

博物館資料情報統合検索のためのコアメタデータ

山本 泰則* 安達 文夫†

* 国立民族学博物館 † 国立歴史民俗博物館

本稿では、博物館がもつ人文科学系の実在資料情報を統合的に検索・共有するために「博物館コア」という小さなメタデータ要素セットを新たに提案している。また、人文科学のより広い分野の研究資源の中で博物館の資料情報を共有・活用するために、インターネット上の情報資源の発見や記述に用いられる Dublin Core Metadata Element Set と博物館コアメタデータとの対応関係について議論している。

Core Metadata to Search for Museum Object Information across Databases

Yasunori YAMAMOTO* Fumio ADACHI†

* National Museum of Ethnology † National Museum of Japanese History

In this paper, we propose *museum core*, a small set of metadata elements used to search for information of non-digital museum objects across databases maintained by the separate museums. We also discuss the crosswalk from the *museum core* to Dublin Core Element Set so as to retrieve the museum object information together with research resources in the humanities.

1 はじめに

博物館が収蔵する資料の目録情報は、従来は台帳として記録されてきたが、近年これを電子化してデータベースを作成し、インターネットに公開することが広く行われている。また、資料の管理や博物館どうしの資料情報の交換・共有を目的として、情報の標準的な記述法やそのモデルが提案されている [1, 2, 3, 5]。しかし博物館では、すでに独自の様式で相当数の資料情報を蓄積しているため、それらを新たな様式で記述し直すことは必ずしも容易ではない。むしろ機械的な対応づけ—マッピング—程度の操作が現実的である。

一方、情報の記述と発見に比較的汎用的に用いられているメタデータとして Dublin Core Metadata Element Set [7] (以下、Dublin Core と

呼ぶ) がある。これを共通メタデータとして、人文科学分野のデータベースを横断的に検索するシステムが公開されているが [8, 9], Dublin Core は元来インターネット上の電子的な情報資源の記述・発見を目的に規定されたものであるため、博物館のもつ実在資料に適用するにはなじまない面がある。

本稿では、博物館の特に人文科学系の実在資料情報の統合的な検索・共有を目的として「博物館コアメタデータ」(以下、「博物館コア」と呼ぶ) を新たに提案する。さらに、人文科学のより広い分野の研究資源の中で、博物館の資料情報を共有・活用するために、博物館コアメタデータと Dublin Core との対応関係について検討を加える。

2 実在資料記述における Dublin Core の問題

人間文化研究機構では、機構を構成する各機関がもつ合計 100 以上にのぼる人文科学系データベースを横断的に検索する「統合検索システム」[9]を作成し公開しているが、それを実現するための共通メタデータとして Dublin Core を用いている。しかし、人文科学分野の実在資料に関する情報を Dublin Core に変換する場合、さまざまな困難が生じる [10, 11]。

まず、元の資料データベースの項目を Dublin Core のどの要素に割り当てればよいか判断に迷う場面がある。たとえば、楽器の種類情報はふつう type に割り当てるのが自然であるが、データベースによっては subject あるいは format に割り当てた方が適切な場合もある。土器資料の形状情報も type, discription など、割り当て先は複数の可能性がある。

次に、実在資料は、製作地、使用地、出版地など資源に関わる場所の情報が重要である。しかし、Dublin Core では、資源のライフサイクル上で生じた事象の時間情報（たとえば、製作時期、使用年代、出版年月日など）を記述する要素 date はあっても、これに対応する空間的情報を記述する要素はない。coverage は資源の内容に関する時間的・空間的範囲を記述する要素であるので、ここに製作地、使用地などの情報を入れることは適切ではない。インターネットは資源の空間的な位置がほとんど意味をなさない世界であるが、実在資料では場所は重要な情報のひとつである。

以上のような問題は、実在する資料、特に「もの」資料で顕著に生じる。それは「もの」資料が必ずしも何かを明示的に記述しているわけではないからである。それと同時に、「もの」資料は形状や材質、その資料が置かれていた状況や文脈など、さまざまな暗黙の情報をもっている。

また、手紙のように本来は紙に書かれた記述内容が重要であるが、時代を経て古文書になると、紙の大きさや紙質、折り目など、「もの」としての性質も重要性を増す場合が少なくない。

その他、Dublin Core は何を記述対象と考えるかよって、各要素の内容が非常に敏感に変わるメタデータであるといえる。たとえば、音楽演奏を撮影したビデオテープがあったとき、録画したビデオテープを記述対象とするなら、ビデオ撮影者は creator、音楽の演奏者は contributor となる。一方、音楽演奏そのものを記述するなら、演奏者が creator、作曲者などが contributor、ビデオ撮影者は、どの要素にも記述されない可能性もある。Dublin Core を使った情報検索では、何を記述対象にしているかを常に意識している必要があり、このことはデータベースの検索利用者にとって必ずしも容易なことではない。

3 博物館資料情報の標準化

博物館資料情報の共有や情報交換のため、情報記述の標準化の提案がさまざまになされている。

標準化を積極的に推進する組織のひとつに国際博物館会議 (ICOM) の国際ドキュメンテーション委員会 (CIDOC) がある。CIDOC は 1994 年に「博物館資料の最小限情報分類報告」(MICMO) [1] を提案し、その中で博物館資料の記述に必要な最低限の情報として 17 のカテゴリを規定した。また翌年の 1995 年には、MICMO を発展、詳細化した「博物館資料情報のための国際指針」(IGMOI) を提案している。IGMOI は 22 の情報グループと 74 の情報カテゴリから構成されている。

一方、日本では、2005 年に東京国立博物館が「ミュージアム資料情報統合化モデル」[6] を発表している。これは、美術館・博物館における歴史・民俗・考古・美術の各種資料がもつ情報とその記述をモデル化したもので、4 種類の性格に分類された 34 の属性から構成されている。

以上は、博物館資料情報を記述するためのなんらかの標準的なメタデータを提案するものであったが、これとは別のアプローチとして CIDOC が 1998 年に発表した「概念参照モデル」(CRM: CIDOC Conceptual Reference Model) [3] がある。これは博物館資料にかかわるさまざまな基本概念とその関係をオントロジを用いて記述したもので、基本概念と各博物館のもつ資料情報を対応づけることにより、博物館ごとに構造の異なるデータベースをオントロジを介して橋渡しをしようとする試みである。

4 博物館コアメタデータの提案

前節のようにすでに提案されている博物館資料情報のためのメタデータの利用を検討したが、どれもわれわれの目的に適したものとはいえなかった。MICMO はメタデータの要素数は妥当であるが、管理的な情報を除けば資料の基本的な物理的記述に限定されている。逆に、IGMOI や「ミュージアム資料情報構造化モデル」は、要

表 1 博物館コアメタデータの要素

要素	意味
1. 資料ID	資料を識別する記号、番号。
2. 名称 *	資料の名称、題名。
3. 時間 *	資料に関わる時間的情報。 日付、年代、時代など。
4. 空間 *	資料に関わる空間的地理的情報。 地名、住所、年代、緯度経度など。
5. 人物 *	資料に関わる人、団体、製作者、 使用者、収集者など。
6. 分類 *	資料の分類、種別。
7. 主題・ キーワード *	資料に関わる主題、キーワード。
8. 大きさ・重量	資料の大きさ、重量。
9. 数量	資料の構成物の数。
10. 形状	資料の物理的形状の特徴。
11. 材質	資料の材料、材質。
12. 記述	資料に関する記述的な情報。
13. 関係	他の資料との関係。セット、 全体・部分、レプリカなど。
14. その他	上記要素に該当しない情報。

(*印の要素は検索に利用)

素数が多く、内容が詳しすぎる。また、CRM は、情報科学的には興味ある枠組みであるが、内容が複雑で、だれもが容易に理解して使えるものではないと判断した。

以上のことを考慮し、新たに提案する博物館資料情報のためのコアメタデータに求める要件を、以下のように定めた。

1. 資料情報の共有、横断的な検索、検索結果の一覧表示を目的とする。(資料の管理や資料記述の新たな標準を提案するものではない。)
2. 資料の細分化した詳細記述ではなく、Dublin Core 程度の要素数で、既存の目録情報から簡単に機械的に変換(マッピング)できる。
3. 博物館の人文科学分野の資料に適した要素を含む。
4. 必要に応じて、元の(各博物館固有の)データベースのレコードが容易に参照できる環境で利用される。

これらを念頭に置きながら、われわれが所属する国立民族学博物館、国立歴史民俗博物館の資料データベースを中心に資料情報を分析し、前節の標準メタデータも参考にしながら要素を抽出し、表 1 のように博物館コアを定義した。

要素 3~5 には、資料そのものおよび資料に記載された内容について、「時間」、「空間」、「人物」の視点から情報を記述する。要素 8~11 には、資料の物理的な属性を記述するもの、要素 14 には、共通メタデータとして収容しきれない各データベース固有の情報を収容するためのものである。

5 Dublin Core との関係

本稿では、Dublin Core になじまない博物館の実在資料に適したメタデータとして博物館コアを提案した。しかし、世の中のさまざまな情報に Dublin Core でメタデータが付与されている以上、それらとの対応を考察し、より広範囲な人文科学の研究資源の中で、博物館資料情報を

含めた統合的検索の可能性を探ることは有用である。

博物館コアと Dublin Core の要素は表 2 のように対応づけることができる。Dublin Core への変換を機械的に行うためには、いくつかの課題がある。

ひとつは、たとえば博物館コアの「人物」には、Dublin Core の creator, contributor, publisher に相当する情報が含まれる可能性があるが、博物館コアではこれらを区別していない。「時間」、「空間」の要素についても同様のことがいえる。もし、博物館コアに Dublin Core を意識した限定子を導入し、より詳細な記述をしたならば、これらの情報をスムーズに Dublin Core の各要素へ対応づけることができるが、その場合 simple なメタデータを目ざした博物館コアの基本方針に反することになる。

一方、「形状」、「材質」といった要素は、「もの」の属性として明確な情報である。Dublin Core のどの要素に対応づけるかはいくつかの選択肢が

表 2 博物館コアメタデータと Dublin Core との対応

博物館コアメタデータ	Dublin Core
1. 資料ID	identifier
2. 名称	title
3. 時間	date (製作、使用、他) coverage (記載内容)
4. 空間	description (製作、使用、他) coverage (記載内容)
5. 人物	creator (制作) contributor (寄与) publisher (公開) description (記載内容)
6. 分類	type ?, description ?
7. 主題・キーワード	subject
8. 大きさ・重量	format
9. 数量	format
10. 形状	type ?, description ?
11. 材質	format ?, description ?
12. 記述	description
13. 関係	relation
14. その他	description

考えられるが、ルールを定めてそれに従えば、機械的な対応づけは可能である。

最後に、博物館コアの「分類」には、原データベースにおいて、さまざまな基準でカテゴリ化された情報項目が対応づけられる。それを Dublin Core の type 要素に対応づけるのが適切かどうかは、原データベースにおける情報項目の定義や意味、内容に依存する。

6 むすび

本稿では、博物館の特に人文科学系の実在資料情報の統合的な検索・共有を目的として「博物館コアメタデータ」を新たに提案した。また、このような目的に、インターネット上の情報資源の記述や発見のために開発された Dublin Core Metadata Element Set を用いた場合の問題点について議論を行った。それと同時に、人文科学のより広範囲の研究資源の中で、博物館の資料情報を共有・活用するために、博物館コアメタデータを Dublin Core に対応づけるときの課題についても検討した。

今回は、われわれが所属する国立民族学博物館と国立歴史民族博物館の資料情報を中心に分析し、博物館コアメタデータを定義した。これは、人間文化研究機構がもつ資料情報の統合検索を考える過程で、ボトムアップ的に生じた問題意識から生まれたメタデータであるため、既存の博物館資料情報のための標準メタデータとの比較検討が不十分な面がある。今後はそれらとの関係、博物館コアメタデータの位置づけをより明確にしなければならない。また、他館が提供している博物館資料情報についても詳細に調査し、博物館コアメタデータの適合性を検証し、要素構成を改善していくことが必要である。

参考文献、URL

- [1] Grant, A., Nieuwenhuis, J., and Petersen, T.: *MICMO: proposed guideline for an international standards: minimum information cat-*

- egories for museum objects, International Documentation Committee, International Council of Museums, 1994.
- [2] ICOM/CIDOC: *International guidelines for museum object information: the CIDOC information categories*, International Committee for Documentation of the International Council of Museums, 1995.
(<http://cidoc.mediahost.org/guidelines1995.pdf>)
- [3] ICOM/CIDOC: The CIDOC Conceptual Reference Model.
(<http://cidoc.ics.forth.gr/index.html>)
- [4] Orna, E. and Pettitt, Ch. 編: 博物館情報学入門, 勉誠出版, p.181, 2003.
- [5] 村田 良二: ミュージアム資料情報構造化モデルの開発, 情報処理学会シンポジウム論文集, vol.2005, No.21, pp.63-70, 2005.
- [6] 東京国立博物館 博物館情報処理に関する調査研究プロジェクトチーム: ミュージアム資料情報構造化モデル, 東京国立博物館 2005.
(<http://webarchives.tnm.jp/docs/informatics/smmoi/>)
- [7] Dublin Core Metadata Initiative.
(<http://dublincore.org/>)
- [8] 安達, 山本, 関野, 原, 柴山, 安永: 人文科学研究データベースの統合検索システムの構築, 2008年画像電子学会第36回年次大会予稿集, R.2-1, 2008.
- [9] 人間文化研究機構: 研究資源共有化統合検索システム, 2008.
(<http://www.nihu.jp/kyoyuka/tougou/>)
- [10] 安達 文夫: 人文科学情報共有化のための博物館資料情報のマッピングの検討, 2006年画像電子学会第34回年次大会予稿集, pp.179-186, 2006.
- [11] 山本 泰則: 人文科学系データベース横断検索のための共通メタデータ - 統合検索システムにおける定義と課題 -, 2009年画像電子学会第37回年次大会予稿集, T1-3, 2009. (CD-ROM)

付録A 歴博, 民博のデータベースから博物館コアメタデータへのマッピング

国立歴史民俗博物館					
要素名	記述対象	館蔵資料	館蔵中世古文書	館蔵『懐溜諸周』	館蔵高松宮家伝来禁裏本
資料ID	資料	資料番号	資料番号, 卷子・冊子番号, 文書番号	資料番号	資料番号
名称	資料	資料名称	文書名	資料名称	資料名称, 外題, 内題, その他の題, 別書名・合綴書名
時間	資料	製作年代	時代		
	記載内容		日付(和暦, 西暦)	年代, 西暦	時期
空間	資料	使用地			
	記載内容				
人物	資料			作成, 作成, 宛名	著編者, 刊写年代・刊写者
	記載内容		差出, 充所	宛名	
分類	資料	実物・模造		分類c, 形態	
	記載内容				
主題・キーワード	記載内容			分類a, 分類b	
大きさ・重量	資料	法量, 尺度	法量 縦 横	法量 縦 横	法量
数量	資料	数量		点数	頁数, 紙数
形状	資料		形式		装訂
材質	資料	材質			料紙
記述	資料 / 記載内容				
関係	資料	コレクション名	区分, 所収文書群		
その他	資料	区分	欠簡・断簡		
	記載内容		端裏書, 本文書出, 本文書止	位置, 備考1, 備考2	巻数, 表紙等, 奥書等, 本文, 界線・匡郭, 付訓, 印記, 注記

国立歴史民俗博物館					
要素名	記述対象	館蔵錦絵	館蔵縄文時代遺物	館蔵装身具	館蔵野村正治郎衣装
資料ID	資料	資料番号	資料番号	資料番号	資料番号
名称	資料	資料名称	資料名称	資料名	資料名称
時間	資料	和暦, 西暦	時期, 型式	時代, 和暦, 西暦, 世紀	時代 (和暦, 西暦)
	記載内容				
空間	資料	画工・版元住所	遺跡名, 所在地		
	記載内容	国名, 地名			
人物	資料	画工名, 彫師, 摺師, 版元名		作者	
	記載内容	人名			
分類	資料		種類	種別	種類
	記載内容	主題分類			
主題・キーワード	記載内容	内容分類, その他固有件名		意匠	モチーフ
大きさ・重量	資料	法量	法量, 重量	法量 縦 横 高さ	法量 縦 横
数量	資料	員数		数量	
形状	資料	判型, 形態			
材質	資料		材質	材質	材質
記述	資料 / 記載内容				
関係	資料	コレクション名	コレクション名	コレクション名	
その他	資料	彩色	出土状況, 備考	技法, 備考	技法, 糸込み, 地色, 伝来, 備考
	記載内容	画工署名 改印		銘文	

国立民族学博物館							
要素名	記述対象	標本資料 詳細情報	衣服・ アクセサリ	中西コレクション	音楽・芸能の 映像	音響資料目録	日本昔話資料
資料ID	資料	標本番号	標本番号	資料番号	コンテンツ番号	資料番号	データ番号
名称	資料	標本名, 現地名, 訳名	標準衣服名, 現地名, 現地名英字	親書誌名, 資料名, 現地での名称	タイトル(現地名), タイトル(日本語)	タイトル	タイトル
時間	資料	受入年度, 収集日, 使用年代, 製作年代		出版年, 使用年代	撮影時期	発行年	収録日, 概要採録日
	記載内容						
空間	資料	収集地, 使用地, 製作地	地域・民族名	出版国, 出版地, 地域, 使用地	地域, 国, 地名	発行者	収録地
	記載内容						
人物	資料	原収集者, 使用者, 使用民族, 製作者		責任表示, 出版者, 使用民族,	担当教員, 作曲者	製作者	収録者, 話者, 概要採録者
	記載内容				演奏者, 民族		話者?
分類	資料	OWCコード, OCMコード	データ区分, 身装概念, OWC	種別			OWCコード
	記載内容			文字種	OWCコード, ジャンル, 声楽・舞踊・器楽	OWCコード	日本昔話タイプ, 話形分類
主題・ キーワード	記載内容						
大きさ・重量	資料	寸法・重量	丈	大きさ・形状			
数量	資料			枚数, ページ数			
形状	資料		形態	大きさ・形状			
材質	資料	製作法・材料	素材	材質			
記述	資料/ 記載内容			資料解説		トラック情報	概要
関係	資料	受入番号, 収集番号, 展示番号, 関係品情報, 文献, 図, 表, 添付資料・ 関連情報	收藏場所	所蔵_配架場所	素材テープ番号	製品番号, シリーズ名, 付属資料	
その他	資料	使用状況, 用途・ 使用法, 制作状況, 製作法・材料, 変 遷・分布,	布地特性, 構造技術	版表示, 巻冊次, 言語	再生時間, 著作権(注)	再生時間, 規格, 利用条件	(再生) 時間
	記載内容				演奏形態, 楽器(種別), 楽器名		