

# XML の文書作成規格による教材作成と学習・教授スタイル

内海 淳

Email: utsumi@cc.hirosaki-u.ac.jp

弘前大学人文学部

©Key Words XML, DocBook, DITA

## 1. はじめに

DocBook や DITA (Darwin Information Typing Architecture)などの XML の文書作成規格を用いることにより、多様なフォーマットや学習者に最適化された教材やマニュアルを単一のソース文書から作成することが可能となる。しかし、XML 文書作成規格には、DocBook や DITA の他にも多くの規格が存在している。にもかかわらず、日本においては、教材やマニュアルなどを作成する場合に、どの XML 文書作成規格を選んだらよいかを判断する情報がほとんどない。

本発表では、DocBook と DITA という比較的取り上げられることが多い規格を中心に、それぞれの規格の特性とそれぞれに適した学習・教授スタイルがどのようなものであるかを考察する。実際に、同一の教材・マニュアルを DocBook と DITA などの規格で作成し、それぞれの長所短所を具体的に検討し、それぞれがどのような学習・教授スタイルに向いているかを考察する。

## 2. XML 文書作成規格

### 2.1 XML 文書作成規格とは

本発表で取り上げる XML 文書作成規格とは、XML に基づいて定義された文書作成のためのマークアップ言語を指す。XML 文書作成規格の代表的なものとして、DocBook, DITA, TEI (Text Encoding Initiative), S1000D などがある。XML 文書作成規格のほとんどは、コンピュータのハードウェアやソフトウェア、航空機などのマニュアルといった技術文書の作成のために開発されたものである。しかし、最近では、技術系の文書に限らず、より広範囲の文書の作成にも利用されてきている。例えば、上述の TEI は、人文学、社会科学、言語学を中心とした様々な文書を記述するための XML 文書作成規格である。

XML 文書作成規格を使って(ソースとなる)文書を作成する場合、原則として、レイアウトやフォントの書式等の、文書の見栄えに関する情報を排除し、その文書の論理的構造のみを記述することが求められる。ソースとなる文書は、XSLT (Extensible Stylesheet Language Transformation)や XSL-FO (Extensible Stylesheet Language Formatting Object)などの処理系によって html, ePub, pdf などのフォーマットへと変換される。この変換の際に、ソースとなる文書の論理的情報に基づいて、それぞれのフォーマットに適したレイアウト処理など

が施される。

### 2.2 XML 文書作成規格の長所

XML 文書作成規格の利点の一つは、html, ePub, pdf などの多様なフォーマットを統一的に管理することが可能になることである。ワープロや TeX などで文書を作成する場合は、特定のフォーマットを前提にして文書を作成するため、そのフォーマットに特有のレイアウトやスタイルの情報が入り込んでしまう。そのため、いったん作成した文書を他のフォーマットの文書に変換する際に手直しが必要になる。これに対し、XML 文書作成規格を使う場合、特定のフォーマットを前提にしたレイアウトなどの処理は XSLT などによって行われるため、ソースとなる文書の手直しは、論理的な内容の変更に限られる。

XML 文書作成規格には、上記の特定のフォーマットにとらわれないこと以上に重要な機能がある。それはターゲットごとに異なる情報を統一的に管理する機能である。この機能をプロファイリング(profiling)と呼ぶ。教材やマニュアルの場合には、文書の基本的な構成は共通であるが、ユーザのグループごとに異なる形で提供する場合がある。例えば、同一の内容のマニュアルを日本語、英語、中国語で別々に提供する場合もあるし、学部や学科などの所属に応じて細部を変更する場合がある。また、学習教材をその習熟度に応じて書き方を変える場合なども考えられる。

XML 文書作成規格では、ソースとなる文書内に、これらのターゲットごとに異なる情報をタグの属性(attribute)として指定する。ソースとなる文書を XSLT などに変換して出力する際に、ターゲットにしているグループの属性を指定することにより、そのグループ向けの情報のみを含んだ文書を出力することができる。

### 2.3 XML 文書作成規格の短所

XML 文書作成規格の短所としては、レイアウト重視の文書作成には向いていないことが挙げられる。上述のように、ソースとなる文書は、原則として、レイアウトなどに関する情報を含まない。レイアウトの処理などは XSLT に任されるが、XSLT では複雑な見栄えに関する処理を行うことは難しいため、複雑なレイアウトなどが必要な場合は、InDesign や TeX などの、組版ソフトを介在させることが必要になる。

### 3. DocBook と DITA

#### 3.1 DocBook

DocBook は 1991 年から Norman Walsh らによって開始された XML 文書作成規格であり、現在はバージョン 5 となる DocBook5 が公開されている。Walsh (2010)<sup>(1)</sup> に DocBook5 の詳細な仕様が解説されている。また、DocBook から XSLT を使って様々な表示形式へ変換に関し詳しく解説しているものとして、Stayton (2007)<sup>(2)</sup> がある。DocBook は OASIS(Organization for the Advancement of Structured Information Standard)の下で管理されており、誰でも自由に無償で利用することができる。

DocBook は、拡張しない限り、Book と Article の 2 つの形式で文書を作成する。このため、この 2 つの文書形式に関連する多くのタグがあらかじめ用意されている。

#### 3.2 DITA

DITA は IBM が開発した XML 文書作成規格であるが、現在は、DocBook と同様に、OASIS によって管理されており、誰でも自由に無償で利用することができる。

DITA は、企業向けの XML 文書作成規格として最近特に注目されてきている。日本語での解説書も既に 3 冊<sup>(3)(4)(5)</sup>出版されており、英語圏ではさらに多くの解説書があり、セミナーなども頻繁に開催されている。

DITA では Topic と Map という概念を使って文書を作成する。Topic とはひとつの主題に関する自己完結したコンテンツの単位である。Map はこれらの Topic を組み合わせてどのような文書を構成するかを定義する要素である。書籍を作成する場合は、Bookmap という特殊化された Map があらかじめ用意されている。

#### 3.3 学習・教授スタイル

DocBook は、これまでよく使われている書籍・論文の形式の文書が基本であるため、従来からよく行われてきた、基本的な事項からより高度な知識へと進む、「知識積み上げ型」の学習・教授スタイルに適している。

DITA は、Topic という独立した要素を、Map を使って結びつけている。したがって、個々の授業内容がある程度独立してはいるが、授業間の関連性がある程度と持たれているオムニバス形式などの「並列的な」学習・教授スタイルに適している。そのような学習の一例として、外国語の文法を挙げることができる。例えば、英文法の「能動態・受動態」と「過去時制・現在時制」を学習する場合、両者の間には、基礎対上級の関係はないが、相互に密接な関連がある。このような事項を扱う場合、要素を並列的に配置し、その要素間で相互参照を多用することになる。このような形の文書作成には DITA が力を発揮する。

#### 3.4 コスト

DocBook も DITA も無償で利用できるため、この点でのコストの優劣はない。しかし、導入のために必要な「労力」の違いがある。

まず、意識改革が求められるかどうかの問題がある。上述の通り、DocBook は伝統的な Book と Article の形式の文書を作成するため、比較的抵抗なく導入することができる。

これに対し、DITA は Topic と Map という、まったく新しい概念で文書を作成することが要求される。DITA でも Book 形式の文書を作成することはできるが、そのためには、Topic と Map を使いこなせなければならない。DITA に関するセミナーが多い理由の一つは、このようにまったく新しい考え方をしなければならぬことにあると考えられる。

次に、拡張の必要性に関する違いを挙げることができる。DocBook は文書作成に必要とされるタグ要素を数多く用意しているため、ユーザが自らタグ要素を拡張・作成する必要はあまりない。

これに対し、DITA では、あらかじめ用意されているタグ要素は、DocBook に比較して、かなり少ない。これは、DITA がオブジェクト指向モデルを採用し、ユーザが自分の利用目的に合わせた文書作成のための要素をみずから拡張・作成することを前提としているからである。これらの要素の拡張・作成は、比較的容易にできるようになっているとはいえ、ある程度のプログラミングの知識が必要とされる。

最期に、情報の得やすさの問題がある。上述の 2 点に関しては、DocBook の方が DITA に対して有利であると考えられるが、情報の得やすさに関しては逆である。インターネット上のリソース以外で、DocBook に関する日本語で書かれた情報はほとんどない。これに対し、DITA は既に 3 冊も参考書が出版されており、セミナーなども活発に行われている。

### 4. おわりに

本発表では、XML 文書作成規格を用いた教材作成の問題について取り上げ、代表的な XML 文書作成規格として、DocBook および DITA を比較した。しかし、どちらの規格もまだまだ教育現場で使われる機会が少ない。PC やタブレット端末が教育現場に浸透していくにつれて、このような XML 文書作成規格の必要性が高まっていくことが期待される。

#### 参考文献

- (1) Norman Walsh : DocBook 5, O'Reilly (2010).  
(<http://www.oasis-open.org/docbook/documentation/reference/html/docbook.html> にオンライン版が公開されている)
- (2) Bob Stayton : DocBook XSL: The Complete Guide, 4<sup>th</sup> ed., Sagehill Enterprises (2007).  
(<http://www.sagehill.net/docbookxsl/> にオンライン版が公開されている)
- (3) Comtech Services, Inc : “DITA 概説書”, エスアイビー・アクセス (2011).
- (4) Ann Rockley, Steve Manning, Charles Cooper : “DITA 101”, エスアイビー・アクセス (2011).
- (5) Julio Vazquez : “DITA の実践”, エスアイビー・アクセス (2011).