

# セマンティックWebとオントロジー

## Semantic Web and Ontology

武田 英明<sup>\*1\*2</sup>

Hideaki Takeda

<sup>\*1</sup> 国立情報学研究所  
National Institute of Informatics

<sup>\*2</sup> 東京大学 人工物工学研究センター  
RACE, The University of Tokyo

In this article, I show the overview of research on Semantic Web and Ontology. Ontology in computer science was introduced in 90's as the extension of Knowledge Based Systems. It is mainly investigated for knowledge sharing systems. On the other hand, Tim Berners-Lee, the inventor of Web showed the concept of Semantic Web as the extension of Web and started the activity for Semantic Web. Semantic Web uses ontology as the structure of Web information so that Semantic Web Languages such as RDFS and OWL are invented for representation for ontology in Web. Now ontology is used variously from purely vocabulary to knowledge systemization.

### 1. はじめに

本タイトルは研究会の名称であるが、「A and B」というちょっと変わった名称である。セマンティックWebとオントロジーの研究会であることは間違いないが、では両者はどういう関係にあるのかを疑問にもたれる人もあるかもしれない。本稿ではセマンティックWebおよびオントロジーに関する一般的説明と最近の研究動向などを紹介するとともに、両者の関係についても考えてみたい。

### 2. オントロジー

オントロジーとは本来は哲学の存在論のことであるが、90年代の人工知能においてコンピュータ科学に導入された。Tom Gruber はオントロジーとは「概念化の形式的な仕様」と定義しており、具体的にはコンピュータの中で扱われる概念の体系を指している。

オントロジーは形式論理を用いて書かれることが多く、この場合は概念自体や概念間の関係は論理の公理として書かれる。また一般に階層関係をもつことからオブジェクト指向パラダイムにおける抽象化に近い。

#### 2.1 オントロジーに関わる研究

オントロジーに関わる研究としては、オントロジーの構成に関する基礎的な研究、具体的なオントロジー構築に関する研究、オントロジー構築・利用ツールに関する研究、オントロジーを利用した研究などが挙げられる。

オントロジーの構成に関する基礎的な研究としては上位オントロジーあるいは形式オントロジーの構築に関する研究が挙げられる[武田]。上位オントロジーとは全ての他のオントロジーの元になるもので、世界にどのようなものが存在するかについての体系である。有名なものとしてはイタリアの Guarino らの DOLCE[Gangemi2009]、アメリカの Smith らの Basic Formal Ontology (BFO)[Greon2004]がある。近年、日本の溝口らも YATO[Mizoguchi 2009]というオントロジーを提案している。

具体的なオントロジー構築に関して、様々な分野での分野オントロジーが試みられている。本研究会でも、航空宇宙、薬機能、パスイオ(バイオ)、ナノテクなどのオントロジー構築が報告さ

れている。

オントロジーに関連するツールとしては、オントロジー構築ツール(「法造」や DODDLE など)、オントロジー自動構築ツールなどの研究が行われている。

オントロジー利用研究については、教育から行政手続きまで幅広く試みられている。

#### 2.2 オントロジーに関わる会議

形式オントロジーに関しては FOIS (International Conference on Formal Ontology in Information Systems)という会議がほぼ隔年で開かれている(1998, 2001, 2004, 2006, 2008)。

また JCOR: the Japanese Center for Ontological Research が主催して InterOntology08, InterOntology09 という会議が東京で開かれている。この JCOR は NCOR: the (US) National Center for Ontological Research, ECOR: the European Center for Ontological Research に対応する組織して 2007 年に岡田教授が慶應大学内に作った組織である。

### 3. セマンティックWeb

Web は Tim Berners-Lee は研究者間の情報交換のために考案したものであり、その後広く人々の情報交換のために使われるようになった。その経緯についてはいまさら語る必要もないであろう[Bernes-Lee2000]。

しかし、彼は現在の標準となった http と html に満足していなかった。彼はもっと知的なサービスをその当初から考えていた。Web の標準化が一旦なされたとき、それをセマンティック Web として提示した[Bernes-Lee 2001]。セマンティック Web は Web の標準化団体である W3C(World Wide Web Consortium)での標準化活動と情報科学の研究者の研究の両輪で進められている。

#### 3.1 セマンティックWebとは

セマンティック Web の方向性を理解するのは”layered Cake”と呼ばれる図(図1)に基づくとわかりやすい。この図はかなり初期のもので、その後さまざまな変化をしているが、基本構造はあまり変わっていない。いわゆる Web は XML と namespace, URI の層である。これは基本的に現在の Web を支える仕組みであると同時に、この上に Web の拡張として構築されるセマンティック Web の基盤としても重要である。とくに URI はオントロジー共有において重要な役割を果たす。

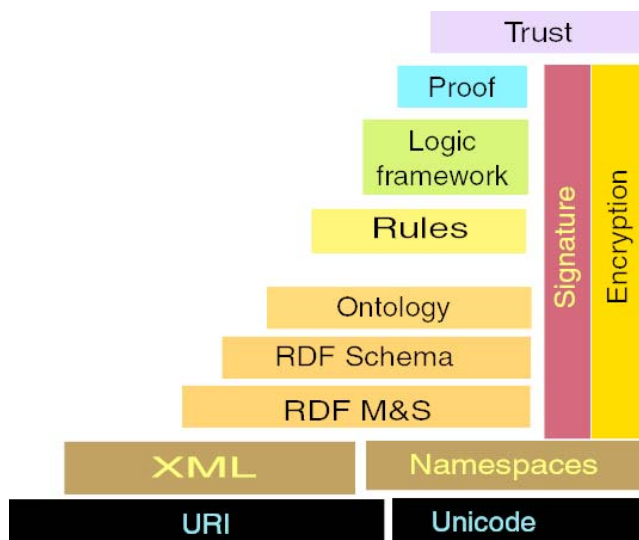


図 1: セマンティック Web の階層図

<http://www.w3.org/2002/Talks/09-lcs-sweb-tbl/>

RDF, RDF Schema (RDFS), OWL はセマンティック Web のための言語である。セマンティック Web として標準化されているのはここまでで、これより上の部分は研究途上である。

セマンティック Web は名前からすると「意味」を直接扱っているようにみえるが、実際には自然言語処理やこれまでの知識処理で扱ってきた「意味」とは異なる扱い方をしている。一番のポイントは URI による意味の表記あるいは限定というアイデアである。

ある概念がなんであるかを定義するのは本質的に困難である。しかし、現実には人々はそういった定義を気にすることなく、概念を使ってコミュニケーションをしている。大まかに概念を理解しており、その理解が両者にとって同じであればコミュニケーションは成立する。

セマンティック Web ではこの異なる人が一つの概念を共有しているということ、同じ URI を共有することによって実現する。もしある概念に関して別の人が別の意味を共有しているなら、別の URI を共有すればよい。このような仕組みを用意することで意味定義の泥沼にはまることなく、プラグマティックな意味共有を実現している。

とはいえ、たとえ部分的であっても URI を参照したときにその概念の定義があることは望ましい。その定義を可能とするのがオントロジーの考え方であり、実際の記述はセマンティック Web のために開発されたオントロジー記述言語 RDFS, OWL によって行われる。

### 3.2 最近の傾向

セマンティック Web は初期にはトイプロブレム研究が多かったが、近年は Web に実際にあるデータ(往々にして巨大データ)を扱う研究が主流になっている。ことにこの 2 年ぐらいにおいては Linked Data というアプローチが注目を集めている[武田 2008]。Linked Data とは Web 情報を HTML ではなく RDF でアクセス可能にしておき、多様なサイトの情報を RDF レベルで相互利用するというものである。

ISWC においては Semantic Web Challenge というコンテスト型企画が毎年行われている。2008 年からは Billion Triples Track というものが別途おかれ、10 億以上のデータを使うシステムのみ参加して競争するというプログラムが行われている。

また人間間のネットワークをメタデータとした Web の研究や Wikipedia を対象とした研究も広い意味でのセマンティック Web<sup>1</sup> と考えられ、盛んに研究されている。

### 3.3 セマンティック Web に関する会議

現在、セマンティック Web に関しては国際的な会議が多数開かれている。International Semantic Web Conference (ISWC) は 2002 年より毎年開催されており、2008 年までで 7 回開催されている。European Semantic Web Conference (ESWC) は 2004 年に始まった European Semantic Web Symposium を前身として発足して毎年開催されている。Asian Semantic Web Conference (ASWC) は 2006 年から原則毎年で開催されている。

これとは別によりビジネスに近い会議も行われている。Semantic Technology Conference は 2005 年から毎年アメリカ西海岸で開催されているもので、IT 系のビジネス関係者を参加者とする会議である。また Linked Data Planet Conference という先に述べた Linked Data だけを対象とするビジネス系の会議も開かれている。

### 4. セマンティック Web とオントロジー

先に述べたようにセマンティック Web にとってオントロジーは Web 情報の構造化を担う重要なレイヤーである。もちろんセマンティック Web にはオントロジー以外にも重要な技術はある。一方、オントロジー研究からみるとセマンティック Web は一つの応用であり、オントロジー研究が常にセマンティック Web を出口とする訳ではない。しかし、RDFS や OWL というセマンティック Web 言語の開発によって、オントロジー研究がより実践的になり、その点ではセマンティック Web との関係は強くなっている。

表 1 と表 2 に古崎ら[古崎 2006]によるオントロジー研究の分析結果を示す。まず、オントロジー利用目的を 7 段階に分けている。これは上が表層的な利用であり、下が内容的な利用である。次に利用方法として知識モデルの利用法とメタデータの利用法に分けている。前者は人工知能における知識処理であるが、後者はセマンティック Web 的利用でオントロジーを利用した情報処理である。さらに扱うオントロジーの中身を 0) 語彙, 1) is-a, 2) 関係, 3) 意味, 4) 強い公理に分けている。

表 2 では以上の 3 つの軸で、本研究会発表および国内論文 (J-Stage) でオントロジーに言及している 97 の論文を分類している。表 2 の上段が知識モデルの利用、下段がメタデータの利用である。この表からわかるようにいまやオントロジーの利用は多岐に渡り、表層的な扱いから深層的な扱いまでであることがわかる。

表 3 には本研究会の 1 回から 20 回までの一般発表のタイトル一覧を示している。基礎から多様な応用まで様々なタイプの研究がこの分野でなされていることがわかるであろう。

### 5. 終わりに

本稿ではセマンティック Web とオントロジーに関わる研究状況を概観した。Web は未だ拡張の一途をたどっており、Web に関する研究の必要性は高い。本研究会のアプローチは人工知能と Web 技術の融合を目指すものであり、筆者は有望な方法であると信じている。

#### 参考文献

[Gangemi2009] A. Gangemi, N. Guarino, C. Masolo, A. Ultramari and L. Schneider: Sweetening Ontologies with

<sup>1</sup> いわゆるセマンティック Web に対して lower-case semantic web と称することもある。

DOLCE, Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web, pp. 223-233, 2002.

[Grenon2004] P. Grenon, B. Smith, and L. Goldberg: Biodynamic ontology: applying BFO in the biomedical domain, D.M. Pisanelli (ed.), Ontologies in Medicine, Amsterdam, pp 20-38, IOS Press, 2004

[Mizoguchi2009] R. Mizoguchi: Yet Another Top-level Ontology: YATO, Proc. of the Second Interdisciplinary Ontology Meeting, pp.91-101, Tokyo, Japan, February 28 - March 1, 2009

[Berners-Lee2000] T. Berners-Lee, Weaving the Web, Texere Publishing, US, 2000. (訳書: 高橋徹訳, Webの創成, 毎日コミュニケーションズ, 2001)

[Berners-Lee2001] T. Berners-Lee, J. Hendler, O. Lassila, The Semantic Web, Scientific American, 2001

[武田 2008] 武田英明: セマンティック Web と Linked Data, 技術研究報告, SWIM2008-22, [ソフトウェアインタプライズモデリング], pp. 25-28 電子情報通信学会 (2008).

[古崎 2006] 古崎晃司, 來村徳信, 溝口理一郎: Web2.0 時代のオントロジー利用雑感ーライトウェイトからヘビーウェイトまでー, 第 14 回セマンティック Web とオントロジー研究会, SIG-SWO-A602-06, 2006

表 1 オントロジーの利用目的と利用法による分類 [古崎 2006]

	知識モデルの利用法	メタデータの利用法
0)辞書的な利用	自然言語処理技術における利用	
1)共通語彙としての利用	キーワード解析	文献へのキーワード付加
2)インデックスとしての利用	Web検索用ディレクトリ	ドキュメントへのタグ付け
3)データスキーマとしての利用	データベーススキーマ	文献 DBの書誌情報
4)知識共有の媒体としての利用	参照オントロジーを用いたマッピング	コミュニケーション支援
5)知識モデルの規約としての利用	知識ベースシステム	セマンティックウェブ
6)知識の体系化への利用	知識の体系的管理	コンテンツの統合管理

表 2 オントロジー応用研究事例の分析[古崎 2006]

		オントロジーの種類(意味的構成要素)				合計	
		0)語彙	1) is-a	2)関係	3)意味		4)強い公理
オントロジー 利用形態	0)辞書的		1		2	3	
	1)共通語彙	1	1	1	1	4	
		11		9	1	21	
	2)インデックス		2	1		3	
			2	4		1	7
	3)データスキーマ	1		4		5	
		1				1	
	4)知識共有		3	5	4	12	
		1	9	3	1	14	
	5)モデル規約	2	1	12	7	22	
			4		4		
6)体系化				1	1		
		18	18	45	15	1	97

表3 セマンティック Web とオントロジー研究会 発表タイトル一覧

1	Webシラバス統合のためのレコード解析 人間行動と空間情報のセマンティクスの構築にむけて トピックマップとPublished Subject 時空間構造に基づくWeb検索の拡張	オントロジーエディタDODDLEファミリーの開発と利用 RDFデータを対象としたグラフ検索におけるクエリ生成方式の検討 部分順序プランによるセマンティックWebサービス合成
3	不均質なオントロジーに対する問合せ近似変換 「法道」におけるオントロジー分散開発 MMM (Meta-Model Management) プロジェクトの構想 DAML-Sに基づくサービス連携エージェントの実現について オントロジーの自動構築に関する基礎的研究 企業教育におけるセマンティックウェブの応用	12 Web Service Composition based on Task Repository Ubiquitous Metadata Scouter Semantic Web Services and OOP toward Unified Services
4	自然言語処理と蓋然的なセマンティックウェブ 水平統合とオントロジー モバイル環境下におけるサービス利用のためのタスク知識記述および実行環 Boostingを用いた評判の信頼性評価方法 RDF検索言語xRQLの提案と実装 バイオロジーにおける高次知識の体系化:パスウェイデータベースとオントロ デバイスオントロジーに基づくシグナル伝達機能オントロジーの構築 Webコンテンツの分析に基づくオントロジー構築および情報整理の試み 宇宙システムの開発・運用支援のためのオントロジーによる知識共有・知識連	13 ISO/IEC MFI Ontology registrationについて 感染症発見/監視システム BioCasterにおける専門用語アノテーションスキーマの構 WordNetとEDRのOWL表現 ビジネスオントロジーに基づく情報システム構築方法 Semantics Driven Development of Enterprise Software Based on Ontologies
5	空間の意味表現と空間機能検索 文書からのキーワード抽出と関連情報の収集 ナノテク材料技術の知識の構造化プロジェクトにおけるオントロジーの利用 利用者からのフィードバック情報を用いたオントロジー拡充技術 セマンティックグループウェア: RDFを用いたKnowWhoの実現 製品の状態モデルを利用したSemantic Webアプリケーションの構想 楽機能オントロジー リスプによるRDFSPプロセッサ セマンティック・ウェブサービス用エージェント Webコンテンツの機能に着目した検索結果の構造化に関する提案 セマンティックWebに向けてのWebブラウザによる意味的処理のイメージ 知識交流を支援する文書管理ツール 'CoreLibrary' テキストマイニングのための意味アノテーション 科学知マネジメントのための組織知メモリの構成	14 情報コンテンツの信頼性とその評価技術 プログラマ: 生活オントロジーを用いた留守番ロボット 「Web 2.0」からセマンティック・ウェブへの跳躍 概念化アスペクト: オントロジー構築の手掛かり 自然言語処理のための上位オントロジーと大規模オントロジーの試作 情報家電オントロジー構築に向けた取組み Web2.0時代のオントロジー利用雑感 →ライトウェイトからヘビーウェイトまで Folksonomyの3部グラフ構造を利用したタグクラスタリング ユーザのロールに応じたモバイルサービスナビゲーションシステム ユーザ興味オントロジーを用いたイノベーション検出手法
6	blogページの自動収集と監視に基づくテキストマイニング WebLogを対象とした評価表現抽出 Weblog情報を融合したコンテンツブラウジング Weblog解析に基づくWeb情報検索の信頼性向上技術 Weblogにおけるエゴセントリック検索の提案と実装 場log:Weblog環境における位置情報利用の提案 Blogコミュニティの抽出と分析	15 程度表現オントロジーの提案(1)コンセプトと設計 程度表現オントロジーの提案(2) 情報家電での適用例 共通テストベッドとしての家系図トピックマップ 「法道」によるオントロジーの分散共同構築支援 CGM分析技術の現状と課題 →メタデータ オントロジーの応用可能性について - オントロジー検索エンジンを用いた領域オントロジー構築支援環境DODDLE-OWLの リスプによるOWL-Fullプロセッサ 軽量オントロジーに基づくコンテンツ駆動型プラットフォーム
7	述語に基づく文書の木構造の解釈 複数の学会支援サービス群からのパーソナルネットワークの抽出及び視覚化 コンテキストウェアコンピュティングとコンテキストの定式化 組織知の発信支援~知の創造・継承活動モデルに基づく機能設計~ Integrating Deep and Shallow Semantic Structures in Open Ontology Forge Weblog研究の現状	16 感染症関連イベントオントロジーの構築 学術分野動向把握のためのオントロジー構築 学習・教授理論オントロジーの構築と利用~Theory-awareなオーサリングツールの A Method for Modeling Business Processes with OWL-T Language 固有名詞抽出技術を用いたオントロジー・メンテナンスツールの設計 競争製品や競合するサービスに関するオントロジーの構築とその利用 ナレッジ階層に基づく信頼オントロジーによる地域情報網の実現に関する一考察 許認可申請手続における行政手続オントロジーに基づく作業計画支援アプローチ EKOSS - An Ontology-based Semantic Web System for Knowledge Sharing
8	環境の共起情報を用いたWeb上からの個人メタデータ抽出 環境の多様な利用のための実世界セマンティクスに関する考察 ナレッジ階層に基づく概念化による思考過程支援の構想 Realization of seamless knowledge: connecting distributed RDF and Topic コンテキストの再利用に関する考察 柔軟なシステム連携を実現するためのインタフェースマッピングの提案と検証	17 OWLによる型付き単一化代入 ロール概念の成立条件とis-a関係の依存性についての基礎的分析 ロール理論の拡張によるインスタンス生成・消滅問題の考察 Network-Centric Healthcare Ontology Model 特許情報における発明者群と単語群の階層構造 インスタンス管理問題におけるidentityの取り扱いに関する考察 Semantics in QoS for Web Services: A Survey URI Context Databaseの提案 オントロジー拡充のための固有表現の共起情報を用いた語彙獲得 評判分析のための感性表現オントロジーメンテナンスツールの開発 汎用オントロジー構築における日本語Wikipediaの適用可能性 Report on LinkedData Planat Conference (tentative)
9	潜在的ウェブログコミュニティ抽出のための二部グラフ分割アルゴリズム セマンティックウェブサービスのためのタスク処理言語 表形式データからのオントロジーの獲得 設計の思考過程を表現するための概念処理モデルに関する検討 RDF/Topic Maps マッピングとアプリケーションへの適用 意味情報を活用した自動システム連携手法の提案と連携支援ツールの実装 オントロジーを利用した情報パーソナライズ方法の提案 オントロジーを用いたソフトウェア資産再利用法の検討 Semantic MediaWikiの構築に向けて SCORM2004準拠学習コンテンツの知識レベル設計支援 -知的学習支援プラットフォームに関するオントロジー工学的考察- 学会会議における位置情報コンテキストを用いたWeblog作成支援システムの 開発と運用	18 軽量オントロジーによるセマンティックプログラミング 言語に基づく推論のための事象間関係知識データベースの構築 臨床医学オントロジーの構築とその基本思想 機能的知識外化ツールOntoGearを用いた失敗知識の構造化と活用 Wikipediaを用いた用語説明のモデル化と事典的検索への応用 Wikipediaエントリ構造抽出ツール: Wik-IE WikipediaとFolksonomyタグに基づくドメインオントロジー構築支援環境の実現と評価
10	Semantic WikiによるRDF自動生成 なんでもRSS - HTML文書からのRSS Feed自動生成 Blogを基盤とした情報流通システムの提案と実装 オントロジー構築・アノテーション支援ツールOpen Ontology Forge 人間関係オントロジー サーベイ: Blog研究の現在 2005	19 Webサイトナビゲーションのためのオープンカードソートを用いた分類体系構築 レシピの構造を反映したメタデータに基づく部分レシピの再利用法 イベント構造の抽出に基づく画像管理法 タクソミを活用したメールに潜むKnowHow, KnowWho可視化技術 Use of Ontology in Text Classification セマンティック・ウェブ上の語彙構築機構 軽量オントロジーによるセマンティックプログラミング 言語に基づく推論のための事象間関係知識データベースの構築 臨床医学オントロジーの構築とその基本思想 機能的知識外化ツールOntoGearを用いた失敗知識の構造化と活用 Wikipediaを用いた用語説明のモデル化と事典的検索への応用 Wikipediaエントリ構造抽出ツール: Wik-IE WikipediaとFolksonomyタグに基づくドメインオントロジー構築支援環境の実現と評価
11	Context間の関連性に着目した民俗学のためのOntology構築 ローマ法の現代的慣用時代の法学位論文における師弟関係と主題のメタ A Consideration of Real-World Semantics for Various Utilization of Environment(2) システム開発プロセスにおけるオントロジー利用の検討 オブジェクト抽出のためのオントロジー構築法の提案 コラボレーション環境におけるメタデータ共有システムの開発	20 Wikipediaを用いた異分野知識ベース群の連携 Wikipediaからの連想シソーラス構築プロジェクト ウィキペディアの成長における秩序と多様性 マスコラボレーションにおけるコンテンツ形成プロセスの分析 意外性のある知識発見のためのWikipediaカテゴリ間の関係分析 Wikipediaの編集履歴を用いた記事の信頼性導出 An Information Recommendation Model Based on Concept Classes Extracted from Wikipedia Categories Wikipediaを利用した音声認識用語モデルの構築および評価 Wikipedia概念体系を用いた日本語ブログ空間のトピック分布推定 Wikipediaからの拡張クエリ生成によるWeb検索とその評価 多型トピックモデルを用いたWikipedia検索 Wikipediaを用いた多言語情報アクセスに関する研究: 言語間リンクの分析と応用