

# 図書館目録とメタデータ

渡邊隆弘（整理技術研究グループ）

当グループでは 2000 年度の 1 年間、インターネットの急激な発展の中で普及が進むデータ処理技術である「マークアップ言語およびメタデータ」をテーマとして研究活動を行ってきた。文書などのデータの構造的記述を追求するマークアップ言語と、データに関する二次情報を意味するメタデータは、図書館における整理技術・情報組織化活動にも密接な関連をもっている。

本発表では、近年あらわれた「電子図書館資料」に対するメタデータはどうあるべきか、さらには電子図書館を視野に入れたとき図書館目録は今後どのように変わっていくべきか、について若干の考察を試みる。

## 1 . はじめに

電子図書館、インターネットと整理技術

- ・ 一次情報のデジタル化
- ・ 図書館の枠を超えた情報のやりとり

「電子図書館資料」

- ・ 図書館資料のデジタル化
- ・ WWW ページなどのネットワーク情報資源

電子図書館資料のメタデータ<sup>1</sup>表現はどうあるべきか？ を中心に  
これからの図書館目録はどうあるべきか？

## 2 . 電子図書館資料とメタデータ

メタデータ（書誌情報）の必要性

- ・ 文字以外の情報表現
- ・ 全文検索の限界
- ・ 一次情報構造化の限界

---

<sup>1</sup> メタデータ

「データに関するデータ」「データに関する構造的データ」と定義される。データに対する二次的情報がすべて含まれ、定義・用途とも極めて広範にわたる概念であるが、ここでは情報の発見や同定識別を目的とする「書誌情報」を念頭におく。なお、対象を WWW ページなどのネットワーク情報資源に限定して言及されることもあるが、本発表ではそのように限定した用い方はしない。

## 電子図書館資料の特性と目録規則

- ・ 情報表現の多様性
- ・ 媒体と内容の結びつきが強固でない
- ・ 内容変更・改訂が容易
- ・ 必要な部分だけを検索・抽出しての利用  
メタデータを作成する情報単位の問題

## メタデータ記述規則の検討

- ・ 「意味」の問題と「構文」の問題は分けて考えるのが合理的  
意味の側面（どのような情報を記述すればよいか）  
構文の側面（どのような形式で記述すればよいか）  
記述単位と構造表現

## 3 . メタデータにどのような情報を記述するか（意味の側面）

### 必要の観点

- ・ なるべく詳細なデータ記述
- ・ なるべく詳細なデータ要素の弁別

### 標準化の観点

- ・ 図書館の枠を超えた情報交換・情報検索
- ・ 世界標準に合わせるべきだという意識

### Dublin Core<sup>2</sup>

- ・ 情報作成者でも作れるようにコアエレメントだけを定義  
情報専門家が作るメタデータの標準としては不十分
- ・ 拡張可能性の評価  
Qualifier によるサブエレメント定義に限界  
構造をもつ値の表現不可
- ・ 図書館を超えた広いレベルでの情報交換には有効

---

<sup>2</sup> Dublin Core Metadata Element Set

ネットワーク情報資源の発見を目的として提唱されたメタデータ記述規則であり、国際標準として普及しつつある。

リソースの著作者がメタデータを同時に作成することを想定して、15 項目の基本エレメントに限定され、またすべての項目がオプションとなっている。一方で各機関等が独自に項目を付加していくなどの拡張可能性もあり、完全に自己完結した規則ではない。

エレメントの意味定義に特化された規則で具体的な構文定義は持たない。

図書館その他の機関によるメタデータ作成にも広く用いられており、日本にもいくつかの実用システム例がみられる。

## 4 . メタデータをどのような形式で記述するか ( 構文の側面 )

現在の目録における表現形式

目録規則	ISBD 区切り記号
MARC フォーマット	タグ、インディケータ、サブフィールド識別子
内部フォーマット	多くは RDBMS に格納

「詳細なデータ要素の弁別」に限界

繰り返し、入れ子構造をもったデータ

マークアップ言語<sup>3</sup> ( SGML<sup>4</sup>, XML<sup>5</sup> ) と RDF<sup>6</sup>

データベース記述言語として有効

- ・ より詳細・柔軟な構造表現が容易に可能
  - ・ 図書館の枠を超えた標準性
- ただし、実際の表現能力はメタデータ規則に依存

## 5 . 記述の単位と関係構造の表現

図書館目録

物理単位の利用

単行書誌単位、逐次刊行書誌単位

電子図書館資料

必要部分を検索・抽出する利用

内容に則した情報単位

アクセス可能性に応じて、より細かな単位にも

---

<sup>3</sup> マークアップ言語

テキスト中にタグ挿入などの「マークアップ」を行なって文書を記述するための言語。WWW 文書の記述に使われる HTML ( HyperText Markup Language ) により、一般に広く普及した。

<sup>4</sup> SGML ( Standard Generalized Markup Language )

汎用的なマークアップ言語として 1980 年代に確立された。タグで文書構造を明確に表現し、文書構造情報とレイアウト情報を分離して扱うことを基本理念とする。対象とする文書の構造に応じてタグは自由に設定でき、文書本体に先立つ DTD ( Document Type Definition 文書型定義 ) にタグと文書構造を定義する。

HTML は WWW 文書表示用に定義された DTD を背後に持ち、SGML のサブセットである。

<sup>5</sup> XML ( eXtended Markup Language )

インターネットの発達を背景に作られた SGML のサブセット。自由なタグ設定による十分な表現能力という SGML の利点を残したうえで、簡略化やネットワーク対応の拡張等が行なわれている。

<sup>6</sup> RDF ( Resource Description Framework )

メタデータを記述する汎用的枠組みの規約。データ要素を定義する ( Dublin Core のような ) メタデータ規則ではなく、データの表現方法・構文を規定するものである。どんなメタデータ規則にも適用できる。

「リソース」に対して ( 複数の ) 「属性」と「属性値」を記述していくモデルであるが、構文記述には XML を使用している。

#### 構成要素（構成部分）の記述

- ・ 記述方法
- ・ 資料構造・文書構造の表現（階層構造表現）

#### その他の関係構造

- ・ 様々な情報単位の交錯
- ・ 様々なバージョンの発生

## 6 . おわりに

#### 電子図書館資料に対するメタデータ標準化

- ・ 単一規則を定められるか？
- ・ データ交換の範囲とメタデータ規則の精度
- ・ 十分な記述内容と論理的な構造

#### 図書館目録の今後

- ・ 「電子図書館資料」だけの問題ではない