

Web マーケティングにおけるデータ収集及び分析手法

豊谷 純[†] 三浦和真[‡]

日本大学[†] 日本大学[‡]

1. はじめに

インターネットが普及して、世界中の各サイトで様々なサービスが開発及びリリースが行われている。これだけ情報の入手先やその内容が多種多様で多岐に渡るようになると、その一つ一つを Web ブラウザを経由して利用するのは煩わしいものとなる。

例えば天気予報や最新のニュースなどは、ブラウザを利用しなくとも、パソコンのデスクトップ上等都合の良い場所に表示可能である。また天気予報や、ニュース、為替情報、株式情報などを、それぞれ別々のウィンドウを表示するのでは無く、1つの画面に欲しい情報を全て好きな場所に表示させることも可能である。

これを容易にするのが、APIであり、開発言語に依存せず、インターフェースだけ書式を合わせれば、お互いに全く異なるシステム間で情報の交換が可能になる。

本論では世界中の製品評価サイトや、ブログやつぶやきページから、情報を収集して、マーケティングを行う手法を確立する事を目標として、どの手法が適切であるのかを調査する。

2. 情報収集方法

インターネット上の情報検索サイトやブログ、つぶやき、商品評価サイトなど、ありとあらゆるページを参照すれば、情報は入手可能である。しかしながら、膨大に増大する情報を手作業で全て行うには非効率的であり、現実的な方法とは言えない。

従って Web 上から情報を取得すると共に、プログラムでデータマイニングなどの統計解析を行うのが望ましい。これに対しては、最近では、様々なサイトが情報を Web サービス API によって提供するのが流行のようになっている状況であり、この API を利用した情報収集を行うことにする。

3. 単純な API による情報取得技術⁴⁾

基本的には API を利用する際には、入力パラメータをサーバへ送信すると、その結果が XML データにて提供される。それを例えば DOM 等を利用す

Data collection and analysis technique in Web marketing

[†]Jun TYOTANI · Nihon University

[‡]Kazuma MIURA · Nihon University

れば、コンピュータ内のメモリに展開し、データの操作を簡単に行う事ができる。

Web サービス API を利用するには、具体的にどのようなすれば良いかを示すために、比較的シンプルな天気情報の表示を例に解説する。

様々な企業が情報提供しているが、天気予報の Web サービス API¹⁾ を例に取り上げる。

この API では都道府県別に、天気予報を文字だけでは無く画像と併せて提供しており、東京都を例にすると図 1 のような情報が得られる。

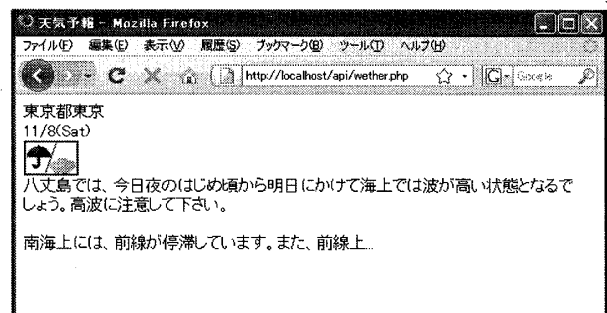


図 1 天気予報の取得画面

4. 主要な Web サービスの API

まずインターネット上で提供されている API には、どのような種類のものが存在して、どのような情報が利用できるのかを把握することにする。

特にこの分野の先駆者である Google は、著名な Google Map や情報検索の API 等、50 以上に及ぶ充実したサービスを提供している。

また Amazon Web サービスでは、EC2/S3 等の名称で、仮想マシンやデータストレージの時間貸しを行うクラウドコンピューティングサービスを提供している状況である。このようにありとあらゆる情報をプログラムで操作可能な環境が整っている事が分かる。

代表的な Web サービス API 一覧 (2009 年 11 月)

Google Code

Google Maps, Youtube, カレンダー等

Yahoo! デベロッパーネットワーク

各種検索, Yahoo! オークション等

Amazon Web サービス

商品検索, 商品情報取得等

- 楽天ウェブサービス
 - 楽天市場, 楽天トラベル等
- ぐるなびWebサービス
 - お店検索, クーポン等
- じゃらんWebサービス
 - ツアー旅行や宿泊予約等
- お天気Webサービス
 - 全国および都道府県別の天気概況情報等
- はてなWebサービス
 - 人力はてな, はてな検索, ランキング等
- イースト辞書ウェブサービス
 - 英和辞典, 和英辞書, ウィキペディア等
- 価格.com Webサービス
 - 価格.com, アキバ総研, 食べログ.com等
- @nifty Web Service
 - @niftyクリップ, ビデオ共有等
- リクルートWebサービス
 - ホットペッパー, フロム・エー等
- Web ad Fortune 無料API
 - 年・月・日単位で古いデータが取得可能
- ケンコーコムウェブサービス
 - 医薬品の商品情報提供

他

5. ブログからのデータ取得

本研究では、アンケート調査を行わずに、インターネット上のありとあらゆる情報を収集してマーケティングを行う手法の確立を目標としている。

そのために、まずはWeb上のブログから情報の収集を行ない、マーケティングの適用を検討する事にした。情報の収集先としては、Yahoo!ブログ²⁾ やGoogleブログ³⁾ の検索エンジンのAPIを利用する事にした。そして通常はブラウザで操作するが、手作業でデータマイニングや統計処理を行うことは現実的では無い。

現在ではデータは殆どの場合で、XML形式で収集することが可能になっており、そうで無いサイトもRSSは対応している場合が多い。ただし、RSSは、その性質から上位10程度しか表示されないため、マーケティングを行うには、得られたHTML形式のデータから、必要なデータのみを読み取る工夫が必要になると考えられる。

最も簡単な方法は、数時間毎など、定期的に時間間隔をおいて、データを収集する方法である。具体的には、プログラム自身にタイマー機能を持たせるか、またシステムのバックアップツールやソフトウェアテストツールの自動実行機能、あるいはネットサーフィンの自動巡回ツールやOSのタスク機能等をプログラムと併用して利用する必要がある。

またもう一つ、注意すべき事項があるが、RSSでは、一覧表示を見やすくするために、ブログ本文が表示されず、タイトルの他に要約あるいはブログの冒頭部分が数十文字表示されるに留まる事が多い。従って、ブログで検索した結果を受け取り、そのURLを取得して、さらにその人のブログのサイトにアクセスし直して本文を取得する必要がある。

その一例となるが、Yahoo!ブログ²⁾ の検索結果を図2に示す。この方法は、ブログのみならず、口コミサイトや商品の評価サイトからのデータ取得でも利用できる。このことから、プログラムで自動的にマーケティングのためのデータ収集が行えることが確認できる。

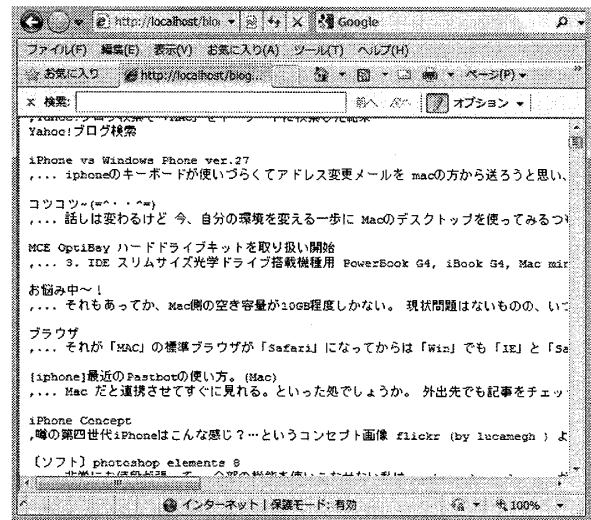


図2 ブログの検索結果画面

6. おわりに

このようにAPIを利用して、Web上のありとあらゆるサイトにアクセスして、テキスト形式として取得できることが示された。それによって得られたテキストデータをテキストマイニングすれば、マーケティングに利用可能であることが示されたと考える。

7. 参考文献

- 1) livedoor Weather Hacks(天気予報), http://weather.livedoor.com/weather_hacks/
- 2) Yahoo!ブログ検索: <http://blog-search.yahoo.co.jp/>
- 3) Googleブログ検索: <http://blog-search.yahoo.co.jp/>
- 4) 豊谷, 若林, 渡邊, 大澤: 「WebサービスAPIによる情報ディレクトリとマッシュアップ技術」, 日本情報ディレクトリ学会誌, 2009年, p. 67-72