

博物館情報の相互運用と CIDOC CRM の役割

村田良二

東京藝術大学美術学部先端芸術表現科

概要：博物館の情報化はデータベース化からネットワーク化へと進みつつある。複数の博物館が協調してそれぞれの持つ情報を有機的に関連づけていくことができれば、その有用性は計りしれない。この目標の実現のためには複数の技術を組みあわせて運用する必要がある。現在、特に中心的な課題である資料ドキュメンテーション形式の国際標準として CIDOC CRM の開発が進められている。本稿では CIDOC CRM の概要と機能を紹介し、相互運用に向けて関連技術のなかでどのように位置付けられるのかを考察する。

Interoperation of Museum Object Information and the role of the CIDOC CRM

MURATA Ryoji

Department of Inter Media Art, Tokyo National University of Fine Arts and Music

Abstract: Development of museum object information is now focusing on mediation and interchange of diverse cultural data. For this purpose, it is necessary to integrate multiple technologies, and the most important point is using common data model. ICOM/CIDOC has been developing the CRM as a common language for discription of museum objects. This paper gives outline of CRM with examples, and then we discuss integration of CRM and other relevant technologies

収蔵品目録をデータベース化して運用するという従来からの情報化が徐々に達成されつつある現在、博物館における情報化に関する議論は次の段階であるネットワーク化・相互運用へ中心が移ってきている。国際博物館会議 (ICOM) のドキュメンテーション委員会 (CIDOC) は、博物館資料のドキュメンテーション標準に向けて、これまで『博物館資料情報のための国際ガイドライン』や『CIDOC リレーショナル・データモデル』を開発してきたが、現在の延長上に CIDOC Conceptual Reference Model (CRM) の開発が進められている。これは多様な博物館情報を記述するための共通の枠組みである。今後の博物館情報の相互運用の基盤になることが期待されており、新しい ISO 標準にするための準備も進められている。

博物館資料情報の相互運用の必要性

博物館に蓄積された情報は各館個別のシステムで運用されているが、例えば図書館における NACSIS-CAT のように協調して運用できれば大変利益が大きい。一般利用者にとって有用な

情報は、例えばある人物の作品にどのようなものがあるのか、あるいはある時代・地域の美術家、さらに特定の技法による作品にどのようなものがあるか、などといったことであろう。つまり収蔵情報のような「博物館・美術館に関する情報」よりも「特定の人物・作品に関する情報」のほうが利用価値が高いと考えられるのである。研究者や教育関係者でも同様であろう。また博物館の情報担当者にとっては、互いに持っている情報をスムーズに交換できればシステムの利便性が高まるであろうし、企画展覧会の際に貸借される物品の情報を電子的に交換できれば記録管理や展覧会目録の作成などにも役立つだろう。

現状と問題・障害

こうした相互運用のためには何らかの共通の枠組みが必要である。最も重要な点は資料の情報をどのように記述するかということである。博物館資料ドキュメンテーションの標準化については ICOM / CIDOC が 1995 年に発表した『博物館資料情報のための国際ガイドライン』をはじめ、様々な提案がなされている。しかし博物館という場所に集められるものはきわめて多種多様である。また一口に博物館といっても、その多くは歴史博物館や科学博物館、美術館のように各々専門的な活動をしている。従って情報システムに対しても当然各館ごとに違ったニーズがあり、異なる概念や用語が用いられている。また同じ分野でも情報の詳細度などはまちまちになっている。

CIDOC CRM

CIDOC CRM はこのような問題を解決して博物館情報の相互運用を実現するための手段として、開発がすすめられている。CRM は知識工学でいうところのドメイン・オントロジーである。博物館の領域で用いられる事物や概念間の関係を分析して定義付けを行い、それらをオブジェクト指向の方法で形式化する。従って概念や事物はエンティティ(クラス)として定義され、エンティティの間に関係付けるプロパティが与えられる。

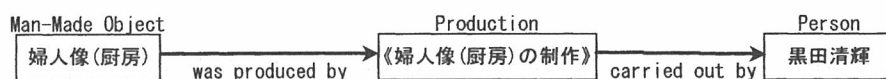


図1. 簡単なモデル

図1. に簡単なモデルの例を示す。これは「婦人像(厨房)」という作品を黒田清輝が制作したことを示している。図中の四角は Man-Made Object や Person といったエンティティと「婦人像(厨房)」のようなそのインスタンスを表しており、矢印はそれぞれのエンティティを関係付けるプロパティである。CRM の Version3.2 では約 70 のエンティティと 100 余りのプロパティが定義されている。エンティティもプロパティもともに継承関係をもつ階層として組織化され、それぞれの概念を適切に表すように注意深く設計されている。

これらのエンティティとプロパティを使って情報をモデル化していくことで、それぞれの情報が互いにどのように関連しあっているかを表現するのである。このような意味ネットワ

ークによる表現は、データ交換のために単純に共通のフィールドを定義する方法に比べて、異なる構造・詳細度のデータベース間で情報を交換する目的に適している。また CRM は必要に応じて拡張可能であるため、共通のフィールドを定める方法と違って情報のソースから詳細な部分を失うことなく表現できるという利点もある。

さらに CRM では多重インスタンス化を許容している。博物館資料のようなものは本来にその資料の種別を決定するのが難しい。多重インスタンス化によってそのようなシステムデザイン上の困難は軽減される。あるいは CRM でモデル化された情報に矛盾があっても、それをどう処理するかは CRM では定めていない。このような方針は、CRM が「正しいデータを提供する」ことよりも「コミュニケーションを可能にする」ことに重点を置いていることを示すものといえよう。

相互運用のための基盤と CRM

異なるデータ構造・フィールド項目をもつデータベース間でデータを交換するために必要な変換（マッピング）を最小限にするには、共通のフォーマットを用いればよい。しかしフラットな属性-値ペアでこれを実現しようとする最大公約数として共通の属性だけを用いなければならない。だが博物館資料は多様である。A 館で記録している属性はほとんどの館では使われていないが、Z 館にデータを移す際には有用である、といった事態が多分に予想される。CRM のようにデータ間の意味を保持する構造であれば、フィールド名の違いを吸収するだけでなく、詳細度のレベルが異なるシステム間での交換も容易になる。たとえば資料の保存状態は《A. 資料》に対して行なわれる《B. 検査》の結果《C. 状態》が記録され、CRM では A、B、C がそのような関係にあることが表現される。《B. 検査》という出来事の情報（日付や担当者）をデータとしてもたないシステムでは、B. を飛ばしての《C. 状態》を取り込めばよい。《B. 検査》が複数回あればその中から最新のものを選べば現在の状態が得られる。逆に《B. 検査》が表現されていないデータからは、デフォルト値として「少なくともデータを受け取った時点以前に検査が行なわれた」とすることができるだろう。もちろんこの場合はデフォルト値として推測されたものであることをシステム内で明示する必要があるだろう。このように博物館のような多様な情報システム抱える分野においては CRM のような記述力の高いモデルが必要なのである。

ところで現実に CRM を利用するためにはモデル化した情報を何らかの形式にする必要がある。以前は XML 表現のための DTD が提案されていたこともあったが、現在では Model and Syntax 仕様が W3C 勧告になった RDF (Resource Description Framework) が有力視されている。RDF は Web 上のリソース相互の関係を記述するための方法であり、メタデータの記述などに用いられる。この RDF ではデータモデルの基礎が CRM と同様の意味ネットワークであるため、XML 構文、スキーマ定義に RDF を用いることができるだけでなく、RDF 用のツールやライブラリも使うことができる。

RDF を利用するにせよ他のデータ形式を利用するにせよ、現在の Web 関連技術の進展とそ

の可能性の広がり、また利用の容易さやコスト面を考えると、博物館間でのデータ交換は基本的には専用プロトコルなどより HTTP を使うのが現実的だと思われる。インターネットを利用するためにはしっかりとしたセキュリティ基盤が必要だが、利用範囲が拡大しつつある公開鍵基盤 (PKI) を用いることでこれを確保できる。TLS / SSL でもよいし、RDF を利用する前提であれば XML Signature や XML Encryption などの XML 用の PKI アプリケーションを利用することもできる。公開鍵証明書は利用各館の間で承認された機関が発行することになるだろう。

一般利用者が高い利益のあるデータベースの横断検索に関しては、すでに図書館の領域で Z39.50 が実用されている。Z39.50 の検索属性集合やレコードフォーマットに CRM を基礎にしたものを使うことが考えられるだろう。モデルが違うため CRM をそのまま利用することはできないが、例えば資料検索であれば Man-Made Object エンティティのプロパティを属性として使うといった方法が考えられる。CIMI プロジェクトの文化財のための Z39.50 プロファイルも参考になるだろう。

また CRM を RDF で表現するならば、いわゆる Semantic Web に博物館の資料情報を組みこんでゆく可能性が開かれてくる。Web を巡回するエージェントが RDF 表現された CRM モデルを蓄積・結合してゆけば、博物館領域だけでなく他分野のデータと組みあわせて、より一般的な領域でのデータの利用が可能になる。このような場面を考えると、語彙の解釈を支援する仕組みが必要になる。CRM そのものは用語に関しては対象範囲外としており、直接用語統制やシソーラスを扱わないが、ある言葉がシソーラスや辞書で定義されているものであることを示すことはできる。従って CRM モデルの解釈の際にエージェントがシソーラス・サーバを参照する、といった構成によってより高い精度を得ることができるだろう。もちろんそのためにはそれぞれの分野で承認されたシソーラスなり辞書なりが必要である。これには Getty Research Institute の「The Art & Architecture Thesaurus」などがあるが、日本語で利用できるものが望まれる。

以上のように、博物館情報の相互運用に向けて CIDOC CRM が果たすであろう役割はまさに中心的と言ってよい。しかし相互運用は CRM だけで可能になるのではなく、関連技術との連携の実証など運用に向けた課題は少なくない。とはいえ、それだけの価値ある目標であることは確かであろう。

参考文献：

1. Nick Crofts, Ifigenia Dionissiadou, Martin Doerr, Matthew Stiff (editors), *Definition of the CIDOC object-oriented Conceptual Reference Model, Version 3.2*, ICOM/CIDOC CRM Special Interest Group, July 2001 (http://cidoc.ics.forth.gr/docs/cidoc_crm_version_3.2.rtf)
2. Martin Doerr and Nicholas Crofts, *Electronic Communication on Diverse Data - The Role of the oo CIDOC Reference Model*, 1998 (<http://cidoc.ics.forth.gr/docs/crmrole.htm>)