

図書館における自動レファレンスサービスシステムの実現 ～Web上の二次情報と図書館の一次情報の統合～

田村 悟之[†], 清田 陽司^{††}, 増田 英孝[†], 中川 裕志^{††}

[†] 東京電機大学大学院工学研究科情報メディア学専攻

^{††} 東京大学情報基盤センター図書館電子化研究部門

現在の情報検索において、Webサーチエンジンの利用は中心的な役割を果たしている。しかしながら、検索結果において必要な情報の発見が困難、信頼性に欠ける情報が少なくないなど様々な問題も存在する。この問題は、図書館の利用によって解消することができる。現在の図書館を利用する上での欠点として、提供されている情報やリソースが利用者に分かり易い形で整備されていないこと、Web上の情報と比較すると図書館の情報は即時性に欠けることなどが挙げられる。この状況を改善するため、WikipediaやWebディレクトリなどのWeb資源、図書分類体系(NDC:日本十進分類法, BSH:基本件名表目表)や情報検索メタ知識などの図書館資源を統合的に利用することにより、図書館員が行っているレファレンス応答サービスの自動化の実現を図ることを目指す。最初の取り組みとして、本研究では分類体系の統合ブラウジングインタフェースを実装した。さらに、Webディレクトリと図書館資源での分類体系統合に対し、先行研究であるカテゴリマッチングツールHICALを適用し評価を行った。

Toward the realization of an automatic reference service system for libraries ～The integration of secondary sources in the Web and primary sources in libraries～

Noriyuki Tamura[†] Youji Kiyota^{††} Hidetaka Masuda[†] Hiroshi Nakagawa^{††}

[†] School of Engineering, Tokyo Denki University, Tokyo Denki University, Tokyo

^{††} Information Technology Center, University of Tokyo

While using Web search engines plays a central role in information retrieval, the outputs of the Web search engines have several problems: for example, it is difficult to find necessary information: and there are not a few information with low reliability. These problems can be solved with using libraries. However, the information and resources provided by libraries have the following problems: those are not properly organized for ordinary users; and those are less contemporary than the information on the Web. To improve the situation, we aim at realization of automatization of reference service that a librarian provides by using library resources integratively such as information retrieval meta-knowledge, book classification system (NDC:Nippon Decimal Classification, BSH:Basic Subject Headings), and Web resource(for example, Wikipedia and Web Directory). As the first step, we implemented an integrated browsing-interface of classification systems. Furthermore, we applied a category matching method HICAL to classification system integration of the Web directory and library resources.

1 はじめに

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。現在では図書館に代わってWebが情報探索の中心的な役割を果たしつつある。これにはGoogleやYahooなどの存在が大きく影響しているといえる。一般的なWebサーチエンジンのインタフェースはどの利用者でも直感的に利用することができ、たいていの検索質問に対し、数百億ものWebページから何らかの解答となるページを見つけることが容易になっている。Wikipediaなどのオンライン百科事典や人力検索サービスを利用すると、学問の基礎的なことや人物、商品検索などの簡単な調べ

物では、ものの数分で一通りのことは調べることが可能である。このため、図書館を利用することなく調べ物を済ませてしまう人も多い。

しかしながら、情報探索の全てがWeb検索で解消できるというわけではない。例えば、個人が発信する情報には、編集者の主観、間違った情報の発信などのため信頼性、正確性に欠けるものも少なくない。さらに、現在のWebサーチエンジンの検索結果というのは、ほとんど分類されておらず検索結果同士の関連性が薄いため、情報の広がりがなく知識を深めるに至らず、学術目的の情報探索では致命的となる。一方で図書館側では、日本十進分類法(NDC)や基本件名表目表(BSH4)などの分類体系を用いて、膨

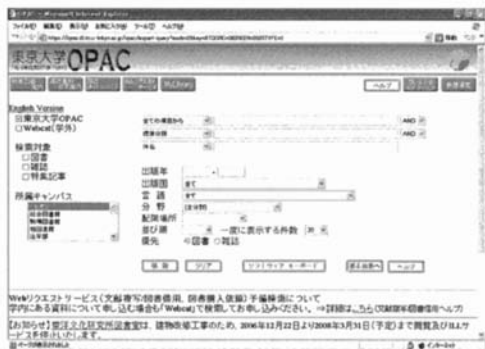


図: 1 東京大学付属図書館 OPAC インタフェース

大量の書籍、雑誌、論文などを図書館情報学の専門家が分類、組織化を行い管理しているため、信頼性、正確性を重視した情報探索が可能である。例えば、雑誌で興味を持ったことを書籍で再度調べるなどすることによって知識を深めていくことができる。

反面、利点とされる情報の専門性や組織化が情報探索初学者にとって欠点ともなってしまう。専門性が深くなると、情報探索の難しさが増し、組織化が深くなると、新しい情報の追加が困難となり即時性が低くなってしまふ。図書館には電子目録検索システムである OPAC が存在し、Web サーチエンジンと同様なキーワード検索が可能であるが、ある程度の知識が必要であり、敷居の高いものである。例えば、探したい書籍の「書名」や「著者名」が明確になっていない場合に適当なキーワードを入力してみても、検索結果から必要な書籍を見つけ出すことは難しい。このような場合は「件名」(図書館資料の統制キーワード)や「分類記号」を活用することが求められるが、「件名」や「分類記号」といった専門用語を知っている利用者は少ない(図 1)。

情報探索の信頼性の向上や情報資源の有効活用などの問題点において Web 資源と図書館資源の連携が必要とされている。そこで、我々は「図書館と Web の長短所はお互いに補完し合える関係であること」と「Wikipedia は図書館と Web の特徴を併せ持ち、双方の仲立ちになる可能性がある」ことの二つに着目し、Web 情報資源と図書館情報資源を統合した情報探索インフラの構築を目指す。その最初の取り組みとして、OPAC、Web サーチエンジンの情報探索ツールとしての問題点を改良するために、一次情報資源と二次情報資源の行き来が可能となるインタフェースを備えた分類体系統合ブラウジングツールを提案する。最終的には、図書館員が行なっている

レファレンスサービスの自動化を目標とする。これにより、Web を出発点とした情報探索を、図書館を利用した深い探索に誘導していくことが可能になり、図書館の存在意義の再発見にも繋がっていくと考えている。

本稿の構成を以下に示す。2 節において、図書館情報資源と Web 情報資源の現状について述べた後、4 節でそれぞれの情報資源における代表的な分類体系の性質を示す。4 節では自動レファレンスサービスに向けての最初の取り組みとして試作した Reference Navigator を取り上げる。5 節では、情報資源の統合のため、先行研究であるカテゴリマッチングツール HICAL を適用し評価を行う。最後にまとめを行う。

2 一次情報と二次情報の連携の必要性

近年、Web が普及し情報が爆発的に増加してきており、それに伴い我々はあらゆる情報を瞬時に獲得することができる。情報探索において Web サーチエンジンが大いに貢献していることは間違いないことであるが、出力される情報が必ずしも信頼できるものであるとは言いがたい。

以下では、実際に起きた事例を挙げる。2002 年 11 月頃、次のような話題が Web や多くのメディアによって取り上げられた。「JM 社が携帯電話のコネクタに差し込む『モブデム』というアダプタの開発に成功したと発表した。JM 社の発表によると、このモブデムを携帯電話に挿すだけで、音声 IP パケットに変換できるという。さらにこのシステムを使えば、日本中どこへでも、月額 4,500 円で携帯電話がかけほうだいになる」。その結果として、多数の個人投資家が JM 社の親会社の株式を購入し、株価は急騰した。しかし、実際には定額通話は実現せず、大型経済事件に発展したその業界や通信分野に関して予備知識のあるユーザがこの情報を見た場合、少なからず疑問点が浮かぶ。反面、まったく知識のないユーザにとって当時としては夢のような話だと感じるだろう。

この例のように、現在の Web 情報というのはどれだけ情報を多くのユーザに注目させられるかに焦点が当てられていて、信頼性の欠けることが少なくない。Web において信頼性のある情報検索を行うためには、利用者自身が当該分野に関する基礎的な知識を有している必要がある。基礎的な知識があるならば、疑問点に関して図書館資源などの信頼性の高い情報を照らし合わせてその情報が正しいものであるか確認することができるが、基礎的な知識がない場合確認することが困難である。また、Web では過

去に関する情報を得ることが難しいという問題がある。例えば、「A市における過去100年間の犯罪発生率」を調べたい場合、数年前までの統計データを発見できる可能性は高いが、過去100年間となると困難である。これらの問題が起こるのは、二次情報を多く含むWeb情報資源と、一次情報を含む図書館情報資源の連携が図れていないため、情報の信頼性を確認する手段が少なくなっているためであると考えられる。図2は、情報資源の現状のWeb情報資源と図書館情報資源の関連性の薄さを示している。Web情報資源はさまざまな分野を網羅しており、キーワードを用いて簡単に検索することができるため、情報探索スタート時の手がかりとしては最適であるが、情報の専門性・信頼性という観点では欠点を有している。一方、図書館の情報資源は探索を深めていくにつれて専門性・信頼性のある情報にたどりつくことが可能であるが、探索の端緒となる手がかりに乏しい。現状では両者の情報資源の連携がとれていないため、両者の長所を生かした情報探索を行うことは簡単ではない。

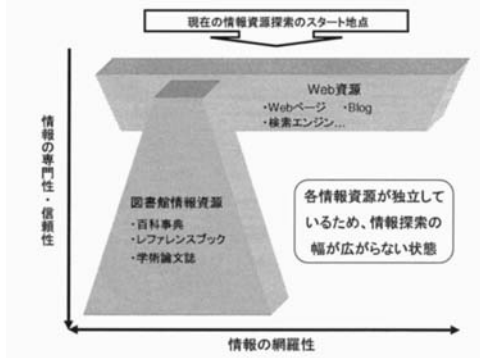


図: 2 現在の情報資源関連性モデル

一方、2003年ごろからWikipediaが急速に成長をしており、ブリタニカとの比較調査で示されたように信頼性が一般のWebページよりも高く、各情報間での組織化もされている。そこで、Wikipediaの情報の網羅性、組織化性に着目し利用することにより、Web情報資源と図書館情報資源の組織化の仲立ちをさせることによって、図書館資源を有効活用できるようになると考えている。本研究は各情報資源の統合を分類体系レベルでの対応付けとして取り組む。清田ら¹⁾は図書館情報資源とWikipediaの分類体系の統合に関する考察を述べているが、本稿では、Web情報資源とWikipediaの分類体系の統合を市瀬ら^{2) 3)}によって提案されたカテゴリマッピング

手法 HICAL によって実験的に試みる (図3参照)。

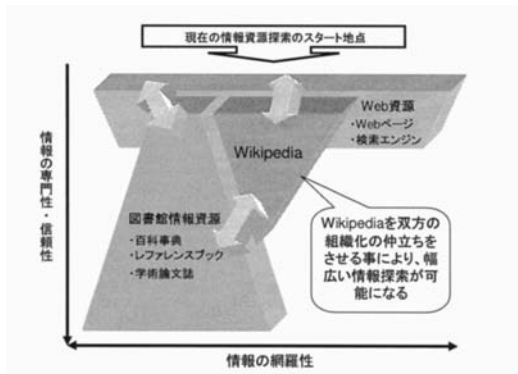


図: 3 組織化後の情報資源関連性モデル

3 各情報資源における分類体系

本節では、本研究で対象とする分類体系の特徴を述べる。表1に、各分類体系の特徴を簡潔にまとめた。

3.1 図書館分類体系

一般的に図書館情報学分野の専門家によって管理されている。例えば日本図書館協会では、委員会を組織してNDCやBSHの改定作業を行っており、現在はNDC第9版(NDC9)、BSH第4版(BSH4)が最新版である。

専門家による慎重な作業を経てトップダウン的に構築されていることから、安定的に利用することができる。深い組織化がなされているという長所がある。また、特にBSH、LCSH(米国議会図書館件名標目表)、NDLSH(国立国会図書館件名標目表)などの各種件名標目表は、情報探索に拡がりをもたせるためにきわめて有用なツールである。

3.2 Wikipedia分類体系

Web上でWikiシステムによって共同編集されているオンライン百科事典であり、誰でもどこからでも自由に編集することができる。現在、200以上の言語で作成されており、日本語版では30万項目以上の記事が存在する。

多数の人が編集に参加していることから、世界中の多くのカテゴリを網羅している。また、ハイパーリンクを容易に張ることができるというWikiの特徴を生かし、項目の組織化の取り組みも行われている。膨大な項目から効率的に一覧を生成するために「Wikipediaカテゴリ」というボトムアップ的な組織化の仕組みが導入されており、項目を多様な観点から分類することが可能となっている。

また、カテゴリ自体にも上位カテゴリ（親）のカテゴリを付与することが可能である。複数の親をもたせることができるのは、従来の分類体系にはない大きな特徴である。

3.3 OpenDirectory 分類体系

OpenDirectory（以下、ODP）に登録されているウェブサイトはカテゴリに分類され、カテゴリは担当のエディタによって管理されている。エディタをボランティアに委任するという方法はオックスフォード英語辞典での先例を World Wide Web での索引作りに取り入れたものである。

ODP は、オープンコンテンツであり、その成果は RDF の形で毎週供給され、ライセンスを遵守する限り、無料で利用することができる。言語ごとやカテゴリごとなど、自由に改変して利用することが認められている。

4 Reference Navigator

図書館にはレファレンスサービスと呼ばれる、利用者の資料・情報探索を図書館員が援助するサービスが提供されている。また、レファレンスサービスをオンライン上で利用できるサービスを DRS（デジタルレファレンスサービス）という。このサービスの利点として、図書館員が直接利用者とはやり取りを行なうため、利用者の漠然とした質問の具体化を図ることができることがあげられる。

1 節で述べたが、現在の図書館オンラインサービスにはインタフェース上の問題点以外にも、即時性や人件費などの問題点があり、それらを解決するための取り組みが行われている。京都大学附属図書館の自動レファレンスサービスシステム⁴⁾では、オンライン上で NDLSH（国立国会図書館分類法）や知識ベースを利用して典型的な質問への回答を行なう。また、ユーザとの対話を行なう機能により、質問の具体化も図っている。

本研究では図書館情報資源と Web 情報資源を統合的に扱える自動レファレンスサービスの自動化に向け、その取り組みの初期段階として一次情報資源と二次情報資源を並列的にブラウジングできる Reference Navigator の試作を行った。

4.1 基本機能

Reference Navigator の軸となる目的は、情報探索補助ツールである。そのため、Reference Navigator 単体で情報探索を完結するのではなく、外部サービスと連携して始めて情報探索が行なえる。

Reference Navigator では主に、検索対象周辺の関連カテゴリ（項目）間や、複数分類体系のカテゴリ

間を移動しながら必要な情報を探索する。その後、NDC 番号や件名を得て外部サービスへ誘導する。図 4 が内部構成図となる。現在のカテゴリマッチング処理部では、各分類体系に文字列が一致するカテゴリ名が存在する点に着目し、文字列が完全一致したカテゴリを同一カテゴリとしている。コンテンツデータ処理部では、各カテゴリに対応する情報を OPAC や AmazonAPI などの外部サービスから引き出し処理する。

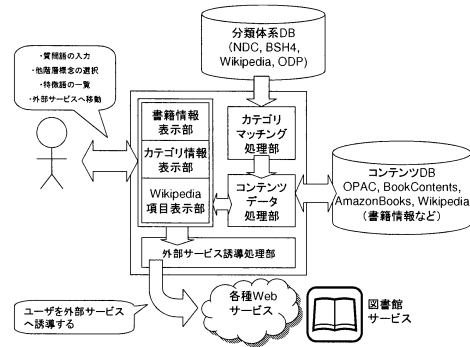


図: 4 Reference Navigator 内部構成図

4.2 画面構成

4.2.1 分類体系一覧画面

画面構成は図 5 で示すように、各分類体系を色分けし、一覧として確認できるようになっている。分類体系は Wikipedia, NDC9, BSH4, ODP のオントロジー構造を統合的に探索することが可能である。上のテキストボックスに探索主題 (Wikipedia の項目名) を入力すると、その上位カテゴリ名などがハイパーリンクとして表示される。ここでいずれかのカテゴリ名をクリックすると、さらにそのカテゴリの上位カテゴリ名、下位カテゴリ名などが表示される。同一のカテゴリ名が Wikipedia 以外にも含まれる場合には、それらの上位カテゴリ名、下位カテゴリ名も共に表示される。クリックを繰り返していくことで、情報探索を深めていき、図書館の情報資源を探すための手掛かりを得ることができる。図 5 では、カテゴリ「人工知能」の Wikipedia, BSH, Open Directory における上位・下位カテゴリ、Wikipedia において「人工知能」カテゴリに属する項目名、「人工知能」に対応する NDC 分類記号などが一覧表示されている。

表: 1 各分類体系詳細表

	図書館情報資源		中立の情報資源	Web 情報資源
分類体系	NDC	BSH	Wikipedia	OpenDirectory
カテゴリ数	10150	11184	15532	18336
親カテゴリ	1	1	複数	1+α (Symboliclink)
構造	Top-Down	Top-Down	Bottom-Up	基本は Top-Down
利点	管理が容易	管理が容易	柔軟性	柔軟性, 管理面
欠点	柔軟性に欠ける	柔軟性に欠ける	複雑さ	曖昧さ



図: 5 Reference Navigator 分類体系一覧画面

4.2.2 書籍情報詳細表示画面

分類体系一覧画面で選択された書籍情報の詳細を表示している。現在の図書館書籍検索では、主に書籍名、目次、概要などが検索結果で表示されているだけなので、利用者にとってその書籍を選択する上での手がかかりが少ない。そこで、東大 BookContents と Amazon カスタマレビューを並行して表示することで、書籍選択の幅を広げている (図 6)。

4.3 活用例

Wikipedia カテゴリと図書館の分類体系 (BSH, NDC) の対応付けと、図書館の分類体系の活用について説明する。Wikipedia カテゴリと図書館の分類



図: 6 Reference Navigator 書籍情報詳細画面

体系の間には、後述するようにカテゴリ名が一致するものが存在する。図 7 では、「経済史」「災害」「地震」が一致している。この一致を利用することにより、Wikipedia の活用によって一般化された探索主題を、図書館の分類体系によってさらに掘り下げていくことが可能となる。例えば、「経済史」「経済史-日本」→「経済史-日本-平成時代」とたどっていくと、NDC 分類記号 332.107 によって「日本経済事典」(日本経済新聞社, 1996 年)などを探することができるし、「地震」→「耐震建築」とたどっていくと、件名「耐震建築」によって「地震とマンション」(西澤英和ほか, 筑摩書房, 2000 年)などを探することができる。

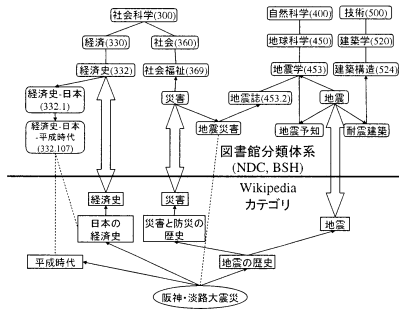


図: 7 図書館の分類体系と Wikipedia カテゴリの対応付け

5 複数分類体系の統合

Web 2.0 とよばれる枠組みの普及に伴い、様々な組織や個人により作成されたオントロジーが爆発的に増加している。代表例として Wikipedia や、各種の Folksonomy サービス (Flickr など) が挙げられる。しかし、これらの組織化された情報は独立に存在し、相互に関係を持っていない。そこで、複数のオントロジーや分類体系を統合する技術として、複数の異なる分類体系を対応付けるカテゴリマッチング (category matching) や、複数の異なるオントロジー間のクラス・プロパティを対応付けるオントロジーアライメント (ontology alignment) が必要とされる。

本研究では、図書館分類体系と Web 分類体系の統合をカテゴリマッチングの問題として取り組むこととする。

5.1 本研究とカテゴリマッチングの関係性

カテゴリマッチング手法の先行研究として、星合ら⁵⁾の統計的アプローチによるカテゴリマッチングツール SCM (Semantic Category Matching) を提案している。また、本研究で取り扱う市瀬らの研究は Web カタログに共通する Web ページの URL の一意性に着目したアプローチであり、カテゴリ基準の類似性の判定に、 κ 統計量を用いているのが特徴である。

本研究においてカテゴリマッチングを考えた場合、Wikipedia は一つのカテゴリが複数の親を持つことが許されているため、一般的なカテゴリマッチング手法の適用が困難であると考えられる。そこで、本稿では市瀬らによる HICAL を Wikipedia と OpenDirectory に対して実験的に適用し、そこから得られる結果を今後の統合手法の研究に反映させることを目的とする。

ここで、市瀬らによる実験との相違点について述べておく。HICAL では、2つのインターネットディレクトリに含まれる Web ページを基にカテゴリマッチングを行なった。しかし、本実験で扱う Wikipedia は、カテゴリ自体が外部 Web ページの情報を持っているわけではなく、項目ページに外部ページへのリンクが与えられている。そこで、各々のカテゴリに含まれる項目ページに含まれるすべての外部ページへのリンクを、そのカテゴリに含まれる Web ページのリストとして扱う (図 8)。

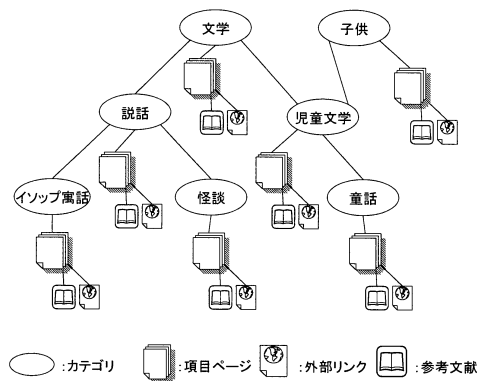


図: 8 Wikipedia における外部リンクと参考文献

5.2 評価実験

本実験では、既存のカテゴリマッチング技術をどの程度本研究に応用できるかを評価する。市瀬らの論文では一つのカテゴリに対し複数の類似カテゴリペアが作成され、その中で一番類似度の高いペア間で Web ページの移動を行っている。しかしながら、本研究では分類体系に含まれる Web ページの対応付けが目的ではなく、情報探索で利用することを目的としているため、全ての類似カテゴリペアを対象に評価を行う。なお、本実験では有意水準を 5% に設定して実験を行った。

5.2.1 実験環境

Wikipedia, OpenDirectory 共に 2006 年 7 月時点で公開されているものを利用し、その前処理としてカテゴリ名と外部 Web ページへのリンクリストの組み合わせを抽出した。実験対象とするカテゴリは以下のカテゴリとその下位に属するカテゴリとし、詳細は表 2 に示す。

- Wikipedia : 芸術, 学問, 文書作成, 人文科学, 文化/文学
ODP : Top/World/Japanese/アート/文学

表: 2 実験に利用したデータ

	Wikipedia		OpenDirectory		共有 Web ページ
	カテゴリ数	Web ページ数	カテゴリ数	Web ページ数	
文学	536	13224	290	2620	1474
ゲーム	180	6360	243	1998	538
宗教	1311	13652	335	2884	883

表: 3 実験結果

		類似カテゴリ数 (カバー率)	
		Wikipedia	OpenDirectory
文学	標準	43(8.0%)	24(8.3%)
	子カテゴリまで許容	196(36.6%)	252(86.9%)
	孫カテゴリまで許容	365(68.1%)	271(93.4%)
ゲーム	標準	36(20.0%)	23(9.4%)
	子カテゴリまで許容	124(68.9%)	140(57.6%)
	孫カテゴリまで許容	148(82.2%)	219(90.1%)
宗教	標準	66(5.0%)	38(11.3%)
	子カテゴリまで許容	346(26.4%)	162(48.4%)
	孫カテゴリまで許容	539(41.1%)	212(63.3%)

- Wikipedia : 娯楽/ゲーム
ODP : Top/World/Japanese/ゲーム
- Wikipedia : 人の行動, 倫理, 人の一生, 生活, 文化, 思想, 心/宗教
ODP : Top/World/Japanese/社会/宗教・精神世界

5.2.2 結果と考察

表3に, HICALのアルゴリズムが類似していると判定したカテゴリの個数と全体に対する割合を示している. また, カバー率はHICALによって対応づけられたカテゴリ数の全体(実験対象となっているもの)に占める割合を示す.「標準」は, HICALが類似カテゴリと判定した個数である.「子カテゴリまで許容」は, HICALが類似カテゴリと判定したカテゴリが持つ下位カテゴリも含めて対応できたとみなした場合の個数である.例えば, Wikipediaの「文学/詩/詩人」とODPの「Japanese/アート/文学/詩/詩人」の類似カテゴリペアがあったとき, Wikipediaの下位カテゴリの「文学/詩/詩人/象徴派の詩人」も類似カテゴリとする.「孫カテゴリまで許容」も同様に, 2階層下までの下位カテゴリを考慮した結果を

示している.

「ゲーム」のマッチング結果では, 副カテゴリが他のカテゴリと比較すると少なく, また, 共通Webページも多い. そのため, 類似カテゴリと判定された個数が多くなっている. 反対に, 「宗教」では, 全体の副カテゴリも多く, 共通Webページも少ないので, 類似カテゴリも少ない. 結果, 「宗教」は分類体系全体が深く類似カテゴリの許容範囲を広げても, カバーできていないことがわかる. しかしながら, 出力された類似カテゴリペアの全てが, 本研究の目的とする情報探索支援において利用可能かどうかを示すものではない.

そこで, HICALの出力結果である類似カテゴリペアの質を手で判定することとする. この評価で対象とする類似カテゴリは許容範囲を広げたものではない(すなわち, 表3における「標準」が評価対象となる). また, 対応付けた分類体系を, 今後Reference Navigatorで利用するため, 対応付けたカテゴリ間を違和感なく移動できるかを想定して判定を行う. 判定基準は類似カテゴリペアをラベル上で比較してもらい, 0から2の3段階で行い, 「2」は違和感なく分類体系間の移動が可能, 「1」は若干カテゴリ間の

表: 4 評価結果

	文学			ゲーム			宗教		
類似カテゴリペア数	183			93			171		
判定基準	0	1	2	0	1	2	0	1	2
比率	30%	39%	31%	22%	49%	29%	23%	39%	38%

差があるが移動可能,「0」はカテゴリの大きさにあまりに差があるため移動不可とする。なお,本実験の評価では被験者の持っている知識や主観が大きく影響されるため,工学系大学生5人に判定を行ってもらった。

結果を表4に示す通り,「0」と判定された割合は全体の3割弱となっていた。判定内容を確認してみると,Wikipediaの書物,出版/出版物」とODPの「Japanese/アート/文学」の類似カテゴリペアのように明らかにカテゴリの大きさが異なるペアが「0」と判定されている。また,そのほかには「宗教,アブラハムの宗教/ユダヤ教」と「Japanese/社会/宗教・精神世界/キリスト教」のようにラベル面では異なるが関連性の強いペアに関しても「0」と判定されていることがわかった。

この実験では,類似カテゴリペアのラベルのみを比較対象に評価を行った。しかしながら,ラベル面だけではユーザの知識によって,類似カテゴリを提示しても利用価値がなくなってしまう場合もある。現在のオントロジーアライメントやカテゴリーマッチングでは,評価方法が提案手法ごとに乱立しており,評価方法の統一が必要となっている。Avesani⁶⁾らは,3つのカテゴリーマッチング技術の評価手法として,同一のデータセットを用意し評価を行っている。本研究においても,分類体系の対応付けに関しては評価をよく考慮して進めていかなければならない。

6 おわりに

本稿では,Web情報探索が抱える現状の問題点を述べ,それらの解決手法として,Wikipediaの情報の信頼性,組織化に着目し,Wikipediaを図書館情報資源とWeb情報資源の仲立ちをさせることを提案した。そして,本研究の初期段階の取り組みとして情報探索補助ツールReference Navigatorの実装を行った。また,分類体系の対応付けの取り組みとして,先行研究であるカテゴリーマッチングツールHICALを実験的にWikipediaカテゴリとODPカテゴリに適用し評価を行い,既存のカテゴリマッ

ング技術が本研究において適用可能であることを確認した。

今後は,本研究の目的に合わせた分類体系の対応付けの提案とReference Navigatorのインタフェースの改良が必要となってくる。

参考文献

- 1) 清田陽司, 田村悟之, 中川裕志, 増田英孝. Reference Navigator: 異種オントロジーの統合ブラウジングツール～図書館の分類体系とWikipediaカテゴリの対応付け～. 言語処理学会第13回年次大会ワークショップ「言語的オントロジーの構築・連携・利用」論文集 (2007, to appear)
- 2) 市瀬龍太郎, 武田英明, 本位田真一. 階層的知識間の調整規則の学習. 人工知能学会論文誌, vol.17, No.3, pp.230-238 (2002)
- 3) 市瀬龍太郎, 濱崎雅弘, 武田英明. 階層的分類データを統合するための規則学習機構 人工知能学会論文誌, vol.19, No.6, pp.521-529 (2004)
- 4) 黒橋禎夫, 日笠亘. 京都大学附属図書館における自動レファレンス・サービス・システム. 情報管理, Vol.47, No.1, pp.184-189, (2001)
- 5) 星合 忠, 山根 康男, 津田 宏 カテゴリマッチング技術に基づくオントロジーアラインメント問題への取り組み 人工知能学会論文誌, Vol.20, No.6, pp.437-447 (2005)
- 6) Avesani, Paolo and Giunchiglia, Fausto and Yatskevich, Mikalai A Large Scale Taxonomy Mapping Evaluation. Technical Report DIT-05-044 (2005)