

インターネット時代におけるオントロジーと 業務システムへの適用について

Ontologies in the era of Internet and their application to business information systems

岡部雅夫[†]

[†]東京電力株式会社 システム企画部

要旨

Web 上の膨大な情報のより有効な活用のためにセマンティック Web に代表されるオントロジー化が注目されている。ただし、業務システムの連係・統合のための活用を考えた場合、記述言語・用語に依存しない緩やかなオントロジーのハーモナイゼーションが求められていると同時に、信頼性を担保できる枠組みが必要である。MMF Ontology Registration は、その支援のために、ISO/IEC において検討されているものである。

1. はじめに

インターネットの普及に伴いユーザが利用できる情報は飛躍的に増大し、その情報のより高度な活用を目指して、セマンティック Web に代表されるように、オントロジー化が試みられている。ただし、無統制なインターネット上の情報は、特に、業務システムでの活用を考えると、その信用性を如何に担保するか等の課題も多い。本稿では、インターネット時代における業務システムの連係・統合という観点から、現在のセマンティック Web に代表されるオントロジーの課題を考察するとともに、その解決のための方策として、ISO/IEC の場で検討されている MMF Ontology Registration について簡単に紹介する。

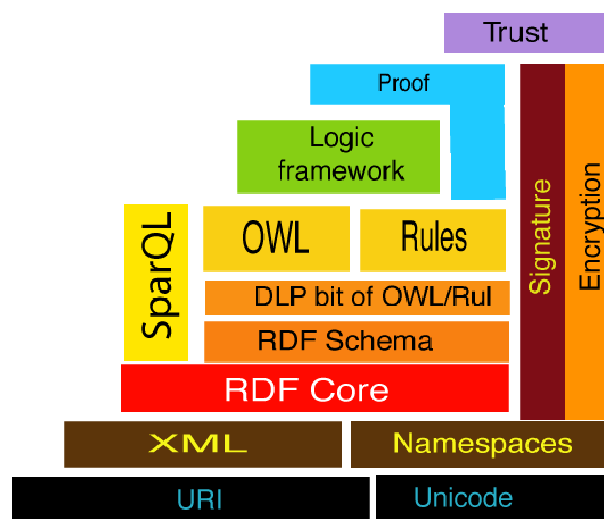
2. セマンティック Web について

1988 年ごろに Web の創始者である Tim Berners-Lee により提唱されたとされるセマンティック Web は、その後、大きなうねりとなり、現在のオントロジー・ブームの根幹を成している。よく引き合いに出される例に、以下の2つがある。

- ・ エージェントが、ディナーパーティの料理に適したワインを教えてくれる[1]。
- ・ エージェントが、母の病気に際し、適切な病院を見つけ、アポイントを取ってくれる[2]。

ただし、必ずしもこのようなパーソナル・ユースのみを意識したものではなく、その主旨は、コンピュータが Web 上の情報を理解し、処理することにあると思われ、いわゆる業務システムによる処理・活用も包含したものであると考えるのが自然である。セマンティック Web の全体構造は、図 1 のように示される。これらの内、下位の要素技術から固まり

つつあり、W3C では、記述言語の基本となる RDF や OWL の標準は既に定め、また、問い合わせ言語として SPARQL の仕様も概ね固まっている。ただし、これだけで、オントロジーがすべて記述できるわけではなく、SWSL(Semantic Web Service Language)等の様々な拡張が検討されている。さらに、より根源的には、そのオントロジーの信頼性をどのように担保するかという問題があり、そのための機能は、Trust という最上位の層において提供されることになっているが、その具体像は明らかではない。



出所：<http://www.w3.org/2005/Talks/0511-keynote-tbl/>

図 1 セマンティック Web の階層構造

信頼性を担保するための一つのアプローチとして、McGuinness 他が提唱する InferenceWeb[3]がある。これは、得られた結論が、どのような情報を信頼し、また、どのような推論アルゴリズムを利用することにより得られたものであるかを明示的に示そうというものである。そのための基本となるものが IW base と呼ばれるレジストリで、前提とした情報や推論アルゴリズムに関するメタデータが登録されている。その基本的な考え方として、あくまで結論を得るための基となった情報を明示的に示すことを目的としており、それを以ってその結論を信頼するか否かの判断はあくまでユーザにゆだねられている。本来的には、確かに最終的な判断はユーザにゆだねられるべきものである。ただし、実務的に考えると、例えば、料理にあったワインの選定であれば、多少リスクのある情報を使っても、もしかすると飛び切り良いワインが選択できれば、ためしにやってみるかもしれない。ただし、病院の選択であれば、通常は、リスクのある情報をもとにした選択はおこなわないであろう。また、業務システムからの利用であれば、画一的に信頼性を担保できるものでなければ、実際的には利用は困難であろう。

3. 業務システムの連携・統合について

一方、業務システムの連携・統合は、古くからの課題である。インターネットが普及し、Web サービス等も含めると、連携・統合が求められる世界は、限りなく広範になっている。

従来、業務システムの連携・統合を実現するためには、データ辞書の整備が中心となってきた。業務システム間でやり取りされる用語（データ項目名等）の標準化を行い、それをデータ辞書の登録管理しようというものである。ただし、現実には、統一的なデータ辞書の整備は、1企業においても、多様な業務を扱う大きな企業体になると、なかなか困難であるのが実態である。

そのような実態を踏まえ、1企業体ないし特定業務領域に於いても、多様な用語の存在を認め、用語の標準化の代わりに、その用語の意味をフォーマルに記述することにより、業務システムの連携・統合を実現しようとする動きがある。即ち、連携・統合を目的とした業務システムの仕様のオントロジーによる記述である。

例えば、米国・国防総省では、かつて統一的なデータ辞書を構築しようとしたものの、現実には組織別のデータ辞書の構築にならざるを得なかった反省を踏まえ、大規模な商用オントロジーである Cyc を基に、データベースのスキーマを中心にアプリケーションの仕様の記述を試みている[4]。各業務システムで扱う用語が統一されていなくとも、それらが Cyc をベースに定義されているために、他の業務システムにおいても解釈が可能で、利用可能となる。

また、ISO TC184/SC4 では、製造業におけるアプリケーションソフトウェアの統合を目的とし、PSL と呼ばれる KIF に基づく製造業の分散型プロセスのオントロジーないしその基礎となる用語の定義を、膨大な量の ISO18629 シリーズとして発行しつつある[5]。

即ち、業務システム間の連携・統合を図るため、業務システムの仕様をオントロジーとしてフォーマルに記述しようとする試みは、地道にはあるが、着実に進みつつある。ただし、そのオントロジーを記述する言語は、必ずしも、RDF/OWLに限らない。国防総省の例であれば、CycL (Cyc を記述している独自の言語) であり、PSL の場合は KIF であり、その内容に応じて適したものが選ばれている。

4. MMF Ontology Registration

結局、インターネット時代の開かれた環境を前提に、業務システムの連携・統合という観点からオントロジーを見た場合、2つの問題点が存在する。

第1に、オントロジーが必ずしも RDF/OWL に限らずさまざまな言語で作成されていること。

第2に、さまざまなオントロジーが無統制に作成され、その信頼性が定かではないこと。

このような問題を解決するための一方策として、現在、ISO/IEC において検討されている MMF Ontology Registration[6]の概要を紹介する。このプロジェクトは、現在、ISO/IEC JTC1 SC32/WG2 において国際標準の策定が進められている MMF プロジェクト[7]の1サブプロジェクトをなすものである。

1 番目の問題に関しては、すでに見たように W3C のコミュニティーにおいてもすでにさまざまな括

張がなされようとしていること、CycL, KIF 等それぞれの目的・程度に応じて様々な言語が使用されていることを考えると、言語の多様性を認めることが必要であろう。実際、OMG においては、Common Logic、RDF, OWL, Topic Maps 等の幾つかの言語に対応したメタモデルおよびそれらの間のマッピングを定めようとしている。そこで、MMF Ontology Registration では、このようなりポジトリの存在を前提に、言語に極力依存しない形でオントロジーの登録・管理を行う。図 2 に示されるように、一般にオントロジーは様々な文の連言として構成され、また、それぞれの文は、様々なシンボルから構成される。MMF Ontology Registration は、単純に、このレベルでの構成管理を行い、それぞれに対して、管理上必要となるメタデータを保持する。オントロジーの実際のセマンティックスが必要な場合は、基本的には、URI をベースに、ODM 等に則り実際にオントロジーを保持するリポジトリにアクセスすることになる。その全体像を示すと図 3 のようになる。即ち、セマンティックスに対応するシンタックスを保持しないことにより、セマンティックスは保持できないが、最低限の構成情報をシンタックスに独立に保持することが可能になる。

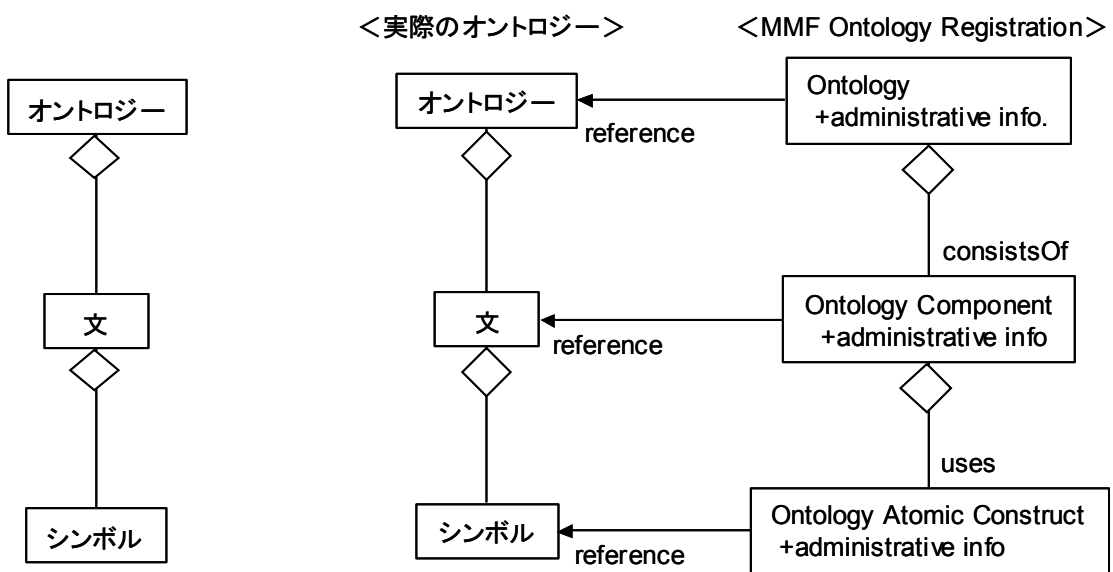


図2 オントロジーの一般的な構造

図3 実際のオントロジーとMMF Ontology Registrationの関係

2番目の問題に関しては、MMF Ontology Registration は、基本的には、登録・管理によりその信頼性を担保する。登録・管理されるオントロジーは、参照オントロジーおよびローカル・オントロジーに大別される。各業界において標準化され、業務システムから見て常に信頼にたるオントロジーであると認められた場合のみ、参照オントロジーとして登録される。これにより、各業務システムからみて、画一的に「正しい」と判断されることになる。一方、ローカル・オントロジーは、参照オントロジーをベースに、内容の追加・削除を含め、ある業務システムのためにローカライズされたものである。ローカル・オントロジーは、参照オントロジーとの関係においてその内容が保証される部分もあるが、基本的には、他の業務システムがそれを利用しようとする場合、それを信頼するか否かは、その業務システムの責任と判断による。図 3 の汎用的な構造を前提に、参照オントロジーとローカル・オントロジーを導入すると、MMF Ontology Registration のメタモデルの概要は、図 4 のようになる。結局、MMF Ontology Registration は、オントロジーの記述言語・用語に依存しない緩やかなハーモナイゼーションを支援する一方で、オントロジーの信頼性に関しては画一的な判断が可能な枠組みを提供することにより、業務システムによる連係・統合のためのオントロジーの利用を促進しようとするものである。

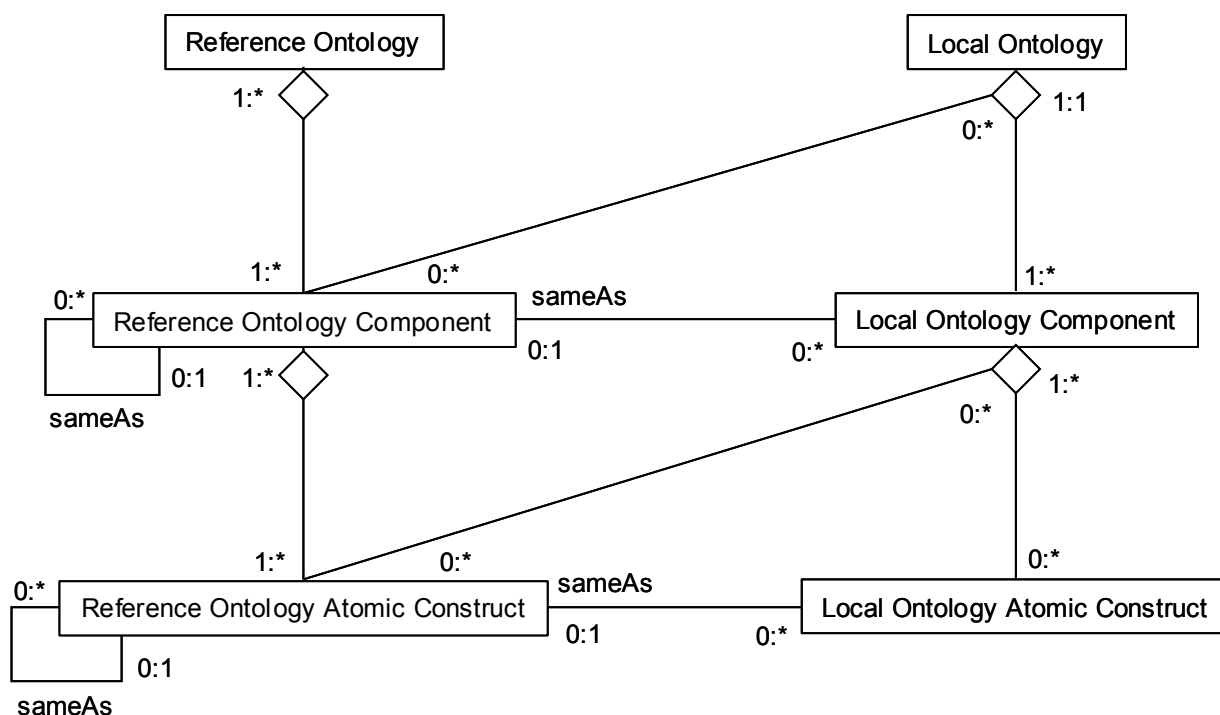


図4 MMF Ontology Registrationのメタモデルの概要

5. 今後の課題

業務システムの連携・統合という観点から現状のセマンティック Web に代表されるオントロジーの課題を述べ、その一つの解決策として、MMF Ontology Registration の概要を紹介した。現状、MMF Ontology Registration は、緒に着いたばかりの非常にプリミティブな仕様であり、直ちに業務システムの連携・統合を可能にするようなものではない。オントロジーの保持するセマンティクスを実際に活用するには、ODM 等に則った実際にオントロジーの内容を保持するリポジトリが必要になる。また、参照オントロジー間に、相矛盾する記述があり、それが異なる用語で記述されている場合に、それをどうやってチェックするかといった本質的な問題も存在する。とはいえ、MMF Ontology Registration のように記述言語に依存しない汎用的なレジストリは、必ず必要になると思われる。実質的に、MMF Ontology Registration が機能するか否かは、その仕様以上に、標準化されたオントロジーを登録・管理を担う機構が立ち上がり、かつ、認知されるか否かに係わっているように思う。

参考文献

- [1] <http://www.w3.org/TR/2004/REC-owl-guide-20040210/#Usage>
- [2] Berners-Lee, T., Hendler, J., Lassila, O., "The Semantic Web, A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities," Scientific American, May 2001
- [3] McGuinness, D.L. and da Silva P.P., "Explaining Answers from the Semantic Web: The Inference Web Approach," Journal of Web Semantics, Vol1 Issue 4, 2004
- [4] Peterson, B.J., Anderson, W. A., and Engel, J., "Knowledge Bus: Generating Application-Focused Databases from Large Ontologies", Proceedings of the 5th KRDB Workshop Seattle, WA. 31-May-1998, pp.2-1 – 2-10
- [5] 西岡靖之, "離散型プロセス記述のためのオントロジー : PSL", 人工知能学会誌 19 巻 2 号(2004 年 3 月) pp167-171
- [6] ISO/IEC FCD 19763-3 Information technology – Framework for metamodel interoperability – Part 3: Metamodel for ontology registration
- [7] ISO/IEC FCD 19763-1 Information technology – Framework for metamodel interoperability – Part 1: Reference model