MULTI-lingual Documentation

Jörg Schütz

IAI
Martin-Luther-Str. 14
D-66111 Saarbrücken
Tel: +49 (0)681 38951 32
Fax: +49 (0)681 38951 40
Internet: joerg@iai.uni-sb.de

Lean production and flat decision making hierarchies are the predominant characteristics of a successful, modern, globally operating enterprise. Within the last years, this form of enterprise management has led to the decrease of costs of complex manufacturing processes and the increase of productivity and product quality. Thus, it has also an important impact on the satisfaction of product users and product customers, as well as of the employees in the various value cycles of the manufacturing processes.

The documentation of technical products in its different forms (product data, marketing data, process descriptions, repair and service manuals, owner manuals, training materials, and so forth) and formats (databases, print media, CDs, online information, and so forth) is getting more and more an enterprise's bottleneck. This, since the employed production processes in the documentation field are not integrated into the restructuring and reorganisation procedures of the production and management work-flows of lean production. Therefore, technical documentation cannot keep pace with the new production requirements (time pressure for the document production, better quality of documentation, and so forth). However, technical documentation is more and more recognised as being a strategic added value of an enterprise; this in particular by enterprises operating in global market domains, and which have to deliver the product documentation simultaneously in several languages, such as the automotive industry and the Pharmaceuticals industry.

In these industries, management is looking for techniques and technologies to improve the quality and accuracy of documents, to speed up the production of documents, to find ways to automate some of the production processes, and to reduce cost and time-to-market of particularly multilingual documents. These techniques concern, on the one hand, the modelling of the documentation processes and their integration with product data environments, and on the other hand, the modelling of software architectures and their implementation for the specification of a computational framework for advanced networked digital documentation environments. They also seek techniques to manage the complexity of such advanced information technology systems as they increase in scope and scale. In particular, they have recognised the need to solve recurring problems, such as physical distribution, concurrency, replication, security, load balancing, fault tolerance, and above all the efficient and effective handling of natural language and multilinguality. The complexity will vary by application domain and process phase. One of the key motivations is to create software and lingware packages that can adequately address all scales of documentation complexity across all application domains. An additional motivation is to integrate the best practices in industry, encompassing widely varying views based on levels of abstraction, domains, architectures, life cycle stages, implementation methodologies and implementation technologies.

During the implementation of the MULTIDOC Concerted Action (a kind of feasibility study for the definition of a an industrial strength project), it became clear that the automotive organisations recognised MULTIDOC as strategic to their business, particularly within the domain of multilingual documentation. A MULTIDOC partners consortium was established with several automotive organisations (BMW, Bertone, Jaguar, Rolls-Royce, Renault and Volvo) willing to dedicate resources to work toward a strong MULTIDOC system solution definition and implementation. The intended project investigation cycles are characterised by: Fragmentation that describes the situation at the start time of MULTIDOC where different systems and methodologies are deployed in the different European automotive enterprises. The next step is to unify the different approaches, and to aim at a harmonisation of the existing product data, in particular the terminological data. These efforts then should lead to

a standardised software architecture and the employment of multilingual language technologies. Obviously, as a first step toward standardisation one particular application domain was selected (service and repair methods). The MULTIDOC partners contribute a variety of expert perspectives, including, but not limited to the following: information technology perspectives, business modelling, state machine semantics, documentation types, components, collaboration, refinement, software frameworks, distribution, and meta-modelling. The final result will be a collaborative team effort. The future of this effort shall result in the standardisation of the chosen approach and its industrialisation. Product documentation, particularly multilingual product documentation, is highly time consuming, expensive and not co-ordinated with the new production cycles of lean production and agile competition. For example, product data environments (technical data, design data, construction data, manufacturing data, repair and maintenance data, and so forth) and product documentation environments (different kinds of manuals, technical notes, service bulletins, and so forth, as well as their distribution on several media such as paper, CD, on-line, and so forth) are separated from each other, but they share similar knowledge production steps. The production steps in product documentation concern:

- **Data gathering** that is essential for building the basis of technical information and know-ledge repositories.
- **Data selection and categorisation** that is needed for the identification of the information and knowledge units as elements of these repositories.
- **Authoring** that is the document creation process, i.e. the articulation of information and knowledge units by means of natural language expressions.
- **Editing** that is the structuring of a technical document for a specific form of delivery, including possible navigation capabilities (paper, micro fiche, CD, on-line, and so forth).
- **Translation** that is concerned with the all multilingual aspects of product documentation on the language level and on the cultural level.
- **Managing** that is concerned with all service and maintenance operations of product documentation (update, revision, cataloguing, etc.).

Although both environments share the basic building blocks of their processing cycles, product documentation is not an accompanying process of the product data operations, and it starts at a quite late time within the product manufacturing cycle. This obviously causes time delays of the simultaneous shipment of the product and the product documentation, particularly when multilingual documentation is required. Today, the production of multilingual documentation normally starts after the release of the source language documentation, which causes yet another time delay.

Possible solutions to circumvent the product documentation bottleneck can be found in three domains:

1. **Co-ordination and harmonisation** of product data environments and product documentation environments as, for example, demonstrated by the recent activities to couple the STEP standard with the SGML standard.
2. Deployment of **linguistic intelligence** put into operation by language technology in both environments.
3. Introduction of **advanced information technology** to permit an optimal integration based on object orientedness and interoperability and the concept of continuous improvements (aka "kaizen") as facilitated in product manufacturing. The latter allows for the compliance with standards, the handling of legacy data and systems, and the integration with other information technology domains of the enterprise such as office automation and work flow management, as well as Web-based technologies such as intranets and extranets.

The investments (research and development, introduction of operational systems, and so forth) into these domains are justified by the multiple return-of-investment (ROT) possibilities that concern not only the core business (automotive industry), but in addition the European (and international) component manufacturing industry, the European (and international) scientific community and last but not least the European (and international) citizens on their way to the knowledge society. The demonstration of MULTIDOC will concentrate on the authoring step where several control capabilities for the source language support the technical writer in defining information objects. This also includes first steps toward "translation engineering" that is concerned with the optimisation of the translation production chain and the linking of source language information and target language information.