High-value document generation: a common methodology proposal

N.PERRY, A.BERNARD, A.CANDLOT, S.A.KHODJA,
IRCCyN - Ecole Centrale de Nantes
1, rue de la Noë - B.P. 92101 - 44321 Nantes CEDEX 03 – France
Nicolas.Perry@irccyn.ec-nantes.fr ; ivgi@irccyn.ec-nantes.fr

Key Words: Document Management, Methodology, Knowledge, Integration

I. Introduction

Lean extended enterprise and build-to-order induce better integration of PLM activities that go through computer aided systems and knowledge-based information environments [1]. This change of landmarks from physical document to electronics claims to redefine information support functionalities [2]. These information have to be shared and exchanged between the actors of the process. Specifications or constraints are usually transmitted from one expert to the other in a global convergence [3]. The differences between their competencies limit the global understanding of problems. Computer integration in the expertise chain aims to optimize this kind of relations and thus the use of enterprise knowledge. The number of different enterprise concepts and complexities caused by different interpretations of these concepts encourages enterprises to standardize concepts and formalize behaviors. These efforts build re-usable and adaptable platforms and imply deep business architecture redeployments [4]. The rapidly changing environment requires convenient collaboration knowledge integration tools [5] and interoperability between different information sources [6].

The main difficulty encountered is to control the complexity of information quantity and informality. A reduction of the work-structure diversity helps in this regard to optimize the information efficiency. Seen from the process of collaborative document writing, it is hard to guarantee the consistency of the document's subject. The contents of shared documents may deal with many subjects and fields, and every collaborator would compose the text by its own understanding, different from other's because of different knowledge backgrounds. As a result the subject of the shared documents will be inconsistent among many copies. Furthermore, the communications between collaborative systems, the systems and environments will also influence the consistency of the subject. An approach focusing on the semantic and syntax distinction can help to resolve the consistency problems involved in collaborative writing [7]. This work is part of a collaboration project between two research teams supported by NRF in South Africa and CNRS in France. It aims at the identification of possible synergy around performance indicators for knowledge management improvement. This collaboration starts from a global observation. When benefits from productivity optimization become harder to obtain, the market expectancies are changing faster. Enterprises have to analyze and control their core competencies to react efficiently to this new challenge. In the following sections, two different approaches on two different application fields prove the interest of a global methodology for the creation of information consolidation tools in order to build structured knowledge-based information environments to interoperate between all the partners. This paper presents a three-phased methodology to optimize and ensure coherent enterprise documentation:

1. At first must be identified the fundamental elements of the structure. It corresponds to the Infrastructure Definition Phase.
2. Relationships between these elements are then identified, and the elements are deployed in a coherent manner to optimize their efficiency. It is the Architecture Phase.
3. The third phase is document generation by a validated knowledge-based application

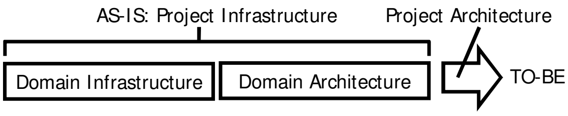

Figure 1: Two levels of Infrastructure and Architecture.

First, the Enterprise Infrastructure is made off elementary concepts that can be classed among process, products, resources and external effects [8]. These concepts specify all enterprise objects relating to three points of view: functional, behavioral and structural (FBS) [9]. Their boundaries can be retrieved through the perception ability of stakeholders [10]. Each person naturally does this division, but the formalization of a common understanding is harder to accomplish. The reason is that knowledge and meaning can't be externalized from humans to computers [11] or other documents. Meaning contained in representations has to be internally rebuilt by users [10]. Documents are considered as the inscription of knowledge and the problem is to analyze and propose structured concepts as a base for their management. Ontology is one of the possible ways to achieve this goal. Research on ontology [12] seeks to provide enterprises with concept definition and management tools [13] [14]. The common main steps are: domain limit definition, manual or automatic corpus analysis, concept extraction and organization.

The aim of this first phase is to differentiate concepts. The analysis of their relationship is part of a second phase focusing on Architecture. Concepts "behave" differently according to their context. The modeling of this structure and the analysis of its possible evolution constitute the Architecture phase. The maturity of the infrastructure knowledge leads to a restricted number of concepts. They are more relevant and meaningful for defining the specifics of the studied domain. They are usually formed by a general name (corresponding for example to UML Class) and a limited numbered of typological instances (corresponding to UML Object) [15] [16].

The sum of their behaviour constitutes an as-is platform from which all the business outputs are derived. Usually the platform is build informally according to the enterprise evolution. It raises incoherencies in concept levels or typology definitions. Concretely it can be illustrated by a misuse of a machine (unclear relationship between a process and a resource), an inefficient procedure (confusion in processes), or an unsatisfactory product (unconsidered external effects, badly defined core product concepts). In order to optimise platform efficiency, the roles of the Architecture phase can be defined as follow:

- Ensure the coherency between concepts
- Optimise relationships and build working environments
- Evaluate model maturity and complexity reduction
- Define a coherent integration method of knowledge in final products

The two following examples highlight how a domain Infrastructure and Architecture form an enterprise management Infrastructure. This global as-is state can be then redeployed through an enterprise management Architecture to ensure a better use of enterprise knowledge.

III. Example analysis in insurance industry domain

This example focuses on complex contract documentation found in insurance companies. It illustrates how the methodology introduced in this paper helped to analyse and improve the current Master Contract between the insurance company and its policyholders by first getting a good understanding of the current contract and its impact on the enterprise infrastructure. From this, improvements to the existing enterprise infrastructure could be made, and a new improved contract could be generated. A Master Contract governs all benefits and requirements between the insurance company and its policyholders. The complexity manifesting in the contract, also leads to complexity in the IT systems.

It illustrates how the methodology introduced in this paper helped to analyze and improve the current Master Contract between the insurance company and its policyholders by first getting a good understanding of the current contract and its impact on the enterprise infrastructure. In the event of a claim, the contract may in some cases be interpreted differently by the client and the insurer, leading to disputes between the two parties. The complexity manifesting in the contract, also leads to complexity in the IT systems. In order to configure the IT systems, the contract must be interpreted. Ambiguous interpretations may lead to inconsistencies between the contractual terms and the IT systems, creating a legal risk to the company. Modeling of As-Is enterprise infrastructure, and developing of To-Be improvements.

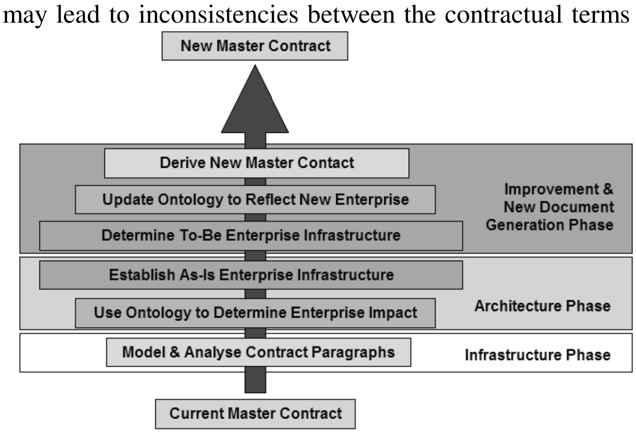

Figure 2 : Master Contract Improvement Process.

The analysis of the paragraphs and the extraction of the paragraph concepts therefore correspond to understanding of the Domain Infrastructure. The next step is to extract specific concepts.

From these contract paragraphs and to establish relationships to other enterprise concepts contained in overall enterprise ontology. In order to decrease the complexity and ambiguity of the Master Contract, it was necessary to firstly understand the contract, and secondly its impact on the enterprise is a whole. To summarize, ontology first help creating the as-is picture of the enterprise, giving an understanding of the enterprise, highlighting its incoherencies and guiding the to-be improvements formulation. In a second phase, the to-be changes to the enterprise have been implemented (cf. Figure 2). Consequently, the first ontology had to be updated in order to incorporate the new evolutions.

IV. Example analysis in aircraft manufacturing domain

In the world today, companies' computerization forces to assume that computer aided systems support design and manufacturing preparation phases [16]. Even if a global integration of the whole product and process life cycle is deployed, the harmonisation of the semantics associated to each expertise included in the life cycle remains difficult. It aims at the specification and development of a knowledge-based engineering tool to help the definition of process plan for small-size high-specificity production batches. The process plan is the complex document considered here. This tool is specified either by the sum of diagrams (graphical representation) or by the sum of objects and their ties (informatics representation). Their evolution depends on two knowledge axes: the project maturity and the refinement of domain knowledge.

The first task of the Infrastructure Phase is to systematically identify them through documents, presentations and meetings. The justification of this work lies in the need of visibility and understanding. The number of concepts and processes rapidly increases [17].

The spine of the data model is progressively defined. Models are simplified by the identification of main concepts that are kept for rapid communication in the project team. The last category is induced by the introduction of a temporal link between process model and data model. The choices in scenarios become determining on the final platform efficiency.

The first uses entity classes to map the studied mechanical part according to the formalized knowledge.

The second uses tool classes to sort out relevant items and to open the structure to equipment evolutions. At the end of the pre-competitive phase of the development project, the expected gain is a ten factor. The resulting tool must still be completely deployed and validated by the industrialization phase.

V. Conclusions

The conceptual similarities between these two very different application examples reminded on Table 1 encourage both teams to structure their work in a similar methodological approach. Naturally similar kinds of problematic are encountered, implying the search for common knowledge-based methods to solve them. It indicated that completely different tools (EDEN© and Roadmap towards MEGA© and UML diagrams) have deep similar concerns, i.e. to ensure coherency, consistency and a unified understanding in multi-faceted project teams. The integration of both methodological aspects and technical solutions leads to a skeleton strategy reusable for further identical problems. The main experience outputs are summarized in Table 2.

	Insurance Industry Example	Aircraft Manufacturing Example
Domain Infrastructure & Architecture Specifications	Master Contract	Expert Booklet
Expected KBE Tool	CRM Tool	Computer Aided Tool for Process Plan Definition
Final Complexe Document	Premiums, cover, claim payments	Process Plan

Table 1: Example-Context Comparison

Principles	Insurance Industry Example	Aircraft Manufacturing Example
Collect and Evaluate Information	Corpus Analysis / Meetings / Mindmaps	Corpus Analysis / Meetings
Sustainable Traceable Updating	Doc. Managemnt / Versioning System	MEGA Database Management
Accurate Overview	3D Solution Space	MEGA Referential
Common Environment	EDEN Tools	MEGA
Syntaxic Choices		
Data Representation	Metadata / Keywords	UML Class Standard
Process Representation	Moogo	UML Activity Standard
Semantic Choices		
Domain Representation	Moogo Diagrams / Life Cycle Roadmap	UML Diagrams
Project Representation	Masterplan Roadmap (PERA)	UML Sequence Diagrams
Share Concepts between Users	Ontology Building	XMi / XML automatic generation
Data Analysis / Performance Indicators	Database Use Statistics	Referential Size Statistics

Table 2: General to specific phase identification

Informatics research works on algorithms to automatically analyse a document corpus in order to create a first cut of a structuring ontology. The other difficulty lies in compiling statistical data in relevant indicators. It can be maturity phases or the respect of the methodological principles highlighted here. The Architecture phase goals are thus valuable hints for such indicators. Next works should focus on how to monitor the global data evolution. Moreover the main advantage felt by the teams is the distinction between project and domain elements that is revealed in the need of clearly specifying the project syntaxes for a better construction of the domain semantics. In a nutshell, this paper proposes a methodology to refine unorganised information complexity to semantically enriched relevant concepts.

This reduction of the work structure complexity and heterogeneity helps to optimise the complex document generation by creating more coherent applications or work environments and helps interoperability between experts and applications. The maturity of knowledge contained in these structures contributes to a better agility towards output expectations. The introduced distinctions between Infrastructure and Architecture in one hand and Domain and Project in the other induce de facto an awareness of stakeholders on their position on a knowledge refinement scale. It avoids confusions in concepts considerations and allows the identification of global project risk.

VII. References

[1] M. Thannhuber, M. M. Tseng (1), H.-J Bullinger, An Autopoietic Approach for Building Knowledge Management Systems in Manufacturing Enterprises, Annals of the CIRP, 48/1/1999, p.89.

[2] S. Ha, G. Pahng, M. Chang, S. Park, H.M. Rho (2), Managing design knowledge : active document system, Annals of the CIRP, 50/1/2001, p.313.

[3] A. Candlot, ND. Du Preez, A. Bernard (1), Synergy and Knowledge in an Innovative Project between Academia and Industry, International Conference on Competitive Manufacturing, Stellenbosch, 2004, ISBN 0-7972-1018-0

[4] C. Midler, L'Auto qui n'existait pas, Dunod, 1998, ISBN : 2-10-004228-9.

[5] H. Rozenfeld / W. Eversheim (1), An Architecture for Shared Management of Explicit Knowledge Applied to Product Development Processes, Annals of the CIRP, 51/1/2002, p.413.

[6] A. Sheth, Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics, Interoperating Geographic Information Systems, Vol.25, Kluwer, 1998.

[7] Jianshe, Z., Liyi, W., Using Ontology to Solve the Consistency Problems in Collaborative Writing, Proceedings of the Sixth International Workshop on Collaborative Editing Systems, Chicago, Nov. 6th 2004, http://dsonline.computer.org/collaborative/events/iwces-6/

[8] M. Labrousse, A. Bernard (1), P. Véron, Generic FBS concept for Process/Product/Resource integration, in Tools and Methods of Competitive Engineering, Edited by Imre Horwvath, Paul Xirouchakis, Volume1, pp.384-394, Millpress Rotterdam Netherlands, ISBN 90-5966-018-8, 2004.

[9] M. Labrousse, A. Bernard (1), P. Véron, Enterprise processes traceability based on a FBS approach, 5e Congrès International de Génie Industriel, Québec, Canada, 26-29 octobre 2003.

[10] B. Cassin, Vocabulaire Européen des Philosophies, 2004, ISBN 2-85-036-580-7, ISBN 2-02-030730.

[11] J. R. Searle, Minds, Brains, and Programs, The Behavioral and Brain Sciences, vol. 3. Copyright 1980 Cambridge University Press.

[12] Guarino N., Formal Ontology, Conceptual Analysis and Knowledge Representation, International, Journal of Human and Computer Studies, 43(5/6): 625-640, 1995.

[13] Troncy R., Isaac A., "DOE : une mise en oeuvre d'une méthode de structuration différentielle pour les ontologies.", 13èmes Journées Francophones d'Ingénierie des Connaissances, IC'2002, 2002.

[14] Bachimont B., Isaac A., Troncy R., "Semantic Commitment for Designing Ontologies: a Proposal. ", 13th International Conference EKAW'02, 2002.

[15] G. L. Zuniga, Ontology: Its Transformation From Philosophy to Information Systems, FOIS'01, Ogunquit, Maine, USA, October 2001.

[16] Gruber T. R., Toward Principles for the Design of Ontologies Used fo Knowledge Sharing, International Workshop on Ontology, Padova, Italy, March 1993.

[17] A. Bernard (1) & N. Perry, Fundamental concepts of product/technology/process informational integration for process modeling and process planning, Int. Journal of Computer Integrated Manufacturing, vol. 16,(7).pp. 557-565.2003,N° ISSN 0951-192X.