

令和元年6月28日現在

機関番号：34310

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K00450

研究課題名(和文)人の真の情報ニーズを汲み取るコンシェルジュ型資料検索システムの構築

研究課題名(英文) Development of Concierge-type Material Search System to Capture People's True Information Needs

研究代表者

原田 隆史 (HARADA, Takashi)

同志社大学・免許資格課程センター・教授

研究者番号：30218648

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：現在図書館で使われている本の配架場所は、知識を体系化した日本十進分類表などをもとに行われる。しかし、実際に行われる資料の利用は学術的なニーズのみに基づいて行われているわけではない。そこで本研究では、図書館での資料の利用状況や人々の図書から受ける読後感などを元にして人々の真の情報要求を把握することを試みるとともに、Wikipedia中の記述をもとにして検索式を拡張したり類似資料を同時に提示するなどが可能な資料提供システムを試作して評価した。その結果、検索式の自動拡張や関連する資料の提示の有効性が明らかになるとともに、利用者の持つ知識や検索経験の影響が大きいことも明らかとなった。

研究成果の学術的意義や社会的意義

本研究の成果は、検索語との対応関係のみに基づいて資料を提示する従来の図書館OPACや、学問としての知の体系のみに基づく図書館の書架のブラウジングでは実現できない利用者要求を満たすものであり、人手によるパスファインダーやレファレンス・サービスの欠点をシステム化によって補うための基礎的な材料になりうるものと考えられる。また、同時に図書館だけではなく書店や類似施設での資料配置にもつながるものと考えられる。

研究成果の概要(英文)：The location of books currently used in libraries is based on the Japanese Decimal Classification system to organize knowledge. However, the actual use of materials is not based solely on academic needs. Therefore, in this research, we try to capture people's true information needs based on the usage situation of the materials in the library and the people's feeling after reading. In addition, we develop and evaluate a material provision system which can extend search formulas and present similar materials simultaneously based on the descriptions in Wikipedia. As a result, it become clear that the automatic expansion of the search formulas and the presentation of the related materials are effective, and the influence of the user's knowledge and search experiences are significant.

研究分野：図書館情報学

キーワード：図書館 OPAC 情報要求

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

図書館では、書架の配置を工夫することで体系化した知識を表示することが行われており、ブラウジングなどの行為を通じて利用者はこの効果を楽しんでいる。これら図書館の書架が礎とする分類は、学問としての知の体系に基づいて構築されたものである。しかし人は学術的なニーズのみに基づいて資料を検索するわけではない。例えば「フランス料理」は日本十進分類法では 592.23 (西洋料理) に分類され、書架ではその近くにイギリスやイタリアの料理法やレシピに関する資料が並ぶ。西洋料理の研究者のニーズはこの資料提示法で満たすことができる。しかし「フランス料理店を新たに开店したい」という利用者であれば、あわせて提示すべきは店舗経営や内装等に関する資料である。あるいは「フランス料理店でプロポーズをしたい」利用者に対しては、マナーやプロポーズ指南、デートコース案内、宝飾品やワインに関する資料が必要となる。これらのニーズは、既存の分類に基づく資料提示法では満たすことができない。これは従来の検索システムの検索対象は資料に含まれる語や資料の主題であって、その資料に対しあり得る利用者のニーズを検索することはできなかったためである。

このような、分類でも検索システムでも満たせない、複数の主題が関わる資料を通覧するニーズを満たす役割は、現在はパスファインダーやレファレンス・サービスが果たしている。しかし、パスファインダーはあらかじめ用意したニーズにしか対応できず、レファレンス・サービスは、特に込み入ったニーズに対応する場合には来館して相談する必要がある。そしてそのいずれについても、作成あるいは回答する図書館員によって対応できないニーズ、不得意な分野が存在する。これらに対応した仕組みづくりが求められる。

2. 研究の目的

本研究では図書館利用者の探索行動を検索ログ・貸出記録を元に分析し、入力された検索語から利用者の真の情報ニーズを読み取って結果を提示するという、いわば「コンシェルジュ型資料検索システム」と呼ぶべきシステムを開発するための基礎的研究を行った。このようなシステムは、検索語との対応関係のみに基づいて資料を提示する従来の OPAC や、学問としての知の体系のみに基づく図書館の書架のブラウジングでは実現できない利用者要求を満たすものであり、人手によるパスファインダーやレファレンス・サービスの欠点をシステム化によって補うものともなりうると思われる。

3. 研究の方法

(1) 図書館資料は目的に対応して、どのような組み合わせで利用されているのかの調査

図書館における資料の利用行動について、特にあるニーズに対して資料を探索する際の一連の行動および閲覧資料との対応関係を調査する。具体的には、図書館で同時に貸し出される図書がどのようなものであるのかについて調査を行う。同時に貸し出される図書が同じ分類番号でない場合には、どのような組み合わせで貸し出されるのかを明らかにすることで、学問上の主題が異なる図書の組み合わせを探し出すことができると期待される。

(2) 利用者が図書から受ける感情についての調査

図書から人々が受ける印象は、図書の内容に大きく影響されると考えられる。しかし、図書の内容を示す文章が必ずしも利用者に影響される情報を全て含んでいるとは限らず、図書館 OPAC 上で表示される内容を示す文章を元にした推薦を行っても利用者のニーズとはかけ離れるケースが考えられる。また、図書から受ける印象をより反映するものとしてオンライン書評が存在するが、書評の作成者と利用者との間にズレが生じてしまい、実際には有効ではないことも想定される。そこで、小説を対象して図書の内容を示す文章を示す文章と読後感との関係を調査するとともに、オンライン書評の内容と当該図書を持った読者が付与した印象とが一致するかどうかについて、感性パラメータを付与してもらうことで検証する。

(3) 検索語として用いられる語の拡張

検索語の拡張については、NDL サーチにおける件名標目表の利用や、Webcat Plus における Wikipedia 中の語の利用という例もあるが、前者は対象となる語の数が少なすぎる、後者は図書と関係のない語間の関係の影響が大きすぎるなどの問題がある。本研究では、図書に含まれる語の重みを重視した新たな手法の開発のための調査を行う。具体的には国立国会図書館の全蔵書データおよびテキスト化済の目次・索引データを対象として語の共出現および出現箇所を表示し、図書館の利用者に論理和で検索語を拡張するのに使用することに有効であるかどうかインタビュー調査を行い検証する。

さらに、検索語の拡張を行ったシステムを構築するためには、実際にどのような語が拡張する対象として適切であるかについての検証が必要である。そこで、オンライン書店で公開されている図書の内容紹介文および書評中の語を対象として、どのような語が検索語の拡張に使用できるか、どのようにして語を抽出することが適切かについて調査を行う。

(4) 多様な検索手法を用いる図書館システムの有効性の検討

近年、図書館においては従来の OPAC とは異なり操作性の向上、検索対象範囲の拡大ならびに検索結果の表示内容の充実をはかった「ディスカバリーサービス」が導入されるようになって

できた。このディスカバリーツールを実際に大学生に利用してもらい、その有効性と問題点についてアンケートおよびインタビュー調査を行って分析する。

(5) 検索結果として提示する結果の拡張

図書館 OPAC の検索結果として提示される資料は、あくまで検索語と関連するものであって、利用者ニーズそのものに対応するものではない。そこで、図書館 OPAC の閲覧履歴および図書館の貸出履歴の分析から、新たに追加すべき資料を提示する仕組みを提案する。分析の視点としては共時性と通時性の両面に着目する。共時性の面では、たとえば同時に貸出がなされた図書の中で利用者の利用目的が共通であると考えられる図書の組み合わせの抽出・組み合わせられ方のパターン分析などを行う。また、通時性の面では複数の利用者が特定の資料を共通して閲覧する組み合わせを用いて協調フィルタリングの仕組みを用いた図書の推薦が可能な仕組みを構築し、その有効性を検証する。このような図書推薦の仕組みとしては、オンライン図書販売サイトである Amazon も備えている。そこで、本システムで推薦される図書と Amazon が推薦する図書の違いについても調査する。

4. 研究成果

(1) 図書館における同時貸し出し図書の分析

図書館で同時に多数の図書を貸し出す場合の組み合わせとして日本十進分類表の 1000 区分での共通性を調査した。具体的には大学図書館において 9 冊の図書が同時に貸出されている場合を対象として共通性を調べたところ、半数以上の図書が同じ分類番号である例は全体の 53.3% であり、残りについては異なる分類番号の図書が借りられる結果であった。また、借りられた 9 冊の図書が 6 つ以上の分類番号にまたがってパラバラである割合は 25.3% と高い結果となった。また 6 つ以上の分類番号にまたがる図書が貸し出されたうちの 31.5% はタイトル中に同じ語が含まれていた。この結果から、類似の図書として主題だけではなくタイトル中の語の共通性が利用者の情報ニーズにあった提供に効果があることが示唆される。

(2) 書評中にみられる感情表現と読後感の分析

書評中にみられる形容詞および名詞と読後感との関係を調べたところ、形容詞については読後感と一致するものが多いものの、表現は多様であり、また書評中には表現されない読後感も多いことが明らかとなった。また、名詞についてはさらに多様な表現が使われることが多く、そのまま機械化の基礎データとして使用するの難しいことが明らかとなった。さらに、書評中では相反する感情を示す形容詞などが同時に使われることも多く、どれが内容と関わるのか等についてのより詳細な分析が必要であることが明らかとなった。

(3) 書評中の単語を元にしたパラメータと読者の付与したパラメータとの比較

小説を対象として読者が読後に受ける印象を示すパラメータ(感性パラメータ)の付与を行うにあたり、人手でパラメータ付与を行う場合と書評中にみられる形容詞および名詞を基に機械学習によって付与するとを比較した。その結果、両者が完全に一致する割合は 32.2% にとどまり、プラスマイナス 1 までを正解とみなした場合でも 59.8% であった。これは、人手で付与する場合には極端な評価は行われない傾向が強いことが大きく影響していると考えられる。実際に、0~5 という 6 段階評価の 0 や 5 はほとんど付与されず、1 や 4 もそれほど多くなかった。パラメータの種類別では「興奮した - 冷めた」「温かい - 孤独な」「和む - 辛い」では比較的一致する傾向が見られ、「驚く - 平凡な」「明るい - 暗い」「爽やか - どんより」についてもある程度的一致が見られたが、それ以外のパラメータでは一致する割合が低い結果となった。

(4) Wikipedia や図書のタイトルなどを元にした検索語の拡張

検索語として入力された言葉だけでは情報要求に十分に答えられないことを想定し、Wikipedia や図書のタイトルなどで検索語と共出現回数の多い語を論理和で接続した検索を行って評価した。その結果、Wikipedia を元にした語については情報要求とは異なるもものがほとんどであったのに対して、図書のタイトル中で共出現したものについては関連するものも多く、検索語拡張の候補とできる可能性があることが明らかとなった。ただし、図書のタイトル中の語についてもノイズは極めて多く、実際のシステムへの応用については課題が多いことが明らかとなった。

(5) 料理用語を対象とした検索語追加候補の検討

検索語の拡張を目的として、料理に関わる語を対象にオンライン書店で公開されている図書の内容紹介文および書評中の語を抽出し、それらの語が検索語の拡張に適切かどうかを検討した。具体的には Amazon に掲載された図書の案内やオンライン書評中に含まれる料理に関わる語を抽出して大学生に提示し、それらの語が検索時に追加されることが効果的かどうかを判定してもらった。その際、語の抽出方法としては n-gram と形態素解析の両方で行って両者の比較も行った。その結果、いずれの手法を用いた場合も 8 割以上の語が検索ノイズの原因であり、それ以外のほとんどの語についても追加することで大きな効果はあがらないという結果が得られた。検索語の抽出については、単純に出現する語を選択するのではない手法の開発が望まれ

る結果となった。特に n-gram を用いた場合には「ガレットとマーガレット」、「ドリアとエイドリアン」等のように部分一致の問題でノイズが多く発生する例が見られた。語の抽出時には n-gram ではなく形態素解析を用いる必要があることが示唆されたと考えられる。

(6) ディスカバリーツールに関する評価

従来型 OPAC と高度な機能を備えたディスカバリーツールを被験者が同時に使用・比較する評価実験を行った。100 名超の被験者（大学生）に対し、6 問の課題を課し、いずれも OPAC とディスカバリーツールの双方を用いて回答を作成するよう依頼した。その結果、ディスカバリーツールについて、ツール自体の使用頻度（習熟度）よりも本人の年齢（学年）が正答率に影響することが示された。ツール単体の習熟度よりも本人の経験・知識がツールの利用方法に影響したことから、ツールのみで機能改善を測るのではなく文脈情報を加える本研究課題の有益性が改めて示唆された。

(7) 図書館における同時貸出と NDC を用いた図書の推薦，Amazon の推薦

図書館で同時に多数の図書を貸し出す場合には関連する図書が借りられることが多い。しかし、すべての同時貸出が関連するとは限らないことも確かである。そこで、本研究では同時貸出データと NDC とを組み合わせることで同時貸出だけを使用する場合に混入するノイズの低減を試みた。その結果、単純に同時貸出データと NDC とを組み合わせただけの場合、必要な図書が取り除かれてしまうという負の面が多いことが明らかとなった。しかし同時に、全くジャンルの関係がない NDC が存在することも見いだすことができた。このことは、同時貸出データと NDC との組み合わせにおいては、「検索語 NOT 関連がほとんどない NDC」のように関連がない主題の図書を取り除くという手法が有効である可能性を示唆していると考えられる。

また、図書館における利用データを元にして協調フィルタリングに図書の推薦を行う場合と Amazon の推薦を比較したところ、Amazon が推薦する図書の方が利用者が興味を持つ比率は高かった。ただし、Amazon が推薦する図書と貸出記録を元にして推薦される図書とは一致しないことが多く、図書館の貸出記録を元にした推薦図書の方が Amazon よりも広い範囲の図書が推薦されると利用者が感じる人が多いことが明らかとなった。このことは、図書館の貸出記録が Amazon とは異なる図書の推薦の基礎データとなりえる可能性を示していると考えられる。ただし、このことは全く関係ない図書の利用も多いことをも示しており、より詳細な分析が必要であることが明らかとなった。

5. 主な発表論文等

〔雑誌論文〕(計 7 件)

佐藤 翔, 楠本 千紘, 服部 亮ほか. 日本の大学生は情報源が Wikipedia 日本語版である情報の信憑性を他のオンライン百科事典である情報よりも低く判断する. 情報知識学会誌. 査読有. 28 巻, 2018, 223-252

DOI: https://doi.org/10.2964/jsik_2018_023

Shoko Nakahata, Emiko Sakamoto, Akiho Oda, Noriko Kobata and Sho Sato. Effects of color of book cover and typeface of title and author name on gaze duration and choice behavior for books: Evidence from an eye-tracking experiment. Proceedings of the Annual Meeting of the Association for Information Science and Technology. 査読有. 53 巻, 2016, 1-4
DOI : 10.1002/pra2.2016.14505301100

〔学会発表〕(計 9 件)

Yukinori Okabe, Takashi Harada, Sho Sato et al. The development of the book report creation support game. IFLA 84th World Library and Information Congress. 2018 年

佐藤 翔, 池本 実緒, 小池 敬大ほか. 公共図書館内における利用者の注視行動の傾向と図書館デザインの影響. 日本図書館情報学会 2018 年度春季研究集会, 2018 年

Takashi Harada, Sho Sato, Sakura Yasuda et al, Development of the reading guidance system & Growing game in accordance with impressions after reading. IFLA World Library and Information Congress 83rd IFLA General Conference and Assembly. 2017 年

6. 研究組織

(1) 研究分担者

研究分担者氏名: 兼宗 進

ローマ字氏名: Susumu KANEMUNE

所属研究機関名: 大阪電気通信大学

部局名: 工学部

職名: 教授

研究者番号(8桁): 00377045

研究分担者氏名： 逸村 裕
ローマ字氏名： Hiroshi ITSUMURA
所属研究機関名： 筑波大学
部局名： 図書館情報メディア系
職名： 教授
研究者番号(8桁)： 50232418

研究分担者氏名： 宇陀 則彦
ローマ字氏名： Norihiko UDA
所属研究機関名： 筑波大学
部局名： 図書館情報メディア系
職名： 准教授
研究者番号(8桁)： 50232418

研究分担者氏名： 岡部 晋典
ローマ字氏名： Yukinori OKABE
所属研究機関名： 愛知淑徳大学
部局名： 人間情報学部
職名： 講師
研究者番号(8桁)： 60584555

研究分担者氏名： 佐藤 翔
ローマ字氏名： Sho SATO
所属研究機関名： 同志社大学
部局名： 免許資格課程センター
職名： 准教授
研究者番号(8桁)： 90707168

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。