Customised high-value document generation

Niek du Preez¹, Nicolas Perry², Alexandre Candlot², Alain Bernard (1)², Wilhelm Uys¹, Louis Louw¹
¹University of Stellenbosch – Industrial Engineering Department – South Africa
²IRCCyN (UMR CNRS 6597), Virtual Engineering for Industrial Engineering - Nantes - France

Abstract
Contributions of different experts to innovation projects improve enterprise value, captured in documents. A subset of them is the centre of expert constraint convergence. Their production needs to be tailored case by case. Documents are often considered as knowledge transcription. As the base of a structured knowledge-based information environment, this paper presents a global approach that helps knowledge-integration tool deployment. An example, based on process plan in aircraft manufacturing, indicates how fundamental understanding of domain infrastructure contributes to a more coherent architecture of knowledge-based information environments. A comparison with an experiment in insurance services generalised the application of presented principles.

Keywords:
Methodology, Knowledge, Integration

INTRODUCTION

1.1 Documentation Complexity
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 documents to electronics claims to redefine information support functionalities [2].
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 optimise 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 standardise concepts and formalise behaviour. These efforts build re-usable and adaptable platforms and imply deep business architecture redeployments [4].

The rapidly changing environment requires convenient collaboration and 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 [7]. A reduction in the work-structure diversity helps to optimise information efficiency and meaning. 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 based on his understanding, different from others 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 focussing on the semantic and syntax distinction can help to resolve the consistency problems involved in collaborative writing [8].

1.2 Project Context
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 optimisation become harder to obtain, the market expectancies change faster. Enterprises have to analyse 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.

This paper presents a three-phased methodology to optimise and ensure coherent enterprise documentation:
- At first the fundamental elements of the structure must be identified. It corresponds to the project infrastructure definition phase.
- Relationships between these elements are then identified, and the elements are deployed in a coherent manner to optimise their efficiency. It is the project architecture phase.
- The third phase is document generation by a validated knowledge-based application.

2 KNOWLEDGE BASED COMPLEX DOCUMENT GENERATION

2.1 Understand the Infrastructure
The enterprise infrastructure is made up of elementary concepts that can be classified as processes, products, resources and external effects [9]. These concepts enable the specification of all enterprise objects relating to three
The boundaries of knowledge and meaning cannot be externalised from humans to computers or other documents. Meaning contained in representations has to be internally rebuilt by users. Documents are considered as the inscription of knowledge and the problem is to analyse and propose structurization concepts as a base for their management. Ontology is one of the possible ways to achieve this goal. Research on ontology seeks to provide enterprises with concept definition and management tools. The common main steps are:

- Domain limit definition.
- Manual or automatic corpus analysis.
- Concept extraction and organisation.

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 modelling of this structure and the analysis of its possible evolution constitute the domain architecture phase.

2.2 Construct the Architecture

The maturity of the infrastructure knowledge leads to a limited 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 number of typological instances (corresponding to UML object).

The sum of their behaviour constitutes an as-is platform from which all the business outputs are derived. Usually the platform is built 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 follows:

- 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 following example highlights how a domain infrastructure and architecture form a project infrastructure that helps management. This global as-is state can be then redeployed through project architecture to ensure a better use of enterprise knowledge.

![Figure 1: Two levels of Infrastructure and Architecture.](image)

3 EXAMPLE ANALYSIS IN AIRCRAFT MANUFACTURING DOMAIN

3.1 Definition of the couple Product / Process

In the world today, companies’ computerization forces to assume that computer-aided systems support design and manufacturing preparation phases. 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. In this example, several conceptual worlds co-exist; they force multiple expert reinterpretations.

An aircraft manufacturer, in a partnership with CAD/CAM development leader, five laboratories and a government organisation, proposed the study of the possible integration of design process plan and manufacturing. It has led to a Research Project supported by the French Ministry of Industry. It aims at the specification and development of a knowledge-based engineering (KBE) tool to help the definition of process plan for small-size high-specificity production batches.

The number and diversity of partners have required a common reference to refine specifications from the expressed needs of the aircraft manufacturer to the final knowledge-based engineering tool. The process plan is the complex document considered here. It can be seen as the output of the KBE tool. This tool is specified either by a sum of diagrams (graphical representation) or by a sum of objects and their relationships (computer representation). Their evolution depends on two knowledge axes: the project maturity and the refinement of domain knowledge.

3.2 Project Infrastructure definition phase

The combination of expertises formed an implicit conceptual corpus. 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. It creates an IS ontology that represent the inter-subjectivity of stakeholders, preparing a coherent reference base to build the final KBE application.

A modelling tool, MEGA, helps to perform this systematic exploration. UML-like activity diagrams (Figure 2) and UML class diagrams have been used. The interest of these formalisms is that they propose a strong common syntax but allowed people to rebuild their own semantic interpretation of meta-models. To maintain the advantage of the flexibility, this freedom implies a clear common definition of this semantic before starting to model. Rework of models can regularly occur when deep changes have been highlighted. It is the occasion to renegotiate the semantic interpretation. These reworks are significant leaps to maturity. During the project development, three principal phases were used:

- Brainstorming phase: all partners propose the objects they feel they need for their part of the work. The separation of tasks is difficult and processes are often over detailed and over constrained. The number of concepts and processes rapidly increases.
- Homogeneity phase: the knowledge on the existing software model increases and brainstormed concepts are progressively replaced by corresponding elements

![Figure 2: Part of the Domain Activity Diagram.](image)
of the already-deployed models (industrially or scientifically).
- Refinement phase: the tasks on processes are better distributed and are more homogeneous. 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 remaining items are semantically more relevant and coherently interrelated.

The last category is induced by the introduction of a temporal link between process model and data model, when scenario is discussed. This link has been formalised by a sequence diagram. It creates a justification dynamic that can be illustrated by Figure 3.

It should be noticed that this modelling method only helps to justify the need of elements but does not solve the definition of contained algorithms. At this stage, the work on project infrastructure (concepts referring to process planning infrastructure and architecture) is combined with architecture. The choices concerning the scenarios determine the final platform efficiency.

![Figure 3: Concept Validation with Sequence Diagram.](image)

### 3.3 Project Architecture constitution

As the main items have been identified, the specification maturity process continues in detailing methods that make them interacting. In the presented industrial case, the specifications of the architecture can be summed up as a need for a link between geometry, resources and process capabilities and a resolution of constraint problems [23].

The central place of knowledge capitalisation through databases orientated the solution to an integration of tools, possible manufacturing sequences and classified entities in one main database called OSE (French acronym for Tool, Sequence and Entities [24] – they are the high-value project-infrastructure concepts). For the constraint convergence of these semantically different items, a typology system has been added to help the aggregation of all elements. The new infrastructure organisation is presented in Figure 4.

![Figure 4: OSE architecture organisation.](image)

The process flow using this solution appears in Figure 5. Two main steps can be highlighted. The first uses entity types to map the studied mechanical part according to the formalised knowledge. The second uses tool types to sort out relevant items and to open the structure to equipment evolutions. A last third step consists in a classical cutting condition optimisation [17] and a validation of proposed solutions in the workshop.

![Figure 5: OSE architecture deployment.](image)

### 3.4 Complex Document

Thus, for each new part considered, it becomes possible to first automatically identify the manufacturability of geometrical entities and then prepare the skeleton of the process plan by an organisation of sequences.

Another procedure helps to define the set-ups by calculations of accessibility directions. It is not described here.

With the help of these tools, the CAD document can rapidly be transformed into a CAM document, the expert concentrating his efforts only on difficult points. At the end of the pre-competitive phase of the USIQUCK project [18] the expected gain is a ten factor. The resulting tool must still be completely deployed and validated by an industrialisation phase.

Other fields of application could expect benefits from modelling processes followed here. Cost optimisation in microelectronics components is proposed. A similar approach is deployed in this new context.

### 4 DISCUSSION

When comparing the above example with a very different one in the field of insurance, it shows conceptual similarities depicted in Table 1. It encourages the structuring of a general methodological approach.

<table>
<thead>
<tr>
<th>Tool Type</th>
<th>Sequence</th>
<th>Entity Type</th>
<th>Domain Infrastructure &amp; Architecture Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gauge</td>
<td></td>
<td>Geom. Type</td>
<td>Master Contract</td>
</tr>
<tr>
<td>Mode</td>
<td></td>
<td>Dimension</td>
<td>Expert Booklet</td>
</tr>
<tr>
<td>C Material</td>
<td></td>
<td>Quality</td>
<td></td>
</tr>
<tr>
<td>C Coupe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tool Type</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expected KBE Tool</th>
<th>Domain Infrastructure &amp; Architecture Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRM Tool</td>
<td>Master Contract</td>
</tr>
<tr>
<td></td>
<td>Expert Booklet</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Final Complex Document</th>
<th>Insurance Industry Example</th>
<th>Aircraft Manufacturing Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premiums, cover, claim payments</td>
<td>Insurance Industry Example</td>
<td>Aircraft Manufacturing Example</td>
</tr>
<tr>
<td>Process Plan</td>
<td>Expert Booklet</td>
<td>Computer Aided Tool for Process Plan Definition</td>
</tr>
</tbody>
</table>

Table 1: Example-Context Comparison

Naturally similar kinds of problems are encountered, implying the search for common knowledge-based methods to solve them. It indicated that completely different tools have deep similar objectives, 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 summarised in Table 2.
In this process of specification refinement for a final complex document, strengths and weaknesses appear at the same stages in the two examples. The first drawback is that a relevant collection of terms is difficult to create. This task is mainly performed manually. Computer 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. The creation of a global methodology gives a better understanding of what should be monitored. It can be different maturity phases or the respect of the methodological principles highlighted here. The architecture phase goals are thus valuable guidance for such indicators.

Further research should focus on how to monitor the global data evolution. Moreover the main advantage is the distinction between project and domain elements that is revealed in the need to clearly specify the project syntaxes for a better construction of the domain semantics.

<table>
<thead>
<tr>
<th>Principles</th>
<th>Insurance Industry Example</th>
<th>Aircraft Manufacturing Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collect information</td>
<td>Corpus Analysis / Meetings</td>
<td>Corpus Analysis / Meetings</td>
</tr>
<tr>
<td>Sustainable Traceability</td>
<td>Doc. Management / Upgrading System</td>
<td>MEGA Database Management</td>
</tr>
<tr>
<td>Accurate Overview</td>
<td>3D Solution Space</td>
<td>MEGA Referential</td>
</tr>
<tr>
<td>Common Environment</td>
<td>EDEN Tools</td>
<td>MEGA</td>
</tr>
<tr>
<td>Syntactic Choices</td>
<td>Metadata / Keywords</td>
<td>UML Class Standard</td>
</tr>
<tr>
<td>Process Representation</td>
<td>MooGo</td>
<td>UML Activity Standard</td>
</tr>
<tr>
<td>Semantic Choices</td>
<td>Domain Representation</td>
<td>Moogo Diagrams / Life Cycle Roadmap</td>
</tr>
<tr>
<td></td>
<td>Project Representation</td>
<td>Masterplan Roadmap (PERA)</td>
</tr>
<tr>
<td></td>
<td>Share Concepts between Users</td>
<td>Ontology Building XM/ XML automatic generation</td>
</tr>
<tr>
<td></td>
<td>Data Analysis / Performance Indicators</td>
<td>Database Use Statistics</td>
</tr>
</tbody>
</table>

Table 2: General to specific phase identification

5 CONCLUDING REMARKS

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. The maturity of knowledge contained in these structures contributes to a better agility towards output expectations.

The introduced distinctions between infrastructure and architecture in the one hand and domain and project on the other induce, de facto, an awareness of stakeholders on their position on a knowledge refinement scale. It avoids confusion in concept considerations and enables the global project risk to be identified.

The two specific contributions open ways to deploy these principles in different activity domains. The separation between methodology and tools improves a step further the possible agility that today’s enterprise project culture may require.

6 REFERENCES
