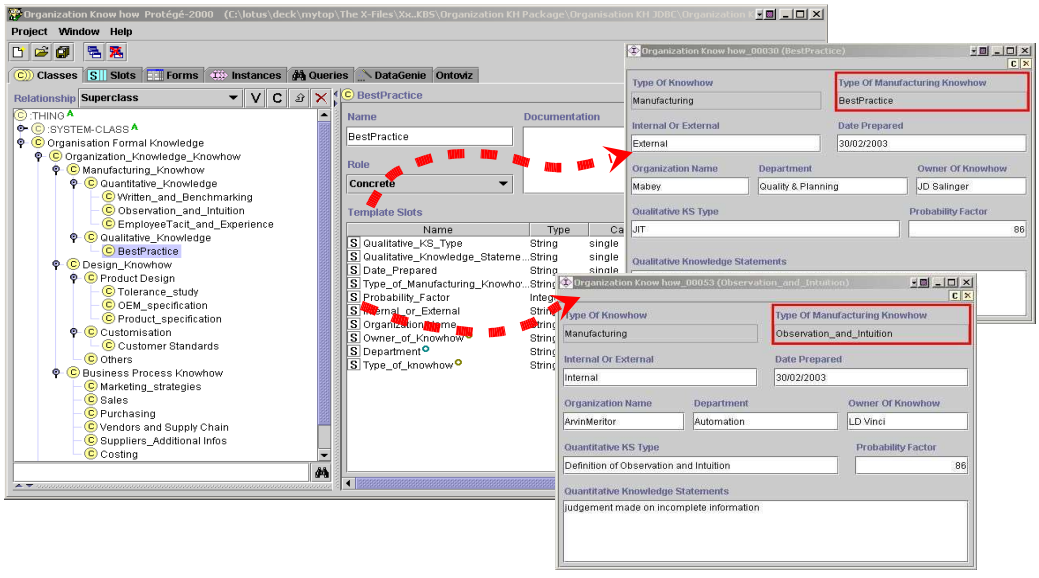

Figure 5.1, Protégé user interface for knowledge acquisition

The retention plate and four item assembly for the door latch (see Figure 5.2) was modelled for the entire lifecycle. In addition the full project data including over 300 parts were modelled into the Bill of Material. New components are entered interactively and work seamlessly throughout the lifecycle. The application proved very simple and intuitive to use. The collaboration tools use proven technologies by SUN and worked seamlessly for many-to-many connections.

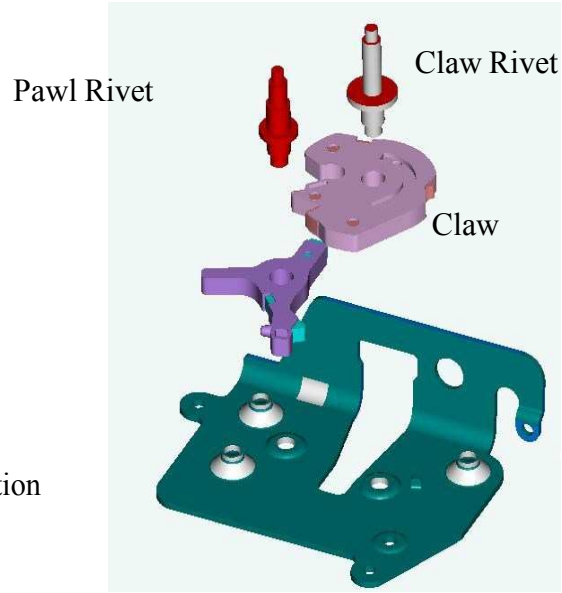

Figure 5.2, Latch assembly view. (Courtesy ArvinMeritor)

6. Benefits

The new methodology fulfils the following criteria for
(i) De-centralisation: Reduce centralised bottlenecks in bandwidth and resources, empowerment of collaborators within networks to "control" the knowledge they create, manage intellectual property rights and enable domain professionals to tailor the system.
(ii) Open standard: reduce interoperability issues for product and project knowledge, easier set up for inter-enterprise collaboration.
(iii) Open source: Elimination of software licence costs, a solution to the problem of vendor lock-in in the long term, elimination of unnecessary complexity and freedom to modify the application.
(iv) Semantic Knowledge: Eliminating the ambiguous context of the knowledge. Efficient query and retrieval mechanisms, intelligent agents to function on context aware information.
(v) Platform and application independence: Enable the enterprise to concentrate on its work and not be tied in to any vendor, rapid migration to future applications.

7. Conclusion

The methodology described is a suitable solution for collaborating enterprises especially SMEs, to create manage and reuse their knowledge, collaborate easily and without expense. There is a lot of work in progress, and final conclusions can only be drawn once the knowledge and communication protocols are integrated more completely.

Acknowledgements

The authors would like to thank the UK Engineering and Physical Science Research Council (EPSRC) for its financial support. The industrial collaborators ArvinMeritor, LSC Group, Mabey & Johnson and PTC Corporation.

References

1. Gao, JX.; Aziz, H; Maropoulos, PG; Cheung, WM., Application of product data management technologies for enterprise integration. *International Journal of Computer Integrated Manufacturing* **2003** 16, 491-500.
2. Aziz. H; Gao, JX; Maropoulos, PG; Cheung, WM., A design environment for product knowledge management and data exchange, *Methods and Tools for Co-operative and Integrated Design*, **2003** (ed) Serge Tichkiewitch and Daniel Brissaud, published by Kluwer Academic Publishers.

3. Lihui; Weiming-Shen; Xie; Neelamkavil; Pardasani, A. Collaborative conceptual design - state of the art and future trends. *Computer-Aided Design* **2002**, 981-996
4. Beckett, Determining the anatomy of business systems for a virtual enterprise. *Computers in Industry* **2003**, *51*, 127-138
5. Camarinha-Matos; Afsarmanesh, Elements of a base VE infrastructure. *Computers in Industry* **2003**, *51*, 139-163
6. Kim; Kang; Lee; Yoo, A distributed, open, intelligent product data management system. *International Journal of Computer Integrated Manufacturing* **2001**, *14*, 224-235
7. Goh; Koh; Domazet, ECA rule-based support for workflows. *Artificial Intelligence in Engineering* **2001**, *15*, 37-46
8. Zha, Du A PDES/STEP-based model and system for concurrent integrated design and assembly planning. *Computer-Aided Design* **2001**, *In Press, Uncorrecte*,
9. Zhou; Chin; Xie; Yarlagadda, Internet-based distributive knowledge integrated system for product design. *Computers in Industry* **2003**, *50*, 195-205
10. Vasara; Krebs; Peuhkuri; Eloranta, Arachne--adaptive network strategy in a business environment. *Computers in Industry* **2003**, *50*, 127-140
11. Sharma; Gao; Bowland, Implementation of STEP Application Protocol 224 in an automated manufacturing planning system. *Journal of Engineering Manufacture (Part B)*. **2003**
12. Cheung, WM; Maropoulos, PG; Gao, JX; Aziz, H., Knowledge-enriched Product Data Management System to Support Aggregate Process Planning, *1st International Conference on Manufacturing Research*, Glasgow **2003**. Advances in Manufacturing Technology XVII, Ed. Y Qin and N Juster, 253-258
13. Cheung, W.M., Maropoulos, P.G., Gao, J.X. and Aziz, H., **2003**. Workflow Activity Task Controller: an Approach to Distribute Knowledge and Information in Collaborative Product Development.
14. Aziz, H., Gao, J.X., Maropoulos, P.G. and Cheung, W.M., **2002**, December. Advanced tools and technologies for collaborative product development and knowledge management. In ADVANCES IN MANUFACTURING TECHNOLOGY-CONFERENCE- (Vol. 18, pp. 115-120). TAYLOR & FRANCIS LTD.
15. Cheung, W.M., Aziz, H., Maropoulos, P.G., Gao, J.X, **2002** Integration of a Manufacturing Model with State-of-the-art PDM Systems, 1st CIRP International Seminar in Digital Enterprise Technology (DET02), 69-72.
16. Cheung, W.M., Maropoulos, P.G., Gao, J.X. and Aziz, H., **2003**. The application of web-based technologies in product data management and manufacturing systems interoperability and data exchange. In *Ecec 2003: 10th European Concurrent Engineering Conference 2003-Concurrent Engineering: Ten Years On*. University of Bath.