

博物館情報を用いた複合的メタデータスキーマ 構築手法「ファジィ・スキーマ」の言語設計

秋元 良仁[†] 亀山 渉[†]

[†] 早稲田大学大学院国際情報通信研究科

近年の情報技術の進展に伴い、利用できるコンテンツの量は飛躍的に増大している。また、これら膨大な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環境も整いつつある。このような状況では、メタデータを用いて多種多様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技術が求められる。そこで、本稿ではまずメタデータの現状と問題点を、特に博物館を事例として整理する。それをもとに、複合的にメタデータスキーマを構築する手法である「ファジィ・スキーマ」という概念を提案する。更に、複数の博物館の資料情報管理用メタデータスキーマを例に、ファジィ・スキーマを構築するための言語設計について検討する。

Language Design of Creating Compound Metadata Schema Method “Fuzzy Schema” Based on Museum Information

Ryoji Akimoto[†] Wataru Kameyama[†]

[†] Graduate School of Global Information and Telecommunication Studies, Waseda University

As information technology has progressed recently, the amount of contents that can be used increase. And, there is an environment in which lots of contents can be accessed at anytime and anywhere. In this situation, a technology is needed to extract useful information from various contents by using metadata in order to meet the requirements. In this paper, we summarize the current state of metadata technology and problems. Based on it, we propose the concept called Fuzzy Schema. Fuzzy Schema can flexibly construct metadata schema if necessary. And we also describe the design of Fuzzy Schema language taking an example of metadata schema of museum.

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加している。特に、近年博物館においては古来からの人類の営みを表現した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アクセスするためのユビキタスな情報環境基盤も整いつつある。

このような環境では、利用者の要求に応じて多様なリソースからコンテンツを横断的に取得し、複合的に活用することが求められる。

そのためには、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。これまでにも、コンテンツの特性を考慮した様々なメタデータが提案されており、また、幾つかの国際標準や業界標準が制定されている状況にある。

しかしながら、これらメタデータには根本的な技術的課題が内在しているため、その活用は不十分と言わざるを得ない。

そこで、本研究ではまずメタデータ、特に博物館

分野におけるメタデータの技術的課題を明確にする。次に、課題解決のため、博物館情報に基づき複合的にメタデータスキーマを構築する手法である「ファジィ・スキーマ」という概念を提案する。更に、ファジィ・スキーマを構築するための言語設計について検討を加える。

2 メタデータが抱える技術的課題

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、DTD や XML Schema に代表される、何をどのように管理すべきか、その項目名や適用範囲、データ型等が定義されるものと、RDFS や OWL に代表される、項目間の関係や意味を記述したものに大別される。

いずれの場合も、通常その分野に精通した専門家が構築を行う。これは膨大な時間と手間を必要とするため、構築と利用の間にタイムラグが生じやすい。また、一度運用が始まると、追加・削除等の変更を加えることが困難となる。

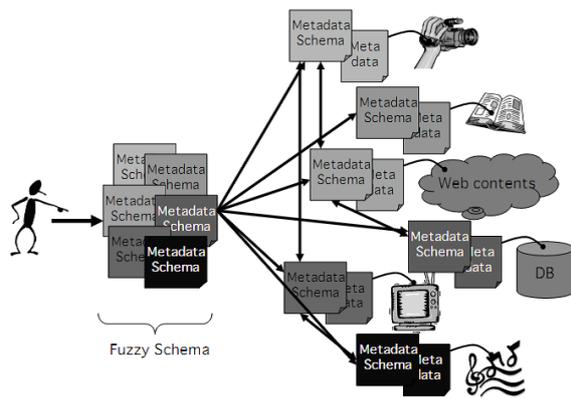


Fig. 1 ファジィ・スキーマに基づくコンテンツ管理

Table 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	CIDOC 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

2.2 メタデータ作成の問題

コンテンツ利用の促進には、メタデータが大量に作成され、インターネット等のネットワーク上に流通する必要がある。メタデータの作成は、一部自動化も見られるが、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。適切なメタデータの作成作業は人的・金銭的・時間的に見て非常に負荷の高い作業と言える。

また、近年のコンテンツ量の増加に伴うメタデータ量の増加も考慮しなければならない問題の一つと言える。

2.3 メタデータ・マッピングの問題

複数のデータベースやアーカイヴ等、多様なコンテンツから意味のある情報を横断的に取得するた

めには、異なったメタデータスキーマに基づくメタデータ間のインターオペラビリティをどのように保証するのかという問題も存在する。

現状、手動による項目間のマッピングやラップによる相互利用が行われているが、項目間の不一致や変更が頻繁に起こるため、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メタデータ (博物館メタデータ) が管理されてきた。1990 年代以降、データの相互利用や業務遂行を目的とした博物館メタデータの記述方式の統一化が見られる (Table1)。

しかし、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おり、いくつかの関連研究も見られる^{2) 3)}。

3 ファジィ・スキーマの検討

3.1 ファジィ・スキーマの提案

あるコンテンツを利用する場合、人間の思考は目的のコンテンツに関するメタデータスキーマを状況や環境、経験に応じて動的に構築し、そこで定義されたメタデータを必要に応じて緩やかにチューニングしていると考えられる。

そこで、本研究では人間の思考のように緩やかにメタデータをマッピングさせるため、複合的にメタデータスキーマを作成する手法「ファジィ・スキーマ」を提案する。

具体的には、まず実際に運用されている博物館用のメタデータスキーマの定性的な特性を調べ、いくつかのマッピング・パターンに分類する。更に、スキーマが定義する各項目が他の項目とどの程度類似しているのか、あるいは類似していないのかを示す尺度「あいまい度」を定義し、マッピング・パターンと組み合わせることでメタデータスキーマを複合的に作成する。

組み合わせによる項目の再利用を行うことで、メタデータスキーマ構築及びメタデータ作成の負荷軽減が図れることが期待できる。併せて、あいまい度を用いて詳細にメタデータスキーマをチューニングすることでメタデータ・マッピングの精度向上が期待できる (Fig1 参照)。

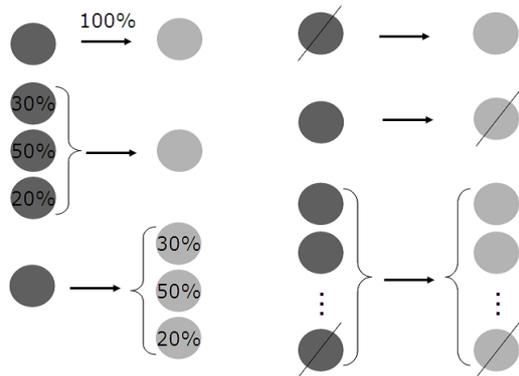


Fig. 2 マッピング・パターン

3.2 ファジィ・スキーマの言語設計

本研究ではファジィ・スキーマを表現するために、マッピング・パターン及びあいまい度を XML 形式で記述表現する言語の設計を検討している。本稿では、マッピング・パターンに関する検討内容を示す。

マッピング・パターンとは、写像のパターンのことを指し、本稿ではメタデータスキーマで定義されるメタデータの項目対応のパターンを指す。前述の通り、これまでも項目の対応付けは行われてきたが、全射や単射が中心であり、かつ人手を介する写像であったため、項目の不一致等が見られた。

そこで、本稿ではまず項目対応時に考えられる写像のパターンを想定し、そのパターンを記述できる言語を設計する。

マッピング・パターンを想定するにあたり、1) 博物館で利用されている収藏品管理システムが管理するデータ項目⁴⁾、2) 国際・業界標準とされるデータ項目 (Table 1)、の 2 点に着目した。

1) の収藏品管理システムは国内外約 20 のシステムを調査したが、いずれも国際・業界標準に準拠しているため、本研究では特に 2) を中心に検討を行った。

Fig2 に想定されるマッピング・パターンの例を示

す。パターンは大きく 3 種に分類される。つまり、1 対 1 でマッピングできるもの、複数個の項目がある割合で 1 つの項目にマッピングできるもの、片方には存在せず (空集合) もう片方に存在するもの、である。特に、複数個の項目がある割合で 1 つの項目にマッピングされる例に関して、割合の配分は項目の類似度合を定量的に表現するあいまい度を適用することで解決を図る。

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技術的課題を整理した。そこでは、メタデータスキーマ構築の問題、メタデータ作成の問題、メタデータ・マッピングの問題が存在することがわかった。そこで、これらの課題を解決するため、複合的にメタデータスキーマを構築する手法である「ファジィ・スキーマ」という概念を提案した。更に、ファジィ・スキーマを構築するための言語設計の検討を行った。

今後は、マッピング・パターンの具体的な記述方法及びあいまい度の定量化アルゴリズムの検討を行っていく予定である。

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の検討", 画像電子学会研究会 (2006-03).
- 4) 秋元: "博物館の収藏品管理におけるメタデータの利用と問題点", 情処研報, Vol. 2004, No. 36, pp. 55-62(2004-03).