

平成24年度修士論文題目一覧

ライブラリーサイエンス専攻

学位番号	学位	フリガナ 学生氏名	修 士 論 文 題 目	論文概要
165	修士 (ライブラリーサイエンス)	イワナガ 絵美 岩永 絵美	プリント教材を共有するためのプラットフォームの設計	公開
166	修士 (ライブラリーサイエンス)	キヨハラ カズユキ 清原 和之	現代アーカイブズとレコード・コンティニューム理論	公開
167	修士 (ライブラリーサイエンス)	サトウ シンジ 佐藤 晋之	変わりゆく電子環境下における音楽アーティスト自身による音情報の管理・提供に関する研究	公開
168	修士 (ライブラリーサイエンス)	タナカ トシエ 田中 要江	大学内システム連携におけるデータ同定の一手法の提案	公開
169	修士 (ライブラリーサイエンス)	チョウ ダイカイ 張 大海	OPAC の検索ログに見られる検索語間の関連性	公開
170	修士 (ライブラリーサイエンス)	ナカオ ヤスロウ 中尾 康朗	レファレンスサービスにおけるユーザーのための検索支援に関する研究	公開

プリント教材を共有するためのプラットフォームの設計

A design of platform for sharing printed text papers among teachers

2FS11055Y 岩永 絵美 IWANAGA Emi

公立学校ではインターネットへの接続やグループウェアの整備が進められてきた。学校内のグループウェアは、校内LANの活用により、教育情報の作成に関する教員のアイデアのデータベース化と情報共有を実現させ、教員の教育活動の効率化を実現させる。教育情報を共有するモデルの一つに、日本の初等中等教育におけるあらゆる教育情報が一元的にデータベース化された教育情報ナショナルセンターNICERがあった。NICERは、我が国におけるあらゆる教育・学習情報を扱う中核的なWebサイトであったが、2011年3月に運用が終了してしまった。これは、教育情報流通の必要性が否定されたわけではなく、初等中等教育に関する教育情報の蓄積や流通に関して、NICERに代わる新たなモデルの構築が検討されている。

教育情報を一元的に集約するシステムには、学習指導要領をもとに作成される様々な教材を集める必要がある。教材には、教員一人一人の授業設計を表した学習指導案や、動画、静止画、音声など授業をおこなうために必要な素材が含まれている。これらを教員の共有財産としてインターネット上で教員等が活用していくためには、システムの利用者となる彼らの意見を反映させたユーザインターフェースを設計することが必要だと考えられる。

教育情報の一つであるプリント教材の体系的な管理に関しては、そのデータベースの活用が小中学校の授業で有効であることが報告されている。本研究は、教育情報データベースに対する高校教員のニーズを調査するとともに、教員が教材を探すのに有効な分類や情報共有するのに有益な付加情報を分析し、それらを反映させたユーザインターフェースの概要設計までおこなう。

第2章では、教員が教材を探すのに有効な分類と情報共有するのに有益な付加情報を分析する。インターネット上で高校の授業向けに作成されたプリントを公開する既存のシステムを分析した結果、教材は学習指導要領に沿うように分類されていた。さらに、プリント教材の対象学年をもとに検索できる機能が設けられているものも確認できた。この分析結果をもとに高校教員にインタビューした結果、英語・数学・社会・理科に関しては学習指導要領の項目によるプリント教材の分類が適切であり、国語は題材とする作品名や作家名が有効な指標であることがわかった。この分類さえ実現されていれば、プリント教材によって学習できる内容を知ることができるため、学年別による検索機能は重要でないということであった。

プリント教材を共有することに対する教員の意識に関しては、否定的な考え方も見られた。社会科の教員には、歴史に対する解釈に個人の思想が反映されることを理由に、国語や英語に関しては、生徒の読解力向上のために新書など新しい文章を用いてプリント教材を作成しているため、著作権処理が手間であることを理由に、プリント教材の共有に消極的な教員が確認された。それに加えて、プリント教材に図画を用いたい場合でも、インターネット上の著作権フリーのものを手に入れられることや、教科ごとの勉強会に参加してベテランにノウハウを伝承してもらうなど、自分のネットワークで解決してしまっているように思われた。しかしながら、生徒の学習力定着のためには多くの演習プリントを提供する必要があることや、ベテランの教材を参考にしたいという欲求があることを理由に、プリント教材を共有するシステムを通して教材作成の効率化やアイデアの伝承を実現させる考え方に肯定的な教員も確認できた。

プリント教材に付加される情報として関心が寄せられたものは、プリント教材に生徒が書き込んだ情報やプリント教材の作成者が普段の授業で使用している教科書等の情報であった。生徒が書き込んだ情報は、生徒の理解度や思考過程を知る一つ的手段として有益であることがわかり、教科書等の情報を記載することに関しては、プリント教材の内容の難易度を判断する一助となることや、同じ教科書を使用している教員のプリント教材を参考にすることができるという意見が寄せられた。

第3章では、プリント教材を共有する際に高校教員が必要とする機能を搭載したユーザインターフェースについて、図を用いながら説明する。

今回は教育情報を一元的に集約するシステムに着目し、教育情報の一つであるプリント教材を対象とした。そして、どのように分類しておけば教材を見つけやすいか、また、プリント教材と併せて参照すると有益な情報は何かという二点について、既存のシステムを分析と高校教員へのインタビューによって調査した。その結果、国語・数学・英語・理科・社会の5教科に適切な分類基準や、生徒の書き込み情報など有益な付加情報を選定し、システムの利用者となる教員の意見を反映させたユーザインターフェースの概要を示すことができた。今後の課題は、検証することができなかった教科の高校教員を対象に分析をおこない、設計に反映させることである。

現代アーカイブズとレコード・コンティニューム理論

Today's Archival Thinking and the Records Continuum Theory

2FS11056G 清原 和之 KIYOHARA Kazuyuki

レコード・コンティニューム理論は、現代のアーカイブズ学に鮮烈なインパクトを与えた、これからの21世紀を展望しうる試金石である。そして、表題のもう一方の「現代アーカイブズ」には、本論文の目的と関わる二つの意味を込めている。まず一つは、近代からポストモダンを経て現代に至るアーカイブズ理論の系譜のなかにコンティニューム理論を位置づけること。そして、二つ目に、組織や集団の流動化、国民国家の退潮、多文化主義、ICT環境の進展といった現代的課題への応答として、この理論の可能性を提示することである。

1990年代後半にフランク・アップワードによって提唱されたレコード・コンティニューム理論は、電子記録時代に適合する新たな概念として迎えられることとなった。現用記録を指すレコードと非現用記録を指すアーカイブズという領域区分を統合した点がこの理論のひとつの特徴であり、こうしたプロセスの統一的把握によって、容易には捕捉不可能な電子記録の記録生成段階からの管理を可能とする、レコードキーピング・システム構築のための理論的基盤が提供されたのであった。しかし、作成の瞬間からレコードキーピング・システムに取り込まれ、社会的に共有しうる記録とするという戦略は、組織母体の観点を特権化するものであり、レコード・コンティニューム理論の本来の特徴である記録への諸主体の観点の多元性という点が十分に反映されないのではないかとと思われる。この記録への多元的な意味付与という視角は、2000年代以降、アーキビストたちがマイノリティの権利獲得やアイデンティティの再生、記憶の創造のためのレコードの価値を新たに発見するなかで強調され、彼らは、これまで専門職集団として蓄積してきた専門知識を自己批判的に再考し、改めてアーカイブズ学の方向性を模索するに至っている。

以上のようなアーカイブズ学の動向を踏まえ、本論文では、レコード・コンティニューム理論がどのように深化、発展させられてきたのかを跡づけ、その将来的な可能性を追求した。

まず、第一章では、近代から現代にいたるアーカイブズ理論の変遷を、客観性と主観性との相克という観点から、批判的に検討した。近代アーカイブズ理論の確立に大きく寄与したオランダの「ダッチ・マニュアル」、サー・ヒラリー・ジェンキンソンの『アーカイブ管理マニュアル』、セオドア・シュレンバーグの『近代アーカイブズ』という三つの古典は、記録資料の客観性を自明の前提としてアーカイブズ管理の方法論を構築してきたが、記録に関わるあらゆる行為主体の主観性の問題は検討の対象とされるこ

とはなかった。この問題を顕在化させることとなったのが、ポストモダン・アーカイブズ理論といわれる潮流である。ポストモダン・アーキビストたちは、記録作成者やその継承者、アーキビストやユーザーといった記録伝達の過程に関わるあらゆる主体による記録への意味付与、解釈の可能性を論及することとなった。

レコード・コンティニューム理論は、そのような記録に関わる様々な諸主体の意味付与、解釈の問題を包括的に捉えるモデルとして構築された。第二章では、コンティニューム理論とはどのようなものなのか、その誕生の土壌となったオーストラリアの実践的経験、アップワードがこの理論を補完するものとして提示したアンソニー・ギデンズの「構造化理論」との関連、さらに、理論的な深化に貢献したスー・マケミッシュによる個人の証拠に関する議論を通して、詳細に明らかにした。その際、強調したのは、レコードキーピングの過程は、それに関わる諸主体のメタナラティブとして理解しうること、そして、その記録に関わる諸行為は記憶の想起と忘却のプロセスと不可分に結びついているということである。

第三章では、2000年代後半以降、盛んに論じられてきているコンティニュームの第4次元「集合的記憶」に関わる問題を考察した。近年では、①記録に対する解釈の多様性、②複数の主体による集合的記憶の共有、③レコードのプロセスの非線形的、循環的な性格に注目する研究がなされており、コンティニューム理論の線形的な理解は批判されてきている。そこで、まず、コンティニュームの各次元間の意味について、記録作成主体とその記録の対象者との合意をめぐるせめぎ合いの観点から再検討した。続いて、記録の当事者間を超えた、他者をも含んだ領域である第4次元の問題を迫った。そして、ジャネット・バステアンの「記録の共同体」概念、及び、エリック・ケテラルの「協同遺産」という概念に注目し、記録作成者とその対象者、さらには、第三者的他者とも共有可能な記録をめぐる、不可避的に生じてくる問題である「記憶のポリティクス」について考察し、その「記録の活性化」の側面とともに、いかに異なる他者間で秩序が維持されうるのかについて、デモクラシーとアーキテクチャに注目して検討した。

最後に、乱立・競合するメタナラティブの調節はいかにして可能か、という問いに対する結論として、第三者的「市民」による、複数性の場における継続的な対話の必要性と、そうした参加や選択への負担を軽減しうるアーキテクチャによるデモクラシーによる解決、という暫定的な回答を提示した。

変わりゆく電子環境下における音楽アーティスト自身による音情報の管理・提供に関する研究

Research into the management and provision of sound information by the artist in the changing electronic environment

2FS11057R 佐藤 晋之 SATO Shinji

1. 背景と目的

効果音や楽曲等の電子データ（音情報）の提供方法は、CD等のパッケージからネットワークでの配信へと移行しつつある。ブロードバンド環境が一般家庭に普及したことや、MP3等のファイル圧縮技術向上等によって、音情報は、インターネットを介し自由に流通させることが可能なコンテンツとなった。

インターネットを利用することで、低コストに音情報を提供できるようになり、個人で活動する音楽アーティストが増加すると予測されるため、本研究では、個人で活動する音楽アーティストを対象にした。

本研究では、個人で音情報を管理・提供している音楽アーティストの現場の声と一般社団法人日本レコード協会が実施したユーザーの利用実態調査から、流通の変化による現状や問題点、課題を明らかにすることで、音楽アーティストとユーザー双方の利益を保証する新しい流通の仕組みを提案することを目的とした。

2. 方法

音楽アーティストとユーザーの利益を定義し、双方の利益を保証する流通の仕組みを提案するため、まず個人で活動中の音楽アーティスト及び関連する分野の者を対象にインタビュー調査を実施し、次に一般社団法人日本レコード協会がユーザーに対して実施した利用実態調査の結果を分析した。

2.1. インタビュー調査

調査対象者は、音楽アーティスト及び関連する分野の者とした。選定基準は、音情報を管理・提供することで生計を営む個人又は小規模組織で活動する者であり、専業で生計を営んでいると考えられる、年収200万円台から3,000万円台の27歳から55歳までの男女20名とした。

インタビュー調査の結果、個人で活動する音楽アーティストは、クライアントワークやコンサート、グッズ販売等を主な収入源としており、マネジメント次第で生計を営むことができる可能性があることが分かった。

2.2. ユーザーの利用実態調査結果分析

ユーザーが求める音情報の聴取方法を明らかにするため、一般社団法人日本レコード協会の「2011年度音楽メディアユーザー実態調査」（2012年2月）及び「2011年度違法配信に関する実態調査」（2011年3月）の結果を分析した。

実態調査結果を分析した結果、ユーザーが求める音情報の聴取方法は、レンタルCDよりも安価に入手でき、嗜好に合った音情報を容易に発見可能であり、何時でもどこでも容易に聴取できる方法であることが分かった。

3. 提案

個人で活動する音楽アーティストの利益は、音情報が認知され、クライアントワークやコンサートの集客等に繋がること、また、正当な評価が反映される市場が形成されることであった。ユーザーの利益は、音情報を何時でもどこでも容易に聴取可能である等の利便性やレンタルCDよりも安価であるといった経済性、嗜好に合った音楽アーティストの音情報を発見できることであった。これらのことから、双方の利益を保証するダウンロード型評価制音楽配信サービスを提案する。

ダウンロード型評価制音楽配信サービスでは、音情報への対価をユーザーからの評価とする。SNSと連携することにより、音情報の聴取履歴や評価といった蓄積された情報が、ユーザー同士で共有されやすくなり、音情報の認知に繋がることから音楽アーティストとユーザー双方の利益を保証したサービスと言える。

4. 結論

現在、音楽産業は転換期にあり、音情報の管理・提供方法や流通の仕組みの変化により、個人で活動する音楽アーティストが増加することが予想される。そのため、個人で活動する音楽アーティストとユーザーの利益に焦点を当て、評価等の情報を蓄積することによって、間接的に収益化する仕組みを提案した。

今後の課題は、個人の音楽アーティストのみではなく、著作権使用料を主な収入源とするレコード会社等の組織や団体にとっても利益のある流通の仕組みを検討することである。

大学内システム連携におけるデータ同定の一手法の提案

A Proposal of Methods for Identifying and Linking Data among Information Systems in University

2FS11059E 田中 要江 TANAKA Toshie

1. はじめに

平成 16 年の国立大学法人化以降, 日本の国立大学を取り巻く環境は大きく変化している. 文部科学省は, 社会との関わりの中で新しい大学づくりに向けた改革として『大学改革実行プラン』を提示している. これは, 今後の大学組織の方向性を示しているものであり, 具体案については各大学の個性や特性を考慮し, 大学が主体的に行うよう任されている. 指針としては, 「大学の質保証の徹底推進」「学業に専念できる環境の整備し, 学生の学修時間を欧米並の水準に」などが挙げられている. また, 「大学改革を促すシステム・基盤整備」では, 大学情報の公表の徹底など説明責任を果たす必要がある. このような業務を進めるうえでは, システム自体を見直す必要があり, 今後の運用に合わせて従来のシステムを変更する必要も出てくる. 現在のシステムは, 目的に合わせて適材適所で運営されているという特性はあるが, 近年の動きにみられる他目的の情報活用や部局を超えた様々なデータ活用には対応していないことが多い. そのため, 学業に専念できる環境の整備や研究成果などの大学情報を外部に発信することをより効果的に行うひとつの方法が, システムの連携だと考えている. 連携することで, 関連情報を提供する導線を確保し情報の不透明性を解消することとなり, 情報提供の環境をより充実させることができる. 以上により, 大学組織内におけるシステム連携についての手法を提案することとする. 本研究では, 特に図書館のサービスと関連している「研究者情報と九州大学学術リポジトリとの連携のための共著分析」と「シラバスシステムと図書館蔵書検索システムの連携」の手法について提案する.

2. 研究者情報と九州大学学術リポジトリの連携のための共著分析

近年, 「大学の質保証の徹底推進」や「学業に専念できる環境の整備」などが進められている. 現在, このような指針に対応するためには, システムのあり方も見直す必要がある. 既存の情報システムは, それぞれの役割のため独立して運用されているため, 部局を超えてシステム連携を行うには適したデータ構造になっていない. 例えば, 教員単位で情報を収集している研究者情報システムや論文単位で収集している学術リポジトリシステム, 講義単位で

管理しているシラバスシステム等, 様々なデータ構造が存在している. 上記で示したシステム連携の利点を実現するためには, 各システムのデータ同士のリンクが必須となる. データをひとつひとつ確認しながら手動で行うことは不可能ではないが, データ量や作業時間, コスト面を考えると現実的ではない. そこで, 異なる二つの文字列を同一のものと判別しリンクさせるといった現実可能性を目指して, その基礎となる手法を提案し, その有効性を検証する.

3. シラバスシステムと図書館蔵書検索システムの連携

現在, 大学ではさまざまな学習支援が行われており, 学習に必要な情報を入手しやすい環境を構築していくことも重要な課題となっている. その機能を果たす役割を担っている部署の一つとして図書館があり学習支援の内容も社会の流れに合わせて充実させていく必要がある. また, 学習支援を行うためにはシステム環境を整えることも不可欠となっており, 過去にさまざまな取組みがなされた結果, 充実したデータ量が蓄積されたと言える. これからは, その情報を利用者が活用しやすい環境にすることが重要だと考えており, 今回は, その方法としてシステム連携を挙げる. 現在は, 大学内にさまざまな情報(データベース)が散在しており, 学生に有用な情報があるにもかかわらず, その情報を見つけれない利用者が多数存在しているといった現状がある. そこで, これらの問題をシステム連携により解決したいと考えている.

4. おわりに

本研究では, 学内で運営しているシステムの中から, 研究者情報システムと九州大学学術リポジトリを連携するための共著分析と, シラバスシステムに掲載されている図書を図書館蔵書検索システムから簡単な操作でその図書を導き出すための連携の提案を行った. その他にも, 学内システムで連携がもためられているものは多く存在している. 具体的には, シラバスの講義情報と研究者情報の担当授業科目を連携させることができれば, 閲覧者に提供できる情報も増える. また, シラバスからデータのインポートが可能となれば, 入力者の手間も減少される. 文字列を利用しデータを同定するシステム連携は, 現段階では, ひとつひとつ丁寧に検討していく必要があり, 連携の目的やシステムの特徴により連携方法も全く異なるため, これからも様々な手法を試す必要があると考えている.

OPAC の検索ログに見られる検索語間の関連性

Relevance between Search Queries in the Access Log of OPAC

2FS11060W 張 大海 ZHANG Dahai

1. はじめに

OPAC 等で論文を検索する際、適切なクエリを思いつかないことにより満足する情報を取得できないという経験をするユーザは多い。検索クエリが適切でないことにより、検索が失敗する状況は以下の二つがある。

- 1) 検索結果が膨大であり、ユーザがすべての情報を閲覧することが困難な場合。この場合、検索結果の中から自分の目的とする情報を見つけ出すのは困難である。
- 2) クエリに該当する検索結果が非常に少ない（あるいはまったくない）場合。

このような検索が失敗する状況でユーザを支援するために、本研究では、九州大学附属図書館における次世代 OPAC である「Cute.Catalog」のアクセスログを利用し、ログデータから「クエリ q_1 による検索のやり直しとなっている検索（本研究ではこれをリトライ検索と呼ぶ）のクエリ q_2 」というクエリ対 (q_1, q_2) を抽出する。これを利用して、クエリ q_1 で検索したユーザに q_2 を検索クエリの候補として推薦することが考えられる

2. 関連研究

本研究と関連するまたは類似する研究は数多く行われており、目的も様々であるが、本研究で用いる「ある検索のリトライとなっている検索のクエリ」を利用したクエリ拡張の研究は著者が知る限り、他では詳細な研究は行われていない。

シンガポールの Nanyang Technological University の図書館 OPAC を対象とする分析及び検討が行われている。この研究によると、図書館 OPAC のユーザが、AND 検索を利用して検索するのは全体の 10%程度である。OR 検索を利用した検索はもっと少なく、1%未満しかない。アメリカのイリノイ大学図書館 OPAC を対象とした研究でも、調査したトランザクションログ 1,400,000 回の検索内わずか 0.22%が OR 検索であったと報告されている。したがって、アクセスログ中の OR 検索を利用してクエリ推薦を行うのは現実的でないことが分かる。

3. 提案手法

本研究では、OPAC のアクセスログからクエリ対 (q_1, q_2) が、「 q_2 はクエリ q_1 のリトライ検索のクエリ

である」という性質を持つかどうかの判定法を提案する。

クエリ q_1 による検索とそのリトライ検索のクエリ q_2 に対して、両検索の時間差は二つの部分から構成されると考えられる。新たなクエリを考える時間とユーザがクエリを修正する時間である（文字列として異なるほど修正に要する時間がかかる）。二つの検索クエリの文字列としての違いは、一種の編集距離を利用して測る。本研究で利用する編集距離はレーベンシュタイン編集距離とは違って、挿入・削除の操作のみを許可し、置換を禁止するタイプの編集距離である。

本手法は q_2 が q_1 のリトライ検索のクエリである場合の二つの検索の時間差を、 q_1 と q_2 の文字列としての相違度から回帰分析を利用して予測し、予測値と実際の時間差との関係から q_2 が q_1 のリトライ検索のクエリであるかどうかを判定するものである。また、検索クエリを修正するときにかかる時間をクエリ q_1 と q_2 に依存せず、一定としたものをベースライン手法とし、提案手法との比較を行う。

4. 評価

本研究で利用するデータは九州大学附属図書館 Cute.Catalog の 2012 年 1 月から 6 月までのアクセスログである。データの総件数は 155,673 件である。本研究は 1 月から 4 月までのログデータを用いてトレーニング用データを作成し、6 月のログデータを用いて評価用データを作成した。

提案したリトライ検索の抽出手法と前述したベースラインの手法により評価用データ (821 件、その内リトライ検索 214 件) に対してリトライ検索を抽出する実験を行った。提案手法の精度は 44%、再現率は 40%であった。また、ベースラインの手法との比較により、提案した手法の有効性を示した。

5. 今後の課題

今回は提案したリトライの抽出手法について有効性の検証を行ったが、抽出された「 q_1 による検索のリトライ検索のクエリ q_2 」というクエリ対 (q_1, q_2) が OPAC のユーザにどの程度役立つか評価していない。抽出されたリトライ検索のクエリを用いたクエリ推薦システムを作成し、その有用性を検証することが今後の課題である。

レファレンスサービスにおけるユーザーのための検索支援に関する研究

A study on the information search support for users in reference service

2FS11061P 中尾 康朗 NAKAO Yasuro

近年、情報通信技術の発展と資料情報のデジタル化の進展に伴ってユーザー自身が自ら情報探索する形態が増えている。ユーザーと図書館と情報源をめぐる関係が変容する中、従来型のレファレンスサービスは広義のレファレンスサービス、情報サービスへ拡張されてきている。本稿ではこのような動向を踏まえた上で今後のレファレンスサービスを「個別支援型レファレンスサービス」「発信型レファレンスサービス」「情報リテラシー教育」の3領域で捉える。そして今後の予測として①ユーザー自身による情報探索へのより効率的な支援、②情報リテラシー教育によるユーザーの探索能力の強化、③それらで対応できないより高度なレファレンスサービスへの人的資源の投入といった重点化を予想する。

ここでユーザーの自立的な情報探索にとってシステムによる支援効果が期待できるのが発信型レファレンスサービスの領域である。発信型レファレンスサービスには電子的なレファレンス事例とパスファインダーの提供がある。本稿ではこれらレファレンス事例とパスファインダーを利用し、ユーザーの情報探索への支援システムを提案する。

提案モデルのプロトタイプを実装するにあたり国立国会図書館「レファレンス協同データベース」の一般公開レファレンス事例と「リサーチ・ナビ」の調べ方案内を対象データとした。

まず、レファレンス事例とパスファインダーデータに関する特性と課題点について考察した。ユーザーの情報探索支援のためには最終的に関連性の高いパスファインダーを探索する必要があるが、質問文に基づくクエリではノーヒットやヒット多数、質問文に基づくクエリでそもそもたどり着けないといったケースがあることがわかった。ユーザーの視点から考えた場合、これは単に検索結果が別々になるというだけでなくユーザーの質問に基づくクエリで検索しても解決につながる関連性のあるパスファインダーにたどり着けない可能性があることとなる。

パスファインダーは、特定のテーマについて組み立てた調べ方の典型的なパターン集である。だがパスファインダーは多様なケースに幅広く対応できるようより一般的な内容にまとめられ縮退化されている。それに対してレファレンス質問というのは非常に個別的なものであるため、個別的なレファレンス質問に基づくクエリで検索しても求めたいパスファインダーにはたどり着けない可能性が出てくる課題点をもっている

一方、レファレンス質問と回答の記録であるレファレン

ス事例についても非常に個別的なものであるため、類似した事例が再現することはあっても、完全に同一の質問が再現することはない。そのため過去のレファレンス事例をそのまま新規の質問に単純に直接適用することはできない。

パスファインダーのもっている課題点を改善するためには、非常に個別性が高いレファレンス事例の特長を利用し、パスファインダーへの探索過程にレファレンス事例を連携させパスファインダーのもつ短所を補完することで課題解決を図ることが考えられる。

そこで両データを連携するシステムを実装する上でレファレンスツールに着目した。レファレンスツールを印刷体のツール、ネットワーク型のツールに分け、それぞれレーベンシュタイン編集距離、階層距離によって、文字列の同定を行って、共通のレファレンスツール集合を特定した。完全一致の事例が多かったが、完全一致だけでは無理なものについては両者の非類似度を測定することとした。印刷体ツールについてはレーベンシュタイン編集距離を、ネットワークツールについては階層距離計算を利用した。集合を取得しそれらに紐付くレファレンス事例、パスファインダー両者のレコードのペアの特定を行った。それをもとに両データを連携させたプロトタイプ検索システムを構築した。

評価実験では、テストデータとしてランダムに選んだレファレンス事例の質問120件について当該事例が持っているレファレンスツールに紐付くパスファインダー群をその質問の正解データとして準備した。現行システムと同じレファレンス事例と連携させずパスファインダーのみを検索した「直接検索」とレファレンス事例を介してパスファインダーを検索した「連携検索」場合を検証した。さらに関連語を追加してクエリ拡張をおこなった「クエリ拡張検索」を検証した。

再現率に関しては、直接検索モデルに比べ連携検索モデルの方が上昇した。提示されたパスファインダー数については、直接検索モデルの提示数に比べ連携検索モデルの提示数が上昇した。クエリ拡張に関しても再現率、提示数に改善がみられた。一方で精度自体が10%未満と低い値にとどまった。連携検索により再現率が上がって検索結果集合が増えても含まれる正解パスファインダーの割合には変化がみられない可能性が考えられる。検索結果集合に含まれる正解パスファインダーがわずかしか含まれないことから、適合率の改善、正解につながる外部知識による拡張、検索結果の絞り込みや並べ替えの工夫が必要であると判断される。