

統計データを通して、地域を知り、日本を知り、世界を知る

1. 背景

社会や経済のグローバル化、少子高齢化の進展、地球温暖化など社会をとりまく環境が大きく変化する中、個々の世帯や企業が的確な意思決定を行っていく上で、統計は重要性を増している。また、統計データを通して、地域を知り、日本を知り、世界を知るということは非常に大きな知的好奇心をくすぐるものである。

この統計データに関して、日本では、統計局において、統計データを国民の共有財産と位置づけ、迅速かつ継続的に提供することで、その利活用を推進している。諸外国、及び国連においても、同様の活動が行われている。

この統計データの利活用に関し、Internet上での利活用が十分に進んでいるとは言えない。例えば、各新聞社が発行するインターネット上の新聞¹に掲載される統計データ、及びグラフの数は、紙の新聞に比べると非常に少ない。Wikipedia²のようなWeb上の百科辞典においても、同様である。

2. 目的

本プロジェクトでは、統計データに容易に、楽しくアクセスし、活用できる仕組みを提供することで、人々の知的好奇心をくすぐり、統計データを介して世の中を覗いてみる、理解する機会を増やす一助となるシステムをInternet上に構築し、Internetにおける統計データの利活用促進に貢献することを目的とする。

将来的には、世界中の統計データを収集し、Webサービスとして公開することで、世界の情報産業の発展に貢献したい。

3. 開発の内容

本システムは、様々な形式(Excel、PDF、XML など)で公開されている統計データを収集し、Mashup 可能なリソースとしてInternetに公開するものである。

本システムのInternet上における位置付け、及び提供する機能を図1に示す。図中、treering.infoは、このシステムの仮称である。本システムは、Internet上において、公開されている統計データと、その利用者の上に位置することになる。

今回の開発では、統計データとして、総務省統計局から公開されている都道府県別データのみを対象とした。

¹ 例えば、<http://www.nikkei.co.jp/>など

² <http://ja.wikipedia.org/>

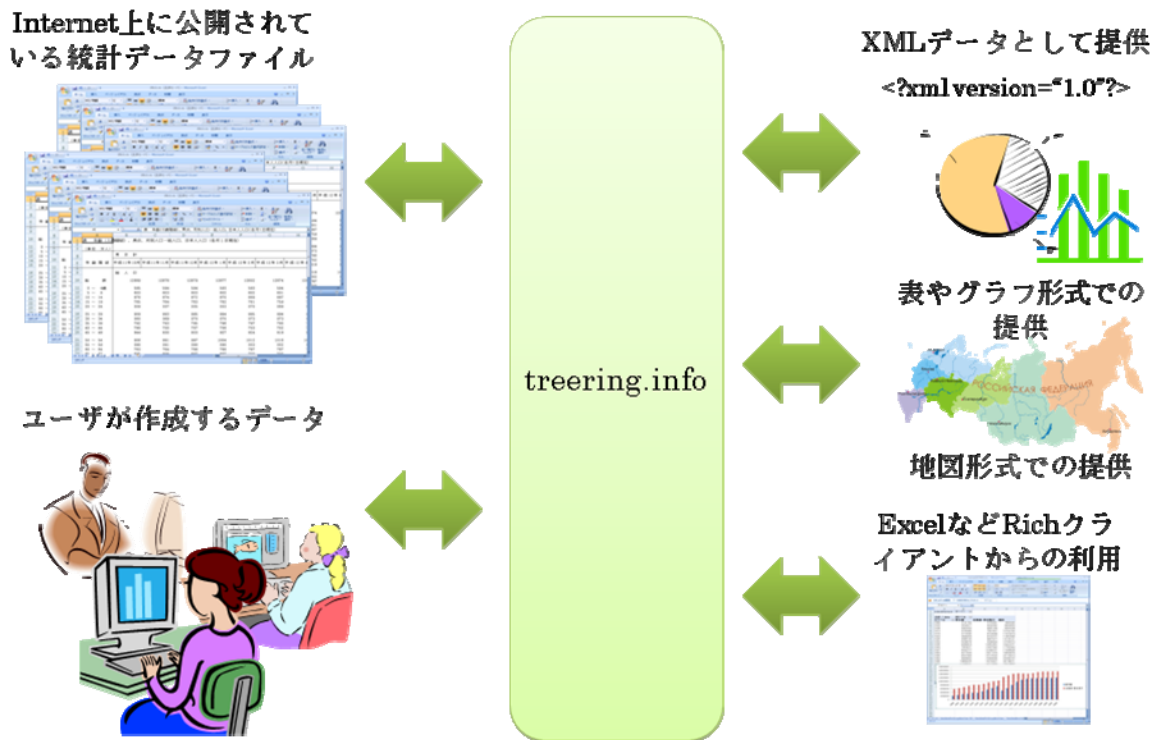


図 1 本システムの位置付け、及び機能概要

作成したWebサービス一覧を表 1に示す。データ取得用とは、Internetに公開されている統計データの生データを取得するためのものである。ランクデータ取得用とは、前記生データに対して、データを昇順/降順並べた際の順位から、データを取得するためのものである。例えば、北海道で降順の3位以内に入る項目は何かといった問い合わせが可能となる。グラフ取得用とは、ある統計データのグラフを取得する際に用いる。例えば、2000年度の都道府県別人口総数について、東京都と、日本の地域(関東地方、近畿地方など)を比較したグラフを図 2に示す。グループ用とは、日本の地域や、1950年度日本の都道府県の100万人単位グループといったようなグループを作成し、そのグループ内やグループ毎のデータを取得するためのものである。その他、簡易なキーワード検索を実行するためのメソッドなども作成した。

この Web サービスのメソッドは、ニーズに合わせて継続して開発を進める予定である。

表 1 作成した Web サービス一覧

Interface の名称	メソッド数	URL
データ取得用	13	http://dev.treering.info/stats/service/Service.aspx
ランクデータ取得用	7	http://dev.treering.info/stats/service/RankedInfo.aspx
グラフ取得用	9	http://dev.treering.info/stats/service/Graph.aspx
グループ用	9	http://dev.treering.info/stats/service/Groups.aspx

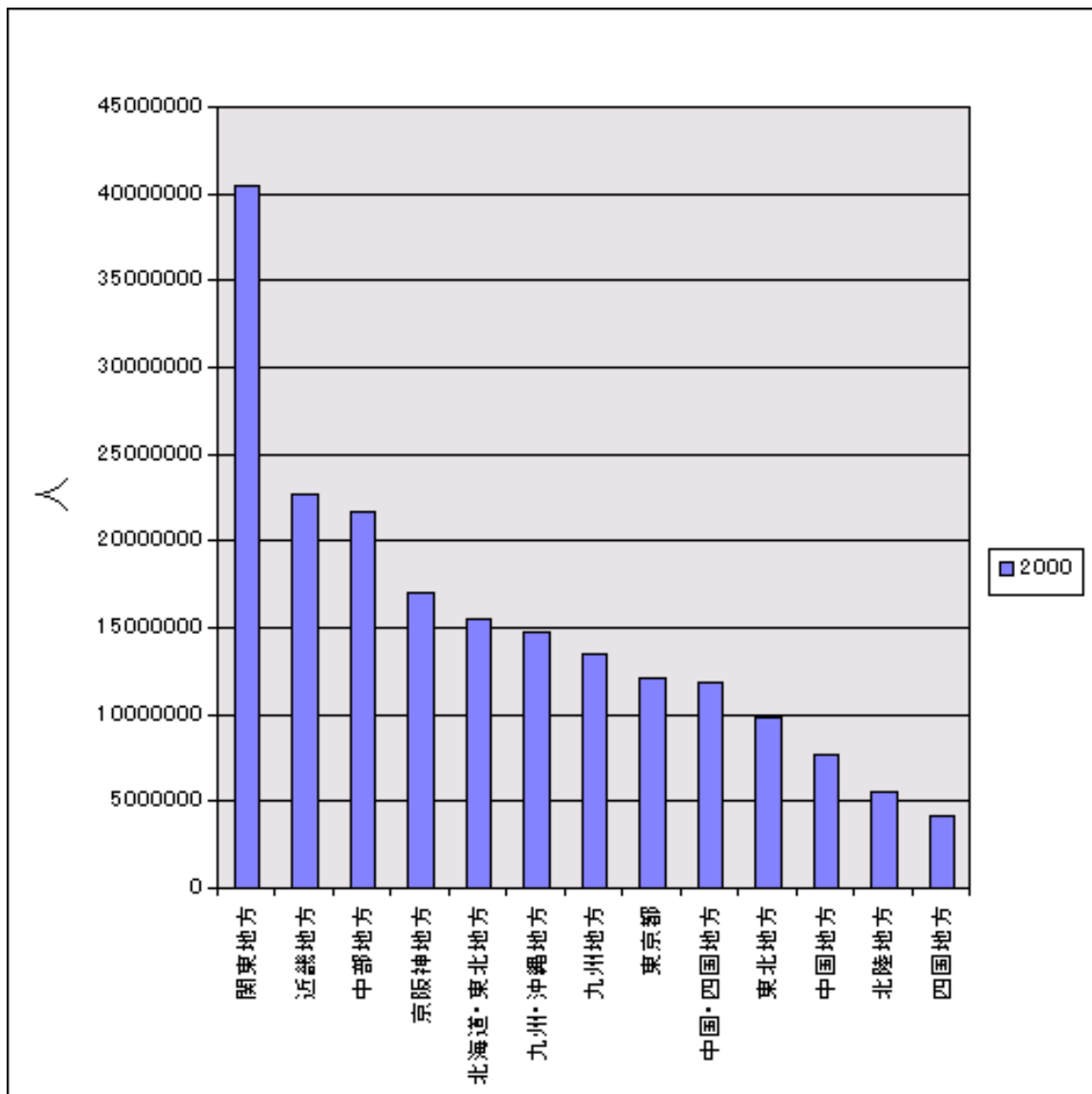


図 2 グラフ取得用 API の例

4. 従来の技術(または機能)との相違

今回開発したソフトウェアの特徴を以下に示す。

- Internet 上に公開されている統計データを収集し、Web サービスとして公開したこと
- 公開されている統計データをそのまま提供するだけでなく、ランキング表示、グループ化したデータの提供、キーワード検索、グラフ作成などの機能を付加したこと
- GoogleMap 上に都道府県別 Polygon を作成し、上記 Web サービスを利用した Mashup サイトを作成したこと

5. 期待される効果

今回作成した Web サービスを公開することで、Internet 上での統計データの利活用を促進することができる。

6. 普及(または活用)の見通し

図 3に本システムの利用イメージを示す。図中、青枠は一般ユーザを、橙枠はフロントエン

ドWebサイトを、黄緑枠はMashupリソース提供サイトを示す。本プロジェクトで開発したシステム(treering.info)の主な役割はフロントエンドWebサイト(各新聞社のWebサイト、Web上の百科事典サイト、情報誌のWebサイト、Blogサイト、コラムニストのサイトなど)に対して、統計データをデータ、表、グラフ、地図等で提供することである。

現在、本システムを活用したフロントエンド Web サイトは存在しないが、今後、活用いただけるサイトを増やす活動を行っていききたい。

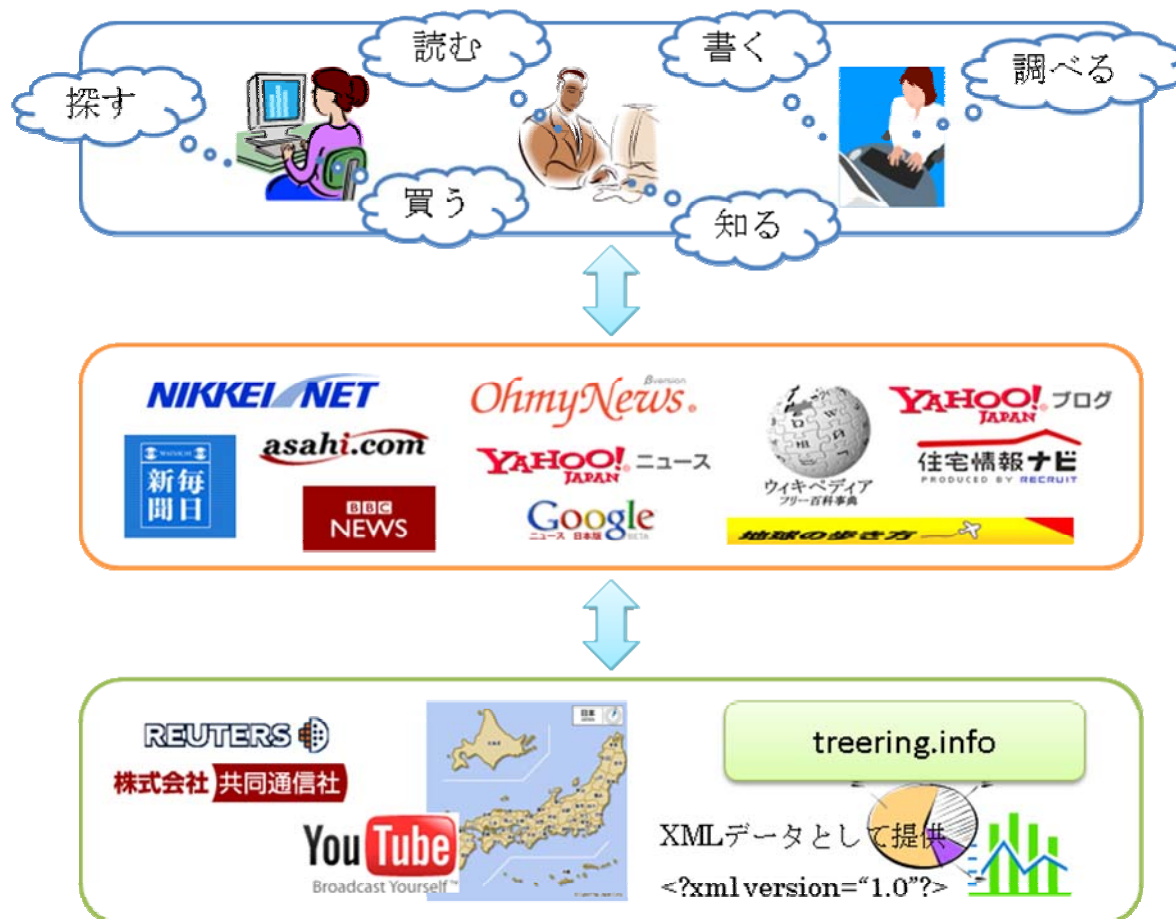


図 3 利用イメージ

7. 開発者名(所属)

吉井 英樹(ソフトバンクテレコム株式会社 研究所)

(参考)<http://dev.treering.info/>

2007年10月10日現在において、利用にはID/パスワードの制限をかけている。ご興味ございましたら、hideki.yoshii@tm.softbank.co.jpまでご連絡ください。