

レポート  
概要新たな融合研究  
セマンティックウェブサービス記述と知識処理を行う  
セマンティックウェブ関連技術

現在、ウェブで提供される情報量は爆発的に増大しており、今後数年で 100 倍規模の増大となることが予想されている。こうして蓄積された文書データを複合的に連携させて、革新的なサービスを提供しようとする技術が「ウェブサービス」である。ウェブサービスが目指すのは、蓄積される情報の表現形式の差異を吸収し、求められるサービス毎にそれぞれのサービスに合致した表現形式を提供することである。

だが、増大する情報量によって、求める情報にたどり着くのが難しくなっている。既に統計学的な操作によって検索情報の順位を定める一般的な情報検索エンジンでは、満足できる情報検索結果が得られないという問題が生じている。こうした問題に対処するため、1998 年にウェブの生みの親である Tim Berners-Lee によって構想が示されたのが、知識処理を実現する「セマンティックウェブ」の概念である。セマンティックウェブは、「意味（セマンティック）」を取り扱う技術である。狭義には国際標準化団体である W3C の定める情報共有のための文書公開手段に係る規格であり、広義には公開文書のオントロジーなど「意味」を取り扱う幅広い技術を指す。

セマンティックウェブ技術を利用することによって、インターネットを介した「知識の共有」など高度な情報処理環境が実現できる可能性がある。現在のウェブ関連技術の潮流によって、より高度な知識処理を実施し、多様なサービスの実現に応用できる環境が整備されつつある。

現在、先進的な科学技術分野では、研究活動に付随して蓄積される文書情報が膨大となっている。例えば創薬研究においては、薬効性能が期待できる反応の膨大な組み合わせ可能性の中から実験による検証を実施するための絞り込みが必要である。セマンティックウェブ技術の確立によって、情報を構造化し、異なる地域・組織に分散して存在する情報に対して、半自動的な絞り込みが実施できる。こうした共通の規格に基づく情報共有によって、宇宙探査、バイオインフォマティクス、ナノテクノロジーなどの分野において、研究開発の生産性が飛躍的に向上する可能性がある。

セマンティックウェブ技術をはじめとするウェブ関連技術の研究推進のためには、融合分野の研究に合致した、ミッションを明確にしたプロジェクト型の推進体制が必要である。

本文は p.22 へ

# （サービス記述と知識処理を行う セマンティックウェブ関連技術）

藤井章博  
情報通信ユニット

## 1 はじめに .....

「ウェブサービス」は、ウェブ上に蓄積された文書データを複合的に連携させて、利用者に何らかの情報サービスを提供する技術である。また、「セマンティックウェブ」とは、「意味（セマンティック）」を取り扱う技術で、狭義には国際標準化団体である W3C の定める情報共有のための文書公開手段に係る規格であり、広義には公開文書のオントロジーなど「意味」を取り扱う技術を指す。こうした技術を利用することによって、インターネットを介した「知識の共有」など高度な情報処理環境が実現できる可能性がある。現在のウェブ関連技術の潮流は、より高度な知識処理を実施し、多様なサービスの実現に応用できる環境が整備されつつある。

この技術は、近年急速に標準化が進展し、産業技術としても重要性が増している。ウェブサービスという名称で示される技術体系は幅があるが、本稿では特に、「セマンティックウェブ」の技術動向を中心に紹介する。セマンティックウェブが目的とする「知識処理」は、科学技術研究のいくつかの領域でその必要性が急速に高まっており、この技術領域における一つの先進的な要素技術である。具体的には、バイオインフォマティクス分野に代表されるように、研究成果を蓄積した膨大な量のデータベースから次の研究指針を導き出すような機能を提供する技術としてその重要性が増している。

まず、2では、知識処理が視野に入った情報流通技術の潮流につ

いて説明する。すなわち、最近のウェブ技術の進展が情報流通を効率的に実現することによって、ビジネス上のサービスの高度化や学術研究分野での情報活用に新しい可能性を生み出している状況を述べる。

3では、ウェブサービス関連技術の中で特に、知識処理を指向して「オントロジー」を利用する「セマンティックウェブ」技術の概要を説明する。この技術は、現在国際的な標準化が進められおり、この体系を利用した科学技術研究が盛んになっている。4では、最近開かれたセマンティックウェブについての重要な国際会議の研究発表を中心に研究動向の概説および研究振興のあり方を述べる。5はむすびである。

## 2 ウェブサービス技術に関する技術の潮流 .....

まずウェブサービス関連技術のこれまでの進展を概観する。つぎに、インターネットを介したサービスが、理想的にはどのような形態を目指しているのか、具体的な例で説明する。そのうえで、ウェブサービスによってもたらされる利点について検討する。

2 - 1

### ウェブによるサービスの提供という潮流

インターネット上での公開情報を取り扱う技術は、これまでいくつかの段階を経て進化してきた。まず、90年代の初頭に電子メールが一般に利用されるようになったことに加えて、HTML (HyperText

Markup Language) によって書式を整えた電子文書の公開、すなわち WWW (ワールドワイドウェブ) が登場した。この HTML 文書の特徴の一つは、他の文書への入り口を示す「ハイパーリンク」を提供することである。しかし、この機能だけでは、公開された文書間の連携をとるという目的のためには不十分である。

そこで、複数の主体によって作

成される電子情報に共通の表現形式を与える機能をもった、XML (eXtensible Markup Language) が登場した。例えば、文書のなかで商品名と価格を表す部分を、XML の機能によってそれぞれ「商品名」「価格」であることを明示することができる。これにより、複数の会社が商品の流通などに関与する場合に、帳票などの関連する文書を電子的に共有できる。さらに、XML は、情報を構造をもって蓄積・利用するのに役立つ。構造化されて蓄積される情報は、その2次加工が容易である。これらの技術進化により、情報を蓄積したデータベースは、多様な用途に利用できるようになった。

文書の共有化・構造化と並行して、文書の加工や情報の相互運用に関する環境も整備されていった。例えば、Java というプログラミング言語は、XML に蓄積された情報を加工し、書面に動きのあるダイナミックな文書表示や、遠隔地からの情報入力を行う機能を提供する。すなわち、ウェブの画面上で、豊富な定型書式の文書の運用が行え、そうした文書情報の大規模なデータベースへアクセスする機能が実現された。

## 2 - 2

### 理想的なサービスの提供

これらの技術を利用することで、電子的に蓄積された情報や文書の統合的な運用環境が実現できる。特に電子商取引分野では、流通過程の管理に電子的な情報共有が積極的に導入されている。そこで実現されるサービスとは、提供される情報の内容にほかならない。そこで、インターネットが提供することができるサービスすなわち情報は、理想的にはどのようなものとなり得るであろうか。やや空想的であるが以下でその理想形を考えてみよう。

例えば、長期の休暇を取ることができ、旅行に行くことにする。旅行代理店で旅行のプランを相談する状況を思い描いてみよう。旅行プランを作成するための「必須事項」は、まず、どこにいつ行くのかということである。つぎに「制約条件」としては予算などが挙げられる。さらに「好ましい条件」として好きな航空会社や座席の種類などがある。旅行プランの作成と種々の申し込みというサービスを受けるためには、これらの情報を旅行代理店のエージェントに説明することになる。

良いエージェントの条件としてまず重要なのは、利用者とのコミュニケーションが活発であることである。しごく当然のことであるが、エージェントと利用者は、同じ言葉や語彙、用語を共有しなければならない。これは人間同士のコミュニケーションでは、暗黙の前提となっている。つぎに、利用者の好みを知った上での「提案」ができなければならない。これを実現できるためには、「お薦め情報の提供」といった一方的な情報提供に留まらず、利用者側の条件を知った上で、例えばクーポン券の利用できるサービスを選択するなど利用者に合致したサービスがありがたい。さらに、利用者の個別の好みなどを記憶し、次のプランの作成に反映できればなお良い。また、利用者が常に見ていなくても必要な作業を自律的に実施

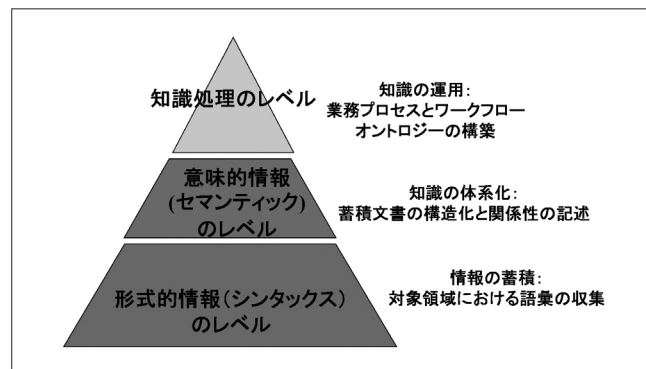
するのが望ましい。ホテルの予約、切符の予約など問い合わせ時間がかかる場合に、利用者に代わって待ち合わせるのである。

すなわち、よい代理人とは、鍵となる情報を収集し、利用者が望むような選択肢を選び、利用者の意向を確認しながら、予約などのサービスを提供してくれる。その際、情報提供者や利用者との語彙の共有や得られている情報に基づいた判断が求められる。

現在、インターネットを介してウェブを通じて提供される情報量は爆発的に増大しており、今後数年で100倍規模の増大が予想されている。情報検索によって、「ヒット」する情報は膨大となりその絞込みに必要な労力が膨大となることは目に見えている。そうすると、現在のように統計学的な操作によって検索情報の順位を定める一般的な情報検索エンジンでは、満足できる情報検索結果が得られないという問題が生じてくる。せっかく有用な情報がネット上に存在しても、それにたどり着けないという矛盾が生じるであろう。

ウェブサービスが目指すのは、蓄積される情報の表現形式の差異を吸収し、求められるサービス毎にそれぞれのサービスに合致した表現形式を提供することである。これによって、膨大な情報空間から利用者のニーズに合致したサービスを選択できる。こうした技術が成熟すると、具体的には次のよ

図表1 セマンティックウェブによる知識処理の概念図



科学技術動向研究センターにて作成



うなウェブサービスが提供できる  
といわれている。

- ① IT 技術に弱い人に優しい情報入出力のための環境が実現する
- ② 巨大な分散型の知識データベースとしてインターネットの世界が利用できる
- ③ 本来別々の目的で構築されたデータを統合して、新しいサービスが生まれる

これらの特徴は、デジタルデバイドへの対処、高齢化社会への対応、生涯学習の支援、多種多様な情報端末とロボットなどの制御、といった機能の提供に役立ち、我々の日常生活に密接した支援に結びつくはずである。

2 - 3

### セマンティックウェブ技術の全体像

図表1は、こうした高度なサービスの提供を実現するための技術の全体像を概念的にあらわしたものである。体系を3つ階層で表現すると、下位層から①形式的情報のレベル、②意味的情報のレベル、③知識処理のレベル、で表される。ソフトウェアの体系を階層的

に記述する場合は次の2点に留意する。第一に、上位の階層はその一つ下の階層の機能を利用して自分の階層に求められている機能を実現するという点である。第二に、異なる階層の間では機能の定義は独立に行われるということである。すなわち、XML等によって構築されたデータベースの機能を利用して「セマンティック」の世界が記述され、それにもとづいて知識処理が実現されるのである。

知識処理に基づくサービスは、最上位の形式的情報レベルで、「オントロジー」と呼ばれる語彙の体系を構築すること、およびそれらの間の関係を規定するサービスの記述によって実現される。現実には、これらの階層は、より細分化された階層に分けられ、標準技術としてその仕様が規定されている。詳細化した階層は、「レイヤケーキ」と呼ばれるもので、以下3でより詳しく解説する。

2 - 4

### 注目される科学技術領域におけるウェブサービスの可能性

いま「サービス」に関する科学がイノベーション政策の観点からも重要であるという指摘がある。

特にITの技術進化の方向として、サービスを指向することは重要であると考えられている<sup>1)</sup>。

例えば、科学技術政策研究所が実施している技術予測調査を取り上げてみよう。科学技術全体の中で、今後10年程度を視野に入れた上で、重要であると考えられる130の「注目科学技術領域」を抽出した。この領域の内容を検討すると、多数の領域でウェブサービスのような大規模な情報検索機能の存在が技術領域の実現にとって有効であるか、あるいはその存在を前提としている。

調査の対象分野の情報通信分野では、「人間の知能支援」に関する領域が将にこの技術に該当し、同じく、ライフサイエンス分野では、「テーラーメイド医療」や「バイオメトリックによる個人認証」などの技術課題がウェブサービスが有効に寄与しそうな課題である。

欧州で最近行われた科学技術予測調査でも、サービス指向が情報技術の進化において重要であるという指摘が、識者による検討の結果として読み取れる<sup>10)</sup>。すなわち、ウェブサービス関連技術は、単に情報技術分野内の技術進化に留まらず、他の分野・領域との融合のあり方を具体的に議論することが求められる技術分野である。

## 3 セマンティックウェブ技術の体系

ウェブサービスを実現するために技術体系は、主要なものとして、「セマンティックウェブ」と後述する「UDDI」が存在する。まず以下では、前者を主に説明する。その技術の中心的な役割を担う「オントロジー」、知識処理の基本となるRDFと呼ばれる言語の概要およびサービス提供のモデルについて解説する。後にUDDIとの関係について説明する。

3 - 1

### オントロジーに基づく連携

実用に供することが出来る規模のデータベースに対して、代理人の例で述べたような高度な情報検索機能を実現するには、以下で述べる「オントロジー」の構築がもっとも重要な課題となる。

オントロジーとは、語彙の体系を現す人工知能分野の専門用語で

ある。例えば、「ニュートン力学」という語彙について検索するとする。検索のために普通に採られる手法は、語彙を表現する文字データの一致／不一致の判定に過ぎない。オントロジーを利用して、より進んだ情報検索を実現するということは、例えば、「ニュートン力学」に対して、その「物理学」との関係や同じ物理学に属する別の知識である「相対性理論」ととの関係を利用するというこ

る。この関係性の記述には、「論理式」と呼ばれる一定のルールに基づいた情報の蓄積が必要となる。すなわち、オントロジーに基づく知識処理とは、こうして蓄積された体系を利用して、多様な問い合わせに対して回答を得られるシステムを構築することになる。

3 - 2

## レイヤケーキ

セマンティックウェブでは、これまで説明したように情報の蓄積と利用は構造化されて行われる。図表2に示す「レイヤケーキ」と呼ばれる体系は、この構造化を規定するものである。この階層構造は、現在まで下位層から順次仕様が定められ、現在5番目のオント

ロジー層まで固まっている。この技術体系に関する国際的は、W3C (World Wide Web コンソーシアム) と呼ばれる標準化団体によって検討されてきた<sup>11)</sup>。

ウェブで取り扱われる情報は、基本的にはテキストとグラフィックスなどの情報であり、「シンタックス」に属するレベルの情報である。これらの情報は表示のための構文によって記述されている。構文として広く利用されているのが、HTML や XML である。

さらに、「セマンティック」を取り扱うために、後述する RDF と呼ばれる階層を設けている。これは、一定のルールに基づいた知識の体系が記述される。RDF については後述する。

さらにその上位に位置するの

が「オントロジー層」である。この層では、注目する知識体系を RDF によって集積された語彙とそれらに対するルールの体系として構築する。以下で解説するゲノムなどの応用分野は、こうして構築されたオントロジーとその利用システムの事例である。

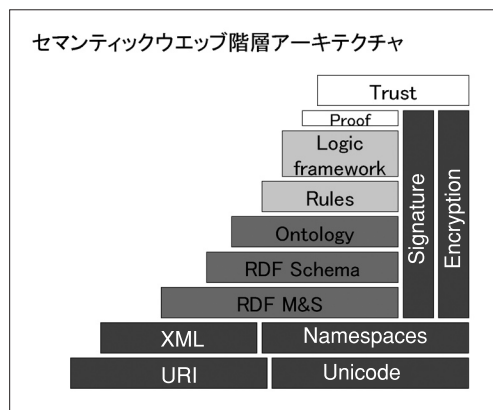
3 - 3

## RDF (Resource Definition Framework)

ここで簡単に「知識の記述」の例を説明する。図表3に示すのは、RDF グラフと呼ばれる三項組みの構文である。RDF は、Resource Definition Framework の略で知識体系の定義を行うのに使われる。知識の断片は、主語 (Subject)、述語 (Predicate)、目的語 (Object) の三つの部分から構成される文法によって記述される。この構造を記述する方法が幾つか存在する。

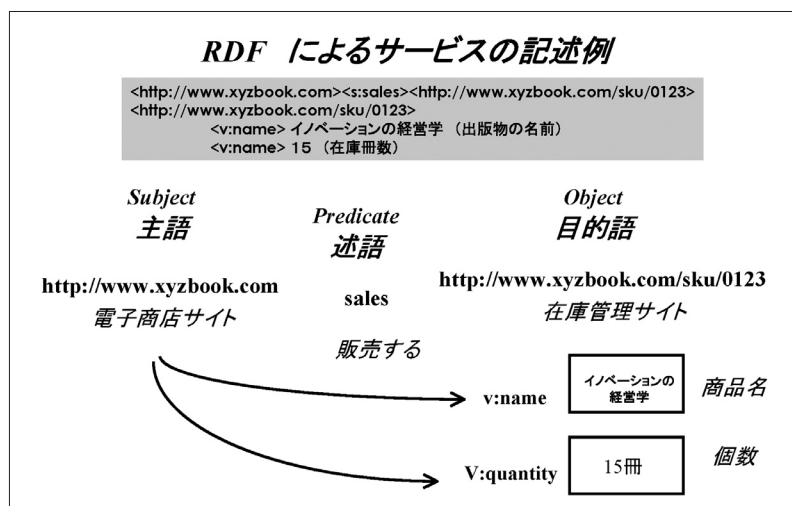
図表に示した例では、書籍の電子店舗の例である。まず、主語で示された書籍の集合が存在する。これは、例えば、店舗における商品の陳列を表す。述語に相当する部分では、陳列されている商品の「在庫情報」をもとに「販売(sales)」を実施するというルールが規定されている。そこで、述語で示されたウェブページに記述されているのは、個別の書籍データの属性となる目的語として、「イノベーションの経営学」という書籍名と「15冊」という在庫数が参照できることを表している。この例は、小売店と卸業者が別々のデータベースを持って商品を管理している状況を表している。この例は、電子商店にとって、「在庫品の引き当て」というサービスをセマンティックウェブの機能を利用して実現したことを表している。

図表2 レイヤケーキ



文献<sup>11)</sup>を参照し科学技術動向研究センターにて作成

図表3 オントロジー構築のためのロジック



科学技術動向研究センターにて作成

## 3 - 4

## サービス提供のモデル

ウェブによるサービスの提供は、単純に情報提供者とその利用者の二者間だけで成り立つのではない。たとえば、ある分野の商品に関する在庫情報などサービスに結びつく情報の提供者は、販売を生業とする事業者と異なる可能性がある。この場合消費者により高度なサービスが提供されるためには、この情報が一定のルールの上で共有される必要がある。在庫情報であれば、販売以外の目的は想定しにくい。が、学術研究のデータは、利用形態があらかじめ特定されていない場合もありうるため、サービス内容の広がり期待できる。

すなわち、ウェブサービスでは、データを提供するもの、サービスを提供するもの、サービスを利用するもの、の三者が独立した主体でありうる。このため、何らかのサービスを販売する「サービス・ブローカ」というというビジネスが成り立つ可能性がある。図表4はこれらの関係を概念的に表現したものである。

現在、電子商取引を推進する立

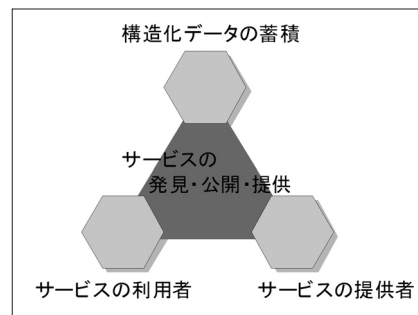
場の技術者は、こうした構造によって、ビジネス上の多様なサービスが効率的に提供できると考えている。同様に、学術研究の分野でも、ある研究分野で別々の目的を持って蓄積された大規模な情報データベースから新しい学術的知見を導き出すという作業をコンピュータが支援することが考えられる。

## 3 - 5

## 関連する技術体系

セマンティックウェブに関しては、W3Cが世界的な標準を規定しており<sup>11)</sup>、周辺の技術仕様とともに勧告が公開されている。セマンティックウェブの規格による体系と共存する形で、UDDI (Universal Description, Discovery and Integration) と呼ばれる規格が存在する<sup>9)</sup>。この体系は、2000年に発足したオアシス・コンソーシアムによって規定されている。この規約は、まず「ビジネスレジストリ」と呼ばれるデータベースに格納するデータの構造を定義する部分 (UDDI Data Structure Reference)、つぎにデータを利用するアプリケーションを作成するためのアプリケーションプログラミングインターフェース (UDDI

図表4 サービスブローカ概念図



科学技術動向研究センターにて作成

Prorammer's API) からなる。プログラマは、このインターフェースの部分にサービス内容に関連するソフトウェアを開発することになる。UDDIは、特に電子商取引の分野において、ウェブによる高度なサービスの提供を目指している。UDDIの規格のもとでは、オントロジーのレベルを規定していない。つまり、サービスの「意味」については、UDDIの体系では自然言語による記述に託されているということである。

やはりこの技術体系でも、データ提供とそのデータに基づくサービスは独立していること、さらに、サービスの構築はそれ自体が一つのビジネスとして成立しうることを想定している。こうした考え方は、ウェブサービス技術全体に共通の特徴である。

## 4 セマンティックウェブに関する研究動向

以下では、セマンティックウェブ技術の研究とこの技術を他分野の研究に応用する研究をそれぞれ紹介し、研究振興について考察する。

## 4 - 1

## セマンティックウェブの理論的研究

理論的な研究として、幾つかの課題を分け、その動向を簡単に述べる。つぎに、応用分野へ適用す

る研究の事例として「バイオインフォマティクス」分野の応用事例を取り上げる。さらに、国内で公的資金によって現在実施されている研究プロジェクトを簡単に紹介する。

## (1) 意味的統合

セマンティックウェブでは、共通の領域に関して複数のオントロジーが作成される場合がある。例えば長年にわたって、複数の研究所で蓄積されてきた文献データ

ベースを統合して運用する場合にこうした状況が生じる。この場合、オントロジー間の対応関係を構築するといったことなど、複数の「セマンティック」の統合技術が重要である。オントロジー間の対応の取り方、大局的なオントロジーと局所的なオントロジーがある場合の対応の取り方、などの研究が重要である。これらの研究は、人工知能分野のデータベース分野におけるデータ統合研究の発展形と位置づけられている。



## (2)記述論理に基づく推論系

オントロジーを記述するためには、「論理式」を構成することになる。これは「述語論理」と呼ばれる文法を利用して実現される。オントロジーを記述する言語は、OWL (Web Ontology Writing Language: オントロジー言語) と呼ばれ、W3Cによって標準化が進んでいる。述語論理によって知識体系を記述することは、「AならばBである」といった関係の集合を構築することに対応している。こうして作られた集合がデータベースとして、セマンティックウェブにおける知識処理の基本となる。

述語論理を利用した知識処理は、「AならばBである」「BならばCである」…という断片的な知識の集合から、「AならばZであるか?」といった問い合わせに対して回答を出力する、という形で実現される。こうした操作は推論機構と呼ばれる。つまり、対象となる知識領域のオントロジーデータベースを構築した上で、OWLを利用した推論機能を領域毎に効果的に実現するための研究がセマンティックウェブでは重要である。

## (3)サービスの発見・連携・実行

必要とするサービスをネット上で発見し、サービスの提供を受けるまでの過程は、利用者が何らかのキーワードによってネット上のサービスを検索し処理を実行することである。このことは、インターネットの検索エンジンにキーワードを記入し、関連するホームページを見つけるという行為を発展させた状況といえる。このとき、要求されるサービスの記述は一定の構造をもつデータである。そこで、求めるサービスが、サービスを提供する側の記述と合致するかどうかの判定方法やその効率など

が研究の対象となっている。

問い合わせの方式の具体的な各種のアプリケーションへの適用やベンチマークなどを進めている研究が行われている。また、推論エンジンを用いて、サービスの組み合わせを動的に構成する試みもある。例えば、サービスに付けられサービスの内容を説明するメタデータ間の関係を判定するアルゴリズムの研究がこれにあたる。他にサービス条件の比較や、文法レベルの比較を併用するものもある。

4 - 2

## 応用研究事例

### (1)eScience

米国メリーランド大学のHendlerらは、セマンティックウェブを用いて、グリッド型のコンピュータシステム上で多様な科学的領域における問題解決を行う研究を「eScience」と呼び、新しい研究の方向を提唱している。念のため、ここで取り上げる「eScience」は、英国で実施中の科学とコンピュータについてのプロジェクトとは別である<sup>7)</sup>。

グリッドコンピューティングとは、汎用のCPUを多数接続し、計算処理を並列的に実行する考え方である。例えば、ある結晶の特性を予想する際、研究者は化学物質の構造を既存のデータベースで調べる必要が生じる。グリッドコンピューティングは、大規模なデータベースの検索とその上での比較的軽い推論処理を実施することに非常に適している。このため、グリッドコンピューティングは、セマンティックウェブにおける処理のためのエンジンとして有望であると考えられている。

一方、専門知識に関するデータベースというものは、従来からの蓄積があるが、これらは必ずしも統一的に構築されてきたわけではない。そうしたデータベースの

データ形式や操作手順は様々である。すなわち、データの規模、サービスを提供するために必要な計算量、他の操作との協調の方法などがデータベース毎に異なる。そこで、「相互運用性」を提供する機能が必要である。数値シミュレーションなどの結果をセマンティックウェブによって記述し、その結果を利用するようなサービスの提供を実施するプラットフォームの構築が重要である。

eScienceでは、幾つかの具体的な科学技術データベースに関して、セマンティックウェブ技術を導入し、共通の情報基盤を構築することを試みている。既存の情報の断片的部分から、研究者が望む知識を取り出すためには、多岐に亘るデータベースを通じて共通のセマンティックスが規定されていなければならない。これによって、幅広い領域にまたがる科学技術研究のためのサービスが提供できるとしている。

### (2)バイオインフォマティクス分野における応用事例

バイオインフォマティクス分野でのセマンティックウェブの応用事例としては、myGrid、MBOY-Service、Semantic-MBOYという三つの「GO:ゲノム・オントロジー」がその代表例である。これらのGOは、データ同士の関係から生体内で起こっている多種の反応経路（例えば、代謝経路など）に関するネットワーク関係を導出するシステムである。

生体内には、多種多様な物質の反応経路が存在し、それらが更にネットワークをもつという複雑な制御が行われていると考えられる。実際、研究により現在約17,000項目の、反応に関するルールが蓄積されている。この反応経路は「パスウェイ」と呼ばれるシグナル伝達経路である。GOに基づくシステムは、パスウェイを文

献から効果的に検索するためのツールといえる。これを用いることで、数100万の化学反応の中から、着目している効果に有効な塩基対や蛋白質に関するルールを検索することができる。

これらの研究プロジェクトで共通に検討されている機能は、自動的な情報サービス、構造化されたメッセージの提供とミドルウェアの存在、情報サービス結果の表示とそのためのインターフェースのあり方、利用者の個別の要求に合致させるための方策、複雑なオントロジーの構築などである。

### (3)国内の研究事例

以下は、人工知能学会の研究発表に基づいて調査した、国内における科学技術分野のオントロジー構築の事例である。2つの事例を簡単に紹介する。

大阪大学の産業科学研究所では、ナノテク材料技術に関する知識の構造化を目指して、オントロジーの構築とその利用が行われている。ナノテク分野の研究では、既存の幾つかの研究領域が幅広く関与するものである。そこで、複数の分野を横断した共通概念の提供を目指して、『概念レベル』のインターフェースの実現を目標としている。現在具体的には、特許情報などの分析により知識の体系化を進めている。

また、ゲノム総合科学研究センターでは、薬のもつ分子レベルでの反応経路（パスウェイ）の知識ベースを構築している。まず、個別の物質間の相互作用をRDFの枠組みにより「薬物」、「生体物質」「相互作用によって生成される出力」3項関係で記述する。このようにして蓄積された関係のデータベース上に、薬機能に関するオントロジーを実装する。相互作用の組み合わせに関する基本的なルールを設定し、データベース上で推

図表5 国内のセマンティックウェブ研究プロジェクトの例

研究内容	研究者の所属機関
ナノテク材料技術の知識構造化	大阪大学 産業科学研究所
宇宙システムの開発・運用支援	宇宙航空研究開発機構、大阪大学
バイオロジーにおける知識体系化	産業技術総合研究所 生命情報科学研究センター
薬機能	理化学研究所 ゲノム総合科学研究センター

科学技術動向研究センターにて作成

論機能を実現する。これによって薬物に関する相互作用の関係に対する効果的な問い合わせが実現できる。研究者は、これを利用して、絞り込まれた相互作用から反応経路を推定することで新たな反応経路を発見できる。

### (4)バイオインフォマティクス分野での、セマンティックウェブ研究の意義

バイオインフォマティクス分野である研究発表によると、彼らが実施したセマンティックウェブの応用という研究活動から、一般的な知見として次のようなことが言えるとしている。

まず、分野固有の知識体系に関するオントロジーの構築は、非常に困難な作業である。特に、すでにデータベース資産が存在しているような場合に、既存のデータ形式との整合性などを考慮する必要があるがさらに困難となる。そこで、この分野の研究には、多くの人材の投入と時間が掛かると述べている。また、セマンティックウェブの利用者が、その求めるサービスをセマンティックウェブの文法で記述しその表記法を通じて、提供されるサービスを発見するという方策は、他の多くの領域に応用できる可能性があり期待される。セマンティックウェブの先駆的な研究事例としてバイオインフォマティクス分野が代表的である。大規模なデータベースを利用して知識処理を伴う検索機能を導入することは、今後他の専門領域に波及し

ていくと思われる。このため、バイオインフォマティクス分野で培われた研究手法は、各種の専門的なデータベースの研究に関して波及効果が期待できるといえよう。

4 - 3

## 研究振興のあり方

セマンティックウェブ関連技術の研究で最も重要な国際会議の一つであるISWC2004 (The 3rd International Semantic Web Conference, 2004) が昨年11月に広島で開催された。参加者は、総勢450名の参加者の内、海外からの参加が300名を越え、日本国内で開催される国際会議としてはめずらしく国際色が豊かであった。発表された論文の質はいずれも高く、20代と思しき研究者の発表が印象深かった。

この会議で基調講演を行ったのは、人工知能の父と呼ばれ、1994年のチューリング賞受賞者であるスタンフォード大学のFeigenbaum教授である。教授の講演では、この分野の研究振興のあり方が明確に示された。教授の言葉を借りると、セマンティックウェブ研究のあり方に関するメッセージとして、「Give me something working!」と強く述べていた。オントロジーの構築と運用・管理に見られるように、セマンティックウェブ研究では、高度な人工知能研究の成果が応用される。教授のメッセージは、ともすると言語モデルなど理論的な研究



に終始しがちなこの研究分野において、実践的応用の重要性を主張したものと思われる。単に実践的な応用の重要性のみを主張しているのではなく、「理論と実践の間の螺旋的な進展を期待する」と補足している。

いま仮に、研究の推進体制をディシプリン型とミッション型に分けて考えてみると、前者は既存の学会活動のなかで自発的に行われるもので、後者は何らかのトップマネージメントが介在し、プロジェクトの推進の中で、既存の専門性の枠を超えて行われるといえよう。教授のいう螺旋的な進展には、ミッション型のプロジェクトの設定が必要であろう。

例えば、セマンティックウェブの有効性を知らしめる実証実験の事例として、Semantic Web Challenge とよばれる応用システムのコンペが国際的に行われている<sup>8)</sup>。これは、セマンティックウェブ技術への一般社会の理解を高めるとともに、研究者の研究活動を刺激してより高い目標設定を促すといった目的で、いくつかの学

会の支援のもとで研究者有志が実施しているコンペである。

開始された2003年度は、10プロジェクトの参加があり、優勝は、英国 University of Southampton の CS AKTive Space と呼ばれるアプリケーションであった。これは、英国のコンピュータ科学研究者を検索するシステムである。2004年のコンペに対しては、12月現在で、18のプロジェクトが参加している。日本からの参加は、2003年の国立情報学研究所からの「Semblog」1件である。今後こうしたプロジェクトに対する、我が国の大学からの積極的な参加が望まれる。

もう一度 Feigenbaum 教授の言葉を借りると、応用分野との接点が重要なこの分野の研究上の進展のためには、研究計画の立案にあたって、研究の「道筋 (path)」の設定が重要である。例えば、教授は、ロボットチームによるサッカー競技を例にあげて道筋の説明をした。こうした競技会を設定することは、当初は研究者の失笑を買ったのかもしれないが、引

き続き実施されていくに従って、「分散人工知能」分野における研究上の蓄積が着実に達成されてきたと説明している。一方、反対の例としてある種のパターン認識の研究を道筋が行き詰った例として挙げた。認識精度のみをコンマ数%のオーダーで追求することは、処理能力が向上した現在では研究上の意義が乏しいと説明した。

知識処理に関する日本の学界活動のレベルは、大きなプロジェクトが推進された80年代の人工知能研究のブーム以来、決して低くないと考えられる。しかし多くの研究活動が、ディシプリン型で実施されているように見受けられる。今後のこの分野の振興には、教授のメッセージにもあるような、ミッションを明確にした、分野融合研究のマネージメントが求められていると考える。すなわち、螺旋的な道筋の先に、理論的にも応用的にも意義のある成果がもたらされるような環境設定・目標設定が重要性であると考えられる。

## 5 むすび .....

本稿で特に、セマンティックウェブ技術に着目した理由は、この技術のもつサービス指向という性格が、今後のITの技術進化の方向のみならず多くの科学技術領域の融合研究テーマとして重要であると考えからである。

現在、第3期基本計画の実施に向けた検討が盛んである。情報通信技術分野は、引き続き科学技術政策上の重点分野の一つとみなされており、特に他分野との融合領域において、多くの重要な技術開発課題が認識されている。

融合領域の研究を振興するにあたっては、「道筋」を意識したミ

ッション型のマネージメントが重要であり、セマンティックウェブ研究には将にこうしたマネージメントが求められている。

### 参考文献

- 1) National Innovation Initiative “Innovate America : Thriving in a World of Challenges and Change”, Dec. 2004
- 2) 赤植淳一、和泉憲明、川村隆浩、武田英明「ISWC2003に見るセマンティックウェブ研究動向」人工知能学会研究会資料、SIG-SW&ONT - A302 - 10
- 3) 財情報処理相互運用技術協会「セ

マンティックウェブ技術の調査研究報告書」、平成15年3月

- 4) 萩野達也、「セマンティックウェブとは」情報処理、Vol.43、No.7、2002年7月
- 5) 清野正樹、「セマンティックウェブとオントロジー記述言語」同上
- 6) Sheila A. McIlraith, Dimitris Plexousakis, Frank van Harmelen (Eds.), “The Semantic Web - ISWC2004”, Lecture Notes in Computer Science, Springer, Nov. 2004
- 7) D. De Roure, Y. Gil, J. Hendler, “E-Science”, IEEE Intelligent

- Systems, 19(1): pp.24 - 25, 2004
- 8) セマンティック・ウェブ・チャレンジのサイト：  
<http://challenge.semanticweb.org/>
- 9) オアシス・コンソーシアムのUDDI 標準化サイト：  
<http://www.uddi.org/>
- 10) 欧州の情報通信技術に関する予測調査プロジェクト：  
<http://fistera.jrc.es/>
- 11) セマンティックウェブに関する標準化サイト：  
<http://www.w3.org/2001/sw/>
- 12) Tim Berners-Lee, James Hendler  
“Publishing on the semantic web” Nature, 410, 1023 - 1024 (26 Apr 2001)
- 

## 執 筆 者



情報通信ユニット

**藤井 章博**

科学技術動向研究センター

<http://www.nistep.go.jp/index-j.html>



工学博士。分散コンピューティングと通信プロトコルの研究に従事した後、電子商取引システムの構築プロジェクトを実施。現在、情報通信技術のイノベーションが経営や政策に与える影響に興味を持つ。