

オブジェクト・ドキュメンテーションへのメタデータの適用

鯨井秀伸

主任学芸員 愛知県美術館

Web を文化資産の情報に対する資源として利用するため、情報システムに蓄積されている量的にも質的にも詳細なドキュメンテーションを統合しディストリビュートするためには、手堅く安定した手段の確立が必要となる。こうした統合への主な障害は、既存システムの意味論上あるいは構造上に互換性が欠けている点である。異なった機関あるいは同一機関内の異なったシステムは、異なった方法で使用するデータを組織し提供し、用語法はしばしば互換性に欠け、詳細のレベルが異なる。互換性を欠く情報システム間のギャップを埋める試みは、アド・ホックな変換規則に基づくか、限定された「コア」データに関心を払う相当量の単純化を手段としている。この試論は、CIDOC CRM によって提案されているオブジェクト指向のデータ・モデルとして提出された、概念参照モデルの適用可能性について検討する。

Application of Metadata to Object Documentation

Hidenobu KUJIRAI

Curator, Aichi Prefectural Museum of Art

In order to achieve the use of web as a global resource for cultural heritage information, we have to establish reliable means for integrating and distributing the detailed documentation contained in the information systems. The semantic and structural incompatibility of existing systems is a barrier to such integration. This report discusses potential issues of conceptual reference model (CIDOC CRM) based on metadata-oriented logical framework, as a basis for providing a semantic reference which enable to render information resources mutually accessible.

■ はじめに

美術・博物館のドキュメンテーションへのアクセスは、広い範囲の利用者の興味を引く可能性を秘めている。オブジェクトの持つ情報は、オブジェクトそのものとそれに人間が応答した様々なアイデアから成り立っている。したがって、それぞれの場合提供される情報は他の資源(sources)を統合する必要がある。特にこれは文化遺産に対して言えることである。個々のオブジェクトについての記述は、限定された興味しか引かない。他のオブジェクトへの付加的な参照、またオブジェクトの歴史的な・地理的な・文化的源への参照は、あるコンテキスト

の中に置くため、あるいは意味を与える支援となる。コレクションを生かすコンテキストに基づく情報は、典型的なことだが複数の機関からまた同一機関内においても複数の部門あるいはファイルからいわば配給されている。従って、異なった情報システムの間になんらかの形式のインターアクションがなければコレクションの潜在的な興味を失われることになる。

情報のプレゼンテーションは、現行の様々な試行が一般的に不適切である分野の一つである。多くの機関は、コレクションの中から、残りの情報の量や性質に関する指摘もなく、部分的な選択的情報を提供するか、「目録」的アプローチから網羅的で簡略な情報を提供しているかのどちらかの方法を取っている。オブジェクトに対し人・場所・出来事などのコンテキスト情報と共に統合的情報を提供しているところは少ない。また、コンテキスト情報を提供しても、それぞれが分離された情報の提供となっている。

異なった要求に見合うためには、異なった形式の提示が予想される。統計的分析やより深い研究内容は、組織的で的確なクエリーのファシリティーを必要とし、それは網羅的なアイテムのリストを発生させる。しかしこの種のアプローチは一般の興味のブラウジングや教育目的には相応しくなく、それらはテキストの説明の多いより「技術的」でない提示や、利用可能な資料へのパスウェイを見出すヘルプ・ガイダンスの形式を好む。主題やコレクションの内容に不案内なユーザーの場合に、初心者には検索基準を入力する可能性を提供することは有効ではない。これらの異なった要求は、異なったインターフェース・デザインを含意し、主題に関する異なったレベルの知識を前提としている。しかしながら、それは両者共に異なった資源からの情報を統合する可能性のあるメカニズムしだいなのである。

EC(European Community)は、美術・博物館、アーカイヴ、ライブラリー情報の統合を現行の戦略的研究・開発(R.D)の目標に定めている。異なったタイプのコレクションである美術館・博物館情報システムやアーカイヴズ、ライブラリーと同様に自然史、美術、民族誌学などは、補完的情報や視点を提供する。それらの組合せ・編集というより結合は、我々の文化遺産への新しい洞察を提供する潜在的可能性を持っている。

主題をたやすく同定することができるよう、意味ある方法でデータを結合し統合することは、直接的な編集に対する必要性よりは、より進んだメカニズムを要求する。異なったドメインの多様な情報要求の実例を参考にする利点はこのにある。個々の制作者を同定することは美術に対し民族誌学ではあまり興味を持たれない。一方自然史では「作者(author)」の概念は分類システムでのみ適用され、収集されたオブジェクトに対しては適用されない。考古学者や古生物学者は習慣的に断片化したオブジェクトを扱い、その後運がよければ、単一の総体に組み合わされる。このプロセスは他の分野には通常なものではない。多様な断片は、同定されプロセスの間追跡される必要がある。歴史学においては多くの情報は仮説的であり、したがって著者(author)によるものとして「署名(signed)」される必要がある。つまり情報と詳細のレベルはある部門に不可欠だが、ある部門には不必要で相互に理解可能というものではないのである。

以上より、異なった源泉からの情報を結合するには、異なった視点が尊重され表現されるための柔軟性を持つ高度の抽象化と各部門にニュートラルな視点を必要とすることが理解される。

#### ■ メディエーション(仲介)処理

**RAMA:** 統一ユーザー・インターフェース

**CIMI:** SGML における構造化テキスト・マークアップ(Z39.50 の利用)  
Z39.50 の短所

**AQUARELLE:** SGML、HTML、Z39.50、HTTP(マルチプル DTDs、多国語ソース検索支援)

**GRASP:** 固定フォーマットへの変換  
ターミノロジーの問題

上記諸例は、3層構造のシステム(中央アプリケーション・サーバーはデータベースとクライアント間のインターフェースとして機能する)を使用し、クエリーとデータのトランスレーションは、個々のデータベースあるいは中央サービスあるいは両者のローカルで処理される。しかしながら、1) トランスレーションは共通点を持たず、特異的で、機能の物理的組み込みによるものである。2) 全ての情報はオブジェクト中心の処理である。人・場所・出来事などの情報は、間接的にしか得られない。

こうした制限に対し、Wiederhold は「メディエーション・サービス(mediation service)」の概念を紹介した。<sup>1)</sup> この概念は他のドメインにおける異なったシステムで手段として使われている。「メディエーションは、アプリケーションの利用に先だち蓄積されたデータを向上させる広範で多様な機能をカバーする。メディエーションは、再現と抽象化の問題に関連して、インターフェースをインテリジェントなものにする。メディエーションは実質的な役割を持ち、変換をさせる知識構造を含んでいる。メディエーションはドメイン専門家によって維持されるものであり、その特徴は、1) ビュー定義を使うデータベース間の変換、2) 多様なデータベース間からのデータのアクセスとマージの方法、3) 基礎をなすデータの抽象化と一般化、4) 不完全なあるいは異なったレベルの詳細あるいは抽象である情報のハンドリング、5) 構造化されたテキストからの情報を統合する方法。6) 誘導データのメンテナンス、である。

メディエーターはソフトウェア・モジュールであって、より高度の層あるいはアプリケーションに対し情報を作り出すため、特定セットのあるいはサブセットのデータについてのコード化された知識を活用する。

<sup>1)</sup> CIDOC Conceptual Reference Model Information Groups, Produced by the ICOM/CIDOC Documentation Standards Group, Editors: Nick Crofts, Ifigenia Dionissiadou, Martin Doerr, Pat Reed. September 1998 Version 2 March 2001

<sup>2)</sup> Gio Wiederhold, "Mediators in the Architecture of Future Information Systems", in IEEE Computer, March 1992.

## ■ 概念参照モデル

美術・博物館の場合について考えると、購入と売却とはそれぞれ別のテーブルに設定されてきた。対象の物質は、「取得 acquisition」や「購入」によって所有者 A から所有者 B へ移り、「処分 deacquisition」や「売却」によって所有者 B から所有者 C へ移る。この情報を「Acquisition」として「正規化」すると、二つの「Actors」を意味させることになり、一方は legal title を譲渡し、一方はそれを得ることを意味する。したがって、「Acquisition」は、「transfer of legal title on an object」と表現されることになる。この視点は、機関に対しニュートラルな概念モデルである。

人・場所・出来事などの概念を想起すると、保存処置に関する場合、機関によっては資料の「状態」を「状態記録」として記録し、ある機関では「状態」を「状態アセスメント」として「人」、「方法」、「文書」と同時に「状態記録」などと関連させているかもしれない。この場合、「状態」へのオブジェクトの直接リンクは得られないか、別の処理となってしまう。この参照の方向性の欠如は、多様な意味論的にオーバーラップしている記述群間のシステムにおけるもう一つの主要な雑多性である。

これらの両者のケースをもシステムの対象として考慮しうるのが概念モデルの特質である。オブジェクトは、「状態記録」を持ち、または「状態アセスメント」によってアセスメントされかつ「状態アセスメント」は「状態記録」によって同定されるという関係にあることになる。このメカニズムは、モデルの拡張と効果的な改善を可能にする。概念参照モデルは、情報カテゴリーにしたがってデータ・システムをデザインし、物理的オブジェクトは単一のテーブルで扱われる。例えば絵画と版画は共に物理的オブジェクトとして扱われる。この関係は「isA」と呼ばれ、そのオブジェクト指向のデータベースへのマッピングは、「generalisation/specialisation」あるいは「superclass/subclass」と呼ばれる。この点は他の理論と大きく変わるところはない。

異なった分野のコレクション間の参照関係について、ルネサンス絵画は考古遺物あるいはデッサン、版画を参照することがある。また一般的に美術作品は同時代の史料との関連性を持つことも十分考えられる。この場合言語の分類概念が大きな影響力を持つことになる。階層的シソーラスである。<sup>iii</sup> ここでシステム化されたデータベースについて考えると、分野毎のシステム全てに対しクエリーを実行する必要がある場合が考えられる。また、それぞれの分野の作品・資料はオーバーラップしないが、関連する人、場所、時などがオーバーラップする可能性がある。特定分野のデータベースから包括的データベースへデータをロードする必要がある場合がある。反対に包括的データベースから適切なデータを特定のデータベースへロードする場合がある。

<sup>iii</sup> AAT 参照。日本語の場合はシソーラス概念の適応は困難かもしれないが、シソーラスの利用に関して情報処理システム内においては、共通語として英語を利用することは可能と思われるし、そうする必要がある。

データベース改善の視点からは、subclass はテーブルと見なすことができ、その superclass の全てのフィールドに加え付加的フィールドを持つ。superclass へのクエリーは、データベースが superclass のインスタンスとして subclass の全インスタンスをみなすことにより処理される。したがって「isA」構成は、複数のデータベースをマージできる機能を持つ。また、概念参照モデルの全てのエンティティーは、「type」を持つ。ここにおいてシソーラス概念が適用され、物理的オブジェクトにおけるような、物理的オブジェクト→美術作品→絵画あるいは版画...という関係性が維持されるのである。このシソーラスは AAT を使用する。

#### ■ 多重性

シソーラスから知られるように、概念は多重の一般化を持ちうるし、実体物は異なった側面から見られうる。概念参照モデルは、「人」を「Actor」とも「Biological Object」とも見る。「銘記(inscription)」を「Mark」および「Linguistic Object」と見るのである。したがって、このモデルはフォーマットというより説明的役割を演じるものである。例えば、昆虫標本は、人工物であり、また生物学的オブジェクトでもある。また、資料の破壊は、常に人の意志に基づくわけではなく、無意志の場合もある。これは出来事概念に属することになる。

#### ■ プロセスとしての限定

文化遺産全般にわたる概念構築のため、終りのない仕事になる可能性が高い。したがって、当面限定的な方向性を示しておく必要がある。ユーザーの概念的フレームワーク。共通の美術・博物館活動。美術館・博物館の収集物。適切な質のレベルを想定した詳細と精度のレベル。技術的複雑性の考慮。

#### ■ メタデータ

美術・博物館ではライブラリーにおけるより狭い意味でメタデータを適用しているように思える。「形式や内容についてのデータを含むオブジェクトやドキュメント、サービスの記述」<sup>iv</sup> 近年この観念は拡張されているように思われ、特に RDF(Resource Description Framework)<sup>v</sup> において、またヘルシンキ・メタデータ・ワークショップにおいて、正確な RDF と Dublin Core の定義に対する共通形式データ・モデルの必要が認識された<sup>vi</sup>。考古学からの批判に見られるように、ライブラリーのメタデータは、美術・博物館のオブジェクトについての考察なしには変換可能ではない<sup>vii</sup>。

<sup>iv</sup> Juha Hakala, Ole Husby, Traugott Koch, "Warwick framework and Dublin core set provide a comprehensive infrastructure for network resource description", Report from the Metadata Workshop II, Warwick, UK, April 1-3, 1996. <http://www.ub2.lu.se/tk/dcwsrept.html>

<sup>v</sup> "Resource Description Framework (RDF)" July 21 1998, <http://www.w3.org/RDF/>

<sup>vi</sup> Stuart Weibel, Juha Hakala, "DC-5: The Helsinki Metadata Workshop, A Report on the Workshop and Subsequent Developments", ISSN 1082-9873, D-Lib Magazine, Feb. 1998.

<sup>vii</sup> Paul Miller, Alicia Wise, "Resource Discovery Workshop: Final report from the Archaeology Data Service", prepared in accordance with guidelines from the Arts & Humanities Data Service (AHDS) and United Kingdom Office for Library & Information Networking (UKOLN) Resource Discovery Workshop series, 4 August 1997.