

Weblog の現在と展望

—セマンティック Web および ソーシャルネットワーキングの基盤として—

国立情報学研究所
武田 英明
takeda@nii.ac.jp

総合研究大学院大学
大向 一輝
i2k@grad.nii.ac.jp

近年の Web におけるさまざまな活動の中で、特に注目されているのが「Weblog (ウェブログ)」である。1998 年頃から登場した Weblog サイトは、すでに全世界で 1 千万に近い数に達しているといわれており、現在も爆発的に増加している。いまや各プロバイダやポータルサイトでは顧客サービスの一環として Weblog のホスティングサービスを提供することが当たり前となった。これに伴う関連サービスも数多い。Weblog は既存のマスメディアやジャーナリズムにも大きな影響を与えており、Weblog 上での議論が世論に反映するような事例も出始めている。このように、Weblog は社会システムとして定着しつつあると思われる。

その一方で、Weblog サイトは一見してそれと判別できるような特殊な形態をしているわけではない。図-1 に示すように、従来と同様の HTML ファイルがハイパーリンクによって接続されたものである。このため、Weblog は新たな名称をつけることによって作為的に起こされた一過性の流行現象でしかないと懐疑的にみる向きも多い。しかしながら、Weblog の普及に際して、コンテンツの記述システムや情報収集の手法など、HTML ファイルを公開するまでのプロセスを支援する技術が飛躍的に進歩している。そして、これらの技術が Web のアーキテクチャを変えつつある。

本稿では Weblog をめぐる技術に着目し、これらの解説を中心として、今後の Web の発展の可能性について述べる。まず Weblog の定義から最近の傾向といった Weblog に関するさまざまな事情について概観する。その上で



図-1 Weblog サイトの例

Weblog にとって重要である Weblog ツールについて説明する。さらに Weblog ツールの技術的基盤であるメタデータやその利用について説明する。その後セマンティック Web と Weblog の関係について述べ、最後にまとめと今後の展望について述べる。

Weblog とは何か

Weblog とは何であるかを一言で説明することは難しい。そこで、本章では発展の経緯や最近の傾向など、さ

さまざまな視点から Weblog を概説することにする。

■ Weblog の定義

Weblog は Web+log からきた造語である。「blog (ブログ・ブログ)」と呼ばれることも多い。現在では、「blogging」のような動詞的表現や、「blogger」といった人を表す言葉も派生している。

Weblog という言葉が定着し、広範なユーザを集めるようになったのは、専用のツールが開発されて簡単にサイトを公開できるようになった 1998 年ごろからである。アメリカで Weblog がまず注目されはじめたのは、新しいジャーナリズムとしての側面であったといわれている。Weblog はマスメディアの制約にとらわれずジャーナリストが自ら意見表明を行うことのできる場であった。その後、9.11 の同時多発テロ後に一般市民が Weblog を通じて意見の交換、議論を行うようになり、そのことが Weblog の認知度を急速に上げた。今日、Weblog はこのような草の根ジャーナリズムのようなコンテンツから、いわゆる日記にいたるまで、さまざまなコンテンツを提供するための基盤となっている。

Weblog の大まかな歴史については以上の通りであるが、いざ Weblog そのものを定義するとすると難しい。詳細な定義については諸説あり、絶えず論争が繰り返されている。各論において文化的、内容的、および操作的定義が乱立し、一般的に適用できるものは少ない。

語源に遡ると、Weblog とは書き手である個人が Web を巡回し、その中で見つけたコンテンツを紹介したサイトであるということができる。最も広義の Weblog とは「日々更新される短いコンテンツの集積」であるとされている。

ただし、この定義に基づく、1998 年以前に作られたサイトも多数該当する。数ある論争の中には「世界初の Weblog サイトは何か」というものがあり、究極の答として Tim Berners-Lee による世界初の Web サイトである「What's New」が挙げられることもある。

比較的多くの人々に受け入れられている定義の例として、Rebecca Blood による著書「Weblog Handbook」¹⁾での説明を紹介する。Blood は Weblog の形態をブログ、ノート、フィルタに分類したうえで、

ブログ：個人的な内容のマイクロコンテンツ（体系化されていない短いコンテンツ）が頻繁に更新されるサイト

ノート：ブログよりも長く、より推敲がなされたコンテンツが更新されるサイト

フィルタ：特定の分野のニュースを取り上げ、コメントを付加する形態のサイト

と定義している。

また、Paquet は Weblog を Personal Knowledge Publishing の一形態であると捉えたうえで、これらの特徴として以下の 5 つを挙げている²⁾。

(1) 個人を編集主体とする (Personal editorship)

基本的には 1 人の個人が編集主体となってサイトを管理し、文書を記述する。したがって記述される情報には一定の個性が反映される。

(2) ハイパーリンクによる接続構造 (Hyperlinked post structure)

1 つの文書 (エントリと呼ぶ) は短い記事からなり、参照する他のサイトの文書へのリンクを含む。また、Weblog サイトに含まれる各エントリは、Permalink と呼ばれる永続的な URI を持つ。

(3) 頻繁な更新と時系列表示 (Frequent updates, displayed in reverse chronological order)

エントリは継続的かつ頻繁に記述され、公開される。新しいエントリは Weblog のメインページの最上部に掲げられ、あとは公開された順番に時系列で並べられる。

(4) コンテンツへの自由なアクセス (Free, public access to the content)

課金などの制限なしに、誰でも自由に Web を通じてコンテンツにアクセスできる。

(5) アーカイブ形式 (Archival)

古くなった記事も削除されることなく蓄積され、Permalink によっていつでもアクセスできる。

これらは定義と呼ぶよりも、むしろ現状の Weblog サイト、とくに後述する Weblog ツールを用いて構築されたサイトの特性を示したものであるといえる。

現在の Weblog の普及において Weblog ツールはなくてはならない存在である。これらのツールが書き手の時間的・心理的コストを下げ、コミュニケーションの形態を変え、結果として先に述べたような数の Weblog サイトを生み出したといつてよい。

■ Weblog をめぐる最近の動向

すでに、Weblog サイトの数が増加するに従い、多くの関連サービスが登場している。対象を Weblog サイトに限定した検索エンジンや、引用関係を分析して最新のトピックを提示するもの、Google や Amazon.com の Web サービスと連携するものなどが挙げられる。

さまざまな分野で大きなインパクトを与えつつある Weblog であるが、その技術によってこれまでは想定されていなかったような事態も起こりつつある。既存の検索エンジンの精度の低下は大きな問題の 1 つである。多

くのWeblogツールは静的なHTMLページを多数生成する。しかも、これらのページは互いに密なリンク構造を有しており、代表的な検索エンジンが持つスコアリングアルゴリズムで高く評価される傾向がある。その結果、検索結果の上位にWeblogサイトが並び、それらのサイトに有益な情報が存在しないといった事態が起こり得る。このように、Weblogによって書くことのコストが劇的に低下したために、膨大なHTMLファイルがWeb上に流通し、すでに「情報過多」と呼ばれている現状をさらに悪化させるのではないかという懸念も見られる。

近年では、Weblogを対象とした学術的研究も始まっている。WeblogはWebの延長線上にあるものの、よりいっそう人間のコミュニケーションに密着したものであるため、計算機上のデータとして分析・利用する計算機科学的アプローチだけでなく、Weblogの内容面に着目した社会学的あるいは心理学的なアプローチもよく採られている。前者の研究としては、Weblogによって構成される空間(Blogspace)が持つ性質をWebグラフの分析手法などを用いて分析した研究³⁾や、Weblogにおけるトピックスの伝播を抽出する研究⁴⁾、Blogに適したランキングアルゴリズムの研究⁵⁾などがある。後者には、Weblog作者に対するインタビューに基づくWeblogの分析^{6), 7)}や、内容分析(content analysis)によるWeblogの分析⁸⁾、社会での影響力の分析⁹⁾などがある。なお、WWW2004(<http://www2004.org/>)ではWeblogに関するワークショップが開催されている¹⁰⁾。

■日本におけるWeblogの位置づけ

日本では、Webの黎明期からいわゆる「Web日記」を公開していた個人サイトが多く、Webコンテンツの1つのジャンルとして定着している。Web日記は、その初期には手作業でHTMLページを更新・公開していたが、近年ではHyper NIKKI System(<http://www.h14m.org/>)やtDiary(<http://www.tdiary.org/>)などの日記ツールや、さるさる日記(<http://www.diary.ne.jp/>)などのホスティングサービスを利用することで簡単にコンテンツを公開できるようになっている。個々のツール・サービスの詳細は文献11)などにまとめられている。

あとに述べるように、ツールのレベルにおいてさまざまな差異はあるものの、Web日記とWeblogは内容面でも形式面でも本質的な部分において共通であり、あえて区別する必要はない。ただ、これまでの歴史的経緯から、アメリカではWeblogという概念の中に日記的コンテンツが含まれるのに対して、日本では日記の拡張としてWeblogが位置づけられるという違いがある。

表現内容		
指向性	事実	心情
自己	「備忘録」 自分のために事実を記録するもの(例: 予定などが書かれた手帳)	「(狭義の)日記」 自分のために自分の心情などを表現するもの
関係	「日誌」 読者を意識して事実を記録するもの(例: 公開日誌, 観察日誌)	「公開日記」 読者を意識して自分の心情などを表現するもの(例: 日記文学, 交換日記)

表-1 日記の分類

■人はなぜWeblog/Web日記を書くのか

本稿はWeblogに関する技術的解説であるが、個々の技術以前に、人はなぜWeblogあるいはWeb日記を書き続けるのかという点について触れておく必要がある。日記という私的コンテンツがWebで公開されるということに違和感を覚える人も多いと思われる。この点については、川浦らの一連の心理学的研究(たとえば文献12), 13))が明らかにしている。表-1に示すように、日記は表現内容と指向性によって4つに分類することができる。川浦らはWeb日記も同様に4つの分類に広く分布することを調査によって明らかにしている。

この表によれば、日記は自分だけのために書くのではなく、他者との関係のためにも書かれるものである。逆にいえば、Web上の日記においては、暗黙的あるいは明示的に他者とのコミュニケーションが期待されていると考えられる。WeblogやWeb日記が相互にリンクされ、Weblogコミュニティ・日記コミュニティを形成しているというのも、このような面からみれば自然な成り行きであり、Weblogツールがさまざまなコミュニケーションの仕掛けを用意していることもWeblogにとって付随的なものではなく本質的なものであることが分かる。

なお、Web上の個人の意見の表明の場としては、「2ちゃんねる(<http://www.2ch.net/>)」をはじめとする掲示板も存在している。コミュニケーションの手段として、話題を中心に複数の書き手が集まる掲示板と、書き手が記述したコンテンツをすべて1ヶ所に集めるWeblogでは、その特性の違いによって用途に応じた相互補完の役割を担っていると思われる。

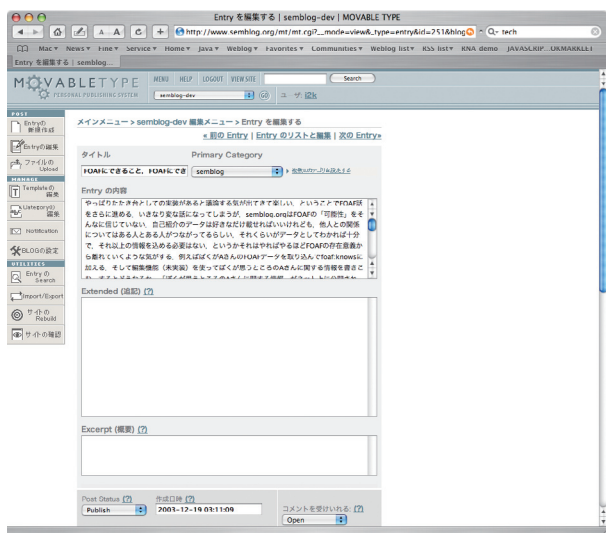


図-2 Movable Type におけるコンテンツ入力画面

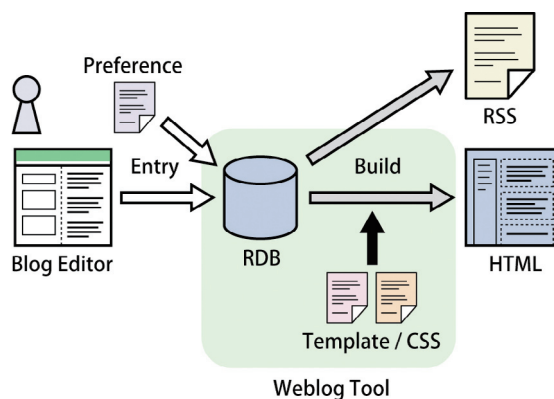


図-3 Weblog ツールのアーキテクチャ

Weblog ツールの概要

Weblog ツールはこれまでに述べた Weblog サイトを容易に構築することのできるソフトウェアである。Weblog サイトが登場し始めた 1998 年前後から、CGI 等を用いてブラウザからコンテンツを記述・公開するためのツールはいくつか存在した。こういったツールは徐々に高機能化し、2000 年から 2001 年にかけて BLOGGER (<http://www.blogger.com/>) や Movable Type (<http://www.movabletype.org/>) といった代表的な Weblog ツールがリリースされるに至り、このようなシステムを利用したサイト管理が一般化した。日本でも、以前から Web 日記のためのツールが多数あり、現在でも多くのユーザが利用している。一方で、海外産の Weblog ツールを導入する際には文字コード等の問題があったが、有志によって日本語化パッチが提供され、初期段階での普及の一助を担った。Weblog ツールの多くは、比較的小規模な、かつ個人を対象としたコンテンツマネジメントシステム (CMS) である。Weblog ツールは以下の 3 つのタイプに大別することができる。個々のツールの詳細は文献 14) などにまとめられている。

・スクリプト型

ユーザが自分でサーバに設置する CGI スクリプトや PHP スクリプトである。カスタマイズの自由度が高い。ツールによっては MySQL 等の DBMS が利用できる環境でなければ動作しないものもある。例：Movable Type, Blossom (<http://www.blossom.com/>), Nucleus (<http://nucleuscms.org/>) など。

・サービス型

アカウントを登録するだけで利用できる Weblog サービス。非常に多くのサービスがあり、それぞれ特徴が異なる。例：ココログ (<http://www.cocolog-nifty.com/>), livedoor Blog (<http://blog.livedoor.com/>), はてなダイアリー (<http://d.hatena.ne.jp/>) など。

・クライアント型

Windows や MacOS X 用のアプリケーションとして提供されるもの。クライアント側でコンテンツ管理を行い、HTML の生成や FTP・WebDAV によるアップロードを自動的に行う。例：Radio UserLand (<http://radio.userland.com/>), iBlog (<http://www.lifli.com/>) など。

Weblog ツールの動作を、代表的なスクリプト型ツールである Movable Type を例に取り説明する。図-1 は Movable Type を用いて作成した Weblog サイトの例である。

ユーザはインターネットブラウザに表示されるフォームにテキストを記述し、保存する (図-2 参照)。

ユーザが記述したコンテンツ (エントリと呼ぶ) は Weblog ツール内のデータベースにて個別に管理されている。サイトの基本設定 (サイト名など) は別に管理されており、ユーザがサイトの再構築を指示すると、Weblog ツールは各データベースより必要な情報を取得し、HTML ページを自動的に生成する (図-3 参照)。その際のサイトのデザインは、ひな形となるテンプレート HTML と CSS スタイルシートによって制御されており、ユーザは容易にカスタマイズすることができる。また、Weblog ツールは 1 つのエントリから単純に 1 つの HTML を生成するだけではなく、最新のエントリの複数個を並べた index ページや月ごとのアーカイブ、カテゴリごとのアーカイブなどを同時に生成する。このように、ワンソース・マルチユースを実現するシステムアーキテク

チャとなっている。この特徴は、後述する RSS 配信にも活かされている。

Weblog ツールには豊富なコミュニケーション支援機能が付加されている。各エントリに対して来訪者がコメントを残すことができる掲示板機能はほぼすべての Weblog ツールに搭載されているほか、来訪者がどのページのリンクを通じてアクセスしたかという Referrer 情報を抽出し、リンクとして表示する機能も一部のツールで用意されている。

また、Weblog ツールに特有のコミュニケーション手段として「TrackBack (トラックバック)」の存在を挙げることができる。TrackBack は Weblog サイト間の逆リンクを生成する機構である。ユーザ A がユーザ B のサイトのコンテンツを引用してエントリを記述した際に、その旨を XML 形式の Ping メッセージとしてサイト B に通知すると、サイト B の Weblog ツールはそのメッセージを受けて B のエントリから A のエントリへのリンクを自動生成する。TrackBack によって、普通に張ることのできる順リンクとは逆方向の、相手から自分の側に来訪者を引き込むようなリンクを生成することが可能になる。この機能を利用して、引用関係を伝えることで議論の流れを時系列的に追うことが可能になる。

これまでの Web は順リンクのみで構成されているために、ページ間のリンクを移動するだけでは特定のサイトに到達できないという蝶ネクタイ構造が形成されるといわれている¹⁵⁾。すべてのページへ到達可能にするには、多くのページおよびリンク情報を格納したアーカイブや検索エンジンが必要となる。Weblog では TrackBack の導入によって、双方向リンクが容易に実現できる。このため複数のサイト間を行き来しながら渡り歩くようなブラウジングや、自分に関連する情報を集めるための小規模なクローリングといった新しい利用法が可能になると思われる。

他にも、Weblog ツールには XML を用いたメッセージの送受信で簡易的な Web サービスを実現する XML-RPC や、標準的な文字コードに UTF-8 を採用するなど、Web における先進的な規格や技術が盛り込まれている。これらは商用サイトにも波及し始めており、今後の Web はこういった新しい技術を下敷きに発展していくものと思われる。

メタデータとアグリゲーション

いくつかの規格は、Weblog ツールに実装されることによって普及したといわれている。Weblog は Web コンテンツの一種であるが、比較的閉じた世界を形成している。すなわち、Weblog サイト間では相互の引用や参照が多く密なコミュニケーションが行われるが、それ以外の Web サイトへは単にハイパーリンクが張られるだけである。このため、Weblog の世界では過去の Web コンテンツとの互換性（過去にさかのぼっての規格のサポートなど）にこだわることなく、新しい規格を取り入れることが容易である。代表的な例としては RSS^{☆1}がある。RSS はサイトの概要を XML 形式で記述するためのメタデータフォーマットである¹⁶⁾。プレーンな RSS は表現力に乏しいが、モジュールという形式を用いて他のメタデータフォーマットの語彙を利用することで、タイトル、更新時間などの追加情報が記述できる。

RSS は現在 0.91, 1.0, 2.0 という 3 つのバージョンが流通しており、それぞれを支持する勢力の間で議論が続いている。また、Weblog に適したサイト概要フォーマットとして新たに Atom (<http://www.atomenabled.org/>) という規格が提案されている。日本では RSS 1.0 (<http://web.resource.org/rss/1.0/spec>) が比較的よく利用されている。以下では RSS 1.0 を取り上げて説明する。RSS 1.0 は RDF (Resource Description Framework)¹⁷⁾ に基づいており、それを XML 形式で記述したものとして表現される。RSS 1.0 での記述例を図-4 に示す。RSS 1.0 は RDF の宣言、channel、item の 3 つのブロックからなる。RDF の宣言部分では、この XML で使用する語彙の体系であるモジュールの名前空間の URI を指定する。この例では RDF 自体の語彙、RSS 1.0 で規定された語彙とともに、書誌情報のメタデータを定義している Dublin Core Metadata Initiative (<http://dublincore.org/>) の語彙を利用する。channel にはサイト全体の概要が記述され、item にはそのサイトに含まれる個々のコンテンツあるいはエントリに関する情報が記述される。channel および item に共通する要素としては RSS 1.0 が持っている title や link (もとのコンテンツの URI、item の場合はエントリごとに生成されたページの URI)、description (サイトの概要。通常は HTML タグを除いた本文の先頭の数十バイト) と、Dublin Core によって定められた dc:date を用いた更新時間情報や、dc:subject を用いたユーザ指定のカテゴリ情報などがある。

☆1 RSS が何の略称であるかもバージョンによって異なる。RSS 0.91 では Rich Site Summary, RSS 1.0 では RDF Site Summary, RSS 2.0 では Really Simple Syndication であるとされる。



図-4 RSS 1.0 の記述例

また、channel すなわちサイト全体のメタデータとしては、エントリがどれだけ含まれているのかを示す items を記述する必要がある。

RSS を利用することにより、ソフトウェアを用いて複数のサイトから情報を収集し、一覧することが可能になった。これをアグリゲーションと呼ぶ。現在ではいくつかの RSS アグリゲーションソフトウェア（アグリゲータ）が提供されている。ここでは筆者らのプロジェクトが提供している Windows 用 RSS アグリゲータ「glucose」(http://www.semlog.org/glucose/) を例にとって説明する。最初に、アグリゲータのユーザは、閲覧したいサイトが配信している RSS の URI を登録する。アグリゲータは HTTP 通信によってファイルを取得する。アグリゲータによって取得された RSS は展開され、3 ペインのインタフェースによって表示される（図-5 参照）。左ペインは RSS を配信するサイトのリスト（チャンネル）である。右上のペインには各コンテンツのタイトル、更新日時、サイト名等のリストが表示されており、各項目によってソートすることが可能である。右下のペインには選択されたコンテンツの内容が表示される。これらのコンテンツは一定時間ごとに自動的に再取得される。ブロードバンド環境の普及にともない、アグリゲータによる情報の閲覧はプッシュ型モデルに近づいているといえる。

日本では、Web 日記の普及と同時に、特定のサイトを登録し、更新状況をチェックするための「アンテナ」アプリケーションが利用されてきた。アンテナは各サイトに対して HTTP 通信を用いてファイルの更新時間を取得

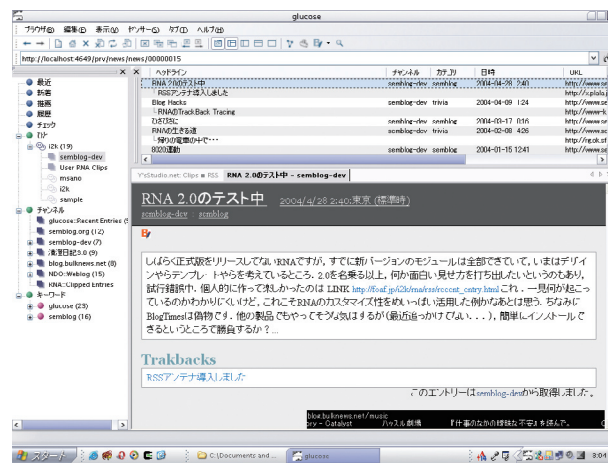


図-5 RSS アグリゲータ「glucose」

し、リストを並べ替える、動的なリンク集であるといえる。アンテナと比較すると、アグリゲータでは規格化された RSS をデータ交換フォーマットとして用いることにより、コンテンツのタイトルや作成者、概要といったよりリッチなコンテンツを取得・表示することができる。

RSS に関連するサービスとして、RSS 全文検索エンジンが存在する。海外では Feedster (http://www.feedster.com/)、日本では Bulkfeeds (http://www.bulkfeeds.net/) や FeedBack (http://naoya.dyndns.org/feedback/) がその代表例である。RSS 全文検索エンジンは従来のものと利用法は変わらないが、検索の対象が RSS のみであり、また検索結果も RSS で提供される。このサービスを利用して、RSS アグリゲータにクエリー（の URI）を登録しておく、キーワードに合致する最新ニュースやエントリを常にチェックするなど、その使用法が可能になる。

また、RSS をアーカイブするサービスの中には、引用回数の多いエントリを推薦する blogdex (http://blogdex.net/) や、ISBN コードを用いて本の批評を抜き出して表示する Technorati (http://www.technorati.com/) などがある。現状では RSS を配信しているサイト数が少ないため、コンテンツに偏りがあることは否めないが、今後さらに Weblog サイトの数が増加していけば信頼性が高まり、より多様なサービスが提供されるものと思われる。

Weblog ツールや RSS の普及と前後して、FOAF (http://www.foaf-project.org/) ¹⁸⁾ というメタデータフォーマットも利用されるようになってきた。FOAF は Friend Of A Friend の略で、友人関係を表現するためのメタデータである。RSS と同様に、FOAF も 2 つの部分から構成されており、最初の部分には本人の名前やメールアドレスなどが記載される。後半の部分には友人の名前、IDなどを knows タグで記述する。FOAF ファイルは各人で分散

的に管理される。これらの FOAF データを横断的に検索し、友人関係ネットワークを可視化する FOAFnaut (<http://jibbering.com/foaf/>) や、情報を集約し、ポータルサイトとして機能する PLINK (<http://beta.plink.org/>) などのサービスが提供されている。現状ではそれ以上の利用法が提案されていないが、今後は後述のソーシャルネットワークとの関連から、このフォーマットが一般に普及する可能性は高い。

セマンティック Web との関係

Web における情報洪水を克服するために、コンテンツに機械可読な意味的タグ（セマンティックタグと称す）を埋め込み、ソフトウェアエージェントによって適切な情報の検索や有効活用を目指す、いわゆるセマンティック Web 技術への期待が高まっている¹⁹⁾。セマンティック Web を実現する要素技術としては先に述べた RDF やオントロジー記述言語が提案されており、基盤は整いつつある。しかしながら、これらの言語が HTML と同様に普及するかは未知数である。コンテンツにセマンティックタグを付加するためには、その文書のコンテキストを把握し、それに応じて適切な語彙を選択しなければならないために簡単ではない。また、タグを埋め込んだことによるメリットを即座に享受することができないといった問題もある。さらには、既存のコンテンツの作者に新たなセマンティックタグの付加を要請することは現実的に不可能であり、人工知能の諸技術や自然言語処理技術による自動付加についても、それらが有効に機能するかどうかはいまのところ不明である。

一方で、RSS 1.0 の基礎となる RDF は、セマンティック Web の基本的な要素技術である。RSS を自動生成・配信する Weblog ツールは個人用のセマンティック Web のためのプラットフォームになり得ると考えられる。筆者らの Semblog (Semantic Weblog <http://www.semblog.org/>) プロジェクト²⁰⁾ では、Weblog ツールおよび補助ツールにより RSS 1.0 を基盤としたメタデータを生成し、セマンティック Web の諸技術が適用可能なプラットフォームの構築を進めている。Weblog ツールを用いることで、一般ユーザにも容易にセマンティックタグを付加できるようになることが期待される。また Weblog のコンテンツが過去の記事の引用を多く含んでいることを利用し、既存のコンテンツへのタグ付けの問題にも取り組んでいる。

Semblog プロジェクトでは、セマンティック Web のビジョンを実現するために必要となる情報流通の活性化を

目指している。現在のセマンティック Web の技術開発では、コンテンツに対する静的な意味記述の厳密化が重要なトピックスとなっているが、ここではコンテンツの動的な側面、すなわちそのコンテンツが誰によって書かれ、また解釈されるのか、もしくはどのような経路を通して流通していくのかといった問題を対象としている。

このプロジェクトでは、Weblog の特性、たとえば TrackBack を用いてリンクの向きを定義することや、RSS アグリゲータの高度化によってコンテンツの流通を推し進めている。その際には、前章で述べた RSS と FOAF を組み合わせる利用することが重要になる。RSS は任意のサイトの概要を記述するものであるが、FOAF と結びつけることにより、その RSS に記載された情報が誰のものであるのかが明確になる。FOAF には友人関係を記述することができるために、「自分に近い友人が記述したコンテンツほど高いスコアを付加する」といった新しい情報検索手法が実現可能になる²¹⁾。

多くの Weblog ツールでは、TrackBack 機能および RSS 配信は個別に実装されており、両者の関連性はあまりない。そこで、Semblog プロジェクトでは RSS に他のコンテンツへのリンク情報や TrackBack 情報を記述するためのメタデータモジュールを提案しており、一部のシステムに採用されている。このような拡張により、あるエントリーに対して TrackBack された他のエントリーの一覧を容易に取得・表示することが可能になる。

RSS を利用することのメリットは、上で述べたような情報流通の基盤がすでに整っていることである。Web における技術開発において、多くの人に好んで利用されるものを目指すことは必須であると思われる。この点で、RSS は大いに可能性があると思筆者らは考えている。

逆に、RSS を基盤とすることの問題は、現状の Weblog ツールによって配信される RSS は情報の種類と量が少ないことである。これをどのように増やしていくのが大きな課題である。

Weblog の今後 - Community Web への展望

ここまで、Weblog やこれを実現するための Weblog ツール、またツール独自の機能について述べてきた。今後の Web は、これらの技術を基礎としてさらに発展していくものと思われる。それでは、どのような方向に進んでいくのだろうか。

1つは、FOAF の普及を足がかりとして、人間関係をいかに表現するかといった技術の発展が見込まれる。欧米

では、近年、「ソーシャルネットワーキング」という概念のもとに Friendster (<http://www.friendster.com/>) や LinkedIn (<https://www.linkedin.com/>) といった個人のプロフィールの収集と、それらに基づく人材のマッチングを行うサイトが急増している。日本でも、Google の社員による Orkut.com (図-6 <http://www.orkut.com/>) が受け入れられたのを皮切りに、GREE (<http://www.gree.jp/>), mixi (<http://mixi.jp/>) といった類似サービスが登場しつつある。これらは、現状では単なる集中型の人間関係登録システムとなっているが、すでに Google はこれらのパーソナルネットワークを新たな検索技術に応用することを表明している。これらのサービス上で得られる情報には限界があるが、個人が持つ Weblog とのリンク付けが可能になれば、各個人についてさらなる情報を得ることができる。Weblog の効用は、個人が一定の場所において意見の表明を続けていくことで来訪者からの信頼 (trust) を得やすいという部分にあるといえる。上記のソーシャルネットワーキング技術と組み合わせることにより、「何を」検索するといった既存の検索手法から、「誰を」検索するという Know-Who 検索を Web 上のオープンな環境で実現することが可能になる。現状のソーシャルネットワーキングサービスは、それぞれが独自のデータフォーマットを利用しており汎用性に乏しいが、今後 FOAF のようなメタデータフォーマットがインフラストラクチャの役割を担い、サービスの統合や新しいサービスの提案につながると期待される。

個人を単位として情報を組織化することにより、これまでは単純に HTML ファイルがリンクで結ばれていた Web の状況から、より情報の粒度が大きくなり、検索性が高まると思われる。また、オーサリング支援技術の向上により、個人が明示的に「書いた」コンテンツだけではなく、カテゴリ分けなどの「編集」プロセスや、どのサイトをチェックしているかといった「興味」を Web 上に表明していくことにより、それらを用いた新たな情報推薦の手法が提案される可能性がある。

このように、Web 上に個人の存在を表明できるようなアーキテクチャを、筆者らは「Community Web」と呼んでいる。Community Web 上での人々のコミュニケーションやコラボレーションの結果が新たなコンテンツとして表現され、他の人によって共有・編集される。筆者らは、このような活動の上に積み上がる「信頼」関係が、セマンティック Web が目指す「信頼の Web」の1つの実現方法であると考えている。



図-6 ソーシャルネットワーキングサービス「Orkut.com」

参考文献

- 1) Blood, R.: The Weblog Handbook, Perseus Publishing, Cambridge, MA (2002).
- 2) Aimeur, E., Brassard, G. and Paquet, S.: Using Personal Knowledge Publishing to Facilitate Sharing Across Communities, in Gurstain, M. ed., Workshop on (Virtual) Community Informatics, Held in Conjunction with the Twelfth International World Wide Web Conference (WWW2003) (2003).
- 3) Kumar, R., Novak, J., Raghavan, P. and Tomkins, A.: On the Bursty Evolution of Blogspace, in Hencsey, G., White, B., Chen, Y.-F. R., Kov'acs, L. and Lawrence, S. eds., Proceedings of the Twelfth International Conference on World Wide Web (WWW2003), pp.568-576. ACM Press (2003).
- 4) Lab, H. I. D.: Blog Epidemic Analyzer: <http://wwwhidl.hpl.hp.com/blogstuff/>
- 5) Adar, E., Zhang, L., Adamic, L. A. and Lukose, R. M.: Implicit Structure and the Dynamics of Blogspace, in Glance, N., Adar, E., Hurst, M. and Adamic, L. eds., WWW 2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics (2004).
- 6) Nardi, B. A., Schiano, D. J., Gumbrecht, M. and Swartz, L.: "I'm Blogging This" A Closer Look at Why People Blog, Communications of ACM (2004) (To appear).
- 7) Schiano, D. J., Nardi, B. A., Gumbrecht, M. and Swartz, L.: Blogging by the Rest of Us, in Proceedings of the CHI2004 (2004).
- 8) Herring, S. C., Scheidt, L. A. and Sabrina Bonus, E. W.: Bridging the Gap: A Genre Analysis of Weblogs., in Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04), Big Island, Hawaii (2004).
- 9) Gill, K.: How Can We Measure the Influence of the Blogosphere?, in Glance, N., Adar, E., Hurst, M. and Adamic, L. eds., WWW 2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics (2004).
- 10) WWW 2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics, <http://www.blogpulse.com/www2004-workshop.html> (2004).
- 11) 日記を認めるシステムたち, <http://kohgushi.fastwave.gr.jp/aboutdiary/>
- 12) Kawaura, Y., Kawakami, Y. and Yamashita, K.: Keeping a Diary in Cyberspace, Japanese Psychological Research, Vol.40, No.4, pp.234-245 (1998).
- 13) 川浦康至, 山下清美, 川上善郎: 人はなぜウェブ日記を書き続けるのか, 社会心理学研究, Vol.14, pp.133-143 (1999).
- 14) Weblog ツールリスト, <http://artifact-jp.com/weblog/>
- 15) Herring, A., Kumar, R., Maghoul, F., Raghavan, P. and Stata, R.: Graph Structure in the Web, in Proceedings of the 9th International World Wide Web Conference (WWW2000), pp.247-256 (2000).
- 16) Hammersley, B.: Content Syndication with RSS, O'Reilly & Associates (2003).
- 17) Manola, F. and Miller, E.: RDF Primer, W3C Recommendation 10 February 2004: <http://www.w3.org/TR/rdf-primer/>
- 18) Brickley, D. and Miller, L.: FOAF Vocabulary Specification, Namespace Document 1 May 2004: <http://xmlns.com/foaf/0.1/>
- 19) Berners-Lee, T., Hendler, J. and Lassila, O.: The Semantic Web, Scientific American (2001).
- 20) Ohmukai, I., Takeda, H. and Numa, K.: Personal Knowledge Publishing Suite with Weblog, in Glance, N., Adar, E., Hurst, M. and Adamic, L. eds., WWW 2004 Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics (2004).
- 21) Ohmukai, I., Numa, K. and Takeda, H.: Egocentric Search Method for Authoring Support in Semantic Weblog, Workshop on Knowledge Markup and Semantic Annotation (Semannot2003), Held in Conjunction with the Second International Conference on Knowledge Capture (K-CAP2003) (2003).

(平成 16 年 4 月 6 日受付)