

情報探索プロセス分析に基づく協調的情報検索支援のための履歴共有システムの構築

著者	松村 敦
著者別名	Matsumura Atsushi
発行年	2013
その他のタイトル	Developing the Search History Sharing System for Collaborative Information Retrieval based on Analysis of Information Search Process
URL	http://hdl.handle.net/2241/120746

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 6月 1日現在

機関番号：12102
 研究種目：基盤研究（C）
 研究期間：2010～2012
 課題番号：22500218
 研究課題名（和文） 情報探索プロセス分析に基づく協調的情報検索支援のための履歴共有システムの構築
 研究課題名（英文） Developing the Search History Sharing System for Collaborative Information Retrieval based on Analysis of Information Search Process
 研究代表者
 松村 敦（MATSUMURA ATSUSHI）
 筑波大学・図書館情報メディア系・助教
 研究者番号：40334073

研究成果の概要（和文）：利用者の情報探索プロセスを実験により取得し、履歴の活用方法に着目して分析した。その結果、履歴は、時系列を考慮し、画像中心で表現し、探索の経路を木構造によって提示する形式が有効であることをが分かった。また、利用者の好みを日本十進分類法を利用して表現する書籍を対象とした履歴共有システムの構築を行った。実験によりベクトル空間法などの従来手法より利用者が有用と思う人物を精度よく推薦できることが分かった。

研究成果の概要（英文）：We analyzed use of the users' search history, paying attention to users' information search process. As a result, it was effective that the search history is expressed visually in consideration of a time series and is shown by a tree structure. Moreover, we build a shared history system for the books that can represent user preference using Nippon Decimal Classification. As a result of the experiment, it turned out that our system can recommend persons useful for the user with accuracy more sufficient than the conventional techniques, such as a vector space method.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	2,000,000	600,000	2,600,000
2011年度	600,000	180,000	780,000
2012年度	600,000	180,000	780,000
年度			
年度			
総計	3,200,000	960,000	4,160,000

研究分野：情報学

科研費の分科・細目：図書館情報学・人文社会情報学

キーワード：情報検索、情報推薦、情報行動、検索履歴、キュレーション、履歴共有、振り返り、学習支援

1. 研究開始当初の背景

技術の発展、利用者のニーズの多様化にもない多機能・高機能な情報サービスが増加

し、利用者の情報探索プロセスもますます複雑になってきた。例えば、NIIの提供する文献情報ナビゲータ CiNii では、探索プロセス

において関連するキーワード、文献、著者など、さまざまな情報が推薦される。また、Google のようにシンプルな検索エンジンにも、関連キーワードや関連情報を提示するサービスが付加されてきている。利用者は、効率よく情報にたどり着くために、これらの情報を適切に利用するスキルを身につける必要がある。

このような複雑な情報探索プロセスを支援するためには、利用者の残した履歴は重要な役割を果たすと考えられる。

Web ブラウザの閲覧履歴を監視して情報検索の過程を抽出してフローチャートとして視覚化した研究がある。この図を参照することによって、情報検索の過程を容易に振り返ることが可能となる。しかしながら、このようなフロー図を一般の人が理解し、検索の効率化につなげるよう活用していく方法は未だ提案されていない。

また、エキスパートの検索履歴が初心者を助ける効果などを狙った協調型の情報検索の研究が盛んに行なわれている。しかし、実際の利用者がどのように履歴を活用し、情報探索を行なっているかは明らかにされていないため、適切な履歴活用システムの実現には至っていない。

このような問題に対して、我々は自分と他人の文献検索履歴として、キーワードとその適合文献のリストを提示する手法を提案している。これを機関リポジトリ横断検索システムとして実装し、大学図書館にて運用を行なっている。このシステムを評価する過程で、

不慣れたタスクでは他人の履歴の活用が活発であるなど、履歴の共有に効果があることが分かってきた。しかし、検索キーワードの履歴にしても文献リストにしても、検索者の検索意図や適合判定の意図といった文脈が明示されないため、有効な利用ができないという問題が生じている。

一方で、我々は「ソーシャルリンク」というコンテンツの組織化の質を向上させる手法の研究を行ってきた。この手法は、ソーシャルタギングで問題となる曖昧性や不完全性を補い、緻密で表現力のあるコンテンツの組織化が可能である。また、ソーシャルリンクは、複数の利用者が活用するための協調的な仕組みも備えており、組織化されたコンテンツを共有することが可能である。そこで、本研究で扱う履歴情報の構造化・共有においてもソーシャルリンクを適用可能と考えた。

2. 研究の目的

(1) 情報探索プロセスにおける利用者の検索履歴の活用行動の解明。

検索履歴をどのように利用しているかを把握するための調査および実験を行なう。こ

れまでに、我々は検索履歴を活用する機関リポジトリ横断検索システムを実装し、検索ログ分析と発話思考法を利用して利用者の検索履歴の活用傾向について分析を行なっている。これによって、他人の履歴と自分の履歴の使い分けがタスクの種類によって大きく異なる傾向を示すことが分かっている。本研究では、さらに利用者の視線情報を測定する実験を追加し、ログや発話に現れない履歴の使い分け意図を明らかにする。具体的には、履歴を「検索キーワード」、「検索結果ページ」、「閲覧ページ」に分け、これらの活用のタイミング（検索ログ）、活用意図（発話、視線）を分析し、いつ、なぜ、どのような履歴活用行動を行なったかを明らかにする。

(2) 検索履歴を共有するための履歴構造化手法とアクセス手法の検討。

検索履歴を複数人で共有するためには、適切な構造化（組織化）とアクセス手法が用意されていなければならない。具体的には、(1)で分析した「検索キーワード」、「検索結果ページ」、「閲覧ページ」のそれぞれの履歴に対して構造化とアクセス手法を検討する。予備的な分析から、利用者の期待する構造化には、時系列的な関係性、意味的な文脈の2点があげられる。時系列的な関係性は、蓄積された検索ログから自動的に抽出・付与できるが、これらを効果的に構造化する方法を検討する。一方、意味的な文脈は、個々の利用者が手動でタグを付与し共有する、いわゆるソーシャルタギングと、我々が提案するソーシャルリンクの適用可能性を検討する。

(3) 履歴の文脈を考慮した検索履歴共有システムの構築。

以上で検討した手法を実装する。ベースとなるシステムは申請当時運用中の機関リポジトリ横断検索システムとし、履歴構造化部分とアクセス部分の拡張を行なう。本システムを、NTCIR などのテストコレクションを利用して客観的な検索精度を評価するとともに、履歴構造化・アクセス部分についてのユーザビリティ評価も行ない、最適な検索履歴の構造化・アクセスシステムの構築を目指す。

3. 研究の方法

(1) 検索履歴をどのように利用しているかを把握するための実験を行なう。C i N i i ウェブAPIを利用して、クエリの履歴と検索結果の重複を表示する論文検索システムを実装する。このシステムによって発話思考法を用いたユーザビリティ評価及び検索実験を行なう。

(2) 予備的システムを実装し、履歴共有のための機能の検討を行う。具体的には、履歴共有の一例として掲示板での議論を題材とし、議論を時系列に沿って構造化するシステムを実装し、その利用方法を調査しする。さらに、検索履歴の表示方法の検討のために、タイトルとスニペットによる表示と画像を中心とする表示の2つのシステムを実装し評価する。

(3) 履歴の提示形式を検討するために、履歴を可視化するシステムを実装し、発話思考法を用いて検索実験を行う。可視化の形式として、木構造の有効性を検討する。

(4) 利用者がどのような履歴に興味を持つかを明らかにするために、利用者の知識構造や好みの推測を行う。研究者と一般利用者の2つの例を対象に行う。具体的には、国文学研究者に焦点をあてた古典籍の構造化システムの実装と評価、および一般利用者に焦点をあてた読書履歴共有システムの実装と評価を行う。

(5) 履歴活用のためのインタフェースの検討のために、文献検索をインタラクティブに行うシステム、履歴をビジュアルに提示するためのまとめシステムを実装し評価する。

(6) 履歴には能動的に操作をし意識に残っているものと、そうでないものがあり、これらの活用のタイミングは異なる。そこで、履歴活用タイミングを検討するために履歴再提示システムの実装を行う。

(7) 履歴共有システムを実装し、予備実験により評価する。対象は書籍とし、20名の被験者を対象に実験を行う。

4. 研究成果

(1) C i n i i ウェブAPIを利用して、クエリの履歴と検索結果の重複を表示する論文検索システムを実装して、検索履歴をどのように利用しているかを把握するための実験を行った。その結果、ユーザビリティ評価によって、履歴の選択的消去を可能とし、ページ番号の表示方法を改良したが、このような微修正が、評価実験において高く評価された。特に、クエリ履歴の消去といった編集操作が利用者によって行なえることは検索行

動と発話から重要であることが明らかとなった。一方、利用者になじみのある課題とない課題とでは履歴の活用方法に大きな違いはなかった。

(2) 時系列に沿って内容を表すキーワードの2次元配置によって議論の内容を把握する手助けとなることが示され、これによって時系列とキーワードによる履歴表現が有用となる可能性を示した。さらに、検索履歴の表示方法として、タイトルとスニペットによる表示と画像を中心とした表紙として提示する方が内容の把握にも貢献すると同時に、印象的な画像を含む表紙を利用者が好むことが明らかになった。

(3) 履歴を木構造として視覚化するプロトタイプシステムを実装し、このシステムで発話思考法を用いた検索実験を行なった結果、ユーザが履歴を活用するのは、過去に閲覧したページや結果一覧を再度見直す場合であり、特に、履歴を可視化しておくことによって、直前よりも過去の検索結果を振り返って利用する行動が起ることが明らかになった。これは、過去の検索行動を利用するには従来のシステムでは不十分であることを示す結果と言える。この結果を利用して、自然に履歴を振り返ることが可能な履歴提示インタフェースを開発・実装した。本システムは、Web閲覧時にも常に履歴を視野に入れることで履歴の構造の記憶と理解を促すものである。実験のために、画面を切り替えて履歴を閲覧するシステムを実装し比較した結果、提案システムの方がより過去の履歴にアクセスする傾向があることを示すことができた。

(4) 国文学研究を例に検索行動を分析したところ、既に権威付けられた古典籍の文献学的分類を辿って探索が行なわれていた。履歴の共有にはこのような共通の知識構造の存在が重要になる。一方、読書履歴の共有では、レビューの長さの好みや読んだ書籍の分類に従って、次の読書行動の選択が行なわれていることが明らかになった。履歴を共有して適切な行動を支援するためには、共通の知識構造の構築と個人の嗜好情報を活用した推薦技術が必要であることが分かった。

(5) 文献検索を継続的に行うためにキャラクターによるインタラクションが効果があることが示された。さらに、履歴の内容を分かりやすく表現するためにWebをビジュアル

にまとめる手法が効果的であることが示された。

(6) 能動的な履歴として、ブラウザブックマーク、twitter の投稿とお気に入り、flickr の写真などを再提示可能な統合的環境を実装した。各システムから自動的にデータを収集し、ユーザに再提示をしてくれる。履歴再提示のタイミングは、完全にランダムに提示する方法をとっているが、今後は、現在の作業から文脈を切り出して提示する手法の実装を考えている。

(7) 履歴を図書館で利用されている分類法である日本十進分類法を用いて構造化する手法を適用し、利用者の知識構造や好みの推測とそれに基づいた人物の推薦が可能となるシステムを実装した。システムを実際に利用してもらい、履歴共有を試みた上で、有用な履歴を持つ利用者の推薦について評価を行った。その結果、従来手法に比べてより有用な利用者を推薦可能であることが示された。

以上により、履歴共有システムの運用の準備が整ったと考える。実運用による履歴取得は倫理的問題の解決に時間がかかり、来年度以降の実験に繰り越すことになったが、これらの研究を統合して実現可能と考えている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 1 件)

- 1) 重田桂誓, 松村敦, 宇陀則彦. 表紙生成エンジンを用いた二次元配置型 Web キュレーションシステムの開発. 知能と情報(日本知能情報ファジィ学会誌), 査読有, 2013, vol. 25, no. 1, p. 609-623. DOI: <http://dx.doi.org/10.3156/jssoft.25.609>

[学会発表] (計 9 件)

- 1) 米島まどか, 松村敦, 宇陀則彦. 過去に蓄積した情報を呼び起こすシステムに関する一検討. ARG 第 1 回 Web インテリジェンスとインタラクション研究会, 2012-12-15, 神奈川近代文学館.
- 2) 益子博貴, 松村敦, 宇陀則彦. キャラクターを活用した学習継続支援ツールの開発. ARG 第 1 回 Web インテリジェンスとインタラクション研究会, 2012-12-14, 神奈川近代文学館.

- 3) 若松望, 常川真央, 松村敦, 宇陀則彦. ユーザの嗜好を考慮したブックレビューランキング手法. 第 4 回データ工学と情報マネジメントに関するフォーラム (DEIM2012), 2012-03-05, シーサイドホテル舞子ビラ神戸.
- 4) 米島まどか, 松村敦, 宇陀則彦. 探索と振り返りの融合を目指した履歴可視化インタフェース. 第 4 回データ工学と情報マネジメントに関するフォーラム (DEIM2012), 2012-03-05, シーサイドホテル舞子ビラ神戸.
- 5) Asuka OTA, Reina HIROSE, Atsushi MATSUMURA and Norihiko UDA. A Bibliographic Search System with a Focus on the User's Knowledge Structure. Proceedings of Osaka Symposium on Digital Humanities, 2011-09-12~14, Osaka University.
- 6) Mao Tsunekawa, Haruki Ono, Kyoji Konishi, Keita Tsuji, Atsushi Matsumura and Norihiko Uda. Shizuku2.0: Cooperative Reading Support System. Proceedings of Asia-Pacific Conference on Library & Information Education and Practice: Issues, Challenges and Opportunities, p. 539-547, 2011-06-23, Pullman Putrajaya Lakeside, Putrajaya, Malaysia.
- 7) 重田桂誓, 松村敦, 宇陀則彦. Web ページを視覚的に要約する「表紙」生成エンジンの開発. 電子情報通信学会, 2011, WI2-2011-01-24, p. 89-90, 2011-03-08, 学術総合センター.
- 8) 山口恭平, 松村敦, 宇陀則彦. 意見情報の時系列を考慮した議論可視化システム. 電子情報通信学会, 2011, WI2-2011-01-24, p. 33-34, 2011-03-07, 学術総合センター.
- 9) 常川真央, 小野永貴, 松村敦, 宇陀則彦. 学習ノウハウの共有を支援するコミュニティ指向型図書館システム. 2010 年度人工知能学会全国大会 (第 24 回), 2B2-1, 2010-06-10, 長崎ブリックホール.

6. 研究組織

(1) 研究代表者

松村 敦 (MATSUMURA ATSUSHI)
筑波大学・図書館情報メディア系・助教
研究者番号: 40334073

(2) 研究分担者

宇陀 則彦 (UDA NORIHIKO)
筑波大学・図書館情報メディア系・准教授
研究者番号: 50261813