

# 検索エンジンがもたらすもの

## Influence of Google Search Engine

永井 武\*

### 1. はじめに

日本で商用インターネットサービスが始まったのは1993年であるから、今年で15年経過した。その間、インターネットは電子商取引、ビジネスメールと並んで300億ページ<sup>(1)</sup>ともいわれるwebコンテンツの検索、閲覧に利用されている。インターネットのアプリケーションとしてwebが出現したのは1990年であり、webコンテンツは年々増加した。

1994年に、当時スタンフォード大学院生であったジェリー・ヤンとデイビッド・ファイロの2人が、興味深いwebコンテンツの索引を人海戦術で作成、ABC順に並べた。後にディレクトリ型の検索エンジンをつけて無料で公開した。当時webコンテンツは5000万ページにも満たなかったが、情報検索は便利だったので、多くのユーザが利用した。1995年、ジェリー・ヤンとデイビッド・ファイロは、ヤフー社を設立した。ポータルサイトと呼ばれるトップページ、およびその次の階層の余白に広告の sponsor を募り、バナー広告を載せ収益を拡大した。米国ヤフーの2005年度売上げは50億ドル強である。

現在のヤフーのサービスは、トップページに買う、知る、楽しむなど6分野に分類された約60個のキーワード、およびトップページの下方にさらに数10個のキーワードを掲載し、合計100個以上のキーワードを並べ、そのうちの1つをクリックすると検索の欄を含む次の階層を表示することから始まる。そこからユーザが必要な情報のキーワードを入力して検索することにより、検索結果を得る。検索結果は現在改良されているが、サービス開始当初は無秩序に表示されるので、ないよりは便利ではあるが、ユーザはコンテンツを読んで、必要な情報であるか否かを判断する必要があり、使いにくかった。ユーザは必要なwebサイトのURL (Uniform Resource Locator) をメモしたり、お気に入りチェックを入れたりして、必要な情報を確保していた。グーグルがサービスを開始する前までは、このようなweb検索が行われていた。

1998年、スタンフォード大学院生であったラリー・ページとサーゲイ・ブリンがグーグルという新しい検索エンジンを開発し、インターネット上に無料で公開した<sup>(2)</sup>。グーグル検索の最大の特徴はユーザが入力したキーワードの検索結果が無秩序に出力されるのではなく、ユーザにとって最も有益なものを1番目に、2番目に有効なものを2番目に出力するというように、ユーザにとって有難い機能をもっている。世界中にある300億ページのwebコンテンツのなかから、ユーザが必要とするあらゆる分野の最も有効な情報を瞬時に検索し出力してくれる。これが日本を除く世界中で最も使用される検索エンジンとなった。数多く輩出したネット企業の中でヤフー、アマゾン抜き、2007年には時価総額24兆円となり、最も成功した企業である。本報告はグーグル検索が使用する技術、ユーザに情報サービスを無料で提供しスポンサから広告料を徴収するビジネスモデル、グーグル検索が人間社会にもたらすものについて報告する。

## 2. グーグルが使用する技術

### 2.1 グーグル検索エンジン出現の前

1997年まで、DEC社のアルタビスタ (AltaVista) が最も多くの人に利用された検索エンジンである。DECアルファを搭載したワークステーション (WS) に、当時 web に存在したコンテンツのインデックスを自動的に作成してサーバとし、ユーザに提供していた。しかし、DEC は本業のミニコンが不振で経営難となり、コンパック、続いて HP に買収され、その技術は現在ヤフーで利用されている。

1998年に最も多くの人々が利用した検索エンジンはヤフー (Yahoo) とエキサイト (Excite) である。エキサイトを起業したのはスタンフォード大学の6人の同窓生である。エキサイトはマイエキサイトという自分だけのトップページにビジネス情報、ニュース、無料メール、ユーザが住む地方の天気予報などに加え、検索の窓をつけユーザに提供し利用者を集めた。検索サーバはインデックス検索と全文検索の機能をもつ。

1999年に最も多くの人々が利用した検索エンジンは、ライコス (Lycos) である。ライコスを開発したのは東部のカーネギーメロン大学である。検索結果に要約および検索結果にリンクを張りアンカーテキストにして、そこをクリックすることで情報を提供することにより多くのユーザを獲得した。

アルタビスタ、エキサイト、ライコス、インフォseek (Infoseek) などの検索エンジンが利用されたが、検索に使用したキーワードを多く含む web サイトを表示することしかできなかった。その理由は、いかがわしいサイトが、車、カメラ、携帯電話、PC、ソフトウェアなど多くのユーザが検索するキーワードをいかがわしいサイトに多数入れることにより、検索エンジンは検索結果の上位にいかがわしい情報を上位にもたらし、ユーザは下位に表示されている必要な情報を探さねばならなかった。

### 2.2 バックリンク (3)

1995年、スタンフォード大学コンピュータサイエンス専攻の大学院生であったラリー・ページとサーゲイ・ブリンは伝統的な学問の世界で研究を行い、金持ちになるのではなく博士号取得を目指していた。博士論文は1つのテーマで学術誌に4,5編の論文を書く必要があるので、テーマ探しは大切でいろいろ模索した。

いくつかの論文を書いたテーマはデータマイニングの研究である。例えば、スーパーなどの店にいくお客がどのような組み合わせで商品を買うのかを分析し、その結果に基づいて商品を陳列するのに利用する研究である。有名な例をあげると、あるスーパーで売上げ額の大きいレシートを分析すると紙おむつとビールを組み合わせで買うお客であることがわかり、分析結果に基づいて紙おむつとビールの陳列棚を近づけたところ、目に見えて両商品の売上げが増加した。このデータマイニングの方法を秩序のない web コンテンツの世界に適用できないかを考えるに至った。

学術論文は多くの学術論文を参考にして書く。すぐれた学術論文は、多くの学術論文に引用され学会に影響を与えている。web コンテンツも他の web コンテンツからリンクを張られている web コンテンツがいい情報源といえると、ページとブリンは思いついた。フリー百科辞典ウィキペディア (Wikipedia) を利用したことがあるであろう。ウィキペディアに書かれている文章の中にブルーのアンカertextがある。ここをクリックすると、この言葉についての詳しい説明のあるサイトの情報がユーザの PC に表示され、これを読むことができる。これはウィキペディアの文章からその文章について記述されているサイトにリンクが張られているためである。ウィキペディアからリンクを張った先のサイトはわかる。しかし、ウィキペディアなどに張っているリンク元は、1995年時点ではわか

らなかった。ページとプリン、データマイニングの方法を秩序のない web コンテンツの世界に適用するためには、このリンク元を知ることが手がかりになるのではないかと推測した。そして、リンク元を知ることがバックリンクと呼んだ。学术论文の大きな評価基準は引用された数であるのと同様に、web コンテンツの価値はバックリンクの数が使えるのではないかと考えた。

ページとプリン、試行錯誤の末、世界中の web サイトをダウンロードしてリンクを調べる以外バックリンクを知る方法はないと決め、それを実行することにした。世界中に存在する web をページとプリンのサーバにダウンロードし、そこに張られているリンクを調べる。ただ数を勘定するのではなく、リンクを張ったサイトがどのくらいのリンクを張られたリンクであるかも勘定する。検索されたキーワードについてのサイトがもつ情報の価値を、多くのまじめなユーザがリンクを張るサイトからのランクは高く、そうでないサイトはむしろペナルティの意味で低いランクにするアルゴリズムを考案し、実装した。それが終了するとその web コンテンツはページとプリンのコンピュータから消去する。これを実行するためには高性能のコンピュータが必要である。ページとプリン、研究費が少なかったため、当時大学で使用されなくなった古いコンピュータを集めてこれをイーサネットに接続し、個々のコンピュータの性能をトータルであたかも高性能のコンピュータとして動作するクラスタ技術を利用した<sup>(4)</sup>。これは、NASA のドナルド・ベッカが開発したベオルフシステムである。ベオルフシステムについては3章で述べる。

### 2.3 ページランク<sup>(3)</sup>

1996年、古いコンピュータをクラスタ化したシステムで web コンテンツをそっくりダウンロードしてバックリンクをつきとめる研究が行われた。自動プログラムを開発し、それを使用してバックリンクの数を測定すると、張られているリンクが多いリンクと少ないリンクがあることもわかった。したがって、張られているリンクが多いリンクからのリンクと、少ないリンクから張られているリンクは重みを加味して評価する方が、web ユーザの評価に近い評価方法となることにも気がつき、これも web コンテンツ評価プログラムに盛り込み、ページランクと命名した。一般ユーザからのリンクより、重要なサイトからのリンクに重みをつける機能をつけた。このようにして、1つのキーワードに対してのページランクをインデックスサーバに収めた。ページランクされている web のコンテンツをドキュメントサーバに収録した。

1997年、ページランクと従来からある検索プログラムを合体させてスタンフォード大学内で試運用を開始した。使用したコンピュータシステムは、3.1で述べるベオルフシステム、および3.2で述べる Linux を OS として使用した。その結果、学内でユーザが求める情報を的確に、迅速に表示してくれるという評価とスタンフォード大学内ユーザが口コミで広めたおかげで、またたく間にスタンフォード大学内に広まった。そのころにこの検索システムに Google の名がつけられた。

1997年には、検索エンジン技術はほぼ完成し、あとはデータベースマシンの能力を格段に上げ、ネットワークの帯域を確保すれば、世界中のユーザにサービスを提供できる目処がついた。マシンとネットワークの帯域の確保は大学院の学生にとっては巨額の資金が必要なため、この技術をライセンスするか売却して、ラリーとプリンは博士論文執筆に戻ろうとした。

この時期、前に述べたようにアルタビスタ、エキサイト、ライコス、インフォシーク (Infoseek) などの検索エンジンが利用されていたが、検索結果は2.1で述べたようにいかがわしいサイトが上位に出力され、インターネット上の検索はビジネスにならない、ベンチャキャピタルが投資しないという時期だったので、検索技術に投資する企業は見つからなかった。

結局、ラリーとプリンは自分達で検索をサービスするベンチャを立ち上げることになったのである。

### 3. クラスタ技術 (ベオルフシステム) <sup>(4)</sup>

#### 3.1 高性能コンピュータ

グーグルが提供する検索エンジンは、300億ページの web から情報を集め、インデックスをつけ、かつそのサイトの持つ情報の価値を、検索されたキーワードに関して最も有益な順にサイトをランク付けし、検索結果を短時間内にアップするために、桁違いに高性能なコンピュータを必要とする。

1965年から、ムーアの法則 <sup>(5)</sup> が続いている。半導体集積回路の集積密度が1年半毎に2倍になることが、約40年続いているのである。今日、科学技術の進歩が早いと感じられるのは、根底にこのムーアの法則が続いているためである。ムーアの法則の結果に最も恩恵を受けているのは、コンピュータである。PCのCPU (Central Processing Unit) は毎年インテルが億の単位で全世界に供給している。一方、大型コンピュータは、世界のシェアトップのIBMでさえ年に1000台は受注していない。計算能力に特化したスーパーコンピュータはさらに受注台数が少ない。その結果、当然高性能コンピュータのコストパフォーマンスは、PCに比較して低い。

1994年、NASAで高性能コンピュータを必要としていたが予算はない。そこで、ドナルド・ベッカはコストパフォーマンスの高いPCを多数ネットワーク接続し、あたかも1台の高性能コンピュータとして動作するコンピュータの開発を始めた。正式な名称はクラスタ技術であるが、IBMなど大型コンピュータ業界がユーザに暴力的価格を要求するので、ユーザを暴力から解放する意味を込めて、暴虐なグレンデルから民衆を解放した英雄ベオルフの名に因んで、プロジェクト名をベオルフとして開発を開始した。

1996年、ベオルフ1号機は完成した。CPUはインテルDX4 (100MHz) 16個を用い、10Mbpsのイーサネットハブで接続し、74MFLOPS (浮動小数点数演算を1秒当たり7400万回演算) を達成した。これは当時、同等の演算速度をもつスーパーコンピュータの価格の1/10で実現した。

2.2で述べたように、ページとプリンがバックリンクの研究を開始し、高性能コンピュータを必要としたが、予算がたりないそのときにこのベオルフシステムが開発されたのである。ページとプリンは大学内で使われなくなった古いPCをイーサネットで接続し、少し高性能なコンピュータを組み立てたのである。

#### 3.2 クラスタコンピュータのOS (Operating System)

PCにWindowsやMacOSのようなOS (基本ソフト) が必要であるのと同様、クラスタコンピュータにもOSは必要である。安価に入手可能で、クラスタコンピュータ用に修正可能な既存のOSとしては、Unix, FreeBSD, Linux <sup>(6)</sup> <sup>(7)</sup> <sup>(8)</sup> に絞られる。1990年来、多くのハッカ達が作り上げたLinuxは処理能力が高く、柔軟性があり、拡張性を有し、堅牢である、という理由でベオルフシステムのOSとして採用され、改良して使用した。

ページとプリンが古いPCを接続して作成したクラスタコンピュータのOSも、2人がオープンソースのLinuxを改良して使用した。ページとプリンはハードウェア、ソフトウェア両面にわたって一流の技術を持ちそれを応用した。

このようにしてページとプリンが組立てたクラスタコンピュータは安価で、拡張性に富み、堅牢なシステムとなった。全体として1台の高性能コンピュータとして動作するが、構成する1台のPCが故障しても全体は停止することなく、十分に機能するという特徴をもつ。現在は、40 - 80台のPC

を1つのラックに収め、2台程度のHDDをつけ、10 - 100のラックを集めたクラスタを世界に分散配置している。世界全体のPCの合計は45万台<sup>(9)</sup>といわれる。

### 3.3 検索サービス会社グーグルの設立

2.3で述べたようにページとプリンが開発した検索エンジンは、これまでの検索エンジンより優れていることをヤフー、AOL、ネットスケープ、マイクロソフトなどのポータルサイトに売り込んだが、検索サービスは機能しないサービスであると思われた時代である。スタンフォード大学の先輩で最も頼りにしていたヤフーのジェリー・ヤンとデイビッド・ファイロへ売込みに行ったとき、そんなに優れた検索エンジンなら自分達でサービスしてはどうか、といわれページとプリンに検索エンジン売却の最後の望みを失った。

スタンフォード大学院の指導教授であるデイビッド・チェリトンに大学院を休学することを相談し、さらに先生知合いのベンチャキャピタルの紹介もお願いした。最初に会ってくれたのはアンディ・ベクトルシャイムである。ベクトルシャイムはページとプリンの検索エンジンの説明を聞くと、その長所を理解し、すぐに10万ドルを出資した。因みにアンディ・ベクトルシャイムはかつてサンマイクロシステムズ社のWS(Work Station)のCPUを開発した人である。サンWSのCPUはスパークチップと呼ばれ、最近20年間WSの世界のトップシェアを持つヒット商品である。その後資金集めは順調に進み、KPCB、セコシアなどの大口や多くの小口合わせて1億ドルの資金を集めた。ページとプリンに資金を節約しながら人を雇い、コンピュータ部品を購入し、自分達で組立てた。1999年終わりに従業員は39人、事務所を借りて、検索エンジンの能力拡大に努めた。しかし、収入らしい収入はなく経費は月に数万ドルかかっていた。

### 3.4 アドワーズ広告

2.3で述べたように1998年頃は、検索結果の上位にいかがわしいサイトで占められていたので、インターネット広告は意義を失いつつあった。これを打ち破る試みがなされていた。

1998年2月、ゴートウドットコム社のビル・グロスがTED (Technology Entertainment Design) 会議において、検索者のキーワードに関連して広告を検索結果に掲載する方法を提案した。これまでの検索キーワードの数が多いサイトをデータベース化する検索エンジンは使用せず、新たにキーワードに関する情報を集めデータベース化し、これに検索エンジンをつけた。ゴートウドットコム社の検索サイトだけでは、ユーザのアクセス数は少ないので、マイクロソフト、AOL、ネットスケープなどに検索窓をおき1クリックあたり10セントを支払ってゴートウドットコム社の検索結果を表示する、というものである。TEDの聴衆の反応はよくなかったが、実際にゴートウドットコム社が広告をとり始めると、広告主は徐々に増加した。理由は、いかがわしい広告を排除できたこと、テレビや新聞に広告を出せない小さい広告主はネットの広告を利用したいなどである。ビル・グロスの方式は、ペイドサーチ(検索連動型広告)<sup>(10)</sup>と名づけられ、ゴートウドットコム社はオーバチュア社と名称を変えた。

2001年ビル・グロスはグーグル社に協力してビジネスをする提案をしたが、グーグルは検索結果を広告料に左右される提案には興味を示さなかった。そうはいてもグーグルは何か収益をもたらす仕組みを考え実行しなければならない。結局、オーバチュアのペイドサーチの考えを取り入れた広告料×ユーザがその広告をクリックした数を広告のランク順にしたアドワーズ広告を2000年から開始した。検索結果はユーザが欲しい情報の順に出力し、出力結果の右のスペースに小さい広告を掲載し

た。広告料は1クリック当り何セントという取決めをネット契約し、自動引落としをした。そのようにして経費を低く抑えた。広告料はテレビ、新聞などの広告料より安く、広告の効果が大きいので広告主は増加した。

### 3.4 アドセンス (11)

2003年頃からブログという簡便なwebサイトが利用され始めた。自分の経験や考えをまとめてwebにアップし世界に公開するものである。多くは日記風の記述であるが、その他に車の乗り心地、DVD装置、カメラ、携帯電話、PC、PC付属装置などありとあらゆる商品の使い心地、便利なところ、メーカーによる違いを自分の経験したことを公開するものもある。それなりに有益な情報がアップされているが、中にはカリスマブログといわれる多くの人に愛読されるブログが現れた。グーグルはブログの空きスペースに、ブログのコンテンツに関連した広告を手数料を取って自動的に掲載し、ブログの作者に広告料を支払うというビジネスモデルを考案し、実行した。これをアドセンスといい、2003年に開始したが、2007年現在アドワーズ広告と同額の収益をグーグルにもたらしている。

### 3.5 その他のグーグルのサービス

#### 3.5.1 グーグル検索

筆者がグーグル検索するのは、教材作成のときである。教材原稿はLatex<sup>(12)</sup>を使用するが、矢印はどのようにするか、スクリーン環境印刷のコマンドは何であったかなど人に聞く前にグーグル検索を利用する。Latexに関する日本語の情報は数多くあり、情報を提供して下さる方々にいつも感謝している。教材の内容は、Linuxを使用したサーバ構築である。DNS, web, FTP, Samba, Mail, FireWall, 動画コンテンツサーバなどを予め設定して教材を書く。Linuxは年に2回バージョンアップするので、教材は毎年書き直す。マニュアル本を読んでもうまく行かないときは、グーグル検索をする。Linuxに関する情報はネットに数多くアップされており、グーグルはそれらの情報を的確に検索してくれるので非常にありがたい。ITの底上げにグーグルは貢献している。

#### 3.5.2 グーグルニュース

グーグルニュースは、各新聞やTVニュースを寄せ集めたものではない。複数の記事をクローラが収集し、比較して総合的に分析し、それらを編集した客観的事実を読者に提供する。グーグルニュースと新聞を読み比べると、読者は報道に違いや偏りがあることを知るであろう。

新聞は、ロイタ通信、AP通信などのキー局から得た情報を、営業が獲得した広告とともに編集し、新聞紙に印刷してトラックで各地に運び、バイクで一軒一軒配達している。読者は、興味がある記事をちょっと読み、広告は全く見ずに一日で捨ててしまう。これは壮大な無駄であり、人的資源、エネルギー資源、森林資源などの無駄は、はかり知れない。2007年10月1日、読売、朝日、日経3社は、RSS (RDF (Resource Description Frame) Site Summary) でネットニュースサマリーを流し、記事本体は紙の新聞を買って読んで下さい、という連合を組んだ。読みたい記事をわざわざ買いに行く読者がいると思っているのであろうか。

立派な記事を書く人は、web上でいくらか書ける。いい記事かどうかは読者が決める。いい記事のサイト、ブログには広告掲載申込みが殺到するので、収入は十分得られる。グーグルは今後もニュースを無料で提供し続けるので、他の報道は常に比較されることになるであろう。

### 3.5.3 グーグルマップ

組織名、住所などをキーワードに検索すると、伸縮自在の地図が表示される。地図は、上下左右スムーズに移動して見える。さらに、グーグルアースを選択すると、道路だけでなく人工衛星から撮影した精密な航空写真を見ることができる。近い将来、地図にも広告が掲載されることになるであろう。

### 3.5.4 G メール

筆者は、自宅でもメールを送受信するために2002年からヤフーメールを学内メールと併用している。学内あてにきたメールをヤフーメールに転送すると、いつのまにかヤフーが割当ててくれた10MBのメッセージストアではあふれてしまい、メールを下さった方々にご迷惑をかけた。Gメールは、メッセージストアの容量が1人当たり1GBと聞いたので、Gメールに変えようとしたところヤフーメールも1GBに拡大されたので、今もヤフーメールを使用している。Gメールがよい影響をおよぼしたと思っている。

### 3.5.5 グーグルドキュメント<sup>(13)</sup>

2007年9月にサービスが開始された。グーグルのトップページを開き、検索窓の上にあるmore>>をクリックする。Google サービス一覧が開き、コミュニケーションと共有の中のドキュメントをクリックする。ここでGoogle アカウントまたはGmail アカウントを要求される。アカウントは無料なので取るとよい。

アカウントを入力すると、MSワードと同じ機能を持つGoogle Docs、MSエクセルと同じ機能を持つGoogle Spreadsheets、MSパワーポイントと同じ機能を持つGoogle Presentationが使える。これまで無料で提供されてきたOpenOfficeと同じであるが、違いはインストールしなくてよいことおよび作成したファイルをオンラインで使えることである。

OpenOfficeは2003年からサービスされている。学生の一部は使用しているが、OpenOfficeの使用はまだ限られている。グーグルドキュメントの使用拡大を期待している。

### 3.5.6 グーグルブックサーチ<sup>(14)</sup>

ページの学部出身大学であるミシガン大学図書館の蔵書約700万冊をデジタルデータにして世界に公開する計画が進められている。同時に、スタンフォード大学図書館の800万冊、ハーバード大学図書館1500万冊、アメリカ議会図書館、ニューヨーク公立図書館、オックスフォード大学図書館の蔵書約250万冊をデジタルデータにして世界に公開する計画も進められている。著作権団体が抗議するので、著作権が残存する書籍は後回しにして着々と進めている。東京大学図書館にも参加要請がきているが、決断が遅いと英語の情報はネット経由で再利用されるが、日本語の情報は紙のまま埋もれてしまうことになる。

### 3.5.7 グーグルベース

住宅、自動車などの販売、料理のレシピ、求人、求職、住宅賃貸、チケット、本、CD販売など生活に密着した情報を提供し、市内程度の狭い範囲で売ります、買いますの場を提供する。日本ではまだサービスされていない。決済は近い将来サービスされるグーグルチェックアウトが使用されるであろう。

#### 4. グーグルによってもたらされるもの

##### 4.1 グーグル以前のマーケティングの常識<sup>(15)</sup>

1897年、イタリア人学者ヴィルフред・パレトが、縦軸を売上額、横軸を商品の種類としてプロットすると、図4.1のように、左から急に下がり、あとはだらだらと延びることを見出した。この図を分析したパレトは、売上額上位20%の商品の売上合計額が80%を占め、下位80%の商品の合計額は20%にしかならない、という法則を導いた。この法則は、あらゆる商品、ソフトウェア、映画、音楽などにあてはまるし、人間社会のあらゆることにあてはまるといわれている。これをパレトの法則、80:20の法則という名前でマーケティング理論で使われている。

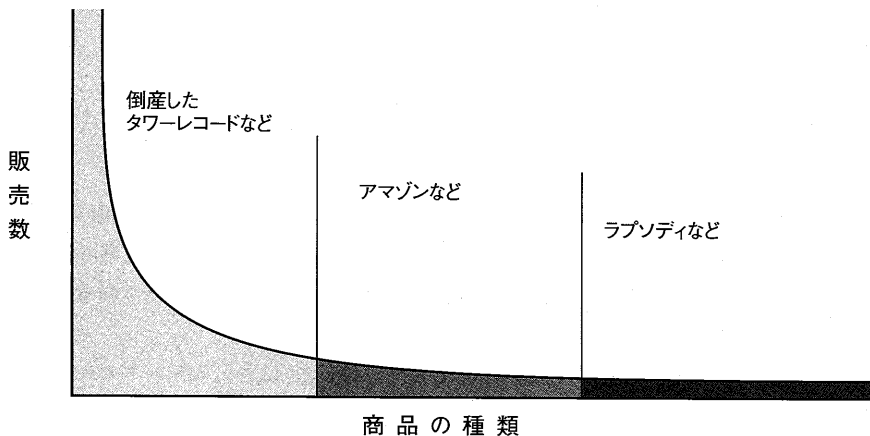


図 4.1 ロングテール

図4.1の左側は、実際の店舗でも収益が出るヒット商品、中央は店舗にはおけないままにしか売れないネット販売に適した商品を示している。最近では、右側のメディアには乗らない、したがって広まらないポピュラソングでもラブソディのようにビジネスにしているケースも見られる。グーグルなどの検索が普及したことにより、図4.1が平坦になりつつある。

##### 4.2 アマゾンドットコム社<sup>(16)</sup>

1995年、ジェフリ・ベゾス夫妻はインターネット上の書店を思いつき、アマゾンドットコム社（以下アマゾンと略記）を設立し、開業した。思いつきというのは、店舗を持たないので資金が不要、在庫を持たないので運転資金が不要であり、最低限の人員とサーバコンピュータがあれば、開業できるであろうということである。本は売れたが、売上シェア全米一位のバーンズアンドノーブルが資金を使って値下げをしたため、アマゾンは長い間赤字の経営が続いた。バーンズアンドノーブルにない売れない本もアマゾンのネット上の店舗にあるので、読者は検索して購入できた。このようにして読まれた本の中でよいものは具体的によいところがアマゾンのサイトにアップされた。書評を読んだ人のうち何人かが購入するといった連鎖がおきている。

200万点を越える出版物があっても、大型書店に陳列できるものはせいぜい10万点である。ネット上なら200万点とその書評をおいても、テキストデータだけなら100GBで十分である。書店の書棚の前で人間が目を探すより、デジタル化された書誌データのデータベースを検索する方が早い。



この便利さを利用した読者に支えられアマゾンに現在収益をあげているのは、大切なことである。なぜなら、今まで出版社、大型書店の宣伝力でベストセラーが作られ、ベストセラーの情報に偏りがちであった情報源が、いい選択肢が多くなるからである。

アマゾンは、CD、医薬品、家庭用品などの有名ではないが、良い商品をネット上で販売している。4.1で述べたパレートの法則でいうと、売上合計額 20%の方の商品をロングテールと呼ぶが、アマゾンはロングテールの中の良い本や商品を届ける流通業者といえる。

#### 4.3 グーグルのビジネス

グーグルのビジネスを端的に言えば、検索エンジン、ニュース、地図、データベース、メール、アプリケーションソフトなど、今まで有料であったものを無料で提供し、それらを提供する画面に小さな広告を掲載するアドワーズ、アドセンスを営む広告代理業であるといえる。広告主が大きくてもかまわないが、小さくてもクリック数に応じた広告料なので、料金に見合った効果がある広告が出せる。会社概要、商品広告を web 上にアップし、広告申込みはすべてネットで行う必要がある。会社概要は紙に印刷するのはやめてネットで配布することにすれば、経費はむしろ安くなる。日本国内の 250 万社の情報がすべて電子化されると、自社が必要としている取引先を的確に検索できる。この取組みは産業が停滞している地方の中小の都市にいい影響をもたらすと考えられる。地方の中小の企業は自社の情報を経費がかからない web 上にアップするだけで業績があがる可能性がでてくる。米国ではもちろん日本でも、今までの方法では広告が出せない中小の企業がアドワーズ、アドセンスを活用して広告を出し、業績を伸ばした例が多数報告されている<sup>(17)(18)(19)</sup>。

グーグル検索を利用すると、物を購入する人は、広告を見て買うのではなく、店頭で店員が売りたい物を買うのではなく、ネットで検索していろいろな人の書き込みをよく読んで、結果としてロングテールまで選択肢に入れて判断するという賢い行動ができる。DVD を買う前に DVD と使用後の感想またはクチコミをキーワードとして検索、上位の情報をクリックすると使用後の感想やクチコミ情報が得られる。さらにキーワードに関連した広告を見ることができる。何回も店頭へ足をこぼなくても商品情報が得られる。

アマゾンが書籍で実現したロングテールから新しいビジネスを発掘したことを、グーグルは広告業で実現したといえる。グーグルはアマゾンよりずっと高収益で実現した。それは賢い広告主と賢い消費者がツールとして選択しているため広告の効果が高いのが理由である。また、2005 年度の日本の媒体別広告費は、テレビ 2 兆円、新聞 1 兆円、雑誌 4 千億円、ネット 3 千億円である。ネットの伸びる余地は十分残っている。なお、米国のネット広告費はグーグルだけで 1 兆円を越えている。

#### 4.4 グーグルの弱点

自分のサイトやブログに広告を取り、その広告を自分で多数回クリックすると自分の広告収入が増える。しかしこれは明らかな詐欺行為である。グーグルは自社にも広告料手数料が入るにもかかわらず、同じアドレスから不自然なクリックがあった場合は自動的に発見し、そのサイトやブログのページランクを最下位にする。このような措置に対する非難の声が聞こえるが<sup>(20)</sup>、広告主や一般利用者から見ると、公平な有難い機能である。決して弱点ではない。もう一つ中国国内で天安門事件をキーワードに検索しても情報は表示されないのはいけない、という非難がある<sup>(21)</sup>。これも中国以外の国では天安門についての情報は得られるので、日本人や米国人がグーグルを非難することはお節介であろう。グーグルを使うと人間が馬鹿になる<sup>(22)</sup>という非難もあるが、それは個人の自由であって馬鹿

になると思う人は使わなければいいだけの話である。ランク付けは別にして、情報選択の決定権はユーザにあることが、賢いユーザに支持されている。

## 5. おわりに

大きな企業はお金を出せば、ネット上で自社の製品を良く書いてもらえる、とあって世界的に有名な日本のS社が実際にアップした。しかし、そのような書込みは、実際に使用した複数のユーザに本当の使用感を書き込まれすぐに否定された。賢いユーザは、賛成意見、否定意見両方を読んで行動する。これをグーグルがサポートしている。

1995年に、ビル・ゲイツ未来を語る、という本が出版された。8章に摩擦ゼロの資本主義というところがある<sup>(23)</sup>。CPU能力が向上し、ネットワークの帯域が広がると、ネット上で自分が欲しいものを探し、ネットで注文することにより、生産者から消費者に直接品物が届くので、取引きのコストが下がり無駄な経費がかからなくなる。それを摩擦ゼロの資本主義といい、それによって人類の生活は向上する、と未来を語った。13年経過したいま摩擦ゼロの資本主義を実現しつつあるのはマイクロソフトではなくグーグルであるといえる。

## 参考文献

- (1) NHK取材班：グーグル革命の衝撃、NHK出版（2007）p.40
- (2) デビッド・ヴァイス、マーク・マルシード著、田村理香訳：Google 誕生、イースト・プレス（2006）
- (3) ジョン・バツェル著、中谷和男訳：ザ・サーチ、日経BP（2005）p.105
- (4) トーマス・L・スターリング、ジョン・サルモン、ドナルド・J・ベッカ、ダニエル・F・サバレーゼ著、北野宏明監訳：PC クラスタ構築法－Linux によるベオルフシステム、産業図書（2001）
- (5) M.Riordan and L.Hoddeson 著、鶴岡雄二、デーン・マツシゲ訳：電子の巨人たち下、ソフトバンク（1998）p.284
- (6) ロバート・ヤング、ウェンディ・G・ローム著、倉骨彰訳：OS 戦線異常あり、日経BP（2000）
- (7) ベッカ・ヒマネン著、安原和見、山形裕生訳：リナックスの革命、河出書房新社（2001）
- (8) 永井武：新潟国際情報大学情報文化学部紀要 2007 年 3 月
- (9) New York Times, 2006 年 6 月 14 日付
- (10) (3) p.162
- (11) (3) p.225
- (12) 奥村晴彦：Latex2 美文書作成入門、第 3 版、技術評論社（2004）
- (13) 日経 Linux、2008 年 1 月号
- (14) (2) p.351
- (15) クリス・アンダーソン著、篠森ゆりこ訳：ロングテール、早川書房（2006）p.17
- (16) 脇英世：ポストゲイツの覇者、講談社（2001）p.243
- (17) 佐々木俊尚：グーグル－既存のビジネスを破壊する、文春新書（2006）
- (18) (1)
- (19) (3) p.228
- (20) (2) p.368
- (21) 文芸春秋、2006 年 6 月号 p.138
- (22) (1) p.213
- (23) ビル・ゲイツ著、西和彦訳：ビル・ゲイツ未来を語る、アスキ（1995）p.258