

# インテリジェントコンテンツ配信のための セマンティックインフォメーション（意味情報）開発

Semantic Information Development for Intelligent Content Delivery

Karlsruhe University of Applied Sciences, Germany

Prof. Dr. Wolfgang Ziegler<sup>1</sup>

大阪大学 CO デザインセンター 招聘准教授

黒田 聡<sup>2</sup>

Satoshi Kuroda

情報受領者の最近の要求に応えるために、知的情報および動的情報提供の概念に導く新しい技術的および方法論的試みがなされてきた。実証データによれば、提供された技術情報の大部分は、依然として紙または電子的な PDF 対応物に基づいている。しかし、同時に、ChatBot、人工知能 (AI) 駆動アプリケーション、インテリジェント検索システム、および製品固有の（構成依存の）情報プロビジョニングを含む、よりインタラクティブでダイナミックなコンテンツにおけるコンテンツの配信が急速に増加している。このような情報サービスを開発・提供することは、情報の作成・管理に直接的な影響を与える。本稿では、インテリジェントな情報開発プロセスによって、ダイナミックな製品開発とモデルベースのコンテンツ作成をどのように統合させることができるかを論じる。<sup>3</sup> また、ドイツと日本の本テーマにおける取り組み状況の違いを整理し、日本のテクニカルコミュニケーターが取り組むべき課題の特定と、その解決のためのアプローチを考察する。<sup>4</sup>

## 1. はじめに

情報マネジメントの観点からは、第 1 に、Web ベースのコンテンツやアニメーション、3D オブジェクト、動画、ChatBot（訳注：自動会話プログラム）などのインタラクティブメディアを含む、より電子的でダイナミックなメディアに向かう傾向が強くなっている。このことは、ドイツ語を話すヨーロッパ市場に関する最近の研究データ[2]によって証明されている。本稿ではこの情報源を参照する。

第 2 に、テクニカルコミュニケーションにおける標準化されたコンテンツ作成から、より精巧な情報モデリングアプローチと情報技術へと、さらに進化している。セマンティックモデリングは、主に情報の品質を向上させ維持するために、内容構造のために以前から使用されてきた。その後、意味論的アプローチが意味論的メタデータモデリングに向けて拡張され、今日ではオントロジーモデリングと情報の意味論的ネットワークに至ってい

る[3,4]。これらの技術は、コンテンツ制作プロセスと組み合わせた製品の複雑性のモデリング、企業内での情報交換のためのセマンティックミドルウェアとしての情報ハブの構築、またはコンテンツ検索プロセスの強化とそれに対応する検索結果の精度の向上のために使用される。さらに、検索および配信のプロセスと機能を強化するために、AI が応用される。

以下では、最近のコンテンツマネジメント・配信の状況を、技術的・方法論的観点から簡単に概観する。また、デジタル化のユースケースとして、また、コンテンツ配信技術のビジネスモデルとして、デジタル情報サービスの創出・実装にこれらのシステムを活用することについても、一般的に検討する。我々はまた、テクニカルコミュニケーションの将来において、より考慮されなければならない、拡張された AI のような、より精巧な側面についても検討する。

## 2. コンテンツマネジメント

コンテンツマネジメントシステム (CMS) は、技術的コミュニケーションの十分に確立された技術的バックボーンを構築する。それらは、自動車産業や航空機産業のような関連分野だけでなく、機械産業やそのサプライヤーにも重点を置いて、ドキュメント部門の内部プロセスを最適化するために使用されてきた。CMS は、製品のバリエーション、製品の変更、およびグローバル化から生じる製品ドキュメントの複雑さを解決し、対処しようとする。この目的のために、それらは、トピックベースのコンテンツ、コンテンツバリエーション、バージョン管理、およびたとえ、ネストされた（訳注：入れ子構造とされた）再利用可能なオブジェクトの粒度レベルでの翻訳マネジメントおよびデータベース処理ならびにそれらの関連するコンテンツライフサイクルをサポートする洗練された機能のような方法論および技術に依存する。コンテンツ構造は、XML ベースの情報モデルのセマンティクス（意味論）に基づいている。セマンティックコンテンツの他にも、多くの企業は、製品ライフサイクルに応じた変更管理や変更管理として、あるいは部品リストによって駆動される自動化されたドキュメント生成として、洗練されたコンテンツマネジメントプロセスに依存している。

## 3. セマンティックメタデータ

コンテンツは、CMS のプロセスを自動化するために、そして後に他のアプリケーションにおいて、より「インテリジェント」になり得る。この目的のために、テクニカルライターはセマンティック（意味）メタデータを用いてトピックや他のコンテンツオブジェクトを充実させなければならない。セマンティックメタデータを定義するための広範な概念の一つは、CMS ユーザーの 10%以上がすでに使用している PI 分類[3,4,5]の方法と、この[6,7]から導出された標準である。この方法論の基本的な実装では、再利用可能な情報単位（トピックまたはコンテンツ断片）は、基本メタデータの 4 象限で構成される仮想情報空間でアドレス指定することができる。それらは、一意に定義された内容プロパティと、トピックの潜在的な用途を二次元で記述する。これらは、製品と情報の 2 つの次元に追加的に割り当てることができる。

上記または比較可能な分類方法を適用することで、トピックをユニークな方法で扱うことができ、たとえば、自動化されたドキュメント集約をサポートすることができる。これは、たとえば、内因性分類セレクターのプレースホルダー構造を有するドキュメントテンプレートによって達成することができる。最近の研究では、CMS ユーザーの約 27%

がすでにこの自動化方法を用いて、各バリエーション固有のドキュメントを作成しており、約 44% がメタデータベースのフィルタリングメカニズムを使用していることが示されている[2]。これまでのところ、PI 分類のすべての分類ブランチと標準メタデータの次元は、企業固有の集合で示され、通常は、それぞれの製品クラス、情報クラスの階層で示されることに留意することが重要である。PI 分類および相関アプローチは、トピックのサイズ、含まれる情報、および内容の区切りに関して、テクニカルライターのための簡潔なスキームを提供することを意図している。最近の産業用途では、製品やトピックのさらなるバリエーション特性、いわゆる機能的メタデータを記述する PI 分類のさらなる拡張が求められる。このことは、たとえば、サービスプロセスにトピックを接続するのに有効である。これを考慮すると、機能的メタデータは、コンテンツをエラーコードやシステムイベントにリンクさせたり、作業時間やツールにリンクさせたりするような概念をカバーすることができる。これらの拡張されたメタデータはすべて、ソフトウェアとハードウェアの複雑な構成が可能なシステムの処理を記述するために必要であるかもしれない。

## 4. コンテントデリバリー

ほとんどの場合、CMS からの製品ドキュメントは、法的要件および国際規格を遵守するために、少なくともエンドユーザーのために、過去から広く作成されている。ドキュメントは、非常に効率的な再利用のシナリオによって作成され、古典的なマニュアル（使用情報）として PDF フォーマットで、あるいは、オンラインヘルプフォーマットや CD のような記憶媒体上のリンクされた HTML のような多かれ少なかれ静的な電子ドキュメントとして提供されてきた。

一方で、複雑な製品をインストールし、保守するための構造化された、詳細な、特定の情報に対する技術者の必要性は、しばしば無視されてきた。より一般的には、インタラクティブでより状況に応じた情報提供のためのあらゆる種類のターゲットグループ（訳注：ターゲットオーディエンス）およびユーザーの要件は、長い間満たされていなかった。

これは近年変化し始めている。現代の Web 技術やシステムアーキテクチャーに合わせて、コンテンツはいわゆるコンテンツ配信ポータル (CDP) によって配信されるようになった。CDP の基本的な定義は、まず[8]で「システムモジュール化されたコンテンツおよび集約されたコンテンツまたは他の情報タイプへのウェブベースのアクセスを、さ

<sup>1</sup> 1, 2, 3, 4, 5, 6, 7 を執筆。提供原文は英語であるが (株) 情報システムエンジニアリング秋山素夫氏の協力を得て日本語化した。

<sup>2</sup> 8, 9, 10 を執筆。

<sup>3</sup> 本稿は、ETLTC 2019 の会議議事録[1]への寄稿の縮小、修正、一部更新版である。Prof. Dr. Wolfgang Ziegler

<sup>4</sup> 一般財団法人テクニカルコミュニケーター協会の実務技術研究会およびワーキンググループ活動の主旨を担う立場で考察した。黒田聡

さまざまなユーザーグループに提供するシステム』として与えられている。

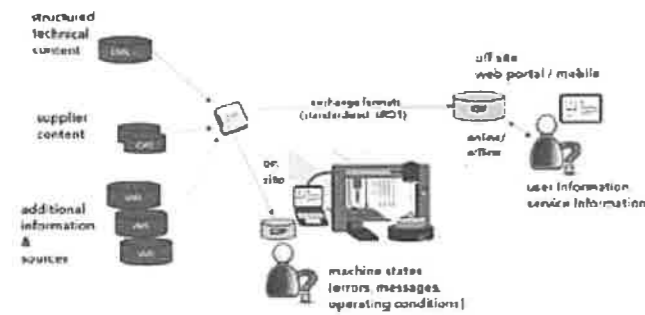


図1. 構造化されたコンテンツソースと構造化されていないコンテンツソース、オンサイト/オフサイトアクセスを含む典型的な CDP 設定

## 5. コンテンツアクセス

CDP および CDP の利用者は、CMS および割り当てられたメタデータに由来するコンテンツが細分化されるトピックベースの恩恵を受けることができる。前述したように、このことは主に CMS におけるドキュメント作成と検索に使用されてきた。今日では、3つの基本的なアクセスタイプが定義できる。

- ・複雑なクラシフィケーションタクソノミーまたは CMS からのより単純なメタデータセットに対応するファセットおよびフィルターによる構造化された検索
- ・ネストされたトピックの古典的なドキュメント構造に沿ったナビゲーション
- ・キーワードやフレーズによる全文データベースインデックスに基づく直接検索

もちろん、対応する検索機能は、基本的なアクセスタイプに基づく結果を提供するが、しばしば、独占的に人間が駆動する、すなわち、手動で実行されるアクセスタイプとみなされる。しかし、CDP は、要求に応じて配信される Web サービスのミドルウェアシステムとしても実装することができる。このアプローチを用いて、コンテンツは、多くの方法で、また多くの使用ケースで、手動または Web サービスを介してアドレス指定できる。対応アプリケーションは、エンジニアリング、生産、製品監視、およびサービス管理からの最近の IoT および Industry 4.0 アプローチに適用されている。コンテンツは、イベント、実際のマシンの状態、および動作条件に応じて、ユーザー（技術者またはエンドユーザー）に配信することができる。このような場合、関連する内容の識別は、標準および拡張（PI）分類によって記述される情報空間におけるトピックの簡潔な位置に決定的に依存することは明らかである。

コンテンツにアクセスするために使用されているデバイスについては、既知のタイプの応答性 Web インターフェイスまたは他のモバイルアプリケー

ションがすべて利用可能である。モバイルデバイス上のアプリケーションの記憶容量が限られているため、同期メカニズムを含むオフライン使用には制限が適用される。したがって、オフライン機能は、特定のベンダーによってのみ提供される。その他のアプリケーションは、セキュリティ上の理由により、ローカル・サービスとしてオンサイト・マシンのみで実行される。

企業内では、テクニカルコミュニケーションとドキュメンテーション・ドメインにおける潜在的に構造化されたコンテンツの量は、構造化されていないコンテンツの量よりもはるかに少ない。構造化されたコンテンツは、要求された、検索された、そして配信されたコンテンツの実際のユースケースを定義し、直接検索プロセスの重みと重要性、および AI 技術対構造化ファセット検索（訳注：ファセットとは切り口、軸を意味する）による可能な強化を定義することを意味する。CDP ユースケースを設定するための追加の決定基準は、オフライン機能、またはさまざまなメディアデータを含むコンテンツオブジェクトの交換、インポート、またはクローリングのための標準化されたデータ処理の容易さである。別の特殊なユースケースは、現時点では、複合現実（訳注：MR。本稿では VR、AR を含む）アプリケーションによって行われる。オブジェクト認識は、種々な異なるユーザーインターフェイスおよび技術によって実行することができる。PI 分類とその標準では、オブジェクト認識は本質的に定義されたコンポーネントを定義するが、ユーザーによる特定のユースケースまたはマニュアル選択は、必要な情報タイプを定義する。これにより、Web リクエストが CDP に送信され、Web サービスを介してコンテンツがインターフェイスに配信される。この場合、CDP に直接アクセスすることはできないが、他の環境やサービスに統合できる。

## 6. インテリジェンス・カスケード

インテリジェントコンテンツの概念は、近年盛んに議論されており、コンテンツ配信ポータルや関連アプリケーションにおける技術的な利用事例が見出されている。前節ではビジネスユースケースに焦点を当てた。テクニカルコミュニケーションの文脈における「インテリジェンス」は、主に以下に基づいていることが広く受け入れられている。

- ・粒状コンテンツ
- ・意味的（XML-）構造化コンテンツ
- ・コンテンツに割り当てられたセマンティックメタデータ

もちろん、標準化された、理解できる書き込みとローカライズから（Web）設計とメディア統合ま

で、テクニカルコミュニケーションの他のあらゆる側面は、依然として基本的に重要である。しかし、コンテンツ配信および関連情報サービスの観点からは、情報空間における検索性とアクセス性、ならびに情報システムによる処理性が最も重要な側面である。最近の進展は、これらの重要な要因を保証し、改善するためにさらなる措置をとったことである。インテリジェンスのカスケード（訳注：数珠つなぎに連結された構造、連鎖的あるいは段階的構造を意味する）という点では、次のようなアプローチを見つけることができる。

- ・ネイティブ・インテリジェンス
- ・知能（Intelligence）の増強
- ・人工知能（AI: Artificial Intelligence）

ネイティブ・インテリジェンスは、CMS または他のオーサリングシステムにおけるメタデータによる分類されたコンテンツの使用と実装をカバーする。議論された方法は、PI 分類のように、ファセットとして直接表示することができ、または CDP で他の方法で処理することができるメタデータを開発し、使用するための簡潔なスキーマにつながる。その結果、最近の CDP は、深いリンクまたは通信 Web サービスに接続された手動または機械誘導検索のためのデジタル情報サービスとして作用することができる。

強化されたインテリジェンスは、セマンティックネットワークのオントロジーモデリングとして近年注目されている。これらのネットワークは、たとえば以下のようなさまざまな方法で使用することができる。

- ・製品の物理モデルの論理から（CMS-）分類を導き、コンテンツの作成をサポートする
- ・コンテンツ間の意味的にモデル化された関係、すなわちユースケースに応じた検索結果の質の改善を用いた CDP における検索と検索の拡張
- ・ドメイン間やデータソース間の情報を管理し、マッピングメカニズムやドメイン間の検索を可能にするセマンティックミドルウェアの実装

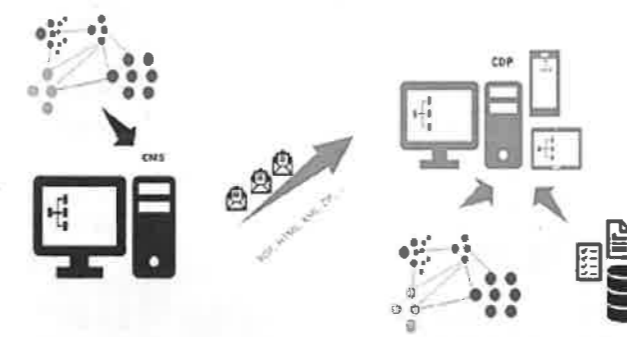


図2. CMS でのコンテンツ作成や CDP でのコンテンツ配信をサポートするオントロジーのアプリケーション領域

最初の2つのケースは図2に描かれており、CMS

（階層メタデータを含む）または CDP（ファセットを含む）のオントロジーのサポートを意味している。最近の実装は、通常、これらの側面の一方に焦点を当てているが、双方を支持するための一般的な制限はない。

オントロジーから導かれる将来のユースケースは、おそらく、意味的にリンクされたトピックのより動的で論理的に駆動される構成であろう。情報要求は、接続されたタスクまたはワークフロー、または他の製品関係としての限定されたコンテキストを伴う特定の問題に適した「マイクロドキュメント」とみなされ得るメソスコピック構造（訳注：マクロとミクロの間に位置づけられる構造を意味する）の事前定義されたセマンティックモデルから成果を得ることができる。これはまた、モノリシックドキュメント（訳注：全体が1つのモジュールでできていて、分割されていないドキュメント）と単一トピック配信との論理的ギャップを埋めることになる。これらは、トピック間の意味関係で説明される論理パターンとして、または情報源間のより一般的な論理パターンとして、事前に定義することができる。論理パターンは、単に静的マイクロドキュメントとして、またはサーチプロセスによって引き起こされるオントロジカル関係の観点から定義することができる。より高い自動化レベルでは、それらは、ユーザー行動の分析（過去の検索プロセス）によって導出されるか、あるいは、基礎となる製品モデルの変化から自動的に導出されることさえある。

AI はまた、テクニカルコミュニケーションおよびコンテンツ関連技術にも普及している。与えられた文脈において、AI は、レガシーデータまたは非構造化データの自動化された内容分類に使用できる[8]。このような場合には、高品質の分類された、したがってネイティブのインテリジェントコンテンツを含む一連の教師データが必要である。さらなるアプリケーションは、コンテンツ生成、すなわち、（CDP-）要求に対応する訓練データから動的に生成されるコンテンツの配信を担う。フロントエンド側では、オブジェクトおよび音声認識のような AI 技術を含み、要求されたオブジェクトおよび情報タイプを識別し、検索プロセスを容易にすることができる。

このことは一般に、インテリジェンス・カスケードのすべてのレベルに当てはまる。情報システム会社やソフトウェアベンダーは、これらのレベルの1つ以上をサポートするために、その方法論とアーキテクチャーを再考しなければならない。

## 7. 概要と展望

最近の技術と情報アーキテクチャーに基づくコンテンツマネジメントとコンテンツ配信に焦点を当てた。著者らは、このコンテンツ関連情報システムの主要な方法と、構造化コンテンツのセマンティックモデル化と分類メタデータを含むインテリジェントコンテンツの関連概念を説明した。デリバリー面では、コンテンツアクセスの種類や、さまざまなテクニカルユースケースに焦点を当てた。また、このような情報シナリオの最近の段階では、CDPからのインテリジェントで高度に検索可能なコンテンツは、インテリジェンス・カスケードの多様なレベルに対処するときに、より高いユーザビリティのための情報サービスを提供し、より効率的に使用できることを示した。このカスケードは、ネイティブ・インテリジェンス、増強インテリジェンス、およびAIの観点から説明できる。実際の実装では、これらの多様なアプローチはすべて、概念的には、製品、関係、プロセスおよび必要な情報空間の量に関する深い知識を必要とする。そのためには、商品開発や情報開発におけるグループ間の連携やコミュニケーションも必要である。一方で、技術コミュニケーションにおける情報管理者の学術的・専門的教育は、これらの新しい分野とレベルの知的情報アーキテクチャーをカバーしなければならない。

## 8. 日本における状況

日本におけるセマンティックインフォメーションの開発に関する取り組みの状況は、ドイツとは異なっている。

一般財団法人テクニカルコミュニケーション協会が公開している技術体系には、本稿の対象となる知識が取り込まれていない（構造化技術の基本理念と用語解説程度に留まる）[9][10]。著者は、2018年に情報処理学会における研究報告の一端として、テクニカルコミュニケーションの知識・技能の動向調査を行った。ここで、2010年以前と2011年後においてテクニカルコミュニケーションシンポジウムで採りあげられた内容に差異があることを見出ししている[11]。

表1の備考に「追加項目」として示す項目が2011年前後から採りあげられるようになった話題である。

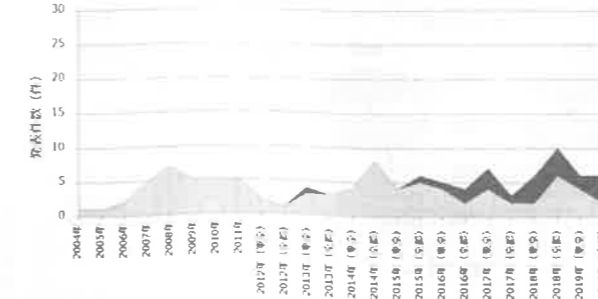
本稿の対象となる知識は「情報デザイン知識と技能」に属する。本稿の1~7で採りあげられている話題は、情報アーキテクチャー（ネストされたトピックの古典的なドキュメント構造に沿ったナビゲーション）、構造化とモジュール化（意味的構造化コンテンツ）、タクソノミー、クラシフィケーション、オントロジーに該当している。

表1. テクニカルコミュニケーションに求められる知識・技能

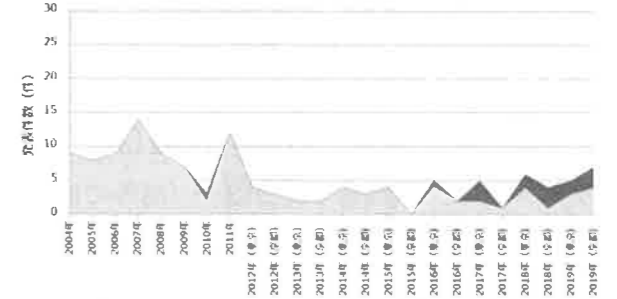
6つの区分	要素	備考
コンプライアンス	製造責任, 製品安全, 消費者保護, 労働安全, 環境保護, キュア&ケアにISO9000, ISO9001, ISO13485 (キュア&ケア対応のため)に基づくQMSに関する知識関する国内法令および標準規格の理解	「キュア&ケア」を追加
	仕向地の法令および標準規格の理解	
	異文化理解	
	校閲力	追加項目
情報デザイン知識と技能	情報アーキテクチャーの知識と技能	
	構造化とモジュール化の知識と技能	
	タクソノミー, クラシフィケーション, オントロジー	追加項目
	表現設計の知識と技能	
	企画書・設計書・または構成案にまとめる能力	
	プロジェクト管理と予算管理	
プロジェクト管理と遂行	情報開発工程の知識と技能	
	テクニカルライティングの知識と技能	
	ユーザーエクスペリエンスライティングの知識と技能	追加項目
	制作ツールの知識と技能	
	翻訳ツールの知識と技能	
	ISO9000, ISO9001, ISO13485 (キュア&ケア対応のため)に基づくQMSに関する知識	「ISO13485 (キュア&ケア対応のため)」を追加
IT基礎知識	コンピューター技術知識	
	データベース技術知識	追加項目
	ソフトウェア開発, プロセス知識	追加項目
	プログラミング知識	追加項目
ユーザー指向	ユーザー理解の知識と技能	
	インクルーシブデザインの知識と技能	追加項目
	ユニバーサルデザイン, アクセシビリティの知識	
	行動観察, ユーザビリティ評価の知識と技能	追加項目
	分析と解析の知識と技能	追加項目
	ユーザーエクスペリエンスのデザインとディレクションに関する知識と技能	追加項目
ニーズの掘り起こしと部門間コミュニケーション	ニーズの掘り起こしの知識と技能	「ニーズの掘り起こし」を追加
	ユーザー調査, 市場調査の知識と技能	追加項目
	異文化コミュニケーション	追加項目
	技術仕様書の読解力と理解力	
	情報収集と整理の知識と技能	

図3は、テクニカルコミュニケーションシンポジウムで話題となった知識・技能の動向を示す。2018年の結果[11]に2019年分を追加している。2010年以前から採りあげられてきた話題の数を淡色で示し、2011年前後に採りあげられるようになった話題の数を濃色で加算して示している。知識・技能の6つの区分のなかで、本稿の対象である「情報デザイン知識と技能」が話題となった回数は少ない。2011年以降にむしろ減少していることが示されており、2011年後に採りあげられるようになった話題を加えても横ばいである。日本でもオーサリングシステムに(DTPではなく)CMSを利用する企業は増えている。この動向と関わりなく、その運用に必要な要素技術が語られることが極めて少ないまま、長年推移している。日本のテクニカルコミュニケーションが、理論的課題よりも実務的課題を好むことは、経験を通じて著者は把握している。実務運用しているCMSの利用に必要な知識・技能が調査に反映されていないのは、調査対象としたシンポジウム(公の場所)で本稿の話題が採りあげられていないことを意味していると受けとめるべきであろう。延べ聴講者数5000名規模の業界最大のイベントでこの結果が示されていることは、考察に値する。このデータは、CMS運用に必要な知識が、基本から実践に至るまで、導入した特定CMSソリューションのみを対象とする技能として、企業内に閉じた状態で教育訓練が行われていることを示唆する。テクニカルコミュニケーション技術として汎化された知識・技能体系としてではなく、業務システムの運用技能として扱われていることを意味する。育成の担い手はCMS提供ベンダーであることも推測できる。特定ソリューションに依存しない一般的な技術としての研究、検討、育成が広く行われているドイツとは、日本は状況が異なる。IT基礎知識が話題になっていないことも含め、日本のテクニカルコミュニケーション教育の特殊性として捉える。

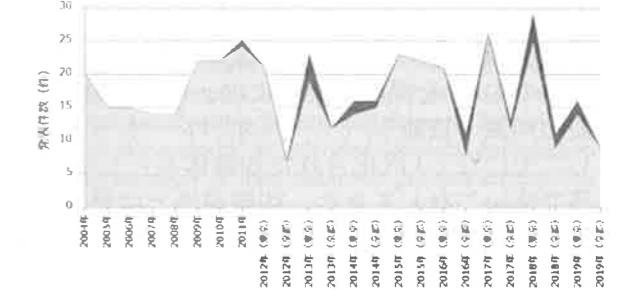
コンプライアンス



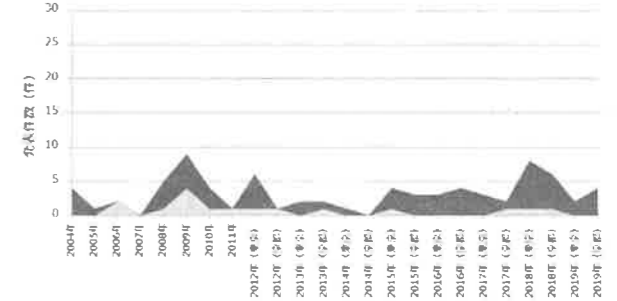
情報デザイン知識と技能



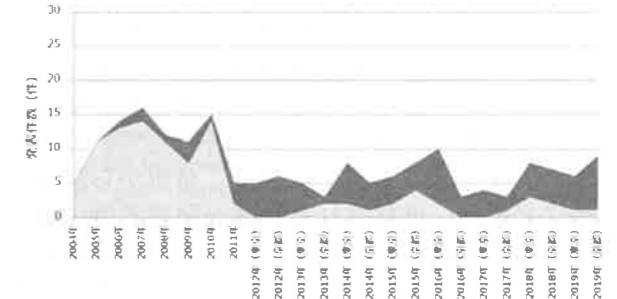
プロジェクト管理と遂行



IT基礎知識



ユーザー指向



部門間コミュニケーション/ニーズの掘り起こしと部門間コミュニケーション

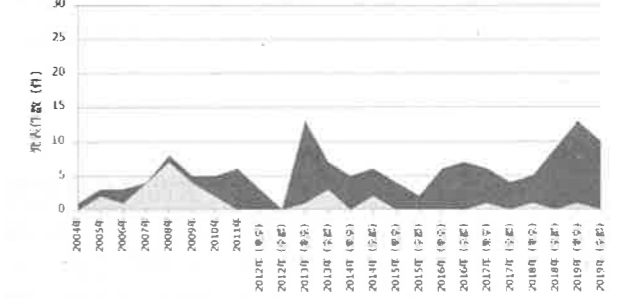


図3. 年別発表件数推移グラフ

## 9. 日本における課題

導入した CMS の運用に必要な知識としてのみ、本稿で話題としている技術が教育訓練されていることには、懸念がある。

ひとつは、テクニカルコミュニケーターという職業の本質に関わる懸念である。テクニカルコミュニケーターは、CMS オペレーター（海外ではテクニカルオパーサーと呼称される）ではない。製品・サービスの提供側が負う責務を関連部門と共に分担し、Information Products (IEC/IEEE 82079-1:2019 上の呼称。製品・サポート情報を意味している)の企画からその維持管理、継続的改善までを総合的に担う職業である。かつてオーサリングシステムとして DTP を採用した際には、その運用管理に必要な知識・技能を（ツールオペレーション技術としてではなく）汎化された知識体系として学び、業務で使いこなしてきた。印刷業界とは異なる使いこなしノウハウや、独自の標準制作工程の定義はこの前提があって生み出されてきている。

もうひとつは、次世代のインテリジェントコンテンツ配信の実現、セマンティックインフォメーションの開発には、本稿で採りあげる知識と技能が欠かせないことだ。この知識なくして次世代の Information Products の企画はできない。運用操作ができることと、技術原理を理解していることは別義である。この違いに起因する支障は今後表面化していく。

その兆しはすでに現れている。テクニカルコミュニケーションシンポジウムでも未来のコンテンツのあるべき姿は盛んに話題となっているが、体系的な技術教育が行き渡っている DTP 的オーサリングシステムを前提とする発想からなかなか抜け出せていない。業界外から登壇者を招いても、ドキュメントを前提としていないその提言を咀嚼し、自発的に業務に応用することに苦勞している。これは、ドキュメント文化に基づく発想から脱するために必要な、汎化された知識が聴講者側に足りないためだ。掲げる未来ビジョンと、実運用してきたオーサリングシステムを分離することが、日本のテクニカルコミュニケーターには困難となっている。結果として、使用しているツールに縛られた改善に留まる傾向が顕著であり、急速に非ドキュメント化とその活用が広がっている情報利用者側との乖離が大きくなっている。

## 10. 展望

テクニカルコミュニケーターは、製品開発者が開発のために作成する技術仕様書には担えない、利用者本位の情報の作成と提供をその存在意義としてきた。その価値を認められているがゆえに、技術仕様書を作り替える業務に予算が割り充てられ

てきている。

しかし、現代の情報利用者の意識と環境は、ドキュメントというパッケージ化された情報提供形態から乖離し始めた。本稿 1~7 で提示されたように非ドキュメント形態での情報配信と利用にどう答えるかが、これからのテクニカルコミュニケーターは求められている。従来のドキュメント、特に目次構造やその配列が担っていた役割を、非ドキュメント形態での配信や利用でどのように担わせるのかを考案しなければならない。

この研究と実践では、情報デザインの知識と技術を深く理解し、応用することが必要だ。このためには、特定 CMS の運用知識ではなく、テクニカルコミュニケーション技術体系に属する汎用的な知識として学び、研鑽することが欠かせない。教材となる文献が現時点では JIS X 0154:2018 [12]しか存在しない状況も改善が必要であろう。一般財団法人テクニカルコミュニケーター協会では、この課題を解消するために、実務技術研究会とワーキンググループを組織して、下記取り組みを進めている。

- ① オーサリングシステムとして、DTP と CMS の両方を想定した標準制作工程の定義
- ② 新たな標準制作工程に基づく人材育成メソッドおよびコンテンツの策定
- ③ 非ドキュメント形態を想定した Information Products の評価技術として解析技術の研究
- ④ 次世代コンテンツマネジメントのためのメタ情報技術の研究

図 4 は、①の成果として定義した標準制作工程である。ソフトウェア開発との親和性を重んじて、国際規格となっている V モデルに基づいている。

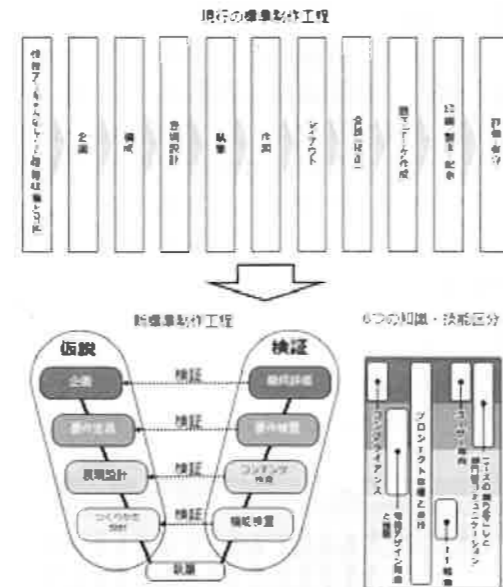


図 4. 標準制作工程の見直し

## [参考文献]

- Ziegler W. Drivers of Digital Information Services: Intelligent Information Architectures in Technical Communication, 2019 ACM Chapter Proceedings on Educational Technology, Language and Technical Communication, ISSN: 2434-9526, (2019) VOLUME 1, S. 48-52 [1]
- Straub D., Ziegler W. tekomp CMS Studie 2018/19, tcworld Stuttgart, Germany [2]
- Ziegler, W. 2017. Metadaten fuer intelligenten Content. In Schriften zur Technischen Kommunikation, Vol 22, p. 51-66, tcworld [3]
- Ziegler, W. 2017. The Evolution of Content Management towards Intelligent Delivery Systems for Technical Communication. In Frontier, Official Journal of Japan Technical Communicators Association JTCA, p. 68-75. [4]
- Ziegler W. 2017. Basic Concepts Which Support the Management and Delivery of Intelligent Content. In Frontier, Official Journal of Japan Technical Communicators Association JTCA, p. 84-91 [5]
- Parson U., Sapara J., Ziegler W. 2017. iiRDS for Technical Writers – Introduction to the Metadata. In Conference Proceedings of the Annual tekomp Conference. [6]
- European Association for Technical Communication - tekomp e.V. 2017. iiRDS Specification - intelligent information Request and Delivery Standard - First Public Working Draft, 03 April 2017. (2017). [7]
- Ziegler W. 2013. Alles muss raus! Content Delivery fuer Informationsportale, In Proceedings of the annual tekomp conference, p. 47 [8]
- トリセツのつくりかた 品質追求編 2015 (新編集版) 一般財団法人テクニカルコミュニケーター協会編著 [9]
- トリセツのつくりかた 制作実務編 2010 一般財団法人テクニカルコミュニケーター協会編著 [10]
- ドキュメントコミュニケーションの4つのリデザイン: AI時代に向けて 情報処理学会研究報告 IPSJ SIG Technical Report Vol.2018-DC-111 No.6 2018/11/26 黒田聡、池田光穂 [11]
- システム及びソフトウェア技術 - 製品ライフサイクル、利用者及びサービスマネジメントの文書化のためのコンテンツ管理 JIS X 0154:2018 (ISO/IEC/IEEE 26531:2015) [12]

[Prof. Dr. Ziegler からの英原文]

## Semantic Information Development for Intelligent Content Delivery

Wolfgang Ziegler, Karlsruhe University of Applied Sciences, Germany

In order to meet the recent requirements of information recipients, there have been new technological and methodological attempts leading to the concept of intelligent information and dynamic information delivery. Empirical data show, that the majority of delivered technical information is still based on paper or their electronic pdf-counterpart. But at the same time, delivery of content in more interactive and dynamic content is increasing rapidly including for example chatbots, AI-driven applications, intelligent search systems and product-specific (i.e. configuration-dependent) information provisioning. Developing and offering such information services to customers and other user groups has direct consequences on the creation and management of information. In this article<sup>1</sup>, we will discuss how dynamic product development and model-based content creation can be aligned by intelligent information development processes.

## INTRODUCTION

From the perspective of information management, there are two major developments: First, there is a strong tendency towards more electronic and dynamic media, including web-based content and interactive media like animations, 3D-objects, movies or chatbot interactions. This can be proven by data of a recent study on the German speaking European market [2]. We will refer to this source throughout the article.

Second, there is a boost from standardized content creation in technical communication towards more elaborated information modeling approaches and information technologies. Semantic modeling has been used before for content structures, mainly to increase and maintain information quality. Semantic approaches have then been extended towards semantic metadata modeling, leading nowadays to ontology modeling and semantic networks of information [3,4]. These technologies are hereby either used for modeling the complexity of products in combination with content creation processes, for building information hubs as a semantic middleware for information interchange within the companies or for the enhancement of content search processes and corresponding increased precision of search results. Finally, applications of artificial intelligence are included in these technologies to enhance search and delivery processes.

In the following, we will take a brief look at the recent situation of content management and delivery from a technical and methodological point of view. Additionally, we will generally discuss the use of these systems for creating and implementing

<sup>1</sup> This article is a reduced, modified and partially updated version of the contribution to the ETLTC 2019 conference proceedings [1].

digital information services as use cases of digitization and as business models of content delivery technologies. We will also consider more elaborated aspects like augmented and artificial intelligence which have to be given more consideration in the future of technical communication.

## CONTENT MANAGEMENT

Content management systems (CMS) build a well-established technological backbone of technical communication. They have been used to optimize the internal processes of documentation departments with a strong focus on machinery industry and their suppliers as well as related areas like automotive or aviation industries. CMS try to solve and handle the complexity of product documentation arising from product variants, product change and globalization. For this purpose, they rely on methodologies and technologies like topic-based content, content variants, version management and sophisticated functionalities supporting, for example, translation management on granular level and database handling of nested reusable object and their correlated content lifecycle.

The content structure is based on the semantics of XML-based information models. Beside semantic content, many companies rely on sophisticated content management processes as variant management and change management depending on product lifecycles or even automated document generation driven by parts lists.

## SEMANTIC METADATA

Content can become more "intelligent" in order to automatize processes in CMS – and later in other applications. For this purpose, technical writers have to enrich topics and other content objects by using semantic metadata. One of the spreading concepts for defining semantic metadata is the method of PI-classification [3,4,5] used already by more than 10% of the CMS-users and standards derived from this [6,7]. In a basic implementation of this methodology, reusable information units (topics or even content fragments) can be addressed in a virtual information space made up by four dimensions of the basic metadata. They describe by two dimensions the uniquely defined content properties and the potential usages of the topics. They can be assigned additionally to two dimensions of products and information.

Applying this, or comparable classification methods, makes it possible to address topics in a unique way and supports, for example, automated document aggregation. This can be achieved, for example, by document templates with a placeholder structure of intrinsic classification selectors. Recent studies show, that approximately 27% of CMS users already use this automatization method to produce variant specific documentation and almost 44% use metadata-based filtering mechanisms [2]. It is important to note that, so far, all classification branches and standard metadata dimensions of the PI-classification are denoted by a company-specific set and usually by hierarchies of product classes, respectively information classes.

The PI-classification and correlated approaches intend to provide a concise scheme for writers in relation to the sizes, the contained information and content delimitations of the topics. In recent industrial applications there are even more dimensions of an extended PI-classification necessary,

describing additional variant properties of products and topics, or, so-called functional metadata. The latter will be relevant to connect topics, for example, to service processes. Considering this, functional metadata can cover concepts like linking content to error codes and system events, to work times and with tools. All these extended metadata might be necessary to describe the handling of complex configurable systems of software and hardware.

## CONTENT DELIVERY

In most cases, product documentation from CMS has been widely produced in the past – at least for end users – to comply with legal requirements and international standards. Documents have been produced by highly efficient reuse scenarios and delivered as classical manuals (information for use) in pdf format or as more or less static electronic documents like online-help formats or linked HTML on storage media like CDs. Even more, the need of technicians for structured, detailed and specific information for installing and maintaining complex products has often been neglected. More generally, the requirements of all kind of target groups and users for an interactive and more situational information delivery have not been met for a long time. This has started to change in recent years. Coming up with modern web technologies and system architectures, content can be delivered increasingly by so-called content delivery portals (CDP). The basic definition of CDP has been given first in [8] as "systems, offering web-based access to modular and aggregated content or other information types for various user groups by related retrieval mechanisms".

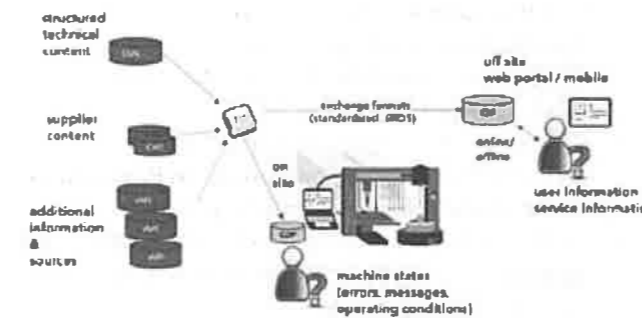


Figure 1. Typical CDP setting including structured and unstructured content sources as well as on-site/off-site access

## CONTENT ACCESS

CDP and users thereof can benefit from the granular topic-based nature of content originating from CMS and also from the assigned metadata. As mentioned, the latter has been used mostly for document creation and retrieval in CMS. Nowadays, three basic access types can be defined:

- structured search by facets and filters corresponding to complex classification taxonomies or more simple metadata sets from CMS
- navigation along classical document structures of nested topics
- direct search based on full-text database indexes by search teams and phrases

The corresponding search functionalities offering, of course, also subsequent mixtures of the basic access types, are often

considered as exclusively human driven, i.e. as manually performed access types. But CDP can also be implemented as a middle ware system of web services delivering on request. Using this approach, content can be addressed manually or via web services in many ways and for many use cases. Corresponding applications are even more initiated by recent IoT and Industry 4.0 approaches from engineering, production, product monitoring, and service management. There, content can be delivered to users (technicians or end users) depending on events, actual machine states and operating conditions. It is clear, that in such a case, the identification of the relevant content crucially depends on the concise location of topics in an information space described by standard and extended (PI-)classifications.

Regarding the devices being used for accessing content, all known types of responsive web interfaces or any other mobile apps are available. Limitations apply for offline usage including synchronizing mechanisms due to the limited storage capacity of applications on mobile devices. Therefore, offline capabilities are offered only by certain vendors. Other applications are run due to security reasons only on-site on machines as local services.

Within companies, the amount of potentially structured content from technical communication and documentation domains is presumably much smaller than the amount of unstructured content. This means, defining the actual use case for requested, retrieved and delivered content will define, the weight and importance of direct search processes and possible enhancements by AI technologies vs. structured faceted search. Additional decision criteria for setting up CDP use cases are offline capabilities or in the easiness and the standardized data handling for interchange, import or crawling of content objects including various media data. Another specialized use case is driven at the moment by mixed reality applications. There, object recognition can be performed by a variety of different user interfaces and technologies. In a PI-formulation, the object recognition defines the intrinsically defined components, whereas the specific use case or manual selection by users defines the required information type. This causes a web request to the CDP, which, in turn, will deliver content via web services to the interface. In such cases, the CDP will therefore not be accessed directly, but can be integrated in other environments and services.

## THE INTELLIGENCE CASCADE

The concepts of intelligent content have been heavily discussed in the recent years and they have found their technical use case in content delivery portals and related applications. The business use cases have been focused on in the previous section. It is widely accepted, that "intelligence" in the context of technical communication is mainly based on:

- granular content
- semantically (XML-)structured content
- semantic metadata assigned to content

Of course, all other aspects of technical communication from standardized, comprehensible writing and localization to (web-)design and media integration, are still of basic importance. But from the point of view of content delivery and related information services, the searchability and accessibility in information space as well as the processability by information systems are the most crucial aspects. But recent developments took further steps to guarantee and to improve

these key factors. In terms of an intelligence cascade, one can find approaches along the lines of

- native intelligence
- augmented intelligence, and
- artificial intelligence

Native intelligence covers the use and implementation of classified content by metadata in CMS or other authoring systems. The discussed methods, like the PI-classification, lead to a concise schema for developing and using metadata which can be displayed directly as facets or processed in other ways in CDP. As a result, recent CDP can act as digital information services for manual or machine-induced search connected to deep linking or communicating web services.

Augmented intelligence has become of recent interest as ontology modeling of semantic networks. These networks can be used in various ways, for example to:

- derive (CMS-)classification from the logical of physical model of products and support content creation
- extend search and retrieval in CDP using semantically modelled relations between content, i.e. improving quality of search results according to use cases
- implement semantic middleware for managing information from and between various domains or data sources, allowing for mapping mechanisms and search across those domains

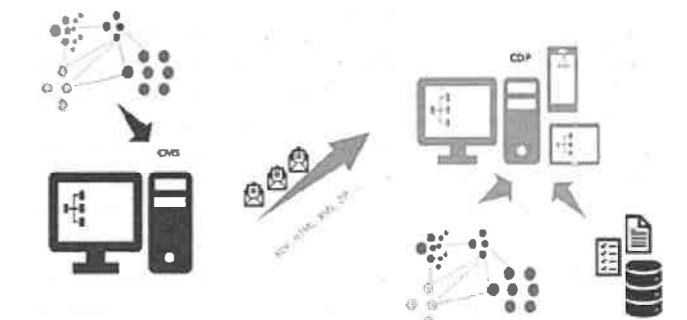


Figure 2. Application areas of ontologies supporting content creation in CMS or content delivery in CDP

The first two cases are sketched in figure 2, implying the support of ontologies for CMS (including hierarchical metadata) or for CDP (including facets). Recent implementations usually focus on one of these sides while there is no general restriction for supporting both sides.

A future use cases derived from ontologies will be presumably more dynamic and logic driven compositions of semantically linked topics. Information requests can benefit from predefined semantic models of mesoscopic structures which could be seen as "micro-documents" suitable for specific problems with limited context as connected tasks or workflows or any other product relations. This would also bridge the logical gap between monolithic documents and single topic delivery. They can be predefined as logical pattern described in semantic relations between topics or, more general, between information resources. The logical patterns could be simply defined as static micro-documents or in terms of ontological relations triggered by search processes. At higher automatization levels, they could

be derived by analytics of user behavior (past search processes) or even derived automatically from changes in the underlying product models.

Artificial intelligence (AI) is also diffusing into technical communication and content-related technologies. In the given context, AI can be used for automated content classification of legacy data or of unstructured data [8]. In these cases, one needs a set of training data including high-quality classified and, therefore, native intelligent content. Further applications deal with content generation, i.e. delivering content which is dynamically created from training data corresponding to (CDP-)requests. On the front-end side, one can include AI technologies like object and speech recognition to identify the requested object and information types and to facilitate search processes

The latter is generally true for all levels of the intelligence cascade. Companies and software vendors of information systems will have to rethink their methodologies and architecture towards support of one or more of these levels.

## SUMMARY AND OUTLOOK

This article focused content management and content delivery based on recent technologies and information architectures. We have explained the key methods of this content-related information systems and the associated concept of intelligent content including semantic modelling of structured content and classification metadata. On the delivery-side, we were focusing on content access types and various technical use cases covering. We have also shown, that in recent stages of such information scenarios, intelligent and highly searchable content from CDP can provide information services for a higher usability. These services can be offered and used more efficiently when the underlying concepts are addressing the diverse levels of an intelligence cascade. This cascade can be described in terms of native, augmented and artificial intelligence.

In actual implementations, all these diverse approaches require, on the conceptional side, a deep knowledge of products, relations, processes and the required volume of information space. This also requires, on the one hand, collaboration and communication between groups in product development and information development. On the other hand, it requires, that academic and professional education of information manager in technical communication will also have to cover these new fields and levels of intelligent information architectures.

## REFERENCES

- [1] Ziegler W. Drivers of Digital Information Services: Intelligent Information Architectures in Technical Communication, 2019 ACM Chapter Proceedings on Educational Technology, Language and Technical Communication, ISSN: 2434-9526, (2019) VOLUME 1, S. 48-52
- [2] Straub D., Ziegler W. tekcom CMS Studie 2018/19, tcworld Stuttgart, Germany
- [3] Ziegler, W. 2017. Metadaten fuer intelligenten Content. In Schriften zur Technischen Kommunikation, Vol 22, p. 51-66, tcworld
- [4] Ziegler, W. 2017. The Evolution of Content Management

towards Intelligent Delivery Systems for Technical Communication. In Frontier, Official Journal of Japan Technical Communicators Association JTCA, p. 68-75.

- [5] Ziegler W. 2017. Basic Concepts Which Support the Management and Delivery of Intelligent Content. In Frontier, Official Journal of Japan Technical Communicators Association JTCA, p. 84-91
- [6] Parson U., Sapara J., Ziegler W. 2017. iirDS for Technical Writers – Introduction to the Metadata. In Conference Proceedings of the Annual tekcom Conference.
- [7] European Association for Technical Communication – tekcom e.V. 2017. iirDS Specification – intelligent information Request and Delivery Standard – First Public Working Draft, 03 April 2017. (2017)
- [8] Ziegler W. 2013. Alles muss raus! Content Delivery fuer Informationsportale, In Proceedings of the annual tekcom conference, p. 47

## 用語表記の統一

### —— 霞が関/霞が関/霞ヶ関/霞ヶ関 ?? ——

Unification of terminology

--- Especially focusing on proper nouns such as place names, company names and person names ---  
KASUMIGASEKI is the center of all ministries in Japan. The address notation of KASUMIGASEKI in Japanese is 霞が関, and the station name of Tokyo Metro (subway) is 霞ヶ関. It is necessary to understand and correctly express these differences. Such mistake in a company name or person's name is impolite to the other party and should not be made.

有限会社 アトリエ・ワン

Atelier Bow-Wow

貝島良太

Ryota KAIJIMA

官庁の中心地は東京都千代田区の「かすみがせき」である。住居表示は「霞が関」だが、東京メトロ(地下鉄)の駅名は「霞ヶ関」になっている。「霞が関」でも「霞ヶ関」でもない。我が国で最初に建てられた超高層ビルの名称は、「霞が関ビルディング」である。したがって、たとえば霞が関ビルへのアクセスを説明するときは、「霞が関ビルディング 住所: 〒100-6090 東京都千代田区霞が関三丁目2番5号 アクセス: 東京メトロ銀座線 虎ノ門駅[11]徒歩2分、東京メトロ丸ノ内線/日比谷線/千代田線 霞ヶ関駅[A13]徒歩6分…」のように書き分けておかないと住居表示名や駅名としての固有名詞の表記と一致しない。「丸ノ内線」も「丸の内線」ではない。分れば構わないという考えもあるが、地名、社名、団体名、氏名、登録商標の商品名などの固有名詞はとくに注意して記載する必要がある。なかでも、社名と氏名などの固有名詞の表記を間違ふことは、相手にたいへん失礼である。また、適切な文字変換や検索を行うためにも、正式名称での表記を調べて、そのとおりに表記することが重要である。

### 1. はじめに

筆者は、マニュアル文をはじめ一般的なビジネス文書の執筆時における疑問点や注意点を、2009年から1年に1テーマ、この場で研究発表をしている。今年はその10回目である。2009年(第1回目)のテーマは、「くさい/下さい」の使い分け<sup>[1]</sup>であった。2010年は、助数詞の「ka」<sup>[2]</sup>、2011年は、研究発表が中止。2012年は、ら抜き言葉<sup>[3]</sup>、2013年は、句点と閉じ括弧<sup>[4]</sup>、2014年は、接続詞の考察<sup>[5]</sup>、2015年は、横書きの文章における句読点についての考察<sup>[6]</sup>、2016年は、複合動詞への送り仮名の付け方<sup>[7]</sup>、2017年は、項目番号の付け方<sup>[8]</sup>、2018年は、用語集の作り方<sup>[9]</sup>であった。これらの発表論文は <http://www.bow-wow.jp/sht3/data>.

htmlの項番021から029に掲載してあるので、参考にいただければ幸甚である。

さて、今年のテーマは、特に地名・社名・人名などの固有名詞に焦点を当てた用語表記の統一である。取説(製品マニュアル)をはじめ、一般的なビジネス文書、規則書、論文、新聞・雑誌の記事、さらには小説のような情緒的なものでさえもこれらの固有名詞の表記が統一されていないと、読み手(ユーザー)をまごつかせたり、イラつかせたり、ときには間違わせたりする。社名・人名などは、間違われた相手方にも不快感が残り、大変失礼なことである。特に買い手市場のとき、売り込み用に作成した資料などで、先方の社名や氏名などを書き損じただけで採用されない場合もあると聞いたことがある。