

特集4

セマンティックWebのツール

森田 幸伯^{†1} 沖電気工業(株)
 津田 宏^{†2} (株)富士通研究所
 清水 昇^{†3} 日本電気(株)
 布目 光生^{†4} (株)東芝
 来間 啓伸^{†5} (株)日立製作所
 佐藤 宏之^{†6} NTT情報流通プラットフォーム研究所

■背景

セマンティックWebは、Web上のデータにメタデータやオントロジを付与することにより、相互運用性やデータの高度利用を可能にする技術である。

セマンティックWebの実現においては、あるコミュニティ(電子政府や業界団体)でメタデータのスキーマ(エレメントセット)を決め、それに従ってメタデータを付与・運用していくという作業が必要となる。HTMLについては、すでにWYSIWYGのオーサリングツールも多く、作成したページを目で見れば、誤りが分かたり満足度が得られたりというフィードバックもある。それに対し、RDFによるメタデータは、フォーマットも複雑で、内容に問題があってもサービスに組み込まれないと分からない場合もある。しかし、膨

大で良質なメタデータを継続的に付与、更新していくことができなければ、セマンティックWeb社会は実現できない。メタデータの付与については、人工工数の削減は大きな課題となる。さらに、上記メタデータの付与は、スキーマを定めた後の支援であるが、セマンティックWebでは、スキーマを含めたオントロジを記述することでより柔軟なシステムの構築が可能になるが、その際のオントロジに対する構築のサポートも重要である。本稿では、メタデータ付与の支援ツールを中心にセマンティックWeb関連ツールの現状を紹介する。

■ツールの概観

図-1にセマンティックWeb関連ツールの階層を示す。ここでは、セマンティックWebの発展段階に従って、大量のデータに対してメタデータを効率的に付与することを目的としたメタデータ構築支援ツールと、(スキーマ構築も含む)オントロジの構築支援ツールに分類している。セマンティックWebでは、オントロジを含むメタデータはRDF(Resource Description Framework)で格納する。機能階層としては、RDF構文を解析する層、RDFの格納管理を行う層、手入力を前提とした内容の視覚化・編集およびメタデータの半自動付与を支援するメタデータ構築支援の層、メタデータ・オントロジを管理・共有する層、およびそれらの上位でエー

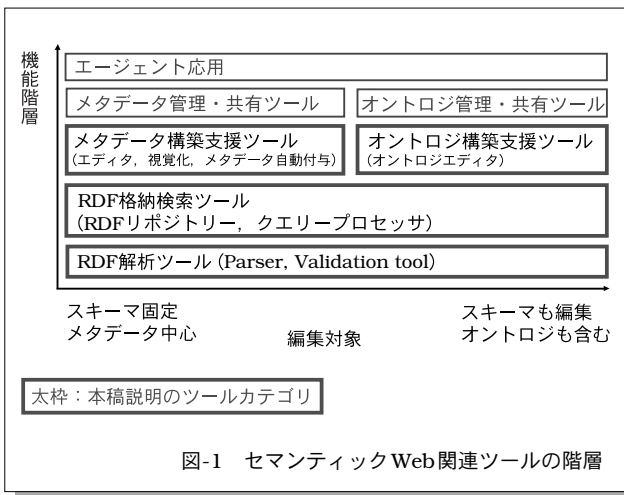


図-1 セマンティックWeb関連ツールの階層

^{†1} E-mail: morita454@oki.com

^{†4} E-mail: kosei.fume@toshiba.co.jp

^{†2} E-mail: htsuda@jp.fujitsu.com

^{†5} E-mail: kuruma@sdl.hitachi.co.jp

^{†3} E-mail: n-shimizu@ce.jp.nec.com

^{†6} E-mail: sato.hiroyuki@lab.ntt.co.jp

名称	開発元	対応メタデータ	出力フォーマット	URL	備考
(メタデータエディタ)					
DC Metadata template	Nordic Metadata Project	DC	HTML META	http://www.ub.lu.se/metadata/DC_creator.html	
Reggie Metadata Editor	Distributed Systems Technology Centre	DC, スキーマファイルによりAGLS等も可	HTML META, RDF, RDF Abbreviated	http://metadata.net/dstc	
MetaWeb Generic Edit Tool	State Library of Tasmania	DC, AGLS, ADMIN Core	HTML META	http://www.dstc.edu.au/RDU/MetaWeb/generic_tool.html	
Metabrowser	Spirit社	DC, GILS, AGLS, NZGLS	HTML META, RSS など	http://metabrowser.sprit.net.au/	統合メタデータ管理環境
(メタデータ生成ツール)					
DC-dot	UKOLN. University of Bath	DC (USMARC, TEI headers, GILS等にも変換可)	HTML META, RDF	http://www.ukoln.ac.uk/metadata/dcdot/	
TagGen	HiSoftware社	GILS, WAGILS, DC	HTML META	http://www.hissoftware.com/taggen.htm	
Klarity	tSA社	DC	HTML META	http://www.klarity.com.au	

表-1 メタデータ構築ツール

ジェントを用いた応用層に分類してみた。

本稿では、紙幅の関係もあり、メタデータ付与ツールを中心に図中の太枠の部分について、主なツールを概観する。まず、メタデータ構築支援ツール(エディタ, 自動付与)およびメタデータを記述するRDFを解析・検証するRDF解析ツールの現状を紹介し、RDFの検証機能を含むツールの例としてRDF Analyzerについて概要を述べる。次にオントロジ構築支援ツールとしてオントロジエディタの現状を紹介し、最後に、メタデータの応用の例として、注釈(アノテーション)付加ツールについて現状を紹介する。また、その際に必要となるRDF格納検索ツールであるRDFリポジトリとクエリープロセッサについて述べる。

■メタデータ構築支援ツール

コンピュータによるメタデータ付与は、膨大な人手工数の削減し、人手による質のバラツキを少なくする意味でも重要となる。従来の自然言語処理、機械学習¹⁾、Webマイニングといったソフトウェア技術の大きなターゲットともなり得る。メタデータ構築支援ツールには、以下のような2つのレベルがある。

(1) メタデータの入力支援(メタデータエディタ)

テンプレートによりメタデータ入力を補助したり、メタデータのフォーマット出力(RDFやHTMLのMETATagなど)を自動化することで、人手によるミスや労力

を減らすものである。

(2) メタデータの自動付与(メタデータ生成ツール)

メタデータの属性値をシステムが自動で入力してくれるものである。メタデータのエレメントにより、自動化の精度はまちまちである。たとえば、Dublin CoreのSubject(関連サブジェクト)の値は、本文に自然言語処理やテキスト分類技術を適用して付与するとか、Publisherの値を情報抽出技術で取り出すなどが考えられる。ただし、こうした自動化は不完全になるのはやむを得ないため、実際はエディタ機能も含んでいるものが多い。

表-1にこのメタデータ構築支援ツールの主なものをまとめた。多くが電子政府のメタデータに関係しており、特にメタデータでは先進的とされるオーストラリアのソフトウェアが目につく。まだRDF出力に対応しているものは少ないが、今後セマンティックWebの普及とともに必須機能となるだろう。

メタデータ自動付与の応用

メタデータ生成ツールとその応用として、自動Webディレクトリ²⁾を紹介しよう。これは、Yahooのようなディレクトリ自動化に向けて、情報の入口となるメニュー的なページで最近話題のページを選別できるようなメタデータを、コンテンツ処理(自動分類、情報抽出)だけでなくハイパーリンクの解析で付与しているのが特徴である。Dublin Coreをベースに、Rel_Location

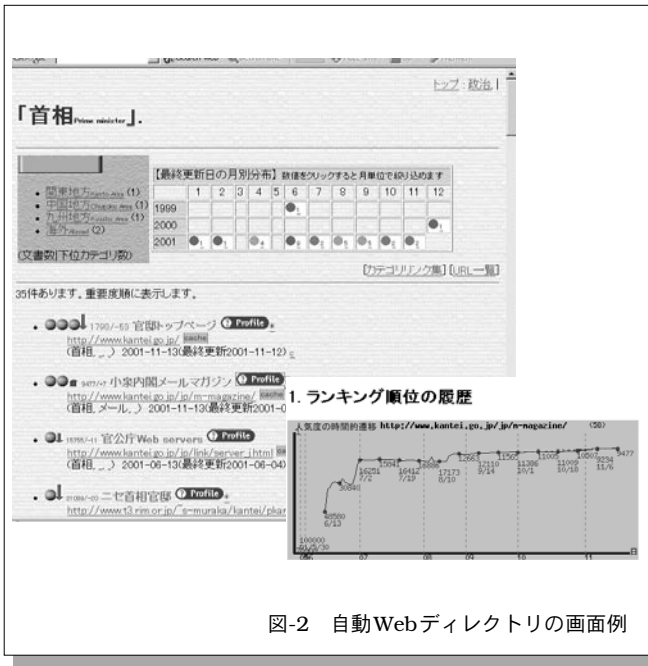


図-2 自動Webディレクトリの画面例

(関連地域：本文から地名情報抽出で実現)，Rank (リンクベースの人気度およびその動きを回帰分析した値)などを付与している。

図-2は、このメタデータを利用することで構築した地域/ジャンル/時間による多観点のディレクトリインタフェース画面である。「政治-首相」カテゴリにおいて、1位の「官邸トップページ」は定番の大きく人気度順位は減少中、2位の「小泉内閣メールマガジン」は人気度が上昇傾向ということが見てとれる。図左下には、この人気度の動きを算出するのに用いた2001年6月から11月までの小泉内閣メールマガジンの人気度順位の推移を表している。メールマガジン発刊(6月中旬)直後の伸びはすさまじいものの、段々伸びは鈍化していることが分かる。いまだに爆発的増加が続いていて、変化も速いWeb世界においては、このようなメタデータ自動生成手法はますます必要といえよう。

■ RDF解析ツール

RDF文書の解析ツールとして、Art Barstowが作成したW3C(World Wide Web Consortium)のWebページにて公開されている「RDF Validation Service」がある³⁾。図-3にその画面例を示す。

これは、テキストエリアにRDF文書を入力すると、出力としてグラフ表示やRDFの3つ組(triple)リストが得られるものである。

入力にはテキストエリアへの直接記述のほかにURIによる指定も可能である。また、出力オプションとしてノ

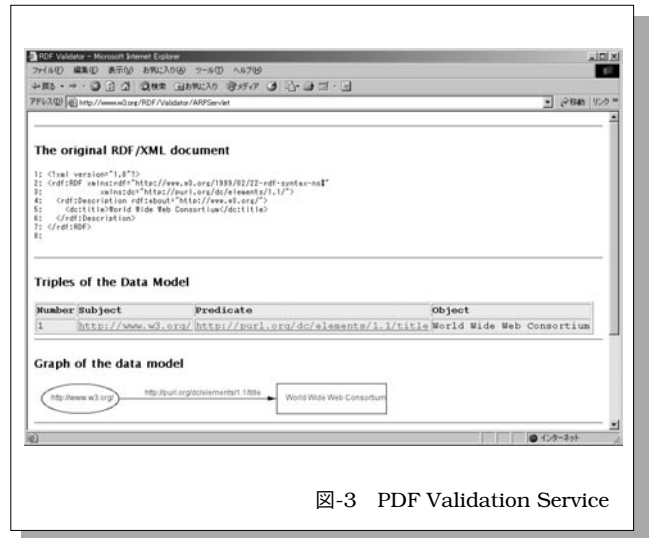


図-3 PDF Validation Service

ードやアークの色やラベルのフォントサイズ、グラフの出力フォーマット(PNG, GIF, PS等)といった細かい指定が可能である。

本ツールはJavaServletとして実装されており、W3CのWebページからブラウザを介して利用できるほか、ソースコードが公開されているためローカルな環境で動作させることもできる。また、従来は解析エンジンとしてSiRPAC(Simple RDF Parser & Compiler)が用いられていたが、現在はARP(Another RDF Parser)⁴⁾を基盤にしている。

ベースとなっている技術は次の2点である。

- RDF Parser (version 1.0.5 of the Another RDF Parser (ARP). Jeremy Carroll at HP-Labs in Bristol.)
- グラフ描画 (1.7.4 of the AT&T Labs GraphViz open source graph drawing software⁵⁾.)

RDF解析エンジンであるARPの特徴としては、

- Javaベースの解析器
- daml:collectionのサポート
- xml:lang, xml:baseのサポート
- URI参照のRFC2396に基づく検証機能
- XML Names XML Names仕様に基づくrdf:IDおよびrdf:BagID'sの検証機能
- Relative Namespace URI references W3C XML Primary decisionに基づくネームスペース内URI参照の検証機能
- aboutEach rdf:aboutEachのサポート(ユーザ定義型を含む)

といった項目が挙げられる。ただ、ツール自体としてRDFスキーマ制約を考慮した検証機能は備えていない。

その他の解析ツールについては表-2にまとめた。

名称	概要
Libwww ⁶⁾	SIRPACと expat (James Clark'sによるXML Parser) をベースとしたC言語ライブラリ。John Puninによる。
Repat ⁷⁾	callbackベースなRDF parserツールキット。こちらも expat を基にC言語で実装されている。Jason Diamondによる。
Perllib ⁸⁾	W3CのPerlコードライブラリ。アクセスコントロールやアノテーションでの利用を想定している。設計・実装はEric Prud'hommeauxによる。
RDF Filter ⁹⁾	JavaとSAX2ベースのシンプルなRDF parser。解析木をメモリ上に保持しないため、大規模文書の扱いに適している。
SWI-Prolog RDF parser ¹⁰⁾	PrologベースのRDF parser。“RDF Validation Service”と同様ブラウザを介してのサービスが提供されている。
VRP ¹¹⁾	ICS-FORTH RDF Suite中のツールの1つでありSofia AlexakiによるJavaベースのparser。RDFスキーマ制約条件に基づいた検証が可能。

表-2 その他の解析ツール

■ RDF Analyzer

RDF Analyzerは、セマンティックWebに関する国産ツールの1つである。図-4にその画面例を示す。

RDF Analyzerの持つ主な機能は、次の5つの機能である。

- ① RDFのBNF¹²⁾, ¹³⁾定義に基づいてRDFデータの解析処理を行う機能
- ② RDFデータの構文の厳密な妥当性検証機能
- ③ RDFデータの意味を人間が理解しやすい文章に翻訳する機能
- ④ RDFのみならずRDFS (Resource Description Framework Schema)¹⁴⁾, Dublin Core¹⁵⁾ および DAML+OIL¹⁶⁾のデータを処理する機能
- ⑤ RDFサーバアプリケーション用にRDFデータ変更用のフォームを生成するCGI機能

RDF Analyzerの開発の狙いを以下に示す。

セマンティックWebは、RDFを用いて、Webリソースに関するマシンリーダブルな情報を作成し、その情報に基づいて自動処理を行うことを目指している。

一般にマシンリーダブルな情報は、Webページ等のヒューマンリーダブルな情報を元に作成される。

生成直後には、マシンリーダブルな情報は、その元になったヒューマンリーダブルな情報と一致していても、お互い独立に情報の更新を行っているため、その一致が崩れる可能性がある。

セマンティックWebは、マシンリーダブルな情報と

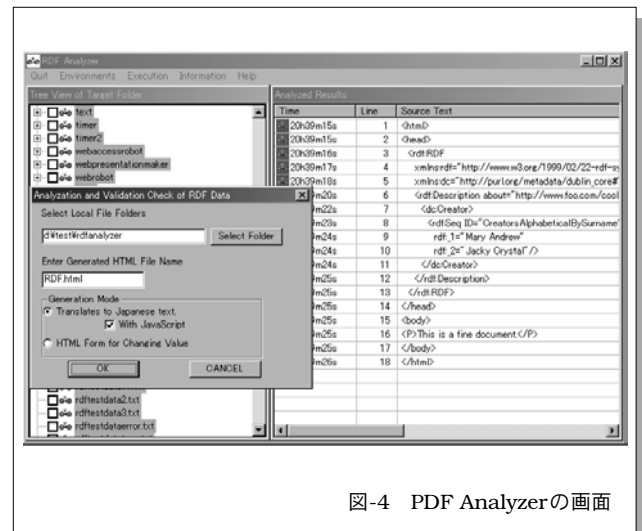


図-4 PDF Analyzerの画面

ヒューマンリーダブルな情報とが一致していることを暗黙の前提としており、その一致が崩れることは、致命的な問題となる。

しかし、RDFにより記述されたマシンリーダブルなデータは、人間が、それを読み理解することが容易な記述形式になっていない。

このため、マシンリーダブルな情報とヒューマンリーダブルな情報との間に乖離が生じて、それを検知することが難しい。また、RDFの記述が誤っていても、なかなか分からないなどの問題がある。

これらの問題に対処するため、RDF Analyzerでは、RDFデータの構文の厳密な検証機能と、RDFの記述内容を人間に理解しやすい文章に翻訳する機能等が実現されている。

- RDFのBNF定義に基づいてRDFデータの解析処理を行う機能

RDFの構文の唯一無二の正式な定義は、Resource Description Framework (RDF) Model and Syntax Specification¹⁷⁾に記述されているFormal Grammar for RDFのEBNF定義である。

したがって、このEBNFの定義に照らして、RDFデータの構文チェックを行うならば、正確なRDFデータの構文チェックが実現できるはずである。

RDF Analyzerは、EBNFをよりソフトウェアで処理しやすい形式に独自に改善したMBNF (Modified Backus-Naur Form) にマッピングし、それに基づいて処理を行うことで、RDFデータの解析処理と厳密な妥当性検証機能とを実現している。

- RDFデータの意味を人間が理解しやすい文章に翻訳する機能

RDFデータの記述は、人間が理解し難いため、RDF

の定義内容をグラフで表示する各種のグラフ化ツール¹⁸⁾が開発されている、しかし、グラフ表示の場合、次の2つの問題がある。

- ①グラフが複雑になると、それもまた人間にとって理解が難しい。
- ②マシンリーダブルの情報の元になったヒューマンリーダブルな情報は、テキストデータであることが多いが、それとの比較がしづらい。

したがって、RDFデータの記述内容を文章データに翻訳できるならば、人間が理解しやすく、オリジナルなヒューマンリーダブルな情報と比較することも、簡単になる。

RDF Analyzerは、日本語のWebページの情報をRDFにより、メタデータ記述することを想定しているので、RDF Analyzerにより、RDFデータを日本語の文章に翻訳する機能を実現している。

- RDFのみならずRDFS, Dublin CoreおよびDAML+OILのデータを処理する機能

現時点において、セマンティックWebで用いる予定のスキーマにはRDFS, Dublin Core, OWLおよびDAML+OILがある。

これらのスキーマで定義された語彙が処理できなければ、RDFデータを文章に翻訳したとしても、分かりやすい文章にすることはできない。

このため、RDF Analyzerは、これらのスキーマを処理する機能を標準機能として具備している。

- RDFサーバアプリケーション用にRDFデータ変更用のフォームを生成するCGI機能

セマンティックWeb技術は、いろいろなアプリケーションで活用されることが期待されている。

それらのシステムでは、当該システムのすべてのデータにメタデータを付加し、そのメタデータに基づいて、処理を行い、より効率的なシステムを実現する。

このようなシステムでは、サーバ側に格納されているRDFで記述されたメタデータの値を参照し、変更することが、しばしば、必要となる。このため、サーバサイドのCGI用に、RDFデータの値を変更可能とするHTML Formを生成できる機能をRDF Analyzerは有している。

なお、RDF Analyzerは、フリーソフトウェアとしてINTAPのWebページ¹⁹⁾で公開されている。

■オントロジエディタ

前述のメタデータ付与ツールが、定められたスキーマに従って、大量のデータに対してメタデータを付与

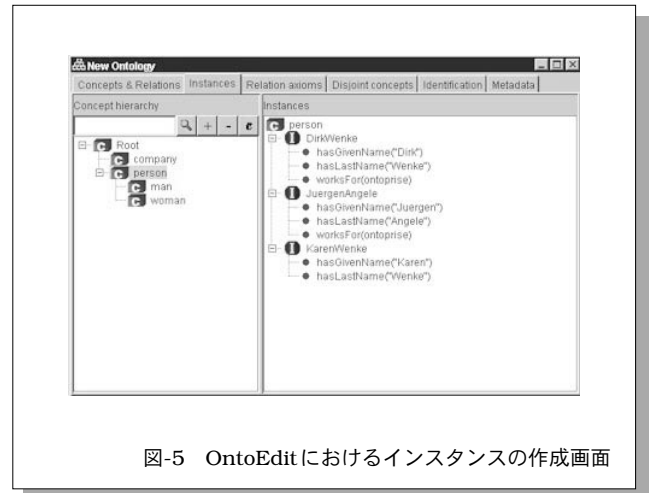


図-5 OntoEditにおけるインスタンスの作成画面

することを主目的としているのに対し、オントロジエディタでは、スキーマを含むオントロジの記述を支援することを主目的にしているツールである。オントロジでは概念階層や、公理系などの記述が可能であるため、概念階層の記述や公理系の記述を支援することが重要である。

オントロジそのものは、セマンティックWeb以前から存在する。Amsterdam大学Social Science Informatics (SWI)のWonderTools²⁰⁾プロジェクトの1999年の報告²¹⁾では、オントロジエディタの実験システムとして、Ontolingua²²⁾、WebOnto²³⁾、ProtégéWin²⁴⁾、OntoSaurus²⁵⁾、ODE²⁶⁾、KADS22²⁷⁾が挙げられている。このうちスタンフォード大学のProtégéWinは、現在はProtégé2000に改版され、RDFやRDFスキーマもサポートするようになっている。

オントロジエディタの例として、セマンティックWebの研究プロジェクトの成果に基づくOntoEdit²⁸⁾を紹介する。OntoEditはカールスルーエ大学のAIFBが開発し、独Ontoprise社が商用化したもので、オントロジの生成、処理等の機能を有し、さらにRDFや、DAML+OIL、F-Logic (Flame-Logic)の各形式に対する入出力機能を有している。OntoEditのオントロジ作成画面を図-5に示す。

OntoEditでは、情報を階層的に表示できる。このため、概念階層を辿りながら必要な情報を編集できる。図-5は、人(person)という概念の下位概念として女性(woman)や男性(man)を定義し、さらに人(person)という概念の実体(インスタンス)として3名を定義している画面である。

また、OntoEditでは、F-Logicに基づく述語論理の処理系を組み込んでおり、制約や推論規則を公理という形で述語論理により記述でき、制約に反した記述の検

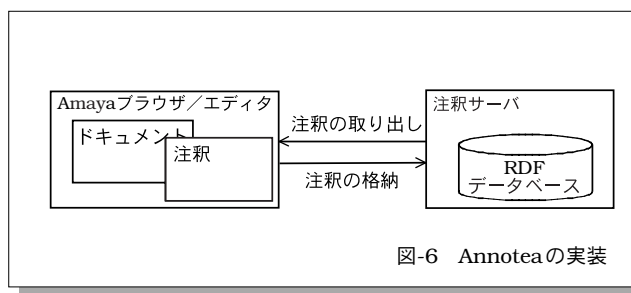


図-6 Annoteaの実装

出や、推論規則に従った記述の導出が可能な Plug-In が用意されている。しかし、F-Logicによる公理の記述は、利用者には難解である場合が多いため、公理を設定するために、「背反 (disjoint) の概念^{☆1}」や「対称的 (symmetric) ^{☆2}」「反対 (inverse) ^{☆3}」「推移的 (transitive) ^{☆4}」な「関係」を指定できるようになっている。たとえば「女性」と「男性」は背反の概念であるという公理を F-Logic で直接記述せずに、「女性」と「男性」の概念間は「背反」であると設定することにより公理とすることができる（すなわち、女性でありかつ男性であることはないという制約を記述できる）。これにより「女性」かつ「男性」を継承する実体を定義した場合、それが公理に反することを検出することができる。

■注釈付加ツール

ドキュメントに加えられた注釈（アノテーション）は、そのドキュメントに対するメタデータの種類と解釈することができる。注釈には、ドキュメントの特定の箇所に対するコメントのほか、一般にはドキュメントの概要や構成、閲覧条件、ドキュメントを表示する際に必要なハードウェア条件なども含まれ、読者に付加的な情報を与えるだけでなく、ドキュメントの検索や表示、再利用において計算機が使用するための情報も与え得る。このような意味で、ドキュメントに注釈を付加するツールはセマンティック Web を構成するためのツールの1つと考えることができるが、注釈の構造や拡張性、他のシステムからの利用可能性などの点においてツールごとによりかなりの違いがある。

現在、商用のものを含めてさまざまな注釈付加ツールが発表されているが、ここではセマンティック Web の枠組みとの親和性を考慮して作られた、Annotea と OntoMat を紹介する。

Annotea

Annotea^{29), 30)} は、Web ドキュメントを対象とした共有型の注釈ツールである。注釈は、ドキュメントに対

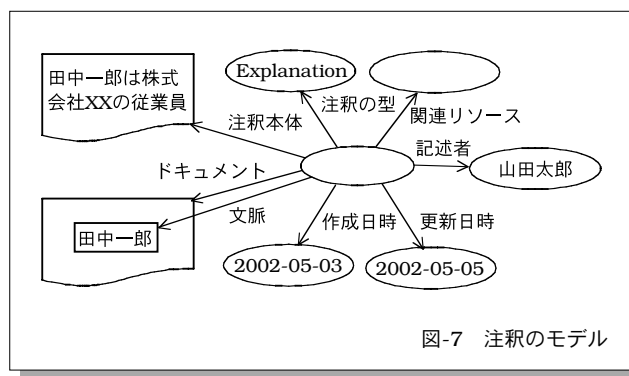


図-7 注釈のモデル

してメタデータとしてモデル化され、RDFに基づいて処理される。RDFを用いることで、セマンティック Web の枠組みに基づいた注釈データの再利用が意図されている。Annoteaでは、注釈をドキュメントに埋め込むのではなく、ドキュメントとは切り離してサーバに格納する。複数の利用者による注釈の共有は、サーバを通じて行われる。また、複数のサーバを利用することができるため、利用者が個別にサーバを置くことも容易である。サーバにはRDFデータベースが用いられる。

Annoteaは、ドキュメントに対する注釈の付加方法とRDFデータベースに対する問合せを定義し、注釈を表示するためのユーザインタフェースは規定しない。Annoteaシステムの実装の1つを図-6に示す。この実装では、注釈の表示と編集にAmayaエディタ/ブラウザを、注釈データの格納に汎用のRDFデータベースを用いている。注釈とドキュメントの結合は、クライアントのAmayaブラウザの中で行われる。

Annoteaの注釈は、次の2つの部分から構成される。

- 注釈本体
- 注釈に対するメタデータ

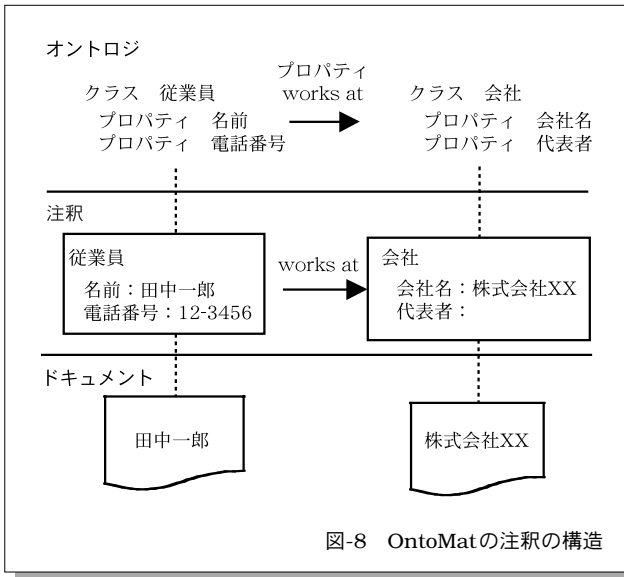
注釈本体は、注釈の内容を記述したXHTMLドキュメントである。注釈に対するメタデータは、付加される注釈に関する情報であり、注釈を付加するドキュメントのURI、ドキュメント中で注釈が付加される部分を示すXPointer、注釈本体のURI、注釈の型、注釈の内容に

☆1 背反の概念：概念Aと概念Bの両方の下位概念になることはないこと。本文中、女性・男性の例参照。

☆2 対称的な関係：1つの関係Rに対してAとBとがある関係Rにあるとすると、BとAとも同じ関係Rにあること。たとえば、AとBとが兄弟であれば、BとAとも兄弟であるといつてよい。

☆3 反対の関係：2つの関係RとSについて、AとBとが関係Rにあるとき、BとAとは関係Sにあること。Rが会員を表し、Sが所属団体を表すとすると、AがBの会員であれば、Bの所属団体にAがあるは真である。

☆4 推移的な関係：1つの関係Rに対して、AとBとが関係Rにあり、BとCとも関係Rにあるとき、AとCも関係Rであること。たとえば、Rが到達可能な関係を表すとき、AからBへ到達可能で、BからCへ到達可能であれば、AからCへ到達可能である。

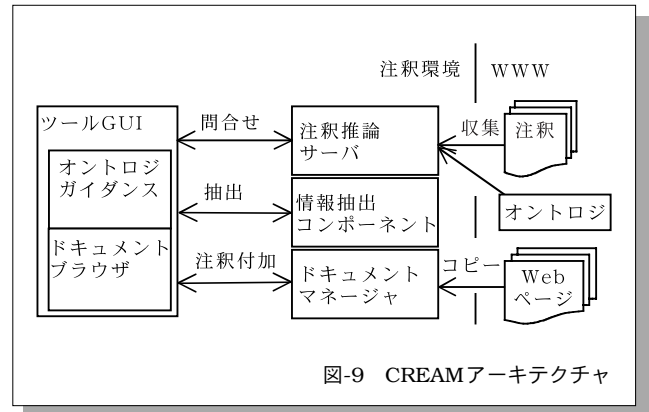


関連するリソース、注釈の作成日時、更新日時、記述者などであり、RDFによってモデル化される。注釈のモデルを図式化して図-7に示す。これは、ドキュメント中の「田中一郎」という部分について、「田中一郎は株式会社XXの従業員である」という説明を「山田太郎」が付加した場合の例である。ここで、注釈の型がとる値はrdf:type Annotationまたはそのサブクラス(Advice, Change, Comment, Example, Explanation, Question, SeeAlsoなど)であり、利用者が定義することも可能である。

OntoMat

OntoMat³¹⁾は、オントロジに基づく注釈をWebページに付与するツールである。OntoMatはOntoAgentプロジェクトの一環として開発が進められ、注釈環境のための枠組みであるCREAM (Creating Relational, Annotation-based Metadata)³²⁾の実装と位置付けられている。注釈付加ツールの多くでは注釈を記述するための一種のテンプレートがあらかじめ用意されているのに対し、CREAMはドメイン・オントロジに基づいてドキュメントに意味情報を付加する点が特徴である。

図-8は、CREAMにおける注釈、オントロジとドキュメント間の関係の例を示している。CREAMではオントロジはDAML+OILのクラスおよびプロパティ定義であり、注釈はそのインスタンスとしてドキュメントに付加される。この例では、ドキュメントの一部である「田中一郎」というテキストに対して、従業員であるという注釈が付加されている。この注釈はオントロジの従業員クラスのインスタンスであり、注釈の中で与えることができる属性(名前、電話番号など)は従業員ク



ラスによって規定される。さらに、オントロジ中で従業員クラスと会社クラスがプロパティ「works at」によって関連付けられているとすると、「works at」によってこの注釈を田中一郎が働く会社のインスタンスと関連付けることができる。

OntoMatは動的にプラグインすることのできるコンポーネントから構成されている。主要なコンポーネントを図-9に示す。オントロジガイダンスは、利用者がドキュメントブラウザで指定した個所の注釈を表示するほか、新規に注釈を付加する際にはオントロジ中のクラスを表示して入力を支援する。注釈推論サーバは、注釈やオントロジを管理してオントロジガイダンスのための情報を提供するほか、それらの間の推論を行う。ドキュメントマネージャは、注釈付加の対象となるドキュメントをWWWからコピーし管理する。ドキュメントを注釈環境の中で管理することにより、同一ドキュメントに対する注釈の重複を避けるとともに、ドキュメントの一部が変更された場合の注釈の半自動的な保守が意図されている。その他、情報抽出技術を利用することによって半自動的な注釈付加を行う、情報抽出コンポーネントが開発されている。

■ RDFのリポジトリおよびクエリープロセッサ

メタデータ付与のためのツールやセマンティックWebのアプリケーションの要素技術として、RDFのデータベースおよびそのクエリープロセッサが急速に発展している。たとえば前述のAnnoteaを利用する場合、サーバにはRDFで記述されたアノテーション用のデータベースが必要であるし、またアノテーションを高度に検索するためにはクエリープロセッサが必要になる。ここではこれらの動向についても少しだけ触れておきたい。

RDFの蓄積は既存のXMLのデータ管理システムも利

用可能である。XMLデータ中のテキストをRDBのデータ構造にマッピングしてデータベース化する手法と、DOMのインタフェースを用いてツリー形式の構造をそのままバイナリ化してデータベース化する手法が存在する。大量のメタデータに複数のエージェントがアクセスする可能性があるセマンティックWebの世界では、高速性の観点からXMLデータをオブジェクトとして管理する後者の技術が今後重要になると考えられる。また、クエリープロセッサはRDFのトリプルの検索を効率的に行うことを目的とし、SQLのような形式でRDFのステートメントを検索できる言語とともに提供されることが求められている。

Sesame³³⁾は、On-To-Knowledgeプロジェクトにおいて、Aidministratoir Nederlandによって開発された、RDF Schemaベースのリポジトリおよびクエリープロセッサである。リポジトリにはRDFデータとRDF Schema情報を格納することができ、ICS-FORTHで開発されたRQL (RDF Schema Query Language) と呼ばれるクエリー言語を用いた検索が可能である。RQLの検索結果はRDFで得られる。HPが提供するJena³⁴⁾というセマンティックWebのツールキットはRDFのモデルを操作するためのJavaのAPIを提供し、RDQL (RDF Data Query Language) によるクエリープロセッサを提供している。この他にも、RDFに特化したデータベースとしてRDFdb³⁵⁾、RDFの構造を解析、格納、管理するPerlのライブラリであるRDFStore (Perl API for RDFStorage)³⁶⁾、MITのセマンティックWebのプロジェクトで検討が進められているクエリープロセッサAlgae³⁷⁾などがある。AlgaeはAnnoteaプロジェクトにおいても利用されている。

■まとめ

セマンティックWebのツールとして、主にメタデータ付与を中心としたツールの概要を述べた。「ツールの概観」の章で、ツールの分類を試みたが、各ツールとも、特徴とする機能を中心に、機能拡大しており、うまく位置付けられないものも多い。また、知的検索エンジンやサイト解析などのより応用よりのツールや推論用のツールなどもあるが、紙幅の関係で割愛する。また応用については、本特集に応用システムに関する章もあるので、参考にさせていただきたい。

セマンティックWebの発展のためには、メタデータを付与・活用する環境が整うことが1つの条件と考えられる。そのために支援ツールがあり、それらのツールは、使われ、改良されて発展してゆく。本稿がその一助になるようなことがあれば望外の幸いである。

参考文献

- 1) 永田, 平: テキスト分類-学習理論の「見本市」-, 情報処理, Vol.42, No.1, pp.32-37 (Jan. 2001).
- 2) 津田, 鶴飼, 三末: Webディレクトリのためのページメタデータの自動付与の試み, 情報学シンポジウム2002, pp.17-24 (2002).
- 3) <http://www.w3.org/RDF/Validator/>
- 4) <http://www-uk.hpl.hp.com/people/jjc/arp/>
- 5) <http://www.research.att.com/sw/tools/graphviz/>
- 6) <http://www.w3.org/Library/src/HTRDF>
- 7) <http://www.injektulo.org/rdf/repata.html>
- 8) <http://www.w3.org/1999/02/26-modules/>
- 9) <http://rdf-filter.sourceforge.net/>
- 10) <http://www.swi.psy.uva.nl/projects/SWI-Prolog/packages/sgml/online.html>
- 11) <http://139.91.183.30:9090/RDF/VRP/>
- 12) ISO/IEC 14977: 1996 (E):
<http://www.cl.cam.ac.uk/~mgk25/iso-14977.pdf>
- 13) 3 A Summary of the ISO EBNF Notation:
<http://www.cl.cam.ac.uk/~mgk25/iso-ebnf.html>
- 14) Brickley, D. and Guha, R. V.: Resource Description Framework Schema (RDF/S) Specification 1.0, W3C Recommendation, Technical Report CR-rdf-schema-20000327, W3C (Mar. 27, 2000).
<http://www.w3.org/TR/rdf-schema>
- 15) DUBLIN CORE: The Dublin Core Metadata Initiative,
<http://dublincore.org/>
- 16) DAML+OIL (Mar. 2001),
<http://www.daml.org/2001/03/daml+oil-index>
- 17) Lassila, O. and Swick, R.: Resource Description Framework (RDF) Model and Syntax Specification, Technical report, W3C (Feb. 1999).
<http://www.w3.org/TR/REC-rdf-syntax>
- 18) GraphViz: <http://www.research.att.com/sw/tools/graphviz/>,
RDFViz: <http://www.ilt.bris.ac.uk/discovery/rdf-dev/rudolf/rdfviz/>,
FRODO: <http://www.dfki.uni-kl.de/frodo/RDFSViz/>
- 19) <http://www.net.intap.or.jp/INTAP/s-web/>
- 20) <http://www.swi.psy.uva.nl/wondertools/>
- 21) Duineveld, A. J., Stoter, R., Weiden, M. R., Kenepa, B. and Benjamins, V. R.: WonderTools? A Comparative Study of Ontological Engineering Tools, In Proc. the 12th International Workshop on Knowledge Acquisition, Modeling and Management (KAW '99) (1999).
- 22) <http://www.ksl.stanford.edu/software/ontolingua/>
- 23) <http://kmi.open.ac.uk/projects/webonto/>
- 24) <http://protege.stanford.edu/index.html>
- 25) <http://www.isi.edu/isd/ontosaurus.html>
- 26) <http://www-sop.inria.fr/acacia/ekaw2000/ode.html>
- 27) <http://www.swi.psy.uva.nl/projects/kads22/>
- 28) <http://www.ontoprise.de/>
- 29) Kahan, J. et al.: Annotea: An Open RDF Infrastructure for Shared Web Annotations, Proc. the 10th International Conference on World Wide Web, ACM Press, pp.623-632 (2001).
- 30) <http://www.w3.org/2001/Annotea/>
- 31) <http://annotation.semanticweb.org/ontomat.html>
- 32) Handschuh, S. et al.: CREAM-Creating Relational Metadata with a Component-based, Ontology-Driven Annotation Framework, Proc. K-CAP '01, pp.76-83 (2001).
- 33) <http://sesame.aidministratoir.nl>
- 34) <http://www.hpl.hp.com/semweb/>
- 35) <http://www.guha.com/rdfdb/>
- 36) <http://rdfstore.sourceforge.net/>
- 37) <http://www.w3.org/2001/Talks/0505-perl-RDF-lib/>

(平成14年5月21日受付)