New Roles and Competencies in Technical Communication Induced by Semantics and Analytics

Wolfgang Ziegler[1*]

[1] Karlsruhe University of Applied Sciences, Faculty of Information Management and Media, 76133 Karlsruhe, Germany

Abstract. Semantic approaches have become a logical extension of the widely-accepted content creation and provision technologies represented in technical communication mostly by component content management systems (CCMS) and content delivery portals (CDP). CCMS are typically used by technical authors trained in the areas of linguistics, structured and standardized topic-based writing and metadata-based content processing. Delivery scenarios and corresponding new technologies widen the demands of organizations towards the role of technical writers as information engineers with competencies in user-centered and use case-driven creation of content. Information engineers cover a wide range of topics being already the core of academic programs and professional education.

More recently, increasing demands for business process integration and process management led to new technological and methodological approaches. Semantic technologies based on ontology modelling have been investigated and introduced for example for CMS processes to improve the handling of product and media variants and the corresponding content planning. Other developments address metadata governance in integrated environments of multiple information management systems. On CDP side, there is a need to improve search support by semantics for example through facetted search, content correlations or AI technologies. Furthermore, in both fields of CMS and CDP, analytics gains significant importance for tracking and improving the corresponding processes of content creation, content search and information access. Therefore, we claim that a clearer definition of a new role of information architects and their competencies will be helpful for differentiating educational and working areas from the field of information engineers in TC.

1 Introduction

In the last decades, the domain of technical communication (TC) has been well-established in academics and industry. Nevertheless, on a global scale, there are huge variations regarding the academic and professional fields of activities and the integration depth of TC departments into organizations, resp. the role of TC as organizational entities. Recently, the areas of TC-related activities are heavily extending, driven by various initiatives in industrial digitalization and architectural changes of information systems.

In this paper, we consider the corresponding change of roles and the development of new responsibilities in TC domains. We use core TC technologies of Content Management and Content Delivery as a starting point and explain the extensions of roles given by new tasks therein. Further responsibilities arise from semantic technologies, AI technologies, organizational issue of data governance and content strategies towards new information services.

The research work and the domain studies were performed in cooperation with industry partners by developing use cases and technical implementations. We covered different industrial areas like machinery industry, consumer products, automation and process industry and software industry.

On academic side, research work with graduate students and undergraduate class experiences integrated in the TC curriculum, was incorporated.

Consequently, we argue in summary for a clearer definition of new roles and competencies in academics and for business professionals. For this purpose, we use the role of information architects being differentiated from information engineers.

2 Content management

The evolution of component content management, i.e., the corresponding processes and software systems, have developed considerably from desktop publishing to database-driven environments and are nowadays one of the core concepts of TC. Their fundamentals cover topic-based content creation and reuse, structured XML-based information structures, semantic metadata enrichment and variant management on topic or sub-topic level [1]. All these concepts for creating so-called “intelligent content” are technically well established and are covered, at least in many specialized educational programs, also in academic curricula and business education. But there are remaining challenges in business applications. We will focus in this context on planning processes as they turned out to be crucial for future tasks of time and quality keeping information delivery on complex products.

On the one hand, the complexity and necessity of content planning capabilities are often underestimated during CCMS introduction and ramp-up phase of system use. Rapid product development and, especially in many software-driven environments, agile processes cause

* Email: wolfgang.ziegler@h-ka.de

frequent content changes and induce time-consuming and complex publication project management. In addition, the wide-spread lack of CCMS support for detailed project and process management capabilities shows a clear need for further development of systems along these lines. Alternative approaches focus a deeper integration of CCMS with other planning and development systems for hardware and software.

On the other hand, the planning tasks we want to focus here, are on a higher conceptional level. As many products underly a sophisticated variant and configuration management, the corresponding information management must reflect the product complexity to a sufficient degree. The degree is set by the defined use cases of information provisioning, i.e., content delivery scenarios and specificity of content accessible by users. Therefore, the management of content variant handling should be strongly systemized. The role of “variant managers” supporting this will typically cover some or all the following tasks:

- Analyzing product configurations and product variants.
- Evaluating drivers for content variants and their connections with product properties from engineering.
- Implementing variant properties and classifications in CCMS to plan content variants of topics and publications.
- Testing, selecting, and defining methods of topic and sub-topic variant handling of content.
- Writing guidelines for variant handling.
- Tracking and measuring use of semantic metadata and variant-building; developing and interpreting performance indicators and analytics for associated processes.
- Consulting and guiding technical writers for the use of semantic meta-data, topic variants, and variant properties.

Finally, the variant managers have to align the mentioned methods and tasks with given digital use cases and further metadata scenarios.

3 Content delivery

With the upcoming of content delivery portals in the past decade, a more dynamic content provisioning came into the focus of content-related digital initiatives [2]. One of the key aspects of CDP is the more user-centered and situational approach for content access. By this, the complexity should be reduced for information users.

- Analyzing and planning of delivery use case and user stories (e.g., required content and document types) together with target groups and process owners.
- Defining user interfaces and access logistics (e.g., sequence of content access and retrieval processes, navigation, pre-knowledge and context detection of information users).
- Developing requirement specifications for delivery applications (GUI-based or headless CDP).
- Organizing information exchange and processes between content sources and content delivery systems. (e.g., packaging types and standards, single topic vs. aggregated delivery, continuous delivery).
- Developing mapping rules from CCMS metadata to CDP search facets.
- Aligning requirements from content consuming applications with CCMS and CDP (i.e., applications from domains outside TC like service apps and IoT services or AR/VR and chatbot applications).
- Tracking and measuring content use and relevance; developing and interpreting use case-related performance indicators and analytics for associated delivery processes.
- Providing feedback from CDP analytics to content and product engineers to improve discovered problems. Moreover, to CCMS owners regarding metadata quality for search and delivery.

Those responsible for the mentioned tasks should still have a solid knowledge of and background in TC processes in order to develop appropriate CDP architectures and apply meaning and helpful content analytics to information engineers.

4 Semantics, AI & data governance

The area of semantics and knowledge graphs has recently gained again much interest in TC and related areas. Ontologies as a concept and corresponding semantic modelling systems (SMS) have already been used before in TC but did not widely spread. The reasons for the revitalized interest are given, beside technological and usability improvements, in the need of handling a strongly increased product and process complexity. This complexity is prevailing in the interdependencies of product parameters, content sourcing and content provisioning. Moreover, the interdependencies and the underlying relations are modelled more naturally in graph technologies and exceed conceptionally the required level given in most CCMS just by taxonomies.

Another reason can be seen in a broader industrial understanding of the necessity for integration layers of cross-domain data and metadata models. By corresponding data governance initiatives, data quality can be controlled, and communication and transparency of data structures can be improved. As a result, organizations aim for tasks associated to graph technologies like:

- Modelling of semantic relations in graph environments (ontologies);
- Developing parts for digital twins, resp. information twins (covering the variants and configurations for products and information products).
- Defining correlations within and between different metadata domains.
- Integration of standard metadata models e.g., iiRDS, VDI 2770.

- Enhance and improve delivery uses cases by graph relations and correlation rules [3].
- Developing and improving of modelling patterns in semantic models (e.g., class constraints, implicit models and explicit relations, ontology extensions).
- Import and export and update of data for content authoring or content consuming applications.
- Developing knowledge graphs queries for considered use cases.
- Provide Web-Services and API interfaces.
- Communication and alignment with metadata domains for data governance.
- Analytics of semantic metadata quality and process quality for governance purposes.
- Developing extension use cases for ontology use.

Semantic technologies like knowledge graphs described so far, are often considered as a weak type of artificial intelligence. But beyond these, AI technologies based on machine learning or deep learning algorithms, are increasingly used in TC domains. Apart from translation technologies we do not consider here, there are applications used for auto-classification, content summarizing, data-driven content generation or similarity analyses. Among those, auto classification and tagging recommender systems are most often used or are under intensive discussion and evaluation. The use case of auto-classification can be connected to migration scenarios of topic based legacy data or to quality assurance purposes in manual metadata enrichment processes of content. As a loop back to graph technologies, AI can support automated class assignments of content objects within ontologies and extract relations among those.

In the investigated context, AI-related activities can cover the following tasks, including, but not limited to, the above-described:

- Developing use cases for AI applications and their prototyping
- Extracting and providing training data for machine learning.
- Analytics of AI results (e.g., recommended classes) and of quality like precision, recall, and other confidence measures.
- Re-evaluating and readjusting model parameters of AI algorithms
- Derive relations and correlations (e.g., for SCR systematics [3]) from AI-based pattern recognition in web analytics data of CDP.

The above stated examples already show the type of competencies, involved personnel has to acquire. They overlaps to some degree with typical topics from information and computer science.

5 Information architecture and services

Beside the aforementioned technical aspects and responsibilities, an increasing number of TC members are involved to some extent in architectural issues of information systems and also in business aspects of content provisioning.

This can be seen in the contribution to data governance teams from the perspective of metadata and their relations to business areas like developer information, sales and marketing information, product lifecycle or service data appearing in all critical customer touchpoints or for internal use. Being part of the governance requires a solid understanding of various business processes and associated data and information. One of the most dominant areas at the moment is, as already statet, the management of configurations and variant of complex products. It has to be aligned with the use cases of content delivery in order to fulfill the goals of situational and more easy access to product-specific information.

Or, in more general words, the definition of use cases set rules for the content strategy of all content-related systems of content creation, management, and delivery. Deriving these use cases is therefore also one of the dedicated tasks of involved TC members.

An associated task is the development of digital information services (DIS). It goes beyond the identification and definition of use cases and defines realistic business cases and their technical system implementation. There are internal services for internal process improvement where, for example, semantic technologies are used do improve findability and accessibility of internal information.

Moreover, there are new business cases of digital services offered to product owners and users. Many of them are data-driven services like product surveillance of operation or failure for cost minimization in the product lifecycle. In these use cases, TC members can contribute to improve customer satisfaction by appropriate content or to support payable content of higher quality through better accessibility in dynamic media

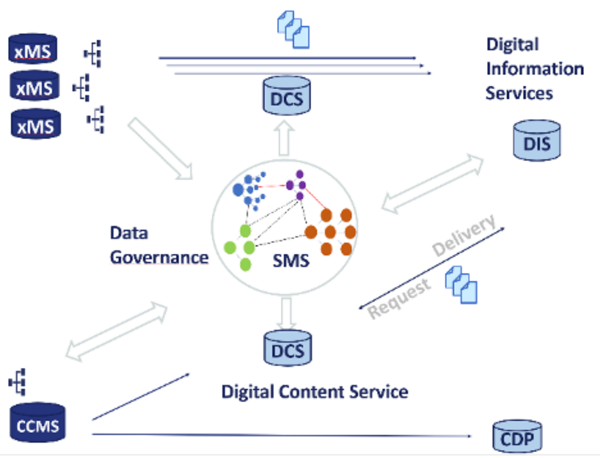

Fig. 1. System architecture and components associated with content authoring, metadata management in semantic modelling systems and content provisioning as the contributions to data governance and to digital information services.

Figure 1 illustrates the system architecture described so far in the previous sections. New tasks and responsibilities are induced by participation in data governance and from contributions to business models build by digital services. Data and metadata governance is depicted as exclusive task of expanding semantic technologies even though, in many companies, this is still a task within traditional database environments. Beside CCMS, there are of course other data and content sources and involved information management systems (xMS) contribution to services and data governance. Digital content services can be realized as headless CMS, resp. headless CDP for the provisioning of content upon request. The graph databases in the semantic modelling systems, SMS can act as metadata services being accessed via web services. CDP can act also as standalone applications having often a focus on content types (here from CCMS) and direct use access via graphical interfaces.

6 Academic experiences

The author has experienced for almost two decades content management issues and underlying XML-related technologies in undergraduate TC classes of approximately 60 students. Content delivery and especially semantic information management has been introduced later to elective graduate courses for a smaller number of 10 students in the average. In general, an estimated number of 30 percent of a students' semester group were typically focusing on semantic information modelling as described in this paper. Others focused on linguistics, communication, or design areas. Subsuming the latter areas for simplicity, at this point, into one field, there was already a clear separation of these field of information engineers, reflecting more classical TC areas.

It turned out, that semantic modelling is a topic of strong students' interest associated with more general questions of knowledge management. Most recent CDP environments and standards make the topic even more attractive within the field of media technologies. A stronger coupling to related output media like AR/VR or chatbots would be desirable being difficult in a limited course time. Corresponding semantic concepts will be transferred subsequently to undergraduate classes but will restricted to conceptional lectures because of a time-consuming project guidance limited to smaller classes.

We discovered a clear need for query knowledge, showing that pre-knowledge in SQL concepts lowered entry thresholds. XML and web technologies were considered as basic pre-knowledge. We also identified a need for introducing REST or other interface services in order to decouple the logical layers of CDP and SMS [3]. The topics of analytics had, so far, not been addressed in the classes of semantic information management. Students have some basic knowledge in statistics and about CCMS metrics. A deeper understanding and experience in web and CDP analytics is still required and would be desirably in future TC curricula.

Problems arose when extensive programming capabilities were needed due to limited background in the TC program. On the other side, abstract concepts of knowledge and metadata modelling are usually strong capabilities of students in these programs. Similar experiences were made in AI areas, where complex algorithms and mathematic formalisms were required.

7 Discussion and Summary

In this paper, we investigated new responsibilities and tasks associated with the introduction of delivery systems, semantic technologies, and corresponding analytics. Subsequently, one can expect the development of new roles in business areas to a broader extent. These roles can be differentiated from information engineers (or i.e. content developers) and can be seen as information architects due the large amount of conceptional and modelling aspect within the mentioned tasks. The wording of information architect is not generally new, but can make it easier to define job descriptions or curricula. In combination with modern dynamic media technologies, one can define the working or educational field of "Information and Media Architecture" (IMA).

The field information architects, described in this paper, requires in summary a variety of competencies at different layers and at different levels of depth.

The technology layer has to cover, beside core knowledge in XML and web technology of CMS and CDP, also semantic standards, and formats (RDF, OWL etc.). Moreover, API and web services technologies.

The methodology layer requires semantic modelling capabilities using different graph patterns, as well as basics in requirement engineering and process modelling for all involved system types. Subsequent analytics are needed for ensuring process and data quality.

On the organizational layer, business processes and their interplay like data governance or lifecycle management issues should be understood. Finally, on the layer of business strategies, information architects should be able to contribute to the development of product services and associated digital information services from a business development perspective.

In summary, dedicated information architects should be able to bridge the gap between content creation, application engineering and business case modelling by architectural, planning and modelling capabilities and by analytical knowledge for process improvement.

We acknowledge cooperation and useful discussions with M. Bamberg, R. Brandl, C. Deschner, M. Kloss, D. Lächler, F. Müller, E. Nauerth, B. Perchutkaite, D. Quilitzsch, J. Scheibe, J. Stähle and many more also on the side of system vendors from CMS, CDP, SMS. The author acknowledges support from the KMM program of HKA.

References

1. Ziegler W. *Basic Concepts Which Support the Management and Delivery of Intelligent Content*, (2017) Journal of Japan Technical Communicators Association JTCA, p. 84-91
2. Ziegler W. *Extending intelligent content delivery in technical communication by semantics: microdocuments and content services*, Proceedings of the ETLTC ACM Chapter International Conference. Aizuwakamatsu, Japan. (2020) https://doi.org/10.1051/shsconf/20207703009
3. Ziegler W. *Semantic Correlation Rules as a Logic Layer between Content Management and Content Delivery*; Proceedings of the ETLTC ACM Chapter International Conference. Aizuwakamatsu, Japan. (2021) (doi.org/10.1051/shsconf/202110202007)