

## 国立国会図書館リサーチ・ナビにおける テーマグラフの生成

坂井 哲<sup>†1</sup> 増田 英孝<sup>†1</sup>  
清田 陽司<sup>†2</sup> 中川 裕志<sup>†2</sup>

2009年5月から国立国会図書館のWebサイトにてサービスを開始したリサーチ・ナビには、さまざまな分野における資料の調べ方を効率的に探せる手段を利用者に提供することを目的とした検索システムが実装されている。本システムでは、資料の調べ方を利用者に提示するために必要となる「カバレッジ」「分類体系の存在」「信頼性」の3つの条件を満たすため、図書館件名標目表とWikipediaを統合的に利用し、入力されたキーワードが図書館分類体系の中でどう位置づけられるのかを示すテーマグラフと呼ばれる図を自動的に生成する機能を実装している。また、ユーザインタフェースにはWebで汎用的に用いられている各種の要素技術を積極的に活用している。本稿では、リサーチ・ナビ検索システムの開発の背景にある考え方を述べた上で、テーマグラフ導出のアルゴリズムと主な技術的ポイントについて解説する。

### Automated Generation of Theme-Graphs at the National Diet Library Research-Navi System

SATOSHI SAKAI,<sup>†1</sup> HIDETAKA MASUDA,<sup>†1</sup> YOJI KIYOTA<sup>†2</sup>  
and HIROSHI NAKAGAWA<sup>†2</sup>

This paper addresses the search engine of the Research-Navi, which starts its service at the web site of the National Diet Library of Japan in May 2009. The goal of the search engine is to enable users to retrieve various methods to get information materials. The following three conditions should be satisfied to meet the goal: coverage for various queries, existence of taxonomy, and reliability. To satisfy these conditions, the search engine generates "theme graphs", which overview relations between query keywords and subject headings of libraries. The algorithm of generating the graphs is based on integrative usage of both subject headings and the Wikipedia category system. Another important feature of the search engines is progressive employment of web-based technologies, including asynchronous communications (Ajax), Web APIs, and

client-side implementation. These technologies enable seamless integration of web information (coverage) and library information (reliability).

#### 1. はじめに

Web検索エンジンの普及は、一般の人々の「情報探し」に対するものの見方に大きな影響を与えている。仕事や買い物、旅行、食事、読書、料理といった日々の行動をするにあたって「まず検索エンジンで探す」という習慣は人々の間で広く共有されている。これにともない、「検索キーワードさえ適切に選べば検索エンジンは答えを与えてくれる」とか、「検索エンジンで何もみつからなければそれ以上探しても仕方ない」という誤解も生じかねない状況が生まれている。

しかし、私たちが日々抱く疑問の中には、検索キーワードを入力するだけでは答えを見いだせないものも数多く存在する。「関東大震災が発生したのはいつか?」「ってどんな病気?」という疑問に対しては、検索エンジンで簡単に答えにたどりつくことができる。一方で、大学の学生の「関東大震災についてレポートを書きなさいいけないんだけど、いったいどんな資料から調べたらいいの?」という疑問や、難病を抱えた患者の「自分の病気について最先端の治療を行っている病院を探す方法は?」という疑問に対しては、検索エンジンは明解な答えは与えてくれない。

「レポートの書くための資料の探し方を知りたい」「病院を探す方法を知りたい」といった疑問を抱くのは、情報探しのテーマがあいまいな場合が多い。このような場合、情報探しのテーマを利用者に推薦することが求められる。情報探しのテーマを推薦するためにはサービス提供側は以下の3つの条件を満たしていることが必要だと考えられる(図1)。

- (1) カバレッジ  
あらゆる分野の質問に対して何らかの情報資源を推薦することが求められる。この条件を満たしていない限り、結局はサービスを利用してもらえない。
- (2) 分類体系の存在  
お薦めすべき情報資源には、情報探しの分野に応じて「定番」とよべるものが存在す

<sup>†1</sup> 東京電機大学未来科学部  
School of Science and Technology for Future Life, Tokyo Denki University

<sup>†2</sup> 東京大学情報基盤センター  
Information Technology Center, The University of Tokyo

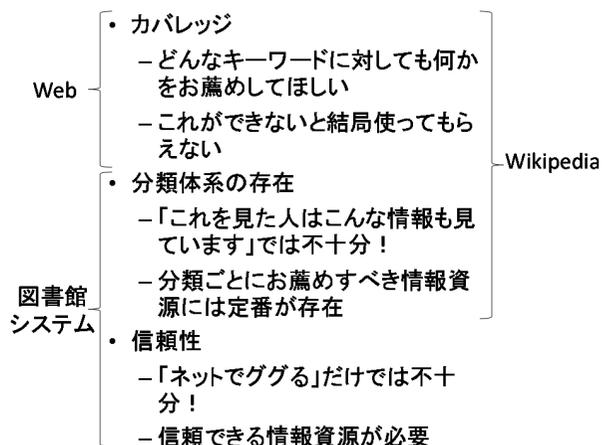


図 1 テーマ推薦の要件

る。例えば、図書館においては入門書(新書など)や参考図書がそれに該当する「定番」となる情報資源を利用者に提示するためには、あらかじめそれらの情報資源を整理しておく分類体系が必要である。

### (3) 信頼性

どんな情報資源を推薦したとしても、利用者側が「このお薦めは信頼できる」という感情を抱かなければ結局は利用されない。「ネットでググる」ことで得られる情報資源を超える信頼性の根拠を利用者に示すことが重要となる。

日常の情報探しにおいて最も広く用いられている Web サーチエンジンは、上記の 3 つの条件のうち「カバレッジ」は満たしているが、「分類体系の存在」「信頼性」においては不十分である。1990 年代の Web では Yahoo! などのディレクトリサービスがこれらの条件を満たす役割を担っていたが、Web 空間の情報量が爆発している現代においては人手によるディレクトリの作成が追いついていない状況である。

一方、図書館の代表的な Web 情報サービスである OPAC では、「分類体系の存在」「信頼性」の条件は満たしているが、「カバレッジ」については不十分である。思いついたキーワードを OPAC に入力しても、何も情報が得られない場合はかなり多い。

オンライン百科事典 Wikipedia は、この 3 つの条件を考慮したとき興味深い位置に存在する。誰でも編集が可能であることから「信頼性」については課題があるが、「カバレッジ」

「分類体系の存在」の条件は満たしていることから、Web と図書館システムとのギャップを埋める架け橋として利用できる可能性がある。

国立国会図書館の Web サイトにて 2009 年 5 月から一般公開されているリサーチ・ナビの検索システムの開発においては、上記の考え方にもとづき、Wikipedia を情報資源の一つとして活用する方針をとった。具体的には、Wikipedia と国立国会件名標目表 (NDLSH) を統合的に用い、入力されたキーワードが図書館分類体系の中でどう位置づけられるのかを示すテーマグラフと呼ばれる図を自動的に生成する機能を実装している。また、リサーチ・ナビ検索エンジンが対象となる情報資源は国立国会図書館が公開しているものが中心であるが、ユーザインタフェースには Web で汎用的に利用されている技術を積極的に活用している。主なものを以下に挙げる。

#### ● タブデザインの採用

限りのある画面領域を有効に活用するため、画面内にタブパネルを配置している。ユーザは 5 つのタブ(「すべて」「調べ方」「本・サイト」「キーワード」「百科事典」)を選択することで表示内容を切り替えることができる。検索結果は、ユーザの用途に沿ってそれぞれのタブ内に分類して表示される。また、「全て」タブには検索結果のサマリを表示させており、ユーザに対するナビゲーションの役割を果たすようにしている。

#### ● Web API の活用とマッシュアップ

リサーチ・ナビ検索システムが対象とする主なコンテンツは、PORTA (国立国会図書館デジタルアーカイブポータル) の API を経由して得ている。これにより、国立国会図書館が所有する多様なコンテンツへの対応を少ない工数で実現することができた。また、PORTA API と書籍販売サイトが提供する Web API を連携させ、検索にヒットした書籍の書影を表示したりレビューを参照したりする機能も実装しており、内部職員向けに試験運用している。

● Ajax 技術によるサーバとの非同期通信リサーチ・ナビ検索システムは Web API の利用によって多様なコンテンツへの対応を実現しているが、コンテンツによって検索にかかる時間にはばらつきがあるため、検索キーワードによってはレスポンスの悪化を招く可能性がある。そこで、ユーザインタフェースに Ajax 技術を採用し、Web ブラウザからコンテンツ毎にサーバへの非同期通信を行わせるシステム構成とした。これにより、検索結果が得られたコンテンツから順次画面に表示させることが可能となった。

● ブラウザ側での検索履歴保持ユーザによる試行錯誤を支援するため、ユーザが過去に入力した検索キーワードをブラウザに保持しておき、プルダウンリストによって簡単に呼

び出せる機能を実装した。

- CMS (コンテンツ管理システム) とのシームレスな連携コンテンツ管理システム Movable Type によって運用されているリサーチ・ナビ本体とハイパーリンクにより連携している。

本稿では、まず 2 節でリサーチ・ナビ検索システムの開発の背景にある考え方について論じ、つづいて 3 節においてシステムの全体構成と技術的なポイントについて詳しく解説する。4 節では、リサーチ・ナビ検索システム独自の特徴であるテーマグラフのアルゴリズムについて述べる。最後に 5 節においてまとめを述べる。

## 2. リサーチ・ナビ検索システム開発の背景

本節では、Web 情報資源と図書館情報資源の特徴を比較し、利用者へのテーマ推薦を実現するための解決策について論じる。

### 2.1 Web 情報資源と図書館情報資源のアクセス手段の現状

インターネットが一般に普及する以前、情報探索を行うにあたって中心的な役割を果たしていたのは図書館であった。図書館には、膨大な資料を整理・組織化し、利用者の情報探索に役立てるためのさまざまなツールが存在する。代表的なツールとしては、日本十進分類法 (NDC) や基本件名標目表 (BSH) などの分類体系があげられる。

現在では図書館にかわって Web が情報探索の中心的な役割を果たしつつある。最初に Google などの Web サーチエンジンを利用することで、たいいていの検索質問に対しては何らかの Web ページが見つかるという状況になっている。しかし、Web 上の膨大な情報は十分に組織化されていないため、情報要求に適したページを見つけることは必ずしも簡単ではない。

近年、Web 技術を活用することで膨大な知識を多数の共同作業によって組織化しようという取り組みが盛んに行われるようになってきている。その結果として、Web 上に一種の情報探索用オントロジーとよべるものが存在するという状況が生まれつつある。これらのオントロジーには、多数の人々の多様な観点が反映されているという大きな利点があるが、限界もある。これらのオントロジーの構築に関わる人々の中心的な動機は、「自分の興味を満たす情報を効率的に探せるようにしたい」というものである。結果として、Web 上に存在するオントロジーはブラウジングを便利にするための組織化にとどまりがちであり、探索を深く掘り下げていくという用途には必ずしも適していない。

一方、図書館の世界で必要な情報にアクセスするための代表的な手段としては OPAC が

存在するが、入力したキーワードに対して何もヒットしないケースはかなり多い。図書館において情報探索の窓口の役割を果たしているレファレンスサービスでは、Wikipedia をレファレンスツールとして活用するという試みがなされている<sup>1)</sup> が、現在はレファレンスサービス自体が一般の人々にあまり認知されていないという状況である<sup>2)</sup>。また、特定のトピックごとに図書館が提供できる関連資料をリスト化したパスファインダーとよばれる情報資源を Web 上で公開する取り組みも行われている<sup>3)</sup> が、図書館が人手をかけて構築する必要があるため、Web 上に存在するオントロジーに比較すればごく一部のトピックをカバーしているにすぎない。

### 2.2 タクソノミーとフォークソノミー

図書館で利用されている分類体系には、NDLC (国立国会図書館分類表) や NDC (日本十進分類法) に代表される図書分類法と、NDLC, LCSH, BSH などの件名標目表が存在する。これらの分類体系は、原則として図 2 左のような構造をもっている。すなわち、頂点となる項目から枝分かれしていく木構造となっている。木構造においては、頂点以外の項目はただ 1 つの親をもっている。また、親項目がさらに親項目をもつことから、多数の階層を構成している。木構造をなす分類体系のことを、本稿ではタクソノミー (Taxonomy) と呼ぶこととする。

一方で、ここ数年 Web 技術を活用することで膨大な知識を多数の共同作業によって組織化しようという取り組みが盛んに行われるようになってきている。その代表例としてフォークソノミー (Folksonomy) と呼ばれるサービスが存在する。例えば Flickr<sup>5)</sup>, delicious<sup>6)</sup>, YouTube<sup>7)</sup>, はてな<sup>8)</sup> などが挙げられる。フォークソノミーとはサービスの一般利用者がコンテンツ (Web ページ, 画像, 動画など) に自身の視点でタグを付与できる仕組みであり、一般的には図 2 右のような構造をしている。フォークソノミーの構造においては、各コンテンツは複数の親を持っている。また、フォークソノミーは項目 (コンテンツ) とタグの 2 階層のみから構成されている。

従来、大量の情報を分類する手法は、タクソノミー的な分類体系によるものがほとんどであった。タクソノミー的な分類体系は分類の専門家によってトップダウン的に管理されており、一貫性のある分類がなされるという長所を持っている。ただし、分類の多様な観点を反映しにくい、語彙数が少ないという短所もある。フォークソノミーは、タクソノミーと対照的な特徴をもっている。多数の参加者がそれぞれの視点からボトムアップ的に分類を作っていくことから、多様な観点からの分類がなされる、多くの語彙をカバーしているという長所をもっているが、一貫性のない分類がなされてしまうという短所もある。

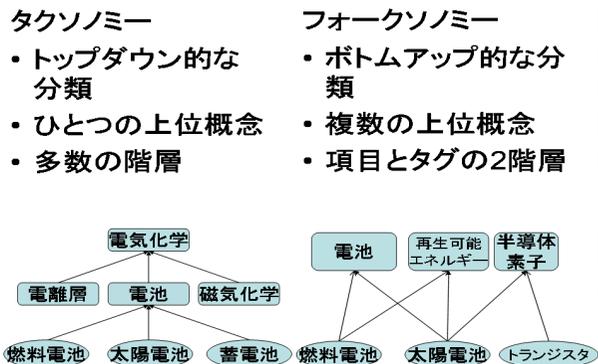


図 2 タクソノミーとフォークソノミー

このように、大量の情報・資料を分類する方法論としては、タクソノミーとフォークソノミーの2つの考え方がある。従来はタクソノミーとフォークソノミーの2つの世界はお互いに独立して存在しており、両者の連携はあまり考慮されてこなかった。

Wikipedia の分類体系である「Wikipedia カテゴリ」は、タクソノミーとフォークソノミーの両者の特徴を持ち合わせている(図3)。例えば Wikipedia 日本語版では「価格」という項目に「マーケティング」「経済学」「市場」などのカテゴリが付与されているが、これらのカテゴリは Wikipedia の編集時に自由に付与することができ、フォークソノミーの特徴を有している。一方で、カテゴリにさらに親カテゴリが付与されて多数の階層をなしている点はタクソノミーの特徴を有している。また、最上位層に属する「社会」「社会科学」「経済」「労働」「産業」などのカテゴリは1つもしくは少数の親カテゴリをもっており、図書分類体系ときわめて類似した構造をしている。さらに Wikipedia は日本語版ですでに60万項目を超える記事が存在しており、カバレッジの面からも優れている。Wikipedia の分類体系構造はタクソノミーとフォークソノミーの2つの世界の橋渡しとなる可能性を持ち合わせている。

### 2.3 解決策

これまで述べてきたように、利用者へのテーマ推薦を実現する上で、Web 情報資源と図書館情報資源にはそれぞれ課題がある。Web 情報資源には、多人数による知識の共有を可能とする技術をベースとした多様で便利なアクセス手段が提供されている反面、テーマ推薦を実現するという観点では十分な組織化がなされていない。一方、図書館情報資源には

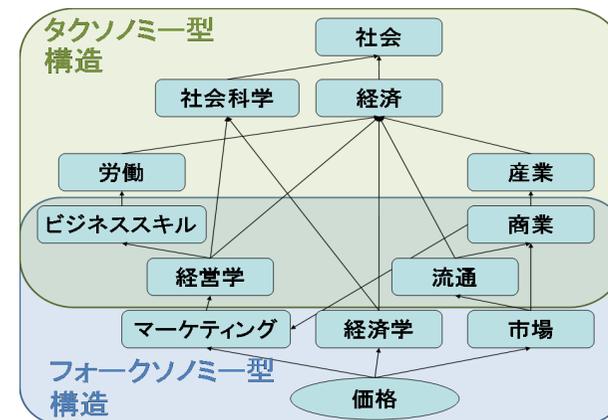


図 3 Wikipedia カテゴリの構造

テーマ推薦に適した厳密な分類体系が存在するが、Web のインターフェースに慣れ親しんだ人々にとってはアクセス手段が十分とはいえない。この課題を解決するために、リサーチ・ナビ検索システムの設計においては、Web で汎用的に利用されている技術を可能な限り取り入れつつ、図書館の分類体系をシステムに取り込むために Wikipedia を媒体として活用する方針をとった。

### 3. システムの構成

本節では、まずリサーチ・ナビ検索システムのユーザインタフェースを示した上で、ユーザインタフェースに表示されるそれぞれのコンテンツが導き出される仕組みを解説する。

#### 3.1 ユーザインタフェースのデザイン

リサーチ・ナビ検索システムのユーザインタフェース(UI)を図4に示す。UIの部品配置は一般的なWebサービスに準拠しており、上部に検索ボックス、下部に検索結果表示エリアを配置している。それぞれの部品の役割を以下に示す。

- (1) 検索キーワードの入力ボックス  
ユーザがこのボックスにキーワードを入力すると、検索サーバにおいて検索処理が行われる。その結果は(3)の検索結果表示エリアに表示される。
- (2) 表示内容切り替えタブ  
ユーザの目的に応じて表示エリアの内容を切り替えることができるようにしている。



図4 リサーチ・ナビのユーザインタフェース(「すべて」タブ)

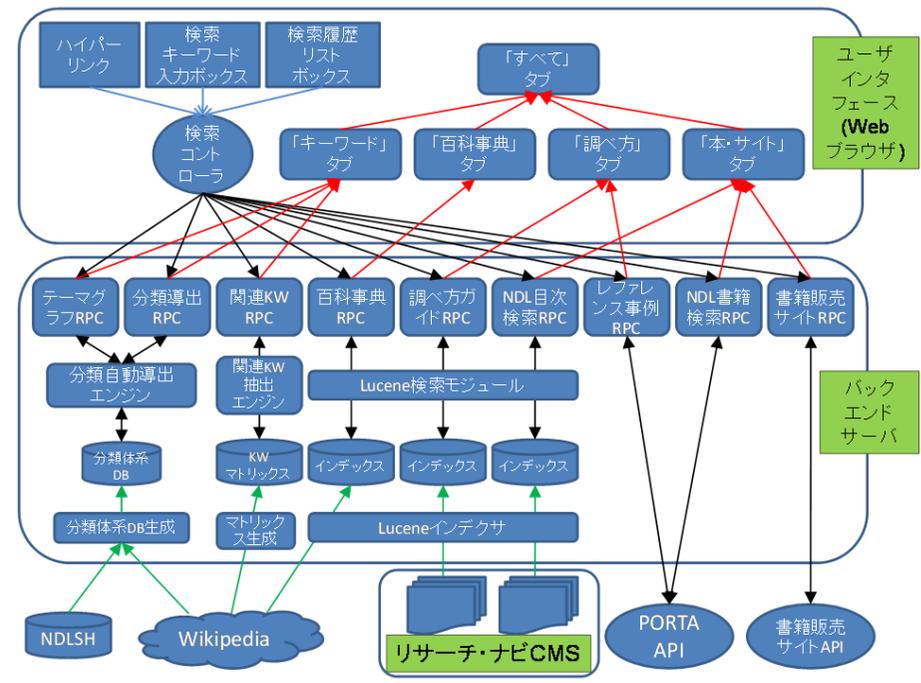


図5 リサーチ・ナビ検索システムの構成

- 現在、「すべて」「調べ方」「本・サイト」「キーワード」「百科事典」の5種類のタブを設けている。それぞれのタブの役割については次節で説明する。
- (3) 検索結果表示エリア  
検索キーワードと選択されたタブに応じて、検索結果アイテムのハイパーリンクのリストを見やすく整理して表示する。ハイパーリンクをクリックすると、当該アイテムを表示するウィンドウが開く。ハイパーリンクにはキーワード入力と同様の役割を果たすものもある(たとえば「燃料電池」で検索した場合に「キーワード」タブ内に関連語として表示される「再生可能エネルギー」をクリックすると、「再生可能エネルギー」で検索した結果の表示画面に遷移する)。エリアの広さは検索結果アイテム数によって縦方向に伸縮する。アイテムが多数の場合には複数のページに分割してページ遷移ボタンを下部に表示し、表示エリアが過度に縦長にならないようにしている。

- (4) 検索履歴リストボックス  
ユーザが以前探したキーワードがプルダウンリスト内に格納されている。任意のキーワードを選択すると、画面状態をそのキーワードによる検索結果に復帰させることができる。

### 3.2 リサーチ・ナビ検索システムの構成

リサーチ・ナビ検索システムは、図5に示すように、ユーザインタフェースとしてのWebブラウザとバックエンドサーバから構成されている。

#### 3.2.1 フレームワーク

リサーチ・ナビ検索システムは、Google Web Toolkit をアプリケーションフレームワークとして採用している。Google Web Toolkit は、Ajax アプリケーションの効率的な開発に特化したツールキットであり、Java 言語により Web ブラウザとバックエンドサーバの双

方の機能を統合的に開発することができる。Java 言語によって記述された Web ブラウザ側のロジックは、コンパイル時に自動的に JavaScript 言語に変換されるため、特別なプラグイン等のインストールを必要とすることなく一般的な Web ブラウザで動作する。また、システム開発時には Java 言語による JavaScript のエミュレーション機能が利用できるため、デバッグ等が容易である。

システムの設計においては、Ajax アプリケーションの大きな特徴である非同期通信を積極的に活用する方針をとった。図 5 に示すように、本システムは外部の Web API (PORTA、書籍販売サイト) を含む多数の情報資源に同時にアクセスする必要がある。もし、全ての情報資源の検索処理が完了した後に検索結果画面をレンダリングする構成をとると、ユーザが検索キーワードを入力してから結果が表示されるまで長い間待たされてしまうことがある。Web ブラウザ側から直接各々の情報資源への問い合わせを非同期通信により行うことで、検索処理が完了した情報資源から順次画面に表示させ、ユーザのストレスを最小限に抑えることが可能となった。

### 3.2.2 ユーザインタフェース

リサーチ・ナビ検索システムのユーザインタフェースは汎用の Web ブラウザによって動作する。リサーチ・ナビの Web サイト <http://rnavi.ndl.go.jp/> にアクセスし、検索ボックスにキーワードを入力することで、リサーチ・ナビ検索システムのユーザインタフェースを構成する HTML ファイル、スタイルシートファイル、画像ファイル、JavaScript ファイルが Web ブラウザに自動的に読み込まれる。読み込み後、JavaScript が自動実行され、ユーザインタフェース内に検索キーワード入力ボックス、検索履歴リストボックス、タブなどの部品を配置する。図 5 中の検索コントローラは、JavaScript によって実装されたモジュールであり、検索の実行において重要な役割を果たす。ユーザが検索要求処理 (検索キーワード入力ボックスへの入力、ハイパーリンクのクリック、検索履歴リストボックスからのキーワード選択のいずれか) を行うと、検索コントローラが起動し、各々の情報資源に対応する RPC メソッド (次節にて述べる) への非同期通信を実行する。非同期通信のコールバックメソッドには、各タブへの検索結果の埋め込み処理が実装されている。これにより、検索処理が終了した情報資源から順次画面に反映されるようになっている。

画面中に配置されているタブのうち、「すべて」タブには特別な役割を持たせている。「すべて」以外の 4 つのタブ (「調べ方」「本・サイト」「キーワード」「百科事典」) に表示された検索結果のうち上位にヒットしたものが、抜粋として「すべて」タブに転送される。

### 3.2.3 バックエンドサーバ

リサーチ・ナビ検索システムのバックエンドサーバは、Java サーブレットとして実装され、Web コンテナ (Apache Tomcat) 内で動作している。Java サーブレットには各々の情報資源に対応した RPC メソッドが実装されている。RPC (Remote Procedure Call) とは、別のコンピュータ上にあるメソッドを遠隔実行することを可能とする技術を指す。各々の RPC メソッドは、ユーザインタフェース側から簡易な手続きにより呼び出すことができる。

各々の RPC メソッドの処理は、大まかに以下の 4 種類に分類される。

- (1) 外部の Web API を直接呼び出すもの  
「NDL 書籍検索 RPC」と「レファレンス事例 RPC」は、国立国会図書館 PORTA の API を直接呼び出している。また、書籍販売サイトの API を呼び出す RPC は、館内の職員向けのバージョンにおいて試験的に利用されている。
- (2) バックエンドサーバ内部に置かれた検索インデックスを利用するもの  
リサーチ・ナビ CMS (Movable Type) に蓄積されているコンテンツ、および Wikipedia の全文データに関しては、全文検索ソフトウェア Lucene によって実装されたインデックスと検索モジュールを利用する。RPC は、定期的にクロールされたデータから生成されるインデックスを参照して検索キーワードにヒットする文書群を抽出し、ユーザインタフェース側に渡す。
- (3) 関連キーワード抽出エンジンを利用するもの  
関連キーワード抽出エンジンは、Wikipedia から生成されたキーワード = 文書マトリックス (行列) を参照して関連キーワード群 (検索キーワードと近い出現パターンをもつもの) を導出する。導出される関連キーワードには、入力キーワードとのベクトル空間におけるコサイン類似度を反映した関連度スコアが付与される。
- (4) 分類自動導出エンジンを利用するもの  
分類導出 RPC は、4.2 節において後述する分類自動導出アルゴリズムによって導出される件名のリストをユーザインタフェース側に渡す。また、テーマグラフ RPC は、同じアルゴリズムによって生成されるテーマグラフの画像ファイルを生成してユーザインタフェース側に渡す。

#### 4. Wikipedia と件名標目表を統合的に活用した分類自動導出

##### 4.1 アルゴリズムの概要

前述したように、Wikipedia は「テーマ推薦の要件」と「カテゴリの構造」の観点から見たときに、きわめてユニークな特徴をもっている。この特徴をうまく用いて、情報探索の出発点として Wikipedia を利用し、そこから概念を一般化することによって図書館の分類体系に導いていくという方法を提案している。

図 6 に導出アルゴリズムの概要を示す。まず、Wikipedia カテゴリの構造について説明する。Wikipedia の記事「阪神・淡路大震災」には、カテゴリとして「日本の経済史」「地震の歴史」が付与されている。さらに、カテゴリ「日本の経済史」には上位カテゴリとして「経済史」が、カテゴリ「地震の歴史」には上位カテゴリとして「災害と防災の歴史」「地震」が付与されている。このように、Wikipedia の記事を一つとりあげてみると、関連するカテゴリ群をツリー構造として取り出せることがわかる。

次に、Wikipedia カテゴリと図書館の分類体系の対応付けについて説明する。Wikipedia カテゴリと図書館の件名の間には、カテゴリ名が一致するものが存在する。図 6 では、「経済史」「災害」「地震」が一致している。

よって、「阪神・淡路大震災」につながるカテゴリが構成する有向グラフの構造を再帰アルゴリズムによってたどることで、「阪神・淡路大震災」に関連する分類を自動的に導出することができる。リサーチ・ナビでは、グラフのエッジに対する重みスコアをノード間の文字列類似度によって定義し、ビームサーチによって重みスコアが相対的に大きい件名を絞り込むアルゴリズムを採用している。

##### 4.2 テーマグラフによる分類の俯瞰

上記に示したように、Wikipedia カテゴリ構造と図書館件名標目表を活用することで、任意の Wikipedia の項目に対して、関連する件名を自動的に導出することができる。リサーチ・ナビ検索システムでは、この導出の過程をグラフ構造として自動的に描画する機能を実装している。このグラフ構造は、与えられた Wikipedia の項目がどのようなテーマを有しているのかを示していることから、テーマグラフと名付けている<sup>9)</sup>。

図 7 に、「燃料電池」という検索キーワードから導き出されたテーマグラフの例を示す。テーマグラフの内容を考察していくことで、「燃料電池」という概念がどのようなテーマと関連を持っているのかを知ることができる。例えば、「再生可能エネルギー」「循環型社会」「環境問題」などの件名と「燃料電池」との関連性を考察すると、「燃料電池は環境問題解決

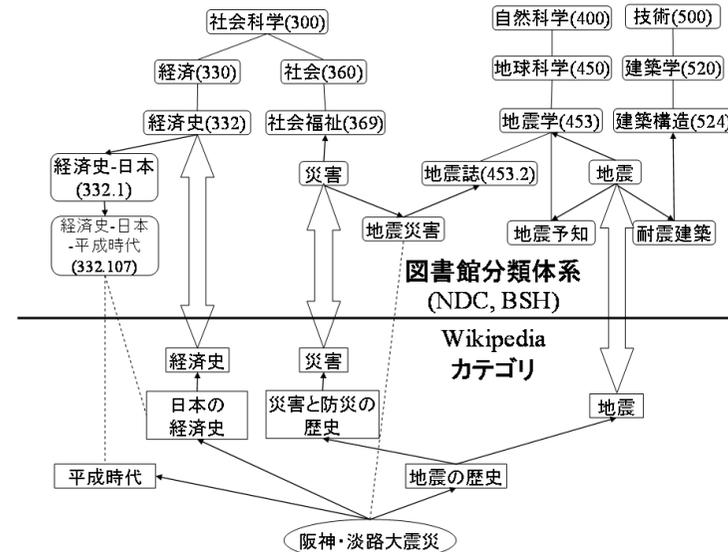


図 6 図書館分類体系と Wikipedia のカテゴリの対応付け

の切り札として注目されている」という背景を見いだすことが可能である。

#### 5. おわりに

リサーチ・ナビ検索システムの開発においては、汎用的に利用されている Web 技術を活用することによって、国立国会図書館がもつさまざまな情報資源への新たなアクセス手段を実現することを目標とした。また、Wikipedia を「情報の架け橋」として活用し、さまざまなキーワードから関連する件名を自動的に導出する機能を実装することで、情報探しのためのヒントの利用者に与えるようにした。

今後は、利用者からのフィードバックやアクセスログの分析を行うことによってシステムの有用性を評価するとともに、さらに使い勝手の向上を目指す予定である。

謝辞 Wikipedia と件名標目表を統合的に活用した分類自動導出の手法の研究に関しては、文部科学省科学研究補助金若手研究 (B)「Web 上の情報資源と図書館情報資源を統合的に利用する情報探索支援システム」(課題番号 18700134) の支援を受けた。リサーチ・ナ

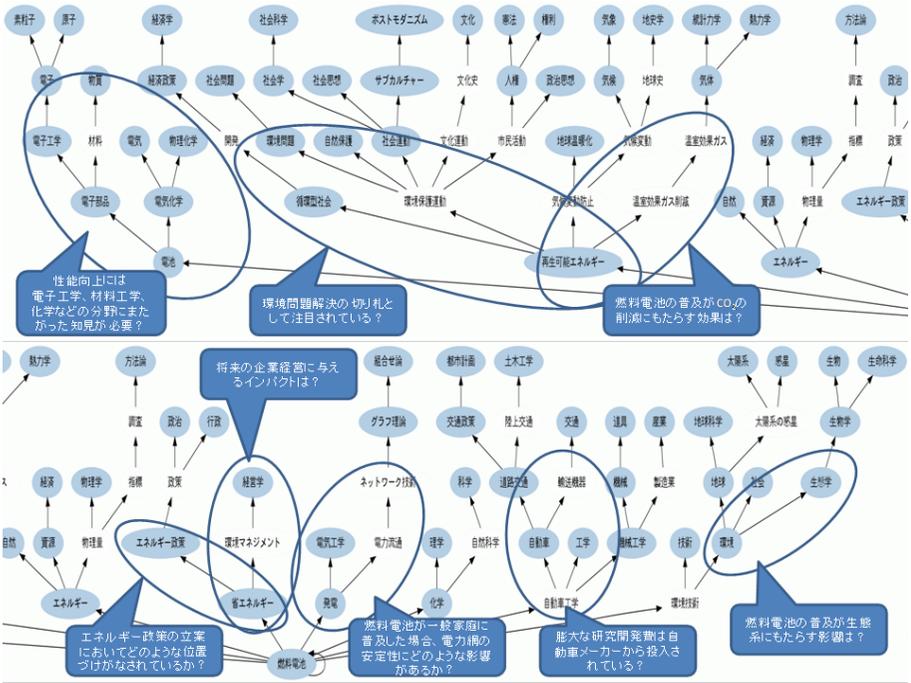


図 7 「燃料電池」のテーマグラフ

ビ検索システムの開発の過程では、国立国会図書館主題情報部参考企画課の方々によくの有益なアドバイスをいただいた。

参 考 文 献

- 1) 兼宗 進：デジタル・レファレンス・ツールとしての Wikipedia，情報の科学と技術，Vol.56, No.3, pp.103-107 (2006).
- 2) 斎藤文男：図書館利用者にとってのレファレンス・サービス，東京都図書館協会報，No.81 (2001).
- 3) Cohen, L.B. and Still, J.M.: A comparison of research university and two-year college library web sites: content, functionality, and form, *College and research libraries*, Vol.60, No.3, pp.275-289 (1999).
- 4) Wikipedia: <http://wikipedia.org/>.

- 5) Flickr: <http://www.flickr.com/>.
- 6) delicious: <http://delicious.com/>.
- 7) YouTube: <http://www.youtube.com/>.
- 8) はてな: <http://www.hatena.ne.jp/>.
- 9) 坂井 哲, 清田陽司, 増田英孝, 中川裕志: 図書館と Web の分類体系を統合的に活用したテーマグラフ可視化インタフェース, 情報処理学会第 70 回全国大会 (2008). 4ZK-9.