

6 L - 4

WWW 上における Metadata の記述と オンラインショッピングへの応用

坂田 毅 多田 浩之 大竹 智久

(株)デジタル・ビジョン・ラボラトリーズ

sakata, tada, otaket@dvl.co.jp

1. はじめに

我々は、WWW 上のオンラインショッピングの商品を対象とした商品検索サービス Commerce Mediator を開発している。Commerce Mediator は、商品の情報をキーワード単独ではなく、キーワードを商品の属性と結びつけて管理し、特徴によって利用者が商品を検索できることを特徴とする。この論文では、Commerce Mediator と、そのために開発した商品特徴 (Metadata) の表現形式 Multi-Schema Metadata Format について述べる。

2. Commerce Mediator

消費者が欲しい商品を探すときには、「色は赤」「サイズは XL」「価格は 5000 円以下」などの商品の特徴に基づく条件を念頭におく。このことは WWW 上においても同じであり商品の特徴を指定した検索サービスが求められるが、現在の WWW 上の検索サービスは「赤」「XL」「5000」といったキーワードのみで検索しており、そのキーワードが商品とどのような関係(色・サイズ・価格)にあるかといったことは検索対象となっていない。Commerce Mediator は、日本国内の Online Shopping Site の商品を対象に、「サイズは XL」「価格は 5000 円以下」といった色や価格など商品の特徴 (Metadata) による検索を利用者に提供するサービスである。Commerce Mediator の全体図を図1に示す。Commerce Mediator は、検索サービスという側面からとらえると、Metadata を管理し仲介する Metadata Mediator と Metadata Mediator と WWW Browser の間の gateway となる Commerce Mediator Proxy によってなりたっている。Commerce Mediator Proxy は検索用の HTML ページや Metadata 登録用のページとそれを処理する CGI プログラム群からなる。Metadata Mediator への Metadata の流

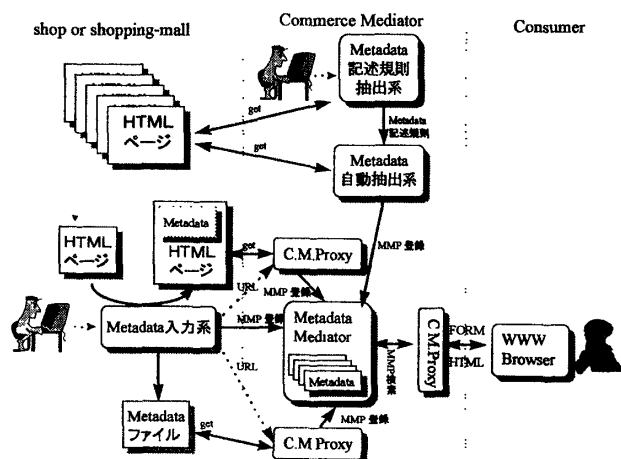


図1 Commerce Mediator での Metadata の流れ

入は三つの経路がある。一つ目は前述の Commerce Mediator Proxy が持つ Metadata 登録用のページに入力する方法。二つ目は、Metadata 入力系を使う方法である。Metadata 入力系は、HTML 文書に Metadata を付加したり、Metadata を記述したファイルを作成する作業を容易にする。我々はこれを無料で配布する。三つ目は、Commerce Mediator 側でロボットを動かして収集する方法である。Metadata 記述規則抽出系は、商品に掲載した HTML ページの中で商品名や価格など商品の特徴情報が記述される位置のパターンを店ごとのルールとして抽出し、それを用いて Metadata 自動収集系が Metadata を集める。Commerce Mediator は 97 年 4 月にサービス開始予定である[1]。

3. Multi-Schema Metadata Format

Internet 上の Resource の特徴を表現する方式は IETF や W3C などを中心に、IAFA Template[2], SOIF[3], Dublin Core[4]などが提案されてきた。これらは Metadata を属性とその値の組で表現しているが、異なる属性の集合(スキーマ)に基づく Metadata 同士の互換性については述べられていない。商品に付与する Metadata のスキーマを考えると、全ての商品をカバーする一つの標準的なメタデータスキーマを制定

The Representation of Metadata for Online Shopping on the WWW

Tsuyoshi Sakata, Hiroyuki Tada, Tomohisa Otake
Digital Vision Laboratories
7-3-37 Akasaka, Minato-Ku, Tokyo 107, Japan

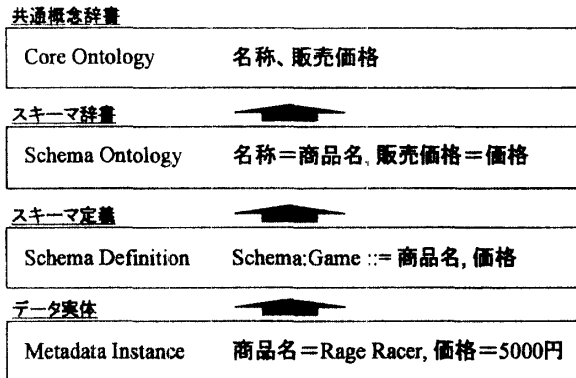


図 2 MMF の構造

するのは難しい。さらにいえば、多くの店が他店との優位性をアピールするために独自にスキーマの改良や独自設計を行うと考えられ、メタデータスキーマの多様化が広がると予測される。これに対応するにはスキーマの異なるメタデータに対しても検索式を有効にして検索対象とする仕組みが必要である。このスキーマ間のメタデータ相互変換性を実現するために、我々は Metadata の論理モデルと表現形式 Multi-Schema Metadata Format(MMF)を設計し前記 Commerce Mediator 内のメタデータの表現に利用している。MMF では Metadata 実体の上に3つの構造を持つ。MMF の構造を図2に示す。

Metadata Instance 部では Metadata の実体が記述される。Metadata は属性とその値のペアのリストで表現され、その表記法は The Warwick Framework[5]の HTML 内表記法を拡張したものであり、HTML ファイルのヘッダ部に META タグを用いて埋め込まれる。値が構造を持つことを考慮し値中の構造要素を Aspect として指定できるようにし、また値が特定の系(通貨単位、靴のサイズなど)に従う場合その情報を System として指定できる。また一つのファイル中に複数の商品の Metadata を区別して書くことができる。例を図3に示す。Schema Definition 部では、Metadata で使われる属性とその Aspect について定義する。属性や Aspect の値が特定の系に限定される場合はここで System 限定宣言を行える。Schema Ontology 部では、スキーマの異なる Metadata や検索式の相互変換性を実現するために、スキーマの属性と他のスキーマの属性の概念的包含関係を階層的に表現している。ここで定義される関係は以下の二つである。

```
<meta name="----" content="begin">
<link rel=SCHEMA.CN href=http://www.dvl.co.jp/scm/cn.scm>
<meta name="CN.商品分類" content="()Game Playstation">
<meta name="CN.商品名称" content="()Formula 1">
<meta name="CN.商品購買" content=
  "(system=RFC1738) http://www.dvl.co.jp/purchase/aaa.htm">
<meta name="CN.商品サンプル" content=
  "(System=RFC1738) http://www.dvl.co.jp/sample/aaa.htm">
<meta name="CN.製造会社" content="(Aspect=名称)(株)DVL">
<meta name="CN.製造会社" content="(Aspect=Tel) XX-XXXX">
<meta name="CN.商品価格" content="(System=JPY)55000">
<meta name="----" content="end">
```

図 3 HTML 中での Metadata Instance の表記例

(1) 同等:二つの属性が同じ関係を意味している場合。商品名称と商品名など。(2) 包含:属性 A が示す関係が属性 B の示す関係の特殊なものである場合に属性 A \subseteq 属性 B とする。出演者と主演女優など。これらの関係を利用して Metadata や検索式を他のスキーマに変換する。変換処理としては、言い換えと拡大解釈の二つがある。言い換えとは同等の属性に変換することであり、拡大解釈とは検索式で指定された属性の内容が含まれているかもしれない属性に変換することである。拡大解釈は、検索サービスにおいて、ユーザの指定された検索式に合致する商品が少ない場合に近似の商品を紹介するなどの用途に使う。Core Ontology 部ではスキーマを新規に作るときに ontology の対象とする標準的な属性の ontology であり、スキーマの数が増大したときもスキーマ同士の互換性を保つのに使う。

参考文献

- [1] <http://www.dvl.co.jp/mediator/index.html>
- [2] Deutsch P., Entage A., "Publishing Information on the Internet with Anonymous FTP", <http://info.webcrawler.com/mak/projects/iafa/iafa.txt>
- [3] Hardy D., Schwartz M., Wessels D., "Harvest User's Manual", <http://harvest.transarc.com/afs/transarc.com/public/trg/Harvest/user-manual/>
- [4] Weibel S., Godby J., et al., "OCLC/NCSA Metadata Workshop Report", http://www.oclc.org:5046/conferences/metadata/dublin_core_report.html
- [5] Lagoze C., Lynch C., et al., "The Warwick Framework A Container Architecture for Aggregating Sets of Metadata", <http://cs-tr.cs.cornell.edu:80/Dienst/Repository/2.0/Body/ncstr1.cornell%2fTR96-1593/html>