

# 欧州におけるテクニカルコミュニケーションの視点：

European Perspectives on Technical Communication:

## グローバル化と多品種大量生産の時代におけるコンテンツ マネジメント システムの原動力と概念

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

ヴォルフガング・ジーグラー博士

By Prof. Dr. Wolfgang Ziegler

テクニカルコミュニケーションの分野において、コンテンツ マネジメント システム (Content Management Systems - CMS) は、ドイツ語圏の国々で長く使用されており、取扱説明書やその他の製品関連ドキュメント制作に幅広く活用されている。本論文では、このシステムの方法論的背景に焦点を当てる。また、CMS が、複雑な技術製品やグローバル化した製品のプロセス志向のコンテンツ作成シナリオにメリットを生じさせることを説明する。昨今の多くの CMS と特に動的コンテンツ配信の発展は、テクニカルコミュニケーションの将来と知的コンテンツの考え方に光を当てることになるだろう。

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.

欧州およびドイツ市場でのテクニカルコミュニケーションの歴史は、特にコンテンツの標準化とそれに関連したコンテンツ作成のプロセスにあると言える。その起点は、ドキュメント構造とそれらに含まれる情報に関する欧州共同体の規制と規格にあった。技術ドキュメントは、納品された製品の機能の一部として取り扱われている。そのため、ドキュメントに対する正確性、完全性と規範的準拠への期待は伝統的に高い。

### ■技術ドキュメントにおける標準化レベル

情報理論の視点から見て、標準化の歴史は、前述したとおり、以下のさまざまなレベルと科学分野を包含している：

#### ■標準化されたコンテンツ

- テクニカルライターは、通常、標準化された共有可能なコンテンツ作成のための特定の執筆ルールに従わなければならない。コミュニケーションに関わる明解さおよび理解度を向上させることを目的としているため、それらのルー

ルは、たとえば、言語学または認知科学に由来したりしている。[1] [2]

- 一貫性を向上させ、誤解や混乱を減らすため、用語定義や技術用語の管理が企業または特定の事業領域に導入される。
- 制限言語：軍事、航空宇宙などの産業は「シンプリファイド イングリッシュ」 (Simplified Technical English - STE) [3] のような制限的な語彙と執筆ルールを採用している。「シンプリファイド イングリッシュ」の目的は、翻訳の必要性を避け、標準化されたコンテンツを通して、読みやすさを増大させることであった。
- CMS においては、この標準化レベルは、用語と制限言語のツールによって向上できることが多い。後者は、文法、綴り字とスタイルルールに従って修正を施す。昨今のライティング支援は、執筆過程ですでに編集ツール中に存在している表現を提案するいわゆる翻訳メモリーシステムに依拠している。その翻訳メモリーシステムを使うことで、一貫性が維持され、翻訳コストが削減される。

## ■標準化されたドキュメントと情報構造

技術的観点から言えば、標準化は、ドキュメントの構造化、より一般的に言えば、事前に定義されたコンテンツ構造によっても実現され得る。以下で、取り上げる技術の概要を説明する：

- MS で使われる XML (Extensible Markup Language) データ形式について説明する。XML は、所与の意味要素(タグ)を用いて、執筆者に内容を構造化させる。XML は、言語を問わず、ドキュメントコンテンツの交換標準としてユニバーサル文字コード(ユニコード)を使用する[4]。したがって、どんな XML ベースのメディア上でも、コンテンツはグローバルに執筆でき、交換でき、パブリッシング(Publishing) [a] することができる。現代の印刷システムとすべてのウェブベースの周辺製品は、ユニコードをサポートしている。
- XML 情報モデルは、テクニカルライターに使用可能な構造化要素のセットを定義する。それらは、一般的に、意味様式をモデル化し、要素相互間で順序と階層間の規則を守るものである。意味要素の一例は「アクション」、「結果」と「データ」などである。
- 再利用可能なコンテンツオブジェクトと分類の定義について説明する。CMS では、コンテンツは、モジュールユニットで構成される。これらのユニットは、ドキュメント間およびドキュメント内での整合性を確保するため、個別にではなく、制限され、標準化された方法で、制作されなければならない。分類規則と明快なメタデータスキームの使用が、モジュールコンテンツの簡潔な定義をサポートする。
- 集合規則(Aggregation rules) [b] が、ドキュメントを形成するためのモジュールコンテンツの(再)利用の整理と標準化に役立っている。規則への適合は、CMS で自動化できるし、分類されたモジュールコンテンツとドキュメント上に出現するモジュールコンテンツの見直しと管理手法によって、手動でも強制的に守らせることができる。

## ■標準化された出力メディア

- 標準化されたレイアウトは、CMS を導入する際の共通の開始点である。企業のさまざまな製品は、自動化されたパブリッシング技術によって確保されるコーポレートデザインに従っていなければならない。XML を使用することで、すべてのクロスメディアパブリッシング用フォーマットと出力メディアのための標準スタイル言語が存在することになる。

## ■プロセス

- 企業内では、通常、CMS が導入される際に、情報の作成、管理、翻訳およびパブリッシングプロセスが再編成される。そして、プロセスは、シングルソースの概念に従い、同時に、以下で述べるほかのビジネスプロセスの影響を受けることになる。

## ■製品工学の原動力としてのコンテンツマネジメント(Content Management - CM)

テクニカルコミュニケーションに特化した CMS の歴史を振り返ることにより、CMS が幅広く採用された理由を簡単に見つけ出すことができる。最適化された生産プロセスとあいまった工学における製品の標準化は、顧客需要に応じていわゆる製品の多品種大量生産に結びついた。特に、以下の製品に関する状況は、結果的に周辺制作物や技術文書の制作とデリバリーに影響を与えた：

### ■製品の複雑さとバリエーション

製品が多種バリエーションとして開発されることで、複雑さが高じることになる。さまざまな部品の組み合わせによって、異なった種類のバリエーションが多数生産されることになる。製品のバリエーションは、個別の顧客のニーズに応じてつくられる(つまり「多品種大量生産」)。

### ■製品ライフサイクルと修正履歴

製品開発や製造の工程では、一般的に多くの仕様変更が生じる。製品や機能や部品は、ダイナミックな製品ライフサイクルにさらされることになり、特に製品の修理や保守管理プロセスで仕様変更を生じさせる原因になったりする。

### ■グローバル化された製品市場

輸出志向の欧州製造産業の対象市場は世界中に広がっている。欧州共同体内で販売されている製品は、通常共通の規格に従っているが、それでも国によってバリエーションが存在している。グローバル市場規模では、それぞれの国の規制や規格に基づいてさまざまな製品対応が成されている。

ここまで述べてきたことは、特に 1990 年代初頭に一般的なテクニカルコミュニケーションおよび CM の概念に影響を与え、関連したソフトウェアシステムの開発へと導いたことである。

## ■複雑なコンテンツを管理するためのソリューション

まず、標準化とコンポーネント化製品が、コンテンツの再利用の概念をもたらした。このことは、テクニカルライターが、あらかじめ定義されたモジュールコンテンツの部品が制御された方法で再利用されるモジュールシングルスソースの考え方に従うことを意味する。初期段階においては、チャプターベースの再利用シナリオであったのに対し、今日の CMS ユーザーは、しばしば、テーマベース、または、いわゆる「トピックベース」のコンテンツコンポーネントを再利用する。本論文では、特定の用語トピックではなく、モジュールコンテンツの最も一般的な定義が使用されている。上述の製品工学のキーとなる事実に関連し、CMS は以下のような技術的な回答を与える：

- 製品の複雑さは、コンテンツモジュールのバリエーション管理を用いて評価される。つまり、このことは、最終製品に使用されるさまざまな部品の種類を表す識別可能なモジュールバリエーションがあることを意味する。そして、多品種大量生産の製品にとってバリエーションごとに適切なドキュメントが構成できるよう、それらは、さまざまな方法で収集される。
- 製品のライフサイクルや変更は、コンテンツモジュールの詳細なバージョン管理によって管理される。バージョン変更は、すべての個別コンテンツオブジェクト（グラフィックス、フラグメント、モジュールとドキュメント）のレビューとリリースのワークフローに基づいている。システムは、ドキュメントの複雑な更新シナリオおよびすべての修正されたコンテンツモジュールやグラフィックスをサポートする。ドキュメントの以前のバージョンへの復元およびその更新は、寿命の長い機械の長期的なコンテンツの保守のために不可欠な機能である。
- CMS は、すべての対象言語でコンテンツオブジェクトを管理する。システムは、すべてのバージョンに関するコンテンツオブジェクトおよび対象言語をトレースする。変更されたか、または、新しく作成されたモジュールのみをエクスポートすることで、翻訳工程は無駄を抑えられ、経費を節約することができる。翻訳者は、グラフィックス、コンテキスト情報と事前に翻訳されたドキュメントや原稿などの補足的なコンテンツ資料によってサポートされる。欧州共同体内では、製品情報はエンドユーザー各々の母国語で提供されることが要求されている。各製品の翻訳コストと翻訳に関わる諸経費は、対応する工程を単純化して、経費の削減を図るうえで、常に主要な論点であった。結果として、ソース言語のコンテンツの再利用は経費を大幅に削減した。最終的にすべてのターゲット言

語の経費も削減された。

## ■ CMS と tekcom 統計の使用

CMS の支援によってコンテンツを制作しているユーザー数は、推定で 1 万人から 2 万人である。そして、そのユーザーのほとんどがドイツ語圏の国々に属している。2005、2008 と 2013 年に tcworld の国際組織とドイツの tekcom と本著者は、共同で CMS に関する調査を行い、報告書を発行した [5]。この調査結果によれば、テクニカルコミュニケーションのすべての分野で、CMS 利用の増加傾向が示された。2013 年について言えば、大企業レベルの会社が約 80%、中規模会社が約 50% から 60%、小規模会社でも約 40% の比率で、CMS を利用していた。データが tekcom メンバー企業からのものであるにせよ、800 社から収集したデータ量は、統計的に信頼ができ、実情に即したデータを提供している。

さまざまな CMS およびそれらへのアプローチ情報を提供する CMS ベンダーが、20 社から 30 社存在していた。約 10 システムが、広範な顧客基盤ツールとして使用されていると見なすことができる。それらの内のいくつかは、たとえば、アジア、米国および英国においても流通している。通常、それらは、メタデータの扱い、自動化、プロセス、クライアント、ウェブ技術をサポートする能力が異なる。

## ■ 情報モデルとプロセス管理

テクニカルコミュニケーションとトレーニング、サービス、またはエンドユーザー使用説明書などに特化した対象を構造化するための標準化された XML 情報モデルは数少ない。ソフトウェアドキュメントの領域（特に米国と他の英語圏の国々）で使われていた DITA（Darwin Information Typing Architecture）情報モデルが、他の領域へ部分的に広がっている [6]。上記のシステムベンダーのコミュニティにおいては、情報モデルはそれほど重要ではない。これらのシステムのほとんどすべては、顧客のプロジェクトを通して開発されたシステムに基づく意味情報モデルを提供し、内部システム規格としてみなされる。いくつかのシステムは、DITA のサポートがある。

「DITA 対システム構造」どちらを使用すべきかについては、世界中の CMS ベンダーとコンサルタントとの間で激しく論争が繰り広げられている。[I] 中立的な技術観点から言えば、どちらの情報モデルをつくるかは、少なくとも中欧市場においては、XML 構造自体ではなく CMS プロセスサポートによる。[II] 今日では、特定の情報モデルは、言語レベル上で標準化されたコンテンツ

よりさほど重要な役割を果たしていない。将来的には、この状況はグローバルな観点とデータ交換とがさらに緊密な関係を持ってくれば変わってくるかもしれない。CMSに部分的に含まれたりサポートされたりしているほかの標準化された情報モデルとしては、DocBook（パブリッシャー構造）[c]、SI1000D（欧州軍事および航空産業）[d]とPI-Mod（Product and Information Model）（ドイツ機械産業）[e]がある。標準化情報モデルに関しては、CMSの実施の約10%の割合を占めている。ソフトウェア産業とCMS機能を持たないアプリケーションの場合においては、その割合はさらに高い[7]。

### ■コンテンツエンジニアリングの方法論

CMSを導入するためには、準備段階で徹底的な計画の立案が必要である。上述のtcworldのCMSの調査研究では、10段階の導入シナリオが考えられている[8]。ITプロジェクト管理が必要であることは言うまでもないが、その計画には社内準備の中心的フェーズとして最も関連性の強い「コンテンツエンジニアリング」作業が含まれる。前述のように、このフェーズには、言語学的規則に基づくコンテンツの標準化作業、モジュール定義およびバリエーションの取り扱い、メタデータと自動化規則の定義、情報モデルの選択、スタイルガイドラインの定義、そして最後にレガシーデータの移行シナリオの決定が含まれる。

これらの作業を行うためには、さまざまな技術とプロセスに関する知識が必要である。したがって、CMSの導入プロセスにおいては、しばしば、外部のCMSコンサルタントまたはCMSベンダーのコンサルティング部門が関わってくる。さらに、情報開発者やコンテンツ技術者として仕事をするための専門的なトレーニングや大学教育を受けたテクニカルライターの人数を常に増員する必要がある。

### ■テクニカルコミュニケーションのためのトレーニング

欧州には、学生にテクニカルコミュニケーションと当該情報管理科目を教えている大学がいくつかある。さらに、専門家向けに民間の団体、大学とtcworldが提供しているトレーニングプログラムがある。CMSに関しては、これらの多くのプログラムにより、以下の分野において、さまざまな学問的、技術レベルに関する教育を受けることができる。

- モジュール構築と再利用シナリオ
- モジュールコンテンツとあらゆる種類のオブジェクトに関する分類

- XML構造化、意味的標準化情報モデル
- バリエーション管理を含むプロセス自動化とドキュメント制作
- グローバル対応のモジュールコンテンツに関する言語学、用語管理および執筆
- 翻訳作業と翻訳プロセス
- データ処理とパブリッシングのためのプログラミング概念と言語
- マルチメディアとウェブパブリッシング

### ■多品種大量生産のための基礎分類

コンテンツモジュールの分類と定義のための重要なメタデータ志向のアプローチは、いわゆる「PI-分類」(Product and Information Classification) [III]方式として定義されている[9]。このアプローチでは、唯一のいわゆる「内部の」メタデータが、当該製品クラス(P)と事前に定義された情報クラスの集合(I)によって関連製品のコンポーネントと結びつく。そして、テクニカルライターは、モジュールコンテンツを(PI-)クラスに割り当て、標準化されたコンテンツ作成のための言語学的規則に従う。この枠組みにおいては、最終製品についてのコンポーネントの目的の用途、または実際の使用を説明する追加のいわゆる「外部の」PI-クラスが存在する。PI-分類の概念によって、データの検索と再利用シナリオと、多くのドキュメント集合、パブリッシングとバリエーション管理に関する多くのプロセスを自動化できる。したがって、この方法は、標準的なモジュールコンテンツから多品種大量生産情報製品を制作するうえでキーとなる。

■知的情報の動的パブリッシングとコンテンツ配信  
ここ数年、テクニカルコミュニケーションのドキュメント視点が、情報の動的配信へと変化してきている。近年注目を浴びているいわゆるコンテンツ配信ポータル(Content Delivery Portals - CDP)は、コンテンツへのトピックベースのアクセスを可能にしている[10]。その意味で、ユーザーは、PDFまたは印刷媒体での膨大なドキュメントを読む代わりに、粒度の細かいレベルでより適切で状況にマッチしたインターネット上でのコンテンツ支援を受けることができる。このような目的に向けて、CDPは、インターネット上の環境に、直接検索、またはファセットナビゲーションなどの異なる検索機能を有している。後者の仕組みでは、CMSからの(PI-)分類によってメリットを広く享受し、検索結果を絞るためにモジュール分類をファセットに直接的にマッピングすることができる。独自の識別可能なコンテンツによって、ソフトウェア製品と組み合わせることで機械は、製品状況



の反応などのエラー処理、またはプロセス制御に関するインタラクティブ情報を直接に発生させたり、または提供したりすることができる。これらのポータルは、デジタル化製品サービスのための「IoT」（Internet of Things）、または「Industry 4.0」構想などの概念である。CDP は、グローバルなインターネット上のインターネットポータルから実施され、ローカルな現場にある機械のポータル、または携帯機器のモバイルアプリケーション、普通のオンラインヘルプシステムに適合したサイズに変換される。

CDP のメリットは、意味情報モデルとモジュールコンテンツの分類によって、情報製品にもたらされる知識であると解釈できる。このような新しい種類の情報は、グローバルに展開する製品の多品種大量生産に対応する高度に構成化が進んだ製品にとっては不可欠なものとなった。

#### ■効率性の証明とコンテンツの関連性

CDP および CMS を将来的に導入した後は、これらのシステムの使用を追跡して管理することになる。CMS には、報告交換 (Report Exchange - Rex)

[IV] 技術によって再利用状況を測定する CMS 用の標準化再利用メトリクス方法が存在する。

REx 測定は、CMS に関する tekomp の研究に記述した 2 つのシステムで利用可能である。効率性の証明は、ドキュメント再利用率とモジュールコンテンツの再利用率によって行われる。システムプロセスを最適化するための主要業績評価指標 (Key Performance Indicators - KPI) による原価モデルも実施可能である [11]。

CDP に関する限り、ウェブ分析ツールを使って、情報が実際にどう使用されているかを追跡することができる。現時点では、コンテンツの関連性または情報の欠落についても、ある分野における情報検索の不成功手法によって知ることができる。続けて、情報機能を向上させ、当該 CMS およびモジュールコンテンツの知識をさらに強化することができる。

#### ■要約

この論文は、中欧、特にドイツ語圏の国々における CMS の現状について説明している。これらの地域においては、多品種大量生産の製品やグローバル化市場に高度に適応できるドキュメントのニーズが高い。このような要求事項は、主に機械産業から成されてきたものであるし、過去 20 年間にわたり CMS 開発に影響を及ぼし CMS を推進してきた。CMS の導入は、しばしば、コンテンツ作成とコンテンツパブリッシングの多くのプロセスの再

検討と再編成の契機となっている。また、複雑なモジュールコンテンツの管理を成功させるためには、CMS 導入前に適切なコンテンツエンジニアリングフェーズが必要である。分類とモジュール化のコンセプトは、CMS を効果的で有益に使用するうえで不可欠なものである。ユーザーのニーズに応じて適応できる知的コンテンツそして製品構成と製品相互の状況的な関係性を提供することにより、近代的なパブリッシングと動的配信のシナリオは、多品種大量生産の製品の価値を高めてくれる。

#### [著者注記 (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](http://www.writec.com/Media_public/IMI_history_and_results.pdf)>
- [3] <<http://www.asd-stel100.org/>>
- [4] <<http://www.unicode.org/>>
- [5] Straub D., Ziegler W. "Effizientes Informationsmanagement durch spezielle Content-Management-Systeme", 3rd edition, tekomp (2014)
- [6] OASIS Darwin Information Typing Architecture (DITA) TC  
<[https://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=dita](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\\_2015\\_DE.pdf](http://www.tekom.de/fileadmin/Dokumente/de/tekom_2015_07_08_Branchenkennzahlen_2015_DE.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\\_Vortrag/CMS\\_II\\_public.pdf](http://i4icm.de/fileadmin/content/HSKA/03_Vortrag/CMS_II_public.pdf)>
- [9] Drewer P., Ziegler W. "Technische Dokumentation - Eine Einführung in die übersetzungsgerechte Texterstellung und in das Content-Management", Vogel Verlag, 2nd ed. (2014)
- [10] Ziegler W., Beier H., "Content delivery portals: The future of modular content" 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/>>

## [著者追加注記]

### I. 本文の解説

ドイツとアメリカのコンサルト間、CMS ベンダーのそれぞれの間で論争がある。

ドイツの CMS ベンダーは、長い間経験を積んだ自分たちの情報モデルがあるので、通常は、DITA を最初に導入することはしない。

### II. 本文の解説

ドイツと中央ヨーロッパでは、CMS を選択する際の最も重要な議論は、システムの機能性であって、本来、システムに内蔵されている情報モデルではないというものである。アメリカにおいては、DITA の情報モデルの視点で決定がなされているというような印象を受ける。すなわち、CMS は DITA をサポートしなければならない、そのうえで、機能が論じられる。

### III. 「PI-分類」 (Product and Information Classification) について

"PI-Classification" は、モジュールコンテンツの生成のためのメタデータの適切な定義のために著者が開発した手法である。これは、多くの CMC システムで利用され、実施されている。

### IV. 報告交換 (Report Exchange -Rex) について

これに関しては、以下の URL でその背景を知ることができる。

<http://www.tcworld.info/rss/article/content-intelligence-for-content-management-systems/>

簡単に言えば、CMS の再利用の効果を測定するためのツールである。

## [訳者注記]

### a. "Publishing" の和訳について

"Publishing" を「出版」と訳すことは、本論文の主旨にそぐわないと判断。日本における出版という言葉は、籍 や雑誌などの出版業における行為を強く想起させるためである。そこで原文を尊重してあえてパブリッシングと表記することにした。本論文で言うパブリッシングは、製造者の法的責務に基づいて行われているデータ変換等による情報生成、印刷、これらを想定対象者に配布・配

信する行為を広く意味している。出版業における行為との差異の 1 つ目は、提供される情報はその説明対象である製品やサービスの一部とされ、提供情報単独で商品として有償提供 されることは国際規格や国内法制度上も業務実態においても想定されていないことである。2 つ目の差異 は、提供する情報やその提供形態に対して国際規格や法令等による干渉（産業区分や法域によっては規制）があり、説明責任の履行状況のエビデンスとして法的役割を担うことにある。

以下の訳者注記は、訳者が著者に質問をして得た回答に基づく。

### b. 集合規則 (Aggregation rules)

ドキュメントは、モジュールコンテンツの収集 (アSEMBル) によって作成される。このプロセスは、"Aggregation (集合)" とも呼ばれる。(これは、"assembly (集合)" と同義語である。) 近年のシステムでは、この集合のプロセスは、自動化されている。

そのため、モジュールコンテンツをある連続性と階層構造に基づいて収集できるよう、ある (ロジカル) ルールが必要になる。

このようなルールが、"Aggregation rules": 「集合規則」である。

もちろん、執筆者がモジュールを手作業でドキュメントに反映する場合でも、あるルールが必要である。ルールは、しばしば、標準と記述される製品の構造に依存する。

### c. DocBook (パブリッシャー構造)

過去においては、出版業界を中心に "DocBook" と言われる情報構造が頻繁に用いられた。これは、XML の情報モデルに基づくもので、あらゆる種類の書籍の執筆および印刷のために O'Reilly 社が定義をしたものである。基本的に、印刷用のソフトウェアマニュアルに使用することを想定して開発されたこともあり、今日に至るまで、ソフトウェア業界におけるマニュアル作成用に使用され続けている。

### d. S1000D (欧州軍事および航空産業)

S1000D (European military and aviation industries) は、「欧州軍事および航空産業」に属する組織によって開発された。同時にこの産業部門が、この情報構造の主要ユーザーでもある。欧州の軍隊と航空産業におけるサプライヤーとベンダーは現在も "S1000D" の構造を使って情報交換をしている。

### e.: PI-Mod (Product and Information Model) (ド

## イツ機械産業)

"PI Model (German machinery industry) "の元となる情報モデルの"PI-Mod"は、LIEBHERR 社によって開発された。その開発の実際はProf. Zieglerが、1997年に手掛けたものである。PI Model のユーザーは、主に機械産業と製造産業の関係企業である。PI Model は、多くの CMS に採用されると同時に、SIEMENS のような企業によって活用されている。PI Model は、市場にはそれほど普及はしていないが、CMS の重要な機能性の向上に影響を与えたとされる。

訳者：(株) 情報システムエンジニアリング (ISE)

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 adaptations 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 tekcom 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 tekcom 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 tekcom 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 automatization 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 automatization 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 faceted 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 these 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 tekomp 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.