

## 創造的情報管理の時代を拓く(!?) DITA への3つの問い

鎌田 博樹

オブジェクトレポート編集長／オブジェクトテクノロジー研究所

### はじめに

情報（文書・コンテンツ…）管理は、ビジネスと IT の永遠のテーマでもあると同時に、焦眉の現実的課題です。情報の総量（大半が複製とゴミ）が爆発的に増大する中で、意味のある情報だけを収集・開発・編集・提供・更新していくには、インテリジェントな管理が絶対的に必要です。管理は、まさに企業の競争力、技術力、法令遵守、危機管理能力など、サステナビリティの根幹に関わる問題となっています。問題はその先。

その1：いくら重要とはいえ、管理を嫌うのも人間の本性であり、創造は管理を超えたところにあります。創造的な人は自分の領域以外での管理を望み、自分の領域では従来の方法や体系を壊そうとするからです。しかし皮肉なことに、壊すためには新たに組み替える必要があり、そのために膨大な時間を費やすこととなります。その負担は、管理的業務に携わる人よりもむしろ多いかもしれません。

その2：管理には何らかの標準が必要です。しかし管理を管理する標準もまた、管理のジレンマから自由ではありません。標準もまた、過去と未来、汎用と独自という反対の世界に足をかける必要のあるもので、唯一正しいというものは存在しないからです。情報という抽象的世界でどんな標準に意味があるのか。過去何十年の間、多くの人が様々な答えを見つけました。しかしそれらを共存させるような仕組みはなかなか有りませんでした。

「脱構築」を本質とする創造的世界が必要とする情報の管理とその標準は、自由に設計でき、知識の組み換えと再利用をサポートし、静的であると同時に動的であるようなものである必要があります。「管理嫌いのための管理システム」「標準嫌いのための標準」とでも言えましょうか。それは無いものねだりのようですが、現実のニーズは、やはり合理的な解を必要としています。DITA への期待はまさにそこにあります。そこで、この問題に長く関心を持ってきたものとして、次の3つの質問を、エキスパートの方に提起してみたいと思います。

1. DITA は創造的情報管理を実現できるか（CMS の方向性と DITA）
2. DITA は誰が、どう使い、どんな利益を得るのか（欧米の展開と日本の課題）
3. DITA は（Web とドキュメントの）CMS の大統一を実現できるか

かなり大きなテーマなので、以下、ディスカッションを御膳立てする意味で、今日に至る「情報」管理のアプローチを整理しておきたいと思います。蛇足・舌足らずはご容赦ください。

### 1. 電子情報管理・文書処理と標準化の壁

DTP から Web に至る今日の電子出版技術の原型は、ほぼすべて 1960 年代後半に米国で生まれました。主として米国海軍が、艦船に装備する（潜水艦でも数トンに達する）マニュアルの軽量化＝電子化を必要としたことに対応したものです。これは実験的技術というだけでなく、

完全に実用されました。軍という閉鎖的な情報環境で生まれた技術でしたが、電子ドキュメントは、ほとんど無限と言える可能性を秘めており、その後のメディア技術の多くはここから派生しています。その意義を要約するなら、次のようなことでしょうか。

- データ／情報の文書化（出版）の電子的ライフサイクルが確立した
- 企業が行う<情報業務>と<配布／出版>との技術的連携が確立した

1980年代に、WYSIWYGのユーザーインターフェースが民生技術化され、パソコンで利用可能になったことで、電子ドキュメントは全面的な発展の可能性を与えられたかに見えました。しかし、当時の電子的環境は、CPU やスループットという以上に、あらゆるところでベンダーの独自フォーマットの壁があり、電子ドキュメントは、わずかにワードプロセッサや、印刷の前工程（プリプレス）としてのDTPという形で実現されただけでした。ユーザーが圧倒的に強力な「軍」や、それに近い大企業は、独自の規格に従った製品をITベンダーに要求することができますが、一般のユーザーではそうはいきません。結局、標準がないところでは、電子ドキュメント（出版）は普及しないわけです。

電子ドキュメント標準化の努力は、1980年代以降、ねばり強く続けられてきました。初期の代表的なものは、ODA (Open Document Architecture) と SGML (Standard Generalized Markup Language) です。前者は窮屈ながらもほぼすべての側面をカバーし、後者は汎用的な構造定義のみにフォーカスする、というまったく逆方向からのアプローチだったのですが、後者が生き残って、今日のマークアップ言語 (HTML/XML/...)の母体となったことは周知の通りです。SGMLが残ったのは、表現に対しても意味 (的構造) に対しても自由に使えるためですが、電子ドキュメントの標準としてはとても不完全なものでした。ワープロやDTPアプリケーションの独自フォーマットの壁も厚く、企業ドキュメントの世界での標準化の進展と社会化は期待外れに終わりました。

しかし、1990年代後半にはWebという、HTMLを使ったハイパーテキスト空間がインターネットとともに拡大し、文書の利用においては従来のようなワープロや紙に代わってWeb (ブラウザ) がプラットフォームとして普及しました。これは革命的な変化でした。1970年前後に生まれた技術のうち、ハイパーテキストがインターネットというグローバルな環境を得て、電子ドキュメント (出版) の主流となったのです。ハードウェアや通信環境は、Webをサポートする方向で大きく躍進し、さらに標準ベースのマルチメディア化、ダイナミック化を押し進めました。これらをあえて要約するならば、IPが前提となったことによって、

- 文書ライフサイクルをサポートする標準的プラットフォームが生まれた
- 文書形式でなくユーザーニーズとUIに合わせて環境が形成された

ということが言えると思います。

## 2. CMSにおける「軽い」構造と「重い」構造

標準の淘汰圧は、利用環境 (ブラウザ) から提供環境 (サーバ・アプリケーション) へと逆に進み、データからコンテンツ管理へと進んでいきました。従来、作成側の環境で停滞していた標準化が、紙や冊子を迂回することでライフサイクルをサポートする方向で一気に進展していったわけですね。近年のCMSの成長は、ほとんどWebサイト管理の成長が牽引していま

す。標準は、ユーザーのためにあるべきものですが、文書の作成と管理に携わるユーザーは、数は多いもののニーズはバラバラで、標準化するにはとても「重たい」世界でした。ドキュメントの標準は、生産側からではなく、真のユーザーである読者側の環境から進むしかなかったということでしょう。

CMS は、爆発的に増大し、複雑化する Web ドキュメントの表現的要素を構造的に管理するために、きわめて自然に使われていきました。これらは非常に客観的、汎用的という意味で「軽い」構造と言えます。しかし、構造化された知識の結晶体としてのドキュメントには、作成／編集者にとって死活的に重要な「意味的構造」が含まれています。コミュニケーションという観点では、表現構造は意味的構造をサポートするためにあり、両者は整合性を持つべきものです。CMS は本来、そのために生まれた技術です。今日の Web が、さらに進化するためには、意味と表現の一致という、ODA 以来の標準化の課題を解決しなくてはなりません。

意味という「重い」構造が、標準化のボトルネックとなってきたことは、上述した通りです。意味的構造が「重い」のは、もちろんあまりに多様で「一般」化しづらいからです。学術論文などは昔から型がきっちり決められています。行政文書などを除いてインタフェースはあまり形式化されていませんでした。筆者はその昔、ドキュメントの分類や構造定義による標準開発を試みたことがあるのですが、使用されるコンテキストによってよく似た構造で違う意味を持ち、異なる構造でも同じ意味を持つという性質を扱うには、利用ドメインによって情報型を定義するしかないと思いました。

### 3. 新世代標準としての DITA

標準化の圧力は、やはり Web、つまりユーザーのほうから生まれました。金融や医療など、専門的情報の大部分は、共通のコンテキストの下で意味的構造はかなり明確で、ドキュメント間での情報要素の関連性も形式化できています。XML という汎用言語を使って型を定義することで、文書に収められた情報は、人間が読解する以外にも、システムの大きな「意味」を持ち、様々なインテリジェンスをサポートできるようになるわけです。例えば、XBRL (eXtensible Business Reporting Language) は、各種財務報告用の情報を作成・流通・利用できるように標準化された XML ベースの言語です。それに対して、今回のフォーラムの主要なテーマである DITA は、テクニカル・コミュニケーションを含む、創造的情報を型付けするもので、XBRL のアプローチとは異なりますが、同様のインパクトを持つと思われます。

財務報告などの「手続き指向」的世界と異なり、テクニカル・コミュニケーションのためのドキュメントは、複雑でしかもたえず変化し、融合し、増殖するテクノロジーの世界を反映して、意味的構造も動的に変化します。携帯電話のマニュアルを考えてみてください。機能の解説に振り回されると、作成する人にも読む人にも悪夢のようなものになってしまいます。静的な意味構造を扱う方法では、いくら頑張っても利用者のニーズに対応できず、ロスとリスクを大きくしてしまうでしょう。

DITA のアーキテクチャは、継承／特殊化というオブジェクト指向のコンセプトを採用している点が画期的です。トピックとマップという基本要素は、現実のテクニカル・コミュニケーションの世界を的確に表現することができます。従来の方法との違いに戸惑う人がいても

無理はありませんが、少なくともこのドメインでは、肥大化する業務と増大するコミュニケーション・ギャップという問題を解決する最善の手段と言えるでしょう。

DITA の最も重要なドメインの一つが、テクニカル・コミュニケーションであることは確かでしょう。しかしそれは「マニュアル」に完結するものではなく、非常に広い世界に通じています。技術教育はその一つで、学校から職場までの様々な現場で新たに要求される知識やスキルの多様な組合せに対して、従来のような「静的」な情報管理をもとに作られた教科書とマニュアルで教育やトレーニングを行っていたのでは、変化する現実には追いつけません。インテリジェントなテクニカルドキュメントは、21 世紀の教育のプラットフォームであると思います。

#### 4. インテリジェント・コミュニケーションの世界

最初に述べたように、DTP から Web に至るドキュメント技術は、テクニカル・コミュニケーションの現場で生まれました。その後、グループウェアとインターネットという別々の環境で発展してきたドキュメント技術は、標準と CMS が実現する「意味と表現の整合的管理」によって再び統合されようとしています。これはドキュメント技術の終着点でしょうか。私はむしろ出発点であると考えます。これまでテクニカルライターやエディタ、情報デザイナーの努力によって築かれてきたアートとしてのドキュメント（サイト）の DNA の保存・再利用を容易にする DITA は、新しいコミュニケーションのプラットフォームとなる可能性を秘めているからです。

では何に向かっての出発点でしょうか。それは利用者の高度な（多様な）ニーズに容易に応えることができるインテリジェンスへの道です。結論だけを申し上げますが、基本的に

- ドキュメント（コンテンツ）管理のインテリジェント化
- ドキュメントのインテリジェント化（サービス化）

という 2 つの方向が考えられます。前者はサイトの構築とコンテンツ管理、スタイル要素、サービス機能などの管理を視覚化、自動化する環境（インテリジェント CMS）であり、後者は、eBook やスマートフォンなど多様な「メディア」への対応、ユーザーのプロファイルに対応し、外部サービスとの連携を動的に行うインテリジェント・ドキュメントなどが考えられます。

幸いにして、ダイナミックでインテリジェントなドキュメントのための技術は、過去 40 年の間に実験・実証を重ねてきました。Web と DTP、多様なデータ、意味と表現を統合的に駆使できる標準環境が、Web 2.0 のソーシャルな環境と結びついた時、知識情報の容器としてのドキュメントのパワーが全開されることが期待されます。（鎌田博樹、08/22/2009）