Google APIを利用した検索方法の研究

目次

序論　――――――――――――――――――――――――1
第１章 Googleの位置
　第１節 インターネットとGoogleの関係　――――――――――――２
　第２節 Googleのデータベースとしての価値　――――――――――５
第２章　Google APIの現状
　第１節 Google APIの概要　――――――――――――――――７
　第２節 Google APIを利用した検索補助　――――――――――――８
　第３節 Google APIを利用したWeb視覚化　――――――――１０
　第４節 Google APIを利用したランキング、そしてグラフ――１３
　第５節 Google APIを利用した学習補助　――――――――――１６
　第６節 Google APIを利用した言葉遊び　――――――――――１７
第３章　Google APIの展望
　第１節 Google APIの問題点　――――――――――１８
　第２節 Google APIの可能性　――――――――――――――１９
結論　――――――――――――――――――――――――２２

序論

インターネットを活用する上で、検索エンジンは必需品となっている。その検索エンジンで筆者が一番利用しているのが「Google」である。なぜ「Google」を利用しているかというと、単純に検索エンジンとして優れているからである。

ところでGoogle APIをご存知だろうか。「Google」が誇るデータベースをユーザー自らが自由に利用することができるというのだ。

このコンセプトに惹かれ、筆者はGoogle APIが我々にとってどのような存在になりうるかを研究しようと思ったである。そして研究するにあたって以下のような順序を踏むことにしたのだ。

第１章ではインターネットの有用性、そしてその中での「Google」の位置、「Google」がデータベースとしてどのくらいの性能を持っているかを述べる。

第２章ではGoogle APIを実例を挙げて、どのような分野で活用されており、またどのような問題が生じているかなど、Google APIの現状を考察する。

そして第３章では、第２章で考察した事を踏まえて、Google APIの全体の問題点、そしてそれを受けての可能性を模索する。

このような段階を踏んでGoogle APIを利用した検索方法を研究していく所存である。

第１章 Google の位置

第１節　インターネットとGoogleの関係

現在インターネットは我々の生活になくてはならないものになりつつある。総務省のまとめた「平成１５年通信利用動向調査の結果」によると、平成１５年末のインターネット利用者数は、対前年比７８８万人増の７,７３０万人となった。また、人口普及率は対前年比６．１％増の６０．６％である。現時点では、高齢者などがインターネットを利用するには敷居が高いのだろうが、パソコン等の情報機器のバリアフリー化が進めばそのような人達も手軽に利用できるようになるだろう。この事からも、これからますます我々の生活に浸透していくと言えるだろう。

さて、インターネットは主にどう使われているのだろうか。再び「平成１５年通信利用動向調査の結果」によると、パソコンからのインターネット利用内容で、一番多いのが電子メールであり５７．６％、ついで情報検索が５７．４％、ニュース等の情報入手が４８．７％とある。このことから、インターネットを利用する人の半数以上はなんらかの情報を検索するために使用することが多いと言える。

閲覧したい情報を探す場合には、検索エンジンを利用することがほとんどだろう。インターネット上などで一般に公開されている様々な種類の情報を、検索ワードなどを用いて検索できるものを検索エンジンといい、この検索エンジンを Web サイトに組み込んだものを検索サイトという。そして、その検索エンジンには大きく分けて２つのタイプのある。それは「ディレクトリ型検索エンジン」と「ロボット型検索エンジン」だ。前者で代表的なのが「Yahoo!」であり、後者で代表的なのが

「Google」である。

「ディレクトリ型検索エンジン」とは、人手で構築したウェブディレクトリ内を検索するサーバー、システムのことである。特長として、人手で構築しているため、質の高いウェブサイトを検索可能だ。また、サイトの概要を人手で記入しているため、検索結果の一覧から目的のサイトを探しやすいという特長がある。

しかし、検索対象となるサイトは人手で入力するため、検索対象となるサイト数が多くないという欠点がある。WWW(World Wide Web)の爆発的な拡大によって、あらゆるウェブサイトを即時にディレクトリに反映させることが事実上不可能になり、現在では主流ではなくなっている。 このため、「ディレクトリ型検索エンジン」では、検索にヒットするサイトが無かった場合、「ロボット型検索エンジン」を用いて結果を表示するような、併用型のものが多いのだ。そして「Yahoo!」もそのようにしている。

「ロボット型検索エンジン」は、コンピュータ化された「スパイダ」と呼ばれるソフトウェアを使って何百万、時には、何十億ページのWebページをデータベースに登録する。そしてタイトルだけでなくページの内容も検索することができ、「ディレクトリ型検索エンジン」よりはるかに詳細な検索をユーザに提供することが可能である。「Google」は「ロボット型検索エンジン」で今もっとも人気があるのだ。

なぜ人気があるのかというと、「Google」の特長としてよく言われることは、検索が速く探している情報が見つかりやすいといったことだ。検索が速くて探している情報がすぐに見つかるなら、人気が出るのも当然である。ではなぜ他の検索エンジンに比べ、求めている情報が見つか

りやすいのだろうか。それはGoogleが独自の「PageRank」という概念を用いWebページをランク付けし、ランクの高いページから優先的に検索されるようになっているからだ。

「Google」を利用して検索してみれば分かると思うが、会社名や商品名を検索してみるとかなり高い確率で公式ページがトップに表示される。そして、その他の求めている情報も簡単に探し当てることができるはずだ。

日本では「Yahoo!」が現在検索エンジンの圧倒的なシェアを誇っている。しかし、ウェブアクセスのリアルタイム解析技術で知られるオランダの「OneStat.com」による、２００４年３月から４月にかけての世界の検索エンジンのシェアで見ると、Googleが５６．４％で１位である。２位のYahoo!の２１．１％であるので、大きく引き離していると言えるだろう。最近では、「Google」を利用してウェブページを検索することを俗に「ぐぐる」という表現をされることがある。これは、「Google」が日本でもかなり認知されつつあることを表しているだろう。

インターネットにおける検索エンジンの必要性という観点から見ると、インターネットの膨大な情報から自分の欲しい情報を検索エンジンを使用せずに一個人が探し出すというのは至難の業だろう。言ってしまえば一切検索エンジンを使わずに物事を調べる方法がリンクを辿るという方法しかないのだ。当たり前のように使っている検索エンジンが、もし使えなくなるとインターネットの価値が下がるといっても過言ではない。なぜなら情報というのは他人に触れられて初めて価値が出てくるからだ。閲覧されない情報など無意味であり、手軽に情報を入手できるというのがインターネットの長所であろう。この事からも分かるように、

今や検索エンジンはインターネットに無くてはならない重要な存在となっている。

第２節　Googleのデータベースとしての価値

現在「Google」は約８０億ページのウェブページをデータベースに登録している。これはとてつもない情報量である。実際に「Google」で検索してみるとわかるのだが、なんらかの意味を持った言葉であれば検索結果に何件もの、時には何十件、何百件ものウェブページへのリンクが表示されている。はっきりいって検索にかからない言葉を探す方が困難であり、載っていない情報はないのだ。

一昔前は何かについて調べようとすれば、事典や辞書等で調べる事がほとんどであった。しかし、今なら「Google」で検索することによって最新の情報を手に入れることができる。「Google」自体が、とてつもなく巨大な百科事典であると言えるのである。

しかし「Google」が全てのインターネット上のデータを網羅しているわけではなく、またそのような検索エンジンは存在しないのだ。なぜならインターネット上のデータ量はあまりに膨大であり、その変化が著しいからであり、データベースにすべての情報を登録し維持することが不可能だからである。

そして興味を持たれなくなった情報は活発に更新されることはなくなり、そのような情報が載っているWebサイトは増加することはほとんどなく、減少していくこととなる。そのため検索結果として現れることがなくなってしまうのだ。

また、「ロボット型検索エンジン」の仕組み上仕方のないことなの

だが、文中に出現しない単語は検索できないという大きな欠点も持っている。

インターネットでは文字ばかりでなく画像や音声、動画等も送ることができる。しかし、現在そのようなマルチメディアを駆使したウェブページがそれほど存在するわけではない。なぜなら、そうしたファイルは容量が大きくなりダウンロードに時間がかかるためだ。

しかしながら現在インターネットのブロードバンド化が進んでおり、最終的に各家庭に光ファイバーが張り巡らされるような状況になると、マルチメディアを駆使したウェブページがもっと増加することになるだあろう。

ただし、注意しなければならないのはインターネット上に流れている全てが正しい情報ではないということだ。このことはなにもインターネットだけに限ったことではないが、インターネットではそのような誤った情報が流れやすい状況にある。

なぜかというと、どんな立場の人でも時間を気にせずに情報を発信することができるというインターネットの長所を利用し、悪意を持って誤った情報を流すというような人も存在するからだ。

また単純に間違っている情報もある。情報の受け手側は、正しい情報と正しくない情報を正しく取捨選択していく必要があるのだ。そうしなければインターネット、ひいては「Google」を有効なデータベースとして扱うことは難しいだろう。

第２章　Google APIの現状

第１節　Google APIの概要

Google APIとは、２００２年４月に「Google Web APIs」よりリリースされた、Web サービスを使用して「Google」の検索を行うためのAPI（Application Programming Interface）のことである。

Webサービスとは，クライアント側の任意のアプリケーションからWebブラブザを介することなくWebサーバ上のサービスへアクセスするための標準的な枠組みのことを指すのだ。要するにWebブラウザを使わずに、プログラム中からWebにアクセスして必要な情報を取り出せるサービスなのだ。

『GOOGLE HACKS』によれば、Google APIは開発者が自分専用のインターフェースを作り、Googleの豊富で膨大な量のデータベースを利用して、検索結果を好きなように利用するのに十分な機能を持っているとのことだ。何ができるかというと、例えばGoogleの持つ検索の仕組み、スペルチェックの仕組み、Googleキャッシュを外部から利用できるのである

Googleキャッシュとは、ウェブページのリンク先のサーバーがダウンした場合に備えて、ページの内容が見られるように多数のウェブページをGoogleが独自に保存しているデータのことである。注意しなければならないのは、キャッシュに保存されたデータは必ずしも最新のものとは限らない。なぜならGoogleがデータベースに登録した時点のページを保存しているからである。

Google APIを利用するには、Google APIへのログインともいえる、

デベロッパーズキーを入手するための登録が必要だ。このデベロッパーズキーを使用して１日に１０００回までのGoogle APIを使った検索が可能であり、それを超える場合は翌日以降に再び検索しなければならない。なお２００４年１２月現在ベータ版であり、利用するにあたっては無料となっている。

どんなに素晴らしいプログラムを作っても、利用されるデータベースがそれに伴っていなければ意味がない。しかし、その点に関しては問題ないと今までで述べてきたことから言えるだろう。

第２節　Google APIを利用した検索補助

このGoogle APIには検索補助というべきものAPIがある。それが「My Google Search」と「GoogleAPISearchTool」である。

「My Google Search」は、まず検索をすると検索結果が１０件表示される（図１－１参照）。この中で自分が探している情報とは違うページにチェックを入れて画面下の「Refine by Example」ボタンをクリックすると、自動的に検索キーワードが追加された上で新たな検索結果が表示される（図１－２参照）。

つまり、チェックを入れたページ、すなわち不必要なページに関係する情報を検索結果から排除し、チェックを入れなかったページ、すなわち必要なページが検索結果に表れるようなキーワードを「My Google Search」が追加して再検索することで、よりよい検索結果を表示しているのだ。

「GoogleAPISearchTool」には、画面右下にある「Cache...」ボタン（図２－１参照）を押すことによって、自動的に記録された過去に使用

した検索キーワードとその検索結果のログを見ることが出来る、という特徴がある（図２－２参照）。

例えば、ある日検索して見つけたページを数日経ってからもう一度閲覧しようと思った時に、もしどのような言葉で検索したか忘れていても、「GoogleAPISearchTool」を使っておけば検索結果のログが残っているためすぐにそのWebページを閲覧することができるのだ。

検索したWebページが数ページならば思い出すのも簡単であるが、数１０ページを超えると、どの検索キーワードでどのWebページが検索されたかわからなくなるだろう。そのような時にこのAPIは真価を発揮するのだ。

現在の検索エンジンは主に検索対象ページの増加と検索精度の向上に目が向けられている。しかし、時間が無いときなどに大量のWebページが検索にヒットすると、そのなかから目的の情報を探し出すのはユーザーにとってかなりの負担になる。

基本的には上から順にクリックしていくか、Webページのタイトルやその他の情報を検索ページから取得し、それを参考にWebページに飛べばいいのだが、ヒットしたWebページが多いときにはかなりの時間と労力を使う。そのような時のために絞込検索というテクニックがあるが、ある一つの問題がある。

その問題とは、絞込検索に用いる検索キーワードの選定には、検索しようとする分野に対するある程度の知識や、検索エンジンに関する知識が求められ、インターネット初心者や検索しようとする分野に対してあまり知識の無いユーザーには大きな負担となることだ。

そこで「My Google Search」である。一見この問題を「My Google Search」は解消してくれているようだが、残念ながら実際にはそうで

もない。なぜなら日本語で検索をした時に検索結果が文字化けを起こしてしまうのだ。これでは簡単にページを選別することができず、「My Google Search」の長所を生かすことができない。

そしてこれはGoogle API全般にいえることなのだが、日本語で検索をすると、検索結果のサイトが日本の「Google」を普通に使う場合と大分違い、ヒット件数が少なかったり、最近作成されたページがひっかからないケースが多いのだ。残念なことにGoogle APIが検索するデータベースと、通常の日本の「Google」が検索するデータベースが違うようなのである。

「Search Guide」によれば、Google APIが検索するデータベースは、Google社が「Yahoo.com」などにアプリケーションサービスプロバイダとして提供している検索データベースのようである。

つまり日本の「Google」で検索できるページが、この海外の検索データベースだと検索することができない場合があるのだ。そしてこの検索データベースは日本のサイト、ページはかなり少ない量しかないようである。日本語で検索するにあたって、残念ながらデータベースとしてはレベルが落ちるようだ。

これは検索補助のために作成されたAPIにとっては致命的である。日本語のサイトを検索するのが日常的な我々にとってこの事実はAPIを使用することを敬遠させる要素になりうるのだ。この事から、残念ながら検索補助を目的としたAPIは現時点では活用しづらいといえる。

第３節　Google APIを利用したWeb視覚化

Webページ同士の関係図を表示する、いわばWeb視覚化とでもいうべ

きAPIがある。それが「HotSpot」と「TouchGraph GoogleBrowser V1.01」である。

「HotSpot」とは

WWWに散らばった情報を、関係性、重要度（ページランク）といったパラメーターを基に抽出しヴィジュアライズします。WEBの関係マップツール（「HotSpot」より引用）

である。

使い方は「URL」の項目に興味のあるWebページのURLを入力し「Search!」ボタンをクリックすると検索開始となる。そして画面に結果が表示される（図３参照）。関連URLが最終的にこのように表示されるのだが、この表示のされ方が重要なのだ。

具体的に説明すると、入力したURL先のWebページに関係が深く重要度が高いと判断されたWebページを表すURLが明るく大きくなり画面の中央に集まってくるのである。その結果としてユーザーが必要としている情報が含まれている可能性が高いWebページは一通り画面の中央に集まることになる。集まったWebページの中でも 特に中央に近いものから見ていくことで探している情報が見つかる可能性が高くなるのだ。

注意として、あくまでも中央に集まったWebページに必ずしもユーザーの欲するものと合致する情報が含まれているとは限らないのである。そして、最終的に集まったWebページを自力で一つ一つ見ていく作業が必要となるのだ。

また、このAPIは日本語で文字化けを起こす事がないのでまだ実用的であるといえるだろう。

「TouchGraphGoogleBrowser V1.01」はURLを入力し「Graph IT」ボタンをクリックすることによって、「Google」から得られる情報を

もとに,他のサイトとの関連を視覚化するAPIである（図４参照）。特定のURLに関連するページから始まり、さらにそれらのページに関連するページという具合に拡大していき、画面上に巨大なURLの地図を表示するのだ。

また表示されている各URLをダブルクリックすることによって、一つにつきさらに最高３０件の検索結果が返されURLの地図が書き足されるのである。これによって関係図がさらに巨大化し、新たなWeb関係図というものを見ることができる。そのことにより、全く異質なWebページ集合の間にある意外なつながりを発見することもあるだろう。

残念ながらこちらは日本語のサイトのURLは文字化けして表示されてしまう。

「HotSpot」と「TouchGraphGoogleBrowser V1.01」に共通して存在する長所として、多くのURL間の関連性が一目で認識できるということが挙げられる。単なる検索結果を見ただけではこのようにいかないだろう。そして表示された結果から異なる事柄の新たな繋がりを発見する可能性も存在するのである。そしてその事を目的として使用するのが最も適しているだろう。

ところで、上記２つのAPIとは一線を隔したWeb視覚化を目的として作成されたAPIに「Memorium」というものがある。なぜ一線を隔しているのかというと、「Memorium」は望んでいる情報を手に入れることを目的とせずに作成されたものだからだ。本来情報を検索するという行為は能動的であるが、この「Memorium」は受動的に情報を入手しようとしている点が決定的に違うといえる。

「Memorium」の画面上では、ユーザーが興味をもつキーワードとその説明を表現するカードが常に浮遊して動いており、 URLとサイトの

説明をユーザーが眺めることができるようになっている。その板が衝突すると、衝突した２つのキーワードを使って検索が実行され、 その結果が新しいキーワードをもつカードとして新たに出現する（図５参照）。そしてカードをクリックするとそのWebページを閲覧することができるのだ。

このままだとカードは増え続けてしまうのだが、カードには寿命があり、２回接触すると画面上から消えるようになっている。また、新たなキーワードを入力したり浮遊速度を制御するなどして「Memorium」の動きを制御することもできるようになっている。

全く関連性の無いキーワードを持ったカード同士で検索をし、その結果がまたカードとなり他のカードと検索をするというわけだ。これでは短時間で求めている情報が手に入ることは、ほとんど無いのだが「Memorium」はそのような用途で作成されてないのでこれでよいのである。

インターネットへの常時接続の普及に伴って、現在ではパソコンの電源が一日中付いていることも少なくない。そこで「Memorium」をパソコンが使用されていない時に起動させておくのだ。自分が興味を持っている事柄がキーワードとして提示し、その中から思わぬ関係を持つものが発見しようとするという使用法が考えられる。そしてこのような使用法が適しているのだろう。

第４節　Google APIを利用したランキング、そしてグラフ

Google APIを利用して、Webの動向を探ることを目的として作成されたAPIがある。それは「Googleランキング測定」と「Googleplot」で

ある。APIを利用して、世間の動向を知ろうとするわけだ。

「Googleランキング測定」は特定のサイトをある検索キーワードで検索をした場合にそのサイトがGoogleで何番目に表示されるかを調べるAPIである。

「対象URL」と「検索キーワード」を入力すると、調査したいサイトのある検索キーワードで検索された時の順位が表示されるのだ。（図6参照）

これはWebページを運営していない人には、あまり価値が見出せないかもしれない。なぜならWebサイトを運営している人のためにこのAPIは作られているといっても過言ではないからだ。

インターネットにWebサイトを公開しているということは、ほとんどの場合そのWebサイトに自らが発信したい情報を載せておき、知ってもらうことが目的なのである。そして自らが発信したい情報を多数に知ってもらうためには、検索エンジンで容易に検索されるということが必要だろう。このことからWebサイトを運営している人にとって、自分のサイトがGoogleでどのような順位になっているかはWebサイトを運営するにあたって重要な要素であるといえる。

例えば、特定の検索キーワードで10位以内にランキングしていれば、よく閲覧されるだろうが、100位を越えてくるとはっきりいって閲覧されることはほとんどないといってもよい。これはWebサイトの運営者からすれば由々しき事態である。

このため、自らが発信する情報をより多くの人に知ってもらうには、検索エンジンの上位に自分の運営するWebページが表示されるように工夫する必要がある。そしてその行為そのもの、またはそのための技術のことを「検索エンジン最適化」と呼ぶ。

一般に「検索エンジン最適化」と言った場合、「Google」の持つアルゴリズムに合わせた対策のことを指す。これはインターネット上での「Google 」のシェアを考えると当然といえ、それは前章で述べたとおりだ。

また、「検索エンジン最適化」は個人のWebサイトに用いられる他に企業にも浸透しつつあるようだ。なぜ「検索エンジン最適化」が企業に注目されているかというと、ネットビジネスの手法として重要視されているからだ。

膨大なWebページの中から目的の情報に辿り着くためには、今や検索エンジンの存在が欠かせないのは今までで述べたとおりである。まず検索エンジンに登録されていることが第一であるが、しかし、検索エンジンに登録されているWebページの数は途方もない数になるので、単なる登録だけでは、アクセスアップは望めないのだ。

つまり、いかに適切なキーワードで上位に表示されるかが、重要なのである。更に検索エンジンの上位に表示されることで、不思議と信頼感や安心感などが生まれ検索エンジンを利用した宣伝には、アクセス数の増加に加え、広告とは違った安心感などの付加価値も付く手法なのだ。

「Googleplot」はある特定の検索キーワードで取得した結果をグラフ化している。要するに検索されて表示される件数が多いほどグラフとしての値が大きくなるということである（図７参照）。

これは検索された情報を入手するのではなく、検索結果数をグラフにすることによって、調査する時に使用するのである。例えば、これを利用すると、ある特定の検索キーワードや概念がどの程度流行しているのかを大体知ることができる。 また、「Google」の検索ヒット数はデータベースが更新される度に変化するので、それを並べてグラフにすれば、

ある特定の検索キーワードや概念が流行りつつあるのか、もしくは廃れつつあるのか一目でわかるのだ。

全ての事柄においてこのように言えるとは一概には言えないが、大抵の事柄に対してこのことは通用するだろう。例えば、これを利用することにより、流行っている検索キーワードや概念を知る事が可能となり、その結果世間のニーズに合わせた情報を発信することができるのだ。

第５節　Google APIを利用した学習補助

Google APIを利用した学習補助をいうべきAPIがある。それが「WebLEAP」だ。「WebLEAP」は入力された文や表現に含まれる単語の列の、WWW上の出現頻度をグラフィカルに表示するものだ。「WebLEAP」を使うことによって、その表現が世界中でどのくらい使われているかを知ることができる。この情報を、文の訂正や修正に役立てることができるのだ。文章中の各部の出現頻度が色分けされてグラフィカルに表示されるので、一目で修正すべき場所などを知ることができる（図８参照）。残念ながら日本語には対応していない。

このAPIを使用するメリットは、いわゆる生きている文章から表現を参照できるということだろう。単語の用法は考えるときは、本来ならば辞書を引き、辞書に用意された例文を利用するところだ。しかしインターネットには実際に使われている文章がそれこそ膨大にあるのだから、そこから適切な表現を探し出せばいいのだ。

辞書は単語の意味を調べるためのものであり、単語は実際にどのような場面でどのように使われるかが重要である。 そして、多くの辞書では単語の意味ごとに様々な例文を用意している。 しかし、普通の辞

書の中の例文のほとんどは生きた文章から抜き出したものではなく、用例の説明のために辞書編纂者が作成したものも多いように思われる。

その点このAPIを利用すれば、どのような表現がより一般的かを知ることができるのだ。日本語が使用できないのが非常に残念である。

第６節　Google APIを利用した言葉遊び

Google APIには現在の世相を風刺した言葉遊びとでもいうべきAPIがある。それが「Googlismもどき」と「Google Smackdown」である。

「Googlismもどき」はある特定の事柄が「Google」上でどのように解釈、理解されているかを調査するAPIだ。調べたい事柄をフォームに入力し、「送信」ボタンをクリックすることで結果が表示される（図９参照）。

例えばある人物を「Google」で検索する。するとその人物が有名であればあるほど、様々な検索結果が表示され、どのように思われているか知ることができる。ただし、そうした作業は「Google」で検索して、検索結果のリストからそのWebサイトに飛んで、その後違うWebサイトに行くことを繰り返すことになる。

そんな時に「Googlismもどき」を使用するとスムーズにその作業を行うことができるのだ。

「Googlismもどき」は入力した検索キーワードが「Google」でどんな風に語られているかを文章ごと直接抽出して表示してくれるAPIである。つまり興味を持っているある人物、ある物、ある場所などがインターネット上でどのように思われているかを知ることができるのである。

「Google Smackdown」は二つの検索キーワードを同時に入力し、結

果を表示するAPIです。使用方法は二つのフォームに比較して検索したい検索キーワードを入力し、「throw down!」ボタンをクリックするのだ。（図１０－１参照）すると結果が表示される。（図１０－２参照）

例えば、似て非なる二つの事柄、様々な類義語などのどちらが「Google」でより多く検索されるかを調査する時に使用するという用途がある。その結果を元によりインターネット上ではどちらの言葉がより一般的かを知ることができるのだ。

どちらも日本語で検索することによって文字化けを起こすことがないので、一見正しく結果が表示されているように見える。しかし今まで述べてきたとおり、日本の「Google」のデータベースで検索されるわけでないので、必ずしも相応しい結果が表示されているとは限らないのだ。

この事から本節で考察した二つのAPIで表示される結果は、日本語で検索した場合、参考程度に留めておくのが適当であるといえる。

第３章　Google APIの展望

第１節　Google APIの問題点

Google APIについて、前章でいくつかのAPIの実例を挙げて考察してきたが、その中で様々な問題があることが明らかになった。

まず、日本語を使用することを前提に作成されたものが少なく、日本語で利用した場合に文字化けを起こすものが多い。また、日本語に対応しているAPIで日本語で検索をしたとしても、Google APIが検索するデータベースと、通常の日本の「Google」が検索するデータベースが違うため、検索結果が日本の「Google」を普通に使う場合と比べ明ら

かに質が落ちるという事態に陥ってしまうのである。

この事実は、「Google」がデータベースとして優れていても、日本語を利用するとGoogle APIではその能力を完全に生かす事が不可能であるということを表している。これはユーザーからすれば不親切極まりない出来であるといえる。

前章で日本語を使用できるAPIをいくつか説明したが、それですら充分にデータベースを活用できない中途半端な状態なのである。これでは少なくとも日本語を利用した場合Google APIは充分に機能しているとはいえない。

そして長い間更新されておらず、ベータ版のままである。これはあまり積極的に開発されていないという事の裏返しではないだろうか。

またこれはインターネット全体にいえることなのだが、旬を過ぎてしまった情報は活発に更新されることはなく、そのような情報が載っているWebサイトは急激な増加は望めない。そのため検索結果として現れず、結果Google APIにもその情報は反映され辛くなってしまうのだ。

また、日本ではGoogle APIの認知度は低いを言わざるを得ない。皮肉なことに「Google」でGoogle APIを検索しても、日本語に対応しているGoogle APIが数えるほどしか結果として表示されないのだ。このことはとりもなおさず少なくとも日本では開発が活発ではないことを現しているのだ。

第２節　Google APIの可能性

今まで本論文で考察してきた上で、Google APIがこれから先どのように展開していくのだろうか。

まず、検索補助は使い方によっては活用することができるだろう。やはり情報を検索する上で、情報の取捨選択はもはや必須であり、その補助をしてくれるAPIは活用される価値があるといえるだろう。ただしそれも日本語に対応しているという前提での話である。

実際には前章で述べたとおりに日本語への対応が不十分なのである。この事は情報を検索するという行為にとって大きな障害であろう。このような事実から現状のままではGoogle APIを利用した検索補助はあまり発展することがないといえる。

次に、Web視覚化については、どうであろうか。これについては、まず大きく違う２種類のものがあるのは前述したとおりである。

「HotSpot」と「TouchGraphGoogleBrowser V1.01」の共通の特長として、Web関係図とでもいうべき図を画面上に表示し、多くのURL間の関連性を一目で認識することができる。単なる検索結果を見ただけではここまで認識することはできないだろう。そしてそのように表示された結果から、全く異なる事柄の新たな繋がりを発見するという使い方が活用されていると言える。

この二つも前述したように、日本語への対応が不十分なためその機能を充分に発揮することができなくなっている。

ところで「Memorium」は望んでいる情報を手に入れることを目的とせずに作成されたもののため、この点ではあまり影響を受けてはいないといえる。情報を検索するという行為は能動的であるのが本来の形であるが、この「Memorium」は受動的に情報を入手しようというコンセプトのもとに作成されているので、その点が前者の二つのAPIとは決定的に違うのだ。

そしてランキングとグラフについては、ランキングは「検索エンジン

最適化」を主とした目的とした場合利用価値はあるといえる。また、グラフは世間の動向を調べるために使用すればよいだろう。

ただし、くどいようだが前述のとおりこれらのAPIは「Google」のデータベースを充分に活用できないので、精密な結果を必要としている場合に信頼できるデータを取得することはできないのだ。この事からある程度の参考にはなるだろうが、そのデータを丸々信用するということはしてはいけないだろう。

言葉遊びについても同様である。特にこの分野はインターネット上での世相の反応が重要視されるので、たとえ結果を取得してもますます慎重に扱わなくてはならなくなるのだ。

最後に学習補助について、「WebLEAP」を英語で使用する場合には全く問題なく使用できる。また、これの日本語版とでもいうべきAPIが作成されれば興味深い。しかし、ここにも日本語に対応しきれていない問題があるため、そのようなAPIが作成される可能性は低く、この分野はこれ以上の発展は望めないだろう。

Google APIを日本語で利用する場合、日本語への対応が充分でないことから、様々な問題が生じている。これからの見通しは正直悪いとしか言いようがない。

しかし、この状況は日本語への対応が充分になった場合、打開される可能性もあるだろう。

結論

Google APIを研究した結果として、現状のままでは状況は厳しいと言わざるを得ないだろう。なぜなら、日本語に対応し切れてないのが最大の理由である。第２に知名度が低いというのもあるだろう。

もし完全に日本語に対応してくれれば、この状況を打破できる可能性はある。なぜなら、やはり日本語で使用できなければ単純に使いづらいのだ。また、今のままでは日本語で利用すると正しい検索結果が表示されないため、せっかく作成された様々なAPIが真価を発揮することができないのだ。

また、知名度が低いというのも大きなマイナスだ。情報を検索する時に使用する検索エンジンの最大手がGoogle APIを公開しているのに、この状況は少し寂しいものがある。このままでは埋もれていってしまう可能性が非常に高いと思われるのだ。

他者に閲覧されない情報は無意味であり、その事柄についての議論が起こるはずもない。この観点からも今のままの認知度であるならば、これから先の展望はないだろうが、もし認知度が上るような事が起これば今の状況を打開する可能性もないとは言えないのだ。

参考文献

『GOOGLE HACKS』 Tara Calishain　Real Dornfest　著
山名　早人　監訳　　田中　裕子　訳
株式会社オライリー・ジャパン　出版
２００３月9月6日　初版第２刷発行

『GOOGLE ポケットガイド』
Tara Calishain　Real Dornfest　DJ Adams　著
山名　早人　訳
株式会社オライリー・ジャパン　出版
２００３年１０月２４日　初版第１刷発行

参項URL

『HotSpot』　管理人　野上　大輔
(http://www.m5g.org/google/hotspot.html)

『TouchGraph GoogleBrowser V1.01』
(http://www.touchgraph.com/TGGoogleBrowser.html)

『Memorium』　管理人　渡邊　恵太
(http://www.persistent.org/memorium/hiki.cgi?FrontPage)

『Googleランキング測定』　管理人　Geekなページ
(http://www003.upp.so-net.ne.jp/geek/index.html)

『googleplot』　管理人　高林　哲
(http://namazu.org/~satoru/mingplot/#googleplot)

『WebLEAP』　管理人　山之上　卓
(http://yama-linux.cc.kagoshima-u.ac.jp/~webleap/index-j.html)

『Googlismもどき』　管理人　ちゃ(CHARMY)
(http://www.casebook.jp/index.html)

『Google Smackdown』　管理人　Paul Bausch
(http://www.onfocus.com/googlesmack/down.asp)
「平成１５年通信利用動向調査の結果」　総務省
(http://www.soumu.go.jp/s-news/2004/pdf/040414_1_a.pdf)
「OneStat.com」
(http://onestat.com)