

セマンティック Web 技術と次世代電子政府での活用方法

に関する調査研究

調査報告書

平成14年3月

財団法人 ニューメディア開発協会

目 次

1. 調査の背景と目的.....	3
1.1 背景.....	3
1.2 目的.....	3
2. セマンティック Web の研究開発・標準化の動向と社会的影響に関する調査.....	4
2.1 セマンティック Web の現状.....	4
2.2 米欧の研究動向.....	7
2.2.1 米国の研究動向.....	7
2.2.2 欧州のセマンティック Web 動向.....	15
2.2.3 欧州の電子政府におけるメタデータ応用の動向.....	19
2.2.4 欧州におけるオントロジー技術の動向.....	27
2.2.5 欧州におけるセマンティック Web 関連システムの動向.....	31
2.3 標準化機関の動向.....	33
2.3.1 W3C の Semantic Web Activity.....	33
2.3.2 WOWG のオントロジー記述言語の動向.....	33
2.3.3 OWL の利用事例.....	37
2.4 活用事例と課題.....	43
2.4.1 ナレッジマネジメント.....	43
2.4.2 e-learning.....	45
2.4.3 e-ビジネス.....	49
2.4.4 コミュニティー支援.....	50
2.4.5 課題.....	52
2.5 応用システム開発時の考慮点.....	54
2.6 社会におよぼす影響.....	56
3. 次世代電子政府での活用方法に関する調査研究.....	59
3.1 セマンティック情報（メタデータ）の付与に関する調査.....	59
3.1.1 国内の行政公開情報の調査.....	59
3.1.2 海外政府のメタデータ付与に関する調査.....	71
3.1.3 メタデータ階層モデル調査.....	73
3.1.4 メタデータ定義調査.....	77
3.1.5 利用者プロフィールのメタデータ定義に関する調査.....	86
3.2 情報のセマンティック検索に関する調査.....	94
3.2.1 メタデータ管理システム調査.....	94
3.2.2 半自動メタデータ生成システム調査.....	98
3.2.3 対話型メタデータ生成システム調査.....	106
3.2.4 利用者プロフィール定義システム調査.....	110
3.2.5 セマンティック検索システム調査.....	114

1 . 調査の背景と目的

1 . 1 背景

来るべき高度情報処理社会のあり方を抜本的に変える可能性を秘めた新たな Web 技術の開発が、欧米において急速に進行中である。その技術は、セマンティック Web 技術である。

現在の Web 技術が、Web 情報を人間が読んで、人間が理解し、人間が操作することを前提にしているのに対し、セマンティック Web は、Web 情報をエージェントにより自動処理させることを目指している。

セマンティック Web の基本原理は、Web を含むあらゆるデータと情報の意味をマシンリーダブルなメタデータで記述し、人間の代わりにするソフトウェア（エージェント）で自動処理させることである。

セマンティック Web は、物理的な事物をも含めた膨大な情報、ハードウェア、ソフトウェア及び機能と言ったあらゆるものを記述可能とし、人間が簡単な指示を行なうだけで、エージェントにより、自律的且つ自動的に処理をすることを目指している。

セマンティック Web はインターネット上での単調で機械的な作業を自動処理してくれるので、経済活動の効率化と社会や家庭の利便性の向上とを飛躍的にもたらす。

米国や EU では、政府の財政的補助の下にセマンティック Web の研究開発プロジェクトが推進されており、その中でセマンティック Web の国際標準規格の検討も進められている。

このセマンティック Web の国際標準規格が、米国と欧州によって、我が国の事情を考慮することなく決められてしまう事のないように、我が国も早急にこうした研究活動に取り組み、その国際標準規格作りへのコミットメントを行なう必要がある。

また、セマンティック Web の研究開発プロジェクトの成果を先取りした新たなビジネスが、欧米では生まれつつある。、我が国の情報技術が世界をリードするためにも、セマンティック Web で拓ける新しい Web の世界への研究に注力することが肝要である。

1 . 2 目的

上記の背景を踏まえて、当調査研究では、平成 13 年度にセマンティック Web 技術の現状と動向とを調査すると共に、次世代の電子政府にセマンティック Web 技術を活用する視点から、セマンティック Web 技術の本質的な利点の究明、利点を生かす上での基盤技術、ツール技術、アプリケーション技術等について調査研究する。セマンティック Web 技術の実用化に向けた問題点や必要な技術要素の洗い出しを行ない、将来の利用に向けた課題を明らかにする。

2. セマンティック Web の研究開発・標準化の動向と社会的影響に関する調査

2.1 セマンティック Web の現状

Web の発明により、インターネットを中心とする IT 革命が引き起こされた。この IT 革命により空間的・時間的制約が取り払われ、経済活動の効率化や新規産業の創出など経済社会構造の転換が生じてきている。

この IT 革命のキーとなった現在の Web 技術は、Web 上の情報を人間が目で見、人間が理解し、人間が指示することを基本原理としている。このため人間が細かな指示を与えなければならないし、人間が理解できない場合は有効に活用することができない。

このような人間だけが理解できる情報をヒューマンリーダブルな情報というが、さらに Web の情報を、コンピュータなどの機械がその意味を理解できる情報（これをマシンリーダブルな情報という）にし、コンピュータで自動処理させる研究とその応用が、次世代 Web 技術として、W3C（World Wide Web Consortium）や DARPA（米国国防高等研究計画局）を中心に米欧で強力に推進されている。これがすなわちセマンティック Web である。

セマンティック Web の基本原理は、あらゆるデータと情報について、マシンリーダブルかつ構造化されたメタデータ¹を付加することによってその意味を記述し、エージェント²で自動処理させることである。このように、コンピュータがエージェントなどを使って、自動的な意思決定を行なえるような情報空間のことをセマンティック Web という。

Web 技術の標準化団体である W3C の設立者である Tim Berners-Lee は、WWW 立ち上げ時からセマンティック Web を構想しており、2 年ほど前からはかなり具体的にセマンティック Web を提唱している。W3C では 2001 年 2 月に Semantic Web Activity を立ち上げており、現在の W3C の中心的な活動は、このセマンティック Web である。Tim Berners-Lee、James Hendler、及び Ora Lassila の言葉によれば、「セマンティック Web は現在の Web を拡張したものだ。Web に記述される情報に明確な意味の定義を与え、コンピュータと人間とがうまく協力して作業できるようにする。現在の Web 構造に意味関係を織り込んでいく初歩的な作業がすでに進みつつある。近い将来、今は画面に表示されているだけの情報をコンピュータがきちんと“理解”してもっと有効に利用し、非常に強力な機能を実現できるようになるだろう³。」ということである。

すでに米国及び EU は調査開発に国家予算を当てており、例えば、米国は DAML プロジェクトに国家予算として 2000 年から 3 年間で 7000 万ドル（約 80 億円）の拠出をしている。EU 委員会は、On-To-Knowledge プロジェクトに 2 年半で 134 万ユーロ（約 1.5 億円）を、IBROW プロジェクトに 3 年間で 110 万ユーロ（約 1.2 億円）の拠出をしてい

¹ メタデータとは、データを記述するデータのことをいう。メタデータを記述するデータのこと、メタデータという。

² エージェントとは、人間の代わりにするソフトウェアのことをいう。

³ Tim Berners-Lee, James Hendler, Ora Lassila. "The Semantic Web." *Scientific American*, May 2001 より。日本語訳は「自分で推論する未来型ウェブ」日経サイエンス 2001 年 8 月号。

る。また、セマンティック Web に関する技術開発において米国と EU はジョイントプログラムをスタートさせ、オントロジー言語の DAML+OIL を作成した。米国では DARPA、MIT、スタンフォード大学、メリーランド大学が、EU ではフライエ大学（オランダ）、カールスルーエ大学（ドイツ）、ブリストル大学（英国）、INRIA（フランス）などが中心となってこれらの研究開発を進めており、多くの成果を出している。

セマンティック Web は、W3C の策定した XML (eXtensible Markup Language) と RDF (Resource Description Framework) を基盤技術とする。Tim Berners-Lee は、2000 年 12 月の講演資料⁴において、セマンティック Web のアーキテクチャとして、図 2-1 のようなレイヤー構造を提示した。

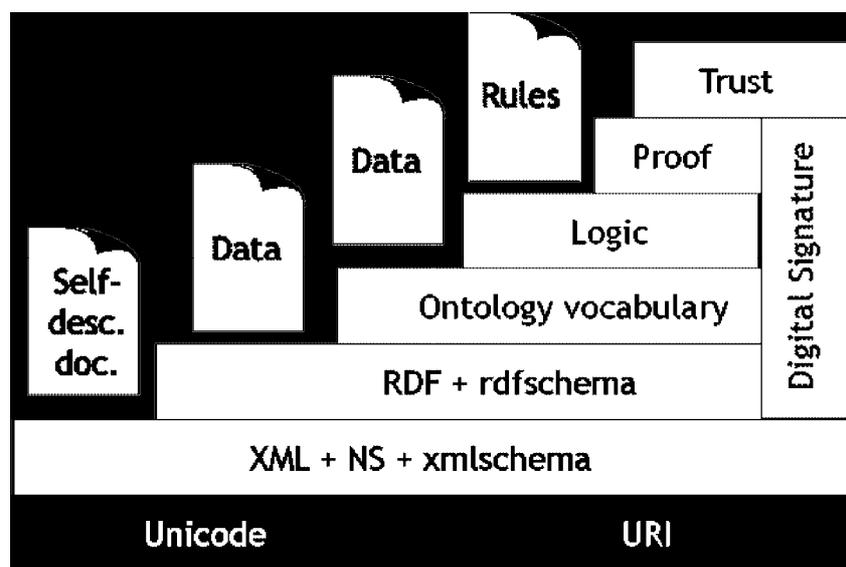


図 2-1 セマンティック Web のアーキテクチャ

以下に、主要なレイヤーについて概要説明する。

XML・・・Web 作成者等が<zip code>等のタグを自ら定義し、Web ページやテキストにタグを付加することにより、それらに注釈を付けることができる。XML によって Web 作成者は文書に任意の構造を付加できるが、その構造が何を意味するかについては XML は何も規定しない。

RDF・・・RDF は XML によって文書に付加された構造に対して、その意味を付加するものである。すなわち、メタデータを記述するための仕様である。RDF では、メタデータは図 2-2 のような主語（リソース）、述語（プロパティ）、目的語（値）の 3 つ組（トリプル）で記述される。すなわち、主語:「`http://www.w3.org/Home/Lassila`」、述語:「`Creator`」、

⁴ <http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html>

目的語：「Ora Lassila」である。これを日本語の文章で述べると、「http://www.w3.org/Home/Lassila の Creator（作者）は Ora Lassila である」となる。

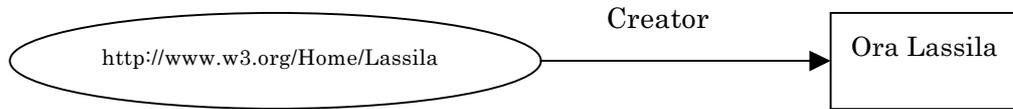


図 2-2 RDF の 3 つ組

オントロジー（Ontology）・・・メタデータの表現方法が異なる場合（すなわち、二つの語彙体系が同じ要素・概念に対して異なった名前を付けている場合）、意味が通じなくなる。複数の語彙体系間で意味の共有化を図るための、「用語間の関係を正式に定義している文書またはファイル」をオントロジーという。

論理（Logic）・・・エージェントはメタデータを処理することで Web ページの内容を理解し、人が発した複雑な質問に的確な回答を導き出すことができる。このような、メタデータに関する処理を「推論」と呼ぶ⁵。論理レイヤーは、推論を行なうための仕組みを提供する。セマンティック Web では、これまでの検索エンジンにおける and と or のみならず、三段論法や否定、述語論理で用いるような量化記号も取扱えるようになる。

⁵ 萩野達也「セマンティック Web の現状と課題」データベースと Web 情報システムに関するシンポジウム論文集、p74（2001）。

2.2 米欧の研究動向

2.2.1 米国の研究動向

2.2.1.1 DAML (DARPA Agent Markup Language)

(1) DAML 概要

DAML はオントロジーを記述するための言語である。RDF スキーマをベースに拡張が施されている。

DARPA から 7000 万ドル (5 年間)⁶の資金提供を受けている。MIT (W3C) を含む複数のプロジェクトからなる。開発は 2000 年から開始され、正式なキックオフは 2000 年 8 月であった。

プロジェクトの目的は、セマンティック Web の実現を容易化するための言語 (オントロジー記述言語) の設計と、ツール (オーサリングツール等) 及びソフトウェア (知的インテリジェント・エージェント) の開発である。

MIT、スタンフォード大学、BBN、Nokia Research Center 等 18 の技術開発チームが参加し、それぞれのプロジェクトを担当している⁷。DAML のプロジェクトマネージャーは米メリーランド大学の James Hendler 教授である。各プロジェクトの年間予算は、平均して 100 万ドル程度であると思われる⁸。

DAML 仕様書

2000 年 10 月 10 日に "DAML-ONT" 仕様書の初版⁹がリリースされた。

2000 年 10 月に Joint US/EU アドホック・エージェント・マークアップ言語コミッティー (Joint Committee)¹⁰が DARPA / メリーランド大学の James Hendler と EU IST プログラムの Hans-Georg Stork によって結成された。

2001 年 1 月 27 日に Joint Committee は DAML+OIL (December 2000)¹¹ という仕様書をリリースし、DAML と OIL との統一化を行なった。2001 年 3 月 27 日には "DAML+OIL (March 2001)"¹² として改版がなされた。DAML+OIL は RDF スキーマをベースとしたオントロジー記述言語であり、RDF にはない機能が追加されている。

DAML-S version0.5

2001 年 5 月 30 日に DAML-S version0.5 がリリースされた¹³。DAML-S は DAML ベースの Web サービス向けオントロジー記述言語であり、Web サービス提供者に対して、曖昧性がなくコンピュータが解釈可能な形式で Web サービスのプロパティとケーパビリティとを記述するためのマークアップ言語を提供するものである。DAML-S の目的は、Web サービスの発見 (discovery)、起動 (invocation)、構成 (Composition) とサービス間

⁶ INTAP 「WWW10 調査報告書」より。

⁷ <http://www.daml.org/researchers>

⁸ INTAP 「BBN Technologies 出張報告」より。

⁹ <http://www.daml.org/2000/10/daml-ont.html>

¹⁰ <http://www.daml.org/committee/>

¹¹ <http://www.daml.org/2000/12/daml+oil-index.html>

¹² <http://www.daml.org/2001/03/daml+oil-index>

¹³ <http://www.daml.org/services/daml-s/2001/05/>

操作（Interoperation）、及び実行モニタリング（Execution Monitoring）などのタスクを自動化するための記述を提供することである¹⁴。DAML-S version0.5 は DAML+OIL（March 2001）の上に構築されている。

（２）DARPA による DAML プロジェクトの位置付け

現在 DARPA において国家予算によって運営される 30 件のリサーチ・プログラムの中に、「PE0602301E Computing Systems and Communications Technology」というプログラムがある（大分類）。本プログラムは 6 件の研究プログラムから成り立つが、このうちの 1 件が「ST-11 Intelligent Systems and Software」である（中分類）。本プログラムは、2001 年計画で 6 件、2002 年計画で 7 件のプログラムを含むが、DAML はそのうちの 1 プログラムに当たる（小分類）。なお、DAML については 2001 年度がプログラム初年度になる（2000 年度の成果報告に DAML の記述はない）。（図 2-3 参照のこと。）

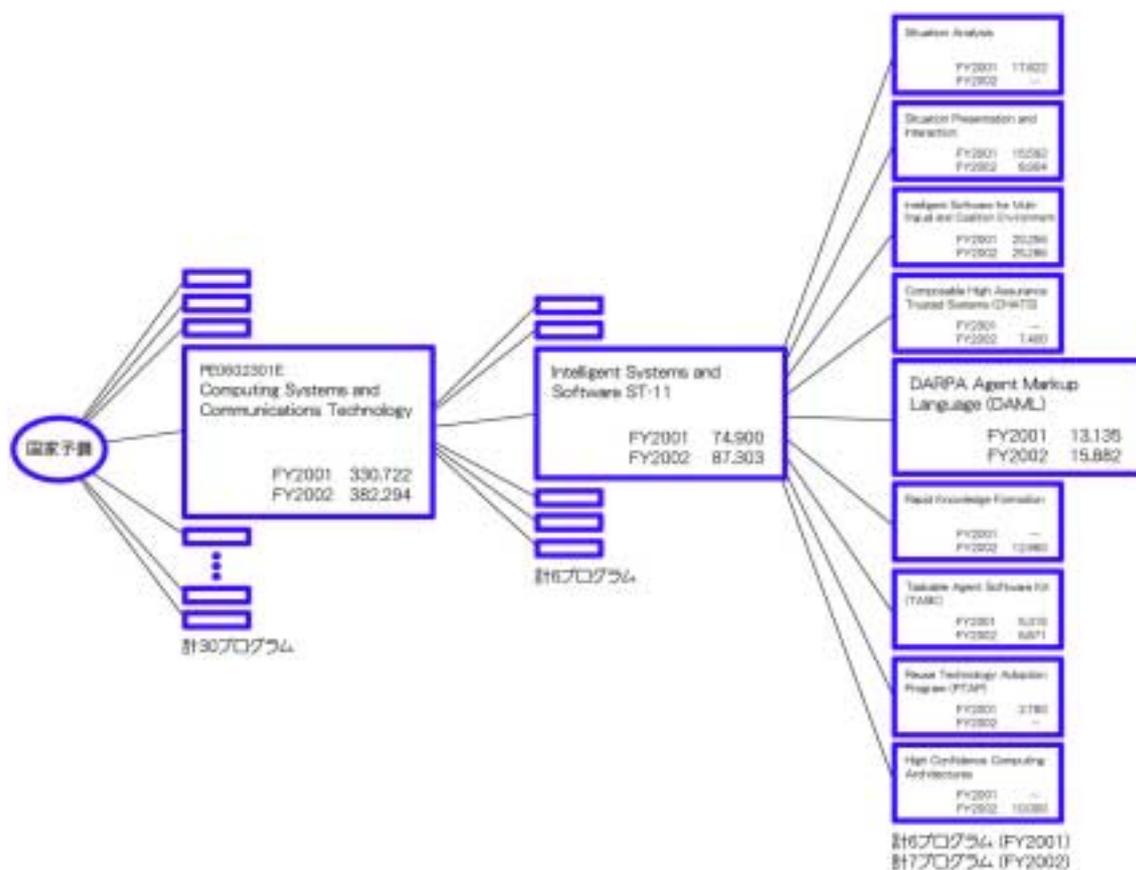


図 2-3 DARPA による DAML プロジェクトの位置付け

「DARPA Agent Markup Language (DAML)」（小分類）の実施計画及び予算（2001 年度）

¹⁴ INTAP「SWWS の聴講メモ（暫定版）」より。

- DAML 言語仕様の完成
- Intelink¹⁵向けブリーフィング・ツールのワーキングバージョン・リリース
- Intelink 上の DAML 検索ツールのワーキングバージョン・リリース
- Intelink 上の DAML オントロジー作成ツールのワーキングバージョン・リリース
- non-pre-planned エージェント間の相互運用をサポートするための DAML 向け要求仕様定義
- 陸軍戦訓センター (Center for Army Lessons Learned) のアーカイブ情報の保存、アクセスや編集能力を向上させる DAML オントロジー作成ツールの効用に関するデモンストレーション
- データ抽出及びリモートでキャッシュされる Web 情報の更新率に関する一階述語論理 (first-order rule) の開発
- Robust Open Source の開発モデルに準拠し、組み立て可能な高アシュアランス・高信頼性システム実現に向けた新しいアプローチの研究
- 高アシュアランス・高信頼性インプリ言語やツール開発に関するフィージビリティと新しいアプローチの研究
- 高アシュアランス・高信頼性を備えたシステム・プロテクション・プロファイル及び高信頼性言語・ツール開発に関する新しいアプローチの研究

計 10 項目 \$M13.135 (約 16 億円弱)

「DARPA Agent Markup Language (DAML)」(小分類)の実施計画及び予算
(2002 年度)

- DAML 技術を C2 に応用したツールセットの定義
- Intelink DAML ブリーフィング・ツールの実験的な分析の実施
- 実務的な Intelink ノード上での DAML 検索ツールの展開
- 軍や国の情報機関向け Web アプリケーションのための DAML オントロジー作成ツールのプロトタイプの実施
- 陸軍戦訓センター (Center for Army Lessons Learned) における検索 (search and retrieval) ツールの能力を向上させた選りすぐりの DAML ツールのプロトタイプ作成
- Millennium Challenge への参加を含む、海軍及びジョイント C2 のインターオペラビリティへの DAML の応用に関する実験的な分析の実施
- World Wide Web の実験的評価や設計に関する 150 万件以上の DAML ステートメントのリポジトリ作成
- dynamic prioritization management 向け技術の開発
- 帯域状況が変化する中でのインテリジェントな情報配信技術の開発

計 9 項目 \$M15.882 (約 19 億円)

¹⁵ FBI、CIA、DEA (米麻薬取締局)、NSA (米国家安全保障局) 等で収集した諜報情報の相互利用ネットワークのこと。Intelink のテストベッド運用は 1994 年に開始された。

(3) DAML 各プロジェクト概要

BBN Technologies

BBN は DAML プロジェクトにおいて、プロジェクト全体を統合する「オーバーオール・コントラクター」の地位を与えられており、商用化の役割も担っている。BBN は DARPA から 2000 万ドルの年間予算を得ている¹⁶。

スタンフォード大学及びドイツ・カールスルーエ大学

Web 上のインテリジェント・エージェントを可能にする OnTo Agents プロジェクト。インフラストラクチャ・コンポーネントの開発(OnTo Agent セマンティック Web ページ・アノテーションツール、オントロジー表現ツールキット、推論システム) や、アノテーションされた Web ページを用いた、貨物輸送スケジューリング用のアプリケーションである OnToCargo-Agent のデモを行なう。

CMU 及び Nokia Research Center

ATLAS (Agent Transaction Language for Advertising Services) プロジェクト。ATLAS とは、広告サービス向けの DAML ベースのエージェント処理言語である。広告したり、要求したり、発見したりするエージェントサービスのための DAML 言語とツールの開発を行なう。また、広告用オーサリングツールや、ミドルエージェント向け仲介アルゴリズムの開発も行なう。

DAML では、各プロジェクトでそれぞれの分野の ontology を構築している¹⁷。上部集合としては、Cyc が無料で公開しているものを使うとのことである¹⁸。

(4) DAML 関連ツール

DAML Viewer

2000 年 11 月 17 日には DAML Viewer がリリースされた¹⁹。同ツールでは、DAML で書かれた Web 上のステートメントを表示することができる。BBN Technologies の Mike Dean 氏と Kelly Barber 氏が開発した。

OntoEdit

2001 年 3 月 7 日にはドイツの Ontoprise 社が OntoEdit v1.03 をリリースした²⁰。同ツールでは、DAML+OIL のフォーマットでオントロジーを作成することができる。OntoEdit については、3 . 2 . 3 節で紹介する。

¹⁶ INTAP 「BBN Technologies 出張報告」より。

¹⁷ <http://www.daml.org/ontologies/>を参照のこと。

¹⁸ 浦本「Semantic Web Working Symposium 参加報告メモ」より。

¹⁹ <http://www.daml.org/viewer/>

²⁰ <http://www.ontoprise.de/>

DAML Crawler

2001年5月4日には DAML Crawler がリリースされた²¹。同ツールでは、DAML で書かれた Web 上のステートメントを収集することができる。BBN Technologies の Mike Dean 氏と Kelly Barber 氏が開発した。

DAML Validator

2001年5月11日には DAML Validator がリリースされた²²。同ツールでは、DAML を使用して作成されたオントロジーファイルの文法をチェックすることができる。BBN Technologies の Dave Rager 氏が開発した。

その他の DAML ツール²³

その他の DAML ツールとして、以下のものがある。

- ・ DAML オントロジー・ライブラリ：クリアリングハウス (150 以上のオントロジー)
- ・ RDF API：Java パーサ及びシリアライザー
- ・ OilEd：オントロジー・エディター
- ・ Chimaera：オントロジー・アナライザー
- ・ DAML Viewer、PalmDAML²⁴、HyperDAML：ナビゲーター GUI
- ・ HAIRCUT：text/DAML 検索エンジン

DAML アプリケーション²⁵

DAML アプリケーションとして、以下のものがある。

- ・ UMBC IT Talks
- ・ ISI todo
- ・ Horus

IT Talks

IT 関係の研究者の講演案内の Web ページに対し、メタデータを付与し、個人のプロフィール、カレンダー、地図情報(物理的にその講演を聴きにいけるか否か)と組み合わせて、興味のある講演を検索し、カレンダーに反映させるシステムである²⁶。

(5) DAML の課題と今後の見通し

DAML の課題としては、DAML オントロジーの作成やマークアップ、アクセスのためのツールがまだ十分でない²⁷。

BBN の意見では、DAML のツールはリサーチレベルのものに留まっている。来年には

²¹ <http://www.daml.org/crawler/>

²² <http://www.daml.org/validator/>

²³ <http://www.daml.org/2001/06/swday-daml/Overview.html>

²⁴ PalmDAML とは、PalmOS 向けに DAML Viewer を実装したもの。

²⁵ <http://www.daml.org/2001/06/swday-daml/Overview.html>

²⁶ 浦本「Semantic Web Working Symposium 参加報告メモ」より。

²⁷ INTAP「BBN Technologies 出張報告」より。

W3Cで標準が出てくるので、民間の開発者が商用のツールを出してくるだろう。また、オントロジー言語も普及するであろう。商用プロダクトは、2、3年で出てくるのではないが、DAMLはXMLと置き換わるものではなく、より豊かな理解のための道具として、エージェントとともに普及するであろう。重要な点は、DAMLを使うことでソフトウェアエージェントを軽量かつ汎用にできることである。FIPA (Federation for Intelligent Physical Agents エージェント技術の国際標準化団体)は、とりわけW3Cで標準化された後には、DAMLの標準を受け入れるであろうということである²⁸。

2.2.1.2 DCMI (Dublin Core Metadata Initiative)

(1) 概要

DCMIは1994年にシカゴで開催されたW3Cの年次コンファレンスであるWWW5から本格的に活動を開始した。

DCMIは700～800の会員組織及び個人の会員を持つ。国家レベルではオーストラリア、デンマーク、フィンランド、英国、アイルランドが加入済みである。ニュージーランド、カナダが地方政府レベルで加入しており近々国家レベルで加入予定である。米国は地方政府ならびにエージェンシーレベルでの加入は行なっているが、連邦レベルということになると最後の加入者になるかもしれない²⁹。

DCMIの目的は、相互運用可能なメタデータ標準の普及促進と、情報資源を記述するための特定のメタデータ・ボキャブラリの開発である。

Web上の各種の情報資源を統一的に記述するためのメタデータ(Title、Subjectといった基本エレメントの集合)を定義することを目指している。メタデータの記述にはRDFを用いる。

DCMIの活動において中心的な役割を担うOCLC(Online Computer Library Center)の年間予算は、116万ドルであり、その内の90%は米国の図書館コンソーシアムのメンバーから得ている。OCLCには25名のフルタイムの研究者がいる。

DCMIが策定したダブリンコアとは、情報資源を記述するための基本的な15エレメントのことである。多くの人たちが共同に使うためのポリティカルな合意の結果15個のメタデータエレメントとなった。これはメタデータ記述の共通になる部分として30ヶ国(25言語)間でコンセンサスが取られている。ダブリンコアの「コア(Core)」とは、このようにメタデータ記述の共通部分、基本部分という意味である。この基本部分に加えて、用途ごとに拡張を行なうこと(拡張エレメントを作成すること)が可能である。

Creator	Title	Subject
Contributor	Date	Description
Publisher	Type	Format
Coverage	Rights	Relation
Source	Language	Identifier

²⁸ INTAP「OCLC(DCMI)出張報告」より。

²⁹ INTAP「OCLC(DCMI)出張報告」より。

各エレメントの詳細については、3.1.4節を参照のこと。

ダブリンコアの拡張性としては、アプリケーションごとに拡張部分を作成している。例えば、「インターナショナル・デジタル・ライブラリ・プロジェクト（バージニア工科大学）」、「Dublin Core・ガバメント・グループ」、「Dublin Core 教育グループ」では15エレメントの上にそれぞれが拡張を行なっている。お互いの互換性は、拡張部分においては必ずしもなくてよい。また、完璧な互換性も必要ない³⁰。

2000年10月、DCMIは活動領域の拡大を行なった。DCMIのミッションは、以下の活動を通じてインターネットを用いた情報資源検索を容易化することとなった。

- ・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みの定義
- ・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

DCMIの2001年のWorkshopは10月22日～26日に東京で開催された。

(2) 今後の展望

ダブリンコアに基づいた技術適用の例としては、RDFを使用したAdobe社のAcrobat製品がある。また、ダブリンコアの技術を使ったビジネスデータの交換・共有は有望である（文書管理システム）。例えば、サンマイクロシステムズ、ノキア、自動車業界はその実現について模索中である³¹。

ダブリンコアの標準化は欧州や米国では進みつつある。例えば、政府文書にダブリンコアに準拠したメタデータを付与しようという取り組みが各国でなされている。各国政府のメタデータ標準については3.1.2節を参照のこと。

³⁰ INTAP「OCLC（DCMI）出張報告」より。

³¹ INTAP「OCLC（DCMI）出張報告」より。

2.2.1.3 RSS (RDF Site Summary)

(1) 概要

コンテンツ・シンディケーションのためのメタデータ記述フォーマットである。メディア・コンテンツ・ポータルを開発するために使用する。標準化されたボキャブラリ(ダブリンコア)の上に構築される。RDF シンタクスを使用している。

RSS は、Netscape が 1999 年に自社ポータル(My Netscape³²)のために開発した RSS0.9 が基となっており、この時点では RDF とは無関係であった。その後 RSS は、Netscape と My Userland の Dave Winer³³によって拡張され、他のポータルサイトでも利用されるようになるが、ボキャブラリの問題などから 2000 年中ごろには使われなくなっていた。しかし、RSS-DEV ワーキンググループが結成され、RDF を中心としてモジュールによる拡張機能を備えた RSS1.0 仕様書³⁴が 2000 年 12 月に発表された。また、RSS1.0 とともに、ダブリンコア・モジュールと、いくつかのサポート用ツールが発表された³⁵。

(2) RSS の応用システム：O'Reilly Network

O'Reilly Network

オンライン・ネットワーク³⁶とは、新しいプラットフォーム、プログラム言語、OS を含むオープンな次世代技術に関心を持つ開発者向けのポータルサイトである。同ネットワークは専門技術者に技術情報を明確かつ統合的に提供する。また、専門技術者のコミュニティーのためのフォーラムを作成し、彼らが互いにコミュニケーションを行ったり、提携サイトとコミュニケーションしたりする機会を提供する。

O'Reilly Network の Meerkat³⁷

Web ベースのシンディケート化された RSS ベースのコンテンツ・リーダーである。RSS は、ニュースや製品紹介、討論のスレッド、その他の多種多様なコンテンツを、いくつかのチャンネルに分類して配信するために用いられる。Meerkat は、これらのストーリー(個々のニュース等)に対して単一のインターフェイスを提供する。Meerkat では、最新のストーリーはそのソースに関わりなく、常にトップの位置に置かれる。

Meerkat はデータベースにチャンネルのリストを保持している。1 時間に 1 回、Meerkat はそれぞれのチャンネルに関連付けられた RSS ファイルを訪問し、データベースに新たなストーリーを追加する。

Meerkat を利用することにより、利用者は自分の関心のあるストーリーを専ら表示することができ、サイトをあちこち飛びまわる必要がなくなる。

³² Netscape の My Netscape はチャンネルラインに沿って組織された大きな「スタート」ページを提供している。チャンネルを好きなように追加したり、削除したりすることで、利用者は自分のページをカスタマイズすることが出来る。

³³ Dave Winer の My.Userland.Com はより広範な焦点を持っており、Userland と EditThisPage.com (Manila の Web サイトの成長するコミュニティー)でホストされた WebLogs からコンテンツを取り入れている。

³⁴ <http://groups.yahoo.com/group/rss-dev/files/namespace.html>

³⁵ INTAP「WWW10 調査報告書」より。

³⁶ <http://www.oreillynet.com/>

³⁷ <http://www.oreillynet.com/meerkat/>

2.2.2 欧州のセマンティック Web 動向

欧州におけるセマンティック Web 技術の研究・開発にあたっては、予算年度が 1998 年から 2002 年の 5 年間にわたる EU の第 5 次 RTD (research and technological development) フレームワーク・プログラム (FP5) によって予算措置が講じられている。FP5 は 4 つの個別プログラムと 3 つの横断的プログラムからなり、全体で 23 の、いわゆる主要なアクションや一般的な研究活動をカバーしている。これらすべての活動のために、5 年間で 149 億 6000 万ユーロが投じられることになっている。FP5 は、エネルギー政策にからむものまで含め、非常に広汎な領域の研究活動を支援しているが、セマンティック Web 技術に関わる研究・開発は、上記 4 つの個別プログラムのうち、最大の予算を持つ「ユーザーフレンドリーな情報社会の創造 (IST)」プログラム (5 年間で 36 億ユーロ) に含まれる。

FP5 はプログラム予算の策定が 1997 年以前に遡るため、当初セマンティック Web に直接的に言及するプログラムはなかったが、2001 年になって、セマンティック Web 技術に関するアクションラインを発表した。しかし、具体的かつ直接的なセマンティック Web 関連技術の研究・開発は後述する第 6 次フレームワーク・プログラム (FP6) において行われることになる。FP5 の IST プログラムにおいては、別途メタデータを用いたプログラムが 20 件以上活動している。

EU 委員会は 2001 年 2 月 21 日、FP6 に向けたプロポーザルを発表した。このプロポーザルにおいても、IST (Information Society Technologies)、いわゆる IT が主要テーマ領域の一つに位置付けられている。FP6 の予算年度は 2002 年から 2006 年までで、5 年間に 175 億ユーロが投じられる予定である (対 FP5 比 16%増)。

IST 領域において実施される予定の様々な活動に備え、EU 委員会の情報社会総合理事会 (DG Information Society) では、一連のミーティングを持った。ミーティングの目的は、FP6 において IT のコンテンツや実装手段に関し産業界や学術的研究機関に対して意見を求めることである。このうち、2001 年 4 月 27 日にブラッセルで開催された知識関連技術に関するミーティングには、学会ならびに産業界をリードする知識技術の専門家 22 名が一堂に会した。

本ミーティングにおいては、知識関連技術の研究・開発を進めるに当たり、主要な取り組みテーマとして以下の 4 つを挙げた。

- ・ 知識システムの可用性 (Usability of knowledge systems)
- ・ 知識としてのコンテンツ (Content as knowledge)
- ・ 標準とインターオペラビリティ (Standards and InterOperability)
- ・ ナレッジ・コミュニティとポータル (Knowledge Communities and Portals)

また、ミーティングの最終報告書において、セマンティック Web が概念的には広範な支持を獲得しつつあり、RDF を採用することが知識システムのためのコンテンツ構築に向けて重要なステップであるといった、現在のマーケット状況や研究状況についての総括がされている。アプリケーションやマルチメディア・モバイル開発環境へのリンクについては、2000 年 11 月にルクセンブルクで開催されたセマンティック Web ワークショップに

において報告された。

欧州でのセマンティック Web 技術に関する理解は総じて以下のようなものである。

- ・ セマンティック Web はマシンが意味情報を含んだドキュメントやデータを理解できるようにする。
- ・ セマンティック Web の主要コンポーネントは、知識表現システム・技術、オントロジー、そしてインテリジェント・エージェントである。
- ・ 学会においては、個別領域向けのオントロジーを開発する動きが高まりつつあるが、強固な競争力を確保するには至っていない。
- ・ 「知識システムの可用性」に関しては、「意思決定支援システム」を、セマンティック Web 技術を用いて構築すれば、ビジネスにおいてユーザーのニーズに直接的に応えることができるのではないかという議論が行われている。
- ・ 意思決定において、マシンに委ねられる部分が多くなれば、人間は本当に重要な部分のみに関与すればいいという、10 年前の SIS (戦略情報システム) で想定されたことが現実味を帯びることになる。
- ・ 人間が重要部分への関与のみで意思決定が遂行されるようになれば、今度はユーザーが「楽しんで」、このプロセスを実行できるようにしなければならない。そのためには、単純にコンテンツやメタデータを作ればいいというわけではない。知識発見プロセスをモデル化し、知識についてのライフサイクル全体の中にユーザーの行動パターンを適用していく必要がある。
- ・ マシン・トゥ・マシン (デバイス・トゥ・デバイス) の環境が重要なのは、最終的なゴールが「ネットワーク化された状態でのヒューマン・トゥ・ヒューマンなインタラクション」だからである。
- ・ 「知識としてのコンテンツ」に関しては、テキスト・ドキュメントに限定されないマルチメディア・コンテンツや、マルチメディア・コンテンツ作成・流通の際に不可欠な認証・公証といったことが議論されている。
- ・ 企業内の知識システム活用については、組織知ないしは集合知の保存ならびに利用といったことが議論されている。
- ・ 「標準とインターオペラビリティ」に関しては、セマンティック Web 研究者にとっての重要領域とされる、別々のセクターもしくはサブセクターで構築されたオントロジー同士をリンクさせるための「上位レベルのオントロジー」といった概念を設定する必要性について論じられている。しかし、こうした議論を展開すると、オントロジー構築に当たってはボトムアップ型にすべきかトップダウン型にすべきかという議論を呼んでしまう。本件については双方のアプローチをハーモナイズすべきであると認識する。
- ・ オントロジーとメタデータは徐々に自動生成されていくべきであるという点は一般的な合意事項となっている。新たなオントロジーを生成し、校正し、相互運用性を持たせるに当たってはデータマイニングの技術が有用である。
- ・ 「ナレッジ・コミュニティとポータル」に関しては、ユーザーインターフェイスやアクセス環境といった議論が交わされているが、これらの議論はナレッジ・コ

コミュニティや知識ベースのポータル環境といったテーマに帰着する。こうした、ユーザーが情報を得ようとすれば必ず見に行く環境においては、必要としないコンテンツをフィルタリングしてくれる機能や、楽しんで見に行けるようなビジュアルイゼーション機能も求められる。こうした機能はeビジネスにおいても活用が可能である。

【参考】カールスルーエ大学のステューダー教授は開発すべき四つのキー・テクノロジー領域として以下のものを挙げている。

- ・ 知識の抽出 (Knowledge Extraction)
- ・ 知識の維持・管理 (Knowledge Maintenance)
- ・ ナレッジマネジメント (Knowledge Management)
- ・ 知識の表現 (Knowledge Presentation)

上記のプロセスを行なうに当たり、オントロジー構築やメタデータ整備を手作業で行なうとなると、すぐにもボトルネック化することになるだろう。したがって、オントロジーやメタデータを半自動的に生成するツールが求められる。個々のオントロジーは不均一に成立していき、異なったビジネス・シナリオやアプリケーション領域に適合する必要があるという前提に立つなら、オントロジー間の連携も喫緊の課題である。オントロジーやメタデータを更新し続け、その進化や拡張をサポートするためにもツールが必要である。

ステューダー教授はまた、たとえばエージェント技術を用いたナレッジマネジメント、セマンティック(メタ)ポータルやインテリジェント・サービスといったアプリケーション候補についても強調している。モバイル・ユーザー向けには、新しい情報機器用のコンテンツに対応させる必要があるが、メタデータとオントロジーによって、伝送データ容量を減らすことができる。もはやドキュメントベースのアプローチは捨てて、相互接続された「知識の断片」を提供していくやり方に集中していく必要がある。また、ビジネスプロセスをナレッジマネジメントに統合していく必要もある。そして、古くなった情報を見つけ最新の知識と関係付けられるようにしなければならない。

セマンティック・ポータルに期待されるのはセマンティック・クエリングを提供することである。同様に、メタデータ・ポータルでは一連のメタデータ・リソースへのアクセスを提供することが求められる。また、インテリジェント・サービスはセマンティック・メタデータを開発していく役割を担う。複数サービスを組み合わせる新しい方法を早急に生み出せるようなソリューションが求められる。

【参考】サレルノ大学のCapuano博士 (Dr. Capuano) は次の三つのアクション

を求めている。すなわち、第一に、分散サービスやセマンティック Web のアプリケーション向けにミドルウェアを提供するセマンティック・グリッドを作ることである。まったく新しい世代の知識ベースの技術やアプリケーションによって、たとえば、ユーザープロファイルに則ったリソース・セレクションやフィルタリングなどを行なうためのインテリジェント・モバイルエージェントやインテリジェント・インターフェイスが可能になるかもしれない。第二に、オントロジー間連携実現の一つの方法としては、あるドメインに特化したオントロジーを構築する場合にベースとするようなスタンダード・コア・オントロジーを通じて連携させるというものである。第三に、インテリジェント Web インターフェイスやそのアプリケーションに関しては更なる研究が必要である。特に、遠隔学習、情報探索、インテリジェント・コンシューマー・サポート・コールセンターのようなアプリケーションに関連する領域においてである。カプアーノ博士はまた、ナレッジ・モデリング、ナレッジ・レプレゼンテーションのための、複雑系システム向けのシミュレーションモデル、最適化モデル、数学的モデルに準拠した手法やツールをサポートするための集中的な努力が必要であると訴えている。

2.2.3 欧州の電子政府におけるメタデータ応用の動向

2.2.3.1 英国の動向

(1) 英国概要

IT 分野で出遅れていた英国では、トニー・ブレア首相のリーダーシップの下、1997 年から本格的に行政情報の電子化に取り組んでいる。1999 年からは、電子政府化を推進する「e-government 担当大臣」を新設し、イアン・マッカートニー大臣を任命した。

英国では 2000 年 7 月の段階で行政サービスの 33% がオンライン化されているが、2005 年までに行政サービスを 100% 電子化することを目標にしている。

2000 年 3 月の段階で、英国で自宅を保有する人のうちインターネットに接続している人の割合は 25% (1 年前は 13%) に過ぎない。そこで政府は 38 億ポンド (約 6000 億円) を投じてインターネット端末を普及させ、それを使いこなせるように国民を教育するプログラムを開始した。図書館や郵便局に端末の導入を始めており、2002 年までに英国内に 6000 ヶ所の公的なアクセス拠点を開設することになっている。

マッカートニー e-government 担当大臣は、「電子政府を導入するには、新しい技術と公的サービスの文化の改革が必要だ。我々のゴールは、国民の指先で 24 時間オープンしている便利な政府を作り出すことだ」と述べている。

(2) e-GMF (e-Government Metadata Framework) の概要

e-GMF は英国内閣府の Office of the e-Envoy が推進している。

2001 年 5 月に、Office of the e-Envoy は e-GMF を発表した³⁸。e-GMF は、公的部門にあまねく適用されるメタデータ標準を開発し、実施するための政府の政策を策定するものである。この標準は公的部門の情報システムのすべてに適用される予定である。同フレームワークでは次の 2 つの主要な決議事項が発表された。

- ・英国政府は、ダブリンコアを英国政府メタデータ標準 (UK Government Metadata Standard (e-GMS)) として採用する。ただし、ダブリンコアだけではニーズに応え切れないので、同フレームワーク内で規定された規則に基づいて、更なる要素の追加と改善を行なう。
- ・汎政府シソーラス (同意語辞典) を開発すること。シソーラスは構造化された用語とキーワードのリストから成り、我々が情報を正確に定義することと、情報をより迅速に発見することを手助けする。汎政府シソーラスは、政府、産業、および市民が利用できる予定である。

e-GMF に従うことは義務であり、段階的に施行される予定である。ダブリンコアの 15 エレメントについては直ちに義務付けられ、他の拡張エレメントと汎政府シソーラスについては、それらが開発され承認された際に義務付けられる。

e-GMF のインプリメンテーションは現在進行中の仕事であり、それには、開発のための中央集権的な (central) サポートとメカニズムを必要とする。そのポータルとなるのが後述の UK GovTalk の Web サイト³⁹である。

³⁸ <http://www.govtalk.gov.uk/documents/UK%20Metadata%20Framework%20v1%202001-05.rtf>

³⁹ <http://www.govtalk.gov.uk>

英国には Citizen Portal⁴⁰という、政府（国と地方）の情報とサービスへのクイックアクセスを提供するポータルサイトがあり、汎政府シソーラスはとりわけこの Citizen Portal のニーズを念頭において開発される予定であった⁴¹。

Office of the e-Envoy は、Metadata Working Group（MWG）を持ち、3ヶ月に1度の頻度でミーティングを行なっている。

e-GMF のミッションは、一般市民がデジタルテレビやキオスクなどあらゆる媒体を通じて24時間365日すべての国家機関の情報にアクセスできる環境を実現することであり、また、そのためにすべての国家機関のみならず地方自治体の参加を仰ぎつつ、国家情報へのメタデータ付与を進め、アクセシビリティと相互運用性の向上を図るためにシステムの標準化と近代化に努めることである。

これまでの成果としては、年間50億トランザクション、20の政府機関と340の地方自治体が参加し、高い評価を得るにいたっている。現状としては、各国の状況を様子見しているが、英国はセマンティック Web といったような新規技術を追うよりは既存技術の活用を通じた展開を志向している。

e-GMF の展開方法としては、各国共通の問題であるが、英国でもネットへの信頼性すなわちセキュリティへの不安が叫ばれていた。そこで、UKOnline.gov という国民による政府情報アクセスならびに省庁間のトランザクションのためのポータルサイトを立ち上げ、約1年間実施してきた。セキュリティを高め認証も行なえるようにするためにゲートウェイサーバも立ち上げた。

UKOnline.gov は二つのサイトを持っており、一つは内閣府内の情報交換を目的とした e-Envoy、もう一つが省庁間の情報交換を目的とした UK GovTalk である。UK GovTalk は国民から広くコメントを求めたり、新しい企画を持ち込んでもらうための Web サイトである。

英国においては通常、省庁間の情報流通は活発とはいえない。その背景には個人情報保護法の規定によって、省庁間の個人情報流通が規制されているという法的側面も影響している。ただし、Knowledge Network という省庁データの閲覧システムは運用されており、簡単な政策概要程度は閲覧することができる。現在は役人しか見ることができないが、年内にはネットで公開される予定である。しかし、現在でも政策立案者が大臣に上げた情報を他省庁の人が見ることができるといった状況に抵抗もあり、そうした文化の改革が求められる。具体的には、当面は強制的に情報公開ならびにメタデータ付与を積極的に推進するよう担当者にプレッシャーをかけたり、広報担当官にクレームを上げたりする方法が考えられる。

Knowledge Network では初期バージョンのメタデータを使っている。データを大中小のカテゴリーに分類するにあたっては、ダブリンコアの Subject エレメントを採用している。メタデータを付与することにより、検索の容易化を意図している。

⁴⁰ <http://www.ukonline.gov.uk/>

⁴¹ しかし、汎政府シソーラスの開発については、実現可能性について検討した結果、中止されることとなった。将来的にアプリケーションがたくさん出てくるとシソーラスが必要になるであろうが、その際には多言語対応している Eurovoc が採用される見込みである。Eurovoc とは、15の言語に対応した多言語シソーラスであり、20年の歴史を持つ。将来は EU 域外の言語も含め言語の標準化が行われていく予定

メタデータの拡張に関するドキュメントには2バージョン(英国及びEU)があり、2002年1月現在で、UK GovTalk上でコメントを集めている。

e-GMFにおいて2001年にメタデータ利用のメリットについて検証したところ、非常に良好な結果が得られた。このため、政府機関においてはメタデータ付与を義務付けているが、すべての機関が使っているわけではない。

上記のメタデータ利用メリットの検証については、具体的には以下の5つの標準検索フォーマットを作って国税庁、地方自治体、e-GMFで検証が行われた。

全自治体の名前

生後6ヶ月の子供がいる場合に親のパスポートに併記することが可能か(もしくは独立して必要か)

PITO (Police Information Technology Center) とは何か(警察関連の団体)

ペットをEU圏内の旅行に帯同した場合に動物用パスポートがあれば検疫免除になるかどうか(昨年、当該法令が施行)

インターネット経由でのオンライン投票実施時期

という5つの項目について、役所で働く50人を選び、メタデータ未使用時と使用の双方について、「検索エンジンは何を用いたか」「検索に成功したか」といった質問をした。また、検索に要した時間、答えの正しさ、結果としてどこまで進行したかなどを調べた。

メタデータの自動生成システムについては、2002年1月現在で国税庁のWebサイトとロンドンのルイシャムという自治体のサイトがメタデータを付与して運用しているが、これらはEDSが開発したメタデータの自動生成システムを使用している。

省庁がメタデータを付与することについては、法律による強制力はないが、ある省庁がこれを無視した場合にはIT関連の予算配分を却下する権限を、内閣府はGuidance(行政指導)という形で持っている。

メタデータ付与が強制されるのは政策、書簡といったすべての政府発行情報についてのみであり、その目的は以下の通りである。

国民による政府情報へのアクセスの容易化

電子データ管理の容易化

データの自動処理の容易化(一部)

メタデータを付与した省内の公文書に関する情報のレポジトリについては、現在はPublic Document Officeが主導する形で各省庁が個別に対応している。

紙の通信文を電子化するような部署はなく、メタデータ付与の対象はすでに電子化された情報のみである。

英国には公文書電子化についての法律がある。2000年末に法案が通過している。この法律により電子ドキュメントにも効力が認められるようになった。これに加え、女王による署名も電子的に行なうことが可能になった。

現状ではWestminster (=英中央政府、England、Wales)政府のみがe-GMFを採用しており、Scotland政府は採用を検討中である。Northern Irelandは特に検討していない。

である。

⁴² 本報告書の付録2と付録3を参照のこと。

2.2.3.2 スウェーデンの動向

(1) スウェーデン概要

スウェーデンでは、北極圏に入っている北部は人口密度が著しく低く、南部の主要都市に人口が集中する傾向があり、全人口の半数以上が、日本と同規模の国土面積中 3%以下の地域に住んでいる。このような自然環境ゆえ、道路や鉄道などの交通網整備には時間とコストがかかるため、より安価に情報だけでも伝達できる電信、電話網は、発明から時をおかずしてスウェーデン社会に広く普及していた。

現在では世界の IT 立国として高い評価を得ているスウェーデンであるが、80 年代後半から 90 年代初めにかけては、「バブル経済とその崩壊」を日本とほぼ同様に経験していた。金融自由化と高い輸出競争力から、80 年代後半にかけて世界的に資本が流入することで景気が過熱、ついに 90 年代初めにバブルが弾け、深刻な景気後退に見舞われた。

しかし、90 年代半ば以降、IT 産業を基幹産業として育成することに成功し、「デジタル経済」へとアメリカ以上に見事に対応することで、バブル経済から立ち直った。

政策面では、バブル経済崩壊後、社会全体が情報・知識の集積・集約に基づく産業活動、社会的活動中心に移行したとの認識を明確にし、情報ネットワークを基軸に社会を再構成することを政策の中心的課題とした。このことは、1994 年に当時のビルト首相が表明した、「万人が、素早く、容易に、安全に、安価に、時と場所を選ばずに、情報を電子的に引き出し、互いにコミュニケーションできるスウェーデンを作る」とのメッセージに具現されている。

1996 年には IT 立国のビジョンを示した「IT 法案」が制定され、Confidence (信頼)、Competence (能力)、Accessibility (アクセス) という情報ネットワークに関する三条件を市民に確保することが重要であるとの見解が示された。こうした基本方針が示され、施策の立案、実行が進んだが、最も影響力の大きな施策が実施されたのは 1998 年であった。

まず、第一に「パソコン法案」(企業が PC および周辺機器を購入し、従業員が家庭で使えるようにリースするプログラム)が成立し、家庭のパソコン保有率が 48%から 67%へと急激に上昇した。特にブルーカラー世帯でのパソコン保有が進んだという。

次いで、初等教育、中等教育の教員 10 万人のうち 4 万人に無償でパソコンを配布する 3 年間のプロジェクトが開始されたことである。これは、一般家庭へのインターネット普及が進むにつれ、父母と子供たちに教員と電子メールで連絡をとることへの要求が強まったことが大きな影響を与えている。

スウェーデンは上記のようにデジタル経済を成長させ、失業率では、98 年以降毎年 1%ずつ低下させることに成功し、2000 年にはついに日本を逆転した。財政面では、93 年以降累積債務も増やさず、景気回復による税収の増大で 98 年から財政収支は黒字に転じ、累積債務も相対的に減らしている。

(2) スウェーデン政府における情報化の概要

情報化の進展を 4 段階で表現すると、Information Interaction Transaction Integration となる。スウェーデンも他国と同様、情報化は緒に着いたばかりであり、Information の段階にある。

スウェーデンの e-government ポリシーとしては、以下の 2 つがある。

「24/7 e-government」（24 時間 365 日いつでもアクセスが可能という意味）

「ライト・タッチ」（中央省庁が地方やエージェンシーに対して強制的な影響力を振るわない、振るえないという意味）

中央政府（Central Government）はゴールだけを決め、課税自主権を持った県（Regional Government）や市（Local Government）レベルが独立して行政を執行する。

（3）行政情報へのメタデータ付与の具体的事例

労働市場局（National Labour Market Board）

企業や政府機関などの求人サイドと求職中の個人との間で行われる求人情報の掲示やそれに応募するにあたっての履歴書作成・提出といった 1 件あたり 2、30 分の処理時間を要し年間 50 万件に上る処理件数をメタデータ付与によって効率化し活発化させることを目的とする。

構築にあたっては HR-XML⁴³のフレームワークを用いた。RDF を使わなかった理由は、たまたま HR-XML があったから使ったというきわめてプラグマティックな理由による。

高等教育庁（National Agency for Higher Education）

スウェーデン国内に 40 の大学があり合計で 14000 のコースが存在する。これをいちいち個々の大学のホームページに見に行く代わりに、ワンストップで検索・閲覧可能なコースカタログを国が作ろうとする試み。

システムの仕組みは国が立ち上げたセンターデータベースに置かれたソフトウェア・ロボットが、各大学が作るデータベースを巡回し、各校のデータベース内のメタデータを収集して回り、センターデータベースを更新していく。各校のメタデータはダブリンコアに加え 5 つの拡張エレメントを採用している。

ライト・タッチの方針では各校にメタデータを付与させるのが難しいとも考えられるが、スウェーデンの教育機関はすべて国立といってよく、国からの予算配分は学生数に応じて行われているため、各校に共生するまでもなくセンターDB に自校の情報が反映されることが死活問題となる。したがって、導入に先駆けて各校の担当者を集めて説明を行なったところ、各校がスムーズに理解してくれたという。

メタデータの付与を行なうためには膨大な工数を要する。各校は具体的にどのように予算や工数を捻出したのかというと、大手 25 校は元々データベースを持っていたため、フィルターを作り一括して移した。フィルターの開発とメタデータ付与は各校の IT 部門が行ない、そのための開発予算は各校の IT 予算から支出された。中小 15 校はゼロから作った。

これらの具体的事例から得られる結論として、スウェーデン ICT 委員会は以下の 2 つ

⁴³ <http://www.hr-xml.org>

を挙げている。

セマンティクスはシンタクスよりも重要である。

- 国が何でもやらなければならない

「自分が持っているものから始めよ」

- HR-XML、ダブリンコア、iCal

また、失敗から学ぶための事例もある。商用車用の運転免許証発行システムの事例である。スウェーデンでは県単位で発行認可手続きを行っており統一フォーマットがない。そこで政府が資金を拠出して”SHS”(スウェーデン語で Spreading & Picking up System) というシステムを作った(取り扱い窓口は郵便局)が、国際標準の SOAP や UDDI といった Web サービスに準拠していないのが問題となった。今度開発する e-ID では同じ轍を踏まないようにすることが必要であるという。

2.2.3.3 アイルランドの動向

(1) アイルランド概要

北海道とほぼ同じ面積に人口わずか 360 万人あまりのアイルランドは、農業従事者が 90 年代後半でも労働人口の 1 割を占めており、主要資源が鉛、泥炭、ジャガイモなど、主要産業が食品加工、織物、自動車組立てといったように、高付加価値を創出できる産業構造が歴史的に見てもなかった。

そこで、73 年の EEC (欧州経済共同体) 加盟以降、徐々に海外との競争力の低さがアイルランド経済を悪化させることになる。80 年代前半にはインフレが年率 20% 程度に跳ね上がり、それに反比例して社会の雇用力は急激に低下、80 年代半ばには失業率が 18% にまで達するほど深刻な不況に見舞われた。

この事態を克服するため、90 年代に入りアイルランド政府は、外資導入、情報通信分野への重点的投資と人材育成などの施策を実行に移し、アイルランド経済は見事に立ち直った。

1992 年から 1998 年までに 20 万人以上の雇用が創出され、失業率は 1993 年の 15.9% から 1998 年の 7.8% にまで減少した。1994 年から GDP は年率 7% で上昇し、国家財政赤字は 1986 年に GDP の 122% に達していたものが、1998 年には 55% へと改善した。

1998 年にアイルランドは米国を抜いて世界第一位のソフトウェア輸出国へと成長した(米国: 30 億ドル弱、アイルランド: 33 億ドル弱。アイルランドは欧州各国にとってのオフショアソーシング先として成長してきた)。これは 90 年代の IT 政策の賜物であり、1100 社以上の外資を導入することに成功した結果といえる。

ここで重要なのは、90 年代の IT 政策を中核とするアイルランドの政策は、1987 年から、労働者、雇業者組織、労働組合、企業、市民などの社会的主体間で結ばれている「社会協約 (Social Partnership Agreement)」の精神に基づく点である。それは 1987 年の『国家再生計画 (Programme for National Recovery)』に始まり、アイルランド国民が戦略的目標を定め、それぞれがどのような役割を担うかを認識し、貧困と不平等に一致して立ち向かうことを目標にしたものである。そこでは、競争を無視することなしに、適切な賃金の上昇と雇用の増大を社会で分かち合うことが謳われている。

(2) アイルランド政府における情報社会実現に向けた政策概要

90年代に国の競争優位確保のために、首相主導でいくつかのイニシアティブができた。内閣官房では様々なエージェンシー横断の e-Strategy を策定するためにハイレベルの組織編制が行われ、歳入、土地登記、教育などに関するイニシアティブが結成された。

2000年7月10日に「電子商取引法 (eCommerce Act) 」がマッカーリース大統領署名によって制定され、ビジネス向けの電子商取引に適した環境整備と電子的な政府サービスの提供を可能にするフレームワークが整備されることになった。

具体的なイニシアティブとしては、以下のものがある。

「 OASIS (On-line Access to Services Information and Support) 」

2000年11月16日にプロトタイプ版の Web サイト⁴⁴の運用が開始された。国民への情報提供を目的とする。ポキャブラリ管理の実現可能性を検証する。

「 BASIS (Business Access to State Information and Services) 」

2001年5月4日に本プロジェクトの第1モジュールが完成した。企業向けの情報提供を目的とする。

「 Metadata 」

中央政府および地方政府の公的サービスへのメタデータ適用を目的とする。

電子政府を推進する EU 各国とのベンチマークの結果、アイルランドが No.1 という結果が出ている。

電子政府の全体政策については、首相直属の機関 (首相府) が策定しており、電子政府政策の立案・普及促進と法律問題について所管している。電子政府政策のほか、情報社会のあり方といった問題も研究している。

(3) メタデータ・イニシアティブ

2000年10月に開始された、公共サービスへのメタデータ適用を目的としたプロジェクトである。ポキャブラリ管理、タクソノミー構築などを行なう。

ダブリンコアのエレメントセットを採用し、オーストラリア政府の策定した AGLS (Australian Government Locator Service) の追加エレメントも採用している。

プロジェクトの最大目的はリソース・ディスカバリーである。

2001年3月に公開された「メタデータ --- 共通の標準を目指して⁴⁵」は、アイルランド公共サービス・メタデータフレームワークのドラフト文書であり、オンラインの情報資源にメタデータを付けるための標準案を提示している。メタデータの普及により、利用者は公共サービスの Web サイト上で入手できる情報をより容易に、より迅速に見つけることができることとされる。

同文書の主眼は、情報とサービスをオンラインで提供する公共サービス機関に対し、アイルランド公共サービスメタデータ標準の基盤としてダブリンコアのメタデータエレメントを適用することを勧告することである。

⁴⁴ <http://www.oasis.gov.ie>.

⁴⁵ <http://www.gov.ie/metaconsult/ipsms.rtf>

同文書は公開され、文書中で提示される複数のプロポーザルに関して、関連する組織からのパブリックコメントが求められた。アイルランド政府のメタデータ・コンサルテーション・グループは最終標準を承認する前に、すべてのコメントを吟味することになっている。パブリックコメントの提出期限は 2001 年 5 月 4 日であった。パブリックコメント集は http://www.irlgov.ie/metaconsult/consultation_responses.htm で公開されている。

メタデータを適用している分野については、プロジェクトの主要目的がリソース・ディスカバリーであるため、Web ベースで記録されたデータへのメタデータ付与がメインである。使用しているツールについては、非常に初歩的なものしか作っていないという。

英国では公文書へのメタデータ適用の効果を検証しているが、アイルランドではメタデータの適用自体が非常に初期的な段階にあるので、具体的な検証は行なっておらず、むしろ国民に対しレガシーデータをデータ化することの意義について教育することに重点を置いているという。

メタデータを適用して政府関係のデータベースを横断的に検索できる検索サービス（Infoseek ベース）を構築する予定もある。

2.2.4 欧州におけるオントロジー技術の動向

2.2.4.1 オランダのフライエ大学

(1) フライエ大学の研究状況

フライエ大学でのセマンティック Web 関連研究は、Artificial Intelligence 学科及び Business Informatics 学科の 2 ヶ所で行われている。AI 学科では基盤技術の研究を、BI 学科では応用の研究を、強く連携を取りながら行なっている。

フライエ大は本分野では比較的先進的な方であり、他には、カールスルーエ大学、英マ国家マンチェスター大などが先進的である。その中でも最も早く手がけている。その他、ブリストル大学は、他大学だけでなく英国 HP 社とも連携して良い研究を行なっているという。

欧州の第 6 次フレームワークプログラムのみならず、トランスアトランティックな研究としては、米国 DARPA と DAMIL + OIL の研究を行なっている。OIL はアムステルダムで開発され、W3C に提言を行なっている。

フライエ大学と共同研究している企業については、初期は Administrator のような大学からスタートアップした小さな会社と共同研究開発を行なってきたが、最近では、BT、ドイツテレコム、ダイムラークライスラー、スイスライフ、オラクル、フィリップスのような大きな企業も参加するようになっている。

(2) オントロジー研究開発状況

フライエ大学ではオントロジーの応用領域として以下の 3 つを検討している。

- (1) 知識管理：電子政府も検討されたが、民間への適用を中心に検討中である。BT、スイスライフなど大企業と共同研究をしており、大規模なイントラネットでの知識流通のための研究を行なっている。
- (2) 文書管理（検索エンジン）：現状の検索エンジンには様々な限界がある。Web のグローバル規模での文書共有の世界で、より知的に管理できるようにしたい。但し、より知的に管理するためには、企業内での利用も重要である。
- (3) 電子商取引(EC)と Web サービス：OBELIX のプロジェクトがある。電子商取引や Web サービス関連の研究では、ORACLE とも提携している。

多くのプロジェクト（およそ 10）があるが、主なものは以下

OnToKnowledge プロジェクト：初期のプロジェクトで OIL と Open Sesame などのツールを作成した。

IBROW：サービス統合 / データ統合のために、Web 上のコンポーネントの知的 Brokering サービスを行なう。

SWAP：セマンティック Web と PtoP 技術を統合する（中心サーバをおかずに直接交換する）。

EU のプロジェクトとして提案を予定しているおり、2002 年 1 月現在検討中である。

WonderWeb：オントロジーマッピングを研究する。複数のオントロジーの対応をとるもので、オントロジーの翻訳に近い。

OntoWeb：

研究プロジェクトではなく、普及のための Web サイトである。大学、企業などとの連携や、ワークショップや会議の開催などがある。OntoWeb については 2.4.4.2 節を参照のこと。

(3) Web サービスや人工知能 (AI) との関係

セマンティック Web と Web サービスの違いについては、フライエ大学の研究者の見解では、Web サービスはセマンティック Web の階層図 (図 2-1 参照のこと) の中で見るとプロトコルや情報バイディングなど低レベルの技術的な統合はできるが、意味的統合はできない。Web サービスは、内容に基づく事はやっておらず、現在の活動は重要であるが、拡張が必要である。オントロジーを導入して、コンテンツに関する意味的な統合を行なうことが重要であるとのことである。

また、電子商取引ではオープン性が特徴の PtoP アプローチがある。しかし、スケールアップサポートが必須であり、ロゼッタネットのようなメデイエーション機能が必要となる。このために、セマンティック Web が必要となる。ビジネスロジックのような多様な動的な構造に対応するためには、データ発見や変換、コンフィグレーションなどを用いて支援する仕組みが必要であるとのことである。

セマンティック Web と人工知能 (AI) との違いについては、フライエ大学の研究者によれば、以下の 2 つの違いがあるという。

(1) オントロジーの交換ができる点。

昔 (AI 研究が盛んであった頃) はゼロから作る必要があった。現在は他のオントロジーを再利用し改良することが現実的な状況になっている。

(2) セマンティック Web のオントロジーは、AI よりももっと単純で、構造は簡単で (場合によってはタクソノミーレベル) あり、そのかわり量は多い。

オントロジーは歴史的には AI から出たが、現在はコミュニケーションや自然言語システム、電子商取引など様々な場面で用いられ多様化しているので、オントロジーでは厳密な (厳密な定義による / 厳密な理論的背景を持つ) オントロジーに限らない。

2.2.4.2 ドイツのカールスルーエ大学

(1) カールスルーエ大学の概要

カールスルーエ大学には AIFB (Applied Informatics and Formal Description Methods) と FZI (Research Center for Information Technologies) という 2 つの研究拠点がある。

AIFB は、セマンティック Web と人工知能関連の研究を行なっている。オントロジーに関しては、欧州最先端の研究所であるとの自負がある。

カールスルーエ大学では、商用化を意識して研究活動を行なっており、開発しているシステムは、商用に十分耐えるものである。AIFB で研究開発した先進技術の産業界への移転は、FZI を通じて行われている。

セマンティック Web 研究開発のキーパーソンであるオランダ・フライエ大学の Dieter Fensel やスタンフォード大学の Stefan Decker もカールスルーエ大学の出身である。

欧州におけるオントロジー及びセマンティック Web 研究のグループには、カールスルーエ大学のグループ(オランダのフライエ大学はこのグループに属する)、英国のマンチェスター大学のグループ、及びイタリアのピサ大学のグループがある。

研究プロジェクトには、EU のプロジェクトとドイツ独自のプロジェクトとがある。以下に AIFB が参画する研究プロジェクトを挙げる。

- 1) セマンティック Web 関連
 - ・ Wonderweb(EU)
 - ・ PHT などの教育プログラム(EU)
- 2) 知識管理関連
 - ・ OntoKnowledge(EU)
 - ・ OntoLogging(EU)
 - ・ Consenser(BMBF)
 - ・ Ontowise(DFG)
 - ・ SWAP(EU)
- 3) ビジネス・インテリジェンス関連
 - ・ ドイツテレコム(DT)
 - ・ クライスラー
- 4) e-learning 関連
 - ・ PADLR
 - ・ Saxony(BMBF/L3S)

(2) 開発ツールの紹介

カールスルーエ大学で開発しているオントロジー関連のツールに OntoEdit と KAON (Karlsruhe Ontology) がある。

OntoEdit

3 . 2 . 3 節を参照のこと。

KAON

KAON は、オントロジーとセマンティック Web の利用基盤であり、RDF をベースとしたオントロジーとメタデータのツール群である。Java,J2EE 及び python を用いて作られている。次の三階層のツールから構成されている。

- (1) アプリケーション&サービス層 (オントロジー作成及びウェブポータル作成ツール等) :
アノテーションツールやオントロジーエディタ、オントロジー/メタデータクローラ、既存データベースのデータを RDF に変換するツール、及びオントロジー/メタデータから HTML データを生成するツールなどがある。
- (2) ミドルウェア層 (API 及びサーバ等) :
RDFS データやオントロジーデータを操作するための API、RDF データサーバなどがある。

(3) データ&リモートサービス層 (P2P、RDB 等) :

データ&リモートサービス層は、KAON を用いて実現されるサービスである。

2.2.5 欧州におけるセマンティック Web 関連システムの動向

2.2.5.1 イタリアの EU 委員会共同開発センター (JRC)

(1) JRC の概要

JRC は、EU 委員会 の一部であり、20 年前に設立して、いろいろなテーマに取り組んでいる。Web 発祥の研究所である CERN(欧州原子核共同研究機関)とは直接関係ないが、CERN 出身の人が多い。

JRC 内部の組織としては、原子力、エネルギー(核、プルトニウム)、安全保障、環境、健康・食品衛生、シンクタンク、情報社会などの分野がある。この中で EU の第 5 次フレームワーク・プログラムは、情報関係など、色々な研究に出資を行っており、この中で、セマンティック Web 関係のテーマも扱われている。

W3C が技術規格を策定する所であるのに対し、JRC はソフトウェアを作って実践的に行なう所という位置付けを行なっている。

(2) セマンティック Web 関連のシステム

ETB (EuropeanTreasuryBrowser)⁴⁶

2.4.2.2 節を参照のこと。

RDFStore (PerlAPI for RDFStorage)⁴⁷

RDFStore のプロジェクトは、2001 年 10 月から開始された。RDFstore は、RDF の構造を解析、格納、管理する Perl のライブラリーである。これは、RDF の主語・述語・目的語という 3 つ組 (図 2-2 参照のこと) の検索を効率的に行なうことを目的としている。検索にはコンテキストが重要であり、コンテキストも考慮した検索が可能である。3.2.1 節を参照のこと。

2.2.5.2 フランスの国立情報処理自動化研究所 (INRIA)

(1) INRIA の概要

2002 年 1 月現在、85 人の研究者、300 科学者 (ポストドクターや海外からの客員研究員) のメンバーで研究を行なっている。W3C の欧州におけるホスト機関である⁴⁸。30 個のプロジェクトが進行しており、以下の 4 つの大きなテーマに大別できる。

ネットワークとシステム

ソフトウェア工学シンボリックコンピューティング

ヒューマンインターフェイス、画像処理、知識システム

複雑系システムのシミュレーションと最適化

これらのプロジェクトからスピノフしたものはいくつか (ISTAR, RealViz など) あり、研究内容が実用化されている。セマンティック Web 関連のプロジェクトとしては、ACACIA があり、これは上記 に含まれる。この ACACIA プロジェクトでは、企業にお

⁴⁶ <http://etbrowse.jrc.it>

⁴⁷ <http://rdfstore.sourceforge.net>

⁴⁸ W3C の米国におけるホスト機関が MIT、アジアにおけるホスト機関が慶應義塾大学 SFC 研究所である。

ける知識管理の方法論とツール作成を研究している。

次世代の電子政府のイメージとしては、市民が必要なフォーマットに書きこめば、必要な情報を得られるというものであり、中央政府と地方政府の情報共有が主な目的だとしている。

(2) セマンティック Web 関連のシステム

ACACIA

2 . 4 . 1 . 1 節を参照のこと。

RDF-PIC

RDF-PIC は、画像データにメタデータを RDF 形式で付加することにより、画像データの検索の際に、実際の画像を検索することなしにメタデータのみで必要とする画像データを獲得する試みである。これによりネットワークの負荷を軽減することが可能となる。

2.3 標準化機関の動向

Web 技術の国際標準化機関である W3C (World Wide Web Consortium) の活動状況について調査を行なった。

2.3.1 W3C の Semantic Web Activity

2001 年 2 月 9 日に Semantic Web Activity グループが設立した。従来の Metadata Activity グループ⁴⁹からの名称変更である。2003 年 2 月に終了予定である。

Semantic Web Activity は、セマンティック Web の実現のために、アプリケーション間のデータの自動化、統合、及び再利用を支援する技術の共同開発と標準仕様化でリーダーシップを担うために設立された。

W3C の Semantic Web Activity は DARPA の DAML プロジェクトから活動支援を受けている。また、EU 委員会からの資金提供を受けている。⁵⁰

Technology & Society Domain⁵¹に属する。Web の技術が社会で実際に利用されるためには、単に技術的な問題を解決しただけでは不十分であり、社会的な様々な問題を解決しなくてはならないが、Technology & Society Domain ではこのような内容に関する活動が行われている⁵²。

Web オントロジー・ワーキンググループ (Web Ontology Working Group, WOWG)

2001 年 8 月に立ち上げられた新しいワーキンググループである。WOWG の活動目的は、各研究開発コミュニティ間のデータの統合と相互運用性を提供するような、構造化された Web ベースのオントロジーを定義する単一の言語を開発することである。

2.3.2 WOWG のオントロジー記述言語の動向

W3C の WOWG では、オントロジー記述言語の標準化検討を実施している。W3C において検討されているオントロジー記述言語としては、2001 年 12 月にノートとして提出された DAML+OIL と、2002 年 1 月の FaceToFace Meeting で提案された DAML+OIL の後継言語である OWL(The Web Ontology Language)がある。OWL には、まだ、詳細な仕様の議論がなされていないので、本資料では、(1) で DAML+OIL の仕様概要について⁵³説明した後、(2) では、現時点での OWL の概要について⁵⁴述べる。

⁴⁹ <http://www.w3.org/Metadata/>

⁵⁰ <http://www.w3.org/Consortium/>

⁵¹ <http://www.w3.org/TandS/>

⁵² 齋藤信男「次世代 Web コンピュータにおけるセマンティック Web 等関連技術動向」(2001 年 3 月) より。

⁵³ Fikes,R and McGuinness,D,: An Axiomatic Semantics for RDF, RDF-S, and DAML+OIL (March 2001) 18 December 2001, Richard Fikes, Deborah L. McGuinness (DAML+OIL Web Ontology Language Submission) <http://www.w3c.org/TR/daml+oil-axioms>, Horrocks,I, Harmelen,F, and Patel-Schneider,P 編集, A Model-Theoretic Semantics for DAML+OIL (March 2001) <http://www.w3c.org/TR/daml+oil-model>, 及び Horrocks,I, Harmelen,F, and Patel-Schneider,P 編集, DAML+OIL (March 2001), <http://www.daml.org/2001/03/daml+oil-index.html> (2001)を参考にした。

⁵⁴ Peter F. Patel-Schneider: The Web Ontology Language (OWL) <http://lists.w3.org/Archives/Public/www-Webont-wg/2002Jan/att-0061/01-swol.text>を参考にした。

(1) DAML+OIL

DAML+OIL は、 DARPA からの資金援助を受けた DAML プロジェクトで開発されていたオントロジー記述言語 DAML-O と、 EU 委員会 IST プログラムからの資金援助を受けた On-to-Knowledge プロジェクトで開発されていた OIL (Ontology Inference Layer) を統合した言語である。 2001 年 10 月 13 日時点での最新版は、 2001 年 3 月の Ver4.1 である。

DAML+OIL は、 RDF スキーマ (RDFS) をベースとしてクラスやプロパティに対する制約の記述能力をさらに向上させた言語であり、言語の数学的基盤である意味論を明確に定義していることを特徴としている。 Ver4.1 では、 RDFS に対して、「クラスのブール結合演算や排他的関係」や「プロパティに対する制約の拡張」などの記述能力の拡張がなされている。以下、順に説明する。

クラスのブール結合演算や排他的関係

クラスに対する以下の演算や関係を表現することができる。

- ・二つのクラスのいずれかに属するというクラスを定義する演算(unionOf)
- ・二つのクラスのいずれにも属するというクラスを定義する演算(intersectionOf)
- ・あるクラスには属さないクラスを定義する演算(complementOf)
- ・二つのクラスに共に属する要素がないという関係(disjointWith)
- ・二つのクラスが等しいという関係(sameClassAs)

図 2-4に、 DAML+OIL の記述例を示す。以下、各々の記述例を順に説明する。

- ・クラス Women の定義では、「女性(Women)は、男性(Man)でない人(Person)である」ことを意味する。
- ・クラス TallMan の定義では、「身長の高い男性(TallMan)は、高いもの(TallThing)かつ男性(Man)である」ことを意味する。
- ・クラス HumanBeing の定義では、「クラス HumanBeing とクラス Person は等しい」ことを意味する。

```
<!-- クラス Women の定義 -->
<daml:Class rdf:ID="Women">
  <rdfs:subClassOf rdf:resource="#Person"/>
  <daml:disjointWith rdf:resource="#Man"/>
</daml:Class>

<!-- クラス TallMan の定義 -->
<daml:Class rdf:ID="TallMan">
  <daml:intersectionOf rdf:parseType="daml:collection">
    <daml:Class rdf:about="#TallThing"/>
    <daml:Class rdf:about="#Man"/>
  </daml:intersectionOf>
</daml:Class>

<!-- クラス HumanBeing の定義 -->
<daml:Class rdf:ID="HumanBeing">
```

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

図 2-4 DAML+OIL の記述例 1

プロパティに対する演算や制約の拡張

RDFS では、プロパティのドメインとレンジしか規定できなかった。すなわち、プロパティをとるリソースや値が属するクラスしか規定できなかったが、プロパティに関する制約記述の向上を目的として、以下に示すような拡張がなされている。

- ・プロパティの値そのものを規定する制約(hasValue, hasClass)
- ・プロパティの値をいくつとることができるかを規定する制約(cardinality, minCardinality, maxCardinality)
- ・プロパティが推移律を満たすという制約(TransitiveProperty)
- ・プロパティのドメインの各々の要素に対して、その値が一つしかないという制約(UniqueProperty)
- ・二つのプロパティが等しいという関係(samePropertyAs)
- ・プロパティのドメインとレンジを交換したプロパティを作る演算(inverseOf)

以下、図 2-5に、DAML+OIL の記述例を示す。以下、各々の記述例を順に説明する。

- ・クラス Animal の定義では、「動物(Animal)は親をもち(hasParent)、かつ、親の数は2つである」ことを意味する。
- ・プロパティ hasChild の定義では、「プロパティ hasChild(子供をもち)と、プロパティ hasParent(親をもち)は逆の関係である」ことを意味する。
- ・プロパティ hasMom の定義では、「プロパティ hasMom とプロパティ hasMother は等しい」ことを意味する。

```
<!--クラス Animal の定義 -->
<daml:Class rdf:about="#Animal">
  <rdfs:subClassOf>
    <daml:Restriction daml:cardinality="2">
      <daml:onProperty rdf:resource="#hasParent"/>
    </daml:Restriction>
  </rdfs:subClassOf>
</daml:Class>

<!--プロパティ hasChild の定義 -->
<daml:ObjectProperty rdf:ID="hasChild">
  <daml:inverseOf rdf:resource="#hasParent"/>
</daml:ObjectProperty>

<!--プロパティ hasMom の定義 -->
```

```
<daml:ObjectProperty rdf:ID="hasMom">
  <daml:samePropertyAs rdf:resource="#hasMother"/>
</daml:ObjectProperty>
```

図 2-5 DAML+OIL の記述例 2 ([3]より)

(2) OWL(The Web Ontology Language)

OWL は、W3C の WOWG において、Peter F. Patel-Schneider を中心として検討されている言語である。また、WOWG の検討フェーズとしては、要求仕様を明確にしている段階と思われるが、以下では OWL の概要について述べる。

OWL の DAML+OIL からの基本的な変更点は、以下の 3 つである。

- (1) OWL の意味は、model theory for RDF⁵⁵と矛盾しないようにする。
- (2) OWL の構文は、the XQuery 1.0 and XPath 2.0 Data Model⁵⁶と矛盾しないようにする。
- (3) データ型の取り扱いを単純にする。

意味の変更の背景にある基本的なアイデアは、サブクラスのような意味論の観点から扱うべきコンストラクトを構文としてもつ RDF アプローチを採用するだけでなく、解釈における関係を明確にすることにある。しかしながら、意味的な矛盾を引き起こすような言語全体の拡張は避けなければならないので、意味的な矛盾が存在しないことを確信できるようにするという観点から、解釈における関係を明確にできるコンストラクトの数を制限する方針とのことである。

また、構文は XQuery Data Models なので、XQuery の標準的な実装を用いることにより、XML Schema のデータ型のほとんどすべてを提供できるようになっているとしている。

55 Patrick Hayes 、 RDF Model Theory 14 February 2002, <http://www.w3c.org/TR/rdf-mt/>

56 Mary Fernández、Jonathan Marsh、Marton Nagy : XQuery 1.0 and XPath 2.0 Data Model W3C Working Draft 20 December 2001. <http://www.w3c.org/TR/query-datamodel/>

2.3.3 OWLの利用事例

(1) コレクション・マネージメント

コレクション・マネージメントは大規模なデータ集合体（テキスト、イメージ、マルチメディア、Web ページ等）の管理のことで、ドメインへの依存性やスケーラビリティの問題を含んでいるため、OWLの有望な応用分野と考えられている。WOWGの2002年1月のFace-to-Face Meetingでは、次の5つの具体的な事例が紹介された。

Use Case 1: "ARKive: catalog of endangered species descriptions"

ARKive⁵⁷は英国 Wildscreen Trust が構築している絶滅危機生物のマルチメディアデータベースで、学校の生徒からこの分野の専門家までが、Web サイトを通じてアクセスできるようになっている。実際のデータベースシステム開発には HP 社が参画している⁵⁸。

Use Case 2: "EDS Web page landfill"

企業の社会とのコミュニケーションやコーポレートメモリの管理に関するもの。具体的な事例は、EDS 社が製品化している PLM(Product Lifecycle Management)ソリューションで、新聞発表、製品提供とケーススタディ、企業内手続き、製品説明と比較、ホワイトペーパー、プロセス記述等を対象としている。EDS 社の代表的な製品には Teamcenter⁵⁹がある。

Use Case 3: "Aerospace Engineering Data Modelling"

航空機産業における企業の社会とのコミュニケーションやコーポレートメモリの管理に関するもの。特に、製造工程管理が中心で、航空機設計ドキュメント、製造工程ドキュメント、テスト工程ドキュメント、メンテナンスドキュメント、イラスト等を対象とし、設計エンジニアが利用する。具体的な活用製品としては、Stilo Technology 社及び子会社の OmniMark Technologies Corporation 社が提供する Content Engineering ソリューション OmniMark⁶⁰がある。

Use Case 4: "Art-image collections"

絵画コレクションのデータベース構築において、オントロジーを利用して、データのインデクシングと検索をサポートすることを目的としたもの。オランダの University of Amsterdam の IBROW プロジェクト⁶¹（Web 上での情報ブローカリング技術の開発が目的）で取り扱っている応用事例。この分野では規約化に関する多くの既存のリソースがあり、ここでも VRA (Visual Resources Association⁶²) で定めている VRA Core Categories, Version 3.0 や、Getty 財団によって構築された AAT (Art and Architecture Thesaurus、芸術作品（芸術カテゴリ、材料、スタイル、カラー等）を記述するための 120,000 用語の

⁵⁷ <http://www.arkive.org.uk/default.asp>

⁵⁸ <http://www.hpl.hp.com/news/arkive.html>

⁵⁹ <http://www.eds.com/products/plm/teamcenter/>

⁶⁰ <http://www.stilo.com/casestudies/cs005.html>

⁶¹ <http://Web.swi.psy.uva.nl/projects/ibrow/home.html>

⁶² <http://www.vraWeb.org/index.html>

階層から成る)などを利用している。

Use Case 5: "Conceptual Open Hypermedia"

学術機関が構築する研究成果データベースのような巨大なハイパーメディア / Webページをブラウジングする際のナビゲーションを対象としたもの。英国 University of SouthamptonのAKTプロジェクト⁶³で取り扱っている事例を参考にしている。

これら 5 つのコレクション・マネージメント事例の分析から、OWL に対する要求事項として、次のような項目がリストアップされている。各項目の右側の数字は、上記のどの事例から導かれた項目であるかを示している。

1) 表現力に関する要求事項

- Class hierarchy with properties/attributes and relations ()
attribute と一般的な relation との明確な区別
- Default knowledge ()
 - 生き物は飛ばない
 - 後期ジョージ王朝風の整理だんすは通常マホガニー製である
- Part-whole relation ()
 - 翼の桁は翼の一部である
 - 整理だんすは独自のスタイルの脚をもつ
- Constraints ()
 - `wing-spar.length < wing.length`
 - `urniture.style = Late-Georgian <=> furniture.culture = British AND furniture.date-created 1760-1811`
 - 利用事例 4 における後の例に対する DAML+OIL の解決策についてのコメントにも注意。
- Instance specification / instances as classes ()
 - Mammal は Species のインスタンスであるが、またそれ自身クラスでもある。
 - A380 は Aircraft のインスタンスであるが、また A380 のインスタンス

⁶³ <http://www.iam.ecs.soton.ac.uk/projects/akt/>

集合も表している。

- ・参照： Protege-2000 の metclass 概念。注： これは RDF の特徴である。
- ・属性値による匿名インスタンスの仕様。
- ・ Mechanism for relation typing ()
さらに、RDF feature (property のために rdf:type を定義することができる。)
- ・ Abstract classes ()
インスタンスのないクラス： 応用事例 4 の color の階層の例を参照。
- ・ Synonyms / lexical term -> concept ()
用語とそれが表すコンセプトを区別する必要がある。 [これは OWL の問題ではないかもしれない。]

2) 他の要求事項

- ・ Provenance (1, 5)
典型的な例： 専門家と非専門家による表記を区別すること、ハイパーリンクによる出典参照。
- ・ Version management (1)
オントロジーを拡張・変更する能力。
- ・ Query support (1)
偽の仮定による推論を行なう能力、例えば「魚としての鯨」。
- ・ Support for content standards (1, 3)
一般的なシソーラス (WordNet、T G N) が、ドメイン固有のシソーラス (AAT、ICONCLASS) と同様、しばしば表記を標準化するために使われる。

(2) 一般的な利用事例

2002 年 1 月に W3C にて行われた WOWG Face-to-Face Meeting において議題となった一般的な利用事例を以下に説明する。なお、この利用事例はドラフトであり、今後はこれらの場面において OWL の利用が期待できるとするものである。

1) デバイス相互運用性

計算機能のあるデバイスでは、デバイスに意味が分かる (セマンティック) レベルで相互運用ができなければならない。こういったデバイスには、PC やラップトップ、PDA、プリンタやその他ネットワークデバイスなどが考えられる。これらのデバイスは旅行の計画などのような同じ目的で使用できる。例えば、ドメインへ適

切にアクセスする場合、オントロジーの同じ集合を使用できるが、デバイス特有の限定された低域幅や制限されたメモリなどを含む特定の要件にマップされたり、カスタマイズされたりするべきではない。

また、デバイスは動的に変化する環境に対してもセマンティック的に採用すべきである。例えば、新しいプリンタを規定のデバイスのネットワークに追加する場合や、PDA が異なる計算や、アクセス、そしてセキュリティー要件がある新しい場所に導入された場合などである。

更に、デバイスは Web サービスによって制御可能である。この Web サービスはオントロジーを使用してタスクの分解とコントロールフローを表示することができるようになる。

2) インテリジェントなエージェント

さまざまなエージェント基準にはオントロジー言語の利用が必要である。この言語は(複数のオントロジーからの)オントロジー表現交換と(あるオントロジーか他のオントロジーへの)オントロジー翻訳をサポートしなければならない。さらに、交換や翻訳されるオントロジーは異なるエージェントプラットフォーム、インフラストラクチャー、あるいはドメインで使用できる。

また、交換/翻訳されるオントロジーは異なる言語で表現可能であるし、異なるレベルであることもできる。例えば、上位オントロジーや下位のドメインオントロジーなどである。複数の上位オントロジーや上位オントロジーの一部が交換/翻訳、そして、判断されなければならない場合もある。同じ個人/インスタンス/データが異なるオントロジーにマップされなければならない場合もある。また、SQL、RQL などの場合、照会言語間でマッピングされる必要がある場合もある。

3) インテリジェントな管理

ビジネス組織、財閥の Web コミュニティーといった場面では、複数の組織の単位で Web 上でお互いに通信を行わなければならない。これらの単位には異なる OM (Organizational Memories) や、知識リポジトリが含まれることがあり、複数のデータベースや文書のセットをサポートしなければならないこともある。これらのセットはすべてこの異なるオントロジーで表示することができる。組織に対してさまざまなレベルがあることがあるため、データ/情報固有のドメインから大企業のレベルまで、さまざまな情報のモデルとそれに相当する語彙がある。

さらに、組織と組織がサポートしているオントロジーは動的に変化することがある。組織の中には、例えば財閥のコミュニティーなど、仮想的であり他の既存の組織の合成や分離であったりすることもある。しかし、こういった組織や組織の単位は、オントロジーマッピングを使用して、異種の語彙やモデルとセマンティック的に相互運用ができなければならない。必要であれば、これらのマッピングを見つけることができ、保持・更新も可能となり、破棄もできることがある。

4) インテリジェントな検索、電子商取引

Web上のB2B、及びB2Cの電子商取引では、買い手はアプリケーションやカタログ、新旧のB2BまたはB2Cエンジンやインフラストラクチャーのさまざまな範囲から将来の製品やサービスを探し、見つけ出すことができる。こういったリソースはすべて、分類や分類学、あるいはオントロジーによってモデル化されることがある。結局、それらはオントロジー言語を使用してモデル化されるのである。

従って、問題となる点は、Web上に異質にモデル化されたり、明示的にまたは黙示的に分類されたりする製品やサービス(そして、他のドメインのオブジェクト)をユーザーが検索できること、及びこれらのモデルや分類化システム、またはオントロジー間を概念的にマッピングすることである。

将来の買い手は、個々の会社のカタログや文書にさまざまに表示されている語彙、及び意味を使用して製品やサービスを探したいと考える。一方で、将来の売り手は、将来の買い手が探すことができるように、製品やサービスを分類し、適切にモデル化するのである。このような概念的な検索/ナビゲーション(分類)には、セマンティック的にマップされるオントロジーが必要である。

5) インテリジェントな統合(異種のオントロジー情報の統合)

WebやWebにアクセス可能なアプリケーションに表示されている情報は、ドメイン毎や、アプリケーションが解決しようとしている問題毎、あるいはモデルの複雑さの度合いに応じてさまざまな方法でセマンティック的に表示されることがある。オントロジーの概念とその概念が特徴付けている個人、インスタンス、データにはセマンティック的な合成・統合・分離の方法、及び、対応する構文上の形式の変換が数多く必要である。

この使用事例はインテリジェントなエージェントの使用事例と関連している。

6) 相互評価

例えば、国防総省の武器、ネットワーク、コマンドと制御システムといったドメインの例においては、多くの場合、特定基準のセマンティックモデル(オントロジー)の基準や適合検査の実現に関してその特定の基準のセマンティックモデル(オントロジー)において難しい点がある。これらの重要な三つのコンポーネントもすべて、お互いに相互関係なしに動的に変化するものであり、原則では、モデルで強制されべきである。基準と一致していると称するアプリケーションは、作成可能な一連の適合検査によって、その適合に関して検査すべきである。これには、モデルと基準と一連の基準の検査間でのセマンティックマッピングが必要である。

さらに、複雑な点としては、基準には異なるバージョンがあるので、モデルと一連の適合検査には異なるバージョンがあることを想定すべきということである。

7) 文脈と表示マッピング

センサー検知、CAD技術、マルチメディア作成といった場面においては、アプ

リケーションは、文脈や表示を元にオントロジーの情報をフィルタリングしたり、表示したりすることによって、ユーザーへのコンテンツ表示を合わせる必要がある。センサー検知と CAD 技術は、ユーザーが関心を持つ特定の表示を表すために文脈上の情報に適合する場面である。これらの表示は、例えば個人化やローカル化といった場合、アプリケーション表示の要件、または、ただ単に与えられた時間、あるいは照会への回答において、オントロジーの中の特定面に重点を置くという要求をもとに抑制やフィルタリングをする必要がある。こういった表示や文脈は合成されたり、操作されたりする必要があるため、オントロジーのアプリケーション理論として持続し、主要なものとなる。

8) 相互運用性

コマンドとコントロールを目的とした防衛関連のアプリケーション、軍事派遣計画、ロジスティックスなどは、サービスと組織ユニット間でマルチドメイン情報を使用し、普及することが必要である。多くの場合、オントロジーは明確で、セマンティック的にほとんど同じ概念に対しても異なる用語や意味を使っている。さまざまなサービスや組織間で命令や情報が行き交う場合や、異なる目標を達成するために同じ情報をしようする場合、セマンティックをできる限り保持するために異なるオントロジーをマップする必要がある。

2.4 活用事例と課題

セマンティック Web 技術、特にオントロジー技術の活用事例と課題について述べる。オントロジーの活用事例に関しては、着手が始まった段階で、事例といっても研究の方向性を示すものや、プロジェクト提案レベルのものに留まっている。現時点では、セマンティック Web の有用性の実験とそのために必要な要素技術の蓄積を行なっている状況で、今後の調査・検証が必要である。調査は主にプロジェクトを中心に行なったが、EU 委員会の PF5 の IST にセマンティック Web 関連プロジェクトがあるため、欧州のものが多くなっている。

セマンティック Web 技術の活用の方向性としては、以下のものが見られた。

- 1) ナレッジマネジメント
- 2) e-learning
- 3) e-ビジネス
- 4) コミュニティー支援

一つは、知識管理の直接的な応用として、ナレッジマネジメントがあげられる。以下の事例では明確に示していないが、統一されたオントロジー/システムを用いることができる企業内での応用と、その後に、個々のオントロジーを持った企業間同士でのマッピングの問題などがある。また、e-learning では、多種多様な教材の蓄積を活用するためにセマンティック Web やメタデータの技術が用いられている。

e-ビジネスへの適用は、オントロジー技術を用いて、取引関係の価値連鎖（バリューチェーン）を記述したり、部品のメタデータの組み合わせから顧客に適した製品を作る際の、カタログの自動生成や製品の自動分類などに活用しようとしている。

また、知識交換の場としてのコミュニティに関しては、セマンティック Web 技術を用いる

2.4.1 ナレッジマネジメント

ここでは、フランス INRIA のナレッジマネジメントシステムである ACACIA とその関連研究、及びナレッジマネジメントの一環としてドイツのダイムラークライスラー社の人材支援システム（HRMore）におけるセマンティック Web の活用事例について述べる。

2.4.1.1 ACACIA

ナレッジマネジメント関連の研究では、フランスの 国立情報処理自動化研究所（INRIA）における ACACIA（Knowledge Acquisition for Explainable, Multi-Expert Systems）⁶⁴プロジェクトがある。ここでは、ACACIA プロジェクトと関連技術研究の概要について述べる。

ACACIA は、企業情報構築・管理・普及のために、暗黙知を形式知として具現化し、情報の共有と再利用を目指す、ナレッジマネジメントシステムである。このためのセマンティック Web のアプローチとしては、まず、企業情報の蓄積のために XML 文書を作成し、オントロジーに基づいた意味的なタグ付けを行ない、Web 上で配布する。タグ付けは RDF

⁶⁴ <http://www.inria.fr/recherche/equipes/acacia.en.html>

を用いて行なうが、その蓄積と検索のために、CORESE (Conceptual Resource Search Engine) という検索エンジンを開発している。これは、RDF の検索エンジンのプロトタイプであり、グラフ表現への変換や RDF 表現の検索を可能にする。もちろん、グラフ表現をそのまま検索の質問として入力することも可能である。推論エンジンはグラフ推論に基づくものを採用している。(図 2-6参照のこと。)

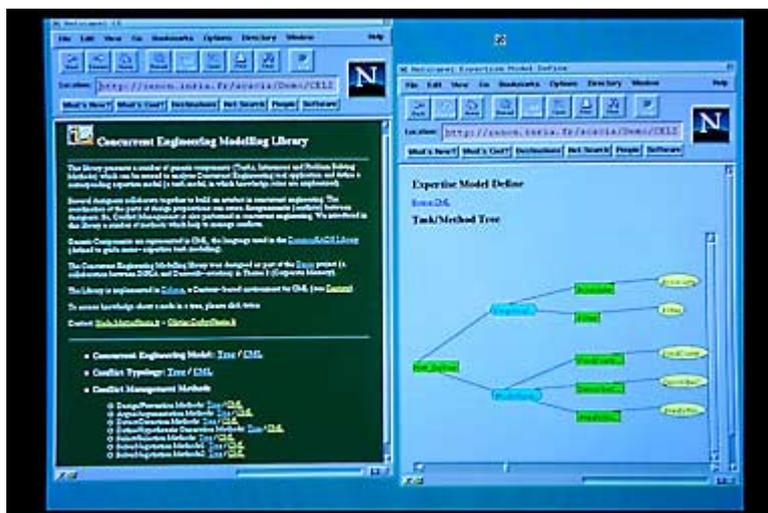


図 2-6 ACACIA における知識記述の例

次に、マルチエージェントの役割を担う、CoMMA(Corporate Memory Management through Agent)がある。CoMMA はオントロジーの作成・管理と CORESE のユーザーインターフェイスとなるマルチエージェントシステムから成る。これらのシステムを用いて新入社員用の企業情報の提供の応用システムを作成している。

さらに、知識表現やオントロジーはさまざまな目的で多種につけることが可能である。これらを再利用して新しい目的に活用することが考えられる。知識表現やオントロジー、タグ付けした情報の表現法として、いろいろな表現方法が考えられるため、それらの言語間の変換が必要となる。これを行なうのが EXMO(Computer mediated exchange of structured knowledge)⁶⁵である。

2.4.1.2 HRMore

ビジネス・インテリジェンス関連の適用事例としてダイムラークライスラー社の HRMore⁶⁶がある。HRMore プロジェクトでは、人材支援システム (Human Resource System) への適用を試みている。HRMore では、スキル活用にあたって、手書きの Job Application を OCR で読み取った記述内容やその他のドキュメントを HR システム用のオントロジーに基づき分析し、スキル情報を抜き出している。そのスキル情報を元にイント

⁶⁵ <http://www.inria.fr/recherche/equipes/exmo.en.html>

⁶⁶ <http://wim.fzi.de/wim/projects/hrmore>

ラネット上で、スキルの需要と供給のマッチングを行なう。実験システムは、OntoPrise社⁶⁷のOntoBrokerを用いており、2000年11月時点で27個の概念、56個の関係、4個の規則(公理)からなるProPerオントロジーが公開されている⁶⁸。ここでの対称となる「概念」は、「ドキュメント」やそのサブクラスである「プロジェクト報告書」や「HomePage」、その他、「人」や「分類」などの概念を定義している。例えば、「ドキュメント」は、ストリングタイプの値を持つ「URL」というプロパティや、「記述言語」、「作成日」、「著者」、「関連するスキル」、「関連する分類」などのプロパティを持つ。関係は、「ドキュメント」とその「URL」、「ドキュメント」とその「記述言語」といった関係を定義している。公理は、例えば[「HomePage」とその「著者(has_autor)」プロパティの値との関係と「作業家」とその「所有 HomoPage(has_Homepage)」プロパティ値との関係は、逆関係(109頁の脚注¹¹⁷「反対の関係」参照)になっている(「ドキュメント」Dの「著者」プロパティがPであれば、「作業家」Pの「所有 Homepage」の値はDである)などの規則を定義している。

2.4.2 e-learning

e-learningはIT技術を用いて教育を行なう事である。教育用メディアを電子化することによりネットワークを通じた配布や遠隔授業などが行なえる。教育用メディアは多種複雑であり、現状包括的で有効な内容検索手段⁶⁹は存在しない。このため、メタデータの付与が期待される分野である。

2.4.2.1 PADLR(Personalized Access to Distributed Learning Repositories)

e-learningへの適用も試みられている。PADLR(Personalized Access to Distributed Learning Repositories)⁷⁰と呼ばれるプロジェクトでは、セマンティックWebの技術を用いて、複数のモジュールから成る教育用メディアの交換基盤の構築を試みる。年度の関係でドイツとスウェーデンでは2001年4月から、米国スタンフォードでは2001年9月から始まった2年間のプロジェクトである。

現代では、科学技術の革新により、生活のさまざまな面においても学習することが要求される。このような環境の中で、教員や教材の不足が起こっている。特に学習者に合わせた教材を的確に効率的に提供することが求められる。このような背景を踏まえ、各機関にある教材などのリソースを相互運用することが重要となる。

実際、大学などの多くの高等教育機関では、教材など教育上のリソースを蓄積して効率化が図られている場合が多い。しかしながら、これらの教育上リソースの管理は、個別に行われることが多く、各大学などの機関がそのコントロールを手放すことはなく、単一の中央サーバに蓄積する形態は実現しないという。

このために、各組織が管理している教育上のリソースを相互運用(活用)する必要がある

⁶⁷ <http://www.ontoprise.com>

⁶⁸ <http://ontobroker.semanticWeb.org/ontos/proper.html>

⁶⁹ 例えば、テキストメディアの内容検索としては全文検索があるが、画像を含むと適切な内容検索手段が存在しない。教育用メディアは基本的にマルチメディアであるので、有効な内容検索手段が存在しない。

⁷⁰ http://www.learninglab.de/pdf/L3S_IWF-cori-4b.pdf

る。教材の共有に関連する話としては、「オープンアーカイブイニシアチブ」がある。ここでは最近、ライブラリー・サーバーからデジタル・ドキュメントに関するメタデータ情報を検索するために HTTP ベースのプロトコルを定義した。しかしながら、それはサーバー検索プロトコル及びその基本的なスキーマとしてはダブリンコアに基づいたもの（教育用に拡張した意義は大きい）に留まっており、十分とはいえない。

PADLR では、ピア・ツー・ピア（P2P）ネットワークに基づいたアプローチで、教育メディア交換用のネットワークを構築しようとしている。しかしながら、ナップスター（Napstar）や Gnutella のような純粋な P2P アプリケーションのアプローチでは不十分である。なぜなら、例えば、Gnutella 中のメタデータはファイル名及びパスに制限されており、「マドンナの Like a Virgin」のようにタイトルが付いて（その文字列が世間一般に知られて）いるファイルにはうまく動くかもしれないが、「代数入門、講義 23」のようにタイトル文字列だけでは中身がわからないファイルではうまく動かないからである。このため、PADLR では、P2P ネットワーク基盤として、Edutella と呼ぶオープンソースのシステムを開発／採用し、その基盤上にオントロジー技術の活用を試みる。

2.4.2.2 ETB (European Treasury Browser)

イタリアにある EU 委員会共同開発センター（JRC）では、セマンティック Web 技術の応用システムとして ETB (European Treasury Browser)⁷¹ という名のプロジェクトにおいて、先生が学校で教えるための教育プログラムを蓄積したデータベースをつなげて相互利用をはかっている。ETB の目的は、欧州域内の学校のための教育用情報資源のためのメタデータネットワーク基盤を構築することである。5～6 年前に XML ベースの教育用オーサリングツールを開発しており、これを利用して、学校にコンテンツの配信を行っていた。初期のものは、NetNews システムを用いて配信されていた。教育用情報資源の蓄積は、これ以前のものも含め古くからあり、欧州域内に 5～6 個の主要なリポジトリデータベースを構築されていたが、歴史的経緯より管理方法は統一されていなかったし、メタデータに抜けも存在する。スキーマの異なるデータベースのためには、RDF を使って項目の相互関係を定義している。教育関係のシソーラス（オントロジー）については、12 言語（EU 以外の言語も含む）から成る約 1,000 語のシソーラスを RDF を用いて作成している。ETB 向けのメタデータ標準はダブリンコアを拡張して利用している。これにより、例えば、英語とイタリア語の同義語をリンク付け、各々の言語から他の言語の同義語を引き検索に利用できる。ETB 自体の最新版は、2002 年 1 月現在、英語、フランス語、ドイツ語、イタリア語、スペイン語、スウェーデン語、デンマーク語、ギリシャ語の 8 つの言語に対応している。

⁷¹ <http://etbrowse.jrc.it>

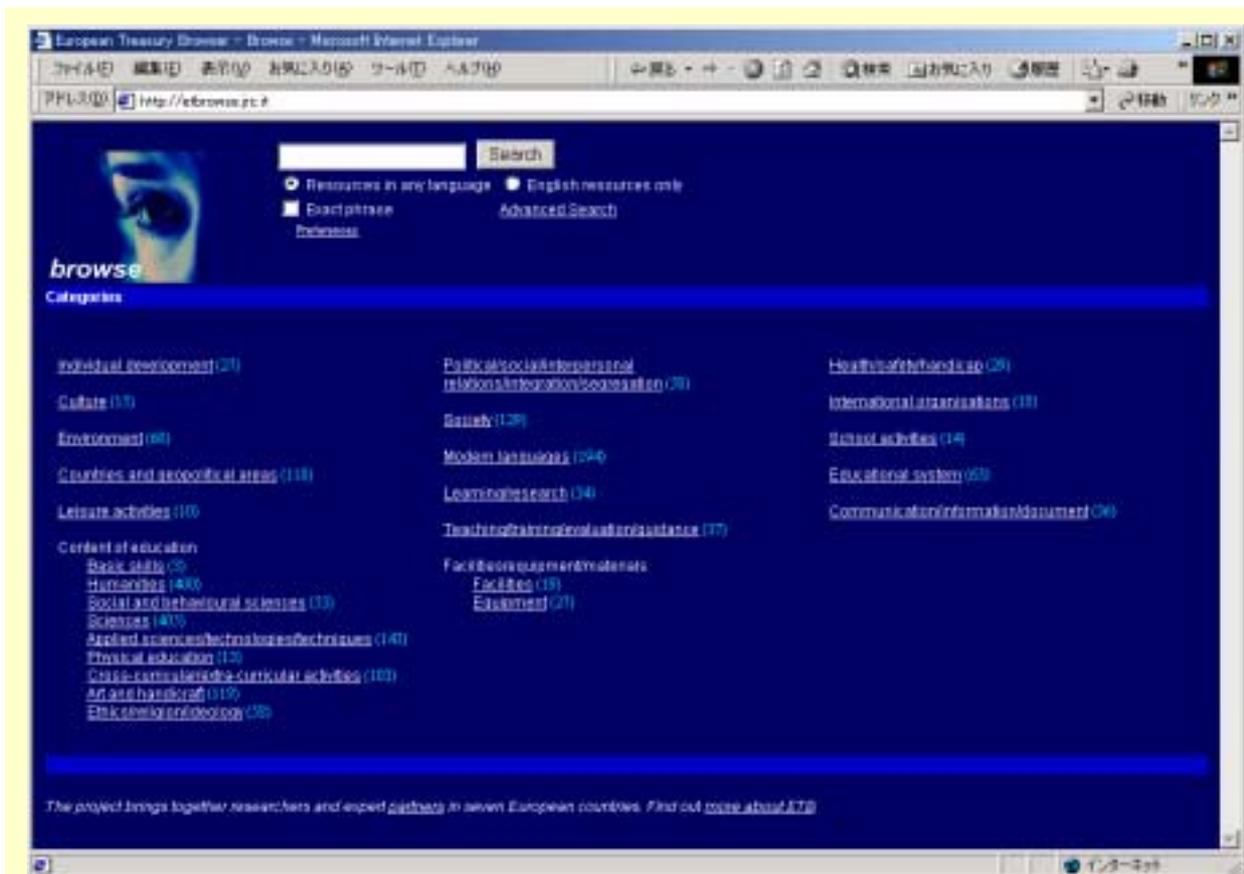


図 2-7 ETB の検索画面

図 2-7は、ETB の検索画面である。検索語として「Music」を入力し search ボタンを押すと、図 2-8のように検索結果が表示される。上部には、music を含むシソーラスの用語として classical music, contemporary classical music, modern classical music, music, music education, music listening などがあることがわかる。

JRC では RDF Site Summary (RSS)⁷²とその Viewer、さらに RDFStore の開発も行なっている。RSS は、手で書いたクエリー (RDQL/Squish⁷³という独自の形式：現時点では自動的に生成できない) から Web サイトの検索を行なうものである。Web サイトの RDF のデータはそれぞれに人間がつけて (自動化はこれから) RDFStore に格納し、検索ができる。現在はデモレベルで実用レベルではないが、セマンティック Web の応用システムを作って動かしている点は注目に値する。(図 2-9参照のこと。)

⁷² <http://java.jrc.it/rssfeeds/doc/scrapper.html>

⁷³ <http://rdfstoredemo.jrc.it/rdql/>

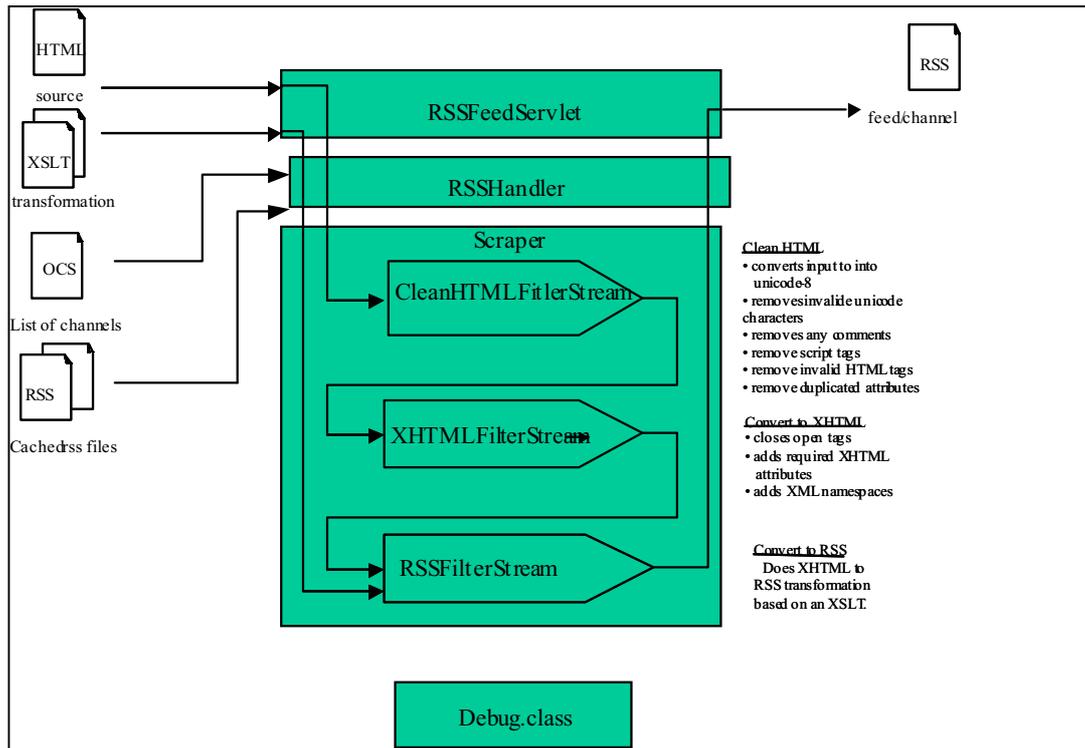


図 2-9 RSS の構成

2.4.3 e-ビジネス

セマンティック Web の目的の一つが、Web 上の情報のマシンによる理解と、それによる自動化とすれば、それをビジネスに展開することも考えられる。特に電子化された製品とその流通過程で、メタデータを活用することにより、どのような事が行なえるのかは興味深い。ここではそのような観点から、e-ビジネスへの適用の試みを行なっている OBELIX⁷⁴プロジェクトについて述べる。

2.4.3.1 OBELIX： e-ビジネスへの適用の試み

OBELIX⁷⁴は、EU 委員会の PF6 の IST プログラムに含まれるプロジェクトの一つである。スペインの LABEIN⁷⁵をコーディネーターとした 2002 年 3 月から 2004 年 8 月末までのプロジェクトであり、従って、2002 年 1 月の調査時には未だ提案段階のものであった。

OBELIX では、XML などの標準化により整備されつつある基盤の上に、よりスマートな e-ビジネスを構築しようとしている。特に、インターネット時代の電子商取引や協働作業には、動的なバリューチェーン (DVC: Dynamic Value Chain) が構成される。オントロジーの技術を活用して、この DVC をモデル化し、解析することを試みようとしている。特に、電子化された複合的な製品 (multi-product) に着目し、個々の部品に付与されたメタデータを、それらを組み合わせた製品に対する分類、設計、流通や解析などに利用しようとしているようである。

OBELIX では、以下のものを開発する。

⁷⁴ <http://www.ontoWeb.org/workshop/ucecOntoWeb/obelix.ppt>

⁷⁵ <http://www.labein.es/eng/index.htm>

- (1) 汎用のオントロジーサーバー及びオントロジーアプリケーションツール
- (2) 複数コンポーネント製品の分類ツール
- (3) 複数コンポーネント製品のコンフィギュレーションツール
- (4) e-ビジネスのモデリングとシナリオ解析ツール

その主な応用としては、以下の4つのものがあげられている。

- 1) デジタル音楽のバリューチェーン
- 2) エネルギー取引
- 3) イベントのオンラインでの企画支援
- 4) エネルギーサービス

例えば、デジタル化された音楽は、インターネット時代の e-プロダクトの典型的な例であり、マーケティングから販売、流通に至るまで電子化することが可能である。電子化されたことによりこれらの機能の担い手は再編成される。一方、著作権などの権利も付随する。ここでは、権利のうち二次利用を中心に検討する。その場合、音楽も通常は(権利者の異なる)複数の音楽を利用するので、複合的な製品と考えることができる。即ち、動的なバリューチェーンと複合的な製品の例と言える。

OBELIX では以下の3つの課題に対する革新的な解決策を開発することにより、最先端の技術の開発を行なうとしている。

- 1) オントロジー及びセマンティック Web 技術で Web 上の情報資源をどのようにシステムが理解できるようにするか?
例えば、XML 文書電子商取引で製品が取り扱い可能になる中で、製品内容情報の自動分類やカタログ化など。
- 2) e-ビジネスにおける Web コラボレーションの新しいアプローチをどのように促進するか?
例えば、イベントや他のサービスコンフィギュレーションの設計をオンラインで協働的に行なうためのシナリオなど。
- 3) 新しい価値の組合せ(value constellation)をいかにモデル化し、解析し、シミュレーションするか?
例えば、カスタマイズされた音楽製品とその権利の提供や、電力の需要側の e-マーケットプレイスでの入札など。

2.4.4 コミュニティー支援

Web 技術を活用したコミュニティの支援は、コミュニケーションから距離と時間を取り除くものとして多くのコミュニティサイトが Web 上に存在することからも明らかのように、魅力ある Web 技術の適用例である。これらを、セマンティック Web 技術を活用して、さらに使いやすくするのは自然な流れである。ここでは、Web サイトへのオントロジー技術の試みとしてカールスルーエ大学 AIFB の Web サイトとオントロジー研究コミュニティの Web サイトである OntoWeb について述べる。

2.4.4.1 AIFBのWeb サイト

カールスルーエ大学のAIFBのWeb サイトでは、SEAL(SEmantic potAL) オントロジーと呼ばれるオントロジーが用いられている。このオントロジーは、54 個の概念、178 個の関係、4 つの規則（公理）からなっている。概念としては、研究領域や学生、プロジェクト、組織、出版物などがあり、それら概念と属性の関係が178 個定義されている。公理系は、例えば、「あるプロジェクトのメンバーであることを記載することは、そのメンバーの所属プロジェクトにそのプロジェクトが含まれていることである」といった規則が定義されている。これらはスキーマを表しており、実際の人物やプロジェクトはこれらのインスタンスとして表現される。

これらのオントロジー定義に従い、AIFBのサイトでは、人（プロジェクト）に対して、研究領域や、プロジェクト（人）、研究グループのメニューを持った検索ページが構築されている。特に研究領域は、概念階層（クラス階層）に従い、階層的に表示されている（図 2-10参照のこと）。表示文字は、属性値になっているので、ドイツ語と英語の二ヶ国語がオントロジーで定義されており、言語を切り替えることによりメニューの表示も変更される。

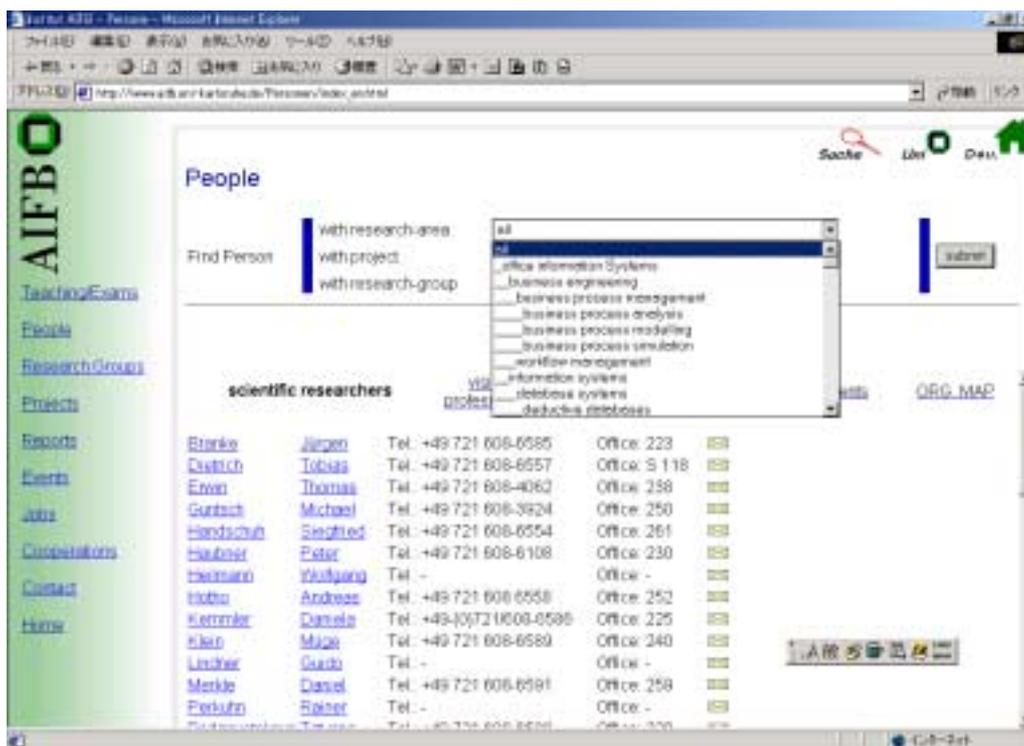


図 2-10 AIFB の研究員検索のページ⁷⁶

⁷⁶ http://www.aifb.uni-karlsruhe.de/Personen/index_en.html

2.4.4.2 OntoWeb：セマンティック Web 技術のコミュニティポータル

OntoWeb⁷⁷は、セマンティック Web 技術のコミュニティのためのポータルである。成果の公開や情報交換により、セマンティック Web 技術の研究開発の促進と、普及を目的にしている。2002 年 2 月現在、まだセマンティック Web 技術をそれ自身に適用してはいないようであり、カールスルーエ大学の AIFB Web サイトにあるような検索システムはないが、SWRC (Semantic Web Research Community) ontology と呼ばれるオントロジーが作成されている⁷⁸。(図 2-11 参照のこと。)



図 2-11 OntoWeb の Home Page

2.4.5 課題

これらの調査を通じて、得られた課題を以下に挙げる。

(1) メタデータの作成・管理方法

メタデータの付与には、多くの場合かなりの労力が必要となる。このために、3.2 で述べられるようなツール類のさらなる整備や、メタデータ付与者への動機付けの工夫など

⁷⁷ OntoWeb : <http://www.ontoWeb.org/>

⁷⁸ SWRC : <http://ontobroker.semanticWeb.org/ontos/swrc.html>

が必要となる。具体的でわかりやすいメタデータの適用事例が望まれる。

(2) 大規模データ・知識への適用

いくつかの事例でオントロジーも作成されているが、まだサイズも小さく、実験レベルに留まっている感がある。それに基づく検索システムも同様であり、より大量の情報を扱う実用レベルの問題への適用が必要である。今後は、実用的なメタデータを使うことによって、コンテキストを考慮した検索が本当に可能かどうかを質・量の双方の観点から検証する必要があるだろう。セマンティック Web の潜在能力は高いが、応用システムが実用レベルに達するのは、ソフトウェアツールが揃うことが大切であり、研究レベルを産業レベルに引き上げることが必要であろう。

2.5 応用システム開発時の考慮点

セマンティック Web を活用した応用システムには、どのようなことを考慮する必要があるだろうか。セマンティック Web を活用した応用システムの多くがメタデータを使ったシステムである。ここでは、メタデータを使った電子政府サービスに限定して応用システムの考慮点を考えていきたい。

第一に考えなければならないことは、電子政府サービス用のメタデータ標準を作ることであるが、全く独自のメタデータ標準を作るべきではない。現時点で、電子政府サービスにおける世界標準メタデータ仕様なるものは、存在しない。しかし、海外各国の政府サービスの電子化におけるメタデータ標準を見渡したとき、3.1.2節で見るように、ダブリンコアを元にしたメタデータ標準を使用している国が多い。また、ダブリンコアを元にしたメタデータ標準を使用していない国でも、ダブリンコアの要素を個々のメタデータ標準の要素にマッピングすることが可能である。新たに作るメタデータ標準は、こういった現状を踏まえて、デファクト・スタンダード（例えばダブリンコアや RDF、XML など）に準拠した形で作るべきである。グローバル化が進んだ今日において、情報資源は、一つの国の中だけで限定して使用されるものではなくなっている。こうしたデファクト・スタンダード技術を使用することで、各国の情報資源を相互運用することを可能にする。また、応用システムもデファクト・スタンダードな技術を使用することで、既存のコンポーネントやライブラリーを使うことが可能になり、開発効率及び信頼性を上げることが期待できる。

一方で、できあがったメタデータ標準が未来永劫同じ形であるとは考えられない。適切な時期に見直され随時更新されるべきである。また、国内あるいは特定の分野の情報では既存のメタデータ標準だけでは要求を満たさない場合も考えられる。その場合、追加の要素を加える必要がある。メタデータ標準では、こういった拡張に関する問題（メタデータ標準のバージョンアップや拡張要素など）に関しても十分に仕様化されている必要がある。メタデータ標準の変更は、そのメタデータ標準を使っている応用システムにとって、多くの時間と資源を消費させる。メタデータ標準の拡張に関しては、安定性と拡張性のバランスを十分考慮して行われるべきである。一方、応用システムは、メタデータ標準の変更に対して最小限の時間と資源で対応できるように考慮されるべきである。

メタデータ標準を用いることで、情報資源が、ウェブページ、電子文書、紙の文書、人（へのポインター）、データベースなどといったものでもシームレスに扱うことが可能になる。応用システムは、この特徴を最大限に利用して開発されるべきである。例えば、オンラインで検索を行なう応用システムを考える。この場合、ネットワークを介してアクセス可能な情報のみを対象とするのではなく、紙の文書や本、CD-ROM、人など物理的な情報資源も対象になる。検索の結果は、ネットワークを介してアクセス可能な情報は、そのままオンラインで表示され、物理的な情報資源は、その情報資源の保管場所、入手方法、連絡先などが表示される。

メタデータ標準を用いることで、組織を超えて情報資源を相互運用することが可能になる。効率的に運用するためには、メタデータを組織横断的に構造化して、全ての応用システムからアクセスできるようにする仕組みが必要である。また一方では、情報資源のセキュリティやアクセス制御も考慮されなければならない。こういった点に関しては、個々

の応用システムが管理するのではなく、全ての応用システムが共通で使う仕組みを作るべきである。

2.6 社会におよぼす影響

セマンティック Web はインターネットの世界を大きく変える技術であると期待されており、それだけに、実用化された暁には社会に及ぼす影響も大きいと考えられる。

あらゆる技術について言えることであるが、技術が単なる技術である限り、または個別技術の集合である限りは、社会への影響は軽微なものにとどまる。市場ニーズが技術を呼ぶのか、技術が市場ニーズを創造するのかは別として、現状ではニッチかつ限定的な状態でしか享受し得ない消費者の満足感が、複数技術の融合によって画期的なサービスに変容し、そのサービスを広く実現するための社会的な制約すらもブレイクスルーしていくムーブメントになったとき、技術はメディアに脱皮し、社会に浸透し、社会に対して影響を与えるようになる。P2P 技術を代表する技術といわれる Napster や Gnutella は、音楽著作権協会から提訴されるなど物議を醸すものの、音楽流通のあり様を一変させてしまうような社会的影響を及ぼすには至っていない。

セマンティック Web という技術も P2P 技術と同じように、それ単独で存在する技術ではなく、複数技術や、技術と技術を組み合わせて画期的サービスを実現するためのビジネスモデルなどの集合体である。北米ならびに欧州を調査した結果、現在「セマンティック Web」という言葉で括られる技術には大きく分けて「メタデータ技術」、「オントロジー技術」と「インテリジェント・エージェント技術」がある。

メタデータ技術とはマシンや後述するインテリジェント・エージェントが理解できるデータの意味を記述するための技術である。メタデータ技術といっても単一的に括れるわけではない。大きくはダブリンコアのように、メタデータを記述する際の標準策定とボキャブラリを作成する（技術というよりは）フレームワークに始まり、ユーザーが各自でメタタグを定義できるようにする技術の XML、用語や概念の意味をマシンリーダブル、マシン・プロセッサブルな形式で表現する技術の RDF といった言語や、RDF に準拠しながら、適用領域毎の標準を記述するための規則である CC/PP（装置やソフト等の能力を記述する）、P3P（個人情報の保護方針を記述する）、PICS（インターネット上のコンテンツ選択に際しての条件を設定する）などがある。また、恐らくは極めて労働集約的な作業になることが想像されるメタデータ付けを行なう作業を（半）自動化するための技術に至るまで多岐にわたる。

次にオントロジー技術であるが、これもメタデータ技術の一角を占める技術であるということもできる。オントロジーとは、概念どうしの関係や、相互関係についての解釈を行なうためのルールを定義した文章の集合のことである。オントロジーを記述するための言語は当初、DARPA の主管する国防領域での適用を企図した米国の標準である DAML と、欧州の学会、産業界が中心となってとりまとめ EU 委員会が予算的に援助して推進された OIL とに二分されていたが、昨年 DAML+OIL に一本化された。ただ、欧州においては、IBROW や On-To-Knowledge のように、産学を横断するオントロジー構築のためのプロジェクトが立ち上げられ、推進されているのに対し、米国独自の動きとしては DARPA での活動しか見当たらない。したがって、オントロジー構築の動きは軍事、諜報領域における戦訓や被疑者の検索精度を向上させるといった目的に閉じるきらいがある。

最後にインテリジェント・エージェント技術であるが、この技術は、究極的には適用領域に依存しない共通エージェント技術であるべきである。しかし、段階論的に言って、本

技術がセマンティック Web 関連技術の中では最も実現までの道のりが長いであろうと考えられる。したがって、初めからフルスペックのエージェント技術の実現を目指すのではなく、当初は市販の検索エンジンに、適用領域に合った機能追加を行なって使う程度のものであってよい（このことは、エージェントに次いで実現の困難なオントロジーについても言える）。

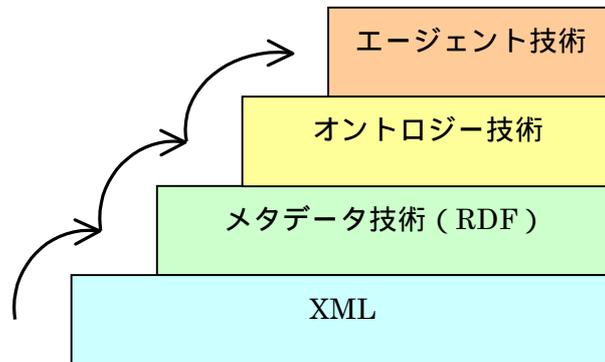


図 2-12 セマンティック Web を構成する技術

重要なことは、当初はこれら三領域にわたる個別技術、個別アプリの寄せ集めに過ぎなかったセマンティック Web が、やがて一部同士が統合され、限られた領域ではあっても、メタデータ オントロジー インテリジェント・エージェントという階段を駆け上がったとき、社会への浸透は加速することになる（図 2-12参照のこと）。また、ここでセマンティック Web 関連技術としては括らなかつた P2P や IPv6 の技術は、セマンティック Web が想定するマシンリーダブルかつマシン・プロセスブルな環境下では不可欠であろう。なぜなら、そうした環境においては、エージェントが問い合わせに行く先は現在想定されるデスクトップ PC やノート PC ばかりではなく、PDA や家電、各種センサーなど、身の回りのありとあらゆるものと通信を行なう可能性があるからである。

そうなれば、あらゆる素材に組み込まれることが見込まれるプロセッサ単位での論議が不可欠となるだろう。そうなったときに初めて、社会におよぼす影響は甚大となる。真の意味においての情報革命と言ってもよいのではないか。ただ、繰り返すが、個別技術の進化が現状では個別に進展し、適用領域も個別に設定されている今日、当初からフルスペックの夢を提示し、そこに直線的に到達するためのマイルストーンを設定するのは、批判や懐疑論ばかりを喚起するだけで現実的ではないであろう。

したがって、次に示す図 2-13のように、実現の容易度に応じ、フェーズ フェーズ フェーズ という順序で段階的に開発を進めていくことが肝要である。

現Webの課題と新しいWebの世界で目指すもの

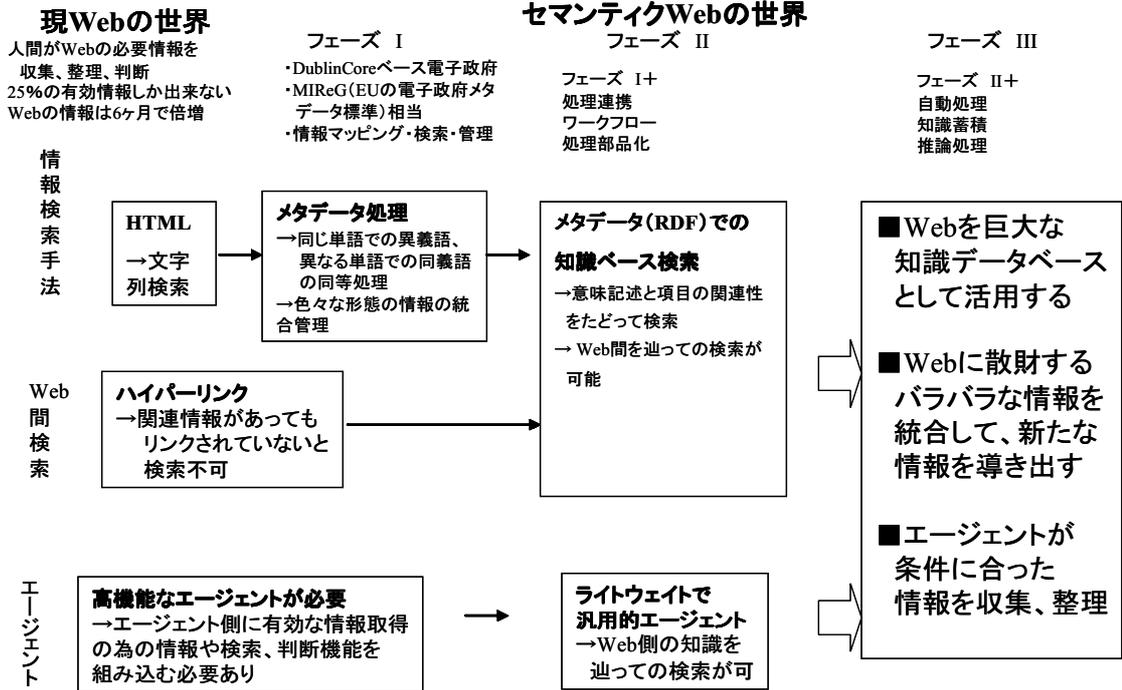


図 2-13 現 Web の課題と新しい Web の世界で目指すもの