

# Community Web プラットフォーム

## Building Community Web Platform

大向 一輝\*1  
Ikki OHMUKAI

松尾 豊\*2  
Yutaka MATSUO

松村 真宏\*3  
Naohiro MATSUMURA

武田 英明\*1  
Hideaki TAKEDA

\*1 国立情報学研究所  
National Institute of Informatics

\*2 産業技術総合研究所  
National Institute of Advanced Industrial Science and Technology

\*3 大阪大学大学院経済学研究科  
Graduate School of Economics, Osaka University

In this paper we propose Web-based communication environment called "Community Web Platform". Our platform provides an easy way to exchange personal knowledge among people with lightweight metadata such like RSS and FOAF. We investigate the nature of "personal trustness" on the environment since it is one and only measure for evaluating subjective information and knowledge. We also discuss how to develop and maintain Community Web applications from our experience.

## 1. はじめに

1989年に誕生して以来、Webの規模は拡大の一途を辿っている。初期のWebは研究者間の情報共有ツールとして用いられていたが、ユーザ数が増加するにつれ、企業の広告メディアとしての利用や、Eコマースなど、新たな用途が考案されてきた。情報共有についても、アカデミック分野の情報のみならず、個人の興味などに関する情報が続々と公開され、検索エンジンやポータルサイトと組み合わせることで多くの情報を容易に得られる環境が構築されてきた。この結果、Webは巨大な辞書と呼べるほどの情報量を誇るまでになった。

一方、Webにアクセスするユーザが増加するにつれ、Webがコミュニケーションの基盤として機能するようになってきている[1]。人々は、現実生活における友人あるいは知人とのコミュニケーションを補完するために、Webを会話や議論、あるいはコラボレーションの場として利用している。

近年のWeb技術の中で発展が著しいのは、このコミュニケーション基盤に関するものである。従来から存在したオンライン掲示板(BBS)に加え、WikiやWeblog(Blogあるいはブログとも称す)、ソーシャルネットワークキングサービス(SNS)など、新たなコミュニケーションのチャンネルが増加しつつある[2]。また、コミュニケーションの結果が辞書としてのWebのクオリティを高めるといったモデルも提案されている。これと同時に、RSSやFOAFといったメタデータの流通によって、ソフトウェアによる情報の集約や検索、編集が容易になった。

このような状況の中で、情報を発信し、受信する個人の存在がクローズアップされている。コミュニケーション基盤としてのWebでは、流通する情報の信頼性を担保するために、その情報が誰のものであるかを明確にする必要がある。そこで、個人のアイデンティティの表出、記名性が求められる。

本研究では、個人の存在がWeb上に表出され、そういった個人間でのコミュニケーションがなされている場を「Community Web プラットフォーム」と定義し、プラットフォームの構築や、プラットフォーム上での情報流通の支援のあり方について考察する。また、Web上に表出された個人の信頼性の評価についても検討する。

連絡先: 大向 一輝, 国立情報学研究所, 〒101-8430 東京都千代田区一ツ橋 2-1-2, Tel: 03-4212-2585, Fax: 03-3556-1916, i2k@nii.ac.jp

具体的な研究手法としては、実世界で運用可能なアプリケーションの構築や、その運用結果の分析を通じて、次世代のWebが進むべき方向性について議論を行う。その過程で、人工知能研究がWebに対して与える影響が明確になることが期待される。

## 2. コミュニケーション基盤としてのWeb

前節で述べたように、Webはコミュニケーションの様式およびそれを支える技術の両面において大きな変化を遂げつつある。ここでは、コミュニケーション基盤としてのWebにおける様相や諸技術について概説する。

### 2.1 コミュニケーション様式の変化

インターネットの黎明期から存在したUsenetや、BBSなどのサービスでは、特定の話題に関する参加者同士の議論がサポートされ、その経過をインターネット上に公開状態にすることで、多くの人々が詳細な情報を得ることを可能にした。これらのサービスは、議論の場に参加者が集まるモデルとなっており、そこで生み出されたコンテンツはすべて場に集約され、場の管理者によって所有されることになる。

一方で、近年急速に普及が進んでいるWeblogは、個々のユーザがそれぞれのWebサイトを持ち、サイト内のコンテンツは個人用コンテンツ管理システム(CMS)によって維持管理されるモデルとなっている[3]。Weblogユーザ間での議論は、自身のサイトにコンテンツを掲示し、相手サイトの該当コンテンツに対してリンクあるいはトラックバックを張ることで行われる。このモデルでは、ユーザが生み出したコンテンツの所有権はすべてそのユーザが持つことが明確になっている。

また、SNSではシステム管理はサービスの提供者が行うが、参加者のそれぞれはシステム内に自身が管理するページを持ち、コンテンツの所有権は参加者にあることが明示化されている[4]。多くのSNSではBBSに類似した議論の仕組みが用意されているが、個々の書き込みは各参加者のページにリンクされるため、コンテンツの記名性が確保されている。

### 2.2 メタデータの普及

CMSを利用したWebコンテンツの作成が一般化するにつれ、コンテンツに付与されるメタデータが目立っている。代表的なメタデータであるRSSは、Webサイト内に含まれるコ

コンテンツのリンクや更新時間、概要を XML あるいは RDF で記述するためのフォーマットである [5]。RSS は、個々のコンテンツをページ単位で生成する Weblog やニュースサイトと親和性が高く、これらの普及に合わせて急速に流通するようになった。

各サイトから配信された RSS は、RSS リーダーあるいは RSS アグリゲータと呼ばれるソフトウェアを用いて自動的に収集し、一括表示を行うことができる。これにより、ユーザは多くの Web サイトを短時間で閲覧することが可能になる。また、RSS のみを対象とする検索エンジンも存在しており、RSS リーダーなどとの組み合わせによって、擬似的なプッシュ配信を実現している。

メタデータの応用としては、個人とその友人関係を RDF によって表現するためのフォーマットである FOAF がある [6]。一部の Weblog では、ユーザのプロフィールを記述するために FOAF が採用されており、具体的な利用法が模索されている。

RDF はセマンティック Web の基礎となるフレームワークである [7]。RSS や FOAF はセマンティック Web 環境を実現するための手段として期待されている。

### 2.3 フォークソノミー

上に述べた Web におけるコミュニケーション様式の変化と、メタデータの普及が融合した事例として、フォークソノミーあるいはソーシャルタギングと呼ばれる技術が生まれ、ユーザの支持を得ている。フォークソノミーとは、多くの人々によって協調的に作られる語彙の体系を意味する [8]。

フォークソノミーを実現しているオンラインブックマークサービスの代表である「del.icio.us」\*1では、ユーザが Web コンテンツのブックマークを作成する際に、タグと呼ばれるキーワードを自由文で入力することができる。ブックマークはタグごとに整理され、表示される。del.icio.us では、タグはブックマークサービスを横断的に検索する際のクエリーとしても機能し、同じテーマで他人が作成したブックマークを閲覧することができる。1 つのコンテンツに対して複数のタグを設定することも可能であり、タグの共起性を利用して形成される単語のネットワークは新たなコンテンツの推薦に利用される。

フォークソノミーはブックマークだけではなく、写真や音楽の共有サービスでも導入されており、さらなる利用法の提案が続くものと思われる。

## 3. Community Web プラットフォーム

Community Web は、実世界における個人の活動を支援するために、他者とのコミュニケーションを可能にする場である。Community Web では、各ユーザの存在とともに、実世界においてユーザ自身を取り巻いている友人関係ないし知人関係が明示化される。本研究では、このような人間関係をパーソナルネットワークと呼ぶ。

Community Web におけるコミュニケーションとは、このパーソナルネットワーク上で情報が移動することであると定義できる。パーソナルネットワーク上の知人から情報を収集することや、知人に向けて新たな情報を発信するといった情報流通に関する活動は、コミュニケーションの一環であると言える。

本テーマでは、Community Web 環境において、個人の情報流通活動を支援し、活性化するための方法論について検討を行う。

先に述べたように、Web 上の個人は、情報の受信者としての側面と、情報の発信者としての側面とを持っている。受信者

の立場から見ると、現在の Web は情報過多状況にあり、目的の情報を得ることが極めて難しくなっている。また、発信者からの立場からは、発信する情報がより多くの人々に届くことが望ましいが、情報を伝えるべき対象に適切に伝わらない事態が生じている。

これまで Web 上の情報流通を支えてきた検索エンジンは、Web ページに含まれるリンクの構造から各ページの信頼度を計算し、順位づけを行っている [9]。しかしながら、検索エンジンのランキング手法に対応して特定のページの順位を上げることを目指す検索エンジン最適化や、多量のリンクが自動的に生成される Weblog サイトの普及により、検索精度の低下傾向が見られる。

これらを回避するためには、パーソナルネットワーク上の個人同士で直接的なコミュニケーションを行うことが望ましい。これを Web 上で実現するにあたっては、以下に示す 3 つの課題が存在すると考えられる。

- 情報流通基盤の構築

現状の Web サイトや Weblog サイトには、パーソナルネットワークの要素が存在しないため、これを付加して情報流通基盤を確立する必要がある。そのためには各サイトの管理者の協力が不可欠であるが、これを実現するには管理者に対するメリットやインセンティブを設計し、提示することが望ましい。

- コミュニケーション対象の選択

パーソナルネットワーク上のユーザの中で、誰から情報を得るか、あるいは誰に対して情報を発信するかを選択することは極めて重要である。とくに、コミュニケーションの対象を直接の知人関係以外から探す場合に、これを支援する必要がある。

- 個人に対する信頼性の評価

Community Web においては、情報の信頼性は内容そのものに起因するのではなく、それを作成し、発信した個人から得られることになる。そこで、個人に対する信頼性を定義し、評価することが必要となる。これを定量的かつ客観的な指標として実現することは極めて困難であるが、各個人が個別に判断する際の参考情報を提示することを目標としたい。

## 4. これまでの研究成果

前節で述べた課題に対して、筆者らはいくつかの研究を行い、一定の成果を挙げた。以下に成果の概要を述べる。

### 4.1 Semblog プラットフォーム

大向らは、Community Web に必須となるパーソナルネットワークと情報流通基盤との統合を、Weblog とメタデータによって実現する「Semblog プラットフォーム」を提案している [10]。ここでは、Weblog による情報流通基盤を、メタデータ層、メタデータ管理層、メタデータ集約層、アプリケーション層の 4 層構造として定義し、それぞれの層に対して拡張を行った。

第 1 のメタデータ層では、コンテンツを表現する RSS に対して、トラックバック情報を記載するなど、Weblog に適した拡張を施したうえで、パーソナルネットワークを表現する FOAF を接続し、コンテンツの所有者に関する情報を明示化した。次に、第 2 のメタデータ管理層では第 1 層の拡張メタデータを効率的に管理するための CMS の提案と、FOAF によるパーソナ

\*1 <http://del.icio.us/>

ルネットワークを容易に拡張するための FOAF TrackBack の実装を行った。メタデータ集約層では、メタデータの編集や再利用を可能にした RSS アグリゲータ「RNA」および「glucose」を提案し、それぞれについて実装を行った。アプリケーション層では、各個人が分散的に運用している RSS アグリゲータを FOAF によって連結し、情報を交換することでコンテンツの検索や推薦を行う「RNA Alliance」の実装を行った。

Semblog プラットフォームを構成するアプリケーションは実際に配布し、現実のコミュニティで運用されている。2004 年度人工知能学会全国大会で行った実証実験では、クローズドな情報支援サービスとは異なる情報流通形態を実現したことが確認された。

#### 4.2 Web からの人間関係の抽出

松尾らは、特定のコミュニティにおけるパーソナルネットワークの構造を、Web 上の情報から自動的に抽出する手法を提案している [11]。この手法では、検索エンジンを用いてコミュニティに所属する任意の 2 名の氏名が Web ページ上でどれだけ共起するかを計測し、関係の強さとして定義するとともに、得られた Web ページの内容を解析して人間関係の種類を判別している。スナップショットを図 1 に示す。

この手法により、直接の知人関係だけでは得られないパーソナルネットワーク全体の構造を知ることが可能になる。また、人間関係の種類や個人間の興味の類似を知ることができるため、Community Web におけるコミュニケーション対象の選択に役立てることが可能であると思われる。

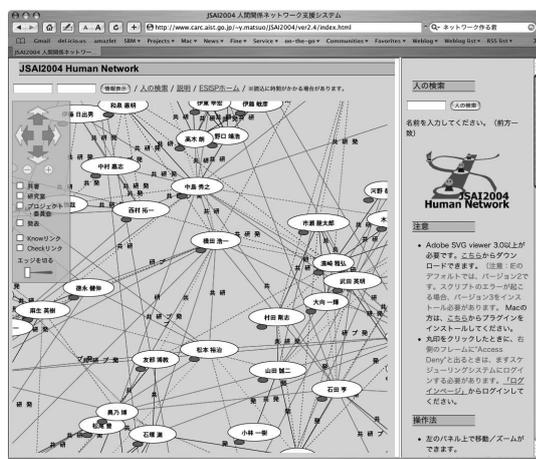


図 1: Personal Network from the Web

#### 4.3 議論における発言者のプロファイリング

松村らは、BBS などのオンラインコミュニティにおいて、議論の履歴を分析し、発言者の発言内容がその後の議論にどのように影響を与えたかによって発言者のプロファイリングを行う手法を提案している [12]。この手法では、発言者間の時系列的関係やキーワードの伝搬関係を利用して、各発言者の特性や議論自体のテーマの抽出を実現している。

この手法は BBS を対象として適用されているが、Community Web 環境においても、各コンテンツの所有者が明確であり、かつそれらのコンテンツには更新情報などのメタデータが付加されていることから、同等の分析が可能であると思われる。

### 5. Community Web プラットフォームの目標

これまで述べてきたように、Community Web プラットフォームは、Web における個人性の表出とパーソナルネットワークによる人間関係の明示化をメタデータによって実現するというアーキテクチャである。そして、このプラットフォームの目的は、情報や知識をそれらが所属するコミュニティにグラウンディングさせることである。価値観が多様化する現代社会では、情報や知識を持つ重要性、あるいは価値基準は、万人に適用できるものではない。それらの価値は、コミュニティ単位か、もしくはさらに微細化された単位である人間関係においてのみ担保されると思われる。

現代における情報あるいは知識の共有は、こういったコミュニティや人間関係をまたいだ形で行われる。2 節で述べたフォークソノミーは、各個人がそれぞれの視点からコンテンツに対してメタデータを付与するものである。フォークソノミーは、そのままでは立場の異なる個人あるいはコミュニティ間が分断されるのみに終わるが、それらのタグが誰によって付加されたものが明示され、共有されることにより、個人間、コミュニティ間の差異が明確になり、相互の知識のマッピングが起こるきっかけになり得る。コミュニティごとに分散管理された知識体系を、多様性を保持したままで再統合することが可能になれば、これまでセマンティック Web の分野で取り組まれてきた大規模なオントロジーの構築につながると期待される。

### 6. プラットフォームの構築に向けて

Community Web プラットフォームがもたらすであろう利点を列挙することは容易であるが、それを実現するにあたっての障壁は数多い。障壁の例としては、個人単位で情報の流通を行う際に生じやすいプライバシーの侵害に対する制御やセキュリティの確保といった社会的な問題が挙げられる [13]。また、認知的な観点からは、フィードバックの問題も重大な障壁として考えることができる。フィードバックが情報共有あるいは情報流通に貢献していることをユーザに感じさせることは極めて難しい。

多くのユーザが情報共有に対して積極的でない理由の 1 つは、ユーザの努力が役に立たないように見えるからであると考えられる [14]。多くの人が Community Web プラットフォームを利用ようになるためには、ユーザの行動をシステムによって常にチェックし、その分析結果をユーザの利用する情報共有サービスに即座に反映させる必要がある。しかしながら、情報や知識の共有は、本質的には長時間を要するものであり、その点が多様性やスケーラビリティにつながるということは無視できない。

そこで、本研究では、実世界における情報・知識共有アプリケーションが満たすべき 2 つの効用 (gratification) を提案する。instant gratification は情報・知識の共有を考慮せずにユーザが得られる効用であり、delayed gratification は情報・知識共有によって得られる効用である。

新たなアプリケーションを普及させるにあたり、instant gratification はユーザをつなぎとめる役割を果たし、その間に情報・知識共有の真の効用である delayed gratification が得られるようになる。あるアプリケーションが普及するかどうかは、それぞれの効用の質よりも、2 種の効用のバランスをどのように設定するかによって決定されると思われる。先に述べたように、情報・知識共有を実現するためには長い時間を要するため、情報・知識共有の結果がアプリケーションの導入直後に反映されるようなモデルは現実的でない。

Web は両方の効用を持っている。ハイパーテキストの編集は、人々が持つ情報や知識を組織化するために適した方法であるために、即座にユーザにとっての利益を生み出す。同様のことをワードプロセッサのような形式で表現することは難しい。そして、Web では編集された文書を公開することに特段のコストを必要としないため、容易に情報を公開することが可能である。情報の発信者は、それを読んだ人々からのフィードバックを得ることにより、delayed gratification を得ることができる。

ここで、重要な問題は、構築すべきシステムにおいてどのように2つの効用を定義し、設計するかということである。本研究では、Web に代表されるシステムの観察結果や、これまでに行ってきた情報・知識共有システム構築の経験をもとに、ユーザを instant gratification の享受者から delayed gratification の享受者に移行させるための translucence strategy を提案する。これは、個々のユーザを近い距離に配置し、delayed gratification の有効性を予感させるという単純な戦略である。これにより、ユーザを delayed gratification を享受できる段階に引き上げる。この戦略を実行するにあたっては、instant gratification を得るためにユーザが取るべきことと、delayed gratification を得るための情報・知識共有行動との差を最小限に留める必要がある。

## 7. まとめと今後の展開

近年普及が進んでいる Weblog や SNS は、個人の存在を Web 上に表出させ、そういった個人間のコミュニケーションを活性化させる働きを持っている。同時に、RSS や FOAF といったメタデータの流通によって、ソフトウェアによる情報の集約や検索、編集が容易になった。このように、Web はコミュニケーションの様式およびそれを支える技術の両面において大きな変化を遂げつつあるが、これらの変化が個人の行動に与える影響については考察がなされておらず、また変化の方向性についての議論も少ない。そこで、本提案では、Web を個人の日常的なコミュニケーションやコミュニティ活動の場として捉え、この機能を有効に支援するための技術的要件および実現可能性について議論を行った。

ここでは、セマンティック Web の分野で議論されてきた情報の信頼性に関する問題や、オンラインコミュニケーションの基盤となる意味体系の整備、情報の伝搬範囲のコントロールなどの課題を、コミュニティの成員が日常的に行うローカルな情報の組織化によって解決することが必要となる。これらが満足されることで、従来のトップダウン型およびボトムアップ型の情報処理パラダイムとは異なる情報流通形態の実現が期待される。

今後の研究の戦略としては、初期は分析研究と提案研究を個別に進める。分析研究では、Web ページや Weblog、BBS 等から抽出可能な人間関係ネットワークについての分析から、個人間のコミュニケーションのモデル化を行う。提案研究では、メタデータ、パーソナルネットワーク、コンテンツマネジメントといった要素技術をもとにシステムを構築し、実証実験等を行う。中期以降はそれぞれの成果を統合した Community Web プラットフォームの提案を中心に、アプリケーションの開発とその有効性の検証を行う。

本提案は、すでに起こりつつある技術の変化、ならびにユーザの行動の変化に適応した Web のアーキテクチャを構築することを目標としている。よって、本提案においては、スケーラビリティを含め、実用可能性を強く意識した人工知能技術の設計に注力する。また、モバイル環境やユビキタス環境といった新たな情報技術との統合についても検討する。

## 参考文献

- [1] B.Shneiderman. *Leonardo's Laptop: Human Needs and the New Computing Technologies*. MIT Press, 2002.
- [2] E.Aimeur, G.Brassard, and S.Paquet. Using Personal Knowledge Publishing to Facilitate Sharing Across Communities. *Workshop on (Virtual) Community Informatics, Held in conjunction with the Twelfth International World Wide Web Conference (WWW2003)*, 2003.
- [3] I.Ohmukai, H.Takeda, K.Numa, M.Hamasaki, and S.Adachi. Metadata-driven Personal Knowledge Publishing. *Proceedings of the Third International Semantic Web Conference (ISWC2004)*, 2004.
- [4] 武田英明, 大向一輝. 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. *情報処理*, Vol. 45, No. 6, pp. 586-593, 2004.
- [5] B.Hammersley. *Content Syndication with RSS*. O'Reilly & Associates, 2003.
- [6] D.Brickley and L.Miller. FOAF Vocabulary Specification. <http://xmlns.com/foaf/0.1/>, 2004.
- [7] T.Berners-Lee, J.Hendler, and O.Lassila. The Semantic Web. *Scientific American*, 2000.
- [8] A.Mathes. Folksonomies - Cooperative Classification and Communication Through Shared Metadata. Technical report, University of Illinois Urbana-Champaign, 2004.
- [9] L.Page, S.Brin, R.Motwani, and T.Winograd. The PageRank Citation Ranking: Bringing Order to the Web. Technical report, Stanford University, 1998.
- [10] I.Ohmukai, H.Takeda, K.Numa, M.Hamasaki, and S.Adachi. Personal Knowledge Publishing Suite with Weblog. *Workshop on Weblogging Ecosystem, Held in conjunction with the Thirteenth International World Wide Web Conference (WWW2004)*, 2004.
- [11] 松尾豊, 友部博教, 橋田浩一, 中島秀之, 石塚満. Web 上の情報からの人間関係ネットワークの抽出. *人工知能学会誌*, Vol. 20, No. 1E, pp. 46-56, 2005.
- [12] 松村真宏, 大澤幸生, 石塚満. 影響の普及モデルに基づくオンラインコミュニティ参加者のプロファイリング. *人工知能学会論文誌*, Vol. 18, No. 4, pp. 165-172, 2003.
- [13] 大向一輝, 武田英明. 人間関係ネットワークに基づく情報フィルタリングを用いた協調的タスクスケジューラ. *電子情報通信学会論文誌*, Vol. J87-D1, No. 11, pp. 1020-1029, 2004.
- [14] L.McDowell, O.Etzioni, S.D.Gribble, A.Halevy, H.Levy, W.Pentney, D.Verma, and S.Vlasheva. Mangrove: Enticing Ordinary People onto the Semantic Web via Instant Gratification. *Proceedings of the Second International Semantic Web Conference (ISWC2003)*, 2003.