

研究開発報告

XML 技術を用いた教材・マニュアルのシングルソース化

人文学部 内海 淳
utsumi@cc.hirosaki-u.ac.jp

1. はじめに

PCやスマートフォンなどの多様な情報機器が学習活動において日常的に使用されている現在、学習教材もこれらの機器などに合わせた多様な形式で提供することが望ましい。同時に、これらの教材を利用する学習者についても、その習熟度や興味の違いに応じて、その学習者に最適化された学習教材を提供することが望ましい。

教材となる文書の作成においては、これまでワープロやTeXなどが使用されてきた。しかし、これらの手段では上記のような要求を満たす多種多様な教材を作成することは、不可能ではないにしても、多大な労力を必要とする。

また、現在の学内の学生向けマニュアル等は、学部・学科ごとに異なる内容が一つの冊子の中に混在しており、学生にとって読みにくく、必要な部分が分かりにくい。しかし、学部・学科ごとに異なるマニュアルを別々に作成することは編集事務の負担を大きくし、実用的ではない。この問題を解決するためには、基本的な構造を共有するが、グループごとの相違を組み込んだ文書を単一のファイルとして編集し、グループごとに異なる出力結果を出すシステムを用意する必要がある。

DocBookやDITA (Darwin Information Typing Architecture) などのXMLの文書作成規格のシステム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。このようなシステムをより利用しやすくすることが本研究の目標である。

2. XML 文書作成規格とは

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) 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指定等の、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。

2.1. 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が異なる場合がある。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。

2.2. 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。

2.3. DocBook

XML 文書作成規格として広く利用されているものとして、DocBook と DITA があるが、本研究では DocBook を採用する。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)⁽¹⁾ 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)⁽²⁾ 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。

本研究でDocBookを採用したのは、DocBookが伝統的な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ため、かなりの意識改革が必要となる。

3. 研究の概要

DocBookは優れた機能を持っているにも関わらず、その利用の仕方についての日本語のマニュアルは存在していない。このことが、DocBookの普及を妨げている大きな要因であると考えられる。そこで、本研究では、DocBookのマニュアルを作成することを第一の目標とした。しかし、DocBookには数百ものタグやそのタグに結びついた属性が存在するため、具体的なターゲットとなる文書を定め、その作成に重要なタグや属性に焦点を絞ることとした。

具体的なターゲットとなる文書として、弘前大学21世紀教育センターの『21世紀教育科目履修マニュアル』を選択した。このマニュアルには、学部・学科ごとに異なる指定が混在している上、将来的には、Web上で公開や、iPadなどでの利用が想定されるため、本研究での試みが有用であると考えたからである。このマニュアルを、紙への印刷を前提としたPDF形式、Web閲覧のためのhtmlおよびxhtml形式、iPad等の新しい電子端末向けのePUB形式の3つの出力形式で、学部・学科・課程ごとのプロファイリングを行う場合を想定して、ソースファイルの作成を行い、その際に必要とされるタグや属性を抽出した。

しおり

21世紀教育科目履修マニュアル

目次

第1章 21世紀教育科目とは

第2章 履修にあたっての基本的知識

第3章 21世紀教育科目のカリキュラム

第4章 人文系の履修事項

履修すべき単位数と履修指定・推薦指定科目

履修できる単位の上限一覧

1年次前期における21世紀教育科目履修登録見本

多言語コミュニケーション実習における履修登録選択の際の注意事項

21世紀教育科目履修マニュアル

目次

第1章 21世紀教育科目とは

第2章 履修にあたっての基本的知識

第3章 21世紀教育科目のカリキュラム

第4章 人文系の履修事項

履修すべき単位数と履修指定・推薦指定科目

履修できる単位の上限一覧

1年次前期における21世紀教育科目履修登録見本

多言語コミュニケーション実習における履修登録選択の際の注意事項

修得すべき単位数と履修指定・推薦指定科目

21世紀教育科目における修得すべき単位数と履修指定・推薦指定されている科目を以下に示します。(必修指定科目は、履修すべき単位数に表中に含まれています。)

表4.1 現代社会課程の修得すべき単位数

科目区分	科目	必要単位数	備考
導入科目	基礎ゼミナール	2単位	
技能系科目	英語コミュニケーション実習	4単位	
	英語コミュニケーション実習	2単位	1言語を選択
	多言語コミュニケーション実習	2単位	
基礎教育科目	文化系基礎	4単位	
	社会系基礎	2単位	
	自然系基礎	4単位	
	情報系基礎	2単位	
	自由選択		情報1・11・111から1科目2単位選択
テーマ科目		14単位	6領域から各1科目2単位、計12単位。その他、適宜選択して1科目2単位、合計14単位
適宜修得単位		2単位	
合計		42単位	

履修できる単位の上限一覧

21世紀教育において、各学期に履修登録できる単位の上限は、科目区分や領域ごとに次のように定められています。

なお、「集中授業」も履修登録できる単位の上限に含まれます。ただし、「スポーツ・体育実技・体育実技」の集中授業は履修登録できる単位の上限に含まれません。

科目区分・領域	各学期の履修登録単位の上限	
導入科目(基礎ゼミナール)	各学期2単位までとする。	
技能系科目	英語コミュニケーション実習	各学期2単位までとする。
	多言語コミュニケーション実習	【特設言語】は各学期1単位までとする。他の科目については上限を設けない。
	スポーツ・体育実技	各学期1単位までとする。ただし、集中授業は除く。
合計	上記の上限を求め、21世紀教育科目全体で各学期24単位まで。	

24

図1 PDF形式の出力

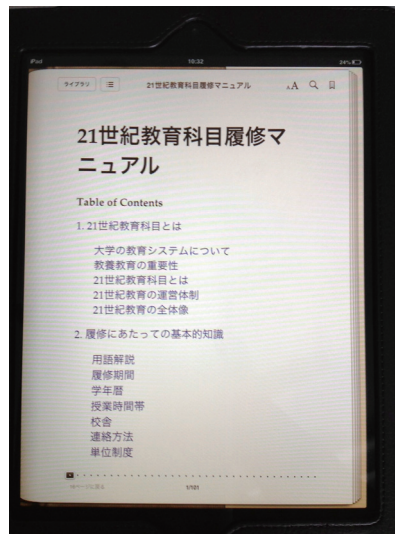


図2 ePUB形式の出力（iPad上のiBooksで表示）

この『21世紀教育科目履修マニュアル』の作成で抽出したタグや属性およびその出力方法について解説したDocBookの簡易マニュアル、『21世紀教育科目履修マニュアル』のソースファイル、プロファイリングを用いた出力例（html、PDF、ePub）は弘前大学総合情報処理センターのWebClass（<http://webclass.stu.hirosaki-u.ac.jp> 学外からは<https://webclass.stu.hirosaki-u.ac.jp/>）上で公開している。

おわりに

本研究でとりあげたDocBook等のXML文書作成規格は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文書作成規格の利用が進むことが期待される。

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、4th ed., Sagehill Enterprises (2007). (<http://www.sagehill.net/docbookxsl/>にオンライン版が公開されている)