

SemanticWebの概要と応用 展望

2003.5.27

野村 直之, 博士(理学)

Justsystem, Technology Advisor

法政大学エクステンションカレッジ & 大学院ITPC

XML Consortium, Evangelist,

XML Consortium, SemanticWeb WG Leader

Int'l Workshop on "NLP and XML/ SemanticWeb " Organizer, Chair

情報処理学会デジタルドキュメント研究会運営委員

FIT2003 Program委員(NLP & ドキュメント領域)

ビジネスモデル学会ナレッジマネジメント研究会 座長

SemanticWebについて (概要)

- 次世代Webの上位規格SemanticWebの階層、言語モジュール群は発展途上にあり、今後も鋭意、拡張・統合・改廃(差替)...
- ..3要素「メタデータ」「オントロジ」「エージェント」を汎用性の高い、モジュラーな構成で実現しようとする方向性は不変...
- メタデータによる情報流通の促進**を中心に、SemanticWebのEC, KM, EAI (with WebServices)への応用を模索...
- ...メタデータ自動抽出... メタデータ標準準拠の意義、視覚化やメタデータDB化...応用実現のための各種要素技術...
- ...「常識検索」に象徴される次世代検索や、XML複合文書の自在な要約など、「ユーザ(消費者)主権」を推し進めた、Web情報流通の一将来像...
- ...RDF-Schemaによる構造化情報抽出やデータ解釈といった、知識による情報加工の仕組みを備えた
- ...高付加価値のドキュメント流通促進...

Today's Talks..

- 2003.5.27 Tue. pm “SemanticWeb Day”
in XML Consortium Week

「SemanticWebの概要と応用展望」

法政大学 野村直之

概要:

次世代Webの基本アーキテクチャSemanticWebが昨年来急速に注目を集めている。「機能の連携」を担うWebServicesと相補的に、SemanticWebは「コンテンツの連携;リポジトリ」を担う。講演では、SemanticWebの全体像から「常識検索」「ノイズレス検索」やEAI、KM、複合情報提供(エージェント)等、有望なアプリケーションのイメージを紹介するとともに、新ビジネスモデルの可能性や、今後の課題についてお話しいたします。

「メタデータ利用の効果

～メタデータ標準の活用事例

～メタデータ自動抽出の要素技術はどこまでいけるか？」

ジャストシステム 川崎洋治

概要:

メタデータは、コンテンツ管理の基盤技術であり、情報流通を大きく促進すると考えられます。伝統的な図書館の10進分類、Dublin Core、MPEG7、SemanticWebのリソース管理の標準規格からオントロジ、Webサービスの各種プロセス記述まで、多彩な分野で注目されています。本講演では、既存の様々なメタデータ標準をどう評価し、取捨選択してアプリに適用すべきかの活用事例、ならびに、メタデータ普及の鍵を握る「既存コンテンツからのメタデータ自動抽出」の要素技術の現状と展望についてお話しいたします。

「XHTML, SVGにRDFを埋め込んでWebサマリー生成コントロール (RSsS)」

日本ユニシス 小林 茂 cf. 5/23 Fri. IPSJ Sig-DD@BusinessShow

概要:

セマンティックWebで最も基本的かつ重要な概念であるメタデータとその応用に焦点を当て、XHTMLに埋め込んだメタデータからRSS標準に準拠したSite Summaryを自動抽出する方法を紹介いたします。これは、コンテンツ主導の情報統合という流れに沿って、XML文書スタイル関係の属性をオブジェクト指向的に統一管理し、著者や編集者の意図を反映したサマリーやデータベース・レコードへの情報抽出の処理内容をSemanticWebのメタデータ標準で統一的に記述する試みであります。

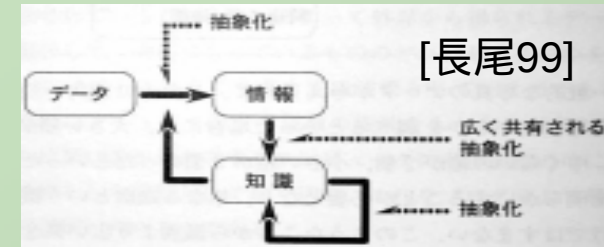
Road Map

1. **今、なぜSemanticWebか？** メタデータ、オントロジ(、エージェント)と積み重ね、サービス[内容]中心アーキテクチャ、アプリへ
 1. メタデータ
 2. オントロジ
 3. エージェント
2. WS vs. SW
3. Web, EC, KM応用とSemanticWeb
4. **試作アプリ1：象徴的アプリ「常識検索」**
5. 試作アプリ2：RSsS for Controlled XML Summarization framework
6. **その他、アプリのアイディア達 (2002.8-9頃発案)**

1. 今なぜSemanticWebか？

Web HTMLの混沌を解決する一連の流れを高度に解決するXML言語
検索, 組織化 ファミリーの正統な嫡子. Contents管理用のWeb上位規格
視覚化..
e.g. この近所で水曜に空いている歯医者とは? (©萩野)

KM 知識管理・検索ソリューションの要請から cf. AI



E-Business EAI B2Bi

WebServices: サービス(機能)結合; インタフェース

↑ ↓

SemanticWeb: 主にコンテンツのオープンな流通のための構造
定義、'翻訳'; 内容

cf. WebServices Architecture :

→ オープン化志向; 属性等に可能な限りSemanticWeb使う

WS = 機能の連携・統合

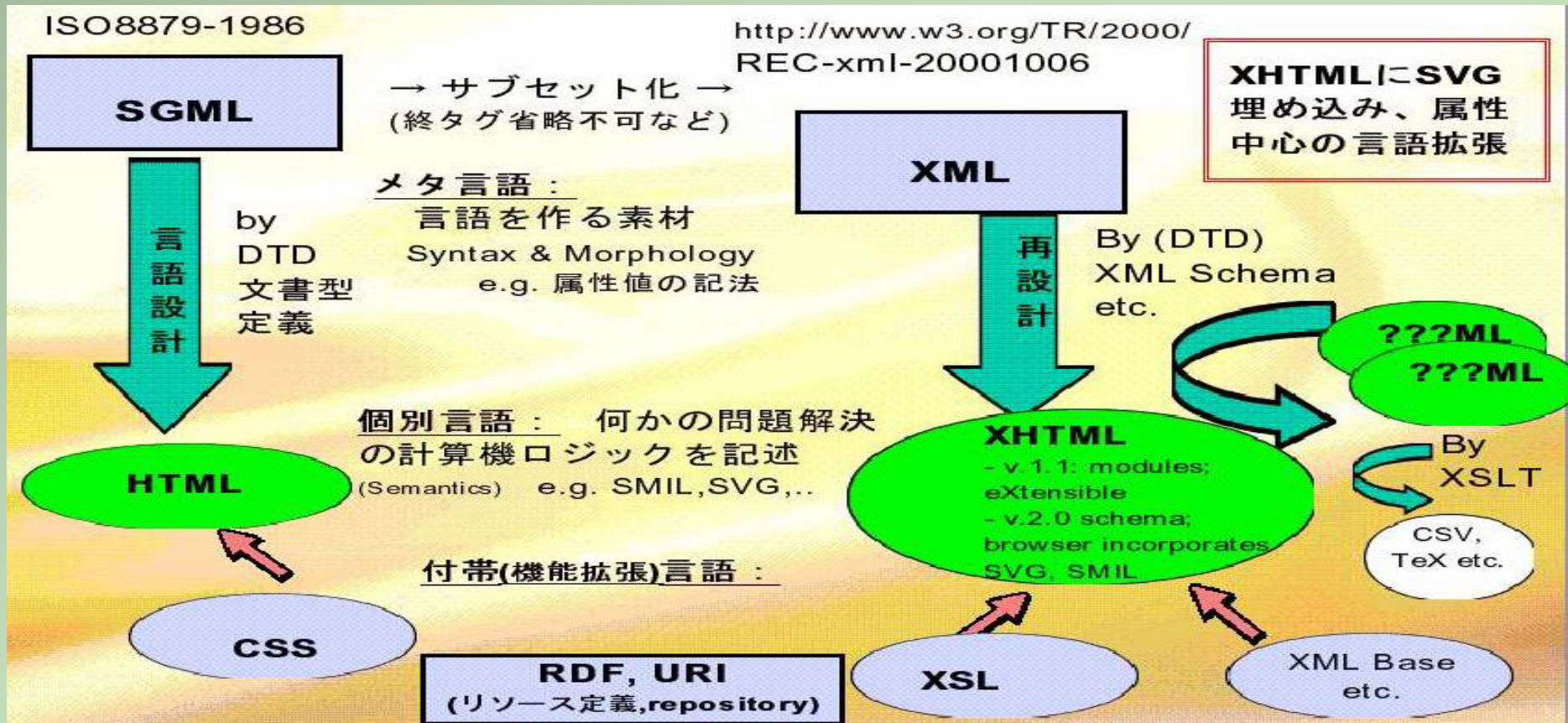
SW = コンテンツの連携・統合

Web文書規格のおさらい

1. 今なぜSemanticWebか？

{過去=SGML/HTML, 現在=XML/複合文書, 未来=SemanticWeb}

XML: メタ言語のSyntax, Morphology



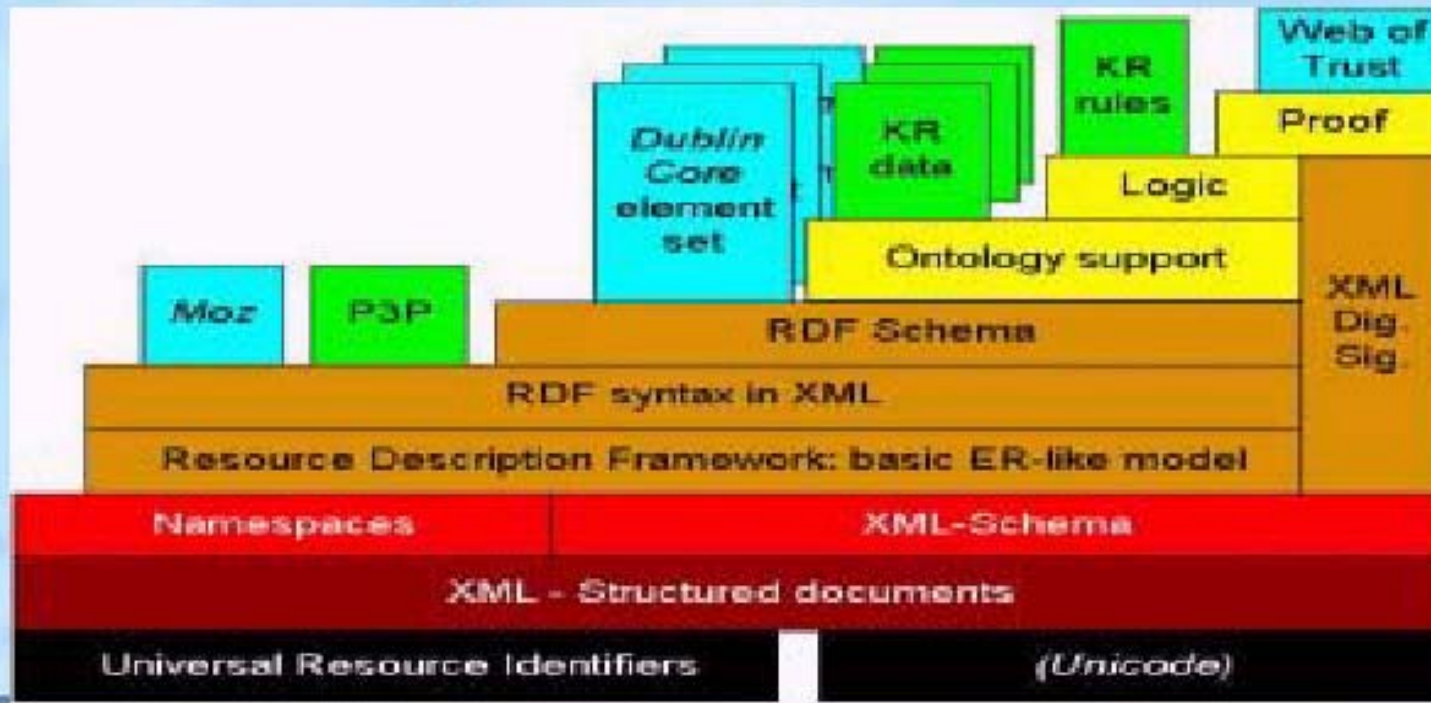
XML: 強力な“共通語”と“方言”を両立。
数値系(基幹系)と文書系(情報系)を統合。
変換スクリプト, 認証プロセス等“動き”も表現可。

WebService
SemanticWeb
http
XML
上位層へ進出。。

1. 今なぜSemanticWebか？ Semantic Webの実質,活用のステップ (1)メタデータ

機械間

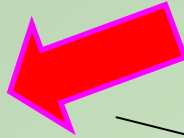
- 直接的には、コミュニケーションのためのもの
- 【間接的に】 KMに貢献- 知識リポジトリは人間に理解しやすい形に(XSLT etc.で) 管理、変換可能
- T.B.Lee:"Builing Blocks of the Future Web", 2000



- 1999年Tim Berners-Lee提唱の次世代のWebビジョン
- 2001年からW3Cの活動として承認

SubRoadMap: Semantic Webの実質,活用のステップ

(1)メタデータ



リソース管理, リポジトリ機能

(2)オントロジ[変換]

コンテンツ管理

別のボキャブラリ空間との相互通訳。探索範囲の拡大。

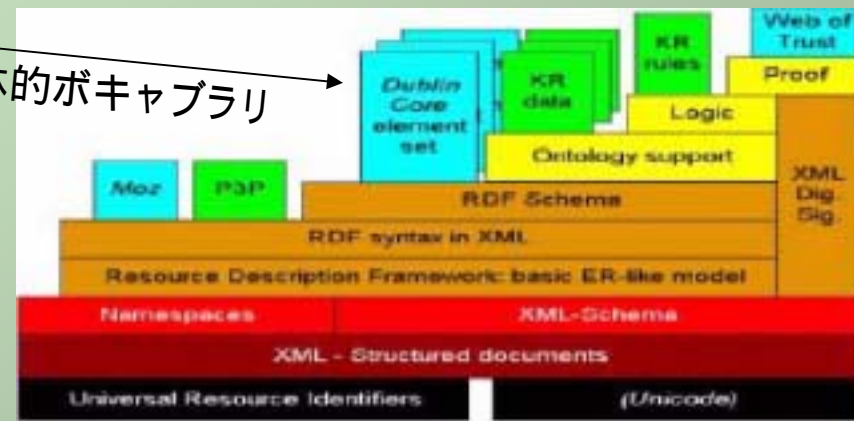
(3)知的エージェント, サービス「内容」の連携

実現 by DAML-S? Semantic Web Service?

メタデータって? HTML, DC

- 書誌情報 5W1H ;書籍(Contents本体)に対する図書カード
- <HTML>
 - <HEAD>
 - <META> - - - -
- DC: Dublin Core 電子図書館の書誌情報15の統一標準
Title, Creator, Subject, Description, Publisher,
Contributor, Date, Type,
Format, Identifier, Source,
Language, Relation,
Coverage, Rights

具体的ボキャブラリ



様々なメタデータ規格

- 汎用
 - DC (Dublin Core) , RDF, RSS
- ユーザプロフィール
 - P3P
- フィルタリング
 - PICS
- マルチメディア
 - MPEG7
- 写真・画像
 - RDFPic <http://jigsaw.w3.org/rdfpic/>
- 地理・観光情報
 - G-XML, JMP, ...
- 電子政府
 - MIREG, e-GMS, AGLS, e-Gov
- eLearning
 - LOM(Learning Object Metadata), ELM, LIP
- News, TV放送
 - XMLNews, NewsML, TV Anytime, ARIB,
- 音楽
 - MusicBrainz, ...
- コンテンツ管理
 - cIDF (Content ID Forum)

.....

WP文書, Web
頁にシャッター速
や絞り値, 緯度・
経度は無い!

小林さん

川崎さん

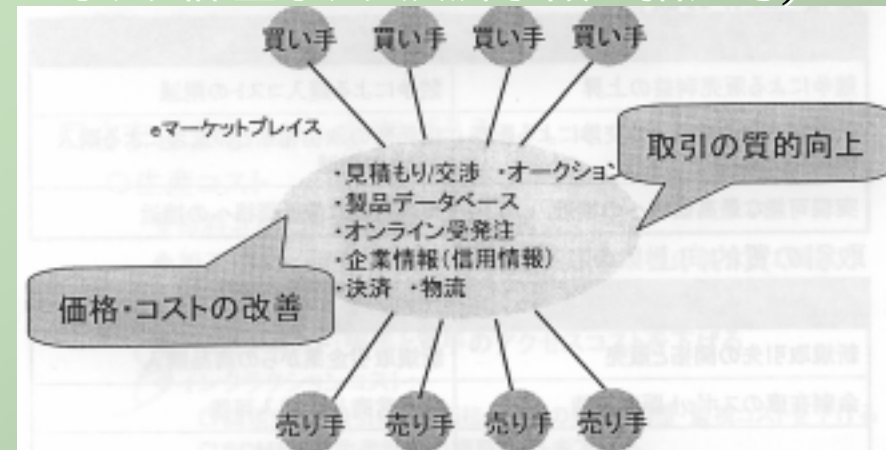
cf. 何が本来「メタ」かを判断するのは案外難しい

eDealML^[野村2002]の4W3H: モノ、サービス実体に付属させればメタデータ

- WHAT = 商品(モノ)、サービス自体の属性 (色,形,大きさ..)
- WHO = 取引相手の属性 (様々な与信)
- WHEN = 注文日時
- WHERE = 商品配送、サービス実現の場所
- HOW = 取引手法 (市場・マッチング手法・課金手法・決済手段の指定等)
- HOW MUCH = 価格
- HOW LONG = 納期

■ 各項が下位構造をもつ

■ 注文は2種類: 売 or 買



観光情報等も、「メタ」か主コンテンツか、目的、視点により異なる。

5W1Hだからといってメタデータとは限らない

- 「取引相手」WHOの品質(5W1Hの下部属性)
- 格付け機関の信頼性のグレード(格付け機関名の集合etc.)
- 財政状況？ 債務超過？ 連続赤字？ キャッシュフロー？メインバンク？
- 過去の自分 / eMPとの取引履歴(全部満足だったか)

必要なだけ木構造を深くして説明能力追加 例: 格付け者格付け者の格付け...

物流にも品質 WHERE, HOW LONG

物流の主属性: availability, 納期, 料金

物流の品質(下位属性):

- 確実につくか(保証性)
- 品物が無事か(安全性)
- 確実につかない場合の対処(補償のグレード)
- 時期が確定(幅がせまいのがよい)

納期 モノ場所 倉庫都合 トラック都合 ペイパーワーク遅延...

SemanticWebでは個々のメタデータでなく、
メタデータの記述形式を規定

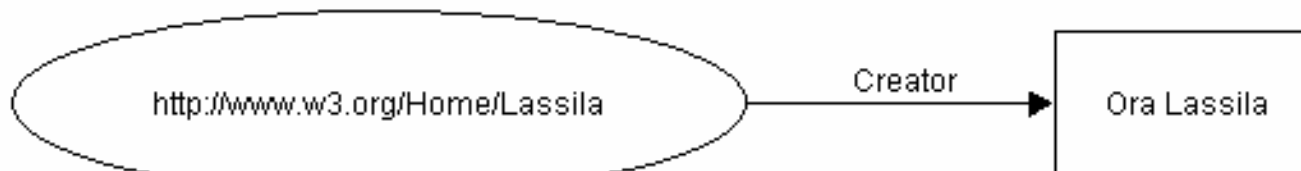
Resource Description Format

RDF: 2つのリソース間の関係を定義するXML言語

Ora Lassila is the creator of the resource <http://www.w3.org/Home/Lassila>

「Ora Lassilaさんはリソース<http://www.w3.org/Home/Lassila>の作者です。」

```
<rdf:RDF>
  <rdf:Description about="http://www.w3.org/Home/Lassila">
    <s:Creator>Ora Lassila</s:Creator>
  </rdf:Description>
</rdf:RDF>
```



メタデータ活用上の問題、 課題が数ある中で..

RDF記述は
信頼できるか？

次世代検索エン
ジン騙しを抑止
可能か？

↓
RDFS (オントロジ),
XMLSignature
Social Filtering..

信頼できそうな情報

```
<?xml version="1.0"?>
```

```
<rdf:RDF
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
```

```
  xmlns:s="http://description.org/schema/">
```

```
  <rdf:Description
```

```
    about="http://www.w3.org/Home/Lassila">
```

```
      <s:Creator>Ora Lassila</s:Creator> ← ここ正直
```

```
    </rdf:Description>
```

```
  </rdf:RDF>
```

嘘の情報

```
<?xml version="1.0"?>
```

```
<rdf:RDF
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
```

```
  xmlns:s="http://description.org/schema/">
```

```
  <rdf:Description
```

```
    about="http://www.w3.org/Home/Lassila">
```

```
      <s:Creator>Ora Gokuu</s:Creator> ← ここうそ
```

```
    </rdf:Description>
```

```
  </rdf:RDF>
```

RDF応用言語 for Web情報流通

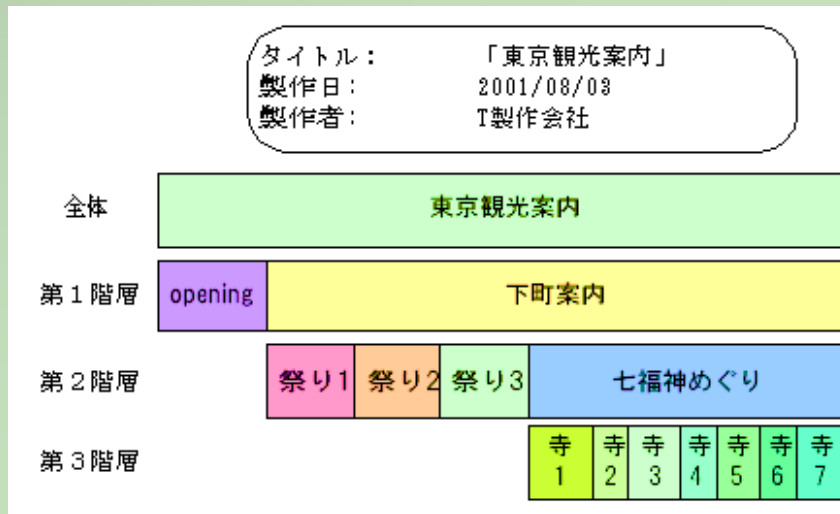
- RSS (RDF Site Summary or Rich Site Summary)
Webサイトの要約をRDFのメタデータとして記述することで、
検索・配信を容易に. <http://purl.org/rss/1.0/spec/>
- PICS (Platform for Internet Content Selection)
有害コンテンツのフィルタリング。ページにメタデータ(PICSラベル)を付与し、クライアントでブロック。
- P3P (Platform for Privacy Preferences Project)
個人のプライバシー情報伝達のための枠組み。各サイトがプライバシーポリシーをメタデータとして公開(XMLベース)。利用者はそれを見て(半自動で)個人情報公開してよいかを判断。

MPEG7: SemanticWeb じゃないけれど重要な マルチメディアコンテンツ・メタデータ構造化標準

<http://www.itscj.ipsj.or.jp/mpeg7/>

http://www.itscj.ipsj.or.jp/mpeg7/mpeg7_body_3.html

日本語FAQページなども充実。
コンテンツ自体は通常は別ファイル。



```
<AudioVisualSegment id="structure-22">
  <StructuralUnit href="urn:ricoh:xxVISION:SegmentTypeCS:4">
    <Name xml:lang="en">structure</Name>
  </StructuralUnit>
  <TextAnnotation type="headline">
    <FreeTextAnnotation xml:lang="en">&quot;Sanja Matsuri&quot;
    Festival
  </FreeTextAnnotation>
  </TextAnnotation>
  <TextAnnotation type="commentary">
    <StructuredAnnotation>
      <Where>
        <Name>Asakusa Jinjya Shrine</Name>
      </Where>
      <When>
        <Name>Saturday and Sunday closest to May 17th and
        18th</Name>
      </When>
    </StructuredAnnotation>
  </TextAnnotation>
  <MediaTime>
    <MediaRelTimePoint mediaTimeBase="ancestor::Description/
    MultimediaContent[1]/AudioVisual[1]/MediaTime[1]">
      P0D7OH1M29S11N3OF
    </MediaRelTimePoint>
    <MediaIncrDuration mediaTimeUnit="P0D7OH0M0S1N3OF">
      1873
    </MediaIncrDuration>
  </MediaTime>
</AudioVisualSegment>
```

図. あるシーンに対するMPEG-7記述の例(抜粋)

http://www.ricoh.co.jp/src/multimedia/MovieTool/about/index_j.html

より引用

MPEG7: シナリオ, サンプルコンテンツ DLサイト@ NII (一部有料)

<http://research.nii.ac.jp/VDB/>

MPEG7:MovieTool

<http://ext.ricoh.co.jp/tech/>



Index of /VDB

Name
[Parent Directory](#)

[1.jpg](#)
[2.jpg](#)
[3.jpg](#)
[4.jpg](#)
[5.jpg](#)
[6.jpg](#)
[7.jpg](#)
[8.jpg](#)
[9.jpg](#)
[10.jpg](#)
[11.jpg](#)
[12.jpg](#)
[13.jpg](#)
[14.jpg](#)
[15.jpg](#)
[16.jpg](#)
[17.jpg](#)
[18.jpg](#)

Name	Size
1.jpg	23-KB
2.jpg	48-KB
3.jpg	23-KB
4.jpg	23-KB
5.jpg	1.6-MB
6.jpg	3.8-MB
7.jpg	3.8-MB
8.jpg	23-KB
9.jpg	13-KB
10.jpg	1.6-MB
11.jpg	18-KB
12.jpg	2.8-MB
13.jpg	14-KB
14.jpg	1.6-MB
15.jpg	72-KB
16.jpg	1.6-MB
17.jpg	14-KB
18.jpg	1.6-MB
19.jpg	18-KB
20.jpg	1.6-MB
21.jpg	18-KB
22.jpg	1.6-MB
23.jpg	18-KB

http://www.ricoh.co.jp/src/multimedia/MovieTool/about/index_j.html

より: MPEG-7関連の参考資料

[descriptionExample000.xml](#) (ZIP 12,807 bytes)

通称Monster Jr.と呼ばれる、ISO MPEG-7 MDS Groupが作成した MPEG-7ファイル(インスタンス)例です。ダウンロードしてご参照ください。

書籍 "Introduction to MPEG-7, Multimedia Content Description Interface",

B.S.Manjunath, Philippe Salembier, Thomas Sikora, John Wiley & Sons, Ltd, ISBN-0-471-48678-7

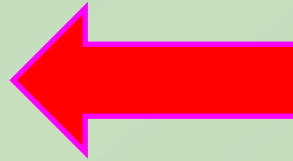
SubRoadMap:

Semantic Webの実質,活用のステップ

(1) メタデータ

(2) オントロジ

(3) エージェント



メタデータから オントロジへ

- ・ RDF (*Resource Description Framework*)

- ・ RDFS (*RDF Schema*)

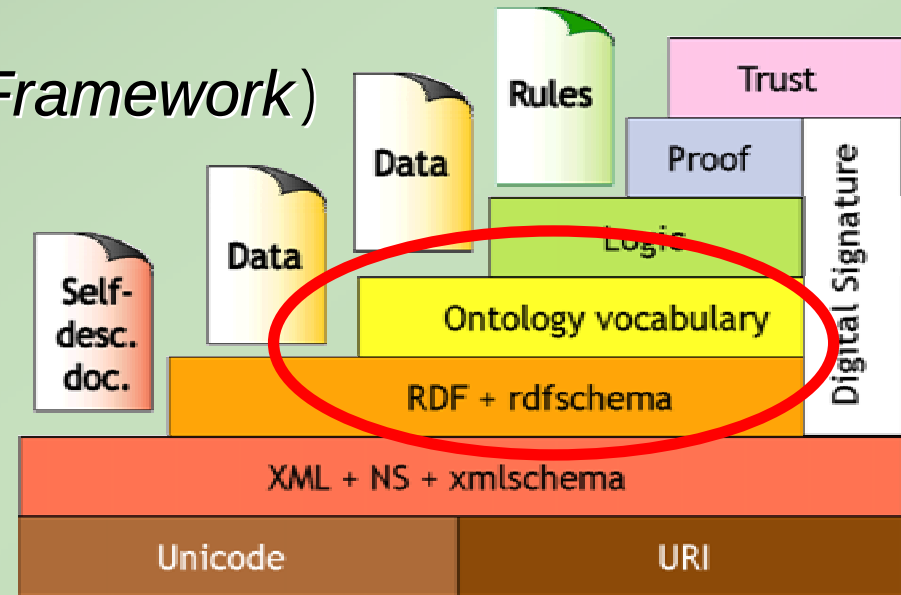
メタデータの構造を定義

- ・ DAML + OIL

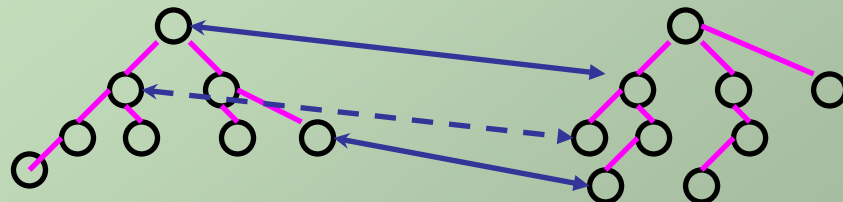
(*DARPA Agent Markup Language + Ontology Inference Layer*)

- ・ DAML-S

(*DAML-Service*)



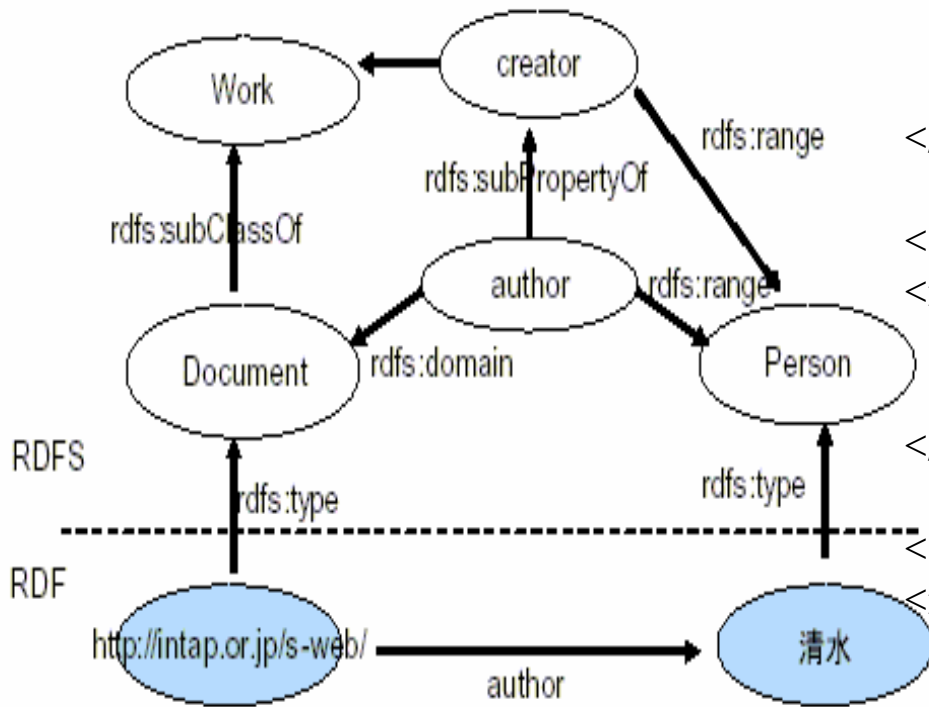
オントロジは異なる意味空間を仲介
cf. namespace: 共存可にするだけ



オントロジの構造定義(RDFS)

→ 意味世界を“通訳”

- ・ クラス階層、プロパティ階層、値の制約 (range, domain)などを定義



<!--例：クラスManの定義-->

```
<rdfs:Class rdf:ID="Man">
```

```
<rdfs:subClassOf rdf:resource="#Person"/>
```

```
<rdfs:subClassOf rdf:resource="#Male">
```

```
</rdfs:Class>
```

<!--例：プロパティhasParentの定義-->

```
<rdfs:Property rdf:ID="hasParent">
```

```
<rdfs:Domain rdf:resource="#Person"/>
```

```
<rdfs:range rdf:resource="#Person">
```

```
</rdfs:Property>
```

<!--例：プロパティhasFatherの定義-->

```
<rdfs:Property rdf:ID="hasFather">
```

```
<rdfs:subPropertyOf rdf:resource="#hasParent"/>
```

```
<rdfs:range rdf:resource="#Man">
```

```
</rdfs:Property>
```

↑ セマンティックWebコンファレンス2002
予稿, Section3 ©INTAP より引用

オントロジ内容記述の例 (DAML+OIL)

<http://www.daml.org/2001/03/daml+oil-ex#MarriedPerson>

```
<daml:Class rdf:ID="MarriedPerson">
```

```
<daml:intersectionOf rdf:parseType="daml:collection">
```

```
<daml:Class rdf:about="#Person"/>
```

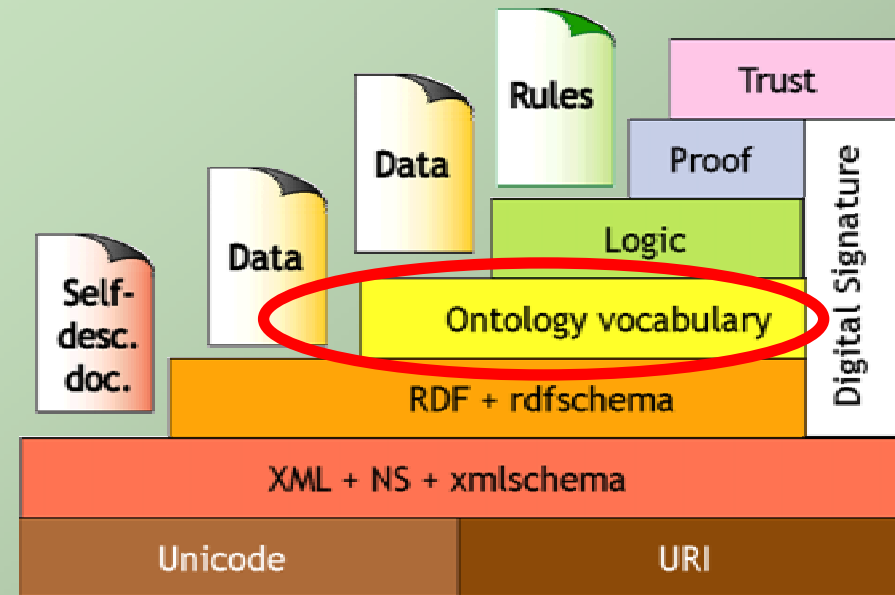
```
<daml:Restriction daml:cardinality="1">
```

```
<daml:onProperty rdf:resource="#hasSpouse"/>
```

```
</daml:Restriction>
```

```
</daml:intersectionOf>
```

```
</daml:Class>
```



推論規則を参照したオントロジ内容記述の例

```
<daml:Class rdf:ID="ソクラテス">
```

```
  <daml:intersectionOf rdf:parseType="daml:collection">
```

```
    <daml:Class rdf:about="#Person"/>
```

```
    <daml:Restriction>
```

```
      <daml:onProperty rdf:resource="#人は必ず死ぬ"/>
```

```
    </daml:Restriction>
```

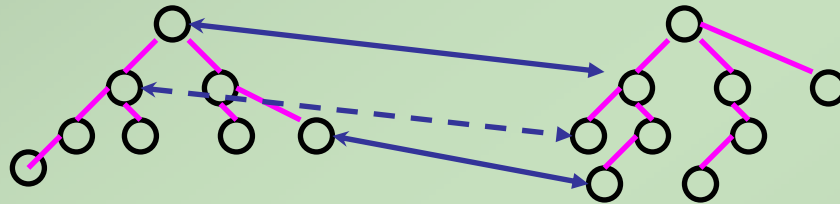
```
  </daml:intersectionOf>
```

```
</daml:Class>
```

OWL : Web Ontology Language

<http://www.w3.org/TR/2002/WD-owl-features-20020729/>

RDF, RDF-Sのリソースの位置付けの仕組みを用いて
Web上の分類体系(クラス)やその間の関係、さらにそれら
の推論ルール群(定理, 公理)を記述する言語。



<http://www.w3.org/TR/webont-req/>

オントロジ定義による知識の相互運用が可能になって初
めて、エージェント(人, DB)が動き回って仕事ができる。

RDFタプルの組等を用いてOWLの語彙を構成

OWL: (つづき)

構成要素 (rdf, rdfs語彙の不足を補い発展させていく。今後も!):

- **ヘッダ要素**: バージョン情報、他のオントロジーのインポートを記述

```
<owl:Ontology rdf:about=""> <owl:versionInfo>webont.html, v.0.1; 2002-07-25 </owl:versionInfo> <owl:imports  
  rdf:resource="http://www.w3.org/2002/07/owl"/> <dc:creator>Lee,  
  Tim, Berners</dc:creator>
```

```
</owl:Ontology>
```

- **クラス要素**: オブジェクトのクラスを定義: 集合演算他
rdfs:subClassOf owl:sameClassAs owl:unionOf etc.

- **プロパティ要素**: プロパティを定義

rdfs:range owl:equivalentTo owl:inverseOf etc.

- **インスタンス**: クラスおよびプロパティ

rdf:type owl:oneOf etc.

```
<owl:ObjectProperty rdf:ID="hasHusband">
```

```
  <owl:inverseOf rdf:resource"#hasWife"/>
```

```
</owl:ObjectProperty>
```

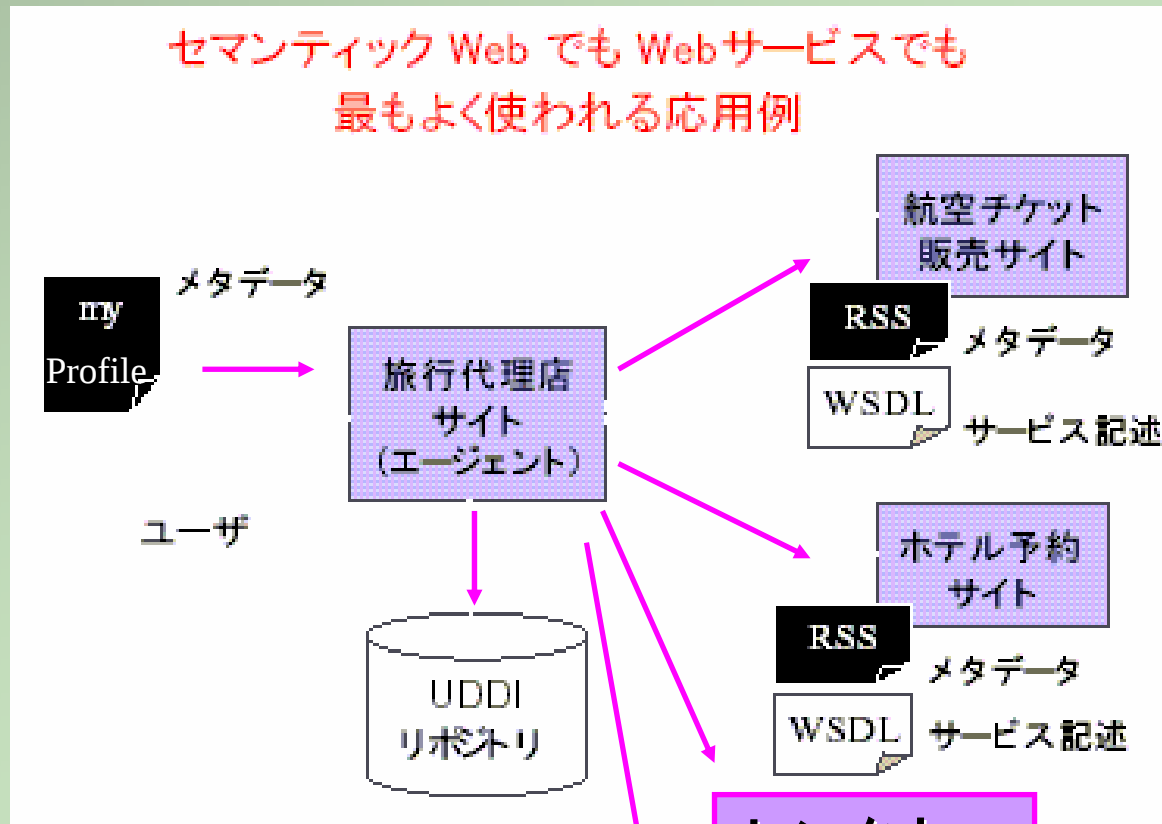
公理(Axioms)の一部

他、プロパティ要素の型を定義する要素 owl:SymmetricProperty etc.

エージェント(swロボット): 非mobileもgood

検索エンジン側の仕事はぐっと楽に？ 豊かに？

日常語"Agent"の例: 旅行代理店



セマンティックWebコンファレンス2002
予稿, Section6 ©INTAPを翻案

これは完成イメージ

実際に作ろうとする
とまだ規格やコンテ
ンツが足りない。。

観光施設#1

観光施設#2

観光施設#n

サブ・エージェンシー

Road Map

1. 今、なぜSemanticWebか？

Semantic Webの実質,活用のステップ

メタデータ, オントロジ (,エージェント)と積み重ね、サービス【内容】中心アーキテクチャ、アプリへ

- メタデータ
- オントロジ
- エージェント
- WS vs. SW 
- Web, EC, KM応用とSemanticWeb
- 試作アプリ1：象徴的アプリ「常識検索」
- 試作アプリ2：RSsS for Controlled XML Summarization framework
- その他、アプリのアイディア達 (2002.8-9頃発案)

WS & SW たかがインフラされどインフラ

- 接続性の拡大、接続コストの低減
- 既存API on WAS (Web Application Server)をWSでラッピング → ASPの低コスト化
- ひいては新奇機能を出現させやすくする触媒
- 日本のASPブーム(2001年前半)で出てきた具体的なASP例: → ASP_DATA2001_05.XLS

WebService = DLLのWeb版(?)

WSDL (Web Services Description Language)

SOAP (Simple Object Access Protocol)

UDDI (Universal Description, Discovery and Information Interface)

WSFL (Workflow)

XLANG (WSDLをビジネスプロセスに変換する言語)

ebXML (EDI標準)

tpaXML (EC標準)

XML Signature

SAML

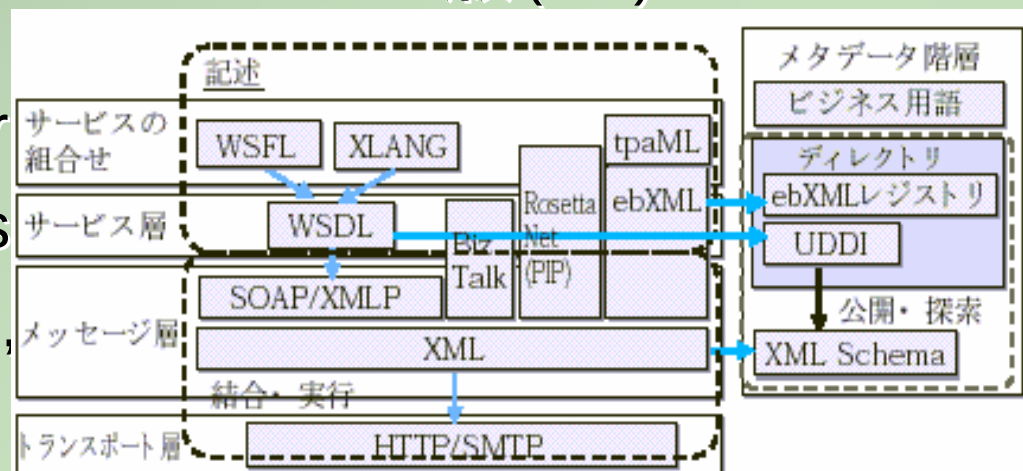


図-5 サービス実行環境のフレームワーク

情報処理2001.9「ソフトウェアサービスへのいざない」
青山幹雄 より引用

DLLとの違い:

Retrieval Agent

Late Binding

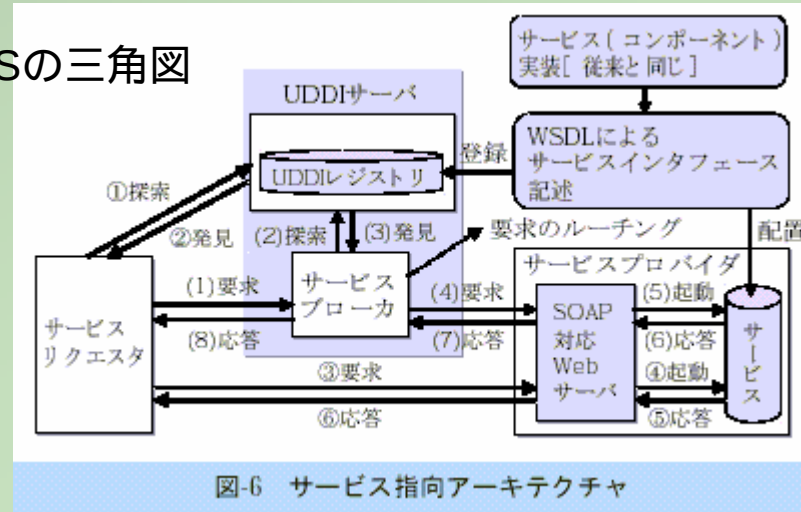
実行保証は「？」

cf. SalCentral : VB プログラム送るとWSDLを返信

cf. XML複合文書 vs. OLE複合文書

新しい定義:「Web サービスとは？」

cf. WSの三角図



- Webサービスとは、
 - URIによって識別され、
 - そのインターフェースおよびバインディング情報は、XMLによって定義・記述され、
 - それらの情報は他システムから検索が可能であり、
 - インターネット・プロトコル上の
 - XML書式のメッセージを主に使用して、他のWebサービスと交信する、
- ソフトウェアシステムである

IBM's SOA: ダイナミックe-businessの原則

http://www-6.ibm.com/jp/developerworks/webservices/010608/j_ws-arc1.pdf

より引用

- B2B統合の複雑さを解決するために必要な条件を明確にする基本原則
 - ソフトウェア・リソース間の統合は、ゆるい結合になっていなければならない。
 - ソフトウェア・リソースのサービス・インターフェースは広く公開し、かつアクセス可能でなければならない。
 - プログラム間メッセージングは、オープン・インターネット規格に準拠していなければならない。
 - アプリケーションの構築は、コア・ビジネス・プロセスを、外部から調達したソフトウェアのコンポーネントおよびリソースとつなぎ合わせることで実現できる。
 - 細分化された多数のソフトウェア・リソースの可用性が増せば、ビジネス・プロセスの柔軟性と個人情報設定が向上するはずである。
 - 再使用可能な、外部から調達したソフトウェアは、お客様にサービスを提供するための費用または生産性(あるいはその両方)の効率化に有効である。
 - ソフトウェアはサービスとして販売できる。

WSとSW: どう、うまく棲み分ける？

どちらも、エージェントや連携・統合を謳うが。

WS = 機能の連携・統合

SW = コンテンツの連携・統合

に加えて、最近、様々な場での議論をきっかけに気づいたのが、XMLの使命2つのうち、各々がどちら寄りであるか、という点:

WS: 「方言」の追求

SW: 「共通語」の追求

ここらを押さえて、ナレッジ・インフラ提供ベンダは、Web標準をうまく活用した商品・サービス展開をしていくべきでしょう。
(N.Nomura)

Web Services本格実用化にあたっての課題をクリアするためにSemanticWebが使える？

- 数ある機能部品候補とその組合せの適否をどう採点？
→ コレオグラフィ(choreography;振り付け)
 - グラフ理論やAIアルゴリズムが必要では？
- だれがどう品質保証するか？
 - Web Servicesインフラの品質保証, Fail Safe機構の導入
 - 機能モジュールの品質
 - 仕様自体の適切さ、豊かさ(品質)
 - 継続的運営の保証(与信管理) → セキュリティ関連の規格群 + 運用規則の社会的合意
 - 悪意、騙しのないWeb Servicesである保証
 - ソーシャルフィルタリングや、信用・格付け機関のBlack Listで管理？
 - 加えて、信用リスク代行(エスクロー)、事故後の保証(保険)も必要

WSとSW: W3Cによる「縄張り」確定

WebService Architecture Requirements

<http://www.w3.org/TR/2002/WD-wsa-reqs-20021114>

AC009 should avoid any unnecessary misalignment with the Semantic Web

- AR009.2 new Web services technologies, developed by W3C Web Services WGs, should be capable of being mapped to RDF/XML.
- AR009.3 all conceptual elements should be addressable directly via a URI.
- AR009.4 the WSA must not preclude the characterization of a Web Service that attempts to make its semantics clear to an automatic system using technologies such as those adopted as part of the Semantic Web. (↑ SWのアプリ系の意味空間を記述するWSを前提としちゃ駄目)
- AR009.5 Web service descriptions should be capable of referencing concepts identified by a URI in an ontology, such as W3C OWL.

WS = 機能の連携・統合

SW = コンテンツの連携・統合

(機能の) プロファイル

(データの) メタデータ

Grid, IS and WSへの拙コメント(某経営者さん宛emailより)

WebServicesの規格群が拡散しかけているのは事実です。しかし、例えば昨秋11月にW3Cから出たWebServices Requirementsに関するNoteによれば、SemanticWebと重複する規格、ボキャブラリについては、SemanticWebを採用すべし、などのガイドラインが示されました:

<http://www.w3.org/TR/2002/WD-wsa-reqs-20021114>

(> AC009 should avoid any unnecessary misalignment with the Semantic Web以下)

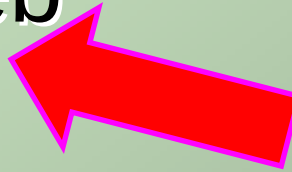
これにより、WebServicesが産業界、商売専用に囲い込まれるのに抑止がかかり、幅広く、リソース、知識活用functionの接続 & 実行手段として、一般Webユーザが汎用的に使える途が残された、と解釈しています。リアルな電子商取引とシームレスにつながる研究プロトタイプが作れる可能性を見出し、とてもエキサイティングな感覚を覚えます。

グラフ理論を駆使して経路の最適化を自動的に行う優れた規格を標準化したり、与信評価・管理、木目細かなセキュリティと利便性を両立させるアイディアを出したりなど、研究要素が多く残されています。かようにWebServicesがTCP/IPやHTML並に広く普及するかは予断を許しません(Google, Amazonの拡張機能で、既に皆様舞台裏で働くWebServicesのお世話になっています)、toy systemに満足しない産業界の研究者、技術者であれば、大いに注目し、活用すべきではないかと思います。

野村直之@法政大EC 2003.02

Road Map

1. 今、なぜSemanticWebか？ メタデータ、オントロジへと積み重ね、サービス【内容】中心アーキテクチャ、アプリへ
2. WS vs. SW ; Semantic Webの実質、活用のステップ
3. Web, EC, KM応用とSemanticWeb
4. 試作アプリ1：象徴的アプリ「常識検索」
5. 試作アプリ2：RSsS for Controlled XML Summarization framework
6. その他、アプリのアイディア達 (2002.8-9頃発案)



EC, Total SCMにおける現場情報サマライズの必要性 ～EAIを超えて社会・経済的コンテンツ連動へ

Total SCMにおけるモノと情報の流れ

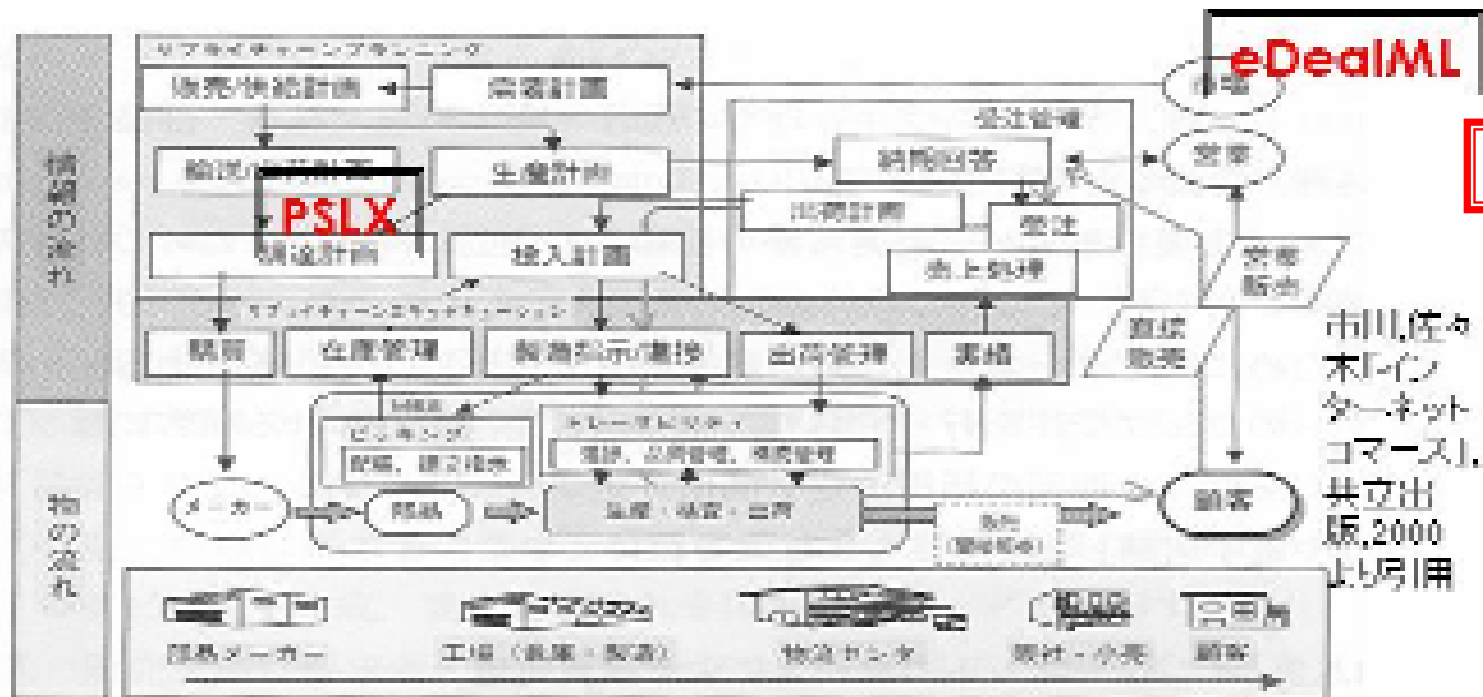


図4.22 サプライチェーンシステムの構成

すべてを統一的に扱うにはXMLが有効そう

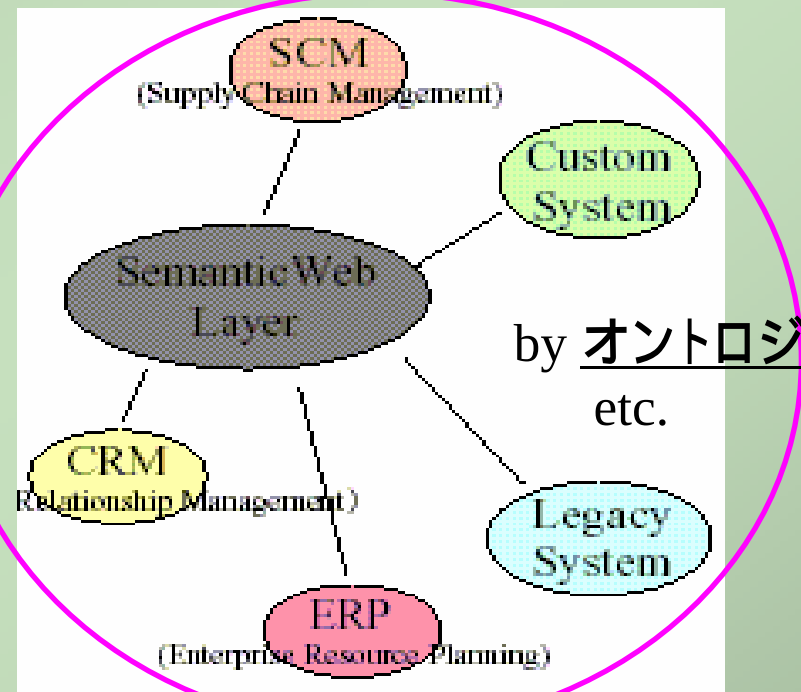
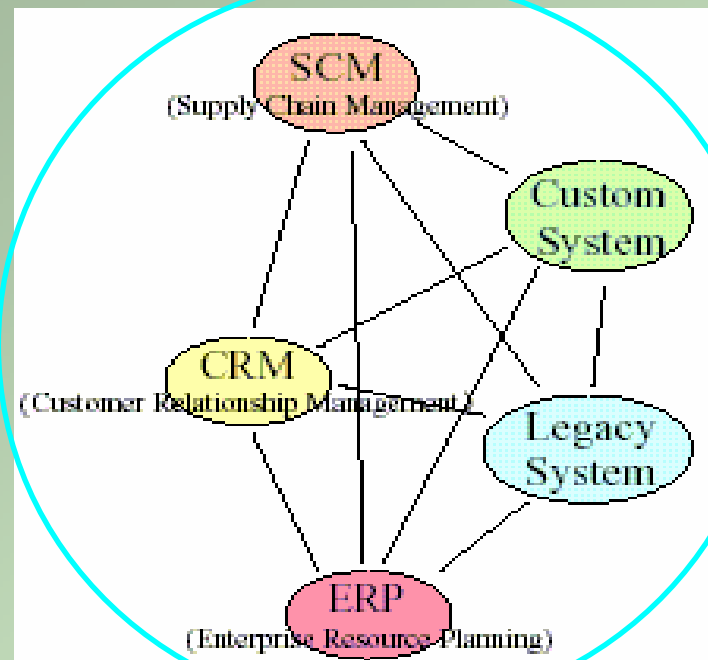
情報、状態遷移の表現、モノの移動・加工のtrigger/trackingもXMLでOK; 組替え、線引きも容易;

EAI とオントロジ変換

EAIの“中間言語”としてのSemanticWeb

3. Web, EC, KM応用とSemanticWeb

セマンティックWebコンファレンス2002 予稿, Section7 ©INTAP を翻案



- mCn 通りの別種インタフェースが必要！
- プログラミング・インタフェース etc. の「意味」は人手で処理: コスト高!

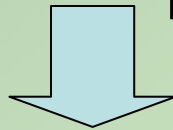
- セマンティックWebレイヤでの統合
- 標準技術 (RDF,RDF-Schema,Dublin Core,...)
- セマンティック・レベル(schema整合)でのプログラミング: データ・タイプ、インスタンス、タスク
- 拡張性(N×1) ●コスト ●メンテナンス

SW:意外に「遅い」現実のEAIの大規模化のために

EAI で、データ量、トランザクション量が増大

現状： 数TPS ～ 100数10TPS

近い将来： 1 ～ 3桁増大が予想される



「データの中身の全てを真面目に処理する必要有？」

「'自分'でなく他の'EAIエージェント'に'横流し'、分担
できないか？」

一部はメクラ判？ (笑)

cf. IBM's "Autonomic Computing" のKnowledgeを
SemanticWeb形式で記述すれば制御系もオープンに。

知識カタログ(リポジトリ; メタ情報)とXML

2つの「意味」(Semantics) :

- 構造で決まる「意味」 Structured IT
- &
- インベントリで決まる「意味」 RDF, URI

e.g. 10進図書分類, リソース間の相対的定義
どちらも XML がベース!

電子図書館コミュニティ制定のDublin Core (5W1Hを拡張した15の書誌情報内容)
<dc: ...> : 名前空間

時間的・空間的範囲 Coverage
内容記述 Description
関係 Relation
出処 Source
主題 Subject
タイトル Title

資源タイプ Type
想定利用者 Audience
作成者 Creator
寄与者 Contributor
公開者 Publisher
権利管理 Rights

日付 Date
形式 Format
識別子 Identifier
言語 Language

KM & XML

‘知’の定義: Minskyは匙投げた! 1992

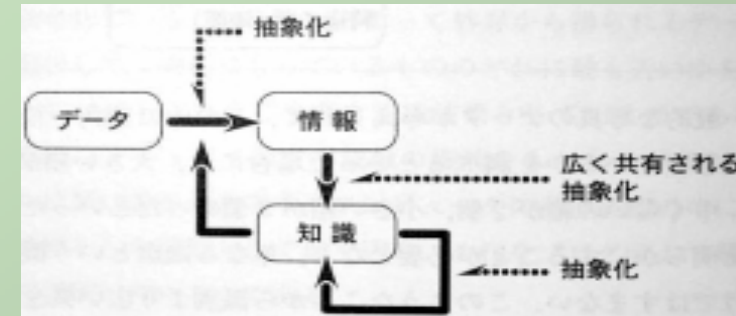
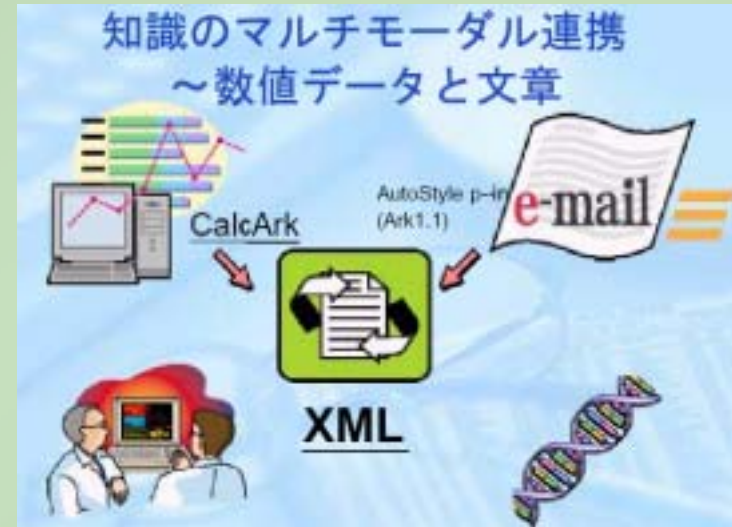
データ : ナレッジ : ノウハウ

- データ:
- “1”, “3”, “4”
“6”, “8”,
“10”, “12”
- 器具で自動生成

- ナレッジ:
- 構造化され、意味付けされた情報 e.g. Web直販戦略レポート12章
- ヒトの頭脳の中で創造

- ノウハウ:
- If 顧客が意味ありげな苦笑2回
Then 回れ右
- 動物集団も得意 Expert_Bottleneck

[野村2001.11] @
知識科学Symposium



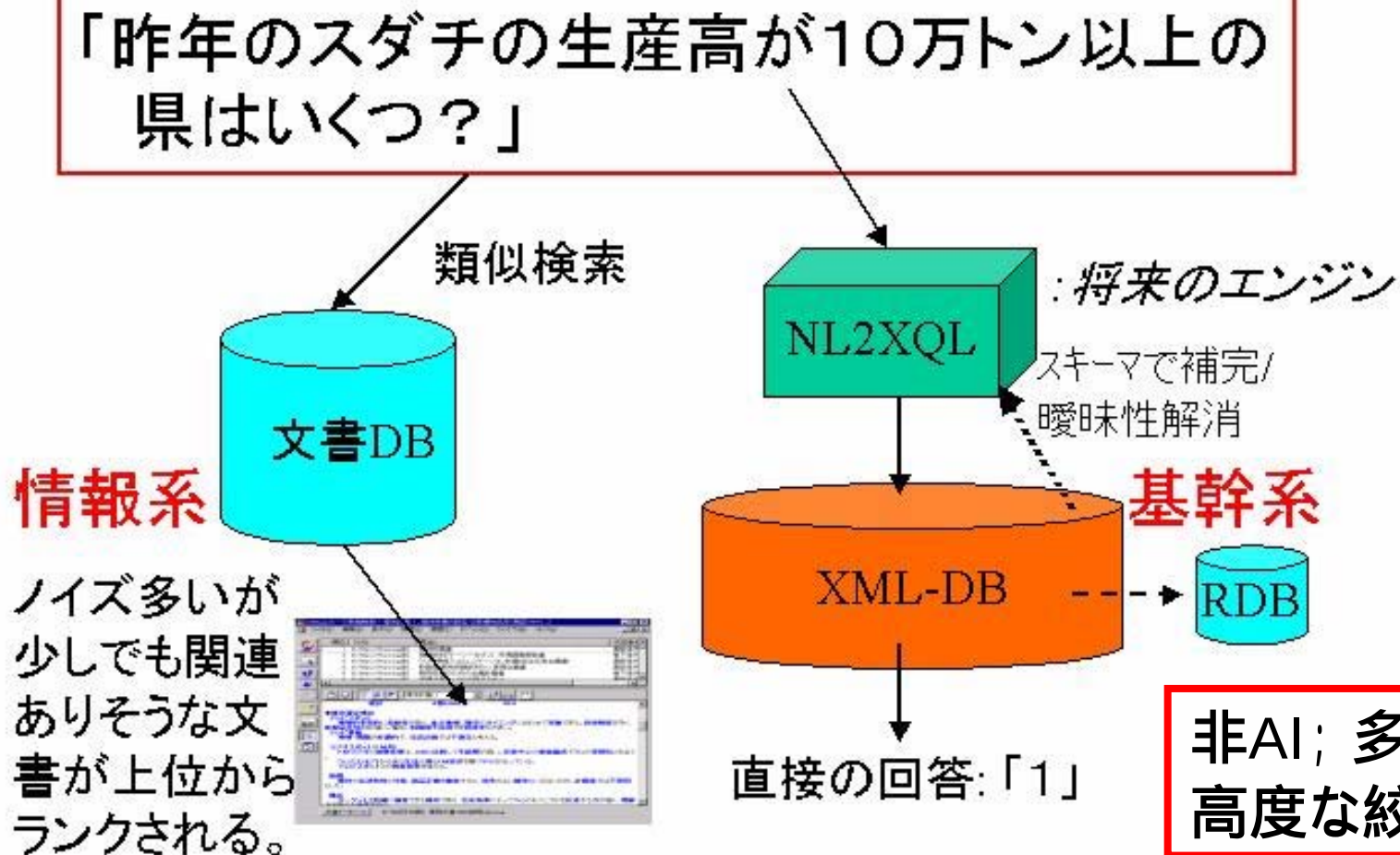
[長尾99]

知識 /≠ 文書ファイル むしろ、
知識 = プログラムとその関連リソース

KM by XML : 情報系と基幹系の統合

(文書系) (勘定系)

類似検索と構造検索の併用による次世代KM

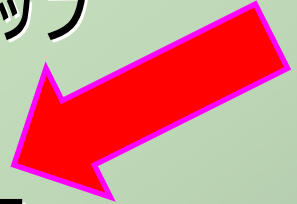


サービス業のKM: 知識創造・運用の具体的な手順を考察

- [野村2003]のメタファ:
文書メタデータ → パケットのヘッダ
類似検索だけでは適合率低いので、
 - ・コンテンツの構造化 (by modular xMLs)
 - ・メタデータ介して異種メディア間の
マルチモーダル連携
 - ・メタデータと本文の間の連携、融通

Road Map

1. **今、なぜSemanticWebか？** メタデータ、オントロジへと積み重ね、サービス【内容】中心アーキテクチャ、アプリへ
2. WS vs. SW ; Semantic Webの実質、活用のステップ
3. Web, EC, KM応用とSemanticWeb
4. **試作アプリ1：象徴的アプリ「常識検索」**
5. 試作アプリ2：RSsS for Controlled XML Summarization framework
6. その他、アプリのアイディア達 (2002.8-9頃発案)



4. 試作アプリ1：常識検索の夢とSemanticWeb @XML Consortium

「常識を備えたSemanticWebのエージェントに検索させる
～「若い奴は演歌は聴かない」って知ってる検索エンジンの試作イメージ」

2002. 6 XML Consortium Weekより

講演者：大泉(JIEC), 野村(法政大)

1. 試作アプリ1：象徴的アプリ「常識検索」

趣旨： 機械間の「理解」のフレームワーク、SemanticWeb
にはどんなキラー・アプリが考えられだろうか。1つの象徴は
「エージェント」技術といわれる。ナレッジWGでは、RDF、DAML
の一部を使って、いわゆる常識知識を複数種類に分け、
各々について、自然な構造表現を施し、いろいろな場所に
オントロジとして配備し、利用者のプロフィールをもとにオン
トロジを調べ、「音楽をDLしたいんだけど最近流行ってる曲は？」
という質問に少量の最適な回答ができるような検討、試作を行ってきた。

今回の講演では、「若者」「中年」といった年齢に関する常識と、
歌の好みに関する常識とを実装した試作システムのデモを行い、
RDF、DAMLを活かす方策を実感して頂くことを目指す。

「常識検索」試作イメージ (SemanticWeb部分の試作ソースのご紹介)

「常識を備えたSemanticWebのエージェントに検索させてみる」

『若い奴は演歌は聴かない』って知ってる検索エンジンの試作イメージ」

http://www.xmlconsortium.org/seminar/w01/prog_4.html

異種の“知識”：分散管理され、権利関係も異なる：

- ・「若さ」：年齢層区分の常識知識 → 検索者との相対関係？
- ・年齢層ごとの好みの音楽ジャンル → 有料調査データをWSで
- ・推薦に際して個人プロフィールや、問い合わせのコンテキスト (e.g. 異性へのプレゼントの楽曲選び?)、目的、一緒に揃えるモノ、サービス(自作の写真etc.) などとメタデータを摺り合せ、適合度を計算。バランスにも配慮。

「常識検索」デモの主要リソース

インターフェース : search.html

オントロジ検索エンジン: rdfsearch.xsl ← XSLT e.g. xalan

Web検索エンジン : search.xsl ←

個人情報ファイル : personinfo.xml

オントロジ記述ファイル: kensaku.rdf

音楽データファイル: data1 (歌、演歌、氷川きよし、きよしのズンドコ節)

: data2 (歌、ポップス、SMAP、freebird、)

: data3 (歌、演歌、水森かおり、東尋坊)

: data4 (歌、ポップス、宇多田ヒカル、SAKURAドロップス)

: data5 (歌、洋楽、マライアキャリー、恋人たちのクリスマス)

: data6 (歌、童謡、不明、うさぎとかめ)

: data7 (歌、童謡、不明、線路はつづくよどこまでも)

2002.6.12 公開の大泉英之氏作成デモシステム より

構造化類似検索, 常識検索への SemanticWebの貢献

- ✧ スキーマ (DBコンテンツのメタデータ) 記述の標準化
- ✧ スキーマ所在探索 w/ WebService
- ✧ スキーマ取得 (Retrieval by Agent) w/ WebService
- ✧ スキーマ照合 w/ XQuery, オントロジ
 - ✧ スキーマ中の必要箇所の絞り込み
 - ✧ スキーマ構造変換 (意味構造の翻訳)
- ✧ 標準化されたエージェント技術の提供

5. 試作アプリ2: RSsS for Controlled Sites Summarizationの試作 @XML Consortium

- RSSがあまり普及しない。 → 手作業で連動させるのが非現実的だからでは？
- “自動要約” 研究も曲がり角。 → 著者、编者、読者の3者がワガママに恣意的にサマライズしたい／して欲しいのを叶える手段が無い！

RSsS:

“Controlled Sites Summary”生成 ↓ cf. RSSは外付け
by RDFをCSS Classに埋め込み

SVG



- XHTML, イメージ, SVG
- 1つのXHTMLに6個の話題 (6省分)

今週の官庁
経済産業省
外務省
環境省
国土交通省
総務省
財務省

Road Map

1. **今、なぜSemanticWebか？** メタデータ,オントロジへと積み重ね、サービス
【内容】中心アーキテクチャ、アプリへ
2. WS vs. SW ; Semantic Webの実質,活用のステップ
3. Web, EC, KM応用とSemanticWeb
4. 試作アプリ1：象徴的アプリ「常識検索」
5. 試作アプリ2：RSsS for Controlled XML Summarization framework
6. **その他、アプリのアイデア達** (2002.8-9頃発案)



6. 他、アプリのアイデア達 (2002.8-9頃発案) 1/8

- ・ あいまい検索+オントロジ・スイッチの常識検索
私鉄乗車案内1: 「急行」「快速」「特急」「快速急行」..
→ 小さな‘認識世界’をスイッチして便利に。
- ・ 私鉄乗車案内2: 「こんど」 vs 「次」
→ コトバの曖昧さが日常生活で問題となる例
- ・ マーケット情報にgood/badの価値判断付けて提供
(Semantic Web Service)
- ・ 天気情報にgood/badの価値判断付けて
例: 農家向け: 「明日は嬉しいことに雨です」
- ・ 観光情報をSemanticWebで付加価値化して提供
個人プロフィール、過去の行動履歴(福岡2回、大阪3回..)に基づき、
旅窓などよりずっと高度に、精密に、パーソナライズド情報を提供。

2/8 (順不同)

一般精度向上 → [山名2003]

あいまい検索+オントロジ利用の常識検索の続き:

・ SemanticWeb活用でGoogleを超える :

例: Naoyuki Nomuraでは現状、同姓同名、別ペルソナ (e.g. ホルン奏者, DAT-heads) で大半が検索ノイズが多い。

→ 検索者の作業コンテキスト (e.g. DeskTop上のバイリンガルhtml文書) から漢字列を推定: 「尚幸」でなく「直之」

ペルソナを特定するRDFを併用して適合率向上。

例: NLP研究者が仕事中に検索している時に、作曲家「のむちゃんの頁」は欲しくない。見たくもない。

- ・ 他、メタデータと、現在時刻・場所を照合して絞込み。
- ・ ‘常識’の活用: 固定的DB: “最近”; 流動的DB: “好みの曲”

3/8 (順不同)

- ・メタデータによる絞り込みをデフォルトで自動実行
5 W 1 H (Boolean), リンク指定 AND 本文内容情報 (体感精度) 適合率向上
 - ・メタデータの信頼性チェックを自動実行
RDF のデータはリソースは実在している、という前提だが。
 - ・高速道路の渋滞情報:
無関係な道路の交通ニュースは聞きたくない;
現在地 & 目的地 経路候補のみの渋滞情報;
- 実現方法の検討: カーナビ内部構造情報の普遍化:
- 点 と 線 (SVGの<path>要素), 距離情報, text情報(caption)
→ いずれも双方向で使えるように、どうSemantic Webの規格で表現するか?
 - テンポラリな情報 (工事・占有許可 / 事故 / 個室トイレの空き状況etc.) をどうoverlapさせるか?

4/8 (順不同)

・組込XML とSemanticWeb:

- 装置の障害情報(log)をSemanticWeb形式で提供できないか?
 - イベントや部品間の関係をRDF[-S]で表現
 - 因果関係(現象と原因?) → オントロジ DAML?
 - コトバと数値の対応付け: 例: 計算式も'リソース'
 - 原因解析・対処用の事例ベースエンジン(異種DB)へexport
- cf 定性推論、パタン学習 → 計算式を分離することで装置の開発工数削減できそう。 = SemanticWebのメリット
- SemanticWeb形式の障害ログ解析ツールで計算量低減。
シスコのIPv6ルータみたいな、SemanticWebのRDF
「アドレス変換」(リソースたぐり寄せ)をハードウェア化?

5/8 (順不同)

SemanticWebは、「機械(同志)が意味を理解する」規格だから、組込系と相性悪いはずないのだ。(笑)

AI, 翻訳支援システム(例: 語彙の分類自体が人によって / 文脈によって違う) → 今のとんちんかんな訳語を産出しまくる機械翻訳エンジンをSemanticWebのオントロジで改善?

- 文脈選択の適合率(精度)向上
- 誤解防止、誤解の危険性(隠れた曖昧性)を指摘。

マーケット情報の解釈オントロジ: ファンドマネージャがどんなニュース見て、銘柄推奨を決めているか? 「ファンダメンタルズ」, 「材料」, 「テクニカル」, 「過去の履歴・経験知識(株価予測の『格言』をSemanticWebで?)」

6/8 (順不同)

メタデータの共通化によりコンテンツ再利用を促進。

デジカメのメタデータで、現地の天候やイベント情報を自動検索

- 赤っぽい文書: 類似画像検索, マルチモーダル検索
- 画像情報のメタデータ利用検索
- RGB, CMKY とコトバ(主観語)との対応付けをオントロジで? cf. MPEG7 画像中の背景と主題
- パタン認識対象の実体のサイズ、推定属性から配送コストを算出 (2002.5 by nomuran)
- **クエリ(検索式)だけの転送による超高压縮率伝送**
hyper-compression

7/8 (順不同)

- 美術館での検索 感性検索
- 医療応用: 緑内障の視野狭窄の度合いを、どう構造表現するか? 色の変換、評価方法によってかわる。評価基準の統一が嬉しい。
- 「音のカタログ」に感性語、個人語(青春の思い出etc.)のメタデータを付加 ~こんな気分の時何を聴く?
- 静脈、網膜、虹彩、指紋のパタンの表現の統一
- 住基ネットをSemanticWebで? (爆笑)
 親族関係のRDF-S; 刑務所に入っていると権利関係のフラグは停止するけど義務関係は残る?
- eJapanのプライバシー保護を技術的に徹底するために、SemanticWeb上位層のセキュリティで?

8/8 (順不同)

メタ・データ自動生成 (文書createや受発信時、雛形利用等)
地道なKM (文書管理)、分散拠点で協同開発するためのツール/インフラ
など。仕様開発、マニュアル開発。

メタ・データ管理用のXML-DB

テキスト・インデクシング型(○)、RDB(), ODB(→ ? 不定形スキーマ
を多種扱うなら) : 構造が一定でシンプルなのが良い?

cf. データ・グリッド ; XQueryが違う形のXMLで取り出せるように、スキーマ解釈付き入出力?

汎用アプリの探索・構築

- 創造性指向のKMにSemanticWebがどう使える?
- ホワイトカラー、普通のビジネスマン、、CRM、マーケティング、SFA、、
などの既存企業情報システムをどう強化できるか

結論: 単一の「キラー・アプリ」でなく多数の既存アプリ、
インフラ全体をメタデータにより活性化、強化すべき

メタデータの地道な活用(だけでも)も大いに有望

- 伝統的な文書管理から発展した知識管理・情報共有
- 商用カタログのリポジトリ標準化
- Webの把握 ・Webのほころびの自動改修
- 情報流通、通信・放送におけるメタデータ流通: EPG
- eLearning Contents共有・再利用,配信・管理標準
 - eLearning技術標準化 LOM (Learning Object Metadata)
MetadataのデータモデルとXML Bindingを規程
コンテンツ自体はアナログも、人のパフォーマンスも指定可
 - 学習者属性を記述するメタデータLIP (Learner Information Package)
氏名、ID、学習目的、保有資格、学習履歴、スキル、知識、Passwd
- 草の根の情報発信の活用拡大 (P2P)

cf.ワークショップ^o: NLP and SemanticWeb

2003.04 Budapest

8 out of 18 submissions :

Ait-Mokhtar et al :Linguistic Parsing of Structured Documents--the example lists in HTML

Buitelaar et al A Multi-layered, XML-based Approach to the Integration of Linguistic and Semantic Annotations

Smrz & Povolny Lexical Databases in XML

Carletta et al The NITE Object Model Library for Handling Structured Linguistic Annotation on Multimodal Data Sets

Freese et al Enhancing XCES to xComForT: An Extensible Modular Archive for the Annotation and Manipulation of Text Resources

Boncheva et al SEMANTIC WEB ENABLED, OPEN SOURCE LANGUAGE TECHNOLOGY

Pareja-Lora OntoTag: XML / RDF(S) / OWL Semantic Web Page Annotation in ContentWeb

Reidsma et al: Cross Document Annotation for Multimedia Retrieval

MetaDataを「より必要とする」XML個別言語、表現媒体がある

- 映像・音声は加工しにくい
 - だからメタデータがtext, html以上に重要！
 - MPEG7 開発、制定の動機
- taskツリー、イベントの論理構造 MOPS (R.Schank)
- ユーザ行為にtriggerされるメタデータ自動生成
 - コンテンツの自動解析・パタン認識困難なメディアで重要。
- 認識結果のメタデータをビジュアルに
 - スケルトン、スクリプト
 - SVG, SMIL 編集GUIで操作単位に。

メタデータとコンテンツ(内容自体)間の自動連携の ニーズと利点

- ・ Schemaによるオントロジ抽出部品
eg. 新聞記事から人物・企業プロフィール
- ・ DCの複合文書での巧みな活用
- ・ eLearning
教材とカリキュラムXMLデータの連動
- ・ 比較,決済などの数値処理と文書処理の融合
例：電子カルテMML.0とレセプト処理

野村直之：‘XMLベースの次世代Web・・・を知識管理の現場が’
どう活用すべきか’ 知識科学シンポジウム, 2001.10.28

メタデータの地道な利用を拡大→ 直近の使命ではないか。 メタデータと情報流通 ～ 私見, idea

- SW&WS: SOAPの外側にRDF, RDF-S ?
- SOAPの外側 開封しなくても配信先を、経由者が独自の価値判断等により切り分けられるようにメタデータを置く。分散【非】管理のエージェント cf. Telescript by MagicCapの失敗
- コンテンツ中心アーキテクチャではメタデータの効果的活用が鍵
- メタデータ照合が(必要ならオントロジ変換を経て)、知識流通・解釈、応用処理の中心的役割の1つを担う。