

# 博物館における収蔵品の解説情報を活用するための テーマ構造の解析と記述方法の提案

鈴木 美咲<sup>†</sup> 清水 恵<sup>†</sup> 杉山 岳弘<sup>†</sup>  
静岡大学<sup>†</sup>

## 1. はじめに

博物館の収蔵品は豊かな情報を持っている。多くの博物館では、これらの情報の一部を、収蔵品のデータベースとして公開している。情報は基本情報と解説情報（収蔵品や関連する物事を詳しく解説するテキスト形式で記述される情報）の大きく2つある。また、学芸員によるガイドや書籍、遺跡の調査報告書などには、博物館が所有する収蔵品のデータベースにまだ登録されていない収蔵品の情報が多く含まれている。

これらの情報をデータベース化し、さらに機械判読可能な状態で公開されれば、さらに有効に活用することができる。特に、解説情報は、ある収蔵品について他の収蔵品や当時の周辺の情報との関連性を示しており、この関連性を計算機で処理可能にできれば、広域的かつ多面的に知識を得る上で重要な情報を提供できるようになる。

本研究では、解説情報について文章の形で記述されている形式から、計算機で処理できるように記述する方法を提案する。

## 2. 解説情報の持つ情報と半構造データ

解説情報は文章で表されており、基本的には、事柄の関連性を示す構造を持っている。この関連性を、文章から抽出し、部分的に構造化する。これを文章と事柄の関連性を組み合わせて、図2のような形式で記述する（図1）。計算機で処理可能な構造化されたデータと、人間が可読な

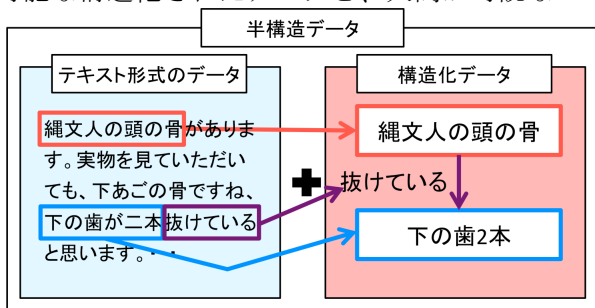


図1：半構造データの形式

A Propose of a Method for Description of Expository Information with Theme Structure Analysing about Museum Collections

<sup>†</sup>SUZUKI Misaki, SHIMIZU Megumi, SUGIYAMA Takahiro /Shizuoka University

テキストの形のデータを結びつけたものを、本研究では半構造データと定義する。

## 3. 解説情報の分析

解説情報の事柄の関連性を抽出するために、テーマ同士の関係性を抽出する会話分析を基にしたテーマ展開パターンの分析手法[1]を用いる。この分析手法を用いて、浜松市博物館の収蔵品を対象に解説情報の分析を行う。

### 3.1 テーマ展開パターン

テーマ展開パターンとは、会話などのテキストにおいて、時系列で話題がどのように展開していくか、話題の関係性を表現したパターンである[1]。テーマ展開パターンには、話題の関係性に応じて図2のような3つのパターンがある。この3つのパターンを応用して、解説情報について事柄の関連性を記述する。

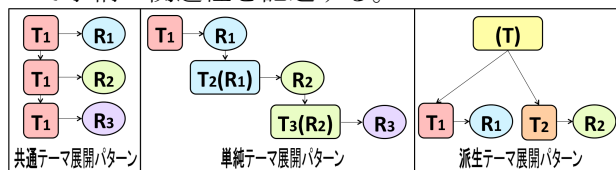


図2：3つのテーマ展開パターン

### 3.2 分析に使用するデータ

分析する解説情報として、浜松市博物館の収蔵品データベース「ある蔵」[2]の解説、学芸員によるガイド、収蔵品に関連した書籍や資料を収集して用いる。

### 3.3 分析の流れ

まず、収集した解説情報について、収蔵品に関する事柄を、語句レベルで抽出していく。例えば、収蔵品が出土した「遺跡」、「縄文時代」、「抜歯」や「屈葬」といった収蔵品に関連する事柄である。次に、抽出した事柄について、主に述語で示される語句同士の関連性を紐付けしていく。語句の関係性において、テーマ展開パターンに従い記述する（図3）。

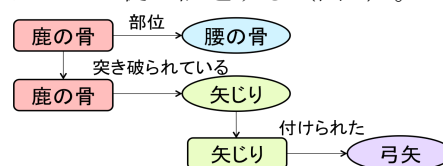


図3：テーマ構造の記述例

#### 4. データの構造の記述方法

ここでは、テーマ展開パターンで記述された解説情報を、RDF 形式で記述する方法を説明する。なお、「ある蔵」の基本情報については、RDF 形式に記述する際、MetaBridge 上の東京国立博物館が作成した「ミュージアム資料情報構造化モデルの RDF」[3]を基本的に用い、述語が定義されていない場合は、ダブリンコア[4]や共通語彙基盤[5]で定義されている述語を使用する。

#### 解説情報の RDF による記述

まず、解説情報で用いられている語句において、「蜆塚遺跡」や「縄文時代」など、LOD としすでに URI が定義されているものがあれば利用する。また、収蔵品の特定の部分を示す場合、あらかじめ収蔵品の構成要素として、RDF で定義しておく。

解説情報は、図 4 の枠囲み①に示されるように、見出しと、オリジナルの解説情報である人間が可読なテキスト形式と、そのテーマ展開パターンを記述した構造化データとによって構成された RDF 形式で記述する。テーマ展開パターンと RDF 形式は相性が良く、一対一で対応させて記述することが可能である。

次に、もともとのデータベースの解説情報には無い、ガイドや文献などの別の情報源からの解説情報について述べる。まず、基本的な記述方法は同様で、図 5 のようにそれぞれ独立したデータとして RDF 形式で記述する。さらに、収蔵品との関係を、図 4 の枠囲み②のように記述して、関連付ける。

#### 5. 検証

本データ構造の検証については、ガイドのシナリオ作成支援システム[6]の実装を行うことで、検証する。

#### 6. まとめ

博物館の収蔵品データベースにおける、解説情報に対して、計算機で処理可能にできるようにするため、テーマ展開パターンを用いてテーマの関係性の分析を行い、半構造化データとして RDF 形式で記述する方法を提案した。

#### 謝辞

収蔵品のデータ収集に協力していただいた浜松市博物館の皆様へ感謝の意を表します。

本研究は科研費基盤研究(C)15K01147 の助成を受けたものである。

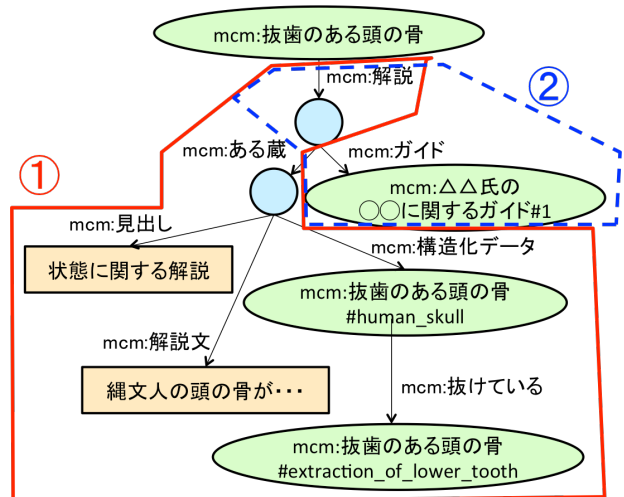


図 4: 解説情報の構造

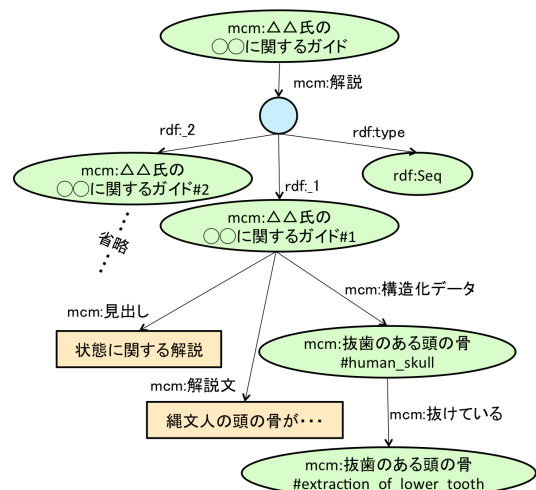


図 5: 他情報源から収集した解説情報の構造

#### 参考文献

- [1] 杉山, “博物館や動物園における Q&A 関係図を基にしたガイドプログラムのデザイン”, Design シンポジウム 2012 論文集, pp. 519-522, 2012.
- [2] 浜松市公式 Web サイト, 「浜松市博物館 ある蔵館蔵資料検索」, 2015, <<http://www.city.hamamatsu.shizuoka.jp/hamahaku/setumei.html>>2016 年 1 月 7 日アクセス.
- [3] 東京国立博物館, 「ミュージアム資料情報構造化モデルの RDF」, META BRIDGE ホームページ, 2013, <<https://www.metabridge.jp/>>2015 年 12 月 26 日アクセス.
- [4] 杉本茂雄: Dublin Core について第 1 回概要, 情報管理, vol. 45, No. 4, pp. 241-254 (2002).
- [5] 独立行政法人 情報処理推進機構, 「共通語彙基盤 | 共通語彙基盤整備事業」, <<http://goikiban.go.jp/>>2016 年 1 月 6 日アクセス.
- [6] 清水他, “博物館展示品の持つ知識のテーマ構造を基にしたシナリオ作成支援システムの開発”, 第 78 回 情処全国大会, 5ZB-02, 2016.