

セマンティックWeb適用システム

セマンティックWebポータルと
ユビキタス環境への適応

萩野達也（慶應義塾大学環境情報学部）

清野正樹（松下電器産業株式会社）

細見格（日本電気株式会社）

Webのポータルサイトとは

- インターネット上、またはイントラネット上の様々なWebサイトやWebコンテンツに対して玄関口(ポータル)となるWebサイト。
- ユーザが求める、または興味のある情報源にたどり着くための案内役。

ポータルサイトの進化

初期のポータル

初期の Yahoo! に代表されるインターネット・ポータルサイトカテゴリ毎に分類された、いわゆる「ディレクトリ型ポータル」

- オープン(誰でも自由にアクセス)
- 万人向け(誰もが同じ構成のページを見る)

最近のポータル

Yahoo! のほか、楽天のようなショップ・ポータル、ぐるなび、関心空間など多彩に企業内の情報ポータル(EIP)も普及

- 利用者毎にパーソナライズされたポータルページを提供可能
- コンテンツだけでなく各種検索などの機能もパーソナライズ

Webコンテンツの個人適応と状況適応

利用者への適応

- 利用者自身によるパーソナライズ(表示項目の選択)
- コンテンツの自動分類とユーザプロフィールや利用履歴によるパーソナライズ

内容にもとづく自動分類は現在も技術的チャレンジ領域

端末デバイスや利用目的への適応

- 端末の種類毎に専用のページやサイトを人手で用意(i-mode用サイト等)
- 自動変換サーバを用いた端末適応(OpenTV Prism, Oracle Portal-to-Go 等)
- 端末プロフィールを用いた任意端末への動的適応(CC/PPなど)
- 電子地図や電子事典から各Webサイトへの人手によるリンク(目的別ポータル)

動的適応は画面サイズ程度, 用途や利用状況への適応には人手がかかる

(c)2003 INTAP. All rights reserved.

ユビキタス環境に適したポータルサイトとは？

ユビキタス環境への適応

モバイル端末への適応に加え、多様な状況への適応も重要

状況への適応

- 不慣れな街中にいる
- ショッピング・モールにいる
- どこに行こうか迷っている



端末への適応

- 特定のモバイル端末のみに適応すれば良い訳ではない

ユビキタス環境では、即効性のある情報を選択し適切な形で素早く参照できるようにすることがより重要。

コンテンツがマシン可読なメタデータを持つセマンティックWebなら複雑な内容解析を行わずに様々な環境にその場で適応可能!?

(c)2003 INTAP. All rights reserved.

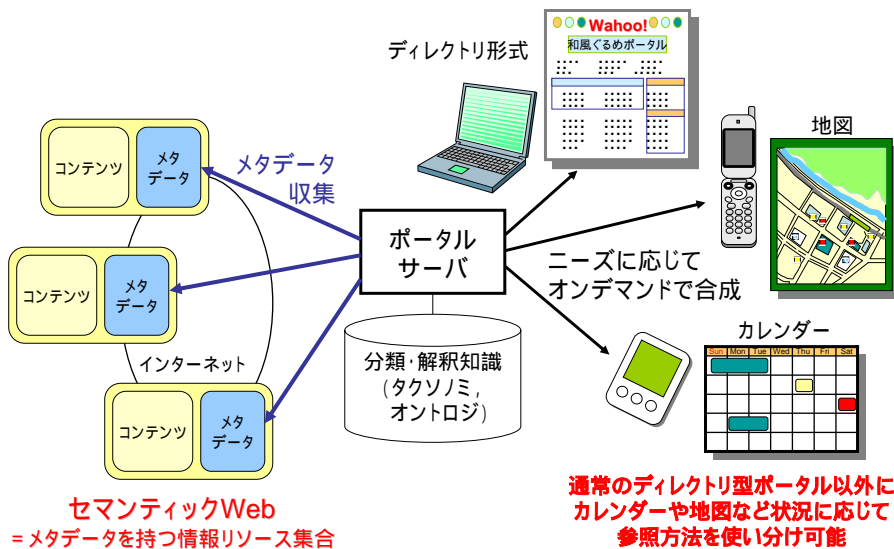
セマンティックWebポータル

メタデータを持つWebコンテンツのポータルサイト

- 各Webコンテンツのメタデータを収集して分類。
- ポータルの用途に応じて直感的に分かり易く表示。
- あらかじめ用意されたメタデータを利用するので、精度の高い分類やWebコンテンツ中にはない情報を利用した分類 / 制御が可能。
- 現在の「RSS リーダ」を原型として発展していく可能性も。

(c)2003 INTAP. All rights reserved.

セマンティックWebポータルのイメージ



(c)2003 INTAP. All rights reserved.

セマンティックWebポータルに対する期待

- 状況や端末に応じた最適なポータル表現
 - 一般的ディレクトリ形式以外に、地図やカレンダー、番組表、グラフなど、TPOに応じた最適な表現形式でWebコンテンツ群を俯瞰、探索可能
- コンテンツの多様な表現への対応
 - コンテンツ毎に異なる単位や呼び方をその場に適した表現で利用可能
- 従来Webとの互換性
 - 従来のブラウザやWebサーバで問題なく扱えるコンテンツを利用可能
- 十分な情報リソースの確保
 - 既存のWebコンテンツに対してポータルに必要なメタデータ付与が容易
 - 新たにコンテンツを作る場合も、誰もが簡単にメタデータを記述可能

(c)2003 INTAP. All rights reserved.

セマンティックWebポールのイメージ(具体例)



地図ビューア上に、関連するWebコンテンツのアンカーを表示。

状況適応の例

例えば、ランチタイムには「ランチを食べられる店」のポータルとして使える。

収集したWebコンテンツのメタデータから、

- ・タイプが料理店に属する
- ・現時間帯で営業している
- ・現地図エリア上にある

等の条件を満たすものを選択し、地図座標に合わせてアンカーを配置。

クリックで**プレビュー生成・表示**

確定(ダブルクリック等)でジャンプ

... など

(c)2003 INTAP. All rights reserved.

セマンティックWebポータル課題

既存の Web ブラウザで正しく表示できる簡単なメタデータ記述方法の確立

XML ベースの RDF 標準シンタックスは、コンテンツ制作者が直接記述するには複雑。

- ・表現可能な範囲を限定しても、より簡単に記述できる標準的な方法を別途用意すべき。
- ・コンテンツに書くのと同じ内容をメタデータとして別に書く手間は最小限にしたい。
- ・コンテンツとメタデータの内容的同期を保証したい。
- ・メタデータを付与したコンテンツも通常のWebコンテンツと同様に見られるようにしたい。



メタデータ記述方法の1案

1. メタデータは HTML 文書内に記述 (コンテンツとメタデータを密に結合)
2. メタデータとしても利用したい本文中の語句にはメタデータ属性のみを付加 (メタデータ記述の省力化と内容的同期の保証)
3. HTML (XHTML) の規格内で記述 (既存のWebブラウザに対応)

(c)2003 INTAP. All rights reserved.

XHTMLコンテンツに対するメタデータ記述例

```
<html xmlns="http://www.w3.org/1999/xhtml" xmlns:erdf="http://www.intap.or.jp/sw/erdf#"
  xmlns:v="http://www.intap.or.jp/sw/restaurant#" lang="ja" xml:lang="ja">
```

```
<head>
```

```
<meta http-equiv="Content-Type" content="text/html; charset=iso-2022-jp" />
```

```
<title>南仏料理 プロヴァンス</title>
```

```
<meta schema="RDF" name="v:type" content="フレンチ" />
```

本文中にないメタデータは meta 要素で記述

```
</head>
```

```
<body>
```

```
<h1 erdf:prop="v:title">南仏料理 プロヴァンス</h1>
```

本文中にあるメタデータは タグに erdf:prop="..." 属性を追加するだけ

```
<table>
```

```
<tr>
```

```
<td></td>
```

画像などにもメタデータで意味付けが可能

```
</tr>
```

```
</table>
```

```
<table border="1">
```

```
<tr>
```

```
<td>住所</td>
```

住所から地図上の位置を (技術的には) 割出し可能

```
<td><span erdf:prop="v:address">東京都港区XXX-X-XX</span></td>
```

対応表やルールが必要

```
</tr>
```

```
<tr>
```

略語から正確な語 (日曜, 祝日) を推定し、実際の日付と関連付け

```
<td>定休日</td>
```

辞書やオントロジが必要

```
<td><span erdf:prop="v:holiday">日曜・祝日</span></td>
```

```
</tr>
```

```
⋮
```

(c)2003 INTAP. All rights reserved.

前記メタデータ記述方法の利点と制約

利点

- 基本的にはHTML (XHTML) 文書の適当な箇所にerdf:prop属性を埋め込むだけ
- ツール開発も容易(メタデータとして宣言したい部分を指定して属性名を選ぶなど)
- RDFデータとして常に「リソース」「プロパティ」「値」の3つ組を書く必要がない
- 本文中に記載のないメタ情報は meta 要素として記述可能
- 本文中に記載のあるメタデータは本文内容との同一性が保証される
- 既存の Internet Explorer や Netscape など主なブラウザでの表示に影響しない

制約

- メタデータの「リソース」は、それが書かれたコンテンツの URI に限定される
- 例えば「名前」というプロパティに対してさらに「日本語表記」と「英語表記」のような階層的プロパティを記述できない(name-j, name-e と分けて対応することは可能)
- 既存のコンテンツにメタデータを付与するには、コンテンツを書き換える必要がある
- RDF標準XMLシンタックスやN3で記述していないため、他システムから直接扱い難い
(この形式のメタデータ記述からRDF/XMLやN3の形式に変換することは容易)

(c)2003 INTAP. All rights reserved.

セマンティックWebポータル課題

位置や時間に関する標準的でパブリックなメタデータ記述手段

例えば地図型ポータルの場合、位置情報(地図上の座標)を簡単に取得できる手段が必要(地図上の位置指定による座標取得、住所情報からの座標計算など)

コンテンツ中のデータ(文字情報)を利用する場合:

- 地図型ポータルの場合、住所からでは空間位置(緯度・経度)への変換が必要
- カレンダー型ポータルの場合、表記のゆれへ対応や数値的日時への変換が必要
(「第3月曜日」、「(定休日は)日・祝・年末年始」など)



オントロジやルール、変換テーブルで対応

セマンティックWebの話では、まるで魔法使いと魔法の杖のごとく登場するエージェントとオントロジ。しかし、これらの実現が最大の課題であり挑戦。すでに実用レベルの機能と規模を持つものも多数。

(c)2003 INTAP. All rights reserved.

セマンティックWebポータル課題

メタデータの信頼性の確立

信頼性 = 真正性・・・内容に虚偽や誤りがないこと
完全性・・・改竄や破損がないこと
可用性・・・正しく利用できること

- メタデータは「人」が見ずに利用されることが多く、誤りや改竄に気付きにくい。
- メタデータを直接利用するマシンは、内容的な正しさを評価する知識に乏しい。



今できること

真正性に対して： オントロジの利用や関連情報の収集・比較による矛盾発見
完全性に対して： デジタル署名による改竄検知
可用性に対して： スキーマの検証

まとめ

- 現Webの主要なアプリケーションであるポータルサイトがセマンティックWebでどうなるかについて考察。
- IT 利用環境のコビキタス化とWebのセマンティック化は相互に協調し合える。
- 現Webのコンテンツ作成者が容易に参加できるように記述が簡単で扱い易いメタデータ付与の1方法を提案。
- 様々な用途、表現形式に対応するためオントロジやルールの整備が必要。
- メタデータの信頼性はそれ自体が重要な課題であるとともにサービスとしてのポータルの信頼性にも繋がる。