

OPAC 利用ログを用いた
文献検索システムに関する研究

筑波大学

図書館情報メディア研究科

2018 年 3 月

小幡 将司

目次

1. はじめに	1
1. 1 研究の背景	1
1. 2 研究目的.....	4
1. 3 論文の構成	5
2. 関連研究	6
2. 1 図書館分野以外における利用者の情報を利用した関連研究.....	6
2. 2 図書館における利用者の情報を用いた関連研究.....	8
2. 2. 1 現実での図書館利用の分析.....	8
2. 2. 2 現実での図書館利用によるシステム.....	8
2. 2. 3 ウェブ上における行動を利用した分析	9
2. 2. 4 ウェブ上の行動を用いたシステム	10
2. 3 他の学生の学習経験を反映した学習支援.....	10
3. 関連データ分析.....	12
3. 1 対象データ	12
3. 2 分析方法.....	12
3. 3 分析結果.....	15
3. 4 レファレンス分析.....	17
4. 研究手法	20
4. 1 概要	20
4. 2 リランキングについて	20
4. 3 システム全体像	22
4. 4 関連科目検索システム	26
4. 5 開講期間重み付けシステム	27
4. 6 利用ログ検索システム	28
4. 7 リランキングシステム	28
4. 8 システム全体のスコア計算	29

4. 9	実装システム	31
5.	結果	33
5. 1	関連科目検索の検証	33
5. 2	月ごとの重み付けの検証	35
5. 3	利用ログ抽出の検証	39
5. 4	リランキングの検証	39
6.	考察	42
6. 1	関連科目の検索	42
6. 2	月ごとの重み付け	44
6. 3	リランキングの検証	44
6. 4	提案手法の限界	45
6. 5	評価手法の限界	46
7.	おわりに	48
	謝辞	50
	参考文献	51
	付録	57

図目次

図 1:利用者行動の履歴を用いる関連研究の分類	6
図 2:利用ログの構造	12
図 3:年間における月ごとのアクセス回数.....	15
図 4:システムへのアクセスに用いられた情報の分析	16
図 5:アクセスされた文献のアクセス回数の分析	17
図 6:アイディアの概要図	22
図 7:システム概要図.....	23
図 8:システム構成図	24

表目次

表 1:除去するログの条件	14
表 2:キーワードに含まれていた単語上位10件.....	18
表 3:リランキング前の例	21
表 4:リランキング結果の例.....	21
表 5:学期と月の対応	26
表 6:リランキングに用いるスコア	29
表 7:語句解析でクエリを含んでいた際の検索結果	33
表 8:語句解析でクエリを完全一致でクエリを含んでいた際の検索結果.....	34
表 9:年間を通じた 5.1 節でヒットした科目による重み付け	35
表 10:質問を行った月を開講期間に含む重み付け	36
表 11:質問を行った月のみによる重み付け.....	37
表 12:質問を行った月±1月の期間による重み付け.....	37
表 13:質問を行った月±2月の期間による重み付け.....	38
表 14:検索結果の適合結果	40
表 15: nDCG による検索評価結果	41
表 16:関連科目の項目を記載していた科目.....	42

1. はじめに

1. 1 研究の背景

大学図書館とは大学において学生の学習・研究を支援するために設立された組織であり、大学に必須の施設である(1)。文部科学省の位置づけでは「大学図書館は、大学における学生の学習や大学が行う高等教育及び学術研究活動全般を支える重要な学術情報基盤の役割を有しており、大学の教育研究にとって不可欠な中核を成し、総合的な機能を担う機関の一つである(2)」とされており、大学図書館はこの役割を果たすために資料および場所の提供、サービスの展開を行っている。その一例として、筑波大学附属図書館においても筑波大学の理念と目標の実現のために、学術情報のアクセス環境を提供することや、快適な施設・設備・支援サービスを提供することなどにより学生の学習を支援することで、学術情報基盤としての機能を果たすことを使命としている(3)。

大学図書館は学術情報の基盤としての役割を果たすために資料や場所の提供のほか、学生の学習・研究を支援するために様々なサービスを展開している。企業はコミュニケーション能力の向上や社会活動体験の増加と言ったものを大学の役割として期待しており、アクティブラーニングの導入や留学やインターンシップ、ボランティアの増加を求めている(4)。その一方で、国はこれからの時代に求められる資質・能力として、主体的に課題に取り組み解決に導く力や未知の課題に挑み解決策を見出すための力、人に対して働きかけ他者と共存していくための力という3つの項目を挙げ、これらの力が極めて重要だとしている(5)。その上で、国はこれからの教育方針として小・中・高等学校から大学までを通じたアクティブラーニングを推進やICTを用いた学習環境の革新、新たな価値観を生み出すための人材育成、特に優れた才能の人材の発掘といったことを指針としている(5)。こういった背景や情報技術の進展などにより、大学図書館は学修環境充実に係る学術情報基盤整備についてコンテンツ・学習空間・人的支援の3要因に整理される要素が存在し、それらの連携を図ることが重要とされている(6)。そこから、図書館の所蔵資料を迅速に利用できるようにする

こととともに、自主的な環境を行うための場であるラーニング commons の整備・連携や学習支援・レファレンスサービス・指導助言などによる学生の支援体制の構築といったことが求められている (6)。大学図書館のサービスは大学における学習・研究上の課題を解決するための手段を提供し、学生が自発的に学ぶための技能を身につけられることを目指している。大学図書館が行う活動では学習環境の整備やラーニング commons を始めとして学生チューターなどの学生が参加し教え合ったりともに学んだりすることやデータベースの利用方法の講座なども存在し、学習・研究について様々な情報により学生を支えている。一例として、筑波大学附属図書館においては、学習支援のサービスとして全学計算機や学生チューターを配備したラーニング commons の設置や大学において学ぶ際に必要な論文を書くための技能や資料を探すための技能などを習得するセミナーの開催など様々なことを展開している (7) (8)。

大学図書館は、場所としての機能および提供しているサービスを、学生の利用に供しており、設立の目的を果たそうとしている。谷らのイェール大学図書館およびマサチューセッツ大学アマースト校大学図書館における調査では、学生は研究テーマに関する資料のブラウジングや授業内容に関する資料を入手するために利用するなど、大学図書館が学生にとって学習する場所の選択肢の一つとして機能している (9)。また、立石の研究では横浜国立中央図書館の大学図書館の利用者の行動内容の分析において、様々な行動の中で学習のために利用している割合が最も多く 45.5% を占めており、図書館が学習のための場所として個人やグループなど利用者にあった形で利用されているとしていた (10)。実際に大学図書館の利用者にどのような目的で利用するかに関して、古橋の調査によれば、東京外国語大学図書館での学習に集中できる場所であるとして資料の検索や学習場所としての利用など学習目的での利用が多いという結果が出ている (11)。以上のことから、大学図書館における資料の入手への支援を始めとする学習支援の機能は学生にとって重要な機能であると考えられる。

大学図書館では様々な観点で利用者にサービスを展開しているが、本研究では文献検索サービスに着目する。大学図書館において情報サービスには利用者からの情報要求に直接対応する直接サービスと情報源やサービスに必要なツールを提供する間接サ

ービスの2種類が存在している(12)。情報サービスとしての文献検索に関わる領域における直接サービスは文献情報に関して図書館に依頼するレファレンスサービスの利用が当てはまり、間接サービスはOPAC（Online Public Access Catalog）や機関リポジトリなどのテクニカルサービスが当てはまる（12）。本研究で着目する文献検索サービスは間接サービスに当たるサービスとする。図書館において文献検索サービスを用いる場合に想定される利用方法としては、利用者が資料を調べる際に自分でクエリを考え検索を行っていくやり方の他に、同級生や先輩、先生にアドバイスを貰ってから資料を調べるといったやり方が考えられる。

本研究では大学図書館が展開している文献検索サービスの中でも特にOPACに着目する。OPACは大学図書館においてオンラインによる対話方式で検索を行い、利用者が直接端末機を操作し、検索したい事柄にあった検索方法を用いて所蔵する資料を検索できるよう設計されたコンピュータ化された閲覧目録のシステム（13）であり、利用者はそこから文献に関する書誌情報や所蔵情報を得ることができる。大学図書館においてOPACは資料の所蔵情報を調べることが可能であるため、図書館を利用して資料を検索する際に所蔵資料を実際に探す際の手がかりとなっている。

OPACは大学図書館の所蔵情報を扱い、学生が学習・研究の際に探している文献のありかを知るために利用することができる。大学図書館は学習・研究のための学術情報基盤としての資料を所蔵しており、OPACは利用者が学習・研究のために大学図書館の資料を検索する際に所蔵情報を確認するために利用できる。実際のOPACの利用のされ方として、筑波大学附属図書館においては図書館が提供する資料検索のアドバイスにおいてOPACの利用を最初に薦めていること（14）や姫路獨協大学附属図書館においてもOPACの利用方法を資料検索の基礎的な知識として挙げている（15）。また、上岡らは情報検索の入門として大学において文献検索を行う調査過程の説明の際に真っ先にOPACに触れていた（16）。さらに、共通科目として図書館の利用方法の科目である「大学図書館の活用と情報探索」（17）を開設している京都大学において参考図書として用いられている講義録では文献を検索するツールとしてOPACを用いるツールとして最初に取り上げ、館内の特定の資料を探す際に用いるとしていた

(18)。大学図書館において OPAC を利用し資料を検索することは基礎的なことであり、資料検索の足がかりとなっている。そのため、大学図書館を内の資料を検索しようと思いついたときにアクセスしやすいサービスとして利用されることが考えられ、利用者が図書館の文献を検索する際に OPAC を最初に利用するケースが考えられる。

しかし、いくつかの研究において OPAC はそれ単独ではあまり利用せず (19)、OPAC 単独の場合には図書館において目的の文献を検索する際に利用者が不便に感じている面が存在している (20) (21) という指摘がある。安蒜らの研究によれば大学生に対してレポート課題が与えられた際の情報収集の際にウェブを利用し、学年を問わず 7 割以上の回答者が OPAC を利用していたが情報検索の際に最初に用いられる検索ツールは Google が最も多くなっており、OPAC の利用もウェブでの情報探索と交互に情報を確認していく形であるとしていた (19)。工藤らは OPAC での資料検索には調べ方に専門的な知識が必要であり、ハードルがあることを指摘していた (20)。その一方で、種市らの研究では OPAC で検索を行う際にはウェブサーチエンジンでの経験が情報探索行動に対し影響を与えており、OPAC における視覚的要素とどのような検索経験があるかという経験的要素で検索は行われ、OPAC で検索を行う際にサーチエンジンのように検索できないと戸惑ってしまうことが言及されていた (21)。このようなことから大学図書館において OPAC は外部で調べてきた資料の所蔵を確認する際に利用されており、単独では文献検索に利用しづらいといったといった課題があると考えられる。

1. 2 研究目的

本研究の目的は OPAC の文献検索結果のランキング結果を改善することにより、利用者がニーズを満たす資料にたどり着きやすくなることである。本研究の目的を達成することにより OPAC による文献検索でニーズを満たしやすくすることで、利用者の文献検索に対するハードルを低くすることを目指す。それにより、OPAC はより学習に貢献しやすくなり、大学図書館における学習・研究の支援に貢献できるようにすることを目指す。学習において必要な情報を見つける際には検索結果から書誌情報を元

に絞込を行う。この時、リランキングを行い、利用者が求める文献に近い情報を検索結果の上位に持ってくる事ができれば、OPAC の検索の利便性が上がり上述の問題点を解決できるのではないかと考えた。OPAC に対して求められていることとしては早く資料を探すようにすることができることである。そこで、より関連性の高い文献を上位に持ってきて探す手間を省くリランキングを行うことで、OPAC での検索がスムーズになり、利用者の利便性が上がり、ひいては大学図書館が求められている学術情報のアクセス環境としての役割に貢献できるのではないかと考えた。

本研究で学習の際に大学図書館の資料を利用する際に OPAC で所蔵情報を確認することから学習の際に確認した情報が利用ログに残っていると考え、学習に関する資料検索の情報として OPAC 利用ログを参照し文献検索を行う手法を提案する。本研究において、検索結果にはリランキングには入力クエリから関連科目を得るシステムと利用ログ中の指定した期間における利用ログの情報を得るシステムを組み合わせる。提案手法では入力したクエリに関連する科目の情報を得ることにより、調べようとしている言葉に関連する開講時期の情報を得る。その上で利用ログから関連科目の開講時期の情報を抽出し、より関連する科目が開講されていた時期の利用をより検索結果の上位になるようにスコアを求める。その上で、元のスコアと組み合わせリランキングを行うことにより、検索者が学ぼうとしていることに関連してよく利用された文献を検索することを試みる。

なお、本研究は筑波大学の卒業研究で行った研究 (22) を拡張して行った研究である。

1. 3 論文の構成

第2章では関連研究を示し、本研究の方向性を示す。第3章では図書館の利用のされ方の分析、第4章では提案手法の説明、第5章で研究の結果を整理し、第6章で考察を行い、第7章でまとめを行った。

2. 関連研究

利用者行動の履歴を用いる観点から本研究に関連する研究としては大きく、図書館分野とそれ以外という軸で分けられ、図書館分野の中も、システムに関する研究と分析に関する研究という軸 (a) と現実での利用を元にした研究とウェブ上の行動を元にした研究と(b)いう2つの軸が存在する。特に2.2.4節で述べるウェブ上での利用を元にしたシステムに関する研究は本研究が属している研究領域である(図1)。

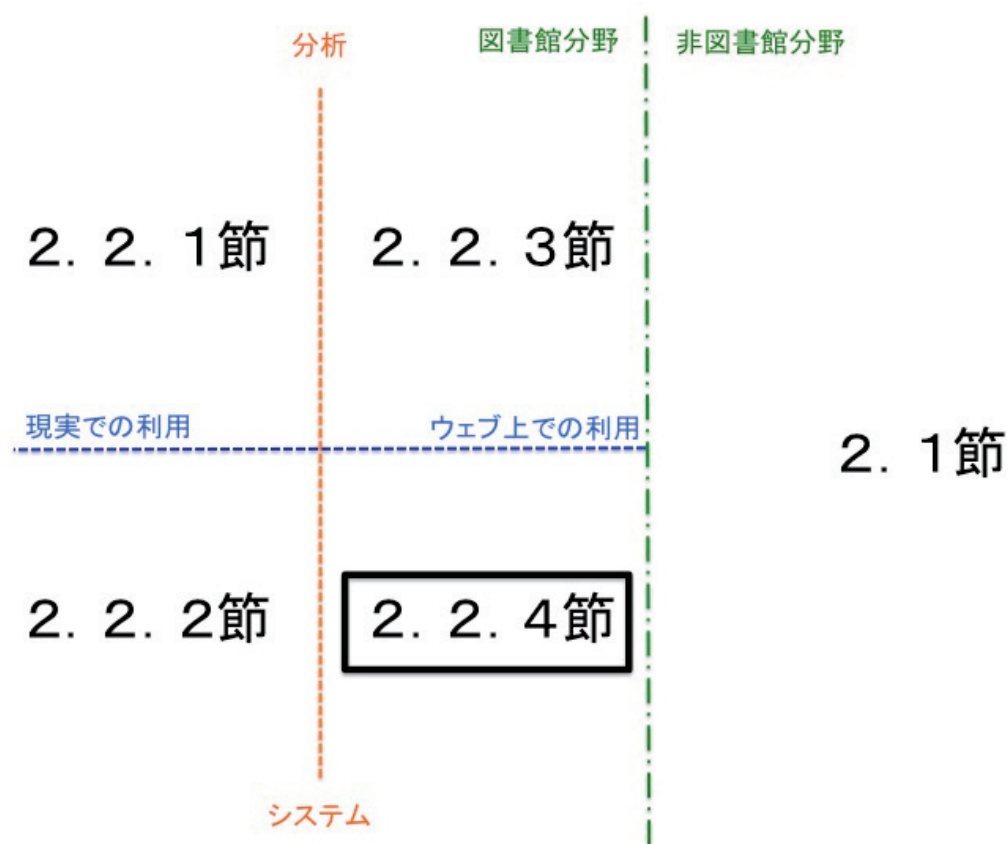


図 1:利用者行動の履歴を用いる関連研究の分類

また、図1で示す領域に加え、利用ログを利用する事に関連して、他の学生の情報を利用した学習に関する研究も関連してくるので2.3節で述べる。

2. 1 図書館分野以外における利用者の情報を利用した関連研究

図書館以外において利用者の情報を用いた研究には利用ログを用いた情報推薦に関する研究が存在している。

映像に関する検索行動の利用には石井ら（2007）の研究（23）が存在する。石井らの研究では映像情報に関して引用しているブログの属性情報による重み付けを行い、視聴者の興味を反映した検索方法の提案を行っている。この研究により、属性情報の量と質による検索結果の比較を行った。

中村ら（2008）の研究（24）ではショッピングサイトにおいて利用ログから利用者の興味を抽出し商品の推薦を行うシステムの研究を行った。山元らの研究では利用者の興味に関する最長共通部分列（LCS）を抽出し学習することにより推薦を行う手法を提案した。すべてのLCSで推薦を行う手法と比較し、利用者の興味に関するLCSによって利用者にあった情報をより早く提供できることがわかった。また、利用ログのLCSを用いた研究には山元ら（2009）によるパターンマイニングを行う研究（25）も存在する。山元らは閲覧行動に関する仮説を立てて推薦情報の検索に用いるパターンマイニングに関して検証を行い、検索対象の特徴を判別することが可能であり、順序情報を用いたほうがより利用者にとって好ましい情報を推薦することができることを示した。

また、石津ら（2010）の研究（26）ではショッピングサイトにおける商品推薦に関して研究が行われた。石津らはショッピングサイトへアクセスする際に取得されるパケットログについて解析し、ユーザの行動を推定・分析し、購入過程を可視化するとともに推薦する支援ツールを提案している。この手法において、パケット中に含まれる情報から購入日時や値段に関しては利用者にあった情報を提供できる一方で、メーカー名や型番に関しては上手く提案できないことを示した。

これらの研究の研究は利用者のニーズに合った情報提供に関する研究である。これらの研究では図書館以外の分野において利用者の情報として利用ログが用いられていることから、図書館における文献検索においても同様に利用ログを用いることで利用者のニーズにあった情報を提供できるのではないかと考えた。

2. 2 図書館における利用者の情報を用いた関連研究

図書館分野において利用者の情報を用いた研究に対して、2つの軸を考える。1つはウェブ上で資料検索を行った際に記録される情報に対する研究と実際の図書館利用に関する研究という軸、もう1つは分析に関する研究であるかシステムに関する研究であるかという軸である。本研究はこの2つの軸のうち、ウェブ上の行動を用いたシステムの研究という位置づけで研究を行う。

2. 2. 1 現実での図書館利用の分析

松井ら（2006）の研究（27）では図書館の貸出データに対し、蔵書回転率と蔵書貸出率を指標として分析が行われた。和洋分類や受け入れ区分、予算単位、主題分野ごとに比較して分析を行い、貸出情報に基づく文献の管理の仕方や配架の仕方・書架の位置などに対する分析を行った。

塩沢ら（2008）の研究（28）では図書館の貸出統計の情報を元に学生がどのように図書館を利用しているか、学生像の分析を行った。図書館への関心のアンケート調査を行い、アンケート結果と貸出冊数の関連性の分析を行った上で、入館者数や年度ごとの貸出冊数の分析を行い、図書館の関心に関する分析を主に行った。

本研究では現実での図書館の利用を OPAC システムの構築という方向でアプローチを行い、現実での図書館を利用しやすくし、図書館利用を促進することを目指す。

2. 2. 2 現実での図書館利用によるシステム

原田（2009）（29）および原田ら（2010）（30）の研究では、貸出履歴を元にした文献推薦手法に関する研究が存在する。原田らの研究では主題や時間などの要素に基づき、貸出履歴の情報の重み付けを変更してフィルタリングを行ったときの検索結果を比較し、最適な重み付けがどのような重み付けであるか、明らかにした。また、辻ら（2012）の研究（31）では原田らの研究の手法と Amazon に対し情報を入力した際に推薦される文献の情報を用いた手法を比較し、図書館において導入すべき図書推薦システムの手法の提案を行った。本研究では図書館内での学習のみで貸し借り

を伴わない文献利用や禁帯出資料といった情報を考慮し、貸出記録に残らない情報を含めて利用者の図書館利用の情報を利用できると考え、OPAC 利用ログを用いた手法を提案する。

2. 2. 3 ウェブ上における行動を利用した分析

野末ら (2004) の研究 (32) は土木学会図書館の書誌データへの利用ログを参照し、利用傾向の分析を行った。利用され方の中でも特に検索に用いるキーワードに着目し、検索キーワードの数の傾向 (熟練度) と 1 回の検索でのセッション件数、検索にかかった時間の傾向 (忍耐度) により分析を行い、熟練度と忍耐度を高いとなしの 2 つに分け、計四分類に対して利用傾向の分析を行った。また、種市ら (2007) の研究 (33) では、名古屋柳城短期大学図書館における情報を分析した。利用月・利用曜日・利用時間・検索キーワード数に対して量的調査法によって分析を行い、大学図書館において利用されやすい時間情報や利用傾向について分析を行った。

Entlich(2010)の研究 (34) では、Cornell 大学図書館における利用傾向の分析を行い、OPCA サイトで利用された言語の傾向や利用された書誌分類の利用について分析し、大学図書館の OPAC サイトの傾向を分析した。Eng ら (2006) の研究 (35) では南洋理工大学図書館の OPAC サイトにおける検索に用いるクエリのパターンの分析が行われた。この分析においてはクエリに用いた要素やクエリの長さ、論理演算子の個数などの分析が行われた。Blecic ら (1998) の研究 (36) では Missouri 大学図書館におけるパラメータを変化させた 2 つのトランザクション処理に対して分析を行った。2 つのトランザクション処理に対して分析を行い、検索パフォーマンスなどを比較し、よりよい処理の考察を行った。

本研究では OPAC ログを分析し、記録されている時間情報やアクセス先といった情報を用いるシステムの構築を目指す。

2. 2. 4 ウェブ上の行動を用いたシステム

Mönnich ら (2008) の研究 (37) では OPAC サイトにおいて同時にウィンドウが開かれた文献の利用ログを用いたシステムの提案が行われた。OPAC サイトを利用した際に記録される利用ログ中において、同時に閲覧していた文献は関連部会として推薦を行うシステムであり、ドイツの OPAC である Biptip (38) の開発に用いられている。一方、Bollen らの研究 (39) では、ネットワーク化された研究者のコミュニティにおけるデータの使用状況を共有し、その情報を元に推薦を行う手法の研究を行った。Bollen らの研究に関連するシステムとしては、ExLibris 社は Bollen らの研究の手法に基づく SFX を用いている学術リコメンデーションサービス bx (40) を提案している。

本研究では利用ログ中に記録されている情報の中でも利用された時間に着目し、アクセスされた時期に開講されている科目に結びつけ、科目に関連する文献であるとし、文献検索に用いる手法を提案する。

2. 3 他の学生の学習経験を反映した学習支援

本研究では利用ログを用いて他者の経験を反映した学習支援を目指している。他の手法でも他者の経験を反映した学習支援が行われている。本節では他者の経験を反映した学習支援の手法に関する研究について紹介する。

半田らの研究 (41) においては、お茶の水女子大学と奈良女子大学において、大学の学習に関して困ったときにどのような相手に相談するかというアンケートを実施した所、それぞれにおいて1番目に高いのは友人、2番目に高いのは先輩となっており、より身近にいる人に相談し、解決を試みることがわかった。また、中林らの調査 (42) においては、愛知みずほ大学における学生の様々な項目に関して相談相手にどのような相手を頼るかというアンケートで調査を行った所、修学・単位の分野において、先輩・チューターをあわせた割合は悩みを抱えている場合は2番目に、悩みを抱えていない場合は1番目に多く存在していた。

このように、他者とりわけ同様の学習を行った経験を持つもののアドバイスは重視されていることがわかる。本研究ではこのような過去の学習者の経験を参考にするアプローチを取る。

3. 関連データ分析

本研究で取り扱う図書館利用ログがどのような情報を含んでいるか、また図書館においてどのような情報が求められているか調査を行うために分析を行った。

3. 1 対象データ

図書館におけるログにどのような情報が含まれるか調査するために、国立教育政策研究所教育図書館（以下教育図書館）の OPAC 利用ログおよび書誌情報、そしてレファレンス協同データベース（43）に記録されたレファレンス事例に対して分析を行った。利用した情報は 2013 年 12 月から 2015 年 2 月にかけての OPAC 利用ログ 10,897,650 件、蔵書情報 462,013 件、レファレンス事例 476 件を対象に分析を行った。

3. 2 分析方法

分析対象である利用ログは図 2 のような構造をしており、1 行につき 1 レコードとして扱う。なお、利用ログのうち、アクセス元の特定を防ぐために一部の情報を X と表記している。

```
165.193.XX.XXX - - [01/Nov/2016:03:39:38 +0900] "GET
/mylimedio/search/book.do?target=local&bibid=XXXXXXXX&lang=jp HTTP/1.1"
200 13014 "-" "XXXX-XXXX/3.1"

165.193.XX.XXX - - [01/Nov/2016:03:39:38 +0900] "GET
/mylimedio/search/book.do?target=local&bibid=XXXXXXXX&lang=jp HTTP/1.1"
200 17251 "-" "XXXX-XXXX/3.1"
```

図 2:利用ログの構造

図2の1項目はどこからアクセスしているかを示す IP アドレス、2項目はいつアクセスしたかを示すタイムスタンプ、3項目はどのような情報に対してアクセスするかを示す要求、4項目はこのアクセスの状態を示すステータスコード、5項目は送受信される情報のバイト数、6項目は直前にアクセスしていた URL、6項目はアクセスに用いていた端末をそれぞれ示す。3章における分析では5項目と6項目が存在しない利用ログを用いる。分析の過程において、利用者が文献情報へアクセスしようとしたログのみを分析対象とするために不要となる情報を削除するために、以下の5点に関するログを削除し分析した。

- ・要求が「GET /」で終了しているログ
- ・サーチエンジンのクローラやロボットからのアクセスログ
- ・png や gif などの画像や css などレイアウト情報を要求し、アクセスする書誌そのものの情報を含んでいないログ
- ・書誌情報以外へのアクセスのログ

要求が「GET /」で終了しているログは情報要求が空であり、どの文献にアクセスしているという情報を含んでいないログであるため、削除した。クローラやロボットからのアクセスは利用者が文献にアクセスした情報ではなく、目的に合っていないため、クローラやロボットの IP アドレスを調査し、IP アドレスを名前引きし、除去対象のドメインと IP アドレスが合致したログを除去した。また、利用者がアクセスした情報でも、メニュー等書誌情報以外へアクセスしたログや表紙やアイコンなどの画像情報、ページレイアウトに関する要求に関しては、直接書誌情報へアクセスしたログではないため、除去した。これらの除去するログの詳細を表1に示す。

表 1:除去するログの条件

フィールド名	条件	理由	
要求	GET /	要求がない	
IP アドレス	stXXXX.nasXXX.gifu.nttpc.ne.jp	検索エンジンのボット	
IP アドレス	googlebot.com		
IP アドレス	b*crawl.yahoo.net		
IP アドレス	msnbot		
IP アドレス	baiduspider*crawl.baidu.jp		
IP アドレス	crawler*ask.com		検索エンジンのボット
IP アドレス	crawl*naver.jp		
IP アドレス	crawl*cuil.com		
IP アドレス	crawl*dotnetdotcom.org		
IP アドレス	spider*yandex.com		
要求	gif、png、jpg、favicon.ico	画像情報	
要求	css、js、jsp、do、mso、xml	ページの構成情報	
要求	shownews.asp	ニュースの表示	
要求	robots.txt	ロボット用ファイル	

除去した結果、利用ログは 63,164 件になり、除去した結果のログをもとに分析を行い、図書館に存在する利用傾向について調査を行った。

分析を行った環境として、OSはWindows7Professional(64bit)、使用言語はPerlとした。

3. 3 分析結果

行った分析すべてを付録Aに示す。はじめに月ごとのアクセス傾向の分析を行った。この分析においては利用ログ中のタイムスタンプの情報に着目して分析を行い、結果を図3に示す。タイムスタンプの書式に合わせてパターンを作成し、月の名前およびその後に出てくる数字の情報を抽出することで年と月の情報を獲得して集計を行うことにより分析を行った。分析の結果、教育図書館において月ごとのアクセスにはばらつきが存在し、2014年の10月にアクセスが最も多いことがわかった。また、同じ1月であっても2014年はアクセスが少ないが2015年はアクセスが多いといった形で、年によって同じ月でもアクセス傾向が異なっていることがわかった。

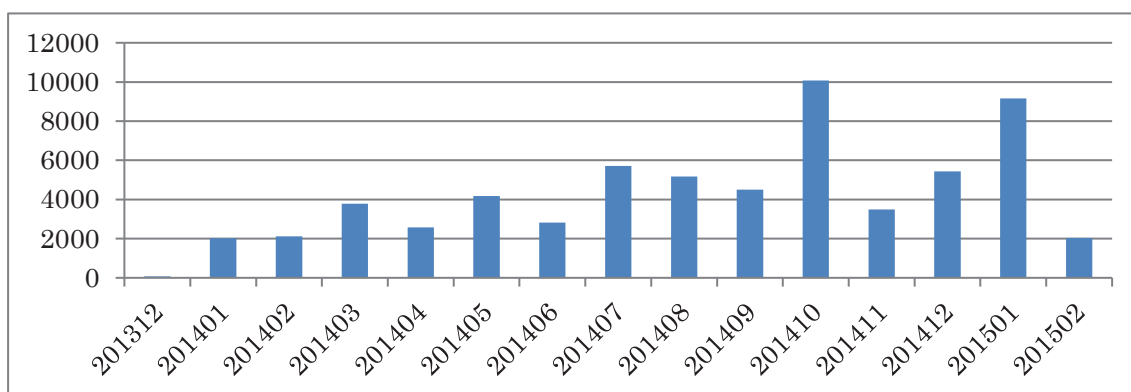


図 3:年間における月ごとのアクセス回数

また、どのような情報に基づき書誌にアクセスしているかについて分析を行った。要求に含まれるURLのパターンを元に分析を行い、書誌情報へのアクセスの際にサイトに対して要求するキーとなったパラメータを分析した。分析の結果は図4のよう

になり、アクセス傾向としては図書の NCID を用いてアクセスを行った「code」によるアクセスと図書館ごとに文献を管理する情報であるローカル ID を用いてアクセスを行った「pkey」を用いたアクセス顔かった。NCID とは NACSIS-CAT における文献の ID のことであり、この分析から教育図書館に関してシステムを構築する際には NCID とローカル ID を用いることとした。詳細なアクセス傾向の分析は付録 A に示す。

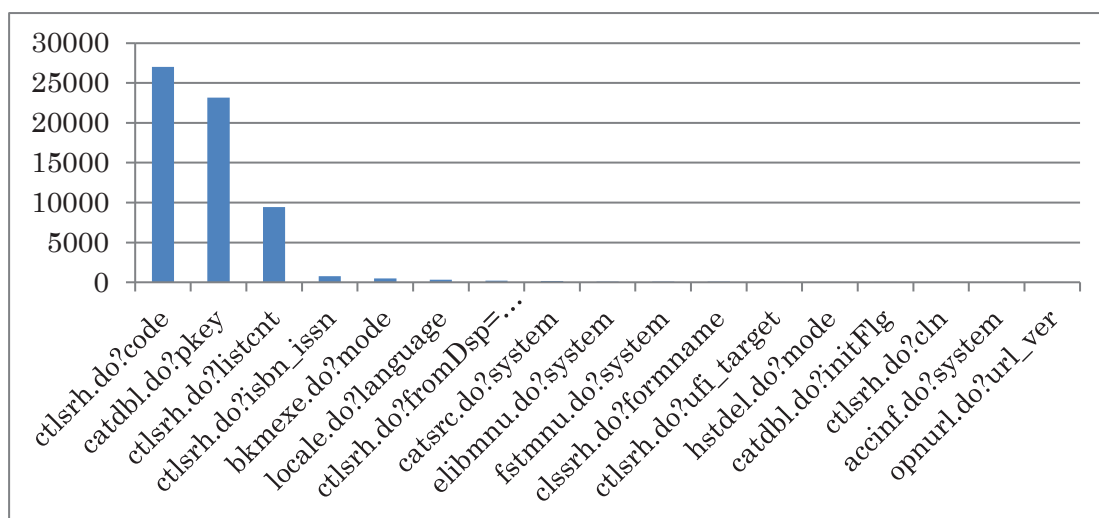


図 4:システムへのアクセスに用いられた情報の分析

続いて、NCID とローカル ID を用いたアクセスのうち、年間でどのような文献へアクセスされているか、また所蔵されている文献がどれほどアクセスされているか上位 2 文献について分析した。利用ログのうち、要求の中でアクセス手法を示すパラメータにおいて、他のアクセス手法のパラメータの詳細の間の値を抽出することでアクセス対象の情報を得ることにより分析を行った。この時、アクセスに最も用いられている「code」パラメータと「pkey」パラメータによってアクセスしているログを対象とした。分析の結果は図 5 のようになり、BB00506833「全国都道府県在京文教担当者連絡協議会編，“文部科学省ひとりあるき”，平成 13 年度版（第 3 4 号），全国都道府県在京文教担当者連絡協議会，2001.」と BB00501993「浅井春夫，松本伊智朗，湯澤直美，“子どもの貧困:子ども時代のしあわせ平等のために”，明

石書店, 2008, p.392.] の2冊は年間で特にアクセスが多いことがわかったが、その反面この2冊以外の文献ではアクセス回数が1/4以下とアクセスが少ない事がわかった。

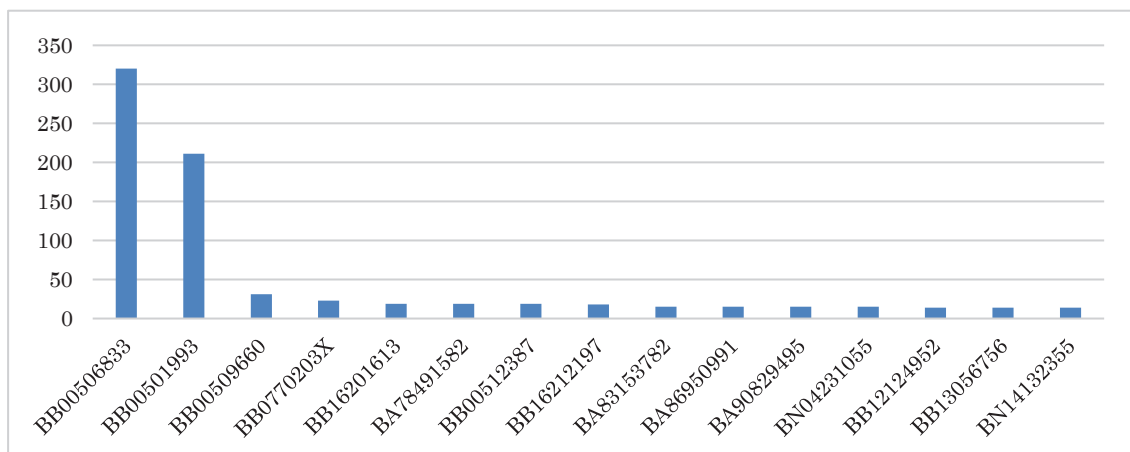


図 5:アクセスされた文献のアクセス回数の分析

そこで、各月で年間と同様に分析を行い、各月での利用傾向を分析した。利用ログのうちタイムスタンプの情報を元に月の情報を得て、各月における上位10件の文献の分析を行った。分析の結果、1月・2月ではBB000621691「Schostak, SohnF., “Maladjusted schooling : deviance, social control, and invididuality”, FalmerPress, 1983, 235p.」、3月では年間でも多くアクセスされていた浅井らの「子どもの貧困」がそれぞれの月で多くアクセスされていることがわかり、月ごとにより需要のある本が異なってくるようになった。これにより、月ごとの利用ログの情報を反映することにより、各月の需要にあった推薦を行うことができるのではないかと考えた。

3. 4 レファレンス分析

図書館はどのような目的で利用されるのかと言った情報を分析する際に図書館に対し寄せられたレファレンスが参照できないかと考え、傾向を分析するために教育図書

館のレファレンス事例の分析を行った。この分析の際にレファレンス協同データベースで付与されている事例ごとのキーワードに基づき、どのようなキーワードに関連した事例が多いのか、多く含まれているキーワードについて分析を行った。分析結果を表2に示す。

表 2:キーワードに含まれていた単語上位10件

含まれる 単語	事例の総数	2回以上レファレンスが行われたキーワード
教科書	155	戦後検定教科書、教科書内容、海外教科書、教科書検定、
		国語教科書掲載作品、高校教科書、教科書、文部省著作教科書、
		外国教科書、教科書採択率、地理教科書、国語教科書、
		教科書掲載作品、音楽教科書、戦前教科書、国定期教科書、
		教科書目録、英語教科書、暫定教科書、国定教科書、
		修身教科書
国	99	国語、国語教科書、国定期教科書、国語科、英国、
		全国学力テスト、国歌斉唱、国定教科書、国語教科書掲載作品、
		国定期、外国教科書
学校	96	尋常学校、小学校、旧制中学校、高等女学校、
		高等小学校、学校制度、学校系統図、高等学校、
		師範学校、中学校
教育	89	教育政策、教育制度史、義務教育、教育史、基本教育法、

		英語教育史、理科教育、音楽教育史、通信教育、戦後教育改革、 教育政策史、戦後教育史、教育制度、中央教育審議会
国語	42	国語、国語教科書、国語科、国語教科書掲載作品、
文部	39	文部省令、文部科学省、文部省通知、文部省
小学	34	尋常小学校、小学校、高等小学校
明治	22	明治期、明治、明治初年、明治初期
高等	21	高等小学校、高等学校、高等女学校
音楽	19	音楽教材、音楽教科書、音楽、音楽教育史

分析の結果、「教科書」を含むキーワードが 155 件と約 1/3 を占めていることが分かった。「教科書」を含むレファレンス事例について詳細な調査を行った所、教育に用いる教材に関する質問や質問者が過去学ぶ際に用いていた教科書に関する質問が多いことからこのキーワードを含むレファレンス事例が多いことがわかった。この分析において、キーワードに含まれている単語でグルーピング下が、グルーピングした単語が関連しないキーワードをまとめてしまうことがあった。今回の分析においては含まれている単語に関して詳細な分析を行い、どのような原因で傾向が存在しているかを分析することを目的としていた為、今回の手法に加え、分析方法として該当単語を含まれている単語の抽出も可能な方法に関しては今後の課題とする。

4. 研究手法

この章では本研究でどの手法を用いるのか、その手法に関して説明を行う。

4. 1 概要

本研究では利用者が過去に学習に利用した情報を反映するための情報源として OPAC 利用ログを用い、OPAC 利用ログを用いた文献検索システムの構築を行う。

本研究では関連科目に関する図書館利用における文献利用が OPAC 利用ログに含まれているとする。科目に関する文献利用はその科目が開講されている時期±1月の範囲で文献が利用され、その期間にその科目に関する文献利用が集中していると考えた。

科目情報に関してシラバスを参照し、入力したクエリがどの科目に関連するかの情報を得る。そのときにクエリに関連する科目が複数ヒットすることを考え、関連科目の開講期間の情報を元に各期間に重み付けを行う。そして、各期間における文献の利用を元に科目に関する利用の情報を抽出し、関連科目が開講されている時期により利用されている文献の情報を元に文献検索結果をリランキングする手法を取る。

OPAC 利用ログを反映する手法を実現するために、入力したクエリに対してシラバスを元に関連科目の情報を得る。その上で関連科目の情報から過去の利用者の利用時期の情報を求めて、過去の開講状況に基づいた重みを求める。求めた重み付けに基づき各月の OPAC 利用ログから抽出した情報に重みを付けてスコアを得て、その情報を元にベースとなる OPAC での検索結果に対してリランキングを行う手法を提案する。

4. 2 リランキングについて

リランキングとは、ある方法で得られたランキング結果を新たな基準でランキングし直すことである。本研究では検索結果に対し用い、順位付けを変更することを指す。

具体的にどのようにリランキングを行うかについては以下の表 3、表 4 を元に説明を行う。表 3 をオリジナルのランキング結果、表 4 をリランキングの結果とする。

表 3 はある基準によってスコア付けされそれに基づきランキングされた結果だとする。この時にすでにある基準によってスコア付けされているが、アクセス回数を ABC の 3 つに分け、A でのアクセス回数が多いものを上位に表示して確認したいとする。この時、A でのアクセス回数が多い順に基づき並び替えを行う処理を行う、処理を行った結果、表 4 のようになる。この時、元々のランキングとは順位が変化しており、目的とする順番に並び替わっている。この並び替えは A のアクセス回数を元にしたリランキングであると言う。

表 3:リランキング前の例

Rank	スコア	対象	アクセス回数 (A、B、C)
1	0.9	A	13、14、10
2	0.7	B	15、15、20
3	0.6	C	14、14、15
4	0.5	D	2、2、6
5	0.3	E	4、3、7

表 4:リランキング結果の例

Rank	スコア	対象	アクセス回数 (A、B、C)
1	0.7	B	<u>15</u> 、15、20
2	0.6	C	<u>14</u> 、14、15
3	0.9	A	<u>13</u> 、14、10

4	0.3	E	4、3、7
5	0.5	D	2、2、6

4. 3 システム全体像

本研究で用いる手法について図 6 にアイデアの概要を示す。

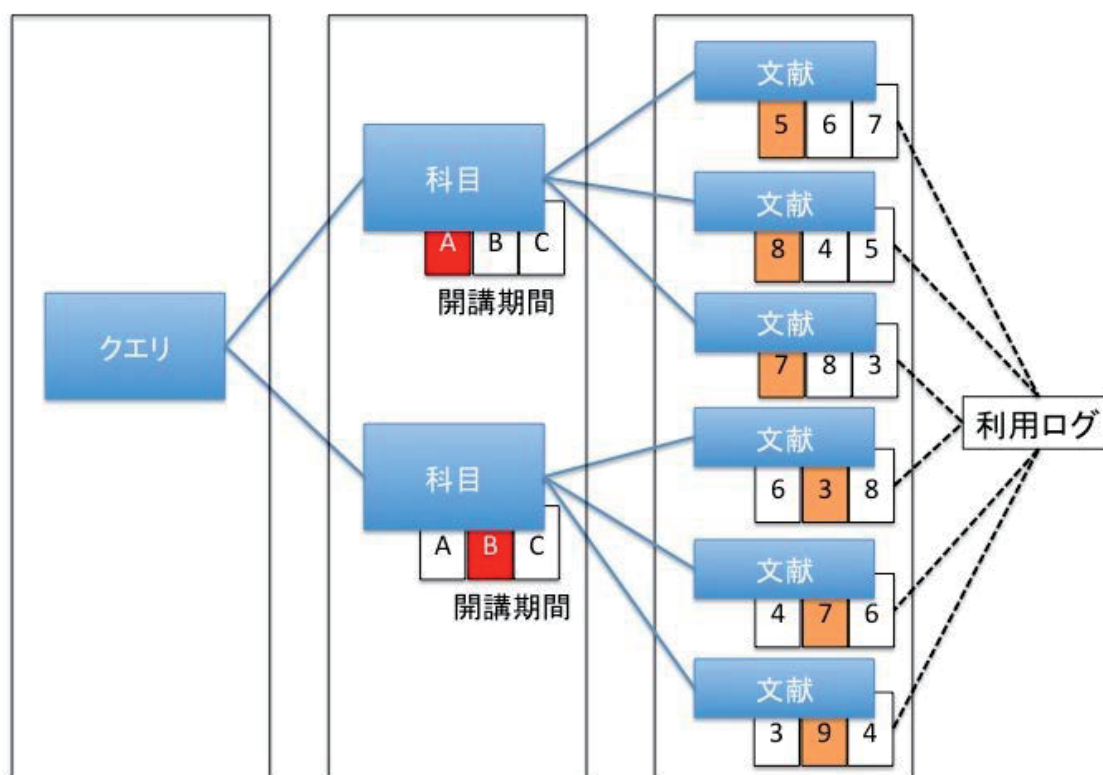


図 6: アイディアの概要図

科目に関連する文献検索を行う際にクエリが入力される場合を想定してシステムの構築を行った。あるクエリによる検索が行われる際にそのクエリに関する科目は複数存在することが考えられ、科目ごとに関連する文献が存在する。科目ごとに関連する文献を発見するための手法として開講時期の情報が参考になると考えられる。開講時期に利用された文献は学習の参考になり、その情報は利用ログに含まれていると考えられる。

本研究で構築するシステムでは関連科目ごとに関連する文献の情報に基づきリランキングを行う手法とした。その上で必要となるシステムとしてベースとなる検索を行う文献検索システムに加え、クエリがどの科目に関連するかというクエリと科目を紐付ける関連科目検索システムと利用ログ中の関連科目が開講された時期に利用された科目に関連する科目の情報を得る利用ログ検索システム、そして関連科目検索システムと利用ログ検索システムを結びつけるために関連する科目の開講期間の情報を得る開講期間重み付けシステムという3つのシステム、そしてリランキングを行うリランキングシステムを構築していく。

リランキングに用いるシステムの概要図を図7に示す。

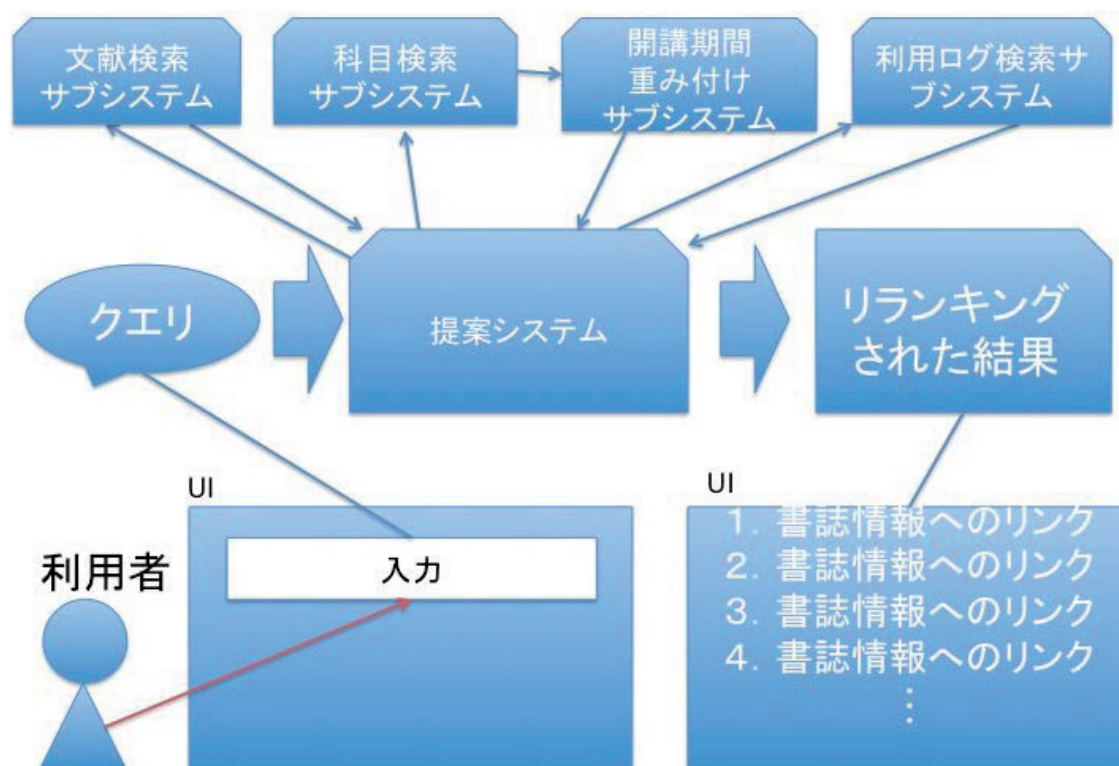


図 7: システム概要図

システム全体の流れとしては、利用者はテキストボックスにクエリを入力して検索を行い、最終的にリランキングされた文献検索結果を受け取る。ユーザにとってはクエリの入力画面と最終的な検索結果の提示のみがユーザインターフェースとして提示される。リランキング処理は提案システムの内部で行われる。

リランキングの処理を行うために、利用者が入力したクエリに基づきシステムは文献検索サブシステムで処理を行い、文献情報を元にしたベースとなるスコアを得る。ベースとなるスコアをリランキングするためのスコアを得るために、クエリを元に関連科目を検索し関連科目の情報を科目検索サブシステムから得られた関連科目に基づいて開講期間による重み付けを行う。関連科目に基づき開講期間重み付けサブシステムにより開講期間による重み付けを行う。そしてベースとなるスコアに載っている文献に対して利用ログ中で関連科目がより存在する時期により利用された文献の情報を利用ログ検索サブシステムから得る。その上で、関連科目がより存在する時期により利用された文献に対する重みが高くなるよう設計されたスコアを算出する。その上で、算出されたスコアを文献検索サブシステムから検索結果の文献群に対するスコアを反映しリランキングシステムでリランキングすることによって、利用者に提供する。

続いてシステムの動作の詳細について説明を行う。システムの構成図を図 8 に示す。

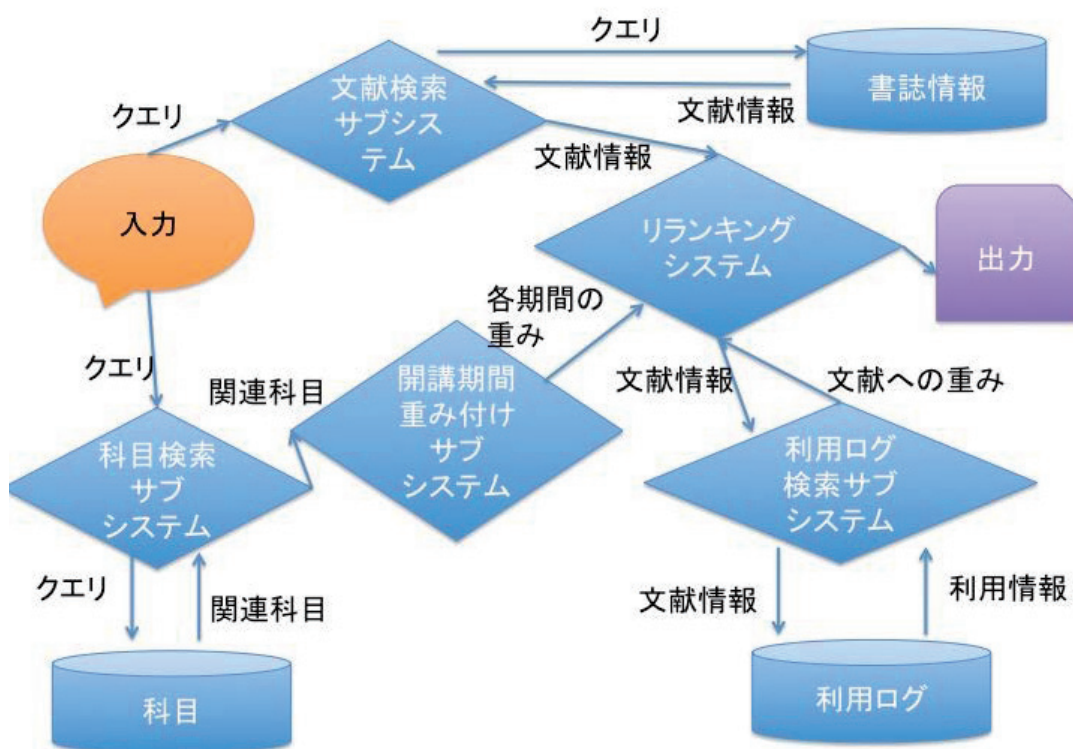


図 8:システム構成図

はじめに入力として利用者がクエリを入力し、その情報を文献検索サブシステムと科目検索サブシステムの2つに渡す。文献検索サブシステムではクエリを入力し書誌情報を元に検索を行い、ベースとなるスコアを出力する。

科目検索サブシステムではクエリ情報と科目情報を入力として渡され、関連科目の情報が出力される。科目検索サブシステムにおいて、シラバスを元にした科目情報を元に検索を行い、入力したクエリに関連する科目の情報を得る。その上で、設定したパラメータに基づきどのような時期の情報を得るか判断し、各月の重み付けの情報を得る。

科目検索サブシステムから得られる関連科目の情報が開講期間重み付けサブシステムに入力として渡る。開講期間重み付けサブシステムにおいて関連科目の情報を入力として渡され、各期間に対する重み付けの情報が出力される。関連科目の開講時期の情報を各月ごとに分け、各月の重み付けとして反映する。

各期間の重みの他に各文献の重みを求めるために利用ログ検索サブシステムで処理を行う。利用ログ検索サブシステムでは利用ログの情報とベースとなるスコアの情報がリランキングシステムを経由して入手した文献の情報を入力として渡され、開講期間ごとの利用に対する重みを反映した各書籍のスコアが出力される。過去の関連科目が開講されている時期に利用されていた文献に対し重み付けを行い、各書籍が関連する科目が開講されている時期に何回アクセスされたかという情報を得る。

科目検索サブシステムに基づく各期間の重み付けの情報と利用ログ検索システムから出力された情報を文献検索サブシステムから得られるベースとなる文献検索の結果に反映し、リランキングを行うことにより最終的なリランキング結果として出力する。

また、本研究ではベースとなるスコアおよびリランキングに用いるスコアを求めるために用いるシステムとしては全文検索システムである Apache Solr (44) を用いる。Apache Solr に対し、クエリを渡すことで書誌情報やシラバスといったインデクスした情報に対し語句検索を行うことによってスコアを得る。

4. 4 関連科目検索システム

本研究で用いるベースとなる元の検索結果に対し、リランキングするために用いるスコアとして、はじめに関連科目に関する検索を行う。

大学において科目に関する情報が記録されている情報源としてシラバスを利用する。シラバスに記録されている情報に対し入力されたクエリが含まれているかどうか語句検索を実行し、入力されたクエリと関連してくる科目がどの科目か、その語句の類似度はどれほどの値かといった情報を得る。この時参照するシラバスはシステムにも用いる利用ログと対応付けて参照する。関連科目の情報を検索する情報として入力されるものとして科目番号、科目名、開講年、開講学期、担当教員、授業概要、学習・教育目標、到達目標、授業計画、履修要件、授業外での学習、教科書・教材、参考書、備考の情報をシラバスから参照する。

この検索において得る情報としてはどの科目がクエリに関連する科目であるかといった情報とそのスコア、そして関連する科目の開講時期の情報を得る。開講時期の情報は筑波大学をもとに考えると、シラバスには春 C 秋 A といった形で学期名が記載されている。システムで用いるときは学期と月を照らし合わせるデータを用意し、シラバスの検索結果を変換することで関連科目が開講されている月の情報を得る。筑波大学における学期と月の対応は表 5 のようにする。また、この時には類似度も同時に参照し、より関連してくる科目が多い月と言った情報を得ることができる。これにより、どの月により多く関連科目が開講されているかという情報を得ることができ、この情報を元に利用ログを参照していく。

表 5:学期と月の対応

学期名	月
春 A	4,5
春 B	5,6
春 C	7,8
秋 A	10,11

秋 B	11,12
秋 C	1,2,3

4. 5 開講期間重み付けシステム

関連科目を検索した結果得られる各月の関連科目の情報を元に利用ログの抽出を行う際に、どのように重み付けを行えば、リランキングの精度が高まるのか調べる必要がある。そこで、本研究ではリランキングの精度を上げることができるか調査するためにどのような範囲で月を扱うか重み付けを複数通り考える。過去の利用者の経験を反映するために利用ログを用いるが、利用ログにおいてどのような期間で参照すれば過去の学習者と対応づくか確かめる必要がある。そのため、本研究ではどのような期間の関連科目を扱うかといった視点でパターンを考え、以下のパターンを提案する。

- ・ 1 年中の関連科目を用いる。
- ・ 実行した月を開講期間に含む科目のみを用いる。
- ・ 実行した月の情報を用いる。
- ・ 実行した月の前後 1 ヶ月を含めた情報を用いる。
- ・ 実行した月の前後 2 ヶ月を含めた情報を用いる。

本研究で提案するパターンは学習に関して調べごとを行う際に予習や復習、課題への取り組みなどを検索行動によって利用する月幅が存在しているため、これらのパターンを切り替えることによってそれらの範囲が広めたり狭めたりすることに対応するシステムを構築し、科目が開講する学期区分を元に分類する。

4. 6 利用ログ検索システム

前節で述べたようにどの月がどれほど関連するかの情報を得た後、関連する月においてどの文献がより利用されるのかという情報により、より関連する期間に利用された文献についてのスコアを得る。

この時、利用ログ中のタイムスタンプに着目し、各文献の各月における利用回数の情報を得る。その上で、各月の重み付けに合わせて各月の利用回数に重み付けしそれを足し合わせることで各文献のスコアを得るとする。この際に利用ログ中に記録される入力されたクエリを参照し、検索に用いているクエリとの類似度を調べ、より検索に用いているクエリに類似している検索において利用された文献の情報の重みをつけるといった手法が考えられるが、同じ科目において調べている事柄に関する検索でも複数のクエリが用いられる可能性があると考え、関連する文献をより広く取り扱いたいと考え、クエリを除いた利用ログの情報によりスコアを付ける。

4. 7 リランキングシステム

4.6 節で得られる情報によって月ごとの関連科目との関連性を反映した各月における OPAC 上で利用された文献のスコアを得ることができる。これによってベースとなる文献検索結果のリランキングを行う。リランキングを行う方法としては、ベースとなる検索結果とリランキングに用いるスコアを正規化してスコア付けし、組み合わせることにより、新たなスコアを得ることで行う。リランキングに用いるスコアは関連科目の開講期間のうち、4.5 節で求める選択したパターンに基づく集計を行う期間内におけるアクセスに基づいたスコアの合計とする。検索結果はスコアが高い順の降順とし、より関連度が高い文献を上位する。スコアの計算は Apache Solr の語句検索のスコアを元にするが、そのスコア計算は書誌情報と検索に用いたクエリの関連度から得られるスコアであり、以下の計算式で導かれる (44) (45)。

$$\text{Score}(q, d) = \text{coord}(q, d) * \text{queryNorm}(q) * \sum_{t \text{ in } q} (\text{tf}(t \text{ in } d) * \text{idf}(t)^2 * \text{norm}(t, d))$$

スコアを構成している関数はドキュメントに関するクエリのスコアを求めている。

関数 $\text{coord}(q,d)$ はドキュメント d においてクエリ中のターム q を変数とし、ドキュメント中におけるタームに合致した数のスコア関数である。

関数 $\text{queryNorm}(q)$ は検索を行う際に様々なクエリでスコアを計算する際にそのスコアが正規化されるために用いられる関数である。

関数 $\text{tf}(t \text{ in } d)$ はドキュメント中にどれほどクエリが出現するのかわを示すスコアである。またそれに関連する関数 $\text{idf}(t)$ はあるドキュメントの内、クエリを含むドキュメントの数がどれほど存在しているかを反映しており、その逆数を取ったスコアである。 $\text{tf}(t \text{ in } d)$ と $\text{idf}(t)$ の2つの関数は文書検索において tf-idf 法という手法で知られているスコアである。

関数 $\text{norm}(t,d)$ はクエリを含むフィールドの長さを示し、クエリを含むフィールドの大きさから、どれほど文書においてクエリが占められている割合を調整するために用いられる。

リランキングの際にはベースとなるスコアに対しどれほどリランキングに用いるスコアを反映するか組み合わせる方法を必要があると考えられる。その為、リランキングにどのように組み合わせる方法をパラメータとし、調整できるようにする。

4. 8 システム全体のスコア計算

利用ログ中にはどのクエリを用い、どの書籍にアクセスされたかという情報が記録される。用いるスコアは表 6 に示す 6 つのスコアとする。

表 6: リランキングに用いるスコア

スコア式	スコアの意味
$S_{base}(d, q)$	ベースとなる文献検索におけるクエリ q に対する文献 d のスコア
$S_{subject}(m, q)$	検索に用いるクエリ q に対する関連科目

	における月 m の比重
$S_{usage}(d, q, m)$	月 m のクエリ q に関連する文献 d のアクセス回数
$Subject_m(q)$	月 m における関連する科目の個数
$Query(q)$	利用ログ中のクエリと入力したクエリ q の類似性
$Usage_r(d)$	利用ログ中の文献 d が利用された回数

$S_{base}(d, q)$ が最終的なスコアのベースとなる文献検索結果のスコアであり、その情報に対し、科目情報と利用ログによってリランキングを行う。この時、 $S_{subject}(m, q)$ と $S_{usage}(d, q, m)$ を掛け合わせることで、各月の関連科目に対する重みを反映したスコアとし、リランキングに用いる。表5および下に示す式においてそれぞれの文字は以下の通りの意味を持っている。

α : 最終的なスコアに含まれるベースのスコアの割合

d : 検索対象となる文献

q : 検索のクエリ

m : 利用ログ中の対象月

r : 重み付けに用いる月の範囲パターン

これらの条件のもと、リランキングの際のスコアの計算式は以下の通りになる。

$$S(d, q) = \alpha \frac{S_{base}(d, q)}{\max S_{base}(q)} + (1 - \alpha) \sum_m^r S_{subject}(m, q) \times S_{usage}(d, q, m)$$

$$S_{subject}(m, q) = \frac{Subject_m(q)}{\sum Subject_{\square}(q)}$$

$$S_{usage}(d, q, m) = \frac{Query(q) \times Usage_m(d)}{\sum \{Query(i) \times Usage_m(i)\}} + \frac{Usage_m(d)}{\sum Usage_m(i)}$$

スコアの計算の際には元となる文献検索のスコアとリランキングを行う為のスコアの2つから構成される。文献検索のスコアはリランキングの元となるスコアとしてリランキングする為の元となる入力したキーワードを元にした書籍情報との類似度検索を行うものとして、元となる情報とする。一方リランキングを行うためのスコアでは月ごとの関連科目に関するスコアと月ごとの文献ごとのアクセス回数のスコアの乗算を用いる。これらのスコアは4.3節で言及したことと関連しており、 r は4.5節で示した月のパターンを意味する。月ごとの関連科目に関するスコアでは月ごとにどれほど関連科目が存在するかを示すスコアとし、文献ごとのアクセス回数はそれを元に利用ログから抽出した値とする。

また、スコアの正規化の方法としては $S_{subject}(m, q)$ と $S_{usage}(d, q, m)$ はそれぞれ全体の利用における割合であるため、 $S_{base}(d, q)$ もスコアの値が最大である結果に対してどれほどの値であるかの割合という形にし、値域を揃えた。

4. 9 実装システム

本研究においてどのようにリランキングすればよりリランキングとして良い結果が出るか考察するためにパラメータを変更し、複数の手法で比較することを考案する。4.6節で示した計算式における α および r をパラメータとし、組み合わせを変更することでパターンを作成し、 α は0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 1.0とする。 α 、 r のパラメータを変更することで適切な値の割合と利用ログを扱う期間について比較を行えるよ

うにする。また、図書館側の OPAC システムにより利用ログに入力クエリが記録されない場合もあるため、本研究では入力クエリを用いない手法を用いる。スコアを計算するために用いる Apache Solr はバージョン 4.10.4 を利用し、システムのプログラミングは Ruby で行った。

開発環境としては二つの環境を用いて実装を行った。データの抽出及びシステムの作成では環境 1 として OS は Windows 7 Professional(64bit)、使用言語としては Ruby を用い、rsolr ライブラリを用い、システム構築を行い、コマンドプロンプトで実行した。システムの作成では、環境 2 として Mac 環境で行った。OS はバージョン 10.10.5、使用言語としては Ruby を用い、rsolr ライブラリを用いてシステム構築を行い、ターミナルで実行した。

5. 結果

5. 1 関連科目検索の検証

筑波大学におけるシラバスを元に関連科目を検索するシステムを構築した。

参照するシラバスとしては 2015 年度の情報学群知識情報・図書館学類の情報を用いた。動作確認の情報源としては 2016 年 7 月 12 日における春日ラーニングコモンズ (46) に寄せられた質問を用いた。寄せられた質問の中から科目の内容に関連しない質問を取り除き、質問件数が特に多かった科目上位 3 件と質問数が少ない順に選ばれていった科目の中からそれぞれ質問を 1 件ずつ取り出し、20 件の科目を選択した。

検索結果を表 7 と表 8 に示す。検索に用いるクエリに一致する科目という一致を条件のもとで実行した所、20 科目中 9 科目でヒットした (表中赤)。その一方で完全一致を条件に検索をした所、20 科目中 7 科目でヒットした。この 2 つの結果を比較した所、一致を条件として検索を行った結果の場合、過剰に関連科目がヒットしていた事例 (表中緑) が存在していた。これは検索クエリとして設定したキーワードを一部含む場合が多く、関連していない科目が多くヒットしてしまうと判断したため、本研究では完全一致を用いることとした。

表 7: 語句解析でクエリを含んでいた際の検索結果

	科目名	クエリ	元とした質問	ヒット数	順位
1	デジタルドキュメント	学術論文情報	2014/5/8	97	4 位
2	ヒューマンインタフェース	processing	2014/6/6	なし	なし
3	マークアップ言語	XSLT	2015/5/12	2	1 位
4	情報サービス構成論	データベース	2014/12/2	17	なし
5	情報行動論	自己組織性	2014/1/28	51	なし

6	知識構造化法	デンドログラム	2012/11/5	なし	なし
7	知識資源組織化論	十進分類法	2014/6/27	42	1位
8	自然言語解析基礎	CaboCha	2013/11/19	なし	なし
9	情報システム概説	2進数	2014/12/15	108	なし
10	テキスト解釈Ⅱ	漢文	2014/4/15	1	1位
11	データベース技術	SQL	2014/10/30	2	2位
12	データベース概説	SQLite3	2016/4/28	95	なし
13	データ構造とアルゴリズム	開発環境	2014/11/20	28	なし
14	基礎数学A	対数関数	2013/8/6	1	1位
15	経営・組織論	表記	2015/6/2	なし	なし
16	質的調査法	5W1H	2012/9/20	122	70位
17	情報検索システム	Python	2015/4/23	なし	なし
18	情報基礎実習	画像	2016/7/8	4	1を1位、2を2位
19	プログラミング演習Ⅰ	文字列	2016/7/8	12	1を5位、2を6位
20	プログラミング演習Ⅱ	条件式	2016/2/4	5	なし

表 8: 語句解析でクエリを完全一致でクエリを含んでいた際の検索結果

	科目名	クエリ	元とした質問	ヒット数	順位
1	デジタルドキュメント	学術論文情報	2014/5/8	なし	なし
2	ヒューマンインタフェース	processing	2014/6/6	なし	なし
3	マークアップ言語	XSLT	2015/5/12	2	1位
4	情報サービス構成論	データベース	2014/12/2	17	なし
5	情報行動論	自己組織性	2014/1/28	なし	なし
6	知識構造化法	デンドログラム	2012/11/5	なし	なし

7	知識資源組織化論	十進分類法	2014/6/27	1	1位
8	自然言語解析基礎	CaboCha	2013/11/19	なし	なし
9	情報システム概説	2進数	2014/12/15	なし	なし
10	テキスト解釈Ⅱ	漢文	2014/4/15	1	1位
11	データベース技術	SQL	2014/10/30	2	2位
12	データベース概説	SQLite3	2016/4/28	なし	なし
13	データ構造とアルゴリズム	開発環境	2014/11/20	2	なし
14	基礎数学 A	対数関数	2013/8/6	1	1位
15	経営・組織論	表記	2015/6/2	なし	なし
16	質的調査法	5W1H	2012/9/20	なし	なし
17	情報検索システム	Python	2015/4/23	なし	なし
18	情報基礎実習	画像	2016/7/8	4	1を1位、2を2位
19	プログラミング演習Ⅰ	文字列	2016/7/8	6	1を5位、2を6位
20	プログラミング演習Ⅱ	条件式	2016/2/4	なし	なし

5. 2 月ごとの重み付けの検証

5.1 節で用いたシステムから得られる情報を元に各月の重み付けを変更するシステムを構築した。5.1 節でヒットした科目を確認することのできた9科目において月ごとの重み付けの変化を確かめるために学期ごとのヒットした科目の開講数に関して質問を行った月で検索を行ったと想定した結果、表9から表13のようになった。オレンジ色に着色している範囲は質問を行った科目の開講時期とする。また、KLCに記録されている質問を行った日を検索を行う日とする。

表 9:年間を通じた 5.1 節でヒットした科目による重み付け

科目名	入力クエリ	ヒットした科目の開講時期					
		春 A	春 B	春 C	秋 A	秋 B	秋 C

マークアップ言語	XSLT	0.250	0.250	0.000	0.250	0.250	0.000
情報サービス構成論	データベース	0.316	0.316	0.105	0.132	0.105	0.026
知識資源組織化論	十進分類法	0.500	0.500	0.000	0.000	0.000	0.000
テキスト解釈Ⅱ	漢文	0.500	0.500	0.000	0.000	0.000	0.000
データベース技術	SQL	0.250	0.250	0.000	0.250	0.250	0.000
データ構造とアルゴリズム	開発環境	0.500	0.500	0.000	0.000	0.000	0.000
基礎数学 A	対数関数	0.000	0.000	0.500	0.500	0.000	0.000
情報基礎実習	画像	0.375	0.375	0.000	0.125	0.125	0.000
プログラミング演習Ⅰ	文字列	0.000	0.000	0.273	0.273	0.273	0.182

表 10:質問を行った月を開講期間に含む重み付け

科目名	入カクエリ	ヒットした科目の開講時期					
		春 A	春 B	春 C	秋 A	秋 B	秋 C
マークアップ言語	XSLT	0.500	0.500	0.000	0.000	0.000	0.000
情報サービス構成論	データベース	0.083	0.083	0.083	0.333	0.333	0.083
知識資源組織化論	十進分類法	0.500	0.500	0.000	0.000	0.000	0.000
テキスト解釈Ⅱ	漢文	0.500	0.500	0.000	0.000	0.000	0.000
データベース技術	SQL	0.000	0.000	0.000	0.500	0.500	0.000
データ構造とアルゴリズム	開発環境	0.000	0.000	0.000	0.000	0.000	0.000
基礎数学 A	対数関数	0.000	0.000	0.500	0.500	0.000	0.000
情報基礎実習	画像	0.500	0.500	0.000	0.000	0.000	0.000
プログラミング演習Ⅰ	文字列	0.000	0.000	0.600	0.400	0.000	0.000

表 11:質問を行った月のみによる重み付け

科目名	入カク エリ	ヒットした科目の開講時期					
		春 A	春 B	春 C	秋 A	秋 B	秋 C
マークアップ言語	XSLT	0.500	0.500	0.000	0.000	0.000	0.000
情報サービス構成論	データ ベース	0.000	0.000	0.000	0.000	1.000	0.000
知識資源組織化論	十進分 類法	0.000	1.000	0.000	0.000	0.000	0.000
テキスト解釈Ⅱ	漢文	1.000	0.000	0.000	0.000	0.000	0.000
データベース技術	SQL	0.000	0.000	0.000	1.000	0.000	0.000
データ構造とアルゴリズム	開発 環境	0.000	0.000	0.000	0.000	1.000	0.000
基礎数学 A	対数 関数	0.000	0.000	1.000	0.000	0.000	0.000
情報基礎実習	画像	0.000	0.000	1.000	0.000	0.000	0.000
プログラミング演習Ⅰ	文字列	0.000	0.000	1.000	0.000	0.000	0.000

表 12:質問を行った月±1月の期間による重み付け

科目名	入カク エリ	ヒットした科目の開講時期					
		春 A	春 B	春 C	秋 A	秋 B	秋 C
マークアップ言語	XSLT	0.500	0.500	0.000	0.000	0.000	0.000
情報サービス構成論	データ ベース	0.000	0.000	0.000	0.333	0.333	0.333
知識資源組織化論	十進分 類法	0.333	0.333	0.333	0.000	0.000	0.000
テキスト解釈Ⅱ	漢文	0.500	0.500	0.000	0.000	0.000	0.000
データベース技術	SQL	0.000	0.000	0.000	0.500	0.500	0.000
データ構造とアルゴリズム	開発 環境	0.000	0.000	0.000	0.500	0.500	0.000
基礎数学 A	対数	0.000	0.000	1.000	0.000	0.000	0.000

	関数						
情報基礎実習	画像	0.000	0.500	0.500	0.000	0.000	0.000
プログラミング演習 I	文字列	0.000	0.500	0.500	0.000	0.000	0.000

表 13:質問を行った月±2月の期間による重み付け

科目名	入力クエリ	ヒットした科目の開講時期					
		春 A	春 B	春 C	秋 A	秋 B	秋 C
マークアップ言語	XSLT	0.333	0.333	0.333	1.000	1.000	0.000
情報サービス構成論	データベース	0.000	0.000	0.000	0.333	0.333	0.333
知識資源組織化論	十進分類法	0.333	0.333	0.333	0.000	0.000	0.000
テキスト解釈 II	漢文	0.500	0.500	0.000	0.000	0.000	0.000
データベース技術	SQL	0.000	0.000	0.333	0.333	0.333	0.000
データ構造とアルゴリズム	開発環境	0.000	0.000	0.000	0.333	0.333	0.333
基礎数学 A	対数関数	0.000	0.333	0.333	0.333	0.000	0.000
情報基礎実習	画像	0.333	0.333	0.333	0.000	0.000	0.000
プログラミング演習 I	文字列	0.333	0.333	0.333	0.000	0.000	0.000

表9から表13の結果から、着色した部分が検索対象の開講月であるが、広い範囲を扱うほど様々な科目を反映できる反面、関連科目が該当月以外に偏っている場合、目的とする科目の情報の重みが軽くなってしまうことが考えられる。

表9を例に考えると、クエリを入力して関連科目がヒットしたのもでも目的とする期間以外にも検索を行った月に開講している科目と同じ重みがついているマークアップ言語やデータベース技術、プログラミング演習 I といった科目が存在している。また、関連科目がヒットしなかった情報サービス構成論に関する検索とデータ構造とアルゴリズムに関する検索において見つかっていなくても該当期間を検索できるか確認した所、目的とする期間以外である春 A および

春 B に対する重みが大きくなっている結果となった。そのため、広く月の幅をとる場合は検索を行っている月の重みをより重くして反映しやすくすべきか検討する必要があると考えられる。また、範囲を限定し、検索を行った月に関連して範囲を変更した場合の変化を確認した所、検索を行った月の±1月、±2月が関連科目の範囲と質問を行った月の期間として両方の情報を含むことができていた。

5. 3 利用ログ抽出の検証

利用ログに関して月で情報を抽出できるようプログラミングを行った。この際に利用ログとして国立教育政策研究所教育図書館の利用ログを用いて動作の検証を行った。

利用ログの抽出により、各月の利用傾向の分析を行った際の情報の基づき各月の利用ログに月ごとの利用傾向があるか確認したところ、月ごとに利用されやすい本の傾向が存在することは関連データの分析でも明らかであったが、1月・2月においては「Maladjusted schooling : deviance, social control, and individuality in secondary schooling」、4月・5月・6月においては「Imprinting : early experience and the developmental psychobiology of attachment」と言った形で上位10件に続けて入っている文献が存在し、まとまった月単位でアクセスされやすい文献が存在していることがわかる(詳細は付録Bを参照)。

このことから、各月には利用傾向が存在し、その傾向をログ中から抽出することでその月の利用にあった文献を検索することができると考えられる。

5. 4 リランキングの検証

筑波大学附属図書館から提供された2016年の11/1から12/1にかけての利用ログのデータに基づきリランキングを行った。

クエリとしては 5.1 節における科目に関する質問の内、利用ログを入手した期間に近い時期に開講されていた 4 つの科目（情報サービス構成論、データベース技術、基礎数学 A、プログラミング演習 I）を対象とした。

その後、Tulips Search（4 7）によりクエリに対する関連度の高い図書上位 20 件を選択し、順位に応じたベースとなるスコアを付けた。その後、利用ログ中でのアクセス回数を求め、順位に応じたスコアを利用ログ中のアクセス回数によりリランキングした。スコアのパターンにおいてベースとなるスコアとリランキングのためのスコアの割合を $\alpha=1.0, 0.5, 0.4, 0.3, 0.2, 0.1, 0.0$ に変更し最終的なスコアの計算を行い、リランキングを行った（詳細は付録 C 参照）。

評価のため検索結果で得られた全文献に対して高適合、適合、部分適合、不適合の 4 段階で適合判定を行った。適合判定作業は著者一人で行った。関連度では高適合 3、適合 2、部分適合 1、不適合 0 とする。検索結果の適合結果を表 14 に示す。

表 14:検索結果の適合結果

	高適合	適合	部分適合	不適合
情報サービス構成論	0	0	1	19
データベース技術	12	5	2	1
基礎数学 A	6	8	2	4
プログラミング演習 I	3	3	2	12

適合判定結果を元に、順位付けの評価指標である nDCG (Normalized Discounted Cumulative Gain) を求めた (4 8)。nDCG は評価対象とする順位の最大数を k 、関連度を rel とすると次の式で求められる (4 8) (4 9)。

$$DCG_k = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log_2(i + 1)}$$

$$nDCG_k = \frac{DCG_k}{idealDCG_k}$$

検索評価結果を表 15 に示す。表 15 に示すように、クエリ「データベース」ではベースラインスコアをそのまま反映する $\alpha=1$ のときに nDCG は最も高くなり、 $\alpha=0.5$ 以下では nDCG は低下していった。クエリ「SQL」ではリランキングのためのスコアをそのまま反映する $\alpha=0$ で nDCG は最も高くなり、それより α が高くなるほど nDCG は低くなった。クエリ「対数関数」では $\alpha=0.4$ で最も nDCG は高くなり、それ以外の α では $\alpha=0.4$ を頂点とした山なりに減少していった。クエリ「文字列」では $\alpha=1$ のときに nDCG は最も高くなり、 $\alpha=0.5$ 以下では nDCG は低下していった。平均では $\alpha=1$ のときに nDCG は最も高くなり、 $\alpha=0.5$ 以下では nDCG は低下していった。

nDCG を比較したところ、適合文献のヒット件数が少なかったクエリ「データベース」とクエリ「文字列」では α が大きいほど nDCG は高くなった。その一方で、適合文献のヒット件数が多かったクエリ「SQL」とクエリ「対数関数」においては α の値が小さいほど nDCG も高くなった。この 2 つの結果から、適合文献のヒット数により結果にばらつきが存在するため、リランキング結果の評価を行うためには、対象とするクエリを増やして詳しく調査を行う必要があることがわかった。

表 15: nDCG による検索評価結果

科目名	入力クエリ	α						
		1	0.5	0.4	0.3	0.2	0.1	0
情報サービス構成論	データベース	0.3155	0.2627	0.2500	0.2500	0.2500	0.2500	0.2500
データベース技術	SQL	0.8549	0.8592	0.8496	0.8956	0.8956	0.9235	0.9335
基礎数学 A	対数関数	0.7551	0.7772	0.7901	0.7717	0.7750	0.7735	0.6810
プログラミング演習 I	文字列	0.5084	0.5019	0.5030	0.4765	0.4377	0.4287	0.4612
平均		0.6084	0.6002	0.5982	0.5984	0.5896	0.5939	0.5814

6. 考察

6. 1 関連科目の検索

関連科目検索を行った場合、語句検索に基づき関連科目検索を行った所、半数以下の科目しかヒットしなかった。この実験の結果ヒットしなかった科目を確認した所、シラバス上に記載されない利用するシステムに関する話題やシラバス中に詳細な内容を記述していないことが原因であると考えられる。科目に関する調べごとの際にはシラバス中の記載されないクエリが用いられるケースも多いと想定することができ、システムを構築するにあたり、科目に関連する詳細な情報を収集し、管理するデータベースが必要になると考えられ、その情報源のサンプルとして今回動作の検証に利用した春日ラーニングコモンズのウェブページに記載されている情報を用いることやmanaba（50）のような学習管理システムの教材そのものの情報を用いることにより検索に用いるクエリに対応できると考えられる。

また、シラバスにおいては関連科目という項目が存在し、その情報に基づき関連科目の関連科目と言う形で関連科目を検索することが考えられた。今回使用したシラバスにおいて記載されていた例は表 15 に示す通りであり、関連科目に関するシステムを今後構築するにあたり利用できる情報であるが情報量が少ないためどのように利用するか検討する必要があると考え、今後の課題としたい。

表 16:関連科目の項目を記載していた科目

科目名	関連科目				
統計	多変量解析	量的調査法			
知識情報演習 I	プログラミング演習 I	プログラミング演習 I I	知識資源組 織化論	データ ベース 概説	
量的調査法	統計	質的調査法			
コンピュータシステムとネットワーク	プログラミング演習 I	プログラミング演習 I I			

メディア社会学	統計	量的調査法	多変量解析	質的調査法	メディア社会文化論
医療情報論	情報探索論				
データ構造とアルゴリズム	プログラミング演習 I	プログラミング演習 I I	情報数学	基礎数学 A	基礎数学 B
メディア社会文化論	メディア社会学				
メディア社会文化論	メディア社会学				
デジタルドキュメント	コンテンツ流通基盤概論	マークアップ言語	テクニカルコミュニケーション	学術メディア論	デジタルライブラリ
web プログラミング	プログラミング演習 I	プログラミング演習 I I (1)	知識情報演習 I		
クラウドコンピューティング	プログラミング演習 I	コンピュータシステムとネットワーク			
ソフトウェア工学	プログラミング演習 I				
アーカイブズの利用	アーカイブズの構築				
情報法	情報社会と法制度				
知識財産概論 A	情報社会と法制度	知的財産概論			
中国図書館学	知識資源組織化論	日本図書館学			

基礎数学 B	情報数学				
--------	------	--	--	--	--

6. 2 月ごとの重み付け

月ごとの重み付けでどのように行うべきか分析した所、範囲を広く取ってしまうと本来目的とする期間より重みが大きくなってしまったことがあった。また、短期間で分析を行う場合でも、今回は月で開講学期を月で判断していたため、複数の学期にまたがってしまう事があった。±1月、±2月の範囲において関連科目がヒットしやすく、学期範囲をより広く授業開始時から取りたい場合は±2月、狭い範囲で確実に近い科目における利用で絞り込みたいなら±1月の範囲が手法として適切ではないかと考えた。

また、検索を行った±1月の範囲で開講されている開講科目という形で利用ログを利用すれば検索している事に近い利用の情報にヒットする可能性が高いと考え、今後の課題としたい。

6. 3 リランキングの検証

リランキング結果の nDCG を比較したところ、ベースとなるランキングスコアとリランキングのためのスコアにおいて、高適合、適合の文献が多い結果ほど α の値を小さく、高適合、適合の文献が少ないほど α の値を大きくするほど nDCG は大きくなり、文献の適合数と α と nDCG の大小に関係性が有るのではないかとということが分かった。

また、本研究では4科目を対象とする4クエリを対象としてリランキングの評価を行ったが、評価用クエリが大きく不足しており、 α のパターンについて検証が不十分であるので、評価用クエリを増やした上で評価することが必要となる。検索結果において適合文献がヒットするパターンやヒットしない場合で検索結果が変化していたことから、今後の課題として、利用ログ検索において適合する適合文献が多くヒットするようなクエリ集合や逆に適合する適合文献が多くヒットしないようなクエリ集合を

設定して、どのような α がよいか検討するために追加して評価することを今後の課題としたい。

また、検索しようとしていることと全く関係のない文献もクエリの選択によって含まれるが、そのアクセスが多い場合に目的と異なるリランキングが行われたケースが存在していた。例えば、「文字列」で検索を行った場合、プログラミングで用いる文献以外で言語に関する文献でも多くの利用を集める文献がヒットしてしまっており、検索結果の評価が低くなる原因となっていた。そのため、リランキングの際には文献の分野など、検索しようとしているニーズに対応するジャンル選択を行えるようにすれば精度を上げることが出来るのではないかと考える。このような手法についても今後の課題としたい。

6. 4 提案手法の限界

提案手法における考え方において、2つのアイデアのギャップが存在していることが手法の限界として考えられる。

一つはアイデアを手法に変換する際にデータの参照が曖昧であることである。クエリと科目と文献を結びつける際にクエリをもとに関連科目検索、文献情報を参照するために利用ログ検索、そしてこの2つを結びつけるために開講期間重み付けをそれぞれ行った。しかし、開講期間重み付けを介して結びつける際に期間の重み付けという形に変換するという手法において OPAC 利用ログには科目に関する利用が含まれているという仮定と関連科目の検索にはシラバスが利用できるという仮定、科目の開講期間に科目に関する利用が集中するという仮定の3つの仮定を用いたが、仮定の検証が一部に留まっている。そのため、これらの仮定を検証することが今後の課題となる。

もう一つは利用するデータにはノイズが含まれていることである。本研究で用いた情報ではシラバス中の単語において本来関連しない科目においてクエリが含まれていることや利用ログ中には科目に関する利用だけでなく様々な利用の情報が含まれてい

る。こういった手法を考えた際にはノイズとなる関連科目以外に関する利用に対する対処法を考慮した上で、考え方が正しいか検証を行うことを今後の課題とする。

また、本研究では時間的都合によりログを入手して十分に検証することができなかったために、クエリに関しては扱わないこととして、クエリログを使用しない手法を用いた。しかし、クエリログを参照することでクエリに対してどのような文献が関連して利用されたかを知ることができると考えられる。そのため、入力したクエリとログ中のクエリの類似度を求めることで類似しているクエリで過去検索した際に用いられた文献を知ることができ、文献検索に役立てることができると考えられる。クエリログの利用は今後の課題としたい。

また、本研究では関連科目検索にシラバスを参照しているため、シラバスに記載されていないような授業に用いているツールや用語といったものでは検索をうまく行うことができない。これにより開講期間の重み付けが上手くいかなくなりランキングのためのスコアを想定通りに計算できなくなる。これを踏まえ、ランキングのためのスコアの計算のために重要である関連科目検索に関して、関連科目検索の改善が今後の課題となる。改善するための案として新たな情報源として 6.1 節で触れたように、シラバス中で利用する情報を拡張することや、春日ラーニングコモンズや manaba といった新たな情報源の追加が考えられる。

6. 5 評価手法の限界

6.3 節で述べたクエリ集合の選定に関してその問題点としてクエリのサンプル数の不足が考えられる。サンプル数の不足について科目選定、質問選定、クエリ集合の作成が存在している。また、適合判定についてもそれぞれ考察する。

科目選定について科目ごとに利用傾向が存在し、クエリ群が異なってくることが考えられる。1, 2 年次では基本的なこと、3, 4 年次で発展したことと言ったように関連する科目でも学年の違いにより用いるクエリが異なってくることが考えられる。

このため、分野の近い関連してくる科目を複数抽出し段階に合わせクエリを設定し検証を行うなど、学年別を考慮した科目選定を行うことを今後の課題としたい。

質問選定のやり方では科目の絞込を行いそこから各科目 1 件ずつとしていた。質問件数が特に多かった科目 3 科目と質問件数が少なかった科目 17 科目を対象としたが、この基準は質問件数が特に多い科目を取り入れたい一方で質問件数が少ない科目から順に取り入れるという基準であった。質問選定のやり方では集計を行う際に除外した質問を除いた最新の質問を用いた。これらの選定基準に対し、入手した利用ログの期間において行われた質問からランダムに選択していく選定を行っていくことを今後の課題としたい。

クエリの作成の際にはクエリ集合の作成方法の見直しとクエリログを用いた手法が今後の課題となる。クエリ集合の作成では選定した質問に応じて一人の判断でクエリの集合を作成したが、複数人で質問を確認し、各々でこの質問の場合どのようなクエリで検索を行うのか集計し、クエリ集合を作成するほうがより多くの場合を比較し検証を行うことが出来るため、クエリ集合の増加を今後の課題としたい。また、クエリを作成するアプローチとしてクエリログ中からランダムに求める手法も考えられる。この手法では科目と紐付けて科目に関連する利用かは断定できないが、クエリのサンプル数を増やし、元々もアプローチとは異なる視点でクエリと科目、クエリと文献を検証していくことが出来る。そのため、新たなクエリの情報源としてクエリログを用いることを今後の課題としたい。

また、文献の適合判定に関して著者個人の一名で行ったため適合判定が安定性を確認する必要があるため、複数人で行い、比較して違いを考慮に入れる必要があると考えられる。それに加え、適合判定を行った文献数も時間の都合の上位 20 件を対象としたがベースランキングとリランキングされた結果において大きな差は生じなかった、より下位の文献を対象とすることで nDCG は変化することが考えられるため、上位 1,000 件を対象として評価を行うことが考えられ、上位 1,000 件を考慮した上でプーリングして上位 100 件のカットオフを行ってそれに対して適合判定を行い、評価を行うことが望ましいと考え、今後の課題としたい。

7. おわりに

本研究は、OPAC の文献検索結果のランキング結果を改善することにより、利用者が目的とする資料にたどり着きやすくなることを目的として利用ログを用いた文献検索手法に対して検討を行った。利用ログ中に含まれる情報から関連科目が開講されていた時期によく利用されていた文献を検索結果の上位にすることにより学習に関連する文献を検索しやすくすることを目指した。関連する科目の検索の際には大学のシラバスを用いて検索を行い、シラバスに記載されている関連科目の開講時期の情報に基づき OPAC 利用ログを抽出してスコアを付けリランキングを行うシステムの構築を行った。

本研究では国立教育政策研究所教育図書館と筑波大学附属図書館の情報を元に利用ログの分析、利用ログと筑波大学情報学群知識情報・図書館学類のシラバスを用いたシステムの検証を行った。それにより、各月における文献の利用のされやすさを分析できた図書館の利用のされ方の調査や、関連科目の検索、リランキングする際に反映する利用ログの検証を行った。

本研究では大学図書館で用いることを想定しシステムを構築したが、データの入手の範囲が限られていたため、月をまたぐパターンを検証や α のパターンと言ったスコアの計算式の検証が制限された。また、検索結果として利用したのが上位20件の評価を行ったため、検証を行う文献の冊数が限られていたため、月の範囲のとり方といったパラメータの検証を行うことができなかった。このため、より広いデータを取り扱っていくこととシステム内のパラメータの調整など細かい部分の調整が今後の課題であると考えられる。

本研究では春日ラーニングコモンズにおける質問を用いて評価用クエリの設定を行った。このデータを用いたことは、春日ラーニングコモンズと図書館利用を結びつける試みとして有効であったと考える。この試みを更に広げて、関連科目の情報源として用いるといった形や単語と科目の紐付けといった形で活用していくことができるのではないかと考え、これらの検証に関しては今後の課題としたい。

また、システムのパラメータの調整などの課題に関しては図書館から情報を入手する難易度を軽視していたという問題を踏まえ、情報の入手において準備と問い合わせが重要であることがわかった。

謝辞

最後に、卒業研究を通じて指導してくださった高久雅生先生、副指導教員の宇陀則彦先生、貴重なデータを提供していただいた国立教育政策研究所教育図書館の関係者の皆様、筑波大学附属図書館の関係者の皆様には深く感謝の意を表し、謝辞とさせていただきます。

参考文献

(1) 日本図書館情報学会用語辞典編集委員会編, “大学図書館”, “図書館情報学会用語辞典”, 第4版, 丸善出版, 2013, p.141.

(2) 科学技術・学術審議会 学術分科会 研究環境基盤部会 学術情報基盤作業部会, “大学図書館の機能・役割及び戦略的位置づけ”, 大学図書館の設備について(審議のまとめ)-変革する大学にあって求められる大学図書館像-, 文部科学省, 2010年12月1日,
http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/toushin/attach/1301607.htm, (閲覧日:2017年12月4日) .

(3) “筑波大学附属図書館とは 附属図書館の使命と目標”, 筑波大学附属図書館,
<https://www.tulips.tsukuba.ac.jp/lib/ja/about/mission>, (2017年12月4日) .

(4) “これからの企業・社会が求める人材像と大学への期待”, 公共社会法人 経済同好会,
https://www.doyukai.or.jp/policyproposals/articles/2015/pdf/150402a_02.pdf,
(閲覧日:2017年12月18日) .

(5) 教育再生実行会議, “これからの時代に求められる資質・能力と、それを培う教育、教師の在り方について(第7次提言)”, 首相官邸,
https://www.kantei.go.jp/jp/singi/kyouikusaisei/pdf/dai7_1.pdf, (閲覧日:2017年12月13日) .

(6) 科学技術・学術審議会 学術分科会 学術情報委員会, “学習環境充実のための学術情報基盤の整備について(審議まとめ)”, 文部科学省,
http://www.mext.go.jp/component/b_menu/shingi/toushin/__icsFiles/afieldfile/2013/08/21/1338889_1.pdf, (閲覧日:2017年12月20日) .

(7) “筑波大学附属図書館 マスサポートプログラム一覧2016”, 筑波大学附属図書館,

<https://www.tulips.tsukuba.ac.jp/lib/sites/default/files/attach/masssupportprogram2016.pdf>, (閲覧日:2017年12月4日)。

(8) “筑波大学附属図書館ライティングセミナー”, 筑波大学附属図書館,
http://www.tulips.tsukuba.ac.jp/writing_seminar/chishikitokotoba.html, (閲覧日:2017年12月8日)。

(9) 谷奈穂, 嶋田晋, “北米の大学図書館におけるニーズ調査とシーズ提供の方策の調査”, 大学図書館研究, 2015, vol.103, p.12-15.

(10) 立石亜紀子, “大学図書館における「場所として図書館」の利用実態”, Library and information science, 2012, vol.67, p.39-61.

(11) 古橋英枝, “大学生の学習実態に基づく大学図書館の役割”, Library and information science, 2014, vol.72, p.95-121.

(12) 木本幸子, “図書館で使える情報源と情報サービス”, 日外アソシエーツ株式会社, 2010, 197p.

(13) 日本図書館情報学会用語辞典編集委員会編, “OPAC”, “図書館情報学用語辞典”, 第4版, 丸善出版, 2013, p.23.

(14) “筑波大学附属図書館 所蔵資料を探す”, 筑波大学附属図書館,
<https://www.tulips.tsukuba.ac.jp/lib/ja/support/ref-tsukubaholdings>, (閲覧日:2017年12月11日)。

(15) “資料探索支援サイト”, 姫路獨協大学,
<http://www.himeji-du.ac.jp/library/tansaku.html>, (閲覧日:2017年12月11日)。

(16) 市古みどり編著, 上岡真紀子, 保坂陸著, “資料検索入門:レポート・論文を書くために”, 慶應義塾大学出版会, 2014, 151p.

- (17) “平成29年度 全学共通科目「大学図書館の活用と情報探索」”，京都大学附属図書館機構，<http://www.kulib.kyoto-u.ac.jp/support/12303>，（閲覧日：2017/12/15）。
- (18) 長尾真監修，川崎良考編集，“情報探索入門 増補版”，2001，197p.
- (19) 安蒜孝政，逸村裕，“図書館における大学生の情報探索行動”，中部図書館情報学会誌，2013，vol. 53，p. 17-34.
- (20) 工藤絵理子，片岡真，“次世代 OPAC の可能性-その特徴と導入への課題-”，情報管理，2008，vol. 51，no. 7，p. 480-498.
- (21) 種市淳子，逸村裕，“エンドユーザーの web 探索行動：短期大学生の実験調査に基づく情報評価モデルの構築，Library and information science，2006，vol. 55，p. 1-23.
- (22) 小幡将司，“OPAC 利用ログを用いた文献検索システムの構築”，2015 年度卒業論文，筑波大学知識情報・図書館学類，2016，35p.
- (23) 石井久治，市川祐介，佐藤宏之，小林透．“映像を話題としたコミュニティ活動支援に基づくアノテーションシステム”，情報処理学会論文誌，2007，vol. 48，no. 12，p. 3624-3636.
- (24) 中村健二，田中成典，吉村智史，細島啓史，北野光一，田中裕一．“ユーザ特性を考慮したアクセスログに基づく情報推薦に関する研究”，情報処理学会全国大会講演論文集第 70 回平成 20 年度，2008，Vol. 1，p. 649-650.
- (25) 山元理絵，小林大，小林隆志，横田治夫，“Web アクセスログからのパターンマイニングによる購買行動の推定”，電子情報通信学会技術研究報告 ライフインテリジェンスとオフィス情報システム，2009，vol. 109，No. 272，p. 89-94.

(26) 石津丈也, 若原俊彦, “パケット解析を用いたショッピング支援手法の検討”, 電子情報通信学会技術研究報告 ライフインテリジェンスとオフィス情報システム, 2010, vol. 109, No. 379, p. 7-12.

(27) 松井朗, 磯野肇, “「蔵書回転率」と「蔵書貸出率」を指標とする貸出データの分析調査”, 奈良大学紀要, 2006, vol. 34, p. 177-190.

(28) 塩沢千文, 玉置すみ子, 翠川舞, 永井貴子, 田中仁, “図書館の貸出統計から見る学生像”, 飯田女子短期大学紀要, 2008, vol. 25, p. 191-200.

(29) 原田 隆史, “図書館の貸出履歴を用いた図書の推薦システム”, デジタル図書館, 2009, vol. 36, p. 22-31.

(30) 原田 隆史, 増田 浩佑, “貸出記録を用いた図書推薦システムにおける重みづけの変更”, デジタル図書館, 2010, vol. 38, p. 54-66.

(31) 辻慶太, 黒尾恵梨香, 佐藤翔, 池内有為, 池内淳, 芳鐘冬樹, 逸村裕, “図書館の貸出履歴を用いた図書推薦システムの有用性検証”, 図書館界, 2012, vol. 63, No. 3, p. 176-189.

(32) 野末道子, 嶋田真智恵, 寺尾洋子, 上田修一, “OPAC ログ分析による検索過程の類型化”, 2004 年度三田図書館情報学会研究大会発表論文集, 2004, p. 41-44.

(33) 種市淳子, 逸村裕, “短期大学図書館における情報探索行動: 文字を付与した OPAC のログ分析と検索実験を基にして”, 名古屋大学附属図書館研究年報, 2007, vol. 5, p. 57-68.

(34) Rich Entlich. “Report of the Collection Development Executive Committee Task Force on Print Collection Usage Cornell University Library”. Submitted to the Collection Development Executive Committee, 2010, p. 1-36,

(http://staffweb.library.cornell.edu/system/files/CollectionUsageTF_ReportFinal11-22-10.pdf).

(35) Eng Pwey Lan, Dion Hoe-Lian Goh, “In search of query patterns: A case study of a university OPAC” , Information Processing and Management, 2006, vol. 42, p. 1316-1329.

(36) Deborah D. Blecic, Nirmala S. Bangalore, Josephine L. Dorsh, Synthia L. Henderson, Melissa H. Koenig, and Ann C. Weller, “Using Transaction Log Analysis to Improve OPAC Retrieval Results” , College and Research Libraries, 1998, vol. 72, no. 2, p. 39-50.

(37) Michael Mönnich, Marcus Spiering, “Adding Value to the Library Catalog by Implementing a Recommendation System” , D-Lib Magazine, 2008, vol. 14, No. 5/6, (<http://www.dlib.org/dlib/may08/monnich/05monnich.html>) .

(38) “BibTip” , Starseite, <http://www.bibtip.com/de>, (閲覧日:2015/12/18).

(39) Johan Bollen, Herbert Van de Sompel, “An architecture for the aggregation and analysis of scholarly usage data.” , JDL’ 06 Proceedings of the 6th ACM/IEEE-CS joint conference on Digital libraries, 2006, p. 298-307.

(40) “ExLibris bx” , ExLibris,
<http://www.exlibrisgroup.com/category/bXUsageBasedServices>,

(閲覧日 : 2015/12/18) .

(41) 半田 智久, “学修支援の情報やシステムに関する大学生のニーズ調査報告 高等教育と学生支援” , お茶の水女子大学教育機構紀要, Vol. 3, p. 66 -83.

(42) 中林恭子, 後藤和史, “本学の学生の援助要請行動について” , 瀬木学園紀要, 2015, vol. 9, p. 82-87.

(43) “レファレンス協同データベース”，国立国会図書館，
<http://crd.ndl.go.jp/reference/>，（閲覧日：2018/01/09）。

(44) 大谷純，阿部慎一郎，大須賀稔，北野太郎，鈴木教嗣，平賀一昭．“改訂新版 Apache Solr 入門 オープンソース検索エンジン”，技術評論社，2014，336p.

(45) ” Class Similarity”，
https://lucene.apache.org/core/2_9_4/api/all/org/apache/lucene/search/Similarity.html.

Apache Software Foundation，（閲覧日：2015/12/19）。

(46) “KLC-春日ラーニングコモンズ”，春日ラーニングコモンズ，
<http://klis.tsukuba.ac.jp/lc/>，（閲覧日：2018/01/09）。

(47) “筑波大学附属図書館 Tulips”，筑波大学附属図書館，
<http://www.tulips.tsukuba.ac.jp/lib/>，（閲覧日：2018/01/19）。

(48) 酒井哲也．“情報アクセス評価方法論”，コロナ社，2015，297p.

(49) Cris Buges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, Gref Hullender, “Learning to Rank using Gradient Descent”, ICML ’05 Proceeding of the 22nd International conference on Machine learning, 2005, p. 89-96.

(50) “学習管理システム”，筑波大学，<https://www.ecloud.tsukuba.ac.jp/manaba>，
（閲覧日：2018/01/22）。

付録

付録A : OPAC 利用ログの分析

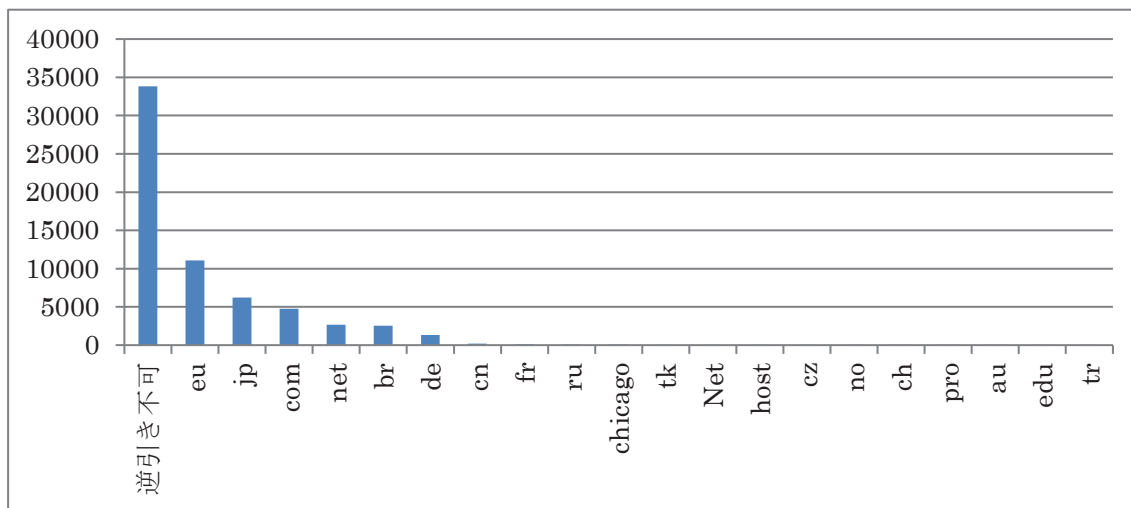


図 A-1 : ドメインごとのアクセス回数

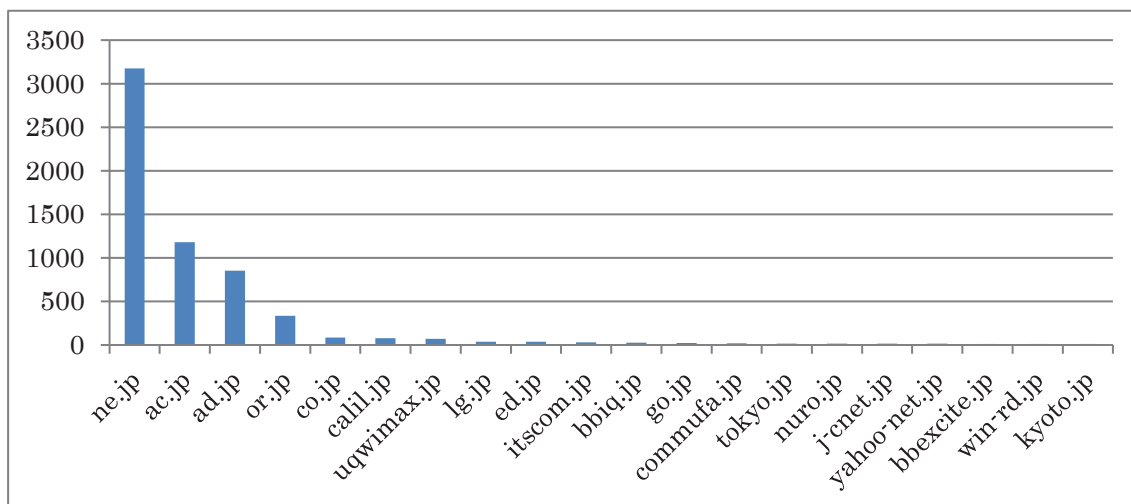
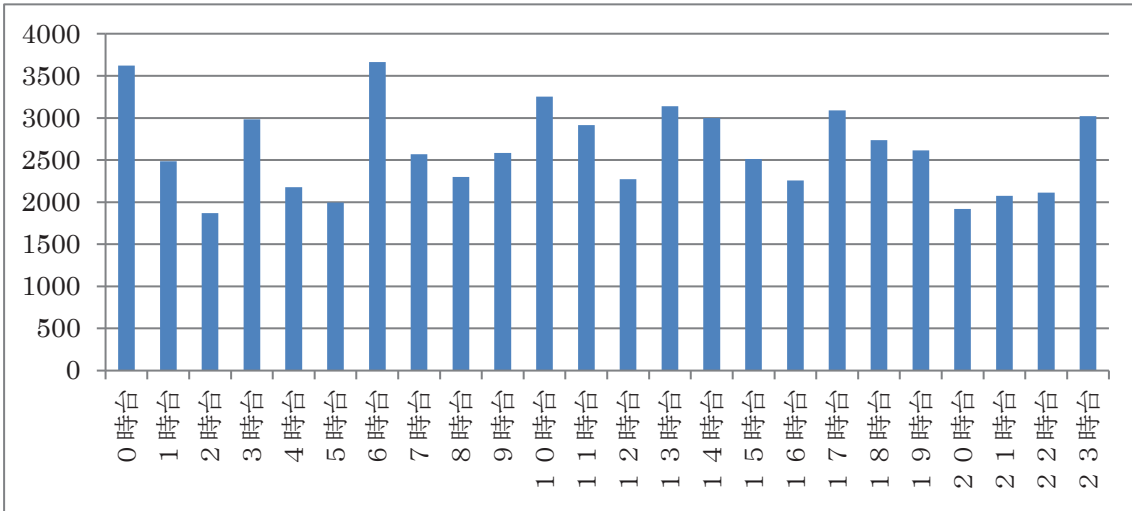
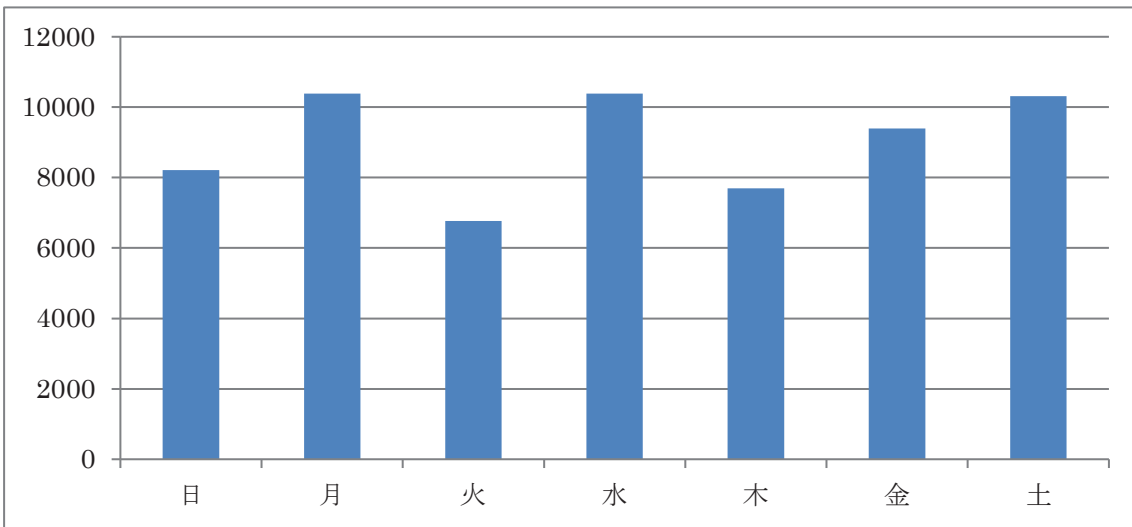


図 A-2 : 国内のドメイン「jp」の内でのアクセス回数



図A-3：時間帯ごとのアクセス回数



図A-4：曜日ごとのアクセス回数

付録B：月ごとのアクセス回数

表 B-1：1月のアクセス回数の多い書籍上位10件

文献 ID	All	Jan	書名
BB00062691	156	86	Maladjusted schooling : deviance, social control, and individuality in secondary schooling
BB00016873	25	13	台湾教育における六士先生の功績
BB00506833	320	9	文部科学省ひとりあるき
BA83161350	35	9	The U.S.S.R. after 50 years : promise and reality
BB00035051	18	9	発展的学習の指導の手引き
AN00361866	29	8	学校運営
BB00054713	10	8	How to improve your memory
BB00087163	44	7	内容分析 (バーナード)
BB00057820	34	7	The U.S.S.R. after 50 years : promise and reality
AN1035926X	20	7	年会論文集

表 B-2：2月のアクセス回数が多い書籍上位10件

文献 ID	All	Feb	書名
BB00062691	156	53	Maladjusted schooling : deviance, social control, and individuality in secondary schooling
BB00036465	21	17	新エゴイズムの若者たち：自己決定主義という価値観

BB00057820	34	11	The U.S.S.R. after 50 years : promise and reality
BB00016873	25	11	台湾教育における六士先生の功績
BB00020214	14	10	教育委員会による地方教育行政
BA72088430	35	8	民国时期社会调查丛编
BB00071382	13	8	中華人民共和国・香港の投資関連税制便覧
AA10509611	45	7	Bulletin of the National Science Museum. Series C, Geology & paleontology
AA12425818	30	7	くらしき作陽大学作陽音楽短期大学研究紀要
BB00088002	20	7	満洲幣制の現在及將來

表 B-3 : 3月のアクセス回数が多い書籍上位 10 件

文献 ID	All	Mar	書名
BB00501993	211	211	子どもの貧困
BB00040868	63	63	Elternwille und staatliches Bestimmungsrecht bei der Wahl der Schullaufbah
EA16210199	56	56	長野大学紀要
BB00023770	24	23	教科書問題 : 今日の焦点とその歴史
BB00506833	320	13	文部科学省ひとりあるき
BB00509660	31	11	三日市小学校草創期における学校名の変遷
BB00054555	71	10	Imprinting : early experience and the developmental

			psychobiology of attachment
BA00306558	18	8	Computation and cognition : toward a foundation for cognitive science
BN04231055	15	8	産業化の時代
BB00070277	8	8	遺傳の綜合研究 : 遺傳の綜合研究委員會研究業績

表 B-4 : 4 月のアクセス回数が多い書籍上位 10 件

文献 ID	All	Apr	書名
BB00054555	71	17	Imprinting : early experience and the developmental psychobiology of attachment
BB00025756	23	15	Educación peruana : un drama en ocho actos
BB00120898	23	8	Globalization and education
BA35621642	22	7	現代科学技術と地球環境学
BN04735955	20	7	昭和財政史 : 昭和 27-48 年度
BA46254601	19	7	大漢和辞典 (1990.5-2000.4)
BA00306558	18	7	Computation and cognition : toward a foundation for cognitive science
BN00193384	15	7	図書館年鑑
BB00506833	320	6	文部科学省ひとりあるき
AA10509611	45	6	Bulletin of the National Science Museum. Series C,

			Geology & paleontology
--	--	--	------------------------

表 B-5 : 5 月のアクセス回数が多い書籍上位 10 件

文献 ID	All	May	書名
BB00506833	320	163	文部科学省ひとりあるき
BB00054555	71	10	Imprinting : early experience and the developmental psychobiology of attachment
BB00025756	23	8	Educación peruana : un drama en ocho actos
BA48803584	16	8	黄昏の哲学 : 脳死臓器移植・原発・ダイオキシン
BA41545163	24	7	21 世紀の持続的発展に向けた環境メッセージ
BA6150778X	23	7	奈良県の地名
BA86950991	15	7	諸外国の教育動向
BA5209558X	12	7	地球と共生する「環の国」日本を目指して
BN00140923	12	7	醍醐寺文書
BN00843581	11	7	地理学文献目録

表 B-6 : 6 月のアクセス回数が多い書籍上位 10 件

文献 ID	All	Jun	書名
BB00506833	320	85	文部科学省ひとりあるき

BB00116448	30	23	児童生徒の学習漢字と語彙の習得に関する基礎的調査 研究
BB00054555	71	13	Imprinting : early experience and the developmental psychobiology of attachment
BN01104878	34	8	我が国の政府開発援助
BB00037527	9	7	現代の診断
BB00103307	8	6	Diagnosis of our time : wartime essays of a sociologist
EA00019264	7	6	適正学級規模に関する研究—学力と学級規模に関する 実験的研究
BB00075472	6	6	雙軒庵鑒賞.
AA10509611	45	5	Bulletin of the National Science Museum. Series C, Geology & paleontology
BB00087163	44	5	内容分析 (バーナード)

表 B-7 : 7月のアクセス回数が多い書籍上位 10 件

文献 ID	All	Jul	書名
EA16241222	102	40	選択性緘黙の児童への関係性を重視した教育相談の取組
BB00062519	20	17	Mein Kampf
BB00506833	320	13	文部科学省ひとりあるき
BB16201613	19	13	The price of government : getting the results we need in an age of permanent fiscal crisis

BB00029642	70	12	Die Durchsetzung des Hochschulselbstverwaltungsrechts vor dem Bundesverfassungsgericht und den Landesverfassungsgerichten
BB00512387	19	11	「日本語を考える」(5年・説明文).
BB00054555	71	10	Imprinting : early experience and the developmental psychobiology of attachment
BB00030329	10	9	日本大学創立七十年記念論文集
BB00086902	9	9	琉球史概観
BB00025829	13	8	The adjunct faculty handbook

表 B-8 : 8月のアクセス回数が多い書籍上位 10 件

文献 ID	All	Sep	書名
BB00029642	70	31	Die Durchsetzung des Hochschulselbstverwaltungsrechts vor dem Bundesverfassungsgericht und den Landesverfassungsgerichten
EA16241222	102	21	選択性緘黙の児童への関係性を重視した教育相談の取 組
BB00054555	71	8	Imprinting : early experience and the developmental psychobiology of attachment
BB00073258	28	5	反応速度論

BB00012266	6	5	小學改正教授法
BB00062004	20	4	Kulturelle Identität und Universalität : interkulturelles Lernen als Bildungsprinzip
BB00104232	8	4	Nouvelles études de sociologie électorale
BA66667772	7	4	教育における「国家」と「個人」
BB02459342	7	4	貧困のダイナミズム : 日本の税社会保障・雇用政策と 家計行動
BB00049190	128	3	On the scope and nature of university education

表 B-9 : 9月のアクセス回数が多い書籍上位 10 件

文献 ID	All	Sep	書名
BB00049190	128	78	On the scope and nature of university education
BB00029642	70	20	Die Durchsetzung des Hochschulselbstverwaltungsrechts vor dem Bundesverfassungsgericht und den Landesverfassungsgerichten
EA16241222	102	12	選択性緘黙の児童への関係性を重視した教育相談の取 組
BB00073258	28	11	反応速度論
BB13749195	11	9	地名符音表 : 一名小學用地圖附録
BB13795454	13	7	日本地誌畧用小學地圖指南譜

BB16229170	9	7	日本地誌畧用小學地圖指南譜
BB00062004	20	6	Kulturelle Identitt und Universalitt : interkulturelles Lernen als Bildungsprinzip
BB00095163	12	6	Jongens vragen
BB00051635	10	5	A model for the flow of students through the Swiss university system

表 B-10 : 10 月のアクセス回数が多い書籍上位 10 件

文献 ID	All	Oct	書名
BB00049190	128	35	On the scope and nature of university education
EA16241222	102	15	選択性緘黙の児童への関係性を重視した教育相談の取 組
EA10127164	34	14	読書指導における読後集団討議の指導計画作成
BB00107289	13	10	The Mahatma : a Marxist symposium
BB00506833	320	9	文部科学省ひとりあるき
BB00025824	14	8	The future of education : perspectives on national standards in America
BB00107297	13	7	The Encyclopadia Britannica : a dictionary of arts, sciences, literature and general information.
BB00055091	11	7	Why men pray
BB00103327	7	7	Languages in contact : findings and problems

BB00062004	20	6	Kulturelle Identität und Universalität : interkulturelles Lernen als Bildungsprinzip
------------	----	---	---

表 B-11 : 11 月のアクセス回数が多い書籍上位 10 件

文献 ID	All	Nov	書名
EA10127164	34	18	読書指導における読後集団討議の指導計画作成
BB00035963	26	11	『早期教育』に対する保護者の意識調査
AN0036198X	6	6	教育研修所紀要
BB00119672	6	6	Initiativpreis Aus- und Weiterbildung
BB00506833	320	5	文部科学省ひとりあるき
AN00084642	27	5	甲南大学紀要. 文学編
BA43949709	6	5	特別活動編(小学校学習指導要領の展開)
BB00049190	128	4	On the scope and nature of university education
AN00361866	29	4	学校運営
AN1035926X	20	4	年会論文集 / 日本科学教育学会

表 B-12 : 12 月のアクセス回数が多い書籍上位 10 件

文献 ID	All	Dec	書名
BB00062691	156	14	Maladjusted schooling : deviance, social control, and individuality in secondary schooling

BB00087163	44	10	内容分析 (バーナード)
BB00506833	320	7	文部科学省ひとりあるき
EA16241222	102	7	選択性緘黙の児童への関係性を重視した教育相談の取組
AN00057901	24	7	教育心理学年報
BB00026777	6	6	Tiempo educativo mexicano / [Pablo Latapi]
BB00063658	7	5	Methodist adventures in Negro education
BB00087443	5	5	海軍航空戦：支那事變戦記
BB00073258	28	4	反応速度論
BN15787418	16	4	文部時報 第1期1回

付録 C: リランキングしたときの検索結果

表 C-1: 情報サービス構成論を元にした「データベース」検索時の上位 20 件の検索結果のスコア

文献		ベーススコア		リランキングスコア		dに応じた各文献のスコア						
タイトル	適合度	bibid	スコア	閲覧回数	スコア	1.0	0.5000	0.4000	0.3000	0.2000	0.1000	0.0000
データベース(平井利明ほか)	0	1640570	1.0000	1	0.0625	1.0000	0.5313	0.4375	0.3438	0.2500	0.1563	0.0625
データベース(橋本大輔編)	0	1491744	0.5000	1	0.0625	0.5000	0.2813	0.2375	0.1938	0.1500	0.1063	0.0625
データベース(石川博)	0	1328821	0.3333	1	0.0625	0.3333	0.1979	0.1708	0.1438	0.1167	0.0896	0.0625
データベース(園田義久ほか)	0	1328327	0.2500	1	0.0625	0.2500	0.1563	0.1375	0.1188	0.1000	0.0813	0.0625
データベース(Jemes)	0	1106040	0.2000	0	0.0000	0.2000	0.1000	0.0800	0.0600	0.0400	0.0200	0.0000

Martin)												
データベース応用	0	1715720	0.1667	0	0.0000	0.1667	0.0833	0.0667	0.0500	0.0333	0.0167	0.0000
7つのデータベース7つ の世界	0	1477520	0.1429	2	0.1250	0.1429	0.1339	0.1321	0.1304	0.1286	0.1268	0.1250
データベース:基礎から ネット社会での応用ま で	1	1676881	0.1250	0	0.0000	0.1250	0.0625	0.0500	0.0375	0.0250	0.0125	0.0000
データベース(速水治 夫ほか)	0	1192179	0.1111	1	0.0625	0.1111	0.0868	0.0819	0.0771	0.0722	0.0674	0.0625
データベース(上島紳 一ほか著)	0	1363797	0.1000	0	0.0000	0.1000	0.0500	0.0400	0.0300	0.0200	0.0100	0.0000
データベース(三石大 ほか編著)	0	1626473	0.0909	1	0.0625	0.0909	0.0767	0.0739	0.0710	0.0682	0.0653	0.0625
リレーショナルデータベ ース入門	0	1731628	0.0833	0	0.0000	0.0833	0.0417	0.0333	0.0250	0.0167	0.0083	0.0000
データベース(魚田勝 臣ほか)	0	1011086	0.0769	0	0.0000	0.0769	0.0385	0.0308	0.0231	0.0154	0.0077	0.0000
データベース(西尾章 治郎ほか)	0	938101	0.0714	1	0.0625	0.0714	0.0670	0.0661	0.0652	0.0643	0.0634	0.0625
データベース(溝口徹 夫)	0	1339316	0.0667	1	0.0625	0.0667	0.0646	0.0642	0.0638	0.0633	0.0629	0.0625
グラフ型データベース 入門(Neo4jを使う)	0	1695074	0.0625	4	0.2500	0.0625	0.1563	0.1750	0.1938	0.2125	0.2313	0.2500
やさしい SQL 入門:	0	1673195	0.0588	0	0.0000	0.0588	0.0294	0.0235	0.0176	0.0118	0.0059	0.0000
使えるデータベース・ウ ェブツール	0	1424480	0.0556	1	0.0625	0.0556	0.0590	0.0597	0.0604	0.0611	0.0618	0.0625
データベース(前田隆 編著)	0	1139897	0.0526	0	0.0000	0.0526	0.0263	0.0211	0.0158	0.0105	0.0053	0.0000
医療現場のデータベ ース活用	0	1408694	0.0500	1	0.0625	0.0500	0.0563	0.0575	0.0588	0.0600	0.0613	0.0625

表 C-2: データベース技術を元にした「SQL」検索時の上位 20 件の検索結果のスコア

文献		ベーススコア		リランキングスコア		dに応じた各文献のスコア						
タイトル	適合度	bibid	スコア	閲覧回数	スコア	1.0	0.5000	0.4000	0.3000	0.2000	0.1000	0.0000
SQL 辞典(堀江美彦著)	2	1717455	1.0000	0	0.0000	1.0000	0.5000	0.4000	0.3000	0.2000	0.1000	0.0000
SQL 辞典(松原澤著)	2	1210838	0.5000	2	0.1818	0.5000	0.3409	0.3091	0.2773	0.2455	0.2136	0.1818
SQL(ミック:2016)	3	1707146	0.3333	3	0.2727	0.3333	0.3030	0.2970	0.2909	0.2848	0.2788	0.2727
Head first SQL	3	1326475	0.2500	1	0.0909	0.2500	0.1705	0.1545	0.1386	0.1227	0.1068	0.0909
SQL(ミック:2010)	3	1382519	0.2000	2	0.1818	0.2000	0.1909	0.1891	0.1873	0.1855	0.1836	0.1818
初めての SQL	3	1280034	0.1667	1	0.0909	0.1667	0.1288	0.1212	0.1136	0.1061	0.0985	0.0909
基礎からわかる PL/SQL	3	1640750	0.1429	0	0.0000	0.1429	0.0714	0.0571	0.0429	0.0286	0.0143	0.0000
SQL(ジム・メルトンほか著)	0	1196303	0.1250	0	0.0000	0.1250	0.0625	0.0500	0.0375	0.0250	0.0125	0.0000
アート・オブ・SQL	1	1300743	0.1111	0	0.0000	0.1111	0.0556	0.0444	0.0333	0.0222	0.0111	0.0000
スキリわかる SQL 入門	3	1485573	0.1000	2	0.1818	0.1000	0.1409	0.1491	0.1573	0.1655	0.1736	0.1818
やさしい SQL 入門	3	1673195	0.0909	0	0.0000	0.0909	0.0455	0.0364	0.0273	0.0182	0.0091	0.0000
SQL アンチパターン	2	1475919	0.0833	0	0.0000	0.0833	0.0417	0.0333	0.0250	0.0167	0.0083	0.0000
SQL 実践入門(ミック)	3	1654126	0.0769	0	0.0000	0.0769	0.0385	0.0308	0.0231	0.0154	0.0077	0.0000
リレーショナルデータベース入門	1	1731628	0.0714	0	0.0000	0.0714	0.0357	0.0286	0.0214	0.0143	0.0071	0.0000
SQL 基礎テキスト	3	954342	0.0667	0	0.0000	0.0667	0.0333	0.0267	0.0200	0.0133	0.0067	0.0000
実践 SQL 教科書	3	1147139	0.0625	0	0.0000	0.0625	0.0313	0.0250	0.0188	0.0125	0.0063	0.0000
標準 SQL(JC デイト 1990)	2	294258	0.0588	0	0.0000	0.0588	0.0294	0.0235	0.0176	0.0118	0.0059	0.0000
標準 SQL(JC デイト 1988)	2	1121621	0.0556	0	0.0000	0.0556	0.0278	0.0222	0.0167	0.0111	0.0056	0.0000
SQL スーパーテキスト	3	1222563	0.0526	0	0.0000	0.0526	0.0263	0.0211	0.0158	0.0105	0.0053	0.0000
独習 SQL	3	1331093	0.0500	0	0.0000	0.0500	0.0250	0.0200	0.0150	0.0100	0.0050	0.0000

表 C-3: 基礎数学 A を元にした「対数関数」検索時の上位 20 件の検索結果のスコア

文献		ベーススコア		リランキングスコア		dに応じた各文献のスコア						
タイトル	適合度	bibid	スコア	閲覧回数	スコア	1.0	0.5000	0.4000	0.3000	0.2000	0.1000	0.0000
おもしろいほどよくわかる高校数学	2	1697156	1.0000	1	0.0769	1.0000	0.5385	0.4462	0.3538	0.2615	0.1692	0.0769

基礎解析入門	3	1100637	0.5000	0	0.0000	0.5000	0.2500	0.2000	0.1500	0.1000	0.0500	0.0000
面積と対数	0	926106	0.3333	0	0.0000	0.3333	0.1667	0.1333	0.1000	0.0667	0.0333	0.0000
不思議な数 e の物 語	0	888973	0.2500	0	0.0000	0.2500	0.1250	0.1000	0.0750	0.0500	0.0250	0.0000
関数とはなんだろ う	0	1359191	0.2000	2	0.1538	0.2000	0.1769	0.1723	0.1677	0.1631	0.1585	0.1538
微分・積分30講	3	1312056	0.1667	1	0.0769	0.1667	0.1218	0.1128	0.1038	0.0949	0.0859	0.0769
反復積分の幾何 学	1	1346752	0.1429	0	0.0000	0.1429	0.0714	0.0571	0.0429	0.0286	0.0143	0.0000
数学公式ハンドブ ック(Alan Jeffreery 著 2013)	2	1490398	0.1250	0	0.0000	0.1250	0.0625	0.0500	0.0375	0.0250	0.0125	0.0000
微積分基礎	3	1462512	0.1111	1	0.0769	0.1111	0.0940	0.0906	0.0872	0.0838	0.0803	0.0769
テキスト複素解析	3	1392118	0.1000	1	0.0769	0.1000	0.0885	0.0862	0.0838	0.0815	0.0792	0.0769
工学のための基礎 数学	2	1460905	0.0909	3	0.2308	0.0909	0.1608	0.1748	0.1888	0.2028	0.2168	0.2308
入門微分積分学1 5章	3	1427456	0.0833	1	0.0769	0.0833	0.0801	0.0795	0.0788	0.0782	0.0776	0.0769
微分と積分	3	1373286	0.0769	0	0.0000	0.0769	0.0385	0.0308	0.0231	0.0154	0.0077	0.0000
数学公式ハンドブ ック(Alan Jeffreery 著 2011)	2	1426790	0.0714	0	0.0000	0.0714	0.0357	0.0286	0.0214	0.0143	0.0071	0.0000
複素関数論と複素 整数論	1	1302019	0.0667	2	0.1538	0.0667	0.1103	0.1190	0.1277	0.1364	0.1451	0.1538
リメディアル大学 の基礎数学	2	1360595	0.0625	0	0.0000	0.0625	0.0313	0.0250	0.0188	0.0125	0.0063	0.0000
大学新入生の数 学	2	1446415	0.0588	0	0.0000	0.0588	0.0294	0.0235	0.0176	0.0118	0.0059	0.0000
現代経済学の数 学基礎	2	1443527	0.0556	0	0.0000	0.0556	0.0278	0.0222	0.0167	0.0111	0.0056	0.0000
算数から数学へ	0	1362741	0.0526	0	0.0000	0.0526	0.0263	0.0211	0.0158	0.0105	0.0053	0.0000
数学読本	2	1316532	0.0500	1	0.0769	0.0500	0.0635	0.0662	0.0688	0.0715	0.0742	0.0769

表 C-4: プログラミング演習 I を元にした「文字列」検索時の上位 20 件の検索結果のスコア

文献		ベーススコア		リランキングスコア		dに応じた各文献のスコア						
タイトル	適合度	bibid	スコア	閲覧回数	スコア	1.0	0.5000	0.4000	0.3000	0.2000	0.1000	0.0000
配列, 文字列	0	1141920	1.0000	0	0.0000	1.0000	0.5000	0.4000	0.3000	0.2000	0.1000	0.0000
探索・文字列・計算機何 (R,セジウィック 1996)	2	726049	0.5000	0	0.0000	0.5000	0.2500	0.2000	0.1500	0.1000	0.0500	0.0000
探索・文字列・計算機何 (R,セジウィック 1992)	2	1135596	0.3333	0	0.0000	0.3333	0.1667	0.1333	0.1000	0.0667	0.0333	0.0000
高速文字列の解析の世 界	1	1465755	0.2500	7	0.0574	0.2500	0.1537	0.1344	0.1152	0.0959	0.0766	0.0574
例題で学ぶ Excel 入門	0	1654098	0.2000	0	0.0000	0.2000	0.1000	0.0800	0.0600	0.0400	0.0200	0.0000
競売ナンバー49の叫び	0	1714214	0.1667	0	0.0000	0.1667	0.0833	0.0667	0.0500	0.0333	0.0167	0.0000
カーネル可変量解析	0	1336842	0.1429	1	0.0082	0.1429	0.0755	0.0621	0.0486	0.0351	0.0217	0.0082
テキストデータの統計化 学入門	0	1348501	0.1250	7	0.0574	0.1250	0.0912	0.0844	0.0777	0.0709	0.0641	0.0574
中国学・日本語学論文 集	0	1220502	0.1111	30	0.2459	0.1111	0.1785	0.1920	0.2055	0.2189	0.2324	0.2459
C 言語教科書	0	1482088	0.1000	3	0.0246	0.1000	0.0623	0.0548	0.0472	0.0397	0.0321	0.0246
Word&Powerpoint	0	1649838	0.0909	2	0.0164	0.0909	0.0537	0.0462	0.0387	0.0313	0.0238	0.0164
入門データ構造とアルゴ リズム	0	1573190	0.0833	7	0.0574	0.0833	0.0704	0.0678	0.0652	0.0626	0.0600	0.0574
開眼! JavaScript	0	1572633	0.0769	4	0.0328	0.0769	0.0549	0.0504	0.0460	0.0416	0.0372	0.0328
コーディングを支える技 術	1	1483101	0.0714	7	0.0574	0.0714	0.0644	0.0630	0.0616	0.0602	0.0588	0.0574
Word のイライラ根こそぎ 解消術	0	1394198	0.0667	6	0.0492	0.0667	0.0579	0.0562	0.0544	0.0527	0.0509	0.0492
Excel のイライラ根こそぎ 解消術	0	1409311	0.0625	0	0.0000	0.0625	0.0313	0.0250	0.0188	0.0125	0.0063	0.0000

明快 Java	0	1300748	0.0588	2	0.0164	0.0588	0.0376	0.0334	0.0291	0.0249	0.0206	0.0164
アート・オブ・R プログラミング	0	1460309	0.0556	14	0.1148	0.0556	0.0852	0.0911	0.0970	0.1029	0.1088	0.1148
Ruby	3	1446815	0.0526	9	0.0738	0.0526	0.0632	0.0653	0.0674	0.0695	0.0717	0.0738
R クックブック	0	1439395	0.0500	23	0.1885	0.0500	0.1193	0.1331	0.1470	0.1608	0.1747	0.1885