

# 博物館情報の分類マッピングを用いた横断検索 Integrated Access for Museum Information using Classification Mapping

画像電子学会 画像ミュージアム研究会 博物館・美術館 DTD-SG

山田 篤<sup>†</sup> 安達 文夫<sup>‡</sup> 小町 祐史<sup>§</sup> 河合 正樹<sup>¶</sup>

Atsushi YAMADA<sup>†</sup> Fumio ADACHI<sup>‡</sup> Yushi KOMACHI<sup>§</sup> and Masaki KAWAI<sup>¶</sup>

<sup>†</sup> 京都高度技術研究所

<sup>†</sup> ASTEM RI/Kyoto

<sup>‡</sup> 国立歴史民俗博物館

<sup>‡</sup> National Museum of Japanese History

<sup>§</sup> パナソニックコミュニケーションズ

<sup>§</sup> Panasonic Communications Co., Ltd.

<sup>¶</sup> インフォコム

<sup>¶</sup> INFOCOM Corporation

E-mail: <sup>†</sup> yamada@astem.or.jp, <sup>‡</sup> adachi@rekihaku.ac.jp, <sup>§</sup> komachi@y-adagio.com, <sup>¶</sup> kawai@infocom.co.jp

## 1. はじめに

画像電子学会画像ミュージアム研究会「博物館・美術館文書の文書型定義 SG (DTD-SG)」では、博物館、美術館が公開する文書情報の横断検索について検討している。本研究グループが提唱する

- I) 情報記述構造レベル
- II) 情報記述内容レベル
- III) 情報ナビゲーションレベル

からなる 3 層フレームワークについては[1][2]で報告し、このうち、レベル I の記述構造レベルに関するプロトタイプと評価については[3][4]で、レベル II の情報記述内容レベルの相互変換の概要については[5]で報告した。本稿では記述内容レベルの相互変換の具体的な実現として、分類マッピングを用いた横断検索について述べる。

## 2. 記述内容レベルの相互変換

### 2.1 相互変換という考え方

我々の提唱するフレームワークでは、各館の独自性、多様性を許容しながら、様々なレベルでの情報の相互変換を行うことにより、複数の館が公開するそれぞれの館蔵品に関する情報(文書)を横断検索し、検索者が望む情報を得ることを可能にすることを目的としている。現実の博物館がおかれている状況と館および館蔵品の多様性を考えたときに、横断検索のための共通のボキャブラリを制定した上で、すべての館がそのボキャブラリを用いて情報発信を行うというモノリシックなアプローチは非現実的であると考える\*。

レベル I で論じた記述構造レベルであれば、ある程度

\* 適用範囲、分野を限れば、共通ボキャブラリの制定が可能なる場合もある。

の標準化は可能かもしれない。実際に我々も、[4]では Dublin Core Metadata Element Set V1.1 (DCMES)[6]を構造変換のターゲットに設定し、スキーマレベルで複数の館の公開する文書情報を横断検索するプロトタイプについて述べた。しかしながら、これは博物館情報の横断検索において、唯一の標準的な記述構造を用いることを主張するものではない。変換ターゲットは必要に応じて異なるものが用いられてもよいと考えている。

これはレベル II の記述内容レベルの相互運用を考えたときにさらに鮮明になる。たとえ共通の構造をもつ文書で情報が提供されたとしても、

- 専門分野によるターミノロジーの違い
- 明示して記述されない背景知識

といった問題は常に生じる。これらを相互変換の仕組みの中で吸収・解決することを考える。このアプローチは、特定分野内において共通ボキャブラリが構築できたとしても、それを分野横断的に利用する際にも有効である。

### 2.2 記述内容としての分類

レベル I のプロトタイプにおいて、DCMES のスロット毎の検索は実現できたが、スロットに書かれる値に関しては何の制約も設けていなかったため、文字列の一致により検索を行っていた。各スロットに書かれる値は、各館毎に異なるポリシーに基づき記述される可能性があり、それらを横断的に取り扱おうとすると、その対応づけが問題となる。

これは様々な記述スロットにおいて問題となる。表記法の違いといった単純なものから、体系的な知識の援用を要するものまで様々な記述の間の対応関係が問題となるが、本稿では「それが何であるか」という点に焦点をあてる。実際に博物館に所蔵されているものに関して見

でも、一般に固有名だけからは、それが何であるかということとはわからない。そこで典型的なユースケースとして、複数の館に対して、ある分類に属するものを横断検索するというタスクを考え、各館の所蔵品に対して、それぞれの館が独自に分類し、それを記述しているときに、どうすれば、それらを横断検索することが可能になるかという問題設定を行った。

このとき、分類の個別性、文脈依存性が問題となる。同じものに対しても、観点が異なれば分類の仕方も異なる。また、館の性質によっても分類の仕方は違ってくるだろう。さらに分類の結果うまれたクラスに与える名称もそれぞれ異なっている可能性がある。これは、

- i. クラスの一致・不一致
- ii. クラス名の一致・不一致

の二つの観点から整理することができる。i について、ある館の分類の結果、生成されたクラスと完全に一致するクラスが別の館の分類結果に存在する保証はない。また ii について、たとえ類似のクラスが存在しても、それらに同じ名称が付与されているとは限らない。

### 2.3 分類マッピング

このような状況の下で、異なる館の情報を横断的に検索しようとするならば、分類体系間の対応関係をとらなければならない。これを分類マッピングと呼ぶ。その過程で背景知識の折り返しも行われる。

分類のマッピングもまた、ある観点から行うことになる。その結果、ある館の所蔵品に対して、異なる観点から作成された複数のマッピングが存在してもよい。博物館自身がこのようなマッピングを行い、運用することもありえるが、多くの場合、これは外部のポータルサイトの役割であると考えられる。そのようなポータルサイトが、それぞれ異なる目的で運営されるとき、複数のマッピングの存在は、ポータル毎の性格づけの結果である。

### 3. 分類マッピングを用いた横断検索の仕組み

複数の館、複数のポータルサイトが存在する場合の分類マッピングを用いた横断検索の概要を以下に示す。

- (0) 各館独自の分類に基づく所蔵品に関する記述は各館毎のレポジトリに格納されているものとする。
- (1) ユーザはあるポータルサイトにアクセスし、そのポータルサイトが提供する分類を用いた検索条件を入力する。
- (2) ポータルサイトではユーザに提供する分類体系と、各館が持つ分類体系のマッピングテーブルを持ち、ユーザからの検索要求を各館の分類体系の

語彙に変換して、各レポジトリに問い合わせる。

- (3) 各レポジトリはそれぞれの語彙で発行された問い合わせに対する検索結果をポータルサイトに返す。
- (4) ポータルサイトは各レポジトリから返ってきた検索結果を整形してユーザに提示する。

### 4. プロトタイピング

分類マッピングを用いた横断検索のサンプル実装と実験を行った。

実験のために、DTD-SG のメンバがそれぞれ仮想的な博物館を定義し、それぞれの所蔵品に関するメタデータを記述した。対応関係を重点的に見るため、今回はテーマを食器に絞り、各館毎に独自の分類体系で分類を行った。各館の分類体系のサンプルを図 1 に示す。

```
- <class_list>
- <class of="Museum_A" id="A.1" name="食器">
- <subclass id="A.1.1" name="皿">
  <subclass id="A.1.1.1" name="大皿" />
  <subclass id="A.1.1.2" name="小皿" />
  <subclass id="A.1.1.3" name="絵皿" />
</subclass>
- <subclass id="A.1.2" name="椀">
  <subclass id="A.1.2.1" name="茶碗" />
  <subclass id="A.1.2.2" name="湯飲み" />
</subclass>
- <subclass id="A.1.3" name="鉢">
  <subclass id="A.1.3.1" name="小鉢" />
</subclass>
```

(a) A 館の分類体系

```
- <class_list>
- <class of="Museum_B" id="B.1" name="食器">
- <subclass id="B.1.1" name="容器">
  <subclass id="B.1.1.1" name="多人数用容器" />
  <subclass id="B.1.1.2" name="個人用容器" />
</subclass>
- <subclass id="B.1.2" name="平板食器">
  <subclass id="B.1.2.1" name="大型平板食器" />
  <subclass id="B.1.2.2" name="個人用平板食器" />
</subclass>
- <subclass id="B.1.3" name="籠状食器">
</subclass>
- <subclass id="B.1.4" name="調理機能食器">
</subclass>
- <subclass id="B.1.5" name="補助食器">
</subclass>
```

(b) B 館の分類体系

```
- <class_list>
- <class of="Museum_C" id="C.1" name="食器">
- <subclass id="C.1.1" name="中華食器">
  <subclass id="C.1.1.1" name="皿" />
  <subclass id="C.1.1.2" name="茶器" />
</subclass>
- <subclass id="C.1.2" name="洋食器">
  <subclass id="C.1.2.1" name="大皿" />
  <subclass id="C.1.2.2" name="カップ&ソーサー" />
  <subclass id="C.1.2.3" name="カトラリー" />
  <subclass id="C.1.2.4" name="グラス" />
  <subclass id="C.1.2.5" name="グラタン皿" />
  <subclass id="C.1.2.6" name="スープ皿" />
</subclass>
```

(c) C 館の分類体系

図 1 各館の分類体系

次に、ポータルサイトで用いる分類体系として分類語彙表[7]を用いた(図2)。そして、このポータルサイトの分類体系と個々の館の分類体系との間のマッピングテーブルを人手で準備した(図3)。

```

- <class_list>
- <class of="BGH" id="1_4520" name="食器">
- <subclass id="1_4520_1">
- <subclass id="1_4520_1_1">
  <subclass id="1_4520_1_1_2" name="食器" />
  <subclass id="1_4520_1_1_3" name="茶器" />
  <subclass id="1_4520_1_1_4" name="茶道具" />
  <subclass id="1_4520_1_1_5" name="酒器" />
</subclass>
- <subclass id="1_4520_1_2">
  <subclass id="1_4520_1_2_1" name="茶道具" />
</subclass>
- <subclass id="1_4520_1_3">

```

図2 分類語彙表に基づく分類体系

```

- <class_list>
- <class of="BGH" id="1_4520" name="食器">
- <subclass id="1_4520_1">
- <subclass id="1_4520_1_1">
  <subclass id="1_4520_1_1_2" corr="食器" />
  <subclass id="1_4520_1_1_3" name="茶器" />
  <subclass id="1_4520_1_1_4" name="茶道具" />
  <subclass id="1_4520_1_1_5" name="酒器" />
</subclass>
- <subclass id="1_4520_1_2">
  <subclass id="1_4520_1_2_1" name="茶道具" />
</subclass>
- <subclass id="1_4520_1_3" corr="皿">

```

(a) A館の分類体系との対応

```

- <class_list>
- <class of="BGH" id="1_4520" name="食器">
- <subclass id="1_4520_1">
- <subclass id="1_4520_1_1">
  <subclass id="1_4520_1_1_2" name="食器" />
  <subclass id="1_4520_1_1_3" name="茶器" />
  <subclass id="1_4520_1_1_4" name="茶道具" />
  <subclass id="1_4520_1_1_5" name="酒器" />
</subclass>
- <subclass id="1_4520_1_2">
  <subclass id="1_4520_1_2_1" name="茶道具" />
</subclass>
- <subclass id="1_4520_1_3">

```

(b) B館の分類体系との対応

```

- <class_list>
- <class of="BGH" id="1_4520" name="食器">
- <subclass id="1_4520_1">
- <subclass id="1_4520_1_1">
  <subclass id="1_4520_1_1_2" name="食器" />
  <subclass id="1_4520_1_1_3" corr="茶器" />
  <subclass id="1_4520_1_1_4" name="茶道具" />
  <subclass id="1_4520_1_1_5" corr="酒器" />
</subclass>
- <subclass id="1_4520_1_2">
  <subclass id="1_4520_1_2_1" name="茶道具" />
</subclass>
- <subclass id="1_4520_1_3">

```

(c) C館の分類体系との対応

図3 各館の分類体系との対応

以上のデータを用いて、ユーザが分類語彙表の食器分

類に従って分類項目を選択すると、各館の所蔵品のうち、当該分類に対応すると考えられる館ローカルの分類に属するものを結果として返すプロトタイプを作成した。

今回は、分類マッピングの動作検証が主な目的であったため、ネットワークは用いず、準備したメタデータ、分類体系、マッピングテーブルをすべてXMLファイルとして持ち、これらのファイルを読み込んで動作するウェブブラウザベースのスタンドアロンのシステムとして実装した。画面イメージを図4に示す。

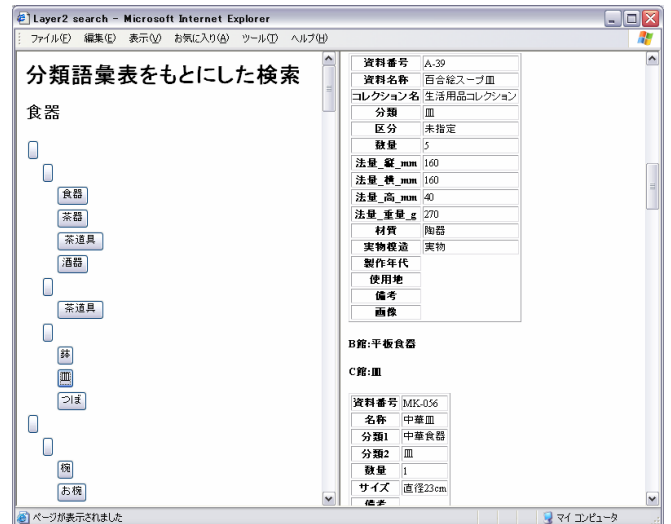


図4 分類マッピングを用いた検索システム

画面左に表示されている検索用語彙からどれかを選択すると、マッピングテーブルから各館毎に対応するクラス名を取り出し、各館のレポジトリに対して検索を行う。各館毎にレポジトリへの格納項目が異なるため、検索結果は各館毎に定められたテーブルとして、画面右に表示されている。

ユーザが選択した分類項目から対応する各館ローカルの分類項目への変換、各館の所蔵品リストの中から対応する分類項目に属するものの選択、各館毎に選択された所蔵品のメタデータの整形は、それぞれXSLTを用いて実現した。

## 5. 考察

分類マッピングを用いることにより、ポータルサイトの語彙を介して複数の館の記述内容の橋渡しをして、それぞれ独自の分類体系を採用している複数館の所蔵品情報を横断検索できる可能性が示された。

今回、マッピングはそれぞれの館の担当者に作成を依頼した。その際に、うまくマッピングできない、対応する適切なクラスがない、といった問題はあった。これは、実験に用いた分類語彙表が、一般的な日本語のシソーラ

スとして構築されたものであり、検索用途に設計されたものではないこと、データとしても若干古かったことも影響していると考えられる<sup>†</sup>。分類語彙表の分類番号をもとに中間層を設定したため、下位区分の総称がつけられるべきところが無名になったことも理解を困難にした。

これはマッピングの品質をいかに担保するかという問題につながる。今回は実験のために手作業で対応付けを行ったが、実際には何らかの自動化ないし半自動化の手だてが必要となる。

また、今回の実験ではクラス間のマッピングのみを取り扱い、クラス階層を上がる (汎化)、下がる (例化) といったオペレーションは実現しなかったが、こういった仕組みが有効な場合もありうる。中間ノードへの対応付けを許すか否か

クラスの多重継承に関しては、それは異なる観点からの分類であるにとらえて、今回は分類を木構造で表すことにした。ただ、実際には検索時にも異なる観点の統合、切り替え (context switching) を要することがある。

より厳密なクラス定義や、等価なクラスの発見、属性値をもとにインスタンスとクラスを関連づけるといった話題は、オントロジ関連の研究成果を援用し、OWL[8]等のオントロジ記述言語を用いて表現することにより、実現の可能性はあるが、それが現実的に運用可能な解であるか否かはさらなる検討を要する。

## 6. おわりに

本稿では、博物館、美術館が公開する文書情報の横断検索を目指して、記述内容レベルでの相互変換の例として、分類マッピングを用いて、各館毎に異なる分類体系を採用していても、それらの間の横断検索を可能とする方式について述べた。記述内容レベルの相互変換のための知識表現としてはオントロジを用いることにより、異なる知識源の統合も視野に入れた記述内容レベルの相互変換を行える可能性がある。

今後、さらにこのフレームワークに関する検討を重ねていく予定である。

## 文 献

- [1] 画像電子学会 VMA 研究会博物館・美術館 DTD-SG, “博物館情報の知的横断検索のためのフレームワーク,” 2002 画像電子学会第 30 回年次大会画像電子ミュージアムテクニカルセッション, pp.75-76, Jun. 2002.
- [2] 画像電子学会 VMA 研究会博物館・美術館 DTD-SG, “博物館情報の知的横断検索の試み,” 2002 画像電子学会第 30 回年次大会画像電子ミュージアムテ

クニカルセッション, pp.77-78, Jun.2002.

- [3] 画像電子学会 VMA 研究会博物館・美術館 DTD-SG, “博物館情報横断検索その役割と課題,” 画像電子学会第 10 回 VMA 研究会, Jan.2003.
- [4] 画像電子学会 VMA 研究会博物館・美術館 DTD-SG, “博物館情報横断検索のための記述構造相互変換のプロトタイピングと評価,” 画像電子学会第 10 回 VMA 研究会, Jan.2003.
- [5] 画像電子学会画像ミュージアム研究会博物館・美術館 DTD-SG, “博物館情報横断検索のための記述内容レベル相互変換,” 2003 画像電子学会第 31 回年次大会, Jun.2003.
- [6] Dublin Core Metadata Element Set (<http://www.dublincore.org/usage/terms/dc/current-elements/>).
- [7] 分類語彙表, 国立国語研究所, 1964.
- [8] OWL Web Ontology Language (<http://www.w3.org/TR/owl-ref/>).

<sup>†</sup> 分類語彙表は 2004 年 1 月に増補改訂版が出版されているが、今回の実験では旧版を用いた。