

1

Webの進化とエージェント、 セマンティック Web

武田 英明 国立情報学研究所／東京大学 人工物工学研究センター

本稿では Web2.0 に代表される近年の Web の進化とエージェントについて、セマンティック Web を仲介として論じていく。セマンティック Web もエージェント技術も Web2.0 の特徴を実現する先行技術であった。しかし、現在の Web の大規模性や多様性に対応することができなかったため、Web2.0 の中に加わることができなかった。もっとも Web2.0 は Web の進化の現在のスナップショットであり、今後は Web は社会としての役割（社会としての Web）が大きくなるであろう。このときは社会環境を構築するセマンティック Web と社会インフラストラクチャを形成するエージェント技術が重要になると期待される。

Web の今とエージェントのゆくえ

本稿では近年話題になった Web2.0 に代表される Web の進化とセマンティック Web 技術、エージェント技術との関連を概観する。その上で Web の進化の今後を考察する。

Web2.0 はビジネスの世界から出てきた概念であるので、情報の研究者にとって縁遠い概念と思われるかもしれない。しかし、我々の情報環境を良くしていくという点においては、情報の研究者とは違った視点を提供しており、学ぶべきことが多いと考えている。Web というのはそもそも実用のための仕組みであり、その実用性を高めようと努力した結果、Web は日々変化していき、それが Web の進化である。そのためにはどんな技術でもテクニックでも貪欲に取り入れている。一方、セマンティック Web もエージェント技術もネットワーク上のソフトウェアの技術として提案されたものである以上、実用的な価値をもってこそ存在意義がある。

これらに共通する特徴はあるのか、相互に影響を与え得る可能性があるのだろうか。本稿ではそのあたりを探っていきたい。

Web の未来予想と現実

以下は 2001 年 5 月号の Scientific American 誌に載った Tim Berners-Lee 他の記事の冒頭である¹⁾。

電話が鳴ったとき、娯楽システムからビートルズの "We Can Work It Out" が流れていた。Pete が電話に出たとき、彼の電話は音量調整を持っている他のローカルデバイスすべての音量を下げるようメッセージを出して、音を低くした。彼の妹、Lucy が医師の部屋から電話をかけていた。「お母さんは専門家に診察してもらい、物理治療を続けてなくてはならないのよ。2週に 1 回ぐらいね。これから私のエージェントにアポイントメントを取ってもらうわ」。Pete はすぐに車での送り迎えを請け負った。

医師の部屋で、Lucy は彼女のハンドヘルドの Web Browser を通して、エージェントに教示していた。エージェントはすぐに医師のエージェントから母親の処方箋に関する情報を入手していた。そしてこのサービスを提供できる提供者をいくつか調べ、加入している保険でカバーでき、母親の家から半径 20 マイル以内で信頼できる評価サービスにおいて優秀と評価されている提供者を探した。次に個々の提供者のエージェントとコンタクトを取り、提供者の可能なアポイントメ

ントの時間と Pete と Lucy のスケジュールのすり合わせをはじめた...

これがセマンティック Web が実現する未来の Web というわけである。Tim Berners-Lee はいわばセマンティック Web に Web の未来を仮託したわけである。ここには Web 上のエージェント、エージェント間メッセージング、エージェント検索、Web の信頼性、エージェント間でのネゴシエーションなど、さまざまな技術要素が含まれている。

これから 5 年以上経った。セマンティック Web はどうなったであろうか。あるいは Web そのものはどうなったであろうか。

Tim Berners-Lee が director である Web の標準化団体である World Wide Web Consortium (W3C) ではワーキンググループなどを作ってセマンティック Web の推進を行ってきた。ここでは RDF (Resource Description Framework), RDF Schema²⁾, OWL (Web Ontology Language)³⁾ といった言語を策定してきた。また研究コミュニティは上記のような言語の開発やその処理系、それを使ったアプリケーションなどを開発してきた。その結果、Web 上の情報にメタデータとして意味を付加する仕組みが構築されている。

さらに Web 上のサービスを標準化する Web サービスと結合して、Web サービスに意味を付加する仕組みが提案、開発されている。

セマンティック Web の歩みの速さは期待外れのものであったかもしれないが、確実に基盤を整えつつある。

ただし、セマンティック Web の技術は RDF など部分的なものは世の中で使われるようになったが、これぞセマンティック Web というアプリケーションはなかなか生まれなかった。すなわち、セマンティック Web は研究室をなかなか出られなかった

一方、エージェント技術もまたなかなか研究室を出ることができなかった。1990 年代にはモバイルエージェント言語 Telescript⁴⁾ が、エージェントコミュニケーション言語としては KQML⁵⁾ が開発されて、エージェントを利用したアプリケーションが普及するかと思えたが、Web の劇的な普及の前に霞んでしまった。その後も Web と連動して動作できる Java ベースのモバイルエージェントも数々開発されてきたが、大規模な利用に至っていない。

Web2.0

この間にも、Web は次々と変化していった。この最近の変化を Tim O'Reilly は "Web2.0" と名付けた。この名付けは絶妙で多くの人が使う概念となった。過去の Web がバージョン 1 なら、今の Web はバージョン 2 であるというわけである。ソフトウェアのバージョン 1 とバージョン 2 では論理的な進展があるわけではないが、なんらかの性能向上や使いやすさでの進歩がある（と期待されている）。今の Web はそんな状態であるというわけである。

彼の記事⁶⁾ によれば、Web2.0 とはまず Web をプラットフォームとして位置付けることである。これは今となっては当然のことなのであるが、あえてその価値を再認識せよということである。次に利用者のモデルとしては「情報の自己コントロール」であるとする。すなわち情報利用者は情報提供者に一方的に従属するといったモデルではなく、相互に関係し合いかつ自立した利用者ということ想定している。特徴的な要素としては

- (1) パッケージソフトウェアではなくてサービス
- (2) 参加のアーキテクチャ
- (3) 高い拡張性とコスト効率
- (4) 再構成可能なデータソースとデータの変換
- (5) 単一デバイスを超えたソフトウェア
- (6) 集合知の活用

を挙げている。また、代表的なサービスとしては

1. フォークソノミー (例 del.icio.us, Flickr)
2. 豊かなユーザ経験 (例 Gmail, Google Map, AJAX)
3. ユーザの貢献 (例 PageRank, eBay, Amazon)
4. ロングテール (ex. AdSense)
5. 公開ではなくて参加 (ex. Blogs)
6. ラディカルな信頼 (ex. Wikipedia)
7. ラディカルな分散化 (例 BitTorrent)

この個々の特徴やその具体化されたサービスなど Web2.0 の詳細については元記事および本誌特集号⁷⁾ を参照されたい。

それでは Web2.0 の出現はセマンティック Web もエージェント技術も不要であるということであろうか。

確かに Web2.0 は、セマンティック Web もエージェント技術でも看過されていた側面を露にしたという点で大いに評価できる。しかし、それによって、他の技術が不要になるのではなくて、相補的な関係であると考えべきであろう。以下では Web2.0 とセマンティック Web、エージェントの関係を詳しく見ていくことにする。

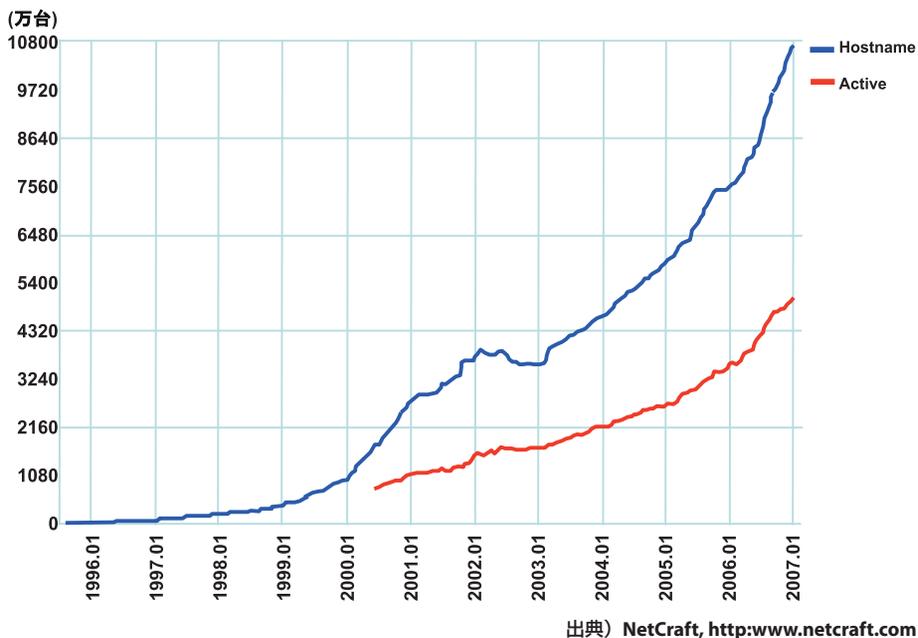


図-1 Webサーバの台数の変化

超大規模情報共有としてのWeb2.0

Webが作られた目的は情報共有の実現であった。Webは自由で安価で簡単な情報共有システムであり、それがゆえにまたたく間に世界中に広がった。

Webが作られて10年あまり経つがこの基本的性質は変わらない。しかし、その環境は大いに変わっている。規模が大幅に変わったのである。Webはいまや億単位の人々が使い、100億近いページがある情報共有システムとなっている。図-1はWebサーバの台数の変化を示したものであるが、この10年の規模の変化を見て取れるであろう。

すなわち、Webは単なる情報共有システムでなく、超大規模情報共有システムとなっている。Webのこの規模は史上初めてのものであり、このスケールがゆえにシステムが質的に変わらざるを得ない。そのボトムアップな試みがWeb2.0である。

先に挙げたWeb2.0の6つの特徴は基本的にこの超大規模性実現のための仕組みであることが分かる。性質2と6からWeb2.0はきわめて多数の人々が能動的に参加することで成立するものであることが分かる。また大規模であるがゆえに、情報の利用も多様であり(性質4)、ユーザも多様なアクセスを自由に行える必要がある(性質5)。そしてその実現の仕組みとしてはサービスであり(性質1)、固定的な仕組みではなく、ダイナミックに実現されるもの(性質3)である。すなわち、大規模性、多様性、インタラクティブ性、可変性を持つ情報共有を

実現しようとしているわけである。

セマンティックWebは情報共有の実現という点でWeb2.0と共通するし、情報交換の仕組みの実現という意味ではエージェント技術と共通している。しかし上記のWeb2.0の性質はこれらにとって新しい要求であり、解決されていない問題である(図-2参照)。それを以下で見ていく。

Web2.0とセマンティックWeb

Web2.0とセマンティックWebは大規模情報共有を実現しようという点においては同じであるが、その注目点は異なる。ある意味、大規模情報共有とは矛盾を内包しているともいえる。情報を共有しようとしているわけであるから、統一された情報交換のフォーマットがなければならない。しかし、大規模性になればなるほど多様な情報共有を認めていかないとけない。

セマンティックWebは元来のWebがそうであったように基本的には情報共有に関する技術である。その方法論としては情報共有の抽象度を上げて知識共有として発展させることで、高度な情報共有を可能とすることを目標としている。知識を共有する仕組みを提供することで高度な情報共有が実現できると考えており、実際、オントロジー言語RDFSやOWLが制定されてきた。

セマンティックWebは情報の標準化に主に注視して、それを発展させようとしている。ある意味、きわめてオーソドックスに情報共有の問題に取り組んでいる。

一方、Web2.0は大規模性に注目している。大規模性

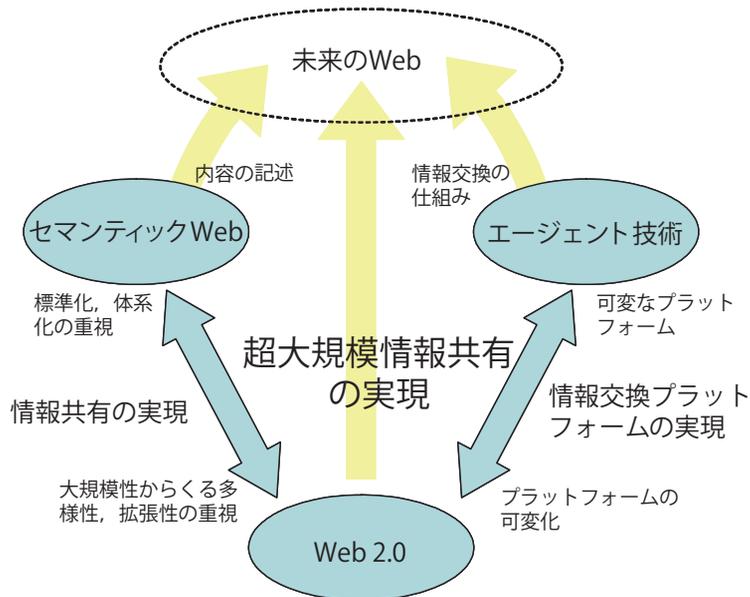


図-2 Web2.0とセマンティックWeb, エージェント技術の関係

はもちろん大量であるということがまず問題になる。が、単に量が多いというだけではとどまらない、というのがポイントである。大規模さが生み出すいわば“創発”的現象(集合知)を積極的に利用すべきであるし、それを可能とするようなフレキシブルな仕組み(高い拡張性)を提供すべきと考えているわけである。

すなわち、両者は補完的である。セマンティックWebは共有の基盤構築に集中しているが、大規模性をもたらす問題は看過している。一方、Web2.0は大規模性やそこからくる多様性を重視するが、共有基盤には注目しない。しかし、安定した大規模知識共有は両方の要素が必要である。

セマンティックWebの方から多様性、分散性を取り込むような研究も現れている。

セマンティックWebにおいてもオントロジーは分散的に開発・利用されると考えられている。そのためにオントロジーのマッピング、アライメント、統合といったものが盛んに研究されている。たとえば、Yahoo!ディレクトリのような巨大オントロジーを自動的にマッピングする仕組みの提案などもある。

またオントロジーを協調して構築するというものも盛んに研究されている。ことに近年、wikiを拡張して、Semantic Wikiとして、オントロジーやオントロジーによるタギングを共同で行う環境として使う動きが盛んである。

さらにオントロジーをWebデータから自動構築するという試みもある。Peter Mikaはフォークソノミーからオントロジーを作るという試みを行っている⁸⁾。ここで

はdel.icio.usのタグに上位下位関係を自動発見することを行っている。

Web2.0とエージェント技術

多くのエージェント技術ではエージェントが実行あるいは通信する環境を普及させて、そのプラットフォームの中でエージェントが活動するということを想定している。しかし、このような環境を普及させることは大変なことであり、それがエージェント技術普及のネックになっている。基本的にhttpとhtmlしかないWebに加えれば、はるかに高度な機能を提供し得るエージェント技術であっても、普及しなければ真価が発揮できない。Webの利用者もこのWebの“低機能”性に満足してはいてわけではない。しかし、すでに普及しているWebを捨てるのではなく、その上に新しい機能を構築しようとしている。それがソフトウェアアーキテクチャとしてのWeb2.0である。

Web2.0の特徴のうち、単一デバイスを超えたソフトウェア(性質5)はコンピュータから携帯機器に至るさまざまな機器を連携させるソフトウェアを指している。これはまさにエージェントによって達成したかったことである。しかし、実現の方法論が異なる。エージェント技術のように新たなプラットフォームを用意して個々のアプリケーションに対応させるのではなく、最低限Webであるということだけを基盤に、その上に用途に応じてさまざまな方法で連携を実現している。

その実現方法の1つが高い拡張性とコスト効率(性質

コラム1: ブログとエージェント

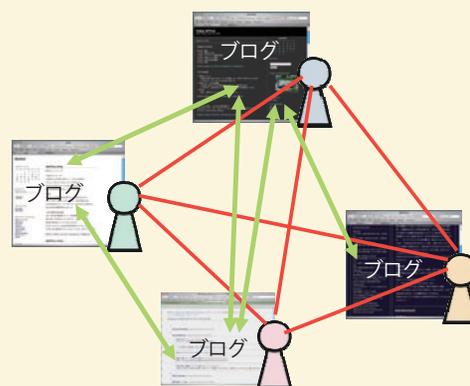
ブログ (Weblog) は Web2.0 でも取り上げられているユーザ参加型 Web の好例である。ブログは単に自己の情報の公開ツールではなく、コミュニケーションツールとしても使われている⁹⁾。ブログは分散システムであるので、コミュニケーションツールとしては、コミュニケーション相手の同定、情報の伝達方法、同期などが問題になる。ブログではこれらをトラックバック、RSS の公開とアグリゲーション、update ping など、独自の技術で解決している。

このような機能を持つブログは視点を変えるとエージェントとしてみることもできる。Semblog プロジェクト

ではこの点を拡張してブログ同士のネットワークを構築して、その間での情報交換を実現する仕組みを構築している¹⁰⁾。エージェント通信と見た場合、XML-RPC を通信プロトコルとして、RSS と RDF を通信内容の言語、エージェント間のリンクの記述には FOAF を用いている。

ブログは Web 世界の個人の存在という意味でもエージェントといえる。

ブログは個人の情報公開を一括して行っているため、いわば Web 世界の個人の存在を示すものとみなされつつある。この意味でもブログはエージェントであり、通信を強化することで Web 上のエージェントとして役割をより積極的に果たすことが可能になる。



エージェントとしてのブログのネットワーク

3) というフレーズに代表され、具体的には ajax や API の公開といったものである。一般に独自の方法で連携といったシステムを作った場合、ユーザを囲い込む方向になりやすい。しかし、ユーザの囲い込みは Web の公開性の原則にも合わないし、またビジネス的に Web の巨大さに対応できないし、さらには技術的にも Web の日々の進歩についていくことができない。

代わりにソフトウェアそのものの公開やソフトウェア API を公開し、多くの人にソースやデータを自由に利用してもらうことで、利用者の拡大とソフトウェアの発展を同時に達成しようとしている。結果として多様な利用方法が開拓されたり、新しい技術が追加されるという柔軟性の高い仕組みとなっている。

現在の ajax や API 公開でできることは固定的な連携など、エージェント技術から見れば基本的な機能である。柔軟な連携や自律性など、エージェントが提供し得る機能は多い。今後は、エージェント研究で出てきた技術はいったん解体され Web 上に再構築されると思われる。

セマンティック Web とエージェント技術の合流点とは?

冒頭に述べたようにセマンティック Web の理想像ではエージェントがいるような Web 世界を想定している。では Web 上の“意味”を持ったエージェントとはどんな実現形態をとるであろうか。

1 つは人間の代理人として人間間のコミュニケーションを代行するエージェントのネットワークが考えられる。後章で述べるように社会活動は Web 上へシフトしていくので、Web 上での代理人としてのエージェントは重要な役割を持つ。この場合、人間とエージェントが意味理解を共有することが重要になる。人間の意味理解を計算機が完全に行うことはまだ困難な問題であるので、エージェントに求められる能力は最低限の意味の共有と良いヒューマン・エージェント・インタフェースである。Web の超大規模性を考え合わせれば、人々が用いる意味を形式的に定義することは困難であるが、先に述べたフォークソノミーのような非形式的な意味を処理できるようなエージェントが必要となってくるだろう。

一方、Web 上でのさまざまな活動を自律的に行うエージェントも考えられる。このようなエージェントには形式的な意味定義と意味理解の能力が求められる。複雑化する Web 上の活動を自動化するには発見や合成といった処理が必要とされるが、そのためにはあいまい性のない意味定義が必要である。

前者の研究紹介としてコラム 1 (ブログとエージェント)、後者の研究紹介としてコラム 2 (Web Services とエージェント)を参照されたい。

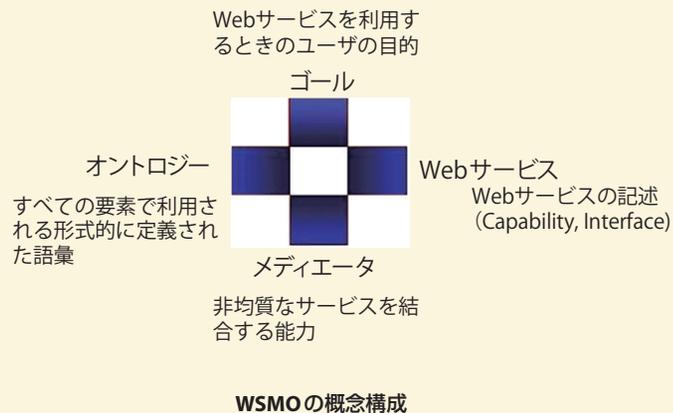
これらの研究を見て分かるように、意味を処理するエージェントというのは Web におけるエージェントの存在として有望かつ期待されている。

コラム 2: Web Services と エージェント

サービス志向アーキテクチャ (SOA) は近年注目されているソフトウェア構築手法であり, SOAP/WSDL/UDDI によって構成される Web Services は Web 上での分散サービスの連携のプラットフォームとして注目を浴びた。しかし, この枠組みでは固定的なサービス連携しか実現できない。というのはサービスの意味を記述する仕組みがないからである。セマンティック Web サービスは Web サービスに意味をつけようという試みである。この中で欧州を中心とするグループでは WSMO^{☆2} というものを提案している。WSMO (Web Service Modeling Ontology) ではオントロジーを用いて, Web サービスの名前, 目

的, 実行の文脈, 入出力などを記述する。これによって意味に基づいた Web サービスの発見, 選択, 合成といったものが可能になる。すなわち, 知的な Web サービスが実現できるというわけである。

これはサービスがエージェント化しているとみることができる。たとえば WSMO に基づいたメディエータ (仲介器) が実装されているが, これはまさにエージェントネットワークのメディエータと同じ役割である。



社会としての Web

Web2.0 は現在の Web のスナップショットにすぎない。さて, 未来の Web は今後どんな方向へ発展していくのであろうか。

Web のこれまでの発展から外挿すれば「Web は社会化」することが予想される。Web2.0 の言い方になぞらえれば, 「社会としての Web」である。「Web の社会化」とは, 我々が日常生活している基盤である社会が Web 上に乗ってしまうということを意味している。この 10 年で我々の生活は確実に Web 上へ移されてきた。その傾向は今後むしろ加速し, 社会の大部分が Web 上へ移動してしまうというわけである。

これが意味することは, 社会に存在するあらゆる要素が Web 上に存在するようになるということである。社会上の存在とは, 人, もの, 人やものの関係, 社会的活動(生活, 教育, ビジネス), コミュニティ, 組織, ルール, モラル, 法律, 犯罪, 政治, 等々である。社会という視点から見ると, 現在の Web は社会化の程度はまだまだ低いことが分かる。しかし, 今後はさまざまな社会要素が次々と「Web 化」されていくであろう。

とはいえ, 現在の社会がそのまま Web にコピーされるわけではない。実空間の持つ制約と Web 空間の持つ

制約は異なっているので, 異なった構造の社会になるであろう。

実空間に比較して Web 空間はどんな特徴を持つであろうか。少し比較してみよう。

まず実空間における「もの」に対応するデータの特徴としては以下のものを挙げる事ができる。

- 複製可能である
- 再利用可能である (何度でも使える)
- 永続性 (ずっと存在する)

複製可能と再利用可能はデジタルデータの一般的特徴であるが, 永続性は少し趣が異なる。確かにデジタルデータは劣化しないという面では永続的であるが, 実際に永続的に存在し続けるかは別問題である。Web 上の情報は消去が簡単で永続性がないように見えるが, 多くの場合, 一度 Web 上に現れた情報はどこかに保存され, 永続的に残り得る^{☆1}。

我々の活動に対応するプロセスの方の特徴としては以下のようなものを挙げる事ができる。

- 時間に依存しない
- 空間に依存しない

☆1 ハードディスクなどデジタルデータの保存技術の発展がデータ量の増大についていけるかぎりという限定であるが, 今のところ, 楽観的に考えてよいだろう。

☆2 <http://www.wsmo.org/>

- 多重化可能である
- 並列化可能である
- 量に依存しない

最初の2つ(時間や空間に非依存)は始めからのWebの特徴である。残りは近年の計算機技術の急速な発達とその普及の恩恵によるものである。計算機資源が潤沢であるということが保証されれば、多重化や並列化が容易であり、さらには実質的に情報の量を気にしなくてもよくなりつつある。

当然、このような特徴を持つWeb上の社会は今までの社会とは異なる構造になるであろう。

たとえば、Web上の取り扱いで混沌とした状況にある著作権問題も、そもそも実空間のための仕組みを複製可能・再利用可能という異なる性質を持つWebに適用しようとすることによって生じている問題である。

こういうことがたくさん生じつつ、社会としてのWebが形成されていくであろう。

ただし、すべてがWebに移行するわけではない。我々は実空間に身体を持ち、依然として実空間で生活し続けるわけである。したがって、一部はWeb空間で一部は実空間でという生活であり、それが今よりずっとWebの比重が増えるということである。

社会としてのWebとエージェント

社会としてのWebにおいてエージェント技術は重要な役割を果たすことが期待されている。そもそも先にあげたWeb空間の特徴は実空間に身体を持つ我々に合わないのである。我々は情報を永続的に記憶したり、大量の情報を扱うこともできないし、ましてや多重や並列に処理することもできない。このギャップは増える一方である。ここに必要なのはまさに代理人としてのエージェントである。

また、社会の要素自身がWeb化するということは我々がインタラクション可能なWeb上の存在物にならないといけない。この身体の代替という意味でもエージェントが必要である。このエージェントは必要に応じてWeb空間のどこへでも行ってインタラクションを行うであろう。

社会としてのWebという視点から見れば、セマンティックWebは社会環境の構築である。すなわち、我々の生活環境や文化をWeb上に構築する仕組みを提供す

る。一方、エージェント技術は社会インフラストラクチャの構築である。その社会環境の中で人々が実際に活動できるような仕組みを提供する。社会としてのWebは両方の取り組みが必要とされる。

未来のWebとエージェントに向けて

本稿では、Webの進化の方向とエージェント技術のかかわり合いについて論じた。関連技術を十分にカバーしているとはいえないし、仮定に基づく議論も多く、粗雑な議論であることは否めない。しかし、Webの急速な発展の中、将来を見越した議論が研究者に必要であると考えて、あえてこのようなかたちで述べさせてもらった。これをもとに多少とも議論が起これば幸いである。

参考文献

- 1) Berners-Lee, T., Handler, J. and Lassila, O. : The Semantic Web, Scientific American (May 2001).
- 2) Brickley, D. and Guha, R. V. (eds.) : RDF Vocabulary Description Language 1.0 : RDF Schema, W3C Recommendation (10 Feb. 2004).
- 3) McGuinness, D. L. and van Harmelen, F. (eds.) : OWL Web Ontology Language Overview, W3C Recommendation (10 Feb. 2004).
- 4) White, J. E. : Telescript Technology : The Foundation for the Electronic Marketplace, White Paper, General Magic, Inc. (1995).
- 5) Finin, T., Labrou, Y. and Mayfield, J. : KQML as an Agent Communication Language, Proceedings of the 3rd International Conference on Information and Knowledge Management (CIKM'94)(1995).
- 6) O'Reilly, T. : What Is Web2.0 Design Patterns and Business Models for the Next Generation of Software (2005). <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html> (Last visited on 4 Oct. 2006).
- 7) 特集 : Web2.0の現状と展望, 情報処理, Vol.47, No.11 (Nov. 2006).
- 8) Mika, P. : Ontologies are Us : A Unified Model of Social Networks and Semantics, Proceedings of the 4th International Semantic Web Conference (ISWC 2005), LNCS 3729, Springer-Verlag (2005).
- 9) 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 (June 2004).
- 10) Ohmukai, I., Takeda, H., Hamasaki, M., Numa, K. and Adachi, S. : Metadata-Driven Personal Knowledge Publishing, in S. A. McIlraith, D. Plexousakis and F. van Harmelen eds., The Semantic Web - ISWC 2004 : Third International Semantic Web Conference, Hiroshima, Japan, November 7-11, 2004., Vol. 3298 of Lecture Notes in Computer Science (LNCS), pp.591-604 (2004).

(平成 19 年 2 月 2 日受付)

武田 英明 (正会員)
takeda@nii.ac.jp

1991年東京大学大学院博士課程修了。2003年より国立情報学研究所教授。同学術コンテンツサービス研究開発センター長(併任)。東京大学人工物工学研究センター寄附部門客員教授(兼務)。大阪大学特任教授。