

# ソーシャルネットを用いたオントロジーの抽出

土屋 吉寛<sup>†1</sup> 和田 雄次<sup>†2</sup>

東京電機大学大学院情報環境学研究所<sup>†</sup>

## 1. はじめに

近年、セマンティックウェブが注目を集めている。昨年の秋にヨーロッパで行われた International Semantic Web Conference (ISWC) が回数を増すごとに参加人数が 1.5 倍ずつ増えていることからその注目度の高さが伺われる。

セマンティックウェブが一般に浸透し、WEB 上に意味ネットワークを確立できれば、様々な分野においてより便利なサービスが誕生すると考えられる。しかしその為の適切なメタデータやオントロジーの生成は完全な自動生成は難しく、またその理解も難解で普及の足かせになっていると考えられる。

しかし一方で近年ブログが一般の人々にも広まり、そのブログの機能の中にセマンティックウェブで使われるメタデータの自動生成機能を標準で搭載しているものが多く、それを用いた研究も大変多く行われている。しかしブログは特に日本では日記用途に使われることも多く、セマンティックウェブを有効に機能させるためのメタデータとしてはブログの情報だけでは不十分だと考えられる。

そこで私はソーシャルネットワークサービス (以下 SNS) を利用することによって一般の人でも簡単にメタデータやオントロジーを半自動生成できるような仕組みを提案する。

## 2. SNS

### 2.1 日本の SNS の現状

SNS とは参加者の実際の世界の間人関係を WEB 上に表現し、新たな交友関係を広げるための WEB でのサービスで、特に日本では mixi<sup>\*1</sup> は既に参加ユーザーが 200 万人を超える日本で最大の SNS となっている。その一方で、最近では次々と新し

いタイプの SNS サービスが増えていて、最近注目されているのは何かの目的に特化したタイプの SNS である。そして同時に wawawa や OpenPNE などのオープンソースの SNS の開発が精力的に行われている。



図 1. mixi

### 2.2 SNS とセマンティックウェブ

SNS とセマンティックウェブには大変親和性の高い部分と相反する部分が同居している。親和性の高い部分としては使用者が自身で更新しやすい仕組みを提供していて、メタデータとして使用されている RDF 形式の一つである FOAF と項目が非常に似通っている。そのため SNS のデータを FOAF 形式で保存し、さらにそれを利用した OWL によるオントロジーを作成することが可能である。

一方で相反する点として、セマンティックウェブはメタデータをオープンにすることを前提にした技術であることに対し、SNS は主に知り合いや共通の趣味を持つもの間で繋がるクローズな空間である。故に情報の制御を厳密にし、さらにどこまでの情報を公開するかを利用者に明示する必要がある。

さらにオントロジーをより効率よく生成するためには各種の SNS を連携させ、そこからオン

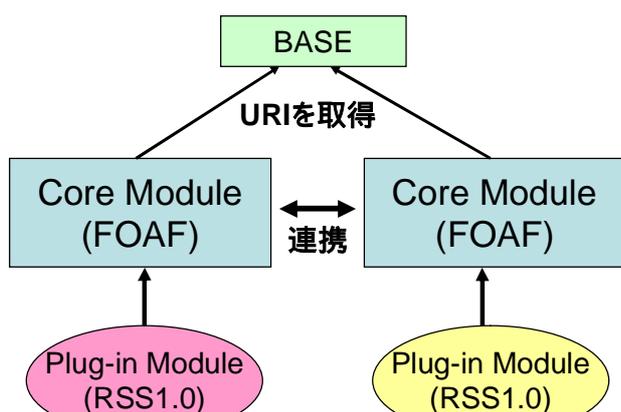
Extraction of ontology with Social Network Service  
<sup>†1</sup> Yoshihito Tsuchiya, Information Environment Engineering, Tokyo Denki University  
<sup>†2</sup> Yuji Wada, Information Environment Engineering, Tokyo Denki University

mixi <http://mixi.jp>  
wawawa <http://wawawa.jp>  
OpenPNE <http://www.openpne.jp/>

トロジーを生成する機能が望まれる。しかし現在の多くの SNS は各 SNS が独立していて連携する機能はない。ユーザーの立場からしても新しい SNS を登録することに様々な情報を再登録するのは大変煩わしい行為である。それを解決するための一つの手段をオープンソースの SNS の wawawa が提供している。wawawa では Base-wawawa と Central-wawawa に分けることによって、SNS 間をつなげ巨大な SNS を作る構想を打ち立てている。しかしこの構造では Central-wawawa に情報が集中しすぎる傾向にある。これは処理が集中する面でも、個人情報を取り扱う面でも好ましくはない。そこで本システムでは中心のサーバーでは個人情報を保持せず、各 SNS の URI 等の最小限情報を保持するにとどめている。

### 3 . システム構成

今回のシステムはより多くの人に使うことによって、よりよいオントロジーが生成できることが期待できる。そこでこのシステムは PHP5.1 を用いて設計されている。PHP5.1 の採用理由として、PHP は近年広く普及している言語で初心者にも学習しやすく、多くのレンタルサーバーにも実装されている。それに加えて PDO の正式採用によりデータベースの違いに影響されにくい設計を行うことが可能であることがあげられる。また PHP ファイルを直接変更することなくデザインの変更がしやすいようにテンプレートエンジンの smarty を採用している。



また今回の SNS では様々な機能に特化した SNS の作成を支援するために SNS としての機能を最小限にしたコア部分に対してプラグインという形で各種のモジュールを追加する形で目的に適した形の SNS を構築できるように設計した。

例えば歯科医用の SNS を構築した場合、技術交換用のフォーラム等の他に、営業時間や曜日を場所と合わせてメタデータとして生成できる機能を持たせることによって、ユーザーが自分の都合に合わせて最適な歯医者を見つけるための検索オントロジーを持った歯科医 SNS を製作することができる。

コアモジュール部分には前章でも述べたように FOAF をベースにしたものを使用し、その他の部分のメタデータに関しては FOAF と同様に RDF 形式であり、ブログ等で広く普及している RSS1.0 形式を採用している。RSS1.0 を採用した理由としては、まずブログ等の先行ツールによって広く世の中に広まっているため既存のツールとの連携が取れやすい。また RDF ベースであるため、名前空間を通してオントロジーに的確に定義を伝えられる。そしてオントロジー言語として採用している OWL が動作することがあげられる。ちなみにオントロジー部分に関して、現在はまだ試作段階だが、オントロジーの使用用途の自由度を上げるために OWL-DL として記述していく方向で進めている。

### 4 . 今後の方針

今回の提案では、ソーシャルネットを用いてオントロジーの半自動抽出を行うための仕組みの提案をしたが、実際にどのようなオントロジーが生成されるかの確認を行うためにも、実際に公開して検証してみる必要がある。その為には検証に適した SNS の考察とそれに必要なプラグインモジュールの開発が必要である。また今回は入力項目からの半自動生成だが、同時に研究している検索エンジンを利用したオントロジーの自動抽出方法に関してもモジュール化を行い、このシステムにプラグインとして取り入れることによって、利用者にとってより便利なオントロジーを構築できる仕組みを提供して、今後のセマンティックウェブを利用した便利なサービスに利用できる形にしたい。

#### 参考文献

- 1) 友部 博教、松尾 豊、武田 英明、安田 雪、橋田 浩一、石塚 満：Semantic Web のための人のネットワーク抽出と利用、情報処理学会、2005
- 2) Robert Tolksdorf, Lyndon J. B. Nixon, Elena Paslaru Bontas, Duc Minh Nguyen, Franziska Liebsch, Enabling real world Semantic Web applications through a coordination middleware, ISWC2005