

Webの持続的発展を支える要因

— セマンティック Web 以降の展開 —

The Factors Supporting Continuous Development of Web

— The Elaboration after the Emergence of Semantic Web —

福井 誠*

Makoto Fukui

従来からのインターネットサービス企業では広告モデルと利用履歴データベースによる競争優位戦略が採用されている。しかし、その後提案された Web2.0 型サービスが有効な収益モデルを示すことができなかったこともあって、インターネットへの全般的な信頼性は低下傾向にある。Webの継続的な発展には、自立運営が可能な新たな事業モデルと、インターネットへの信頼を維持するための手段が必要であることを示した。

キーワード：セマンティック Web、resource based view、インターネットマーケティング、Web2.0

I. はじめに

本稿は、事業モデルの視点から Web の歴史的発展を整理し、その上で将来の姿を展望することを目的とする。改めて述べるまでもなく、Web はインターネットの商業利用とともに発展し、インターネット自体の普及にも大きな役割をはたした。インターネットが社会的な信頼を維持しつつ社会基盤としての役割を果たすために、今後も Web に課せられた役割は大きいと考えられる。ただし、後述するように Web は大きな転換点を迎えており、従来の延長線上での発展が約束されているわけではない。Web が持続的に発展するためには、現在の Web の事業モデルにかわる自律可能な事業モデルが必要であり、一方でインターネットへの信頼を維持するためにはプライバシーを担保する手段が必要である。

このことを明らかにするため、まず商業化以降のインターネットの20年間の歴史の中で、Webサービスの事業モデルが次第に姿を変えてきたことに注目して、過去の変遷を整理することから検討をはじめた。過去の Web の変遷を整理するにあたっては、Tim O'Reilly が2005年に提唱した Web2.0¹⁾ を分岐点として採用することとし、その前後で時代を2区分することとした。

Web2.0の問題点については後で検討することとしたいが、Web2.0が少なくとも事業モデルと

いう視点からは重要なターニングポイントであったと考えられる。また、Webの歴史を時間軸に沿って二分した場合には、時間軸上の中間点はインターネットバブルが崩壊した2000年～2001年頃となる。2009年はBerners-LeeがWebの構想につながる最初の文書“Information Management: a Proposal”²⁾を発表した1989年から20年目となるので、その中間点は1999年頃となり、インターネットバブルの時期と重なる。一方、O'ReillyがWeb2.0を提唱したのは2005年であるが、O'Reillyは単にドットコム・バブル³⁾崩壊後に出現したさまざまなサービスを観察し、整理し、共通する傾向を見いだして、それをWeb2.0と名付けて後に整理したに過ぎない。このように考えると、O'ReillyがWeb2.0という名前を与えた変化は、すでに2000年頃から姿をあらわしはじめていたと考えることができる。以上のことからWeb2.0はWebの歴史を前半の10年と後半の10年とに区分する、まさに分岐点であるといつてよいであろう。

II. 初期のWebの事業モデル

1. インターネットサービス事業の重要成功要因(KSF)とは

後述するように、インターネットの商業利用が解禁されたのは1991年のことであった。これと時期を同じくしてWorld Wide Webというインターネット上の文書管理法が提案される。このWebに事業的な将来性を見出した事業者が、web本来の目的、用途を越えた多様な利用方法をうみだし、様々なアプローチから事業化に取り組む。しかし、多くは事業構築に失敗し、残った事業者もほどなくして収益を確保できる事業モデルがわずかしかないことに気づく。このわずかな可能性にたどりついた企業が最終的に生き残って、現在もサービスを提供する巨大サービスの運営主体となった。事業化に成功したサービスに共通する特徴は、

- 広告によって収益を確保しながら、無償のサービスを提供すること。
- 日常的に利用されることによって、顧客を囲い込み、利用者の行動をすべてログという形式で記録すること。
- これらのログを蓄積し、結果的に利用者の活動に関する巨大なデータベースを構築すること。
- さらに、このデータにより、競争優位を獲得・維持し、収益の拡大を得ること。

と要約できる。

この戦略は、個別企業の経営戦略としては妥当であったと評価できる。実際に、これ以降、Webサービスにはこれにかわる有力な戦略は出現していない。この点については本論の最後で再度検討したい。Webサービス事業の収益は、広告収入、利用料収入、その他の収入に分類可能である。ここで利用料収入とはサービスの受益者が負担する料金であり、その他の収入にはサイトを利用して収益を上げる別の事業者からの利用料あるいはコミッションが含まれる。たとえばポータルサービスであるYahoo! JAPANを運営するヤフー株式会社の事業構造を例にとると、2009年3月期の決算において、広告収入は52%、ビジネスサービスが21%、パーソナルサービスが27%と

報告されている⁴⁾。ここでビジネスサービスはその他の収入、パーソナルサービスは利用料収入にあたる。このように、直接の販売収入や手数料収入による収益構造を持つ EC（電子商取引）サイト以外では、収入の柱となるのは広告収入である。

しかし、広告収入は事業を維持するための手段でしかなく、事業の継続性と競争優勢の持続的発展にとってより重要な要因は、利用者の残したログ情報とそれを集約するデータベースであり、さらにデータベースに蓄積された情報を活用することで得られるサービスの付加価値である。この点においては Yahoo! のようなポータルサイトだけでなく、eBay あるいはわが国では楽天などといった EC サイトも同様であると考えられる。

Web サービスの開発は初期投資額も少なく、開発リードタイムも従来の基幹情報システムとくらべると非常に短い。このため、模倣が容易である。参入障壁が低いことから、少しでも成功しそうなモデルが公開されると、後発の類似サービスが一挙に増加する。ここで過当競争が引き起こされることにより、先行者利益が失われる可能性が常につきまとう。

にもかかわらず、多くのケースで先行するサービスが成功を収め、後発のサービスで成功したケースはきわめて少ない。同種のサービスで生き残れるのは、トップシェアを獲得した 1～2 社だけである。現在ではわが国のインターネットユーザが Google や Yahoo! 以外の検索エンジンを利用することは少なくなったが、わずか 10 年ほど前には数多くの検索エンジンが覇を競っており各々が好みに応じてサービスを使い分けていた。この中で、検索エンジンの Google、ポータルの Yahoo!、EC サイトの Amazon、楽天などだけが現在も主力のサービスとして存在している。たとえば Yahoo! JAPAN の検索サービスでのシェアが特異的に多いとされるわが国では、2008 年のシェアは Yahoo! JAPAN の 44% に対し Google が 41% であり⁵⁾、上位 2 社の合計は 85% に達している。また、社団法人日本通信販売協会（JADMA）が 2008 年 11 月に発表した「インターネット通販利用者実態調査⁶⁾」の結果によると、ネット通販の利用サイトでは楽天市場、Amazon.co.jp、Yahoo! ショッピングの利用率の合計が 94.9%、うち楽天市場だけで 75.7% を占めているとされる。

かつて企業の情報システム構築においては、先行する企業が保有するソースコードや通信ネットワーク構築や端末の整備のために費やした情報システム投資によって、競争優位性を持続的に獲得できた⁷⁾。Web サービスでも同様の現象が生じ、このことによって競争優位の持続的獲得が可能であるとも考えることもできる。しかし、インターネットサービスでは、保有する情報システム資産は競争力の源泉とはならない。インターネットからはじまった「軽量のプログラミングモデル⁸⁾」は、ソースコードに資産価値をもたらさず、かつて顧客囲い込みのための重要な投資と位置づけられていた端末群や通信回線は、インターネットでは利用者側で用意するものとなった。すなわち、Web サービスで観察される寡占化現象は、競争優位を獲得したサービスが競合するサービスの内容とくらべて内容的に優れているからではない。さらにいえば、後発のサービスは先行するサービスの分析を経てサービス構築が可能であるため、後発サービスはより優位な立場にあ

るはずである。

これらの要素が先行者の競争優位を保証しないのに対して、追従する後発のサービスが容易に構築できない要素は利用者の行動履歴データの蓄積にあると考えられる。すなわち、先行するサービス提供者は、後発のサービスが開始されるまでに獲得した利用者の行動履歴データを活用することで後発のサービスを寄せ付けないサービスを構築し、さらにシェアを拡大していくという循環が生まれている。すなわち、蓄積されたデータこそが競争優位獲得の主要成功要因になっていると評価できよう。そして、このような循環は、後述するようにインターネットという仕組みそのものに組み込まれた性質によって生じるものと考えられる。

2. 商業利用によるインターネットの変容

この点についてももう少し詳しく検討することにしよう。インターネットはもともと互酬性を基本的な行動原理として成立し、相互接続のための回線にしても、中継するネットワーク機器についても、相互利益を原則に参加者の役割分担と無償提供とを前提にデザインされた接続形態である。

インターネットの起源が ARPANET であったことはよく知られている。1969 年にアメリカ国防総省の国防高等研究計画局（Advanced Research Project Agency：略称 ARPA）は、冷戦下の緊迫した社会情勢の下でコンピュータネットワークの研究を開始する。その実証実験としておこなわれた大学間ネットワークは、その後のインターネットの特徴である分散型の特性をすでに最初の段階から備えていた。このネットワークが中央集中型ではなく分散型なのは、核攻撃が現実的な脅威と受け取られていた当時の冷戦構造の中で、システム全体の停止を回避するためだったといわれている。

開発のきっかけとなった軍事利用での要請を受けて、部分的な障害に対して全体が停止しない耐障害性が重要な要件となり、そのために冗長構成をとる通信方式が採用された。このような経緯から、現在もインターネットの特徴として冗長構成をことさらに強調する傾向は続いている。しかし、商業化以降のインターネットが、爆発的に普及した原因は冗長構成そのものではない。冗長構成によって耐障害性が高まったことでネットワークの可用性（Availability）⁹⁾が高まり、安定した通信が可能となった。冗長構成が採用されたことで、ISP 事業者は信頼性の高い設備や回線を調達することが不要となり、過剰な投資をする必要がなくなった。このことで利用者が満足できるレベルの通信環境を比較的安価に提供できるようになった。提供側にとっては経営の安定を生み、利用者側にとっては低価格で安心できる環境をえられることにつながった。すなわち商業化以降の利用者は、互酬性の原理に従って応分の役割を担いインターネットの一部分の構成するメンバーとしてではなく、文字通りの利用者として有償で環境を利用するものとなった。このような、かなり迂遠なプロセスがインターネットを商業的に成功させ、普及に導いた構造的

要因であるといえる。以下では、この点について歴史的に振り返ってみたい。

インターネットに商業利用という概念がはじめて持ち込まれたのは、1991年にアメリカのISP事業者が集まってCIX Association（商用インターネット相互接続協会）という業界団体を設立したことが契機となっている。CIXはアメリカのインターネットサービスプロバイダ間を接続する商用の相互接続ポイントの管理団体で、今のネットサービス事業とはかなり異なっているが、ともあれ、インターネットの世界に互酬性ではない原理が初めて持ち込まれた¹⁰⁾。

そして、このCIXという組織がインターネットの互酬性の原理を大きく転換することになる。ただし、実際には商業化以前の互酬性に基づくネットワークと無償の情報利用に慣れ親しんだユーザの行動様式が足かせとなって、初期段階で商業的に成功する事業モデルを描くことはできなかった。

CIXがインターネット接続事業者の団体であることからわかるように、ネットワーク環境をもたない利用者に接続を提供するISP（Internet Service Provider）という形態は容易に思いつく事業形態であり、比較的早い時期から事業化に成功している。

たとえば日本では、アメリカでCIXが設立された翌年の1992年にはAT&T Jemsが日本初となる商用ISPを開設しており、また同年には日本企業として最初の商用ISPとなるIIIが接続サービスをはじめている。インターネット以前の主流だったパソコン通信最大手のニフティサーブ（現@nifty）などもプロバイダ事業を提供し、翌年にはTWICSが日本初の個人向けISPの開始、さらに1995年にはダイヤルQ2を利用したISP事業や無料プロバイダといった様々な事業形態が出現した。さらに、1996年にはNTTによるプロバイダ事業OCNが開始されるなど、わずかな期間にISP事業が次々と出現することになる。

ISPのサービス提供によってインターネットに接続できる環境を得た個人利用者が、利用したサービスは電子メールとWebだった。初期段階ではGopherやネットニュースなど他のプロトコルも利用されていたが、Webの登場とともにこれらのサービスはWebに統合されてゆき、電子メールとWeb以外のサービス利用は次第に減少してゆく。

この中で、電子メールは商業利用以前から存在していたサービスであり、インターネットの誕生とともに提供されたもっとも古いサービスの一つである¹¹⁾。これに対してWebの出現は1990年であり、まさにインターネットの商業利用とWebの発展は軌を一にしている。Webがインターネット普及を決定づけたいわゆるキラーソフトだったとあってよいであろう。しかし、インターネットの商業化とWebの出現がタイミング的に一致しているのはあくまで偶然の産物である。Webはハイパーテキスト構造をタグ混じりテキストによって実装し、インターネット上の文書管理を効果的に行うために開発された仕組みであり、最初から現在のような多様な商業利用を想定して開発されたわけではなかった。そのため、初期のWeb上のコンテンツは無償提供が当然のことと考えられていた。ISPに続く事業性のある領域を模索していた事業者の多くはWebの可能性

を高く評価し、会員制有料サイトなどの方法で有償サービスのあり方を模索したが、実際に有償提供に成功した例はこの当時ほとんどない。

この中で Web を用いた事業化に例外的に成功したかに思われたのが、EC である。EC の場合には、従来からの流通、とくに通信販売などの手法をそのまま援用して事業構築が可能であった。極論すれば初期の EC とはカタログを Web に置き換えただけであり、決済や物流などは既存の通信販売のしくみをそのまま利用していた。その後、インターネット上の電子商取引では、決済方法なども実装され、カタログ通販などとは一線を画するシステムを導入していくことになる。たとえば、後述する行動ターゲティングや協調フィルタリング¹²⁾ などインターネット独自の手法を展開させながら、既存の通信販売とは異なった事業形態を作り上げていく。しかし、最初の段階では上記の通りインターネットあるいは Web 特有の事業モデルを構築したとは評価できない状態であり、ゆえに電子商取引がインターネットの純粋な商業利用のモデルとなったとはいえない。

さらに、電子商取引サイト Amazon.com の事業モデルは、創業期から多くの投資家が有望視していたにもかかわらず、現実には長期にわたって黒字転換することができなかった¹³⁾。これに対して電子商取引サイト以外のポータルや検索サイトなどの業態は創業からわずかな期間で投資家の期待する水準の利益を生み出した。ドットコム・バブル期にはこれら EC 以外の業態に投資家の関心が集中し、EC 以外のサービス事業者が市場から調達した資金をもとに急速に事業規模を拡大していったのである。

3. インターネットの可視化とネットワーク形状の変化の検討

たとえば、Internet Mapping Project¹⁴⁾ や Opte Project¹⁵⁾ など、インターネット可視化の試みの成果から、インターネットの実際の姿は、われわれが普段想像する姿とはかなり異なっていることがわかる。まず、それは均質なネットワークではない。散らばったノードの配置は密度にかなり濃淡があり、巨大なサービスとそれを取り巻く無数のサーバから構成される島宇宙が漆黒の空間に散らばっているかのようなのである。この濃淡の差をもたらしているのは、ユーザのアクセスが集中するノードの存在による。インターネット、そしてその上に展開されるサービスである Web は本来のサーバはそれぞれが対等の関係にある。しかし、実際には中心と周縁をもつ構造になっている。このような島宇宙の中心にあるのは、ひとつには IX (Internet eXchange point) や ISP などトラフィックが集中するノードであるが、さらに、これに加えて巨大サービスが位置する。

さらにこの構造にはもう一つ注意すべき点がある。この図に示されたノードはすべてサーバやルータなどネットワーク機器である、という点である。インターネットはこのようなネットワーク機器から構成されるネットワークであり、パソコンから ISP を介して接続する利用者は、インターネットの構成者ではなく、外部からの参加者に過ぎない。すなわち現状のインターネットと

は、当初の均質なネットワークではすでになく、インターネットに参加した組織や個人はネットワークの中に位置しているわけではない。このことはあたかもインターネット全体がひとつのコンピュータとして機能し、その中心にあるサービスへのアクセス手段と化したかのようにみえる。Nicholas Carr¹⁶⁾の表現を借りるならば、「ワールドワイドネットワーク」が「ワールドワイドコンピュータ」と化したこととなる。このような現象が生じた背景には、サービス提供者が広告による事業モデルを目指したことがあると思われる。次にこの点について整理を進めたい。

4. 広告による事業モデルの構築

初期のインターネットでは、互酬性の文化が残る中で、利用者は無償サービス提供を求めた。一方で提供側にとってサービスの維持運営には巨額の投資と資金が必要であった。とくにサイトの規模が大きくなると、運営費用も莫大な額に達し、運営費用の調達サービス提供者にとっての大きな課題となっていた。

このための資金を調達する手段として、広告モデルが生まれ、広告モデルの要求によって Web サービスは、マスメディアと比肩する規模に成長する。まず Web ページを媒体と考え、このスペースを企業広告に提供することで収益を確保するという事業形態が生まれる。広告モデルは、従来からもマスメディアの事業を成立させるのに不可欠の手法となっていたのだが、これを単純に Web に適用したのが初期の広告モデルによる Web サービスの事業運営であった。

マスメディアの多くは広告収入によって成り立っている。新聞や雑誌は広告以外に販売収入もあるが、紙面の一部は広告に提供され広告収入は大きな収益源である。購読者は広告も含めた紙面を購入しているが、「広告もまた重要な情報である」という論法によって広告のある紙面を購読するという行為は正当化されてきた。一方でテレビやラジオは民放である限り、視聴者に費用負担はもとめられず、ほとんどすべての運用は広告収入により賄われる。

新聞・雑誌とテレビ・ラジオとの事業モデルの違いは、モノとしての媒体の移動が伴うか否かの違いである。紙というモノが購入者の手元に残る新聞雑誌では媒体が販売の対象となり、モノの移動を伴わない電波媒体では無償の提供がなされた。モノの移動を伴わないという点はインターネットサービスも同じである。ここでラジオ・テレビに近い広告モデルが Web の最初の収益モデルとして採用されることになった。

このモデルでは、媒体価値の向上が最大の使命となる。媒体価値を向上させることで、将来のスポンサーに対しては媒体価格を上昇させることができ、既存のスポンサー企業に対しては広告掲載料の費用対効果を向上させることができる。

なお、ここでいう媒体価値とは、いかに多くの閲覧があったかということのみを意味しており、質のいいコンテンツを提供していることを意味しているのではない。その後インターネットでも、顧客の細分化、セグメントという概念が導入されるが、ここでも問題とされるのは、同質の利用

者をより「多く」擁しているかであって、実質的には量の問題がすべてだった。

Web サイトに限らず、ほとんどの広告媒体は利用可能な資源の総量が限定されているがゆえに、希少な資源と見なされる。たとえば新聞の紙面は枚数に制限があり、記事を除くスペースは一定である¹⁷⁾。媒体の希少性はテレビやラジオではさらに顕著となる。新聞や雑誌のページ数はテレビ・ラジオでは時間に相当する。時間という資源は常に一定である。しかも視聴者の行動パターンは似通っている。多くは夜は眠っており、朝起きて、昼間は学校や仕事にでかける。したがって、媒体に接触する可能性のある時間となると、24時間のなかでもさらに限定されたわずかな時間しかない。

このようにつねに広告スペースは希少資源であり、需給関係によって、価値が高いと判断される媒体価格は高騰することとなる。Web サイトではページを増やすコストは新聞にくらべるとはるかに安価であり、容易であるが、Web ページでもトップページ以外の閲覧可能性は低く、さらにアイトラッキング手法を用いた分析からも明らかにされつつあるように、トップページの中でも注視される範囲は限定されている。

この手法が初期の Web に持ち込まれる。サービスを提供するサイトは、利用者のニーズに合った人気の高いコンテンツを提供することで多くのユーザの支持を獲得する。このようなユーザの支持が高い媒体の一部をバナー広告といった形態で広告スペースとして広告主に提供する。

Web サイトはプル型メディアであるがゆえに、視聴者数の正確な把握さえまならないテレビなどとは違って、ログ解析により詳細な利用状況が把握可能であった。後発の広告媒体であるにも関わらずスポンサーや広告代理店が媒体としての可能性に早くから注目したのはこのような理由からである。このようにして一部のサイトは優良なサービスの提供、媒体価値の向上、収益の改善と投資、サービスの充実というサイクルを形成することに成功した。

さらに、この広告モデルによってサイトの維持運営と拡張のための資金確保が可能となった。このことで、多くの広告を集めたサイトは、その後の発展のためにその資金を設備投資にまわす余力がうまれる。このような資金を設備投資にまわすことで、初期の電子商取引がその後インターネット特有の性質をもったサービスへと成長していったように、インターネット広告でも PV (Page View) や Visit¹⁸⁾などを評価指標にして媒体価格が決定される値付けの方法から、実際の効果に近いクリック数に応じた課金など、インターネット独自の様式へと成長していった。インターネットの特性を活用した広告モデルによって、インターネットサービス事業者は収益性のあるモデルを確立したのである。

もちろん、純粋に広告収入だけでそれが可能になったわけではない。このような商業化の可能性が株式市場で評価され、株価の急激な上昇がこれらの企業に巨額の運営資金を供給したことの方が影響は大きい。インターネット関連サービスに対する株式市場での過剰ともいえる評価には、当時のアメリカのクリントン・ゴア政権が1993年に提唱した情報スーパーハイウェイ構想も影響

を与えたことは間違いないであろう。「政策に売りなし」という市場の格言通り、アメリカ政府の政策が、スタートアップ期にあったネットベンチャーとよばれる企業群に巨額の資本を提供したのである。ただし、このことが後にドットコム・バブルの崩壊によってインターネットサービスが打撃を受けることにつながるのであるが、少なくともこの当時は大きな追い風となっていた。

このようにして、巨額の投資をおこなったことで、いくつかのサービスはより多くの利用者を集めるサービスを構築することができた。その結果、多くの利用者が集まることで、その活動を記録した巨大なデータベースを擁するサイトが出現することとなったのである。

Ⅲ. O' Reilly による Web2.0

最初に述べたように Web2.0 は Tim O' Reilly が 2005 年に提唱した概念である。提案者とされる Tim O' Reilly は、動物を木版画で描いた表紙が特徴的な情報処理関連書の出版社オライリーメディア（旧・O' Reilly & Associates）の創設者として知られる。O' Reilly 自身はオープンソースの支援者として知られており、シリコンバレーのインターネット企業家コミュニティの中心人物の一人であるともいわれている。

漠然とした傾向に名前が与えられることでリアリティーが生まれることがある¹⁹⁾。出版業界に身を置く O' Reilly もこのことをよく理解していたと思われる。2001 年のドットコム・バブル崩壊により、市場から豊富な資金を容易に調達し、順調に発展を続けていたインターネット業界は突然の危機に見舞われる。O' Reilly はその後の不況が続くインターネット業界の状況を、間近に見たはずである。そこで O' Reilly は Web やインターネットの重要性を社会に再認識させる目的でカンファレンスを開催し、その成果をまとめた文書を 2005 年に発表する。この“What Is Web 2.0 Design Patterns and Business Models for the Next Generation of Software”のタイトルに用いられたのが Web2.0 という表現だった。

バージョン番号を付けてリリース時期を管理する方法は、ソフトウェアのバージョン管理ではよく用いられる。Web2.0 という命名の意図はそのような時代の空気を実体化して、流行を作ることにあった。

1. Web2.0 の特徴

この論文の中で O' Reilly は、ドットコム・バブル崩壊以降に出現したインターネットサービスを過去のサービスと対比させて相違点を洗い出し、この作業を通してドットコム・バブル以降の Web サービスの特徴を以下 7 項目に要約した。

Folksonomy：ユーザによる情報の分類（タグ付け）

Rich User Experiences：リッチなユーザ体験

User as contributor：ユーザの貢献

The Long Tail：ロングテール

Participation：ユーザ参加

Radical Trust：絶対的な（ユーザへの）信頼

Radical Decentralization：絶対的な分散

このうち、Folksonomy、User as contributor、Participation、Radical Trust はユーザ参加にかかわる内容であり、Rich User Experiences と Radical Decentralization の 2 項目はシステム技術にかかわる要素、The Long Tail は事業モデルにかかわる要素である。このように Web2.0 の定義には雑多な要素が盛り込まれているが、その中心をなす概念は、ユーザ参加と集合知であるといえるだろう。

Web2.0 以前の Web のサービスでは情報の提供側と受け側は厳密に区分されていた。これに対して Web 2.0 では、SNS をはじめとする各種のソーシャルメディアに見られるように、コンテンツをユーザが提供することを前提とするサービスが登場する。また、その整理もユーザにゆだねられる。従来の Web では、Yahoo!のディレクトリ型検索に代表されるようにサービス提供側が情報の整理を行ってきた。これに対して Web 2.0 では、ソーシャルタギングなどの手法によりユーザの手によって、情報の分類・配置が行われる。O' Reilly は、写真共有サイトの Flickr や、ソーシャルブックマークの del.icio.us などこの例として挙げている。このようにユーザがコンテンツの生成から、整理、評価やレビューまでを受け持ち、ユーザの提供する情報が価値を生み出していく。これが Web2.0 がもたらした大きな変化であった。またこれは互酬性文化の復活といってもよい。このような O' Reilly の提案は、巨大化したサービスに囲い込まれることに若干の疑問を感じ始めていた利用者にも、新たな事業モデルを模索していたサービス提供側の事業者にも、大きな期待を抱かせることとなった。

2. Web2.0 の流行と衰退

Web2.0 に関しては、流行を引き起こすことだけに重点が置かれていたとの批判がある。実際に Web2.0 は大きな流行をうみ、流行は短期間のうちに終焉を迎えた。O' Reilly が Web2.0 を提案した背景には、ドットコム・バブル後の IT 業界、とりわけインターネットサービスの沈滞した状況をふたたび活気づけるために、新しい旗印を作ろうという意図があったことは確かであろう。そして、O' Reilly の目算通り Web2.0 という言葉の力によって、ドットコム・バブル後の沈み込んだ IT 業界の空気は一扫され、インターネットユーザは再びネットへと関心を向けはじめた。この点は十分に評価に値すると思われる。たしかに、その影響は O' Reilly の当初の想定をはるかに超えて拡がってしまうこととなり、加熱感は否めない。しかし、これをもって Web2.0 が無意味であったとはいえないと思われる。

Web2.0 の流行が頂点に達した 2006 年頃には、無関係なサービスまでもが Web2.0 を名乗るようになり、明らかに加熱した状況が生まれた。流行はインターネットの世界にとどまることなく、

社会全体に2.0ブームを引き起こすこととなる。2006年7月6日のNIKKEINET IT plusはこの騒動のありさまを様々なケースを紹介しながら伝えている²⁰⁾。たとえば、インターネットサービスの関連では、GMOインターネット証券の「ネット証券2.0」やニフティの「Webメール2.0β」、米シマンテック「セキュリティ2.0」、ディー・エヌ・エーのアフィリエイト（成果報酬型）広告サービスは「アフィリエイト2.0」など、2006年頃のインターネット業界は2.0一色であった。インターネットサービス以外でのケースとしては、たとえばNTT番号情報の「iタウンページ2.0」やトランスコスモスの「マーケティング2.0」などが紹介されている。過剰な期待を集め流行を引き起こした言葉がたどる結末はみな似通っている。今ではWeb2.0はすっかり手垢のついた言葉になってしまった。かけ声だけであまり実態のないことを指すパスワードという言葉がある。Web2.0は短期間のうちにパスワードを代表する言葉のように扱われるようになったのである。

このようなありさまを、「ドットコム・バブル」と対比させて「Web2.0バブル」とよぶことがある。ただし、ドットコム・バブルにくらべると、経済的な影響は限定されていた。早い時期からWeb2.0の事業性が低いことが明らかになってきたことで、投資家の関心も盛り上がりには欠けた。このため株価の高騰などの現象はドットコム・バブル期にくらべると限定的で、社会全体の関心事になったわけでもなかった。「2.0」という言葉が社会的に広範な影響をもたらしたのに対して、本来の意味でのWeb2.0への期待とは、技術者を中心にして、IT業界の内部と利用者の次世代サービスへの期待という範囲にとどまっていたのである。

Web2.0の影響が限定的であったのは、従来の広告モデルから脱却する新たな事業モデルを提案できなかったことにある。これはWeb2.0が特徴のひとつに掲げながら、分散型のシステムとして収益を確保する方法を提示し得なかったことに起因すると考えられる。

IV. インターネットのマーケティング手法に潜む危険性とインターネットへの信頼感の低下

ここまで整理してきたように、インターネットサービス、Web2.0登場以前のサービスは、利用者の行動履歴を内部に蓄積し巨大なデータベースとして構築することで持続的な競争優位を獲得するという戦略を採用してきた。Web2.0以後も、Web2.0が有効な事業モデルを構築できなかったために、この図式は今もなお有効である。

もっとも初期に成立した広告収入による事業運営モデルは今でも採用されている。しかしその手法は常にイノベーションがはかられ、媒体価値のさらなる向上、競合する他のメディアとの差別化のため、より効果的な手法へと展開されている。しかし、より精緻になったマーケティング手法は、ひとつのサービスで蓄積された情報以外の情報を必要とし、個人のより精密な特定を要求する。このことが、時として許容される範囲を超え、個人情報に危機にさらすこともありえる。

インターネットマーケティング手法の危険性については、AOLの検索履歴公開²²⁾などいくつか

の事件、あるいは Hermann Maurer²³⁾ や Nicholas Carr²⁴⁾ らの指摘にもその予兆があらわれはじめている。この問題は、今後インターネットへの信頼を大きく傷つけ、インターネットの継続的な発展を阻害する危険性をはらんでいる。実際に、ヨーロッパを中心に Google が検索市場を独占することに対しては懸念を表明する意見が相次いでおり²⁵⁾、Google Street View の公開ではプライバシー問題を中心に Google に対する批判が世界各地で巻き起こったことは記憶に新しい²⁶⁾。

1. 効果的なインターネット広告のために必要な情報とは

インターネット事業者にとって重要な事業収入である広告をより効果的なものとし、インターネット広告の優位性を示すためには、他のメディアでは採取することのできない利用者情報を得る必要があった。

このために、初期から導入された手法は、利用者のアクセスとともにサービス側に蓄積されるログデータを解析する方法である。ただし、インターネットのログに記録される情報は、ブラウザがはき出す環境変数と呼ばれる情報のみにすぎない。環境変数にはアクセスした時間や、IP アドレス、ブラウザの種類、アクセスする前に見ていた URL などが含まれる。これらの分析からサイト運営に必要な情報を得ることはできるが、マーケティングデータとして考えた場合にはこれだけでは必ずしも十分とはいえない。

たとえば、ログデータの示す情報から同一人物を特定する根拠はない。IP アドレスから推定は可能であるが、常に同じ IP アドレスから同一人物がアクセスするわけではない。プロバイダが動的に割り振った IP アドレスは常に変化しており、職場やモバイル環境など、普段とは違う場所からアクセスした場合も同様である。

さらに、個人属性は環境変数には一切含まれていない。個人属性は広告のパーソナライズのために重要な手がかりを与えてくれるが、趣味・嗜好などはおろか、性別、年齢などの基本的なデモグラフィック属性も環境変数の解析からは得ることができない。Web サーバを運用していれば環境変数はほぼ自動的に入手可能な情報であり、マスメディアで同様の情報を得る場合に必要なコストと比較するとはるかに容易ではあるが、効果的な広告を行うためには用途は限定的であるといえよう。

この点を補うため、また個人情報保護の観点からも、個人情報を利用者本人に自主的に登録してもらう必要があり、このために会員登録という方法が採用されることになる。利用場面はプライバシーポリシーで示した範囲に限定されるが、登録して以降は ID、パスワードで認証された時点で特定の人物特定が可能となる。

ただし毎回 ID とパスワードを要求することは煩瑣な作業であるため、顧客ばなれをまねく可能性があるため、Cookie²⁷⁾ が用いられる。Cookie は、Web の提供者が、利用者のパソコンに情報を保管するためのファイルを作り、そこに必要な情報を保存させておく仕組みであり、もとも

とは、ユーザの認証や、個人別のカスタマイズのために考え出された Cookie はローカル環境上に保護されるため、別の端末からのアクセスでは意味をなさない。ただし、ローカル環境にあることで、利用者がファイルを削除する自由は保障されており、自己コントロールやセキュリティ上のメリットはある。

これに対して、最近問題となっているのが、サードパーティ Cookie²⁸⁾ である。サードパーティ Cookie は複数ドメインをまたがってユーザをトラッキング出来る Cookie であり広告配信に用いられる。この性質からスパイウェア対策ソフトはこれをスパイウェアの一種として扱っている。米国では、このようなサードパーティ Cookie の利用が消費者のプライバシーを侵害するとして、大手広告会社が 2000 年に集団訴訟を提起されるという事態にまで至っている。

この具体例は Google のサービス展開にもみることができる。Google ではもともとユーザ登録をせずとも検索は可能であり、Google の事業モデルの中核にある Google Ad Words 広告は検索した語や句に応じた広告を検索結果と同時に画面に示すことで、利用者の個人情報を取得することなく検索する側・広告主の双方に効果的な広告表示が可能な方法を確立した。

しかし、その後に Google が追加したサービスは、基本的に ID を取得して利用することを前提としている。その代表的なものはブラウザメールの Gmail であり、パーソナライズされたポータルである iGoogle である。現在でも ID 取得を必ずしも必要としないサービスも多いが、一端これらのサービスを利用してログインした状態で他のサービスを利用すると、ID 取得の際に申告した個人情報と利用履歴はひも付けされた状態で Google のサーバに蓄積される。ID の取得に必ずしも正確な個人情報を申告していなくても、Google のサービス間の利用履歴は関連づけて蓄積されることになる。たとえば、検索でどのような言葉をつかったのか、地図でどこを表示したのか、誰からどのような内容のメールを受け取って、どのような内容の返信をしたのか、Google Books でどのような本を選んだのか。このような多様な情報が特定の人物の行動履歴として一元化され蓄積されるのである。

これは Google だけに限らない。初期の段階でデータを蓄積していた多くのサービスは ID を取得することのメリットを利用者に示して ID 取得を促す。そして莫大な開発費を投じたサービスを無料で提供するかわりに、個人の履歴情報の提供を要求する。以前は無料でサービスを提供する見返りに要求されたのは、広告の閲覧であった。あるときから、無償の代価は行動履歴情報の提供へと変化していったのである。

2. 行動ターゲティング広告に見るインターネット広告の暴走

このような経緯をたどって、現在のインターネット広告では、閲覧者のアクセス履歴から嗜好を分析して最適と思われる広告を配信する行動ターゲティング広告に注目が集まっている。この行動ターゲティング広告には、サードパーティ Cookie と同じ問題がある。高木²⁹⁾ は最近の行動

ターゲティング広告、とりわけの自サイトでの閲覧行動だけでなく他のサイトでの閲覧行動まで追跡するタイプのあり方に疑問を投げかけた。高木は楽天が導入した「楽天 ad4U」と「携帯 ID」を用いた行動ターゲティング広告を例にとりながら、その手法に潜む倫理的な問題点を指摘する。「楽天 ad4U」は、訪問の有無を確認したい特定のサイトの URL を画面上には表示されない隠しリンクとして画像などに埋め込んでおき、訪問したことを示す表示色を取得するプログラムを仕掛けるというものである。実際に高木が楽天 ad4U の広告を調べてみたところ、Flash オブジェクトの中に数千個の隠しリンクが埋め込まれており、JavaScript によってそのリンクの訪問の有無を調べ、どんなカテゴリーのサイトに多く訪問しているかを集計し、そのカテゴリーの広告を表示するようになっていた、と報告している。ここで、リンク色の変化を外部から取得できるのはブラウザの欠陥であり、このような欠陥を突くことによって閲覧履歴を取得するという点が問題であるとす。

さらに高木はこの問題は単に倫理上の問題にとどまらないとも指摘する。人がブラウザを使用するに際して「その意図に反する動作をさせる」プログラムであるがゆえに、このような方法で閲覧履歴を参照するプログラムは、刑法改正案として検討が進んでいる「不正指令電磁的記録作成等の罪」のいう「不正な指令」に該当するおそれがあると述べる。現状は法的に規制される対象とはなっていないが、すでにいくつかのウイルスワクチンソフトでは、第三者 Cookie と同様にこの種の広告に仕掛けられたスクリプトをスパイウェアと認識して警告を発するものもある。

さらに高木がもう一つの例としてあげるのは、携帯 ID を用いた行動ターゲティング広告である。この問題については紙幅の関係でここでは詳細に述べないが、このような携帯 ID による個人の特定は、第三者 Cookie よりも危険性が高いと指摘する。なぜなら携帯電話の携帯 ID は携帯キャリアが付与した固定的な番号で、Cookie のように自分の判断で削除することができない。そのため、携帯 ID を用いた広告が普及すると、広告会社により追跡される可能性は携帯電話を解約するまでの何年にも渡って続いてしまうからである。

このような状況について、利用者側にもすでに不信が広まりつつあることは、以下の調査結果からも伺うことができる。USC（南カリフォルニア大）Annenberg School が刊行している 2008 年の Digital Future Report³⁰⁾ では、検索エンジンから得られる情報を信頼しているユーザは 51% であり、2006 年の 62% から 11% 低下したという結果が報告されている。また、マイクロソフトは 2008 年 11 月に個人情報保護に関する説明会を開催したが、そこで同社の最高プライバシー責任者は、プライバシーの脅威によって、インターネットに対する信頼が世界的に低下していると述べ、米国やドイツでは、3 割のネット利用者が ID 盗難などを恐れて、ネットの利用を控えているとした³¹⁾。

このように過度に発達した広告手法やマーケティング手法について、すでに利用者は不信を抱きはじめています。このような状況に対して、提供側は有効な方策を用意しているのだろうか。以下では、Web2.0 以降の Web サービスの動向についてこの点を中心にまとめていきたい。

V. Web2.0以降の展開 — セマンティック Web —

1. セマンティック Web が注目される理由とは

Web2.0の流行が終焉を迎えた後、これからのあり方を再度見直そうという動きがWebにうまれている。その代表的なものにセマンティック Web がある。

セマンティック Web は、Webの開発者である Tim Berners-Lee が 2001 年に発表した論文³²⁾で提唱した概念である。Berners-Lee は、Web が今後さらに高度化していくためには、HTMLにかわる高度に構造化された文書形式が必要であると述べ、Web ページは従来の HTML ではなく、XML や XHTML によって記述されなければならないとした。さらに、構造化された文書だけが存在しても機能しないため、高度に構造化された文書を処理する仕組み、たとえば一群の推論ルールと、オントロジーや構造化された知識表現、人工知能的アプローチと組み合わせなければならないと述べた。具体的には文書の中に RDF (Resource Description Framework) というメタデータの記述言語を用いて情報の意味を埋め込んでおき、メタデータで使われる語彙や分類体系、言葉の相互の関連などの構造は OWL (Web ontology language³³⁾) という言語で記述することで、エージェントが自動的な推論をする手助けが可能であるとした。さらに、セマンティック Web の可能性を存分に発揮するには、それを単一のグローバルなシステムに組み入れなければならないとして、これらの技術を自らが代表を務める Web の標準化団体 W3C によって標準化した。

論文の発表時期が 2001 年であることからわかるように、セマンティック Web という言葉は Web2.0 とほぼ同時期に提案されたことになるが、Web が提案された 1989 年の段階ですでにセマンティック Web というアイデアは存在していたと考えた方が適切かもしれない。これほどにセマンティック Web という考え方は Web の根底にある問題意識である。では、これまでなぜ実現しなかったのだろうか。これはセマンティック Web を実現するために不可欠な構造化された文書が利用できる環境が、Web2.0 によってはじめて可能となったからであると考えられる。

セマンティック Web が想定する高度に構造化されたタグを生成する方法はいくつか考えられる。まず、サービス提供側が既存のプレーンテキストあるいは HTML 文書にメタデータを埋め込む方法が考えられるが、これには限界がある。Yahoo! はかつてディレクトリ型検索エンジンとして出発し、他のロボット型検索エンジン、たとえば後発の Google などと検索エンジンとしての方式で覇を競った。しかし、現在ではロボット型検索の優勢は覆すことができないところまできている。現在でもディレクトリ型検索はサービスとしては継続的に提供しているものの、Yahoo! のトップ画面からは外れている。このことからわかるように、人の目で大量のテキストを読み、それに構造を与えることには限界があった。

これに対して、データが発生した段階で、自動的に、あるいは作成者が自らメタデータを与える方法は、データ生成後にメタデータ化するよりは効率的である。あるいは、データの作成時点でテキストの作成者がデータの構造を自ら指定するという方式も考えられる。従来から形式の

ニュースなどでは RSS によりその種の方法はある程度実用化されている。ブログでも同様に構造化された自動生成することが可能である。Web2.0 と同時期の提案であったにもかかわらず、セマンティック Web が、Web2.0 ブームが沈静化した時期に注目されるのは、Web2.0 以降に実現したユーザー側で提供する豊富なコンテンツをベースとして半自動的に構造化できる環境が整ってきたためと考えられる。

2. セマンティック Web の系譜

ここでセマンティック Web の系譜を振り返っておきたい。

XML の元になった SGML (Standard Generalized Markup Language) は Web の記述言語である HTML よりも長い歴史をもっている。SGML は軍艦や軍用機などのマニュアルなどの文書を電子化するために定められた電子書式の規格である。軍用機や軍用艦船のマニュアルは膨大な量に達するため、納入業者にとってもその編纂は負荷の大きな作業となるばかりでなく、納入された側にとっても保管や管理に膨大な作業を強いることになった。このためマニュアルの電子化は納品業者にとっても発注者である軍隊にとっても重要な課題だったのである。これを実現できる方法として SGML が提案され実用化されることとなった。

対象となる文書はマニュアルなので本文はテキストで書かれている。これにタグによって、文書の構造を与えることで、文書の構造を本文のテキストと一括で管理することが可能となる。この文書进行处理の際に、このようなテキストとタグが混じった文書をマークアップ文書と呼ぶ。SGML は実用的なマークアップ言語としてはじめての試みでもあった。

SGML の歴史は 1979 年に IBM の Charles Goldfarb らが発表した GML (Generalized Markup Language) にまで遡ることができる。GML は概念モデルであったが、Goldfarb らは理想的なマークアップ言語は以下の二つの条件を満たすものでなければならないとの信念から、理想的なマークアップの構築に取り組む。理想的なマークアップ言語の条件とはすなわち、

- ・マークアップは物理的なレイアウトではなく、文書の構造を反映したものでなければならない。
 - ・マークアップはコンピュータと人間の両方が理解可能な形式でなければならない。
- とするものである。

この取り組みの意義を認めた IBM は DCF (Document Composition Facility) として商品化する。Goldfarb 自身はその後、IBM を退職し、GML をさらに発展させた SGML の開発に取り組む。SGML は 1986 年に ISO の規約として出版されることとなり、これが契機となって大きな注目を集めることとなった。

ただし、SGML は ISO の正式承認以前から、アメリカ国防総省や EU の公式出版事務局など数々の公的機関や Goldfarb が在籍していた IBM 社ですでに使用されはじめていた。日本においては、

当時、厚生省への新薬申請、当時の建設省では公共事業の工程管理、特許庁などで利用されている。中央官庁のシステムに採用されたことで、中央官庁の管轄下にある企業、あるいは申請が必要な製薬会社、公共事業を受注するゼネコンなど中央官庁との関連の強い業界では一定の普及をした。

しかし、理想的なマークアップ言語を目指して開発されたことで、SGMLの規約は非常に厳密であり、当初の設定や、文書の作成など、作業は非常に煩瑣なものとなった。このため、SGMLを採用したのは、部品点数が多いことで詳細な添付文書が必要となる航空機産業や防衛産業、自動車産業などの大規模産業にとどまることとなる。

なお、Berners-LeeがWebを開発した当時に在籍していたCERN(欧州原子核研究機構)ではISO承認とともにSGMLによる文書管理に着手している。Berners-Leeは1980年から途中離籍した時期をはさんで1990年のWebを発表した時期までCERNに在籍しており、早い時期からSGMLに接していたはずである。そしてBerners-Leeが提唱したWebとは、SGMLを簡略化したHTMLによって実現したハイパーテキストシステムであった。

SGMLはその精緻な構造ゆえに普及することはなかった。この様子をつぶさにみていたはずのBerners-Leeは、あえて複雑な規約を持ち込まず、最低限の約束事だけで単純化する方向を選択したのではないだろうか。これは推測に過ぎないが、この結果、HTMLは短期間に普及することが可能となったことは事実である。しかし、その代償として、文書の構造を記述するというXMLのもっとも重要な機能は犠牲にならざるをえなかったのである。

なお、普及が進むにつれてHTMLの単純化の弊害が目立ち始める。HTMLが文書の構造を簡略化して、ブラウザに与える表現にのみ特化したため、HTMLの検索には限界があった。Googleなどのロボット型検索エンジンは、この欠点を補う方法として語や句の組み合わせにより最適な情報に到達する手段を提供しているが、これはあくまで代替的な手段に過ぎない。本来はもっと効率的で効果的な情報検索の方法が存在しているはずであった。検索エンジンが見つけてくる膨大な候補は、単に文書に含まれた言葉に一致しているに過ぎない。文書の構造をしめす手段のないHTMLの表記法では、その文字が文書中でいかなる意味をもって書かれているのかは、検索して実際に人が読んでみるまではわからない。

以下ではGoogleなどでも採用されている位置情報をもとにした検索を通してこの問題について説明する。位置検索でもっとも単純なアイデアは文書の中の「住所らしき文字列」に注目する方法である。テキスト文書では、一般に位置情報は住所という表記で示される。そこでこの住所らしき文字列を、一般のロボット型検索エンジンのようにサイトをくまなく訪問し、HTML文書に含まれる言葉を丹念に拾い集める。ここで語よりも少し長いフレーズを検索し、それが住所としての要件を満たしていれば住所と判定する処理を行う。この住所らしき文字列を見いだす作業は、全国の住所を収録したデータベースがあればそれほど困難ではない。このテーブルは後に二

点間の距離を測る際にも必要なので、この処理には必ず必要なデータである。このテーブルを参照し、文書の中から「住所らしき文字列」を見つけ出したら、その都度、位置に変換して上で、この文字列が含まれていたもとの HTML 文書とセットにして保管しておく。住所の位置情報を別のテーブルに持っておけば、「住所らしき文字列」間の距離の計算は容易である。特定の場所で検索をすれば、あらかじめデータベースに格納した位置情報から距離を基準に候補を並べれば位置情報検索が可能になるという仕掛けである。

ただ注意しなくてはならないのは「住所らしき文字列」が必ずしもその HTML 文書が示したい位置を明示的に示しているわけではない。あくまでデータベースに収納されたデータとある種のアルゴリズムによって住所とおぼしき文字列を取り出したに過ぎない。これは、住所録のような形式の書類を HTML 文書化した場合を考えてみればすぐに理解できるだろう。個人情報保護の浸透とともに個人の住所一覧をインターネット上にみつめることは、ほとんど不可能になったが、今でもなお、企業や組織の住所の一覧を記載しているページはたくさん存在する。たとえば、企業の Web サイトでは、自社の組織を示すページに支店一覧があり、その一覧には多くの住所らしき文字列が書かれている。また団体や会員企業名簿に住所が記載されている場合を考えてみればよい。このようなページには「住所らしき文字列」は一つだけでなく、それらしき文字列をたくさん発見することができる。しかし名簿で発見された住所が、その HTML 文書の「位置」であるかは、この情報だけで判定することはできない。

このような問題が起こるのは、文書中の特定の文字列が住所の表示であることを明示的に示す仕組みが存在していないことに起因する。XML のようにコンピュータが理解できる形式で記述されない限り、どのような意図でその文字列が書かれたのかは検索エンジンには理解不能である。そのため、キーワード検索では HTML 文書中の文字列をひたすら検索することになる。またこのことは、アルゴリズムに依存した情報処理の限界と、データ構造の重要性を示唆しているとも考えられる。

3. セマンティック Web は普及するのか

ここで問題となるのは、Berners-Lee 自身が述べるように、現在の Web 全体がセマンティック Web の提唱する構造に置き換わるほどに普及しなければ機能しないことにある。そしてそれが、参入をする事業者に必要な経済的メリットをもたらす仕組みを作れるかにある。

Morville³⁴⁾ は Weinberger³⁵⁾ のセマンティック Web への懸念を引用しながら、「私が危惧するのは、セマンティック Web が SGML と同じ末路をたどるのではないかと、それも基本的に同じ理由で、ということである。メタデータの正規化が真に機能するのは、高い報酬がえられ、管理権限が一方所に集約されており、ユーザに規範を遵守させやすいような、限定的なアプリケーションにおいてである。言い換えれば Web においてはうまく機能しない、ということだ」と述べる。

どの程度普及するかは、セマンティック Web が提供するメリットと、そのための労力のバランスによって決定されることになるだろう。セマンティック Web の目指すところとはなんであろうか。神崎³⁶⁾は、Web は人間が読むための「文書の Web」から様々なデータを自在に発見して利用できる「データの Web」へと向かう、とその性質を簡潔に述べる。テキストに構造を与えるのは、人間のためではなく、コンピュータのソフトウェアがデータを理解し、人間を助けるためなのであるとする。

つまりセマンティック Web とは基本的に AI を指向する発想が根底にある。Google の設立目的は「人類が使う全ての情報を集め整理する³⁷⁾」ことである。Google の創設者である Larry Page と Sergey Brin が過去のインタビューにおいて「情報技術が人間の頭脳と肉体の物理的限界を克服する手段であり、人工知能を実現するために Google はサービスを構築している」と述べているように、サービス全体としての発想の根底に AI 指向があることはよく知られている。現在では Page と Brin が Google の経営に直接関与しているわけではないが、Google の図書館プロジェクトである Google Book Search の担当者は、「人に読んでもらうためにスキャンしているのではない、AI に読ませるためにスキャンしているのだ」と発言しており、AI を指向する方向性は全社で共有されていると考えられる。

このような方向性とセマンティック Web の目指す方向性が一致しているならば、Web2.0 がたどったように、セマンティック Web も旧来の巨大なサービスのための機能として組み込まれることとなり、インターネットへの信頼を回復するための手段とはなり得ないだろう。

これに対して、セマンティック Web の可能性を積極的に評価する意見もある。2007年4月に Cnet が行った Web3.0 の定義コンテストで優勝した Robert O' Brien は、Web1.0 は集中化した彼ら (Centralised Them)、Web2.0 は分散化したわれわれ (Distributed Us)、そして Web3.0 は非集中化したわたし (Decentralised Me) と定義する³⁸⁾。ここで O' Brien は「(Web3.0 は) 世界に参加したくないときのわたしに関するものであり、自分の環境に誰を導き入れるかをより強く制御したいというわたしの側面に関係している。Web3.0 では、わたしの注意の対象が広がって、自分が注意を払うのは誰か、あるいは何か、そして自分を誰に見せるかということにまで及ぶ。それは、わたしにとってのより効率的なコミュニケーションなのだ」と述べる。O' Brien の定義にみられるような方向に展開できれば、セマンティック Web は Web2.0 が特徴として掲げながら実現し得なかった分散化を実現し、インターネットへの信頼を取り戻す手がかりを与えることになるだろう。

VI. まとめ

まずここまでの議論で明らかとなってきたことは、以下のように要約できる。

初期の Web では、広告を主体とする事業モデルにより Web サービスが事業的に成立しうるこ

とが明らかとなったことで、サービス提供者は媒体価値の向上を目的とした事業展開を行うこととなる。

このような事業展開を行うに当たって、参入障壁の低い Web サービスで持続的な競争優位を獲得するため、利用者の履歴データを蓄積し、この活用によって他のサービスとの差別化を図る戦略が選択されることとなる。この結果、競合サービスは駆逐され、トップシェアのサービスだけが生き残ることとなった。トップシェアのサービスだけが残ったことで、利用者のログ情報は巨大なデータベースと化し、インターネットは分散環境という本来の形態から、インターネット全体がひとつのデータベース化する状況、すなわちワールドワイドコンピュータ化してきた。

ただし経営戦略としてとらえた場合には、これらの企業がとった戦略はきわめて妥当であると評価できる。初期の Web サービスが構築した顧客の行動履歴に関するデータベースは、先行企業の競争優位の持続を可能とした。この戦略は Barney³⁹⁾ らによって展開された resource based view による経営戦略の成功例として理解することができる。Barney は競争力の源泉を内部資源に求めた。さらに、内部資源には従来からのヒト・モノ・カネといった有形資源に加えて、ブランド・プロセス・情報・知識などの無形資産が存在することを指摘し、両者の総和が組織の内部資源、すなわちケイパビリティであると定式化した。その上で、企業の内部資源が、持続的競争優位の源泉となるためには VRIO と略される 4 条件を備えていることが必要であると述べる。ここで V は Value であり、経済的価値を創出できること、R は Rareness で稀少であること、I は Imitability で直訳すると模倣可能性であるが、模倣が困難なことを意味し、最後の O は Organization でこれらの内部資源を活用できる組織が存在することである。

ここで初期の Web サービスが採用したデータベースの構築による競争優位の獲得という戦略を VRIO の視点から評価すると、先行企業の保有するユーザの履歴情報は、模倣困難で希少な内部資源であり、経済的な価値をすでにこれらの企業にもたらしている。さらに、このデータを活用するような組織とは、レコメンデーションエンジンなどの機能によって実現する。このように VRIO の要素をすべて満たしていることがわかる。ゆえにこの戦略は、単体の企業のとるべき戦略としては正しい選択であったといえる。

これに対して、初期にはデータベースと同様に希少な内部資源ととらえられていた媒体価値は低下する傾向にある。たとえば Google の Ad Words では、サイトのトップページなど閲覧の多いページを希少な資源としてとらえるのではなく、検索結果やメールの文面などと連動して表示させることで分散化し、広告主にとって最適な広告スペースを多様化している。従って、今日の Web では、ユーザの利用履歴のみが希少な内部資源と見なされ、広告スペースの重要性は相対的に低下しているといえよう。

しかし、このような戦略を主要なサービスがこぞって採用したことが、特定のサービスへの過度の情報の集中を招き、かつてマスメディアが「メディアの権力⁴⁰⁾」と化していったように、あ

る種の権力と化していく現象がみられるようになった。このような寡占化の弊害は、かねてより Google 八分⁴¹⁾ といった現象が指摘され警告されてきたことであるが、インターネット利用者全体にインターネットへの全般的な信頼の低下という現象となってあらわれている。このことは、ひいては Web サービスとインターネット全体への信頼を損ねることにつながりかねない。

本来はこのような方向性への対抗策として提唱されたはずの Web2.0 は、ユーザの参加によってインターネット上のコンテンツの増大と質的向上に寄与したと評価することもできる。しかし、寡占化に対抗するために絶対的な分散を掲げていたにもかかわらず機能上で本質的な分散化を実現できなかったがゆえに、利用履歴データベースにかわる有効な資源を見いだせず、これをもとにした経営戦略を構築できなかった。結果的に Web2.0 は有効な事業モデルを構築できないまま、Web2.0 以前の既存のサービスが Web2.0 の要素を取り入れ、さらに寡占状態を強化するという状況に至った。

このような議論を踏まえ、本稿では、以下の2点を結論として指摘したい。今後インターネットが利用者の支持と信頼を得ながらさらに発展するためには、広告と巨大データベースによる競争優位の維持にかわる収益構造が必要となる。さらに、インターネットが再び利用者の信頼を得るためには、巨大化したサービスにかわる、利用者が自らの個人情報を管理しうる仕組みが必要である。

まず、個人が情報を管理しうる分散化された仕組みとしては、「電子私書箱」と「EUROPASS 履歴書⁴²⁾」が参考となるだろう。ここで、電子私書箱とは、インターネット上に個人にかかわる情報をすべて集約する場所を設け、収納される個人情報の管理を自分自身で行うサービスのことを指す。わが国では、社会保障カードの導入と同時に進められている電子私書箱構想⁴³⁾を想起させる言葉であるが、本来は先に示した iGoogle のような無償のサービスが提供する環境に個人の情報をすべてゆだねるのではなく、インターネット上に自分自身で管理可能な個人情報の集約場所を設置して、自らの意志に基づいてサービスが参照する情報を自己コントロールできる仕組みである。わが国ではこれが、社会保障制度とセットにして議論されることで本来の機能とは別の視点からの議論となっている。本来的に公的機関が運用するものではなく、有償のサービスとして民営で運営されるべきであると考えられる。ただしこの仕組みそのものは Web2.0 が最終的な目標として掲げながらも、実現することを得なかった絶対的な分散を、個人を単位として再集約することで実現する手段として評価されるべきであると考えられる。

もうひとつの事例としては、2005 年から EU 諸国で導入がはじまった共通履歴書様式である EUROPASS 履歴書がある。EUROPASS 履歴書は多言語環境である EU 諸国で職業移動の自由度を保証するために導入された制度であるが、仕組みとしては本人の記載した履歴事項の各項目について第三者が認証し、公的な証明能力を保証すると同時に、記載された項目について本人が提示する内容を自由に編集することのできる仕組みである。この制度そのものは基本的にインター

ネットの利用が前提となっているわけではないが、個人に情報を集約する方法論としては、情報の本来の所有者である利用者個人と情報の信頼性を保証する第三者⁴⁴⁾の組み合わせとして、インターネット上の個人情報の保有形式の参考になる発想と考えられる。

さらに、上記の二つともセマンティック Web を前提としてはいないが、セマンティック Web の構造化されたデータが存在する場合には、その有効性はさらに向上することが期待される。従ってセマンティック Web はワールドワイドコンピュータあるいは AI を指向せず、この種の個人とインターネットを結ぶための手段として再定義されることで、利用者の信頼を得るための方法となり得るであろう。

なお、本稿では事業を継続するための収益モデルについては詳細に検討できなかったが、上記の個人情報を保管する機能が新たなサービスとして出現した場合には、これを有償化することで、分散化された収益構造を作り出せる可能性は指摘できる。この種のサービスを特定のサービス事業者が独占する場合には、その主体が個人情報をすべて閲覧可能な状態にあることになり、現状の課題をそのまま拡張することになる。これは主体が民間企業である場合も、公的機関であっても同じである。従って、この仕組みは公的機関が独占すべきものではないといえるだろう。この種のサービス提供者は官民間問わず、複数存在する必要がある。メタデータの形式だけを標準化しておけば、機能的には単純であるため、大手資本でなければ運用できない仕組みではない。

また有償で提供されることで、利用者 と提供者は契約関係にあり、無償利用の代償として情報の提供を求められることはない。ただ、すでに無償に慣れた利用者が、有償のサービスを受容する素地があるかが課題となるが、これは今後の課題として稿を改めて検討することとしたい。

註と引用文献

- 1) Tim O'Reilly : “What Is Web 2.0 Design Patterns and Business Models for the Next Generation of Software”, (2005/09), online : <http://oreilly.com/web2/archive/what-is-web-20.html> (2009/7/10)
- 2) online : <http://www.w3.org/History/1989/proposal.html> (2009/7/10)
- 3) 米国の株式市場、とりわけ通信関連銘柄が多い NASDAQ の平均株価は 1998 年ごろから急激に上昇し、2000 年 3 月には 6000 ドル近くまで高騰する。その後、連邦準備制度理事会の利上げを契機に株価は崩壊し、2002 年には 1000 ドル台まで下落した。この高騰をドットコム・バブルあるいはインターネットバブルと呼ぶ。
- 4) online : http://ir.yahoo.co.jp/jp/bizres/fy_index.html (2009/7/10)
- 5) 日経パソコン記事 2008/12/22 付けの記事による。Google がグーグルは 2008 年 12 月 22 日に開始した説明会で紹介した、ネットレイティングスの調査結果。2008 年 11 月のページビューを元にした数字である。online : <http://itpro.nikkeibp.co.jp/article/NEWS/20081222/321902/> (2009/7/10)
- 6) online : <http://www.jadma.org/pdf/press/press20081117.pdf> (2009/7/10)
- 7) Charles M. Wiseman : “Strategic Information Systems”, (Irwin, 1988)

- 8) 軽量なプログラミング言語 (Lightweight Languages) は、C 言語のような計算機リソースを多く消費しない言語を指す場合があるが、ここでの軽量なプログラミングとは、Ruby や Python, JavaScript のように Web サービスの開発に多用されている、学習容易でソースコードの修正や作成が容易なスクリプト言語を指す。O'Reilly は “What Is Web 2.0” の中で「軽量なプログラミングモデル」を Web2.0 の特徴のひとつとしてあげている。
- 9) 可用性とは、システムの壊れにくさの意味である。異状発生確率が低く、障害発生時にも修復が速やかに行われるシステムを可用性の高いシステムと呼び、利用の安心感とかかわる概念である。
- 10) 横河デジタルコンピュータ SI 事業本部：「インターネット商業化に向けて (CIX) —アメリカではいま」、トッパン (1993/11)
- 11) 電子メール送信の protocols である SMTP には最初、認証機能すら実装されていなかった。これは、インターネットに接続できる条件をそなえた組織がわずかしか存在しないため、セキュリティに関心が向かなかつたためと考えられる。
- 12) 協調フィルタリングとは、(Collaborative Filtering, CF) は、蓄積された複数のユーザの行動履歴をもとに、嗜好の類似した他のユーザの情報を用いて自動的に推論を行う方法である。Amazon でのリコメンデーションシステムがもっともよく知られた例であるが、他者の行動をデータとして利用する方法であり、利用履歴データベースの存在を前提とした手法である。
- 13) Amazon.com がはじめて単年度黒字を計上したのは 1995 年の創業から 7 年後の 2002 年のことで、この時期にはドットコムバブルはすでに崩壊している。2002.01.29 付けの朝日新聞では「アマゾン・ドット・コムが初黒字 株価 24% 高に。不信の業界に光」と報じられた。
- 14) online : <http://www.cheswick.com/ches/map/index.html> (2009/7/10)
- 15) online : <http://www.opte.org/> (2009/7/10)
- 16) ニコラス・G・カー, 村上 彩 (訳) : 「クラウド化する世界」, 翔泳社 (2008/10/10)
- 17) 広告需要に応じてページ数を臨機応変に増減すればいいようなものだが、新聞というシステム全体はそのようにはできていない。わずか 1 枚、4 ページの増ページでも、記事の作成、版面の構成、印刷帳合、輸送まで、新聞社の抱えているすべての基幹業務を全体的に見直す必要があり、その影響があまりに大きいため、簡単にはページ数を増減することはできない。かつてバブル景気に沸いた 80 年代後半から 90 年代にかけて新聞が増ページに踏み切ったことがあったが、逆に言えばあのような好景気が持続した時期でもなければそのような決断はなされなかったのである。
- 18) Page View は Web サイトの訪問者がアクセスしたページの総計であり、Visit は、そのうち同一であろうと思われる訪問者 (ユニークユーザ) のみを数えたものである。ただし、本文にも述べたとおり、一端セッションが終了した後同一のユーザであることを識別することは困難である。
- 19) 社会構成主義の立場からの narrative approach などこの傾向は顕著である。社会構成主義では現実社会的に構成されるとの立場から、言葉が世界をつくると主張する。
- 20) online : <http://it.nikkei.co.jp/trend/special/itsstyle.aspx?n=MMITba001006072006> (2009/7/10)
- 21) online : http://ja.wikipedia.org/wiki/Web_2.0 (2009/7/10)
- 22) 2006 年 8 月にアメリカの代表的な ISP、ポータル事業者である AOL が、研究目的のために 65 万人の利用者の検索履歴を外部公開した事件。いくつかの報道では流出として報じられているが、実際には AOL が公開して、激しい批判を受けたためにすぐに公開を中止したという経緯がある。AOL は完全な匿名性が保証されているとして、65 万人の顧客の 3 ヶ月にわたるすべての検索履歴 2000 万レコードを公開した。

- しかし New York Times 紙が検証したところ、numb fingers、60 single men、dog that urinates on everything などの検索フレーズから Ms Arnold という女性が AOL の検索者番号 4417749 と判明したと報じ、AOL は最高技術責任者の退任と数名の社員の解雇を実施するなど経営に深刻な影響を与えることとなった。
- 23) H. Maurer : “Report on dangers and opportunities posed by large search engines, particularly Google” 2007/9
online : http://www.iicm.tugraz.at/iicm_papers/dangers_google.pdf (2009/7/10)
- 24) ニコラス・G・カー 前掲書
- 25) ジャン・ノエル・ジャンヌネー : 「Google との闘い—文化の多様性を守るために」、岩波書店 (2007/11)
- 26) たとえば朝日新聞電子版 2010 年 5 月 15 日には、日本を含めた世界各国で、道路沿いの写真をネットで提供する「ストリートビュー」用の撮影車が、無線 LAN を経由して送られた個人データを誤って集めていた、とする記事が掲載されているが、その後各国でこの問題に対し訴訟や行政機関からの指導が相次いだ。<http://www.asahi.com/national/update/0515/TKY201005150179.html> online :
- 27) Cookie は、RFC (Request for Comments) 2965 ドキュメント、“HTTP State Management Mechanism”で定義され正規に認証された Web 技術である。
- 28) Cookie をブラウザに設定する場合、有効なドメインとして、参照している URL と同じものを設定する場合はファーストパーティーCookie、別のドメインを設定する場合にはサードパーティーCookie と呼ぶ。
- 29) 高木浩光 : 「行動ターゲティング広告はどこまで許されるのか」、Nikkei Net IT PLUS , online :
<http://it.nikkei.co.jp/internet/news/index.aspx?n=MMITbe000015102008> (2009/7/10)
- 30) online : <http://www.digitalcenter.org/pdf/2008-Digital-Future-Report-Final-Release.pdf> (2009/7/10)
- 31) online : <http://pc.nikkeibp.co.jp/article/news/20081111/1009612/?f=security> (2009/7/10)
- 32) Tim Berners-Lee, James Hendler and Ora Lassila : “The Semantic Web A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities” , online :
<http://www.scientificamerican.com/article.cfm?id=the-semantic-web>
- 33) そのまま省略すると WOL となるが、すでにこの名前の言語が存在していたため、頭文字を入れ替えて略称としている
- 34) Peter Morville : “Ambient Findability : What We Find Changes Who We Become”, O’Reilly Media, Inc (2005/09)
- 35) David Weinberger : “The Semantic Argument Web” , online :
<http://64.28.79.69/read/swifkick/colum.html?ArticleID=421> (未確認)
- 36) online : <http://www.kanzaki.com/docs/sw/> (2009/7/10)
- 37) online : <http://www.google.com/intl/en/corporate/> (2009/7/10)
- 38) online : http://www.readwriteweb.com/archives/define_web_30_contest_winners.php (2009/7/10)
- 39) Jay Barney : “Gaining and Sustaining Competitive Advantage”, Addison-Wesley (1996/06)
- 40) David Halberstam, 筑紫 哲也, 東郷 茂彦 (訳) : 「メディアの権力」、朝日新聞社 (1999/09)
- 41) 吉本 敏洋 : 「グーグル八分とは何か」、九天社 (2006/12) 吉本によると、グーグル八分 (Google 八分) とは、公正中立とされている Google の検索結果から、特定のホームページが意図的に削除されていることであり、削除が企業や政治家などの「社会権力」からの要請に従い、Google が一方的に行うことが問題であると述べる。
- 42) Europass についてわが国で紹介された例は少なく、国際交流基金 JF 日本語教育スタンダード試行版による解説を以下に示す。2004 年に欧州議会 (European Parliament) と欧州理事会 (European Council) によって議決されたユーロパスは、資格・能力の身分証である。特に、職業領域と教育領域におけるヨーロッ

パの人的流動性を円滑にするために、個人の能力を明確かつ簡潔に示す機能、すなわち欧州各国が独自に持つ資格制度、試験・検定を誰の目にも明らかにする役割を有している。ウェブサイトには5つの資料が記入サンプルとともに掲載されている。これまでの学歴、職歴、業績を示した「ユーロパス：履歴書（Europass CV）」と言語能力を共通の尺度で端的に示す「ユーロパス：言語パスポート（Europass Language Portfolio）」は各人が作成し、留学先や就職先など受け入れ機関に提示するものである。各種資格や検定の補足説明をする「ユーロパス：証明書の補足資料（Europass Certificate Supplement）」、各種卒業資格の補足説明をする「ユーロパス：卒業資格の補足資料（Europass Diploma Supplement）」、他国での教育・職業経験の概要と成果を記録した「ユーロパス：モビリティ証明書（Europass Mobility）」は組織・機関側が作成するものである。「ユーロパス：言語パスポート」は本節で述べてきた ELP の言語パスポート部分をさらに簡略化したものであり、共通参照レベルを用いた自己評価と海外滞在経験を記すことができる。

- 43) この文脈での電子私書箱については、「電子私書箱（仮称）による社会保障サービス等の IT 化に関する検討会【報告書】」に詳しい。online： <http://www.kantei.go.jp/jp/singi/it2/epo-box/index.html>（2009/7/10）
- 44) インターネットでは、この種の役割は CA（Certification Authority）が担っている。CA は、一般的には公開鍵証明書認証局としてデジタル公開鍵証明書を発行する機能を有しているが、これを拡張して個人の履歴事項を証明する機関と位置づけることも可能である。たとえば、大学を卒業したことを証明するのは本人ではなく、当該大学の学長である。ここで大学は CA の役割を果たしている。