

## 2.2 メタデータ

### 2.2.1 はじめに

メタデータとはデータに関するデータであり、情報の特徴や性質を表す二次情報、高次情報である。従来の図書館で使用されてきた例としては、目録、索引、識別子などが挙げられる。また一方で、インターネット上を流れるデータにメタデータを持たせる環境が整ってきた。しかし現在のところ、これらの間をつなぐ技術は存在しない。

本研究では、現在標準化が進みつつあるメタデータ Dublin Core の実装を通じて標準化に貢献すると共に、アプリケーションからデータ配信への要求を伝えるしくみを構築する。また、アプリケーションが要求するネットワーク基盤を提案する。

### 2.2.2 背景

現在の計算機、ネットワークシステムにはデータを統一的に管理するしくみが存在しない。そのため、データの作成者以外はデータの性質、格納場所等の属性情報を知ることができず、再利用や検索を困難にしている。一方で、従来の図書館、博物館は目録を作成することによって各機関の資産を管理してきた。この目録を計算機システムに応用し、メタデータとして定義する動きが現実には稼働を始めてきた。メタデータはインターネットのコミュニティでも認められ標準化が進みつつある。

メタデータの国際的な標準化が進んでいるのは、Dublin Core である。すでにインターネットの標準化として絶大な力をもつ RFC(Request For Comment)が出版され、実装が進められている。アメリカ、オーストラリアなどではすでに国家プロジェクトとして図書館、博物館データの Dublin Core 化が進められている。

### 2.2.3 研究の目的

電子図書館には、データ蓄積機関である図書館という要素とインターネット上の大規模データサーバという要素とがある。Dublin Core は従来の書誌情報を管理してきた図書館員と、XML や WWW のページ作成の研究開発に携ってきたグループとが中心となって、標準化が進められている。

一方、奈良先端大のように計算機、インターネットを専門とするグループが進めている電子図書館の研究では、大規模サーバとしての視点からシステムを検討していくことができる。

電子図書館はデータを蓄積し、効率のよい検索が行えるだけでなく、インターネット上に効率よくデータを配信できるシステムであるべきである。データ配信までを統括したシステムによって、インターネットを効率的に利用できるしくみを構築するこ

とを目的とする。

#### 2.2.4 対象とするデータ

本研究では、前章で述べた講義ライブラリのデータを対象として、メタデータの実装、実験を行う。教育用のコンテンツは、メタデータとしても最も注目されている分野の一つであり、特に標準化を進めているアプリケーションである。

#### 2.2.5 grow metadata

メタデータを付与するにあたっての大きな問題の一つは、メタデータ自体の品質である。特にデジタルデータの場合、作成、デジタル化、コンテンツとして公開、管理などはそれぞれ別の人または部署が行うことが考えられ、データの内容を十分に表現できるメタデータが付与されるとは限らない。特に計算機やインターネットの研究者がデジタル化データを扱う場合、技術に関しては専門であるが、その内容についての専門的知識は充分でないことが多い。

そこで我々は、公開後も変化するメタデータ "grow metadata" を提案する。作者や作成日付などは変化しないメタデータであるが、内容に関する記述は、視点や背景などによって新たな解釈が生まれてくる可能性がある。また、データに対するアクセス回数はそのデータに対する人気を示し、すなわち内容の評価となる可能性がある。

一方で、アクセス状況はネットワークの使用状況を調査する指針ともなり、データやサーバの分散化にあたっての基礎データとなる。

#### 2.2.6 今後の予定

提案した "grow metadata" を実装し評価する。また本サーバの実装にあたり、教育用コンテンツに必要なメタデータを定義し提案する。教育用サーバにはさまざまな形態が考えられるが、その一つのケースとして評価する。

また、電子図書館をインターネット上のサーバとして考えたときのネットワークシステム全体を評価、モデル化し、次世代の情報システムの方向を提案する。