European Perspectives on Technical Communication:

Drivers and Concepts of Content Management Systems in the Age of Globalization and Mass Customization

By Prof. Dr. Wolfgang Ziegler

In the field of technical communication, content management systems (CMS) have a long tradition in German speaking countries and are heavily used for creating manuals and other product related information sets. In this article, we focus on the methodological background of these systems. We also explain how they generate benefits for process-driven scenarios of content creation for complex technical products and globalized customer demands. Most recent developments of content management systems and especially the concept of dynamic content delivery can shed a light on future perspectives of technical communication and the idea of intelligent content.

The history of technical communication on the European and on the German market, in particular, is tied to the standardization of content and the corresponding processes of content creation. Starting point have been the national regulations and standards of the European Community regarding document structures and the contained information therein. Technical documentation is treated as a functional part of the delivered product. Therefore, the expectations towards documentation regarding correctness, completeness and normative compliance are traditionally high.

■ Standardization levels in technical documentation

From the perspective of information theory, the history of standardization, as previously mentioned, covers the different levels and scientific areas as follows:

■ Standardized content

- Technical writers usually have to follow specific linguistic authoring rules for the creation of standardized and exchangeable content. Rules have been derived for example from linguistic or cognitive sciences in order to enhance communicative clarity and understanding [1] [2].
- Terminological definitions and control of technical terms are introduced within companies or business domains in order to improve consistency and reduce misleading information and confusion.
- Controlled Languages: military, aerospace and some other industries use controlled vocabulary and authoring rules like “Simplified Technical English” (STE) [3]. The purpose of STE was to avoid the need for translation and to increase readability through standardized content.
- In CMS, these standardization levels are often enforced by means of tools in terminology and controlled language. The latter provide corrections according to grammar, spelling and linguistic style rules. The most recent language support stems from so-called authoring memory systems suggesting phrases that already exist during the writing process within the editing tool. It also enhances consistency and should reduce translation costs.

■ Standardized document and information structures

From a technological point of view, standardization is also achieved by enforcing rules for document structures and, in a more general way, by predefined content structures. The following list gives an overview of the technologies covered:

- XML data format, which is used in the vast majority of CMS. It forces authors to structure content using a given set of semantic elements (“tags”). XML uses the universal character encoding (Unicode) as an exchange standard of text content in any given language [4]. Therefore, content can be globally authored, exchanged and published in any XML based media. Modern printing systems and all web-based deliverables support Unicode.
- XML Information models define the specific set of structuring elements available for technical writers. Usually, they are modelled in a more or less

semantical manner and obey rules of sequence and hierarchy between each other. Examples of semantic elements are "action", "result", "data" etc.

- Definition of reusable content objects and classification. Within CMS, content is organized in modular units. These units have to be created in a controlled and standardized way, not individually, in order to ensure consistency across and within documents. Classification rules and the use of clear metadata schemes support a concise definition of modular content.
- Aggregation rules help organize and standardize the (re)use of modular content for building documents. Compliance with rules can be automated in certain CMS or enforced manually by means of review and control of classified modular content and its occurrence in documents.

■ Standardized output media

- Standardized layout has been a common starting point for introducing CMS. Different products of enterprises have to follow a corporate design being ensured by automated publishing technologies. Using XML, there are standard style languages for all cross media publishing formats and media.

■ Processes

- Within companies, the processes for creating, managing, translating and publishing information are usually reorganized when CMS are introduced. Processes then follow the idea of single sourcing and are influenced by other business processes as described in the following.

■ Product engineering as a driver of content management

Taking a closer look at the history of content management systems specialized in technical communication, the reasons for their widespread adoption can be easily found. The standardization of products within engineering together with optimized production processes led to the so-called mass-customization of products according to the needs of customers. Especially the following product-related conditions affected subsequently the creation and provisioning of corresponding information sets and technical documentation:

■ Product complexity and variants

Most types of products reveal an enormously high degree of functional complexity. They are often built of components and assembled in many different types of configuration variants. The product variants are often tailored to the needs of individual customers (i.e. "mass customization").

■ Product dynamics and revision history

During product development, production processes and deployment, there are usually many product changes. Products, functions and components are subject to dynamic product lifecycles and cause product revisions especially during service and maintenance processes of machines.

■ Globalized product markets

Target markets of the export driven European manufacturing industries are spread worldwide. Products sold inside the European Community usually follow common standards but there are still different variations depending on the country. On the global scale, there are many product adaptions to individual markets due to local regulations and standards.

The facts mentioned above had an influence on technical communication in general and on content management concepts in particular as early as in the beginning 1990s and led to the development of

corresponding software systems.

■ Solutions for managing complex content

Basically, standardized and componentized products triggered the concept of content reuse. This means, technical writers follow the idea of modular single sourcing where predefined modular content components are reused in a controlled way. While the early concepts were often chapter-based reuse scenarios, today CMS users often reuse theme-based or, so-called "topic-based" content components. Throughout this paper the most general term of modular content is used instead of the more specific term topic.

With reference to the key facts of product engineering as described above, CMS give the corresponding technological answers:

- Product complexity is reflected by means of variant management of content modules. This means, there are identifiable module variants describing different types of components being used in end products. In addition, they can be aggregated in many ways in order to obtain the appropriate variant-specific documentation of the mass customized products.
- Dynamics of products and changes are covered by detailed version management of content modules. Versioning is accompanied by review and release workflows of all individual content objects (graphics, fragments, modules and documents). Systems support complex update scenarios of documents and all revised content modules or graphics referenced therein. Fallback to and updating of former versions of documents are indispensable functions for a long-term content maintenance of long-lasting machines.
- CMS manage content objects in all required languages. Systems keep track of content objects in all versions and their corresponding languages. Translation processes are kept lean and costs are saved by exporting just those modules that have been changed or newly created. Translators are supported with additional content material like graphics, context information and pre-translated publications or drafts.
- Within the European Community it is required to deliver product information in each specific language variant of the end users. The corresponding translation costs and the huge amount of translation overhead has always been an extremely important argument to facilitate the corresponding processes and to minimize costs. Therefore, the reuse of content in source languages reduced costs significantly; costs were also reduced in all target languages.

■ Use of CMS and tekom statistics

An estimate of the number of users creating content with the help of CMS amounts to 10.000 - 20.000 users, mostly in German speaking countries. The international tcworld organization and its German counterpart tekom together with the author performed surveys and publications on CMS in 2005, 2008 and 2013 [5]. These studies show a strong increasing tendency toward the use of CMS in all areas of technical communication. In 2013, companies on enterprise level used CMS with a distribution rate around 80%, mid-size companies with a rate of 50% to 60% and even small-sized companies with a rate around 40%. Even though the results stem from tekom members, the amount of data collected from up to 800 companies deliver reliable and highly representative data on most statistics.

There were 20-30 CMS vendors offering a variety of content management systems and approaches. Roughly 10 systems can be considered as tools being used on a broad customer base. Some of them are also

distributed for example in Asia, the US and the UK. Usually, they differ in the handling of metadata, the capabilities to automate and support processes, client and web-technologies and also in the information model implemented in the CMS.

■ Information models and process management

There are only a few standardized XML information models available for the structuring of technical communication and specialized subjects like training, service or end user instructions. Originating from software documentation domains, especially from the US and other English speaking countries, the DITA information model is spreading partially also into other domains [6]. Within the community of the above-mentioned system vendors, the information model is of minor importance. Almost all of these systems provide system-based semantic information models which have been developed through customer projects and can be seen as internal system standards. Some of the systems also have DITA support. The use of DITA vs. system structures is intensely debated by CMS vendors and consultants all over the world. [I] From a neutral technological perspective, the decision for the one or the other information model is made, at least in the Central European market, by the mentioned process support of the CMS and not by the XML structure itself. [II] These days, the specific information model plays a less important role than standardized content on a linguistic level. This might be changing in the future when global perspectives and data exchange become more relevant. Other specific and standardized information models partially included or supported in the mentioned CMS are DocBook (publisher structure), S1000D (European military and aviation industries) and PI Mod (German machinery industry). Standardized information models are used in roughly 10% of CMS implementations. Within software industry and corresponding applications without CMS, the rate is considerably higher [7].

■ Methodologies for Content Engineering

Introducing CMS requires a thorough planning of all preparatory actions. The mentioned CMS study of the tcworld organization includes a 10-step introduction scenario [8]. Apart from the necessary IT project management it covers most relevant "content engineering" tasks recommended as the core phase of internal preparation. As mentioned above, the task of standardizing content with linguistic rules, module definition and variant handling, the definition of metadata and automization rules, the selection of information models, the definition of style guidelines and finally the determination of migration scenarios of legacy data are included within this phase.
Performing these tasks requires the knowledge of various technologies and processes. Therefore, the process of introducing CMS is often accompanied by independent CMS consultants or consulting departments of CMS vendors. In addition, the number of technical writers is constantly increasing receiving specialized training or academic education in order to work as information developers or content engineers.

■ Training for technical communication

In Europe, there are several universities educating students in technical communication and corresponding information management subjects. In addition, there are training programs for business professionals offered by private organizations, universities and the tcworld organization. Regarding CMS, many of these programs teach the following areas at different academic and technical levels:

- Building modules and reuse scenarios

- Classification of modular content and all types of objects
- XML-structuring, semantic and standardized information models
- Process automization and document creation including variant management
- Linguistics, terminology and authoring of modular content for global use
- Translation tasks and processes
- Programming concepts and languages for data processing and publishing
- Multimedia and web publishing

■ Classification as a basis for mass-customization

An important metadata-driven approach for classifying and defining content modules is defined as the so-called "PI-Classification" [III] method [9]. With this approach, unique and so-called "intrinsic" metadata take into account relevant product components by corresponding product classes (P) and a set of predefined information classes (I). It is then the task of technical writers to assign modular content to unique (PI-)classes and to follow the corresponding linguistic rules for standardized content creation. In this framework, there are additional, so-called "extrinsic", PI-classes describing the intended or actual use of components in end-products. With the aid of the PI-classification concept one can facilitate retrieval and reuse scenarios and automate many processes of document aggregation, publishing and variant management. Therefore, it is a key methodology for creating mass-customized information products from standardized modular content.

■ Dynamic publishing and content delivery of more intelligent information

In the last years, the document-centered view on technical communication has changed towards a more dynamic provisioning of information. So-called content delivery portals (CDP) are under recent development offering topic-based access to content [10]. This means, users are supported with more appropriate and situational web content on a granular level instead of reading monolithic documents on pdf or as printed media. To this end, CDP include different retrieval capabilities like direct search or facetted search in web-based online environments. The latter mechanisms benefit widely from (PI-)classifications from CMS and can directly map module classifications to facets in order to narrow search results. Due to uniquely identifiable content, machines in combination with software products can directly trigger or provide interactive information reacting on present product status like error handling or during process control. Drivers of theses portals are concepts such as the "Internet of Things" (IoT) or the "Industry 4.0" initiative for digitized product services. CDP can be implemented and scaled from global web-portals on the internet, to local on-site portals on machines or mobile applications on portable devices down to ordinary online-help systems.
The benefits of CDP can be understood as the intelligence brought to information products by semantic information models and classification of modular content. This new type of information became necessary for highly configuration-dependent products from mass-customization of globally delivered products.

■ Proof of efficiency and relevance of content

After introducing CMS as well as CDP in the near future, the use of these systems should be tracked and managed. For CMS there is a methodology of standardized reuse metrics which can be measured by the Report Exchange (REx) [IV] technology. REx

measurements are available in a couple of the systems mentioned in the tekom study on CMS. It allows the proof of the efficiency through document reuse rates and reuse counts of modular content. It also permits to implement cost models and to optimize system processes by many key performance indicators (KPI) [11].

As far as CDP are concerned, web analytic tools can be applied to track the actual use of information. At this point, it is possible to find out about the relevance of content or even the lack of information by means of unsuccessful information retrieval in certain areas. It is possible to improve information planning subsequently and further enhance the intelligence of the corresponding CMS and modular content.

■ Summary

This article describes the present situation of CMS in Central Europe and especially in German speaking countries. In these areas, there has been a need for highly adaptable documentation for mass-customized products as well as for globalized markets. Corresponding requirements derive mainly from the machinery industry, influencing and driving CMS development over the past 20 years. Introducing CMS often requires rethinking and reorganizing of many processes of content creation and content publishing. In addition, a successful management of complex modular content requires an appropriate content engineering phase before introducing CMS. Concepts for classification and modularization are indispensable for using CMS in an effective and beneficial way. Modern publishing and dynamic delivery scenarios add value to mass-customized products by providing intelligent content adaptive to user needs, product configurations and their situational relations between each other.

References

[1] Muthig J. “Technical documentation needs standardization” tcworld e-magazine 05/2011 <http://www.tcworld.info/e-magazine/technical-communication/article/technical-documentation-needs-standardization/e-magazine/ >

[2] Information Mapping Inc. “The Information Mapping® Method 30 Years of Research” <http://www.writec.com/Media_public/IMI_history_and_results.pdf >

[3] <http://www.asd-ste100.org/>

[4] <http://www.unicode.org/>

[5] Straub D., Ziegler W. "Effizientes Informationsmanagement durch spezielle Content-Management-Systeme", 3rd edition, tekom (2014)

[6] OASIS Darwin Information Typing Architecture (DITA) TC <https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=dita >

[7] Straub D. “ Branchenkennzahlen für die Technische Dokumentation 2015“ tcworld 06/15 <http://www.tekom.de/fileadmin/Dokumente/de/tekom_2015_07_08_Branchenkennzahlen_2015DE.pdf >

[8] Ziegler W. „State of the Art in Technical Communication Part II: Content Management and Intelligent Delivery“ <http://i4icm.de/fileadmin/content/HSKA/03_Vortraege/CMS_II_public.pdf >

[9] Drewer P., Ziegler W. "Technische Do-kumentation - Eine Einfuehrung in die übersetzungsgerechte Texterstellung und in das Content-Management", Vogel Ver-lag, 2nd ed. (2014)

[10] Ziegler W., Beier H., "Content deli-very portals: The future of modular con-tent" tcworld e-magazine 02/2015 <http://www.tcworld.info/e-magazine/content-strategies/article/content-delivery-portals-the-future-of-modular-content/>

[11] Oberle C., Ziegler W. "Content Intelligence for Content Management Systems" tcworld e-magazine 12/2012 <http://www.tcworld.info/rss/article/content-intelligence-for-content-management-systems/>