

インターネット上の情報検索について

愛知大学法科大学院 伊藤 博文[†]

1. はじめに
2. インターネット上での情報検索とは
 - 2-1. インターネットの普及と情報検索
 - 2-2. 情報検索スタイルの変化
 - 2-3. 検索エンジンの進化
3. 知の偏在から遍在へ
 - 3-1. 集合知
 - 3-2. Web2.0 と情報検索
4. 新しい情報の存在形式
 - 4-1. ブログがもたらすもの
 - 4-2. フィルタリングのない情報発信
 - 4-3. ビジネスモデルとしての情報検索エンジン
 - 4-4. 情報の価値は受け手が決める
5. おわりに

[†] 愛知大学法科大学院教授。本稿と併せて、私の研究用サイト、コンピュータ法学(CaLS)<<http://cals.aichi-u.ac.jp>>をご覧いただき、以下のメールアドレスに忌憚なき意見や批判を送付していただければ幸いです。mailto: hirofumi@lawschool.aichi-u.ac.jp。

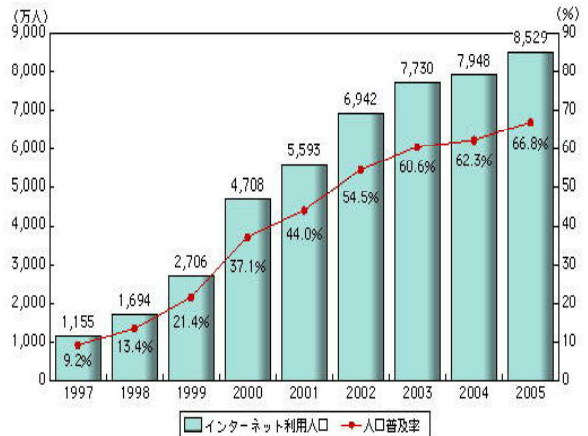
1. はじめに

本稿は、インターネット上の情報検索方法の変遷を辿りながら、情報検索のあり方および情報のあるべき存在形式について検討することを目的とする。インターネット上で繰り広げられるネットワーク技術の進化にはさまざまなものがあり、そこから新しい手法がいくつも生み出されている。たとえば、Webブラウザの進化は、情報を受け取る側のユーザーである閲覧者には、クリックという単純な操作のみを求める操作形式を維持しつつも、情報を提供する側には、文字情報だけでなく画像や動画を組み込んだ電子のパンフレットといった商品提供サイトをビジネスとして展開することを可能とするまで進化してきた。そして今、情報検索の新しいテクノロジーは、情報検索の行い方だけでなく、情報を提供する側にも大きな変革をもたらしている。このような変化を定点から捉えて、その可能性について検討する。

2. インターネット上での情報検索とは

情報検索のやり方は、大きく変化してきている。書籍を中心とした「調査」からインターネット上の「検索」へと大きくシフトしている。これは、パーソナルコンピュータの普及により、これまでの紙媒体から電子のメディアへと情報の存在形式が移っていき、蓄積された電子文字情報がネットワーク化することにより更に、情報検索は、より効率的な手法へと変化してきている。

図1 インターネット利用者数及び人口普及率の動向
総務省通信白書より



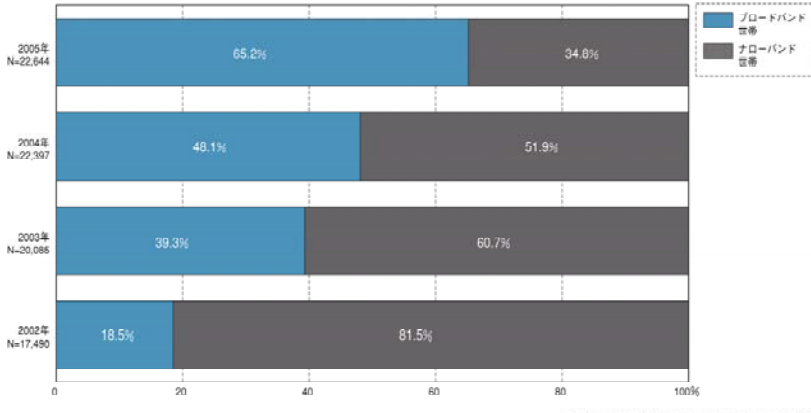
2-1. インターネットの普及と情報検索

インターネットというメディアの普及には、非常に短い間に幅広く一般へ普及・浸透してきたという特徴がある。総務省の情報通信白書によれば、インターネット利用人口は2005年には8,529万人で66.8%の国民がインターネットを利用しているとしている(上記図1参照)。また、インターネットへの常時接続を前提とするブロードバンドの普及(後掲図2参照)は、インターネットユーザーの裾野を広げ、インターネットそのものの使い方に変化をもたらしている。それまで、ダイヤルアップ方式でメールの送受信程度の文字情報のやりとりにしか使ってこなかったインターネットが、音声・画像・動画を常時扱える情報メディアに変化したのである。そして、この普及が、インターネットは日常生活に欠

図2 ブロードバンドの普及率

インターネット白書 2005 より

資料1-1-4 インターネット利用世帯におけるブロードバンドとナローバンドの構成比 (2002年-2005年比較)



インターネット白書2005 ©Access Media/impress,2005

くことのできないメディアとしての不動の地位を獲得させた。溢れんばかりの情報が満載されたインターネットに日々アクセスしなければ日常生活がおくれない社会環境が始まったのである。

この変化につれて、情報検索というあり方が大きく変わったのである。その変化には、インターネット上の情報量(コンテンツ)の爆発と全文検索型検索エンジンの出現が大きなターニングポイントとなっている。はじめは、情報量の爆発からである。インターネットの普及がユーザーの自由な情報発信を可能にしてきた。これにより、インターネット上に存在する情報は既に人間の手作業の力では調べ尽くせない程の量をなすこととなり、その量は今後も増大する一方である。この結果として、情報氾濫の中でユーザーに求められるのは、情報のジャングルから必要な情報だけを選び出す能力である。このニーズが検索エンジンを出現させ、その進化を後押ししているのである。

2-2. 情報検索スタイルの変化

ここで、情報検索のやり方がどのように変化してきたかを私の経験から述べてみたい。

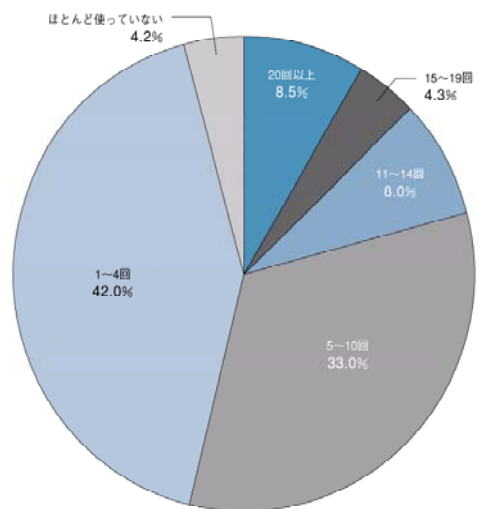
かつて、未だインターネットでの情報検索が一般化していないとき、私の専門とする法律学における法情報調査において、インターネットを使った検索をすることは極まれなことであり、当時は情報検索

という考えもなく「調査」と呼んで紙媒体の書籍などを拾い集めることがもっぱらであった。よって、インターネットで情報検索をするということには何となく抵抗感があった。この抵抗感をもたらせていた一因は、「楽をしている」という後ろめたさである。そし

図3 検索サービスの利用頻度

インターネット白書 2005 より

資料2-4-9 検索サービスの利用頻度 N=1,592



インターネット白書2005 ©impress,2005

でもう一つは、これだけでは完璧な情報検索ができていないという恐怖感であった。

まず前者の後ろめたさである。私が法律学研究の道に入った 1980 年代初頭には当然のこととしてインターネットなど使い物になるものではなかった。法情報を調べる、たとえば論文を書こうと思って必要な文献を集めようとすれば、まずは図書館に出かけ、法律文献の一覧を掲載している法律雑誌などを頼りに、まさに図書館内を這い蹲るようにして、館内階段を上り下りしながら資料を集めた。もちろん、大学図書館で事が足りることはなく、他大学の図書館やはたまた国立国会図書館まで資料を求めたこともある。このような手工業的情報検索を体験してきた者にとって、クリック一つで世界中の図書館情報が得られるというインターネットの効用には、驚愕したものである。しかし、当初はすばらしさに驚くが徐々に自分が手抜きをしているような気にさいなまれることとなる。それは、情報検索は足で稼ぐものであり、汗の対価として得られるものであることが体に染みついてきたからである。この感覚が変わるのには時間を要した。

次に後者の恐怖感である。インターネット上に公開されていた法情報は、インターネットの黎明期にはとてもお粗末なものであった。たとえば日本国憲法の条文一つ見つける

にも苦勞し、無いとわかれば、自分で六法を見ながらキーボードで打ち込み、それをサーバーで公開していたのが実情であった。判例をインターネット上で自在に検索ができるようになるには、クリアすべき問題点が多すぎ、相当の時間がかかると考えていた。インターネット上で調べ得る法情報には限りがあり、とても網羅的な資料集とは呼べなかった。しかし、インターネットの普及とともに、政府、官公庁、裁判所などがインターネット上で情報を公開し出すと状況が一変し、欲しい情報がインターネット上で入手できるようになってきた。情報が紙に印刷されたものから電子文字へとシフトするにつれ、状況は逆転し、図書館に行ってもインターネット上にある情報すべてが入手できない状況になってきた。

そして今や、法情報検索を行うとき、図書館に出かける前に、まずはインターネット上の検索エンジン Google で検索をかける。さらに OPAC で図書館の収蔵状況を確認し、そこを突破口として法情報収集の第一段階を行うのである。その結果を踏まえ、図書館に出向き必要な資料を入手するというスタイルに変わった。更にここ数年で、電子図書館などの電子資料が充実してくると、もはや図書館に出向くことは珍しいこととなり、インターネット上で、つまりコンピュータの前に座ったままで、法情報調査が完結してしまうことが多くなってきたのである。¹⁾ このような変

1) インターネット上で調べられるものが法情報のすべてと言い切ることにはできない状況に変わりはない。歴史的な文書や極めて利用頻度の低い資料などは、インターネット上で公開されておらず、図書館の片隅で眠っている状況に変化はない。この時も、やはりインターネット上の法情報検索では完結しないという恐怖感は払拭できず、完璧を求めれば、やはり図書館の暗い書庫へと行かざるを得ないのである。

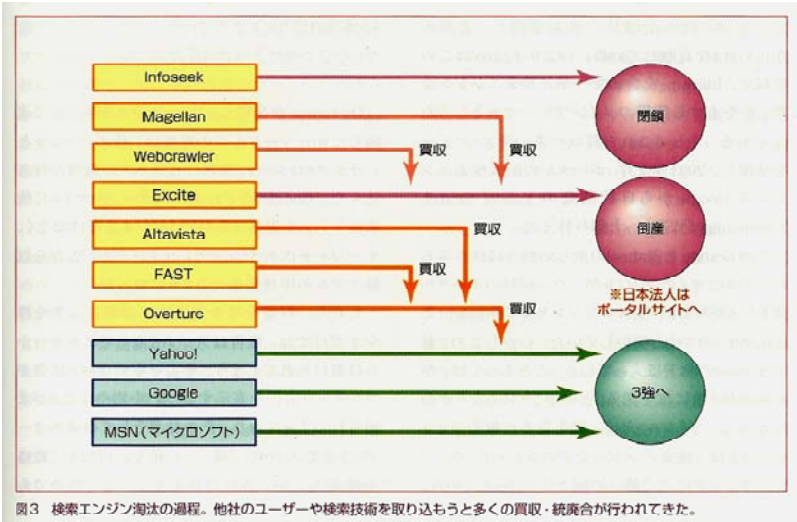
化をもたらしたものが、Google に代表される検索エンジンである。次項でその変遷過程を概観してみる。

2-3. 検索エンジンの進化

今この論文を書いている 2006 年 7 月において、インターネットの検索エンジンといえば、Google、Yahoo!、MSN の 3 大検索エンジンが挙げられる。この 3 社が生き残るまでの過程には数多くの検索エンジンが生まれ消えていったのである。

検索エンジンの歴史は、1994 年の Yahoo! の出現に始まる。Yahoo! は、スタンフォード大学の 2 人の大学院生が作った Web ディレクトリがその発祥であり、初めてのインターネットビジネスの成功例となった。Yahoo! の検索エンジンは当初よりディレクトリ型の検索エンジンであり、当時はインターネット上の情報量も少ないこともあり、スタッフの手作業によるデータベース更新によるものであった。その後、1994 年に始まった Infoseek が端緒となった、機械的に情報を集めるロボット型の検索エンジンが多数出現するが、どれもこれもユーザーにとって新たな変革をもたらすものではなかった。これに一大転機をもたらしたのが Google である。Google は Yahoo! に挑戦するという形で事業展開を行い、独自技術を開発・駆使し、短期間にユーザーの絶

図 4 検索エンジンの盛衰
『Web2.0 への道』インプレス社(2006 年)83 頁より転載。



対的な支持を獲得してしまった。この流れに対抗すべく Yahoo! は、協調路線から対立路線へと戦略を変えながら検索サービスの質的向上を目指しインターネット上の情報検索ポータルサイトの巨人としての地位を守り続けてきたのである。この両者の拮抗を眺めながら、独自のコンピュータ技術で検索エンジンを開発してきたのが、MSN を運営する Microsoft 社である。Microsoft 社は、インターネットでの事業展開が後手後手に回り、この分野での対応が遅れていたが、巨大な資本力をバックに技術開発をすすめ、3 大勢力の一角を形成するまでになっている。

Google に代表されるような全文型検索エンジンの普及は、インターネット上での検索方法を大きく変えた。ショッピング、レジャーの行き先など、何を調べるにも、まずはインターネットで検索し、次の行動を決めると

2) 検索エンジン市場の革命児は、なんと言っても Google であり、その会社の経営方針から生まれる高度な技術開発力からして他社を圧倒している。Google が Web2.0 世界の会社であり、他 2 社が Web1.0 世界の会社と区分けされる所以である。

いうライフスタイルまでも定着させたのである。

3. 知の偏在から遍在へ

検索エンジンがもたらした効用を説明するにあたり、集合知について説明する。

知は情報の集合体である。個別に存在する一定量の情報を一カ所に集積し、有機的に結びつけることにより高度な学問分野を形成することができる。このような集合体をどこに形成するかは様々な要因が働く。学問の最高学府である大学がその集積地である所以は、そのような情報や情報を有機的に結びつけることのできる場や人材を集約できる環境であったからである。そして、そこで知の再生産が行われ、学問が開花し、文化が創造される。しかし、この大学という場を考えても、知の集積地は、誰でもが自由にアクセスできる場ではなかった。また特殊な産業技術を開発する企業の研究所といった知の集積場に一般人がアクセスすることは更に困難であった。

つまり、知は偏在するのである。よって特定の地点において、高度に集約すればするほど知は質の高いものへと変化していくのである。確かに、これがネットワーク社会以前の定説であった。知を求めれば、知のある場所へ出向かなければ知は獲得できない。知の育つ環境にいなければ知の成長はありえなかつ

た。

知が遍在するとは、一地域に偏って集積していた知が、世界中のどこからでも共有できる可能性を持つことを意味する。インターネットはそれを可能にしてくれる。インターネット上に人類の英知が隈無く公開され共有されれば、地球規模で人類がその知を共有できる。今まさに、インターネットはこれまでの定説を覆しつつある。知は遍在することが可能となった。地球規模の知を集約できるプラットフォームが完成したのである。

3-1. 集合知

知の遍在がもたらしてくれるものが、集合知(Collective Intelligence³⁾)である。インターネット上の検索エンジンがもたらした功績の一つは、誰もがインターネットという集合知にアクセスするというライフスタイルを生み出したことである。これまで歴史上存在し得なかった知の集合場所にはインターネットを介して誰でもアクセスすることが可能となっている。

ここで集合知について考えてみる。James Surowiecki⁴⁾によれば、集合知とは、「適切な状況の下では、人々の集団は、その中でも最も優れた個人よりも優れた判断を下すことができる」とするものであり、その「適切な状況」には次の4つの要素を持つとする。

3) The Wisdom of Crowds <http://www.h-yamaguchi.net/2004/06/the_wisdom_of_c.html>(last visited July 4,2006)

4) James Surowiecki “The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations” Doubleday (2004); ジェームズ・スロウィツキー著・小高尚子訳『「みんなの意見」は案外正しい』角川書店(2006年)28頁参照。

(1) 意見の多様性

集合要素である個が多様であること。既知の事実の特異な解釈であっても、個が独自の私的情報を多少なりとも持っている状態であること。

(2) 独立性

他者からの影響に左右されず個が互いに独立していること。

(3) 分散性

個が分散しており、個自身が身近な情報に特化し、それを利用できる。

(4) 集約性

個の要素を集約する優れたメカニズムが存在し、個を集団として一つの判断に集約することができる。

この4原則は、インターネットを形作る基本原理である、自律・分散・協調に呼応するものである。「(4)集約性」を実現する手段が検索エンジンということとなる。

インターネット上の検索エンジンを使った情報検索の前提は、集合知を肯定する立場となる。一見無秩序に各人が個別に情報発信をしているが、それを集積統合して集合知とし、自在にアクセスでき、網羅的に検索ができるシステム、それがインターネット上の検索である。

3-2. Web2.0 と情報検索

最近のインターネット技術の進化を表現する言葉に、Web2.0 がある。Web2.0 は、アメリカの ICT 関連の出版社 O'Reilly Media 社の

CEO、Tim O'Reilly 氏が提唱した用語であり、一つのコンセプトでもある。進化を続けるインターネット上の Web を中心とする技術開発業界の 2004 年以降のトレンドを捉えて命名したものであり、特定の技術を指すものではない。Web2.0 は、それまでの技術 (Web1.0 と位置づける) とは異なる新しい Web 技術やサービス開発姿勢の総称である⁵⁾。

Web2.0 のコンセプトを形作る「7つのWeb2.0の原則」がある。

- (1) Web がプラットフォームとして振舞う
- (2) 集合知を利用する
- (3) データは次世代の「インテル・インサイド」～ サービスの「核」となるデータを持つことで、支配的な立場を得る
- (4) ソフトウェア・リリースサイクルの終焉
- (5) 軽量なユーザーインターフェース、軽量な開発モデル、そして軽量なビジネスモデルを採用する。
- (6) 単一デバイスの枠を超えたソフトウェアを提供する。
- (7) リッチなユーザー経験

このコンセプトが、実際の Web 技術開発や Web を使ったサービスに適用されるのである。その時は、「パッケージソフトウェアではなく、費用効率が高く、拡張性のあるサービスを提供し、独自性があり、同じものを作ることが難しいデータソースをコントロールする。このデータソースは利用者が増える

5) 平成 18 年版 情報通信白書<<http://www.johotsusintokei.soumu.go.jp/whitepaper/ja/h18/index.html>>(last visited July 4,2006)

ほど、充実していくものでなければならない。ユーザーを信頼し、共同開発者として扱う。カスタマーセルフサービスを通して、「ロングテールを取り込む」といった方向性が示されている。この Web2.0 的な検索エンジンが Google であり、Google 社はまさに Web2.0 的な会社と位置づけられている。

4. 新しい情報の存在形式

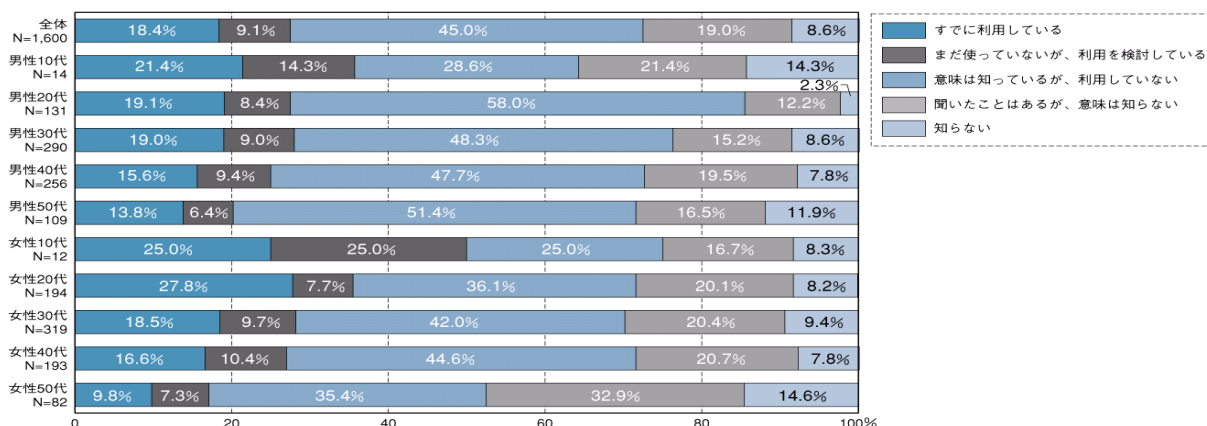
Web2.0 の時代において、インターネット上の情報量の増大と共に、情報の存在形式自体にも新しい変化がみられる。

4-1. ブログがもたらすもの

ブログは急速に普及し、インターネットの個々のユーザーが自由に情報発信を始めている。これは何を意味するか検討してみたい。

まず、情報の発信源が格段に増えることがもたらす情報ソースの多様性である。たとえば世界の或る地域で地域紛争が起きたとすれば、これまでの新聞、テレビというメディアの特派員が発する情報が情報源となり、AP 電や CNN というマスメディア組織を通じて世界に情報が発信される。ここでの情報に偏りや特徴があるか否かは記者の視点で決まる。しかし紛争地域に生きている人々がブログで日々状況を伝えるようになれば当然のこととして、情報源が多様化し情報の偏りが少なくなる。つまり報道されないことがブログで知り得ることとなる。ここで問題となるのは、ブログはそのような場合には効用があるが、果たしてそれを知るためにはどのような仕組みが必要となるのかという点である。つまりインターネットユーザーとなり得る全世界 60 億人が自分のブログで日々情報を提供し出したとしたら、その中から必要な情報をどうやって見つけるのか、その情報をどのように管理するかである。これの答えが RSS

図5 ブログの普及 インターネット白書 2005 より
資料2-4-33 ブログの認知度 [全体と年代別と性別]



インターネット白書2005 ©impress,2005

6) 梅田望夫『ウェブ進化論 本当の大変化はこれから始まる』ちくま新書 582(2006年)。

のような検索エンジンであり、ブログという集合知に対する唯一のアクセス手段となる。

4-2. フィルタリングのない情報発信

情報を公開する場合には一定度の節度が求められる。脳裏に浮かび思いついたことをすべて口に出すのには問題がある。たとえば名誉毀損罪や侮辱罪はそのための法制度⁸⁾である。名誉毀損が我々の社会で犯罪(もしくは民事上の責任)とされる前提には以下の手続きが前提となっている。(a)人は様々なことを思いつく。これは思想信条の自由により担保される。(b)そして人は自由に発言ができる。これは表現の自由により担保される。しかし思ったことを自由に発言する自由には制約が課される。よって(a)から(b)への移行段階でフィルタリングが働く。それは名誉毀損といった犯罪(もしくは損害賠償責任)への配慮、人間関係上の不要な摩擦を避けるなどの配慮がなされ、自制が行われる。この場合通常、人間の行動は、脳に浮かんだ発言内容を紙に書き出す、文章にするなどの一定度の作業を行い時間的経過が必ず存在し、クーリングオフする時間がある。このフィルタリング作業が情報の質的向上をもたらすのである。まさに下書きと清書の違いがある。フィルタリングの無い情報発信は、いわば下書きやメモ書き程度の情報が公開されることになる。

この懸念はまさにブログ社会で始まりつつ

ある。情報の質的低下、不要なコミュニケーション上の軋轢などが生じてしまう。この問題は、ネットワーク社会に依存すればするほど深刻化する問題である。

4-3. ビジネスモデルとしての情報検索エンジン

情報検索の革命をもたらした検索エンジンは、一つのビジネスモデルを構成していることに注目する必要がある。インターネット上の技術はどのような形で使われるにせよ、商業ベースで使われることを念頭においている点は注目に値する。検索エンジンの行っていることは、情報を集めその情報を欲しい人に配信するということであるが、ここにビジネスが介在することにより競争原理が働く。

たとえば、検索エンジンの革命をもたらした Google 社は、はじめはシリコンバレーにいくつも存在するベンチャービジネスの一つであった。Google 社の素晴らしいところは、技術もそうだがビジネスとしての戦略を誤ってこなかった点にある。成功した多くの ICT 企業は当然のこととしてビジネス界で生き残ってきたのであるから、ビジネス戦略において正しい選択を行ってきたことは自明である。如何に優れた技術を持とうともそれを使う場がなければ技術は生かされない。Google は、検索業界の巨人 Yahoo!に挑み、それを凌駕することに成功している。Google

7) RSS とは、RDF Site Summary の略称で、Web 記述用言語の XML 技術を活用して、ブログ等変化の激しい Web ページの情報を一元的に集約できるデータ記述規格である。〈<http://ja.wikipedia.org/wiki/RSS>〉 (last visited July 4,2006)

8) 名誉毀損に対しては、刑事責任として刑法 230 条、民事責任として民法 723 条が規定しており、侮辱罪に対しては、刑法 231 条が規定している。

自体が Web2.0 的な流れを生み出してきており、Yahoo!を時代遅れなものにしてしまったのも事実である。情報検索業とは一体何をしている業界なのかといえば、インターネット上で検索エンジンサービスを提供しているに過ぎない。新しい情報を生み出しているのかと言えばそうではない。今ある情報を集めそれを欲しい人に提供するという基本的な枠組みの中で行っているに過ぎない。そこで行われているのは、自ら情報を集めてきて情報を作り出しているわけではない。インターネット上に公開されている Web 上の膨大な情報を集めて集約し、それを改編整理して欲しい人に欲しい形式で提供しているに過ぎない。Google 自体が、今世界でどのような事件が起きているのかといった情報を生み出しているのではない。このような副次的な情報産業に富が集中するというビジネスモデルを生み出した点が注目に値する。

4-4. 情報の価値は受け手が決める

或る情報に価値が有るのか無いのか、また有るとすればどれだけの価値があるのか。これは情報を扱うときに大きく考慮される点である。情報に対しては、情報発信の段階で取捨選択が行われる。従来型の情報メディアでは、限られた提供媒体の容積中に入る限りの情報だけしか情報を伝達することができなかった。たとえば日刊新聞は紙面に制限があり、伝えられる文字情報・画像情報は有限であった。よって情報の取捨選択が行われ日の目をみない情報が数多く存在した。しかし、インターネット上の情報発信に量的制限はない。

誰でも好きなだけ情報を発信できる。まさにブログ社会のもたらすものである。このような情報発信が可能な環境では、情報の価値は受け手が決めるのである。

よって、インターネット上の情報のあり方は、このような情報を公開してはいけないとか公開する意味は乏しいという発想ではなく、まずは発信する。そして集合知を形作る。そしてその情報の価値は受信者が選択し、必要な情報だけを手に入る。このような情報の存在形式が求められるのである。

5. おわりに

以上のように、インターネット上での情報のあり方及びその問題点について指摘してきたが、ここまでの議論はすべて地球規模で均一のインターネット環境が誰にでも提供されるということを前提としているものである。しかしながら、現実とは異なる。地球規模のインターネットは人類に福音をもたらす可能性のあるものではあるが、それが地球規模で福音をもたらしているのかといえば、そうではない。発展途上国では、未だ食料・住居・電気の安定供給すらままならぬ国情で、どうやってコンピュータを手に入れインターネットにアクセスできるのか悩ましい問題である。インターネットの普及云々を論じている段階の次に考えなければならないのは、デジタル・デバイドの問題である。富める国は知の遍在によりますます富むこととなり、一方で発展途上国との国富差は開く一方である。

インターネットという人類の叡智を誰もが

共用できる環境が整いつつある一方で、知の「偏在」へと再び回帰してしまうような現実がデジタル・ディバイドである。Web2.0に代表されるような新しいトレンドの追求も大事であるが、一方で地域間格差がますます広がっている現実にも視点を向ける必要がある。このような視座をもった新しいネットワーク技術が開発されることが望まれるのである。