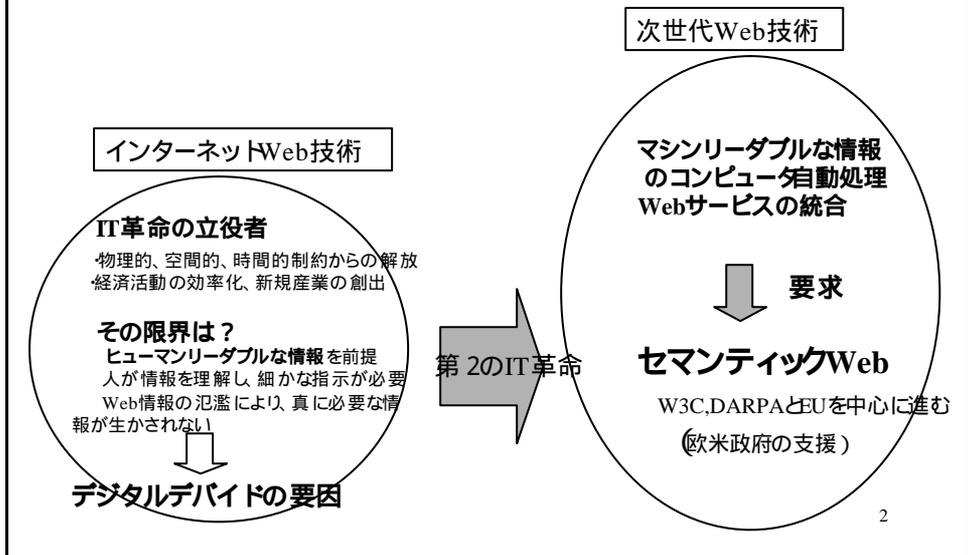


セマンティックWebの概要

平成14年10月30日
財団法人 情報処理相互運用技術協会
セマンティックWeb委員会
委員長 清水 昇
shimizu@intap.or.jp

1

次世代Webを実現するセマンティックWeb技術出現の背景



2

セマンティックWeb^(注1)とは？

あらゆるデータと情報をマシンリーダブルなメタデータ^(注2)でその意味を記述し、人間の代わりにソフトウェア^(注3)で自動処理させること

- (注1) Webを発明したTim Berners-Lee氏により、4年ほど前に提唱されたもの
- (注2) データを記述するデータという意味でメタデータ(超データ)と言う
- (注3) インテリジェントエージェントと言う

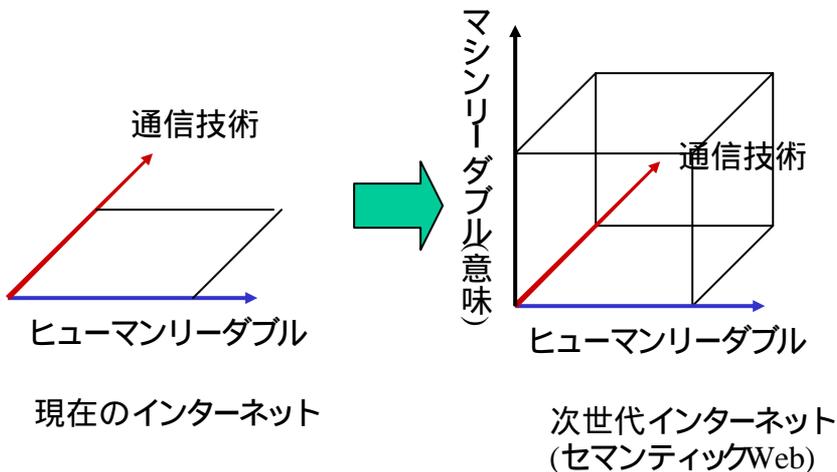
物理的事象も含めた膨大な情報、ハードソフト及び機能等あらゆるものを記述する事が可能となる
人間は簡単な指示のみで、コンピュータが自律的、かつ自動的に処理可能
W3Cが中心になり、欧米で推進、日本は立ち遅れ



経済活動の効率化と、社会や家庭の利便性が飛躍的に向上

3

現在のインターネットとセマンティックWeb



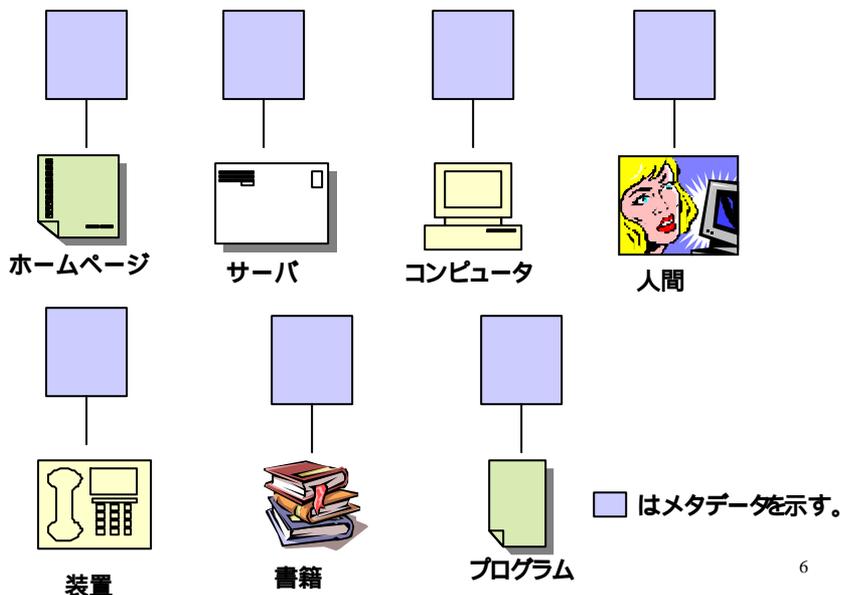
4

メタデータとは

- メタデータとは、データに付けられたデータです。
- 当然、メタデータに付けられたデータも、メタデータです。
- メタデータを付ける対象は、デジタルデータに限定されません、あらゆるものにメタデータをつけることができます。
例、ホームページ、人間、機械、装置、図書館の蔵書、プログラム
- メタデータの実例
フィルタリングにおける有害度を示すラベルデータ、CC/PPの装置プロフィール

5

色々なメタデータ

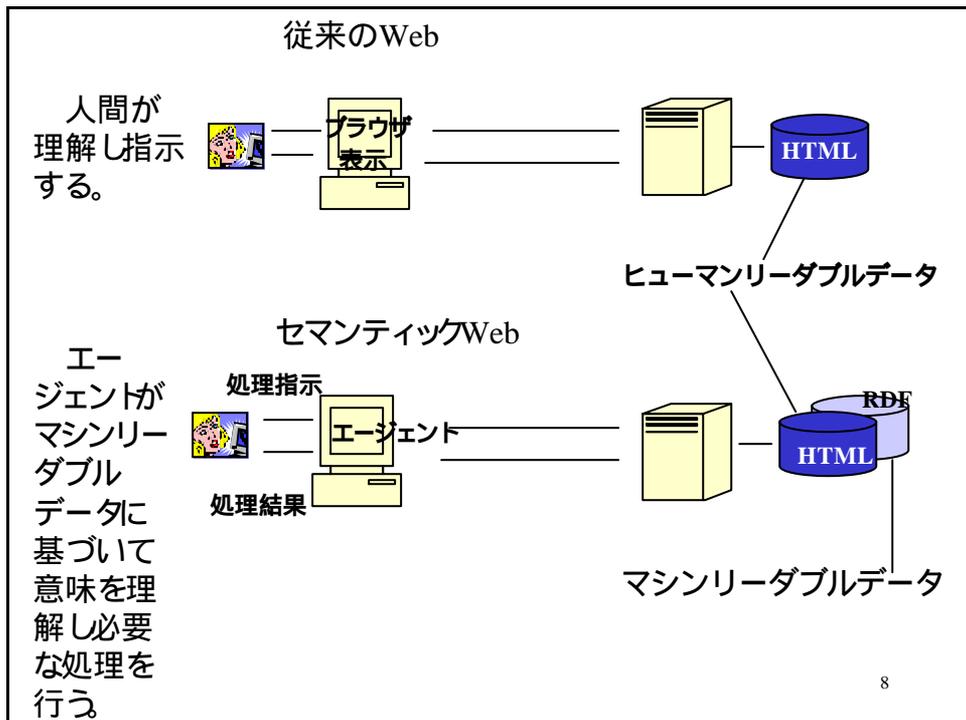


6

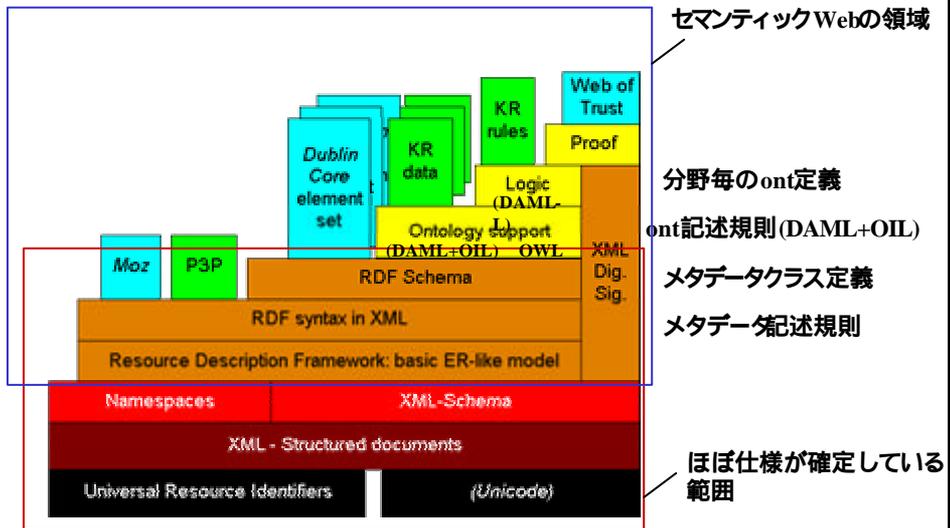
メタデータとセマンティックデータ

- データのデータをメタデータと言う。
- 何らかの意味を示すデータをセマンティックデータと言う。
- セマンティックWebでは、メタデータを用いて基本データのセマンティック(意味)を記述する。
- セマンティックWebでは、メタデータ記述規則としてRDF(資源記述の枠組み)を定めています。

7

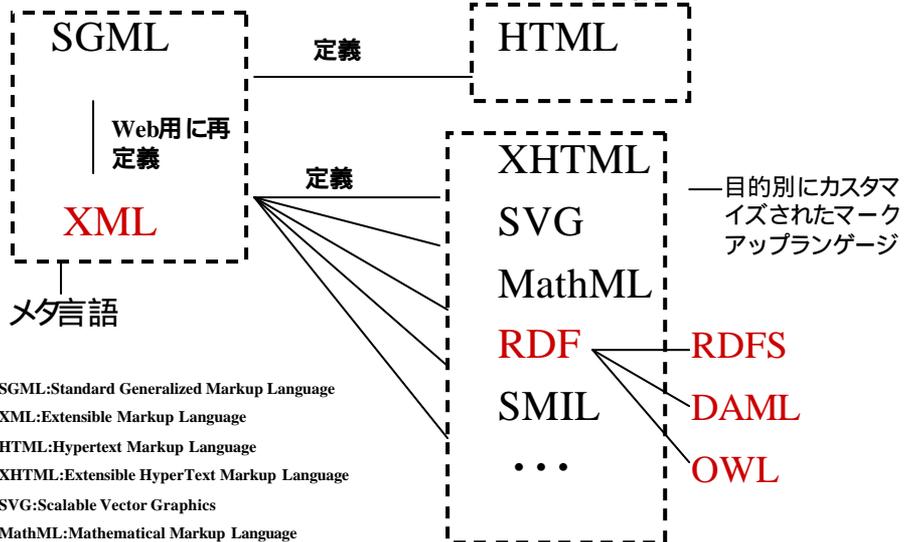


セマンティックWebの階層構造図(by Tim Berners Lee)



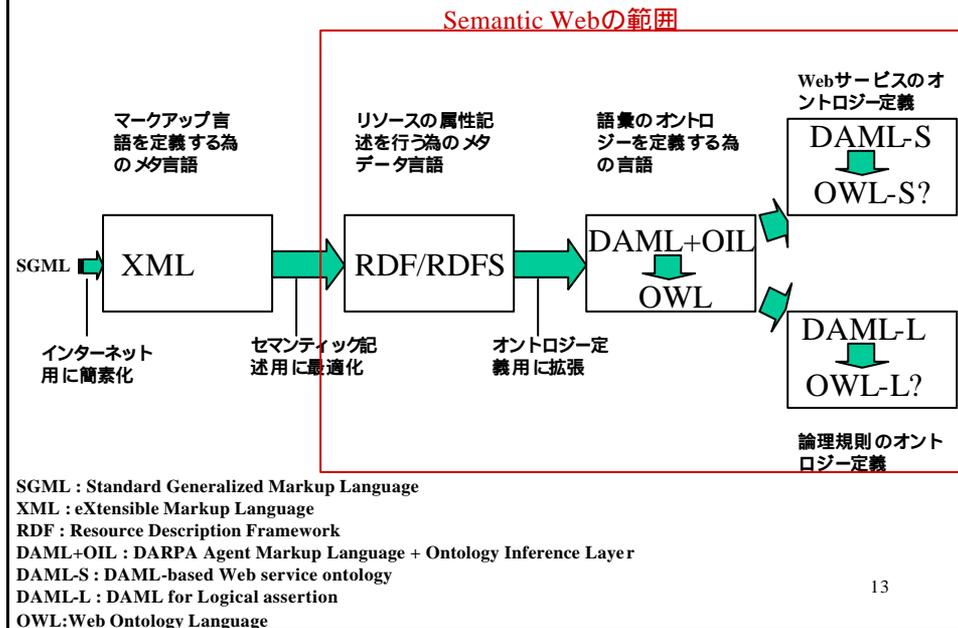
(注)ERモデル=実体関連モデル:実体と実体間の関連とを表現するモデル。 11

RDFとXMLとの関係



SGML:Standard Generalized Markup Language
 XML:Extensible Markup Language
 HTML:Hypertext Markup Language
 XHTML:Extensible HyperText Markup Language
 SVG:Scalable Vector Graphics
 MathML:Mathematical Markup Language
 RDF:Resource Description Framework
 SMIL:Synchronized Multimedia Integration Language

SGML,XML,RDF,DAML+OIL,OWL



XML,RDF,RDFスキーマ,DAML,OWL(1/2)

- XML
言語定義の為の言語
- RDF
属性と関連とを記述するための曖昧さの無い、明快な構文を有する言語
- RDF Schema
異なるコミュニティ間で語彙を共有可能にするため、語彙の意味を定義
- DAML
DARPA開発の基本オントロジー言語

XML,RDF,RDFスキーマ,DAML,OWL(2/2)

- DAML-L

DAML+OIL (or DAML-L)

- DAML-S(DAML for Services)

サービスが如何に動作するか記述する
モデリング オントロジー言語

- OWL

DAMLに代わるセマンティックウェブ用オ
ントロジー言語

15

Semantic Web の狙い

- 1.膨大で、その中の情報を有効に使えなくなっているWebの情報を有効利用可能にする。
- 2.リンク又は単語でのみ関連付けされている情報をメタデータでも関連付け可能にする。
- 3.ソフトウェアで自立的かつ効率的に処理可能にする。
- 4.Web上において、人間が行っている単純で面倒な仕事をソフトに行わせ、人間がより創造的な仕事を行えるようにする。

16

「Webの情報の有効利用」とは

- 現在のWebの検索機能では、全体の25%しか検索できず、検索結果には関係の無い情報が多い、また、Web上の情報量は、6ヶ月で倍になっており、状況は悪化する一途である。



- Webの情報に何が記述されているか示すメタデータを追加し、このメタデータをも用いて最適なデータを検索可能にする。
また、バラバラな情報を統合し新たな情報を得ることも可能にする。

17

「メタデータでも関連付けする」とは

- 現在、Webページの関連性は、リンクポインタと単語とによってのみ示される。



- 関連情報が有ってもリンクされていないと分からない。
- 同じ単語でも意味の異なるものがあり、異なる単語でも意味の同じものがある。

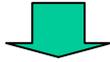


- メタデータとオントロジーとで意味を記述し、それで、関連性を示すことを可能にする。

18

「ソフトウェアで処理可能にする」とは

- HTML及びXMLデータの意味は、曖昧性(自由度)が高く、ソフト処理が難しい。



- 曖昧性の無い、明快な構文のリソース記述言語 RDF を開発し、Webの情報の属性と関連とを記述し、メタデータとして付加。



- 曖昧性が無く、意味が明快な為、ソフトで処理し易い。

19

「人間がより創造的な仕事を行える」とは

- 現在、人間が、Web上から必要な情報を集め、整理し、判断しなければならない。

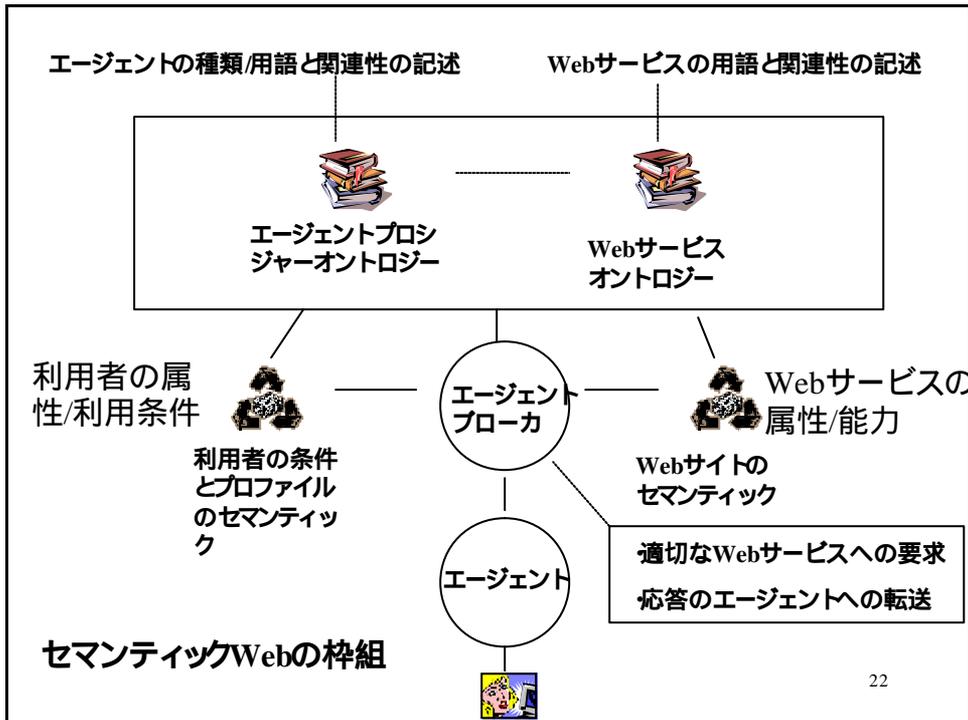
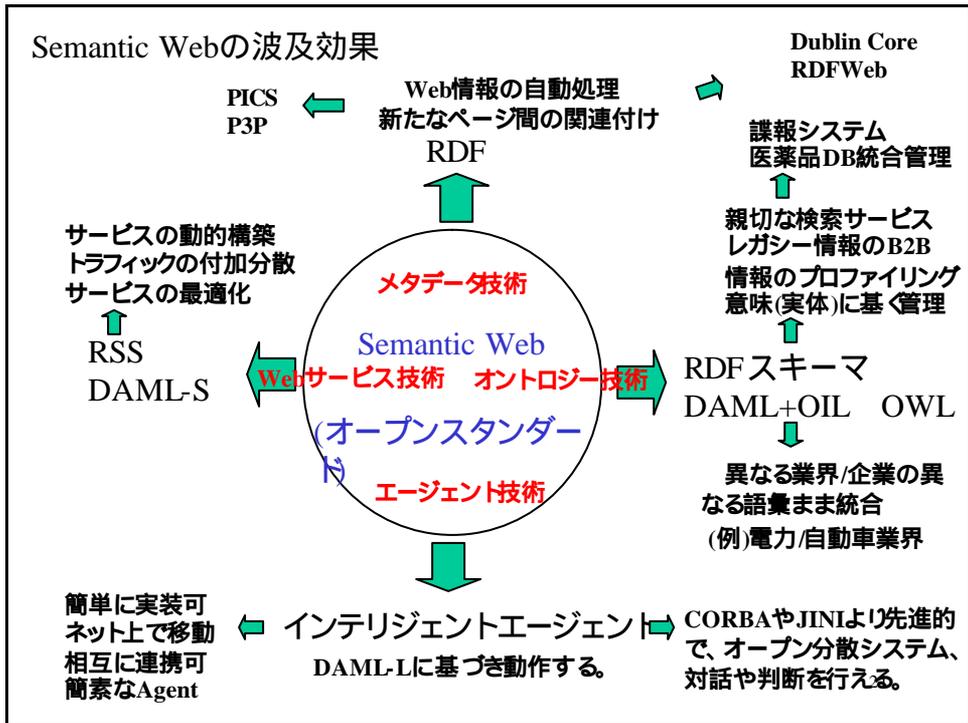


- ソフトウェア(インテリジェントエージェント)が条件に合った情報を集め、整理してくれる。



- 人間は、判断のみを行えば良い。

20



アプリケーションの例(1/2)

1. 利用し易い検索サービス
 - ・日常生活で用いている用語での検索
 - ・「セマンティックWebに詳しい人」と指示すると関連ページから、適任者を見出し表示
2. 携帯端末の能力、通信環境に適したサービス(CC/PPの利用)
能力、通信環境に応じて表示方法やデータの送信方法を変える。
3. TPOに応じたサービス
例えば、地図表示の場合、高速道を走行・・・概略地図、町中を低速で走行・・・詳細地図を表示
4. 次世代マッチングブレース
エージェントが、条件に合致したものを自動検知し、必要な処理を自動的に行う

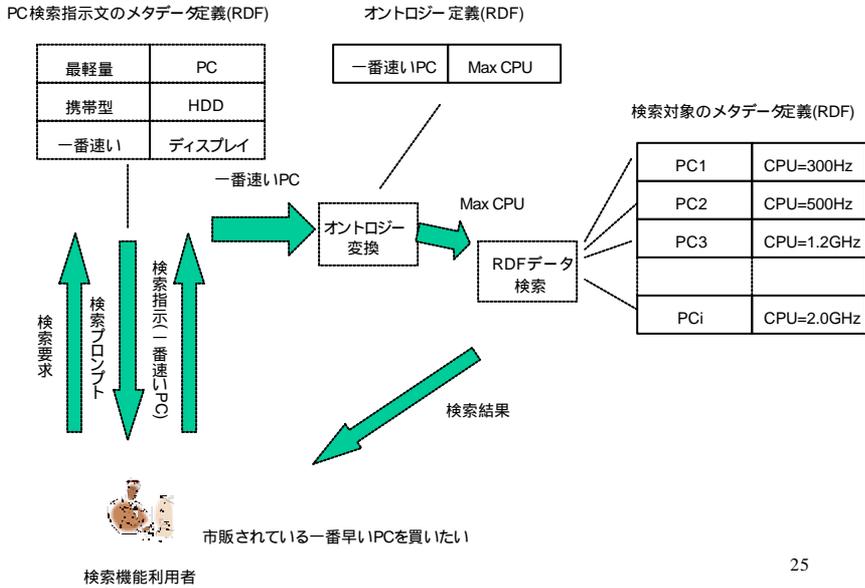
23

アプリケーションの例(2/2)

5. 訪問先の人に付いて知る。(散在情報の統合)
Web上に散在している情報を統合し、その人の興味、業績、経歴、家族構成明らかにする。
6. 景気動向を知る。(迅速、正確かつ総合的な情報処理)
(総ての)企業の決算報告を総合し、日本全体の正確な景気動向を明らかにする。
7. 過去の行動から今後の行動を予測する。(散在情報の統合)
組織(或いは人)に関する情報を統合し、今後の行動を予測する。
8. 個々の利用者に最適化されたサービスを提供する。(最適化)
上記の1.及び3.の情報に基づいて、当該利用者用にカスタマイズしたサービスを提供する。

24

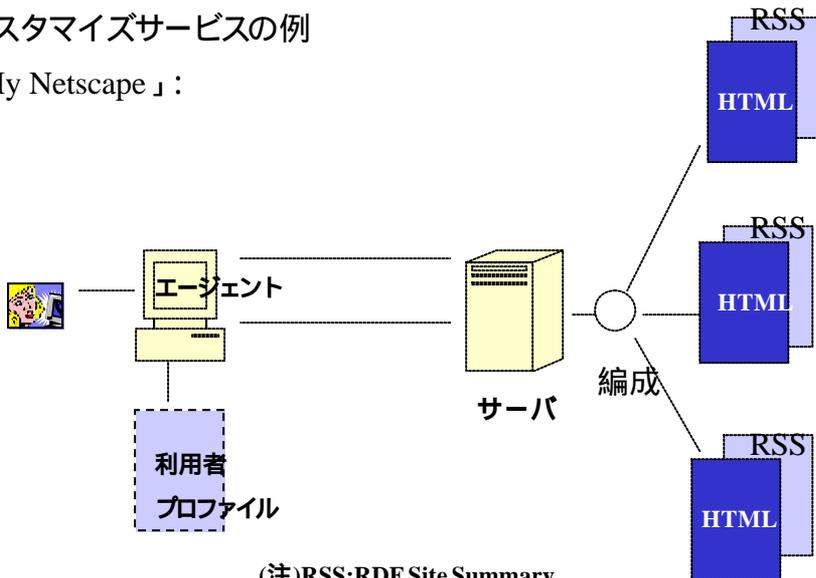
インテリジェント検索処理の流れ



25

カスタマイズサービスの例

「My Netscape」:

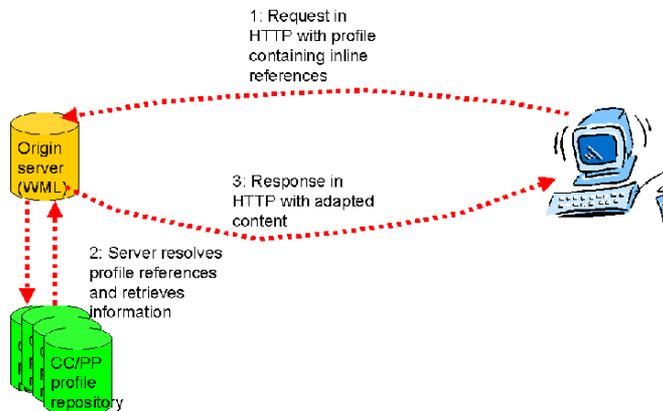


(注)RSS:RDF Site Summary

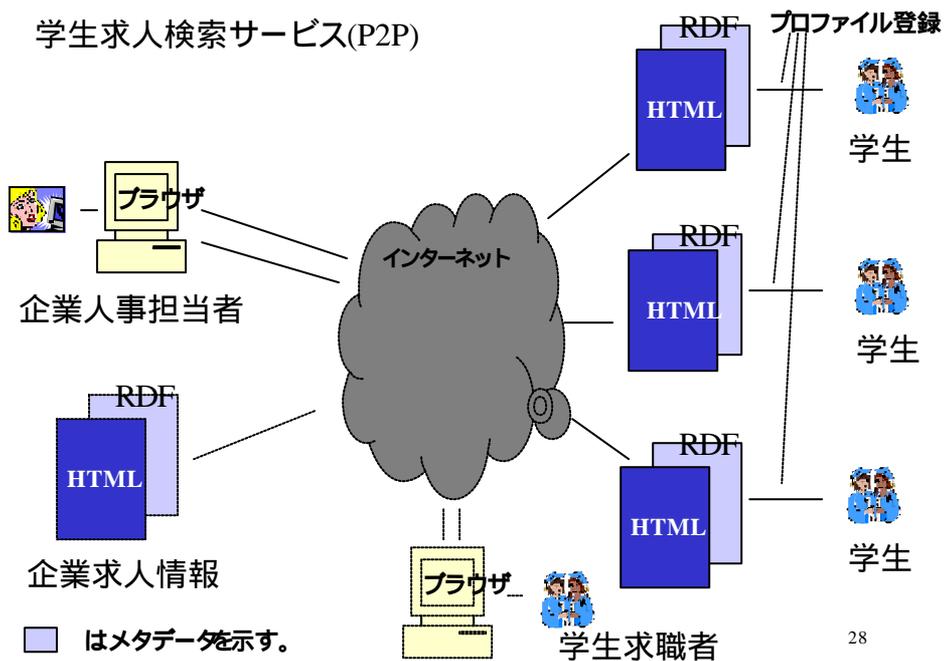
■ はメタデータを示す。

26

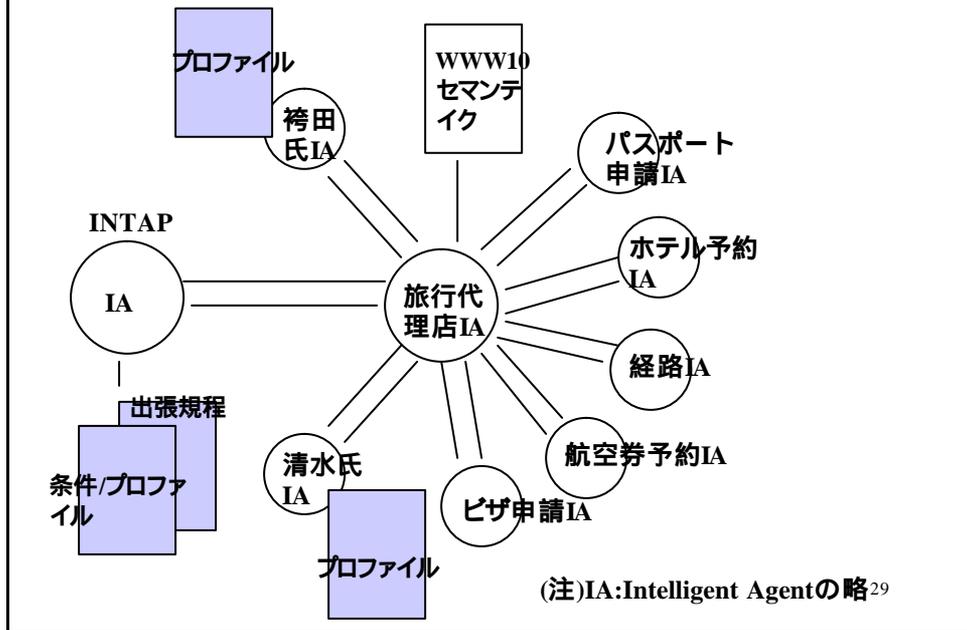
CC/PP (Composite Capabilities/Preference Profiles)



学生求人検索サービス(P2P)



セマンティックWeb調査団の出張手続きを自動処理化した場合



メタデータの活用2(教育分野)

- IEEE LTSC LOM
LTSC(Learning Technologies Standardization Committee)
LOM(Learning Object Metadata)
教材を管、検索、評価、遣り取りする為に必要な情報を定義するメタデータ標準。
非デジタルの教材も対象とする。
- ISO/IEC JTC1 SC36
学習、教育、訓練の為の情報技術の標準化
- ARIADNE
電子教材を作成、管理、再利用する為のEUのプロジェクト
- IMS(Instructional Management Systems)
- ADL (Advanced Distributed Learning)
- DC-E (DCMI-Education)
- EDNA (EDucation Network Australia)
- EUN(EUropeaN Schoolnet)

メタデータの活用3(出版分野)

- DOI(Digital Object Identifiers)
デジタル著作物に識別番号を割り振り、それに基づいて著作権管理を行う
- PRISM(Publishing Requirements for Industry Standard Metadata)
コンテンツの収集、合成、利用者向け加工及び処理を行う為のメタデータ標準

31

メタデータの活用4(行政分野)

- MIREG(Managing Information Resources for e-Government)
- UK-GMF (UK Government Metadata Framework)
e-GMS : e-Government Metadata Standard
- AGLS(the Australian Government Locator Service)
- GoC (Government of Canada)コアメタデータ
- GILS(Government Information Locator Service)
政府情報位置サービス
- NBII (National Biological Information Infrastructure)
国立生物情報基盤
- NEDI (National Environmental Data Index)
国立環境データ索引、
- NARA (National Archives and Records Administration)
米国国立公文書館、NAIL (NARA Archival Information Locator)

32

メタデータの活用5(オーディオビジュアル分野)

- EBU P/META(European Broadcasting Union P/META)
- EBU Panel P/FRA (Future Radio Archives)
- MPEG-7
- MPEG-21

33

SWに関しよくある質問(1/3)

1. セマンティック記述は、XMLを使えば出来るので、RDFは必要ないのでは？
XMLは言語定義の為の言語、メタ言語であり、セマンティック記述を行うことも可能である。しかし、その記述の仕方は複数考えられ、その複数の記述を処理するパーサは複雑になる。当然、複数有るならば、どれか1つに決めようという事になるが、その結果、生み出されたものがRDFである。RDFにより、セマンティックの効率的な記述が可能になると共に、それを処理するパーサも簡単になる。
2. SWはトップダウンアプローチではないですか？
違います。SWはボトムアップアプローチです。セマンティック記述はそれぞれの人が好きに行えば良く、其れがどの位使われるかは、オープンな場での評価による。

34

SWに関しよくある質問(2/3)

3. オントロジーとは何ですか？
語彙の定義です。語彙の指し示す事(例:実体)が何か定義するものです。
[例] 風邪薬 = 総合感冒薬 = 新ルル? A錠
4. SWは画一的なオントロジー定義を押し付けるのではないですか？
違います。それぞれの人それぞれがそれぞれのオントロジー定義を行う事ができます。同じ定義であっても、相反する定義であってもかまいません。また、異なるオントロジー定義を如何に変換するかは、利用者に任せられます。どのオントロジー定義が支配的になるかは、オープンな場での淘汰にまかされます。従来の標準化と異なり、用語を統一しなくても実体と同じか異なるか分かれば良い。
[例] 太郎さんの風邪薬 = 総合感冒薬 = 新ルル? A錠
花子さんの風邪薬 = 解熱鎮痛薬 = ナロンエース

SWに関しよくある質問(3/3)

5. セマンティックWebとAIとはどう異なるのですか？
AIは、閉じた仮想的な世界での自動処理を指向し、其の応用範囲は限られ、特定の人しか関与していないのに対し、セマンティックWebは、現実に存在するオープンなWeb世界での自動処理を対象にし、誰もが自由にあらゆる局面に応用できる。
6. セマンティックWebは4~5年先の技術ではないのでは？
セマンティックWebは、直ぐ使える技術です。
セマンティックWebの活用フェーズには、次の3フェーズがあります。
フェーズ1:メタデータに基づいて処理(例:PICS)
フェーズ2:オントロジを用いた意味処理
フェーズ3:インテリジェントエージェントによる自動処理
フェーズ3に到達するのは、かなり先だと思いますが、フェーズ1は直ぐ活用でき、既に色々な分野で利用されています。

海外の状況(1/3)

- 1.NSF Digital Library Initiative
- 2.Distributed National Electronic Resource(英)
- 3.Global Info(独)
- 4.EU
 - 第5次Framework Programme
 - Delos Network Excellence
- 5.MathNet (独)
- 6.IMS Global Learning Consortium, Inc.
- 7.Federal Geographic Data Committee(FGDC)(米)

37

海外の状況(2/3)

- 8.Visual Resources Association
- 9.Nordic Metadata Project(北欧)
- 10.Renardus Project(EU:蘭)
- 11.Networked Digital Library of Theses and Dissertations(NDLTD)
- 12.CIDOC(博物館での知識表現)
- 13.Publishing Requirements for Industry Standard Metadata(PRISM)
- 14.RDF(Rich) Site Summary(RSS)
- 15.MPEG
 - MPEG4,MPEG7,MPEG21

38

海外の状況(3/3)

- 16.INDECS(知的財産権への応用)
- 17.DOI(科学、技術、医療の書籍への応用)
- 18.Basic Semantic Registry(BSR)(言語概念への応用)
- 19.Government Information Locator Service(GILS)
- 20.Planetary Data System
- 22.IEEE Learning Object Metadata(教育ツールへの応用)
- 23.MARC21(カタログへの応用)
- 24.EPICS Data Dictionary