

B_026

博物館情報を用いた複合的メタデータスキーマの実現手法に関する検討 On Compound Metadata Schema Processing for Museum

秋元 良仁†
Ryoji AKIMOTO

亀山 渉†
Wataru KAMEYAMA

1. はじめに

情報技術の活用が急速に進展し、利用者が取り扱うことのできる情報量は飛躍的に増大している。特に、博物館においては古来から人類の営みを表現した有形・無形の文化財が多数管理されており、近年は文化財のデジタルコンテンツ化が進んでいる。また、これらのコンテンツにいつでもどこでもアクセスできる、ユビキタスな情報環境基盤が整いつつある。

このような環境では、様々なコンテンツから利用者の要求に応じて意味のある情報を横断的に紡ぎ出すことが求められる。そのためには、それぞれのコンテンツが何を表現しているのかを示す情報、即ち、コンテンツに関する情報を記述したデータ(メタデータ [1])を活用する必要がある。メタデータは定義や用途目的が広範であることから、コンテンツの特性を考慮した記述項目と記述形式(メタデータスキーマ)が提案されてきた。また、幾つかの国際標準や業界標準等が制定されている状況にある。しかしながら、これらの標準を利用しても、コンテンツの横断的な利用には様々な技術的課題が存在する。

そこで、本研究では2節でメタデータ、特に博物館分野におけるメタデータの問題点を明確にする。次に3節で横断的なコンテンツ利用を目的とした「ファジー・スキーマ」という概念を提案する。更に、複数の博物館メタデータスキーマを例に、ファジー・スキーマの基本的な言語設計の指針を検討する。

2. メタデータの問題点

2.1 メタデータスキーマ構築の問題

メタデータスキーマは、何をどのように管理すべきか、その項目名や適用範囲、データ型等が厳密に定義される。そのため、通常その分野に精通した専門家が構築を行う。これは時間を要するため、メタデータスキーマの普及とソフトウェアでの利用の間にタイムラグが生じやすい。また、一度メタデータスキーマが構築、普及すると、新たな追加事項等に対して変更等を加えることが困難となる。

2.2 メタデータ作成の問題

コンテンツの属性記述や維持・管理用の情報等、テキスト部分のメタデータ作成は人手により記述を行っている現状がある。現在既に膨大な量のコンテンツが存在しており、これらのコンテンツに対して適切なメタデータを作成、付与する作業は人的・金銭的・時間的に見て非常に負荷の高い作業と言える。

表1 博物館メタデータスキーマの種類

名称	内容
ミュージアム資料情報モデル	東京国立博物館が国内向けに提案。国際標準と互換あり。博物館業務支援及び情報共有を目的とする。
CIDOC/IC	国際博物館会議(ICOM)のドキュメンテーション委員会(CIDOC)が提案する国際標準ガイドライン
CIDOC/CRM	ICの情報共有を目的とした概念参照モデル
SPECTRUM	英国の博物館ドキュメンテーション協会(MDA)が提案する手続き型モデル
CDWA	米国のGetty Research Instituteが中心となり策定された美術情報のメタデータスキーマ
遺物分類標準	韓国国内で利用されるメタデータスキーマ、Dublin Coreベースの基本16項目、オプション116項目を有する

2.3 メタデータ・マッピングの問題

多様なコンテンツから意味のある情報を紡ぎ出すためには、複数の異なるメタデータスキーマ間を横断する形でコンテンツが活用されなければならない。しかし、メタデータスキーマは分野や応用ごとに固定的に形作られているため、相互のメタデータスキーマにおいて知識のマッピングを行う必要がある。これは議論を要する問題であり、異論のない整合の取れたマッピングを実現することは困難である。

2.4 博物館分野におけるメタデータの現状と問題点

博物館は古くから台帳あるいは目録という形で文化財に関する博物館メタデータが整理されてきた。

近年、博物館メタデータの記述方式を整理・モデル化し、システム化する動きがある。これまで、各博物館はそれぞれ独自で博物館メタデータを蓄積し、システム化してきた。しかし、より効率的な博物館の業務遂行や博物館間での情報共有の必要性から、標準的な博物館メタデータの利用が求められている(表1)。

博物館分野においても、前述のメタデータの問題を内在する。特に、博物館間における情報共有の必要性の高まりから、メタデータ・マッピングに関する話題は議論を呼んでいる。

関連研究として、近年ではDublin Coreやミュージアム資料情報構造化モデルへのマッピング研究 [2]、オントロジを用いたメタデータの分類や分散オントロジを用いた博物館情報横断検索 [3] 等がある。

† 早稲田大学大学院国際情報通信研究科

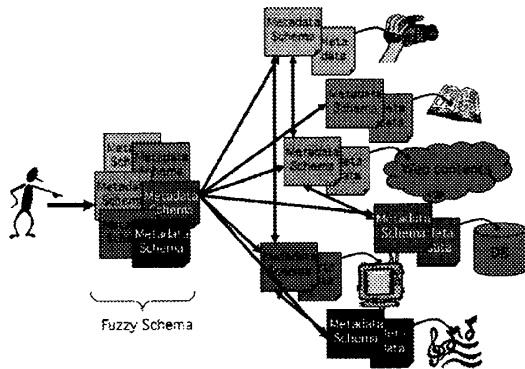


図1 柔軟な知識管理体系に基づくコンテンツ管理

3. ファジー・スキーマの検討

3.1 ファジー・スキーマの提案

メタデータの問題を解決するために本研究では、目的に応じて柔軟に形を変化させる知識管理体系を構築する技術を提案する。

人間は思考をする際、必要とする情報(コンテンツ)に関する知識管理体系(メタデータスキーマ)を状況や環境、経験から都度創り出し、その中で必要に応じて柔軟に取捨選択を行っていると考えられる。これに対し、既存のコンテンツ管理手法では、メタデータスキーマの内容項目の変更や改編等は通常行われない。

既存のコンテンツ管理体系を「静的なメタデータスキーマ」と呼ぶならば、人間が想起する知識管理体系は「動的なメタデータスキーマ」と言うことができよう。本研究では、双方の定性的性質を捉え、どこまでが静的に管理されるべき体系であり、どこからがコンテキストに応じて動的に創り出される体系であるのか、その最適値の算出を定量的に行うことを考える。そして、その最適値算出方法をベースとし、必要に応じて双方を緩やかにチューニングできる「ファジー・スキーマ」という概念を提案する。

ファジー・スキーマを用いることで、ある程度自動的にメタデータスキーマを作成することが可能となるため、メタデータスキーマ構築を簡易化することが期待できる。併せて、これまで人手によっていたメタデータ作成の(半)自動化も見込まれる。更に、メタデータスキーマを複数個連携させることで、各コンテンツのメタデータスキーマの差異を意識することなく、複合的にコンテンツへのアクセスが可能となることが期待できる(図1参照)。

3.2 アプローチ

ファジー・スキーマを実現するいくつかのアプローチが考えられる。

例えば、オントロジーやシソーラスを用いて知識処理する方法がある。これらは意味的にメタデータスキーマを関連付けられるというメリットを持つ反面、知識は整備や整合性の保証が困難であり、記述範囲も狭さも問題と

なる。

また、アノテーションやロコミ情報を利用するアプローチも考えられる。これらの情報の利用は、既に付与されているメタデータ以上の情報が取得できるというメリットがある反面、誰がアノテートするのかという問題と、アノテーション自体の信頼性の保証という問題が存在する。

本研究では、これらに対して統計処理手法によるアプローチを取る。

ファジー・スキーマではあるメタデータスキーマを静的な部分と動的な部分を切り分ける最適値の定量的な算出を行う。その過程において動的に生成できる部分が他のメタデータスキーマのそれとどの程度オーバーラップしているのか、その尺度を計算することで異なるメタデータスキーマ間で確率的なメタデータ・マッピングが可能なのではないかと考えた。尺度としてパターン認識や統計処理で用いられている手法を検討している。また、オントロジーやアノテーションによるアプローチも統計処理的な手法の比較対象として用いることを検討している。

3.3 言語設計の指針

現在、ファジー・スキーマ実現のため、基本的な言語設計を行っている。

具体的には、前述した博物館で利用されているメタデータスキーマを用い、手作業によるメタデータスキーマの分類を行うことで自動化や統計処理の手がかりとする。

特に、様々なメタデータスキーマの手本となっている Dublin Core と、今後日本で中心的なメタデータスキーマになるであろうミュージアム資料情報構造化モデルをベースに、特徴的と思われる部分を重点的に分類を進めて行く予定である。

4. まとめと今後の課題

本稿では、博物館を中心とするメタデータの問題を整理した。そこでは、メタデータスキーマを構築する難しさ、メタデータを作成する手間、既に作られているメタデータスキーマを相互に利用する難しさが問題となっていることがわかった。そこで、これらの課題を解決するために必要に応じて柔軟にメタデータスキーマを構築できる「ファジー・スキーマ」という概念の提案を行った。更に、ファジー・スキーマ実現のためのいくつかのアプローチを提示し、その言語設計の指針を示した。今後は、具体的な言語設計を行い、ファジー・スキーマの構築に取り組んで行く予定である。

参考文献

- [1] Dempsey et al.: "Metadata: A Current View of Practice and Issues", J. of Documentation, Vol. 54, No. 2, pp. 145-172(Mar. 1998).
- [2] 原ほか: "データベースの共有におけるデータマッピングの事例的研究", 情処研報, Vol. 2005, No. 76, pp. 31-38(2005-07)
- [3] 山田ほか: "博物館情報横断検索における分散オントロジーの検討", 画像電子学会研究会 (2006-03)