

第二のIT革命を惹起する
次世代WEB技術研究・開発プログラム
(セマンティックWEB技術を中心にして)

平成13年6月29日
慶應大学SFC研究所/INTAP

INTAPの次世代WEB技術研究活動

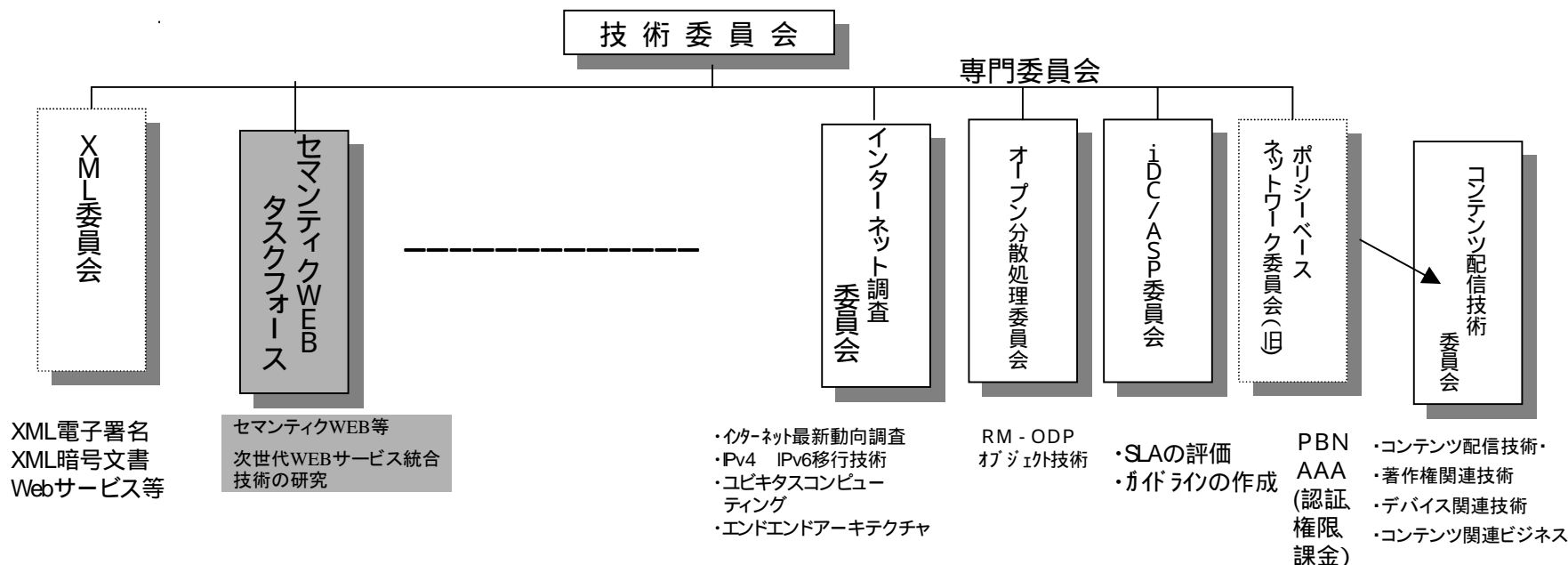
■ WEBサービス技術に関する調査研究:XML委員会

SOAP(プロシージャ呼び出しW3C標準)、UDDI(プロシージャ登録・検索,IBM,MS,Ariba,HP提唱),
E-speak(ネットワークサービス統合環境)等 SOAP:Simple Object Access Protocol,UDDI:Universal Description, Discovery and
Integration)

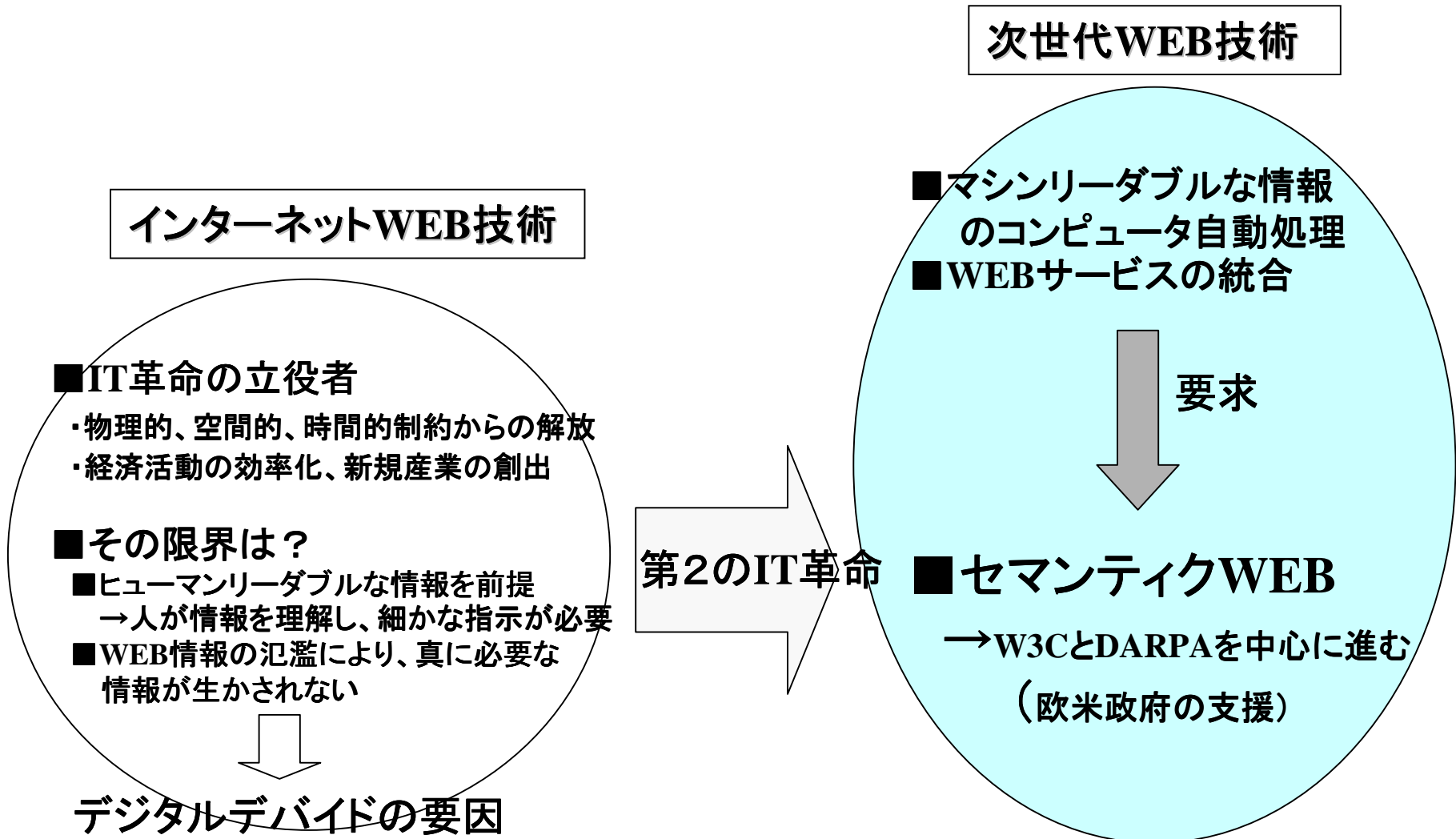
■ セマンティックWEB技術に関する調査研究:セマンティックWEBタスクフォース

セマンティックWEB技術適用の有効性と課題の研究、W3Cホスト組織の慶応大SFC研究所との連携

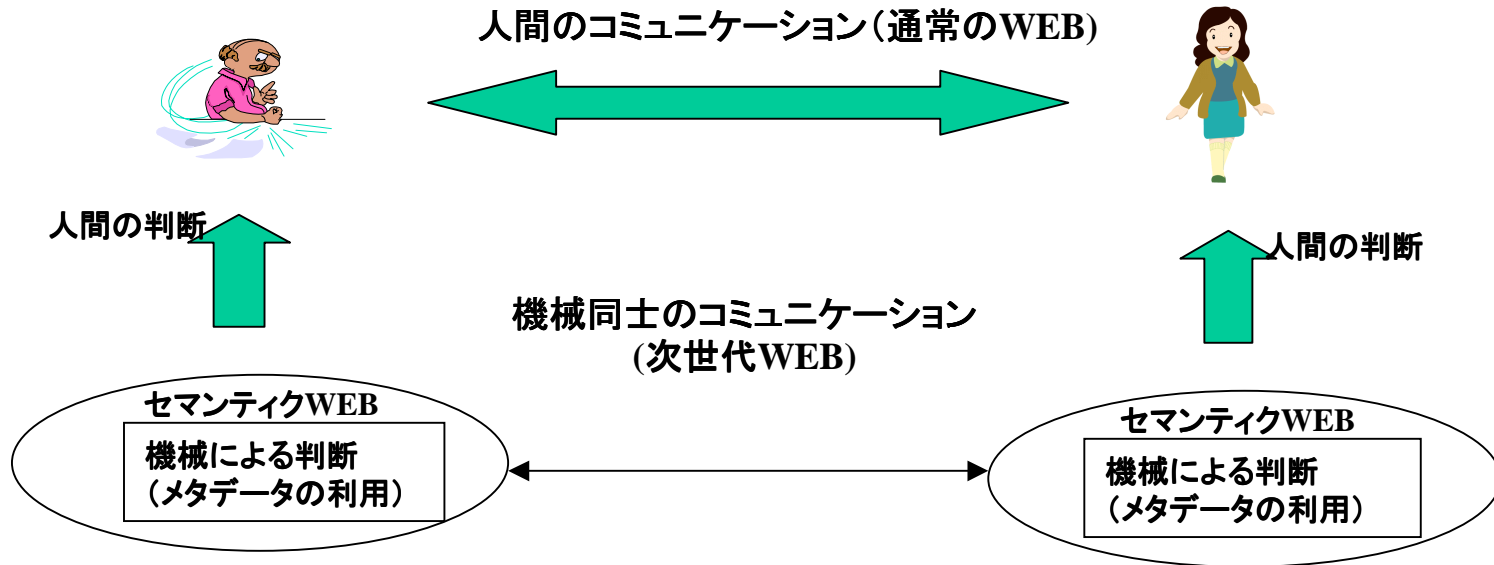
セマンティックWEBタスクフォースの位置付け



次世代WEBを実現するセマンティックWEB 出現の背景

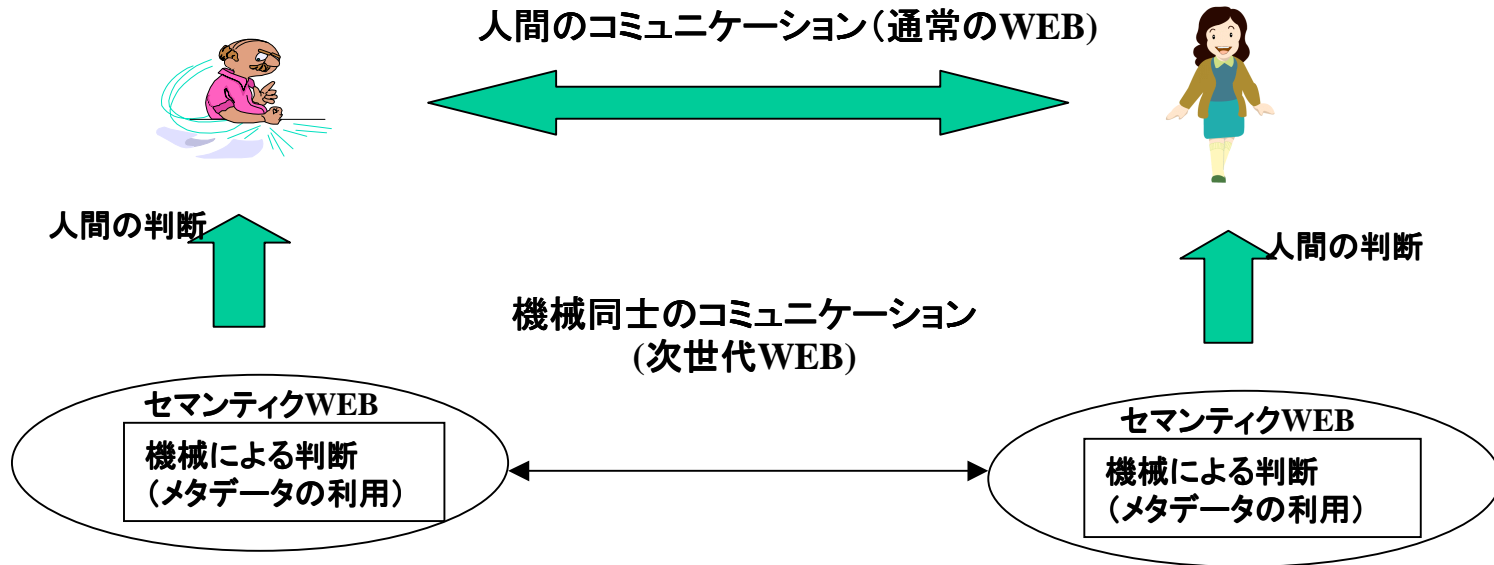


第2のIT革命次世代WEB



- WEB技術はここ10年停止したまま
- 次世代WEB技術胎動のきざし(ワールドワイド)
- バス(次世代WEB)が走り出そうとしている。
- このバスに今乗らないと日本の産業界の痛手となる。
- 評価はバスに乗ってから。

第2のIT革命次世代WEB



- WEB技術はここ10年停止したまま
- 次世代WEB技術胎動のきざし(ワールドワイド)
- バス(次世代WEB)が走り出そうとしている。
- このバスに今乗らないと日本の産業界の痛手となる。
- 評価はバスに乗ってから。

セマンティックWEB***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*でその意味を記述し、人間の代わりにソフトウェア**で自動処理させること

* 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。

** インテリジェントエージェントと言う。

*** Webを発明したTim Berners-Lee氏により、2年ほど前に提唱されたもの

■ 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



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

セマンティックWEBの目標と特徴

セマンティックWEB

→第2のIT革命を惹起する次世代WEB

- 情報の効率的かつ、自動的処理
- シームレスなWEBコンピューティング連携
- 情報の処理プロセスの徹底した標準化と構造化
- 拡張性と柔軟性

- 労働力資源の削減（人手を介さず柔軟にかつ正確に処理）
- 産業競争力の向上
- 画期的効率向上による社会コスト削減

セマンティックWEBの技術開発項目

メタデータ技術

(データ記述の為のデータ)

メタデータ記述の基本枠組み(RDF)

意味及び知識の記述方式と記述規則
(知識記述法、言葉の分類と関連付けのオントロジ)

適用領域毎の標準記述規則

- CC/PP: 装置やソフト等の能力記述
- P3P: 個人情報保護
- PICS: インターネット上のコンテンツ選択
- (将来)→EC用の基準

インテリジェント・エージェント 技術(メタデータを処理)

セマンティックWEB用共通基盤エージェント技術(適用領域に依存しない共通エージェント技術)

- セマンティックWEB用適用領域毎の個別
エージェント技術
- EC用エージェント
 - インテリジェント家電用 エージェント
 - 個人情報保護用 エージェント

ツール技術

- メタデータ記述の基本枠組み(RDF)
- 意味及び知識の記述方式と記述規則
- 適用領域毎の標準記述規則

RDF: Resource Description Framework, CC/PP: Composite Capability/Preference Profiles(装置やソフトウェア等の能力のRDFでの記述)

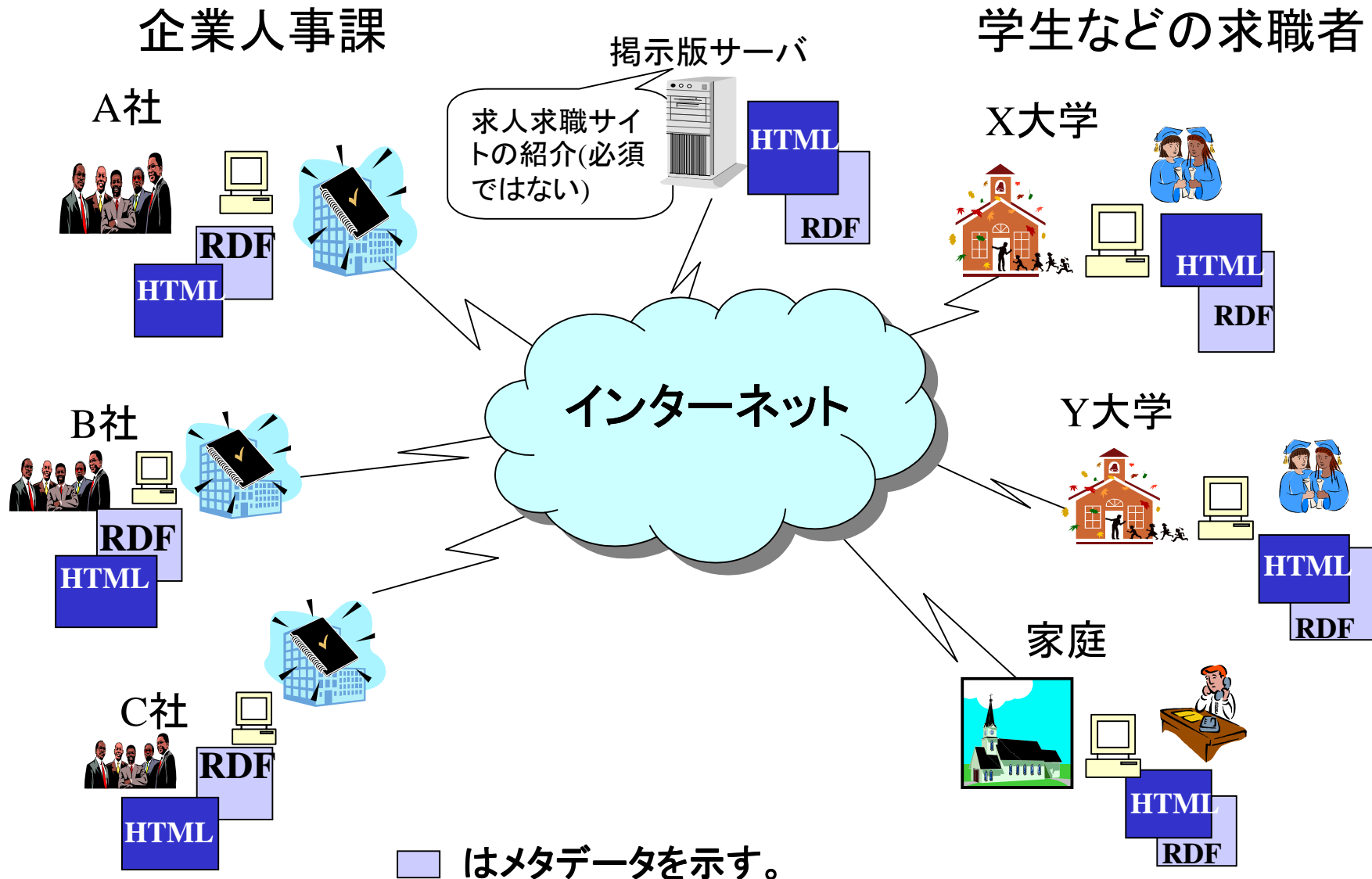
P3P: Platform for Privacy Preference(プライバシー情報提供に関する個人の選択を支援するプラットフォーム),

PICS: Platform for Internet Content Selection(コンテンツ・フィルタリングのためのプラットフォーム)

Copyright © 2001, INTAP, All rights reserved.

応用例：P2P e-マッチング・プレイス（求人、求職）

P2P : Peer to Peer



利用者プロフィールの記述(学生側)

HTML(ヒューマン・リーダブル情報)

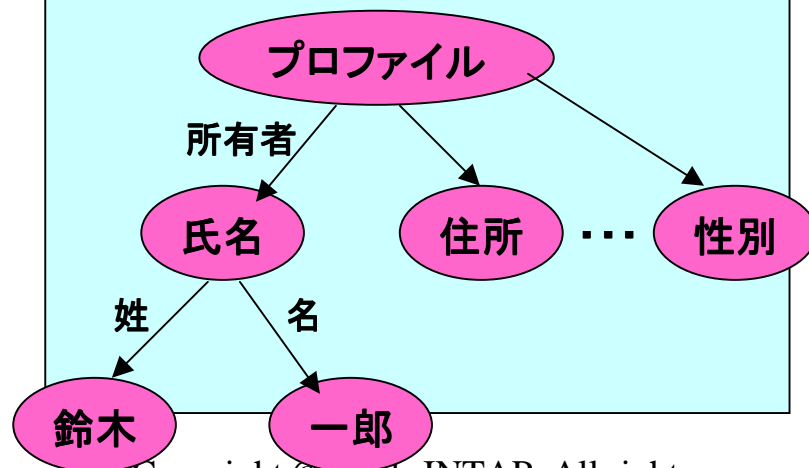
```

<H1>学生プロフィール</H1>
氏名 鈴木 一郎 <BR>
住所 〒220-0112 日本国 東京都...<BR>
生年月日 1980年10月10日<BR>
メール・アドレス、TEL番号
性別
専攻科目
履修科目
論文、分野、キーワード
取得資格
希望の職種
趣味
特技
職歴
アルバイト/ボランティア経験
セミナー受講歴
...
...
  
```

RDF(マシン・リーダブル情報)

```

<プロフィール>
  <氏名>
    <姓> 鈴木 </姓>
    <名> 一郎 </名>
  </氏名>
  <住所>
    <国> 日本国 </国>
    <〒番号> 220-0112 </〒番号>
    ...
  </住所>
  ...
  ...
  </住所>
  
```



利用者プロフィールの記述(企業側)

HTML(ヒューマン・リーダブル情報)

```

<H1>A社ネットワーク事業部プロフィール</H1>
会社名 A社<BR>
住所 東京都千代田区1丁目1番地<BR>
代表取締役社長 小泉太郎<BR>
人事課住所、メール・アドレス、TEL番号
ビジネス種別 ネットワーク・ソリューション提供
求める人材 ネットワーク技術者
                LINUXシステム管理者
期待する技術・知識 ネットワーク管理技術
                    (特にIPv6の知識)
                    OSの知識(特にLINUX)
給与、賞与、年間休日、...
資格 情報処理試験一種、英語検定一級、...
留学制度
以下のアルバイト経験者歓迎
 オフィスのPCやネットワークの管理
 ホームページ、メール・システムの保守
 ウィルス対策などのセキュリティ管理
 ...
 ...
  
```

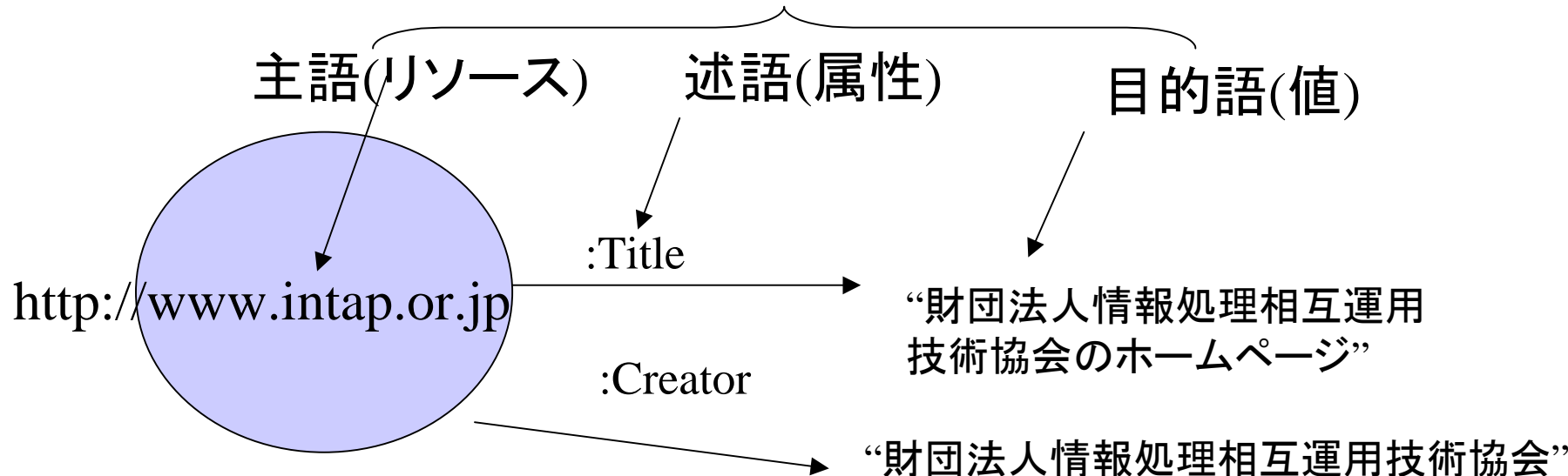
RDF(マシン・リーダブル情報)

```

<プロフィール>
  <会社名> ...
</会社名>
  <住所> ...
  ...
</住所>
  <代表取締役社長>
    <姓> 小泉 </姓>
    <名> 太郎 </名>
  </代表取締役社長>
  <人事課住所>
    ...
    ...
    ...
  
```

RDFの構文例

ステートメント



リソース<http://www.intap.or.jp> の作成者は、財団法人情報処理相互運用技術協会である。`<RDF xmlns = “http://www.w3.org/TR/PR-rdf-syntax#”`

`xmlns:dc = “http://purl.org/dc/elements/1.0/”>`

`<Description about = “http://www.intap.or.jp”>`

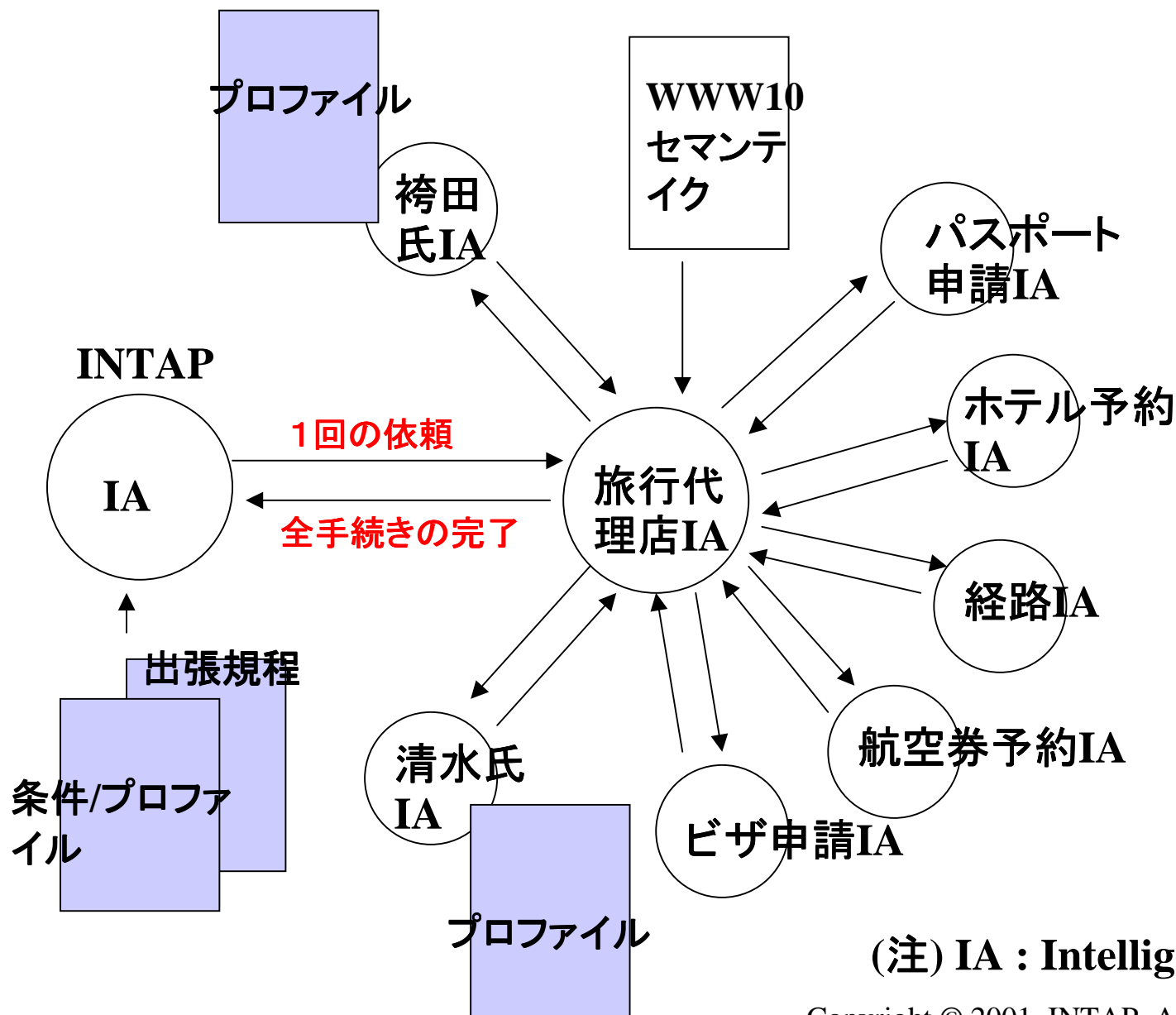
`<dc:Title>(財)情報処理相互運用技術協会のホームページ`
`</dc:Title>`

`<dc:Creator>(財)情報処理相互運用技術協会</dc:Creator>`

`</Description>`

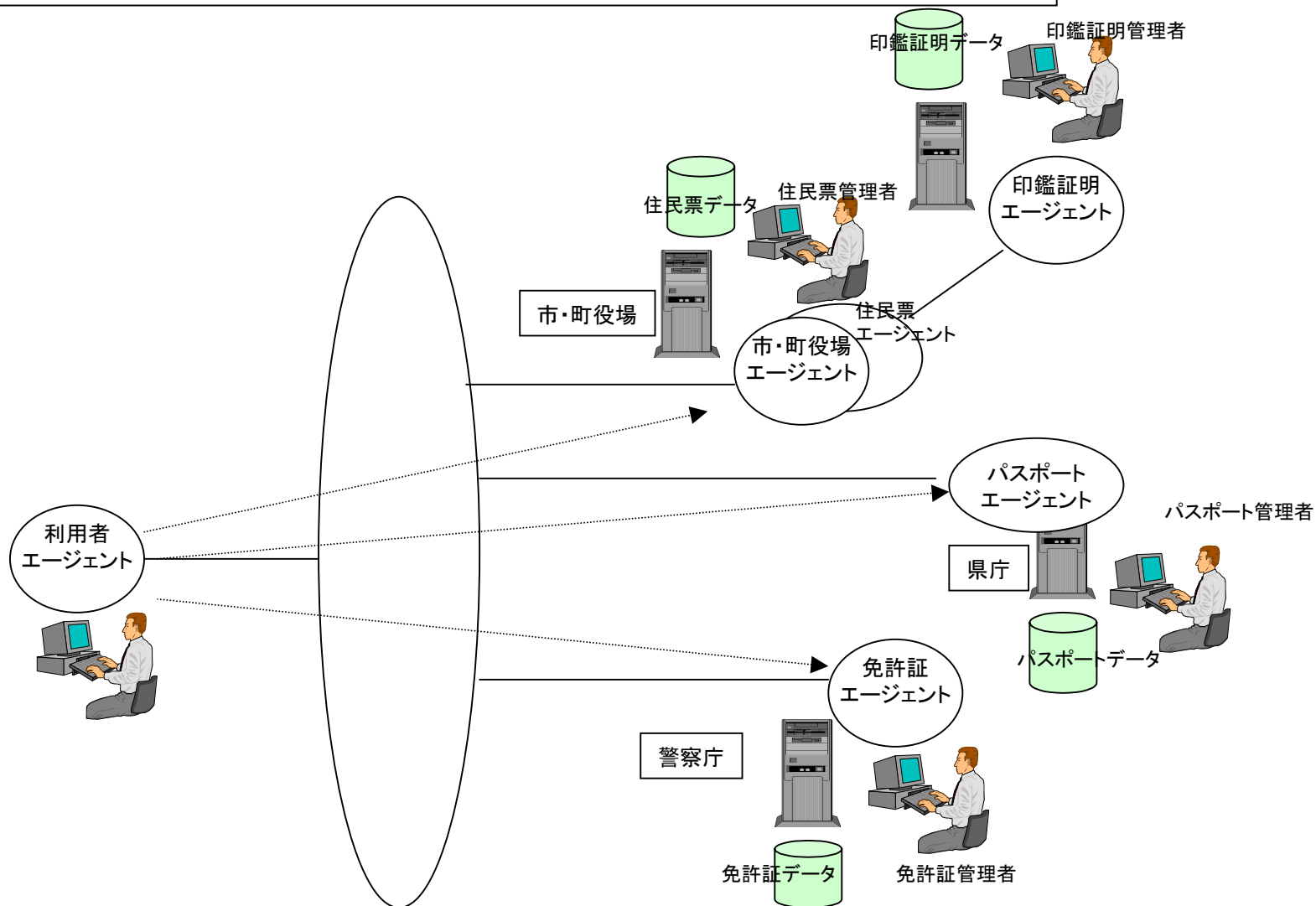
`</RDF>`

セマンティックWeb技術活用の例 (海外出張手配)



電子自治体におけるワン・ストップ・サービス

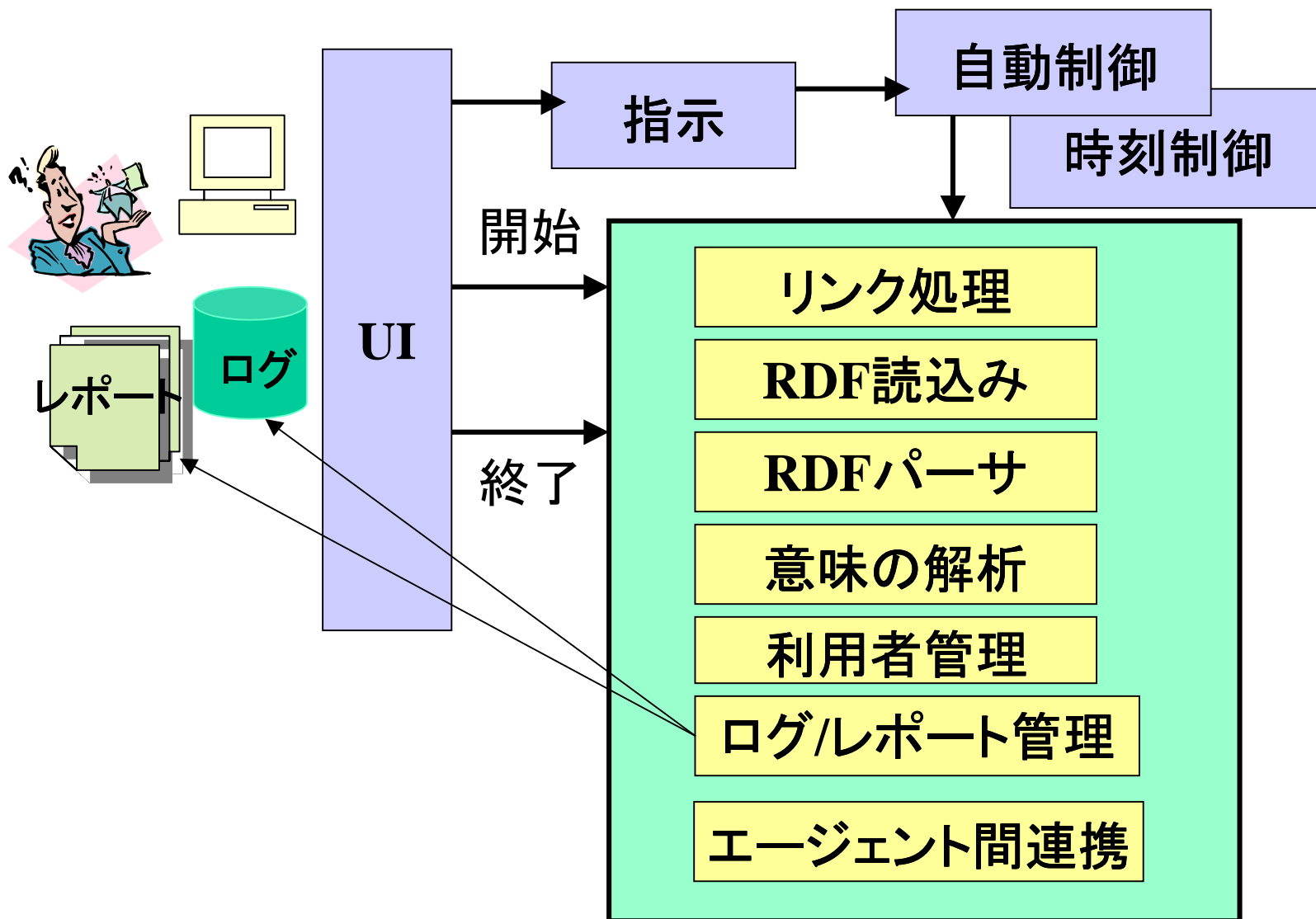
—セマンティックWeb技術を活用した統合サービス例—



メタ・データの定義と保守に必要な機能

- HTMLデータからソフトウェアでRDFメタデータを自動生成する機能
- メタ・データの手動定義(テンプレートの提供)機能
 - ・GUIによるメタ・データの入力。 GUI:グラフィカル・ユーザインタフェース
- HTMLデータに対する注釈定義によるメタデータ生成機能
 - ・例えばW3CのAnnotea。 Annotea:注釈定義ツール
- RDFメタ・データの保守ツール機能
 - ・RDF記述からグラフを表示。
 - ・正しさの検証ツール。
 - ・印刷。
 - ・部分エディタ。

エージェントの構成要素



従来技術との比較

従来

1. DB化による集中管理。
→ DBに載らない人にはチャンスなし。
2. 管理項目にない検索は不可能。
3. 個人プロフィールを預ける。
→ プライバシー情報が漏れる不安あり。
4. 最新情報でないこともある。
5. 特定用途のシステム。

求人側: IPv6に詳しい人が欲しい。

学生1: IPv6ネットワークの構築経験あり。

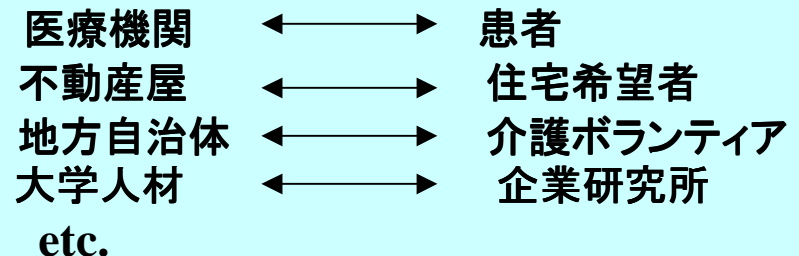
学生2: IPv6の講習会に参加したことあり。

学生3: IPアドレスの枯渇問題の論文を書いた。

エージェントがRDFを解釈して検索し結果を評価 → **学生1、学生3 : Very Good, 学生2: Good**
(他のエージェントに依頼して、論文の記述内容をサーチすることもできる)

今回 (P2P セマンティックWeb技術)

1. DB化は不要。
→ 誰でもすぐ参加できる(**オープン性**)。
2. 柔軟な検索が可能。
3. 個人プロフィールは手元にある。
→ **自分で管理**できるので安心。
4. いつでも最新情報が見られる。
5. 汎用的なシステム(**応用が広い**)。



利点と課題

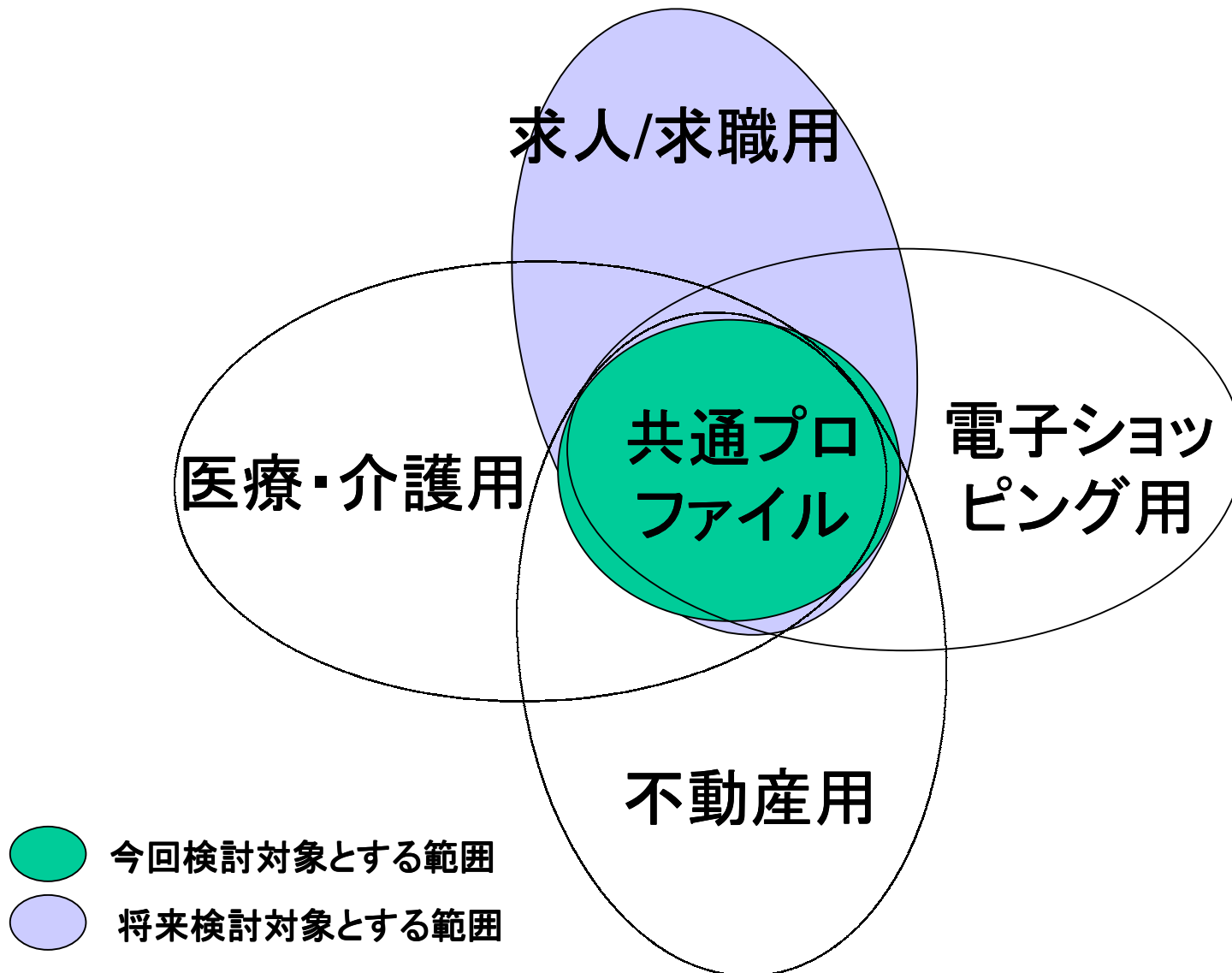
利点

- プロファイル定義の変更や汎用エージェントの組み合わせで多様なアプリケーションに応用することができる。
- メタ・データの持つ意味情報を活用して柔軟な問い合わせができ、情報検索の精度が向上する。
- 人手を介さずコンピュータがマッチする組み合わせを広い範囲から検索するため、作業効率が大幅に向上する。
- ホームページの上に実現されるオープンなシステムであるため、大企業から地方自治体、中小企業経営者や主婦まで誰でも参加することができる。

課題

- **プロファイルの記述法を標準化**する必要がある。インターネットは国際的な広がりを持つので、国際標準化組織への提案が必要。
- 一般の人々が簡単に使えることが大切で、ユーザ・インタフェース技術、エージェント技術、検索・推論技術、用語辞書の整備など、**地道な研究**を継続し、成果をすべての国民が広く利用できるようにする必要がある。
- プライバシー保護、安全性確保、デジタル・デバイド対策などを並行して推進し、社会システムとして健全に機能するよう、**制度面からの検討**が必要である。

利用者プロフィールの標準化の検討



セマンティックWEB技術普及・発展の為に課題

- (RDF)によるメタデータ作成の普及・促進の為に！
 - 業界やコミュニティの規範となる相互運用可能なメタデータ標準の開発促進
- セマンティックWEBの容易化を図るツールの開発
- プライバシ保護
- セキュリティ、信憑性の確保

World-Wide Web Consortium(W3C)

- 1994年に設立されたWorld-Wide Web(WWW)に関する国際標準化組織。
- WWWを考え出したTim Berners-Lee(セマンティックWeb提唱者)が最高責任者(ディレクタ)を務めている。
- 米国MIT、仏INRIA、慶大SFC研究所がホストになり、DARPA(米国防総省国防高等研究計画局)、EC(欧州委員会)などの支援を受け協調して国際活動を推進している。 MIT:マサチューセッツ工科大学 INRIA:フランス国立情報処理自動化研究所
- これまでにHTML, XML, RDFなどの数多くの標準化成果を出している。
- 現在の加盟メンバー数は515社。

Submission

参加組織がWebに関する新たなアイデアを提案する手段で、新しい活動をW3Cに提案する場合によく利用される。

Workshop

新たな活動を開始する前に行われ、どのような問題がありW3Cとして何ができるか議論する。

Working Group

仕様の策定を行う。10～20名程度で構成され、Web, 電子メール、電話会議、face to-face 会議を組み合わせ仕様の作成を行う。この段階から参加していないと出遅れることが多い。

欧米の活動

- DARPA
 - DAML+OIL
- EUと米国Joint Committee
- W3C – XML、RDF, Semantic Web Activity
- DCMI(Dublin Core Metadata Initiative)—Metadata
- SOAP—XML Application
- Ontology, Knowledge Representation

欧州の研究開発プロジェクト On-To-Knowledge

目的:・オントロジーの発展系としてのコンテンツ主導型ナレッジ・マネジメント
ツールの開発

・プロジェクトでの開発成果の一つがOIL (Ontology Interface Layer)

位置付: 第5回欧州フレームワーク・プログラム下のISTプロジェクト
(プロジェクト番号IST-1999-19005)

期間: 2000年1月～2002年6月

予算規模: 総額250万ユーロ (EU委員会からは134万ユーロ)

予算元: EU Information Society Technologies (IST) Programme

研究機関:

- ・Vrije大学(オランダ)
- ・AIFB研究所(ドイツ)
- ・カールスルーヘ大学(ドイツ)
- ・AIdministrator(オランダ)
- ・BT研究所(英国)
- ・スイス・ライフ
- ・CognIT(ノルウェイ)
- ・Enersearch(スウェーデン)

欧州の研究開発プロジェクト IBROW

目的:・複数のデジタル・ライブラリから、利用者の要求に基づいて、ナレッジの検索を可能とする知的ブローカ・サービスの開発。

位置付:第4回欧州フレームワーク下のプレ・フェーズ後、2000年2月から第5回欧州フレームワーク・プロジェクト下でISTプロジェクトとなる。
(プロジェクト番号IST-1999-10132)

期間:2000年2月～2003年1月

予算規模:総額153万ユーロ(EU委員会からは110万ユーロ)

予算元:EU IST Programme

研究機関:・スペイン科学研究協議会 ・ AIFB研究所(ドイツ)

・スタンフォード大学

・インテリジェント・ソフトウェア・コンポーネンツ社

・Vrije大学(オランダ)

及び、次のインダストリー・アドバイザー・ボードのメンバー

・ダイムラー・クライスラー

・ドイツ・テレコム

・Bolesian

・ユニリーバ

・IBM(日本人)

・英国人工知能アプリケーション・インスティテュート

欧州の研究開発プロジェクト IBROW

マイルストーン:

1. 各ライブラリに分散された知識要素の相互運用性のためのインフラの開発
2. 標準的なブラウザによってアクセス可能な知的推論サービスの開発
3. Web／イントラネット上の情報検索用フィルタリング・エージェントを作成・カスタマイズするためのシステムの開発
4. 問題解決およびオントロジーを選択・修正・作成するためのイントラネット／Webサービスの開発
5. 知識要素を記述するための注釈言語およびツールの開発

米国の研究開発DAML Project

DAML: DARPA Agent Markup Language

目的: セマンティックWebのコンセプトを容易化を図る言語とツールの開発

期間: 2000年8月～2005年

予算規模: 7000万ドル

予算元: DARPA (Defense Advanced Research Projects Agency)

研究機関: MIT、Stanford University、University of West Florida、
Teknowledge等 16の技術開発チームが参画。

米国の研究開発 Dublin Core Metadata Initiative

目的: 相互運用可能なメタデータ標準の普及促進と情報資源記述の為の
特定のメタデータ・ボキャブラリ作成

- ・境界を超越した検索のためのメタデータ標準の開発
- ・メタデータ集合の互換性のための枠組の定義
- ・コミュニティの発展を促進する、または規範となる特別なメタデータ集合の開発を促進する

期間: 1994年にDublin Coreの活動立ち上げ

予算規模:

予算元:

研究機関: