

W3Cにおけるメタデータ

慶應義塾大学SFC/W3C

萩野 達也

hagino@w3.org

目次

- World Wide Web Consortium
- Webのこれまでの発展
- Webのメタデータ
- セマンティックWeb

World Wide Web Consortium

W3Cについて



- 1994年に設立
- DirectorはWebの創始者Tim Berners-Lee
- ホスト組織
 - MIT
 - INRIA
 - 慶應義塾大学SFC
- 参加組織500ほど
 - <http://www.w3.org/>

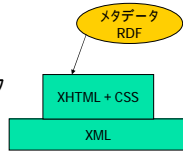
W3Cの活動内容

- Web基盤技術の標準化
 - Leading the Web to its Full Potential...
- Domain
 - Architecture
 - Document Formats
 - Interaction
 - Technology and Society
 - Web Accessibility Initiative (WAI)

Webのこれまでの発展

Webのこれまでの発展

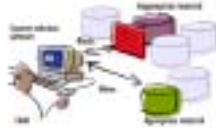
- HTML
 - Web文書
- XML
 - インターネット上の標準データ形式
- セマンティックWeb
 - エージェントの空間



Webのメタデータ

PICS

- Platform for Internet Content Selection
- インターネット上の有害情報へのアクセスを禁止するためのメカニズム



PICS: Internet Access Controls Without Censorship Communications of the ACM 1996, vol. 39(10), pp. 87-93

フィルタリング・ソフトウェア

- ページにつけられたラベルにしたがってアクセスしてよいかを決める



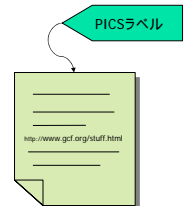
PICS: Internet Access Controls Without Censorship Communications of the ACM 1996, vol. 39(10), pp. 87-93

PICSの構成要素

- PICSレーティング・サービス
 - [Rating Services and Rating Systems \(and Their Machine Readable Descriptions\) Version 1.1, 31 Oct 1996](#)
- PICSラベル
 - [PICS Label Distribution Label Syntax and Communication Protocols Version 1.1, 31 Oct 1996](#)
- PICSフィルタリング規則
 - [PICSRules 1.1, 29 Dec 1997](#)

PICSラベルの例

```
(PICS-1.1 "http://old.rsac.org/v1.0/"
labels
on "1994.11.05T08:15-0500"
until "1995.12.31T23:59-0000" for
http://www.gcf.org/stuff.html
by "John Doe"
ratings (1 3 s 2 v 0))
```



PICSラベルの指定方法

- HTMLのMETAタグとして埋め込む
- HTTPレスポンスに入れる
- レーティング・サービスから別に受け取る

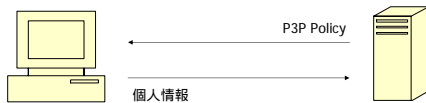
```
<head>
<META http-equiv="PICS-Label" content='
(PICS-1.1 http://www.gcf.org/v2.5
labels on "1994.11.05T08:15-0500"
until "1995.12.31T23:59-0000"
for http://w3.org/PICS/Overview.html
ratings (suds 0.5 density 0 color/hue 1)
'>
</head>
```

PICSメタデータの問題点

- メタデータの値は、数値だけである
- 著者名などにより指定することができない
- PICS-ng
 - 著者名など数字以外も利用できるように
 - 複数の著者などを許すように
- 一般的なメタデータ(RDF)に発展

P3P

- Platform for Privacy Preferences Project
 - [The Platform for Privacy Preferences 1.0 \(P3P1.0\) Specification, 16 April 2002](#)
- 個人情報の取り扱いを記述
 - Webサイトがどのような目的で個人情報を集めているかを表明
 - (今後)個人情報を自動的にサーバに送る



P3Pポリシーの例

```
<POLICIES xmlns="http://www.w3.org/2002/01/P3Pv1">
<POLICY name="forBrowser"
discurl="http://www.catalog.example.com/PrivacyPract
iceBrowsing.html"
xml:lang="en">
<ACCESS><nonident/></ACCESS>
<STATEMENT>
<PURPOSE><admin/><develop/></PURPOSE>
<RECIPIENT><ours/></RECIPIENT>
<DATA-GROUP>
<DATA ref="#dynamic.clickstream"/>
</DATA-GROUP>
</STATEMENT>
</POLICY>
</POLICIES>
```

P3Pの普及の効果

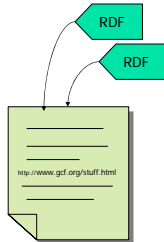
- 企業が個人情報の取り扱いを表明することによって、利用者はその企業を信頼することができる
- 利用者のプライバシー意識の向上
- 電子商取引の効率化の促進

P3Pメタデータの問題点

- RDFではなくXMLで記述されている
 - RDFの普及がまだなのでXMLを利用
 - 将来的にはRDFで記述される
- 個人情報をクライアントからサーバに送る方法を規定していない
 - FORMや既存の方法がすでにあるため
 - 将来的には方法を考える

RDF

- Resource Description Framework
 - Resource Description Framework (RDF) Model and Syntax Specification, 22 Feb 1999*
- Webのメタデータの標準
 - PICSラベルを拡張
 - Web上のリソースに関して任意の記述が可能

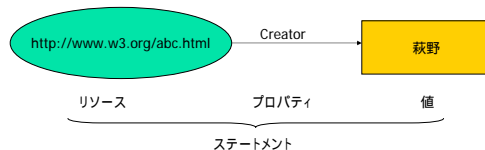


2002/6/5

19

RDFモデル

- 3つの組
 - リソース
 - プロパティ
 - 値



2002/6/5

20

XMLによるRDFシンタックス

- XMLにより表現

```
<rdf:RDF>
  <rdf:Description about="http://www.w3.org/abc.html">
    <s:Creator>萩野</s:Creator>
  </rdf:Description>
</rdf:RDF>
```

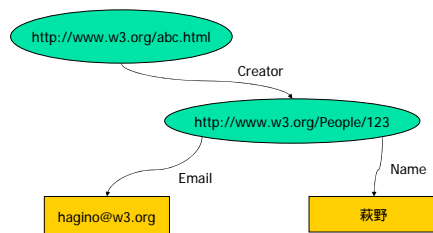
- 省略文法

```
<rdf:RDF>
  <rdf:Description about="http://www.w3.org/abc.html"
    s:Creator="萩野" />
</rdf:RDF>
```

2002/6/5

21

値はリソースでもかまわない



2002/6/5

22

RDFによる記述

```
<rdf:RDF>
  <rdf:Description about="http://www.w3.org/abc.html">
    <s:Creator rdf:resource="http://www.w3.org/People/123"/>
  </rdf:Description>

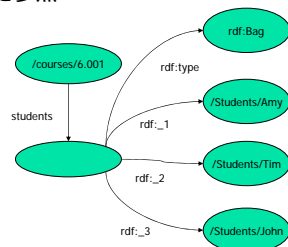
  <rdf:Description about="http://www.w3.org/People/123">
    <v:Name>萩野</v:Name>
    <v:Email>hagino@w3.org</v:Email>
  </rdf:Description>
</rdf:RDF>
```

2002/6/5

23

RDFコンテナ

- 複数のリソースを参照
 - Bag
 - Sequence
 - Alternative



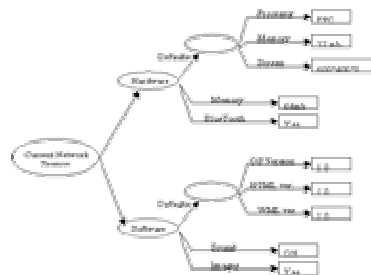
2002/6/5

24

RDFの応用

- CC/PP
 - Composite Capabilities/Preference Profiles
 - 機器情報やユーザの嗜好を記述
- RSS
 - RDF Site Summary
 - Webサイトの要約を記述
 - ニュース配信, 製品紹介などに利用

CC/PP



セマンティックWeb

セマンティックWeb

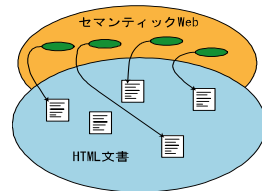
- 次世代のWeb
- 機械的処理を可能とするメタデータ空間
- 賢い検索が可能
- エージェントが処理を行う空間

Webの2つの役割

- 人と人とのコミュニケーション
 - いつでも、だれでも、どこからでも情報のやり取りができる
 - HTMLにより実現
- 人と機械のコミュニケーション
 - 人の問題解決を助ける
 - セマンティックWebで実現

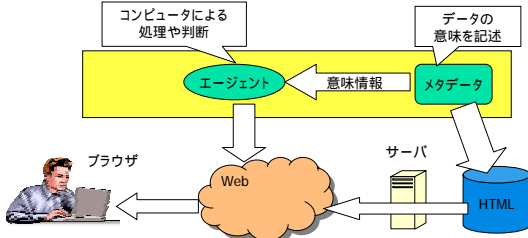
セマンティックWebと現在のWeb

- 別の空間である
- HTMLが作るWebはそのまま
 - HTMLは人が読んで理解する文書
- セマンティックWebはメタデータが作る空間
 - 機械が処理するデータがある



エージェント空間

- エージェントは人に代わりメタデータを処理する



2002/6/5

31

セマンティックWeb利用例(1)

- 「長崎にある旅行代理店を探したい」
 - 「長崎」、「旅行代理店」の2つのキーワードで検索
 - 長崎へのツアーを企画している旅行代理店も見つかる
 - 長崎さんがやっている旅行代理店も見つかる
 - 検索結果から人が判断するしかない
- 旅行代理店のページのメタデータとして住所がつけられていれば良い

2002/6/5

32

セマンティックWeb利用例(2)

- 「水曜日に営業している旅行代理店を探したい」
 - 現在の検索エンジンでは難しい
- 営業日に関するメタデータを付加すればよい
 - 1週間は月火水木金土日からなる
 - 営業日と休業日は相反する

2002/6/5

33

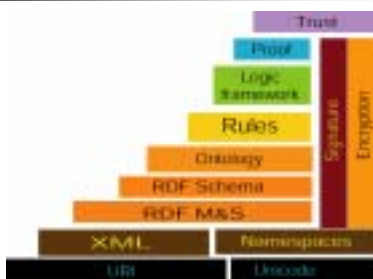
セマンティックWeb利用例(3)

- 「もっともやすい長崎観光ツアーを探したい」
 - 旅行代理店のリストから、次々と選んで長崎観光ツアーを調べていくしかない
 - どこかの旅行代理店の「日本一安いツアー」を信じるしかない
- 旅行代理店のリストに対して、次々と長崎観光ツアーを検索するスクリプトを書くことができると良い
 - 検索結果をつぎの処理にいかす
 - 処理結果がどのようにして導かれたのかを証明する

2002/6/5

34

セマンティックWebのアーキテクチャ

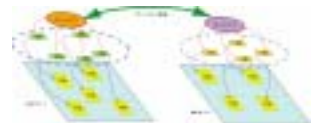


2002/6/5

35

RDF Schemaとオントロジ

- RDF Schema
 - RDFのデータ型
 - クラス階層
- オントロジ
 - 語彙の定義
 - 語彙間の変換を定義



2002/6/5

36

W3C セマンティックWebでの メタデータ

- だれもが何に関しても自由に記述可能
- 矛盾するデータがあるかもしれない
- 記述はすべて文脈において判断されなくてはならない
 - 100%信用はありえない
 - 信用は文脈においてのみできる

2002/6/5

37

W3C まとめ

- World Wide Web Consortium
- Webのこれまでの発展
- Webのメタデータ
- セマンティックWeb

<http://www.w3.org/>

2002/6/5

38

W3C お問い合わせ先

- 慶應義塾大学SFC研究所W3C
- E-mail: keio-contact@w3.org
- 電話: 0466 - 49 - 1170
- FAX: 0466 - 49 - 1171
- 担当: 竹内

お気軽にお問い合わせください

2002/6/5

39