

令和元年6月24日現在

機関番号：32689

研究種目：若手研究(B)

研究期間：2017～2018

課題番号：17K12801

研究課題名（和文）ページ中の可变的要所抽出によるWeb探索学習における模索経験支援

研究課題名（英文）Extraction of Changeful Important Portions in Pages for Enhancing the Meaningful Grouping in Web Search Learning

研究代表者

中山 祐貴（Nakayama, Hiroki）

早稲田大学・グローバルエデュケーションセンター・講師（任期付）

研究者番号：80761569

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：研究活動におけるサーベイや課題解決型学習（PBL）において模索的Web探索の意義と能力育成の重要性が増している。この種のWeb探索の円滑な遂行には、探索目的に応じた適切なクエリ捻出、有用ページ獲得の為に探索結果の吟味、有用ページ中の重要箇所の整理・考察が重要であるが、Web探索に不慣れた学習者には困難な作業である。本研究では、探索すべき事柄の学習者自身による表現の切っ掛けとなるクエリ推薦手法、有用ページ推定手法、探索状況・探索成果物の整理支援手法を開発した。これらの手法に基づく支援システムを実装することで、Web探索の一連のフェーズを見渡した模索経験促進支援の新たな可能性を示した。

研究成果の学術的意義や社会的意義

本研究では、模索的Web探索の一連のフェーズにおける「余計なハードル」を軽減し、学習者の模索経験を促す方法を開発した。これにより、従来手法では対応が困難であった、探索目的に応じたクエリ捻出から、探索成果物の整理・考察に至る模索的Web探索の主要フェーズを見渡した模索経験への傾注困難性に対する新たな支援効果が期待される。

研究成果の概要（英文）：The significance of Web exploration with trial and error has increased in intellectual activities such as group work in Project-Based-Learning (PBL) and research work. To successfully carry out such exploration, it is needed for users but difficult to create suitable queries according to the search purpose, examine the search results for obtaining useful pages and organize important parts in obtained pages. Especially, novices tend to become immersed only in a part of Web exploration such as query creation. In this research, methods for generating the clues to alleviate the extra hurdles in each phase of exploration were developed. Moreover, a support system that implements these methods were developed. This system suggested a novel way for promoting the trial and error experience of novice under considering whole phases of Web exploration.

研究分野：教育学

キーワード：模索的Web探索 模索経験支援 可变的要所抽出 情報視覚化 教育学

1. 研究開始当初の背景

研究活動のサーベイや教育機関での課題解決型学習 (PBL) 等において、模索的 Web 探索を行う機会が増加しており、その能力育成が重要性を増している。この種の Web 探索を上手く実施するためには、「探索目的に応じた適切なクエリ捻出」、「有用 Web ページを得るための探索結果の吟味」、「探索成果物中の重要箇所の整理に基づく考察」が重要である。しかしながら、探索課題に関する知識が薄く、Web 探索に不慣れた学習者は、クエリ捻出や Web ページの閲覧など、1つの行為に没頭し過ぎる傾向が強く、各フェーズを往來しながら実際の模索を経験することが困難である。

これに対して、検索キーワード推薦、有用ページの共有・視覚化など、Web 探索支援に関する研究が数多く報告されている。しかし、殆どが、個別フェーズの支援に止まり、「適切なクエリ捻出から、探索結果の吟味、探索成果物中の重要箇所の整理・考察まで、一連の模索経験」に対する有効な支援は実現されてこなかった。

そのため、Web 探索に不慣れた学習者が模索経験を獲得するための、一連フェーズを見渡した支援が重要であると考えられる。

2. 研究の目的

本研究では、Web 探索の主要フェーズを見渡した模索経験促進支援の実現を目指す。具体的には、上記フェーズに介在する履歴データに着目し、それらの分析に基づいて、各フェーズにおける「手がかり」を抽出・視覚化する手法を開発する。これにより、Web 探索における模索経験を獲得させるための実際的な支援方法の可能性を示す。

3. 研究の方法

本研究では、検索クエリ、ブックマーク、閲覧した Web ページ、Web ページの特徴語、重要語句など Web 探索に介在する履歴データを一体的に管理することが可能な環境を開発する。その上で、Web ページの取舍選択状況の分析に基づいて、探索目的に応じた検索クエリ候補を推薦する手法を開発する。また、ユーザが獲得したページと検索結果ページ間の類似度に着目した分析に基づいて、探索者にとって未知の情報を含むページを推定する手法を開発する。さらに、探索成果物中の重要箇所 (成果物間の関係性や探索不足箇所など) を抽出し、視覚的に示唆する手法を開発する。これらの手法を実装した支援システムを開発することで、模索経験の促進を試みる。

4. 研究成果

4.1 模索的 Web 探索の主要フェーズを見渡した模索経験促進手法

4.1.1 模索経験促進戦略

本研究では、Web 探索の主要フェーズにおいて、次の方針に基づいて Web 探索の手がかりを生成・示唆する。

- ・探索上の行き詰まりが生じている場合には、それを解消する。
- ・行き詰まりが生じていない場合には、出来るだけ余計な影響を与えない。
- ・探索者の意義のある試行錯誤を増加させる。

これにより、学習者が1つの作業だけに留まる状況を軽減し、複数フェーズの往來による意義のある試行錯誤を経験させることを狙う (図1)。



図1 模索経験促進支援の概要

4.1.2 検索クエリ候補推薦手法

クエリの捻出フェーズでは、探索者は、自身の探索目的に応じて検索クエリを捻出し、それを用いて有用ページを得ようとする。このとき、本研究では、探索者の行動履歴を分析し、次のような状況が認められた場合には、クエリ捻出に行き詰まりが生じていると判断する。

- ・一定時間以上、有用ページを獲得できていない状態が続いている
- ・閲覧ページ数に照らして、検索クエリの変更回数が多い
- ・各 Web ページの閲覧時間が短い

行き詰まりが生じていると判断した場合には、探索者のこれまでの探索経緯の分析を基に、次の2つの条件を満たすクエリ候補を探索者に手がかりとして推薦する。

条件1．似て非なるページを獲得できるクエリ

条件2．探索者が気づいていないクエリ

上記の条件を満たすクエリを抽出するために、「獲得済みの有用 Web ページ」と「閲覧したものの有用ではないと判断したページ」の差異に着目する。それぞれのページ中のテキストに対して、形態素解析ツールとテキストマイニングツールを使用して、特徴語の抽出・重み付けを行う。このとき、有用ページの中で、探索者が重要箇所としてマーキングした部分は、既知の情報と見做して対象から除外する。また、抽出した特徴語の重みが閾値よりも低い場合には、その特徴語はノイズと見做して除外する。

次に、有用ページと閲覧のみのページから抽出した特徴語を比較する。閲覧のみのページの特徴語は、探索者が有用と判断しなかった事実を照らして、検索クエリに適さない場合が多いと考え、有用ページのみ出現する特徴語を選出する。さらに、有用ページのみ出現する特徴語のうち、既に検索クエリとして使用した特徴語は、クエリ候補から除外する。その上で、最も重みが高い特徴語をクエリ候補として探索者に推薦する。

4.1.3 有用ページ推定手法

ページの吟味フェーズでは、探索者は、検索結果の一覧から Web ページを閲覧することで、探索目的に照らして有用な情報を含むページ（有用ページ）を探す。このとき本研究では、探索者の行動履歴を分析し、次のような状況が認められた場合には、有用ページの獲得に行き詰まりが生じていると判断する。

- ・一定時間以上、有用ページを獲得できていない状態が続いている
- ・閲覧ページ数に照らして、検索クエリの変更回数が少ない
- ・各 Web ページの閲覧時間が長い

行き詰まりが生じていると判断した場合には、探索者にとって有用と思われるページを推定する。具体的には、探索目的に関する情報と、探索者にとって未知の情報の双方を含んだページ（似て非なるページ）を推定する。そのために、探索成果物と検索エンジンが出力した検索結果ページについて、2段階の類似性分析を行う。まず、探索成果物と検索結果の各ページ間の類似度を算出し、探索成果物に対する類似度が閾値よりも高いページを選出する。次に、選出したページと探索者が重要箇所としてページ中にマーキングした箇所間の類似度を算出する。この類似度が閾値よりも低いページを、有用ページ候補（似て非なるページ）として探索者に提示する。

4.1.4 探索状況・探索成果物の整理支援手法

探索成果物の整理フェーズでは、探索者は、これまでの探索経緯・結果を整理し、今後の探索方針の検討を行う。本研究では、類似する探索成果物とそれらの間の関係、および、探索不足箇所を推定・提示することで、探索状況の把握を支援する。

まず、類似する成果物とその関係を提示するために、探索者が獲得した有用ページ（成果物）間の類似度を、本文全体を分析対象として算出し、類似度の高い成果物を関連付ける。また、成果物中の重要箇所の特徴語を抽出し、成果物を体現するメタデータのタグとする。これにより、成果物の整理を行う際に、探索者の着目に応じて、ある成果物と類似する成果物群を提示し、タグを基に内容の共通性とその要素を示唆する。

次に、探索不足箇所を推定・提示するために、成果物中の重要箇所の特徴語を共起性に基づいてマッピングする。共起性の高い特徴語同士は、同一の話題を示唆する傾向が高く、それらの特徴語をマップ上で近い距離に配置することで、探索状況の把握に役立てる。まず、形態素解析ツールを用いて成果物の重要箇所から特徴語を抽出し、共起行列を作成する。作成した共起行列に基づいて入力ベクトルを作成し、自己組織化マップを用いて2次元平面上に特徴語をマッピングする。さらに、マップに配置した特徴語間の距離が閾値以下であれば、同一の話題としてクラスタリングする。その上で、クラスタ内の特徴語を多く含む重要箇所を選出し、その成果物の獲得に使用した検索クエリを探索者の探索意図としてクラスタにラベリングする。これにより、探索者はマップ上の各クラスタにどのような情報が含まれているか容易に判断できる。同様に、マッピングした特徴語の密度が低いクラスタを探索不足箇所として探索者に示唆する。

4.2 模索経験促進支援システムの開発

開発した手法に基づいて、Web 探索の一連のフェーズに介在する履歴データから探索の手がかりを抽出し、探索者に視覚的に示唆する支援システムを開発した。本システムは、一般的な Web 検索スタイルである検索エンジンを用いたキーワード検索、および、有用 Web ページのブックマークに着目し、これらの行為に介在する履歴データを自動的に蓄積することが可能な環境を提供する。本システムのプロトタイプは、フレームワークと主要モジュールは Visual Studio 2017、C#、Python を用いて実装した。また、データベースは、MySQL を用いて構築した。な

お、本プロトタイプは、模索経験促進支援のためのテストベットとしての意味を併せ持つ。実装したプロトタイプのインタフェースを図2に示す。

これにより、探索者に特別な負担を強いることなく、探索における一連フェーズにおける余計なハードルを軽減し、意義のある試行錯誤への傾注を支援する。



図2 プロトタイプのインタフェース

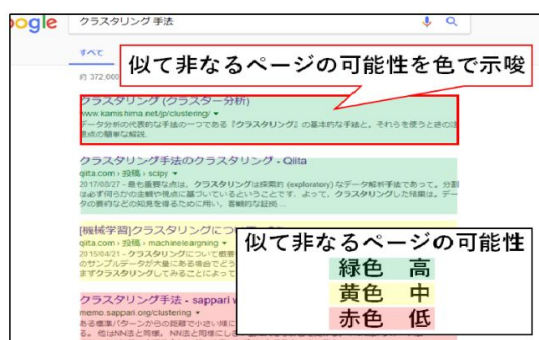


図3 有用ページ示唆機能による検索結果ページの有用可能性の示唆例

実装したプロトタイプの機能群のうち、Web探索における模索経験を促進させるための主要な役割を担う「検索クエリ示唆機能」、「有用ページ示唆機能」、「探索成果物間関係示唆機能」、「探索不足箇所示唆機能」について述べる。

(1) 検索クエリ示唆機能

本機能は、検索クエリ捻出フェーズにおける探索者の行き詰まりを診断する。行き詰まりを検出した場合、その時点までの探索者のページ取捨選択状況の分析に基づいて、検索クエリ候補を生成する。クエリ候補はポップアップ形式で提示され、探索者は簡明な操作でそれらを続く探索に活用することができる。

(2) 有用ページ示唆機能

本機能は、有用ページ推定手法に基づいて、Webページ吟味フェーズにおける探索者の行き詰まりを検出する。行き詰まりを検出した際には、探索者が検索結果を閲覧しているときに、検索結果のページタイトル・スニペットの色分けにより、有用ページである可能性を示唆する(図3)。具体的には、以下の3段階で色分けを施している。

- (緑色) 探索目的に即しており、探索者にとって未知の情報を含むページ
- (黄色) 探索目的に即しているが、探索者にとって既知の情報にあたるページ
- (赤色) 探索目的に即していないページ

(3) 探索成果物間関係示唆機能

本機能は、探索状況・探索成果物の整理支援手法に基づいて探索成果物間の関係性を示唆する。探索者が支援システム上で着目するタグや成果物を選択すると、関連する成果物と介在情報を視覚的に提示する。さらに、探索者が選択した探索成果物と類似度の高い探索成果物を介在情報と共に強調表示する(図4)。

(4) 探索不足箇所示唆機能

本機能は、2次元平面上にマッピングした探索成果物の特徴語群を探索者に視覚的に提示し、支援システムが推定した探索不足箇所を、マップ上の該当する特徴語クラスタの強調表示により示唆する。また、探索者がマップ上の特徴語を選択することにより、選択した特徴語と関連のある探索成果物とその介在情報を共に表示する(図5)。

このようにして、Web探索の主要フェーズにおける作業のハードルを軽減することで、より意義のある模索経験を後押しする。

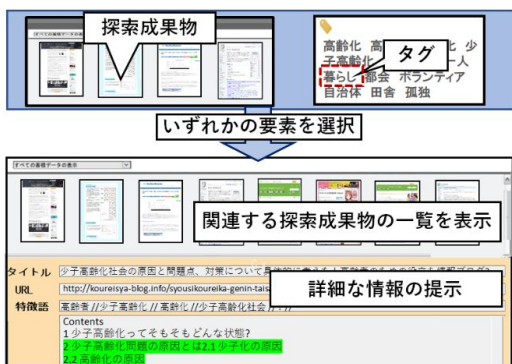


図4 探索成果物間関係示唆機能による成果物と介在情報の示唆例

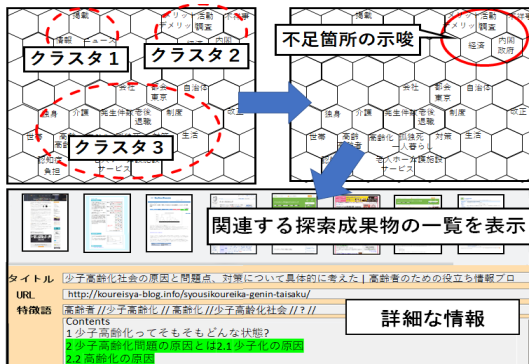


図5 探索不足箇所示唆機能による不足箇所の示唆と関連成果物の示唆例

4.3 実験と考察

4.3.1 実験概要

Web 探索の主要フェーズにおける探索者の模索経験促進に関する有効性の検証と課題抽出を目的として実験を行った。

まず、実験協力者（情報系大学生4名）に「少子高齢化」、「ごみ問題」という2つのテーマに関する Web 探索と、探索結果に基づいたレポート作成を実施して頂いた。作成するレポートは、それぞれのテーマについて「主な問題点（2つ以上）」、「問題点に対する対策・解決策についての考察」の2つを含むよう指示を与えた。また、レポートは、それぞれ A4 サイズで1ページを目安に作成して頂いた。Web 探索は、異なるクエリを使った検索を10回以上実施し、探索成果物（Web ページ中にマーキングしたページ）が20以上になるまで行って頂いた。これらの作業を支援システムによる示唆の有無が異なる環境で各1回（協力者1人あたり2回）実施して頂いた。

上記の作業の終了後に、支援システムが実際に実験協力者に提示した示唆情報の効果等を検証した。具体的には、探索状況を1つ1つ丁寧に観察し、示唆の提供状況と作業履歴を照らし合わせた確認を行った。あわせて、質問紙調査とインタビューを実施した。

4.3.2 結果と考察

Web 探索を経たレポート作成作業の概況を表1に示す。まず、検索クエリの示唆は、4名の協力者のうち協力者Dに対してのみ行われていた。インタビューを含めて実際状況を精査したところ、協力者Dは、テーマの解釈に苦慮しつつも、示唆を参考にしながら、クエリ捻出とページ閲覧で模索している状況が確認できた。このことから、検索クエリ示唆が意義のある試行錯誤を促していることが伺える。それ以外の3名について実際の Web 探索の状況を精査したところ、検索クエリの捻出にあまり時間を使わずに、他の作業に時間を費やしていた。使用した検索クエリ数と照らし合わせると、協力者Dは示唆有の場合にクエリ数の増加が認められ、上記の状況と符合している。また、クエリ捻出における行き詰まりが認められなかった協力者A~Cは、クエリ数だけを見ると必ずしも一様ではないことが分かる。これらのことから、クエリ数の増減だけでは把握できない行き詰まりを、提案手法により上手く推定できているものと考えられる。さらに、協力者Dは、支援システムが示唆したキーワードを基に検索クエリを捻出した場合に有用ページを獲得できていた。このことから、検索クエリ示唆が一定の効果を持っていることが伺える。

表1 Web 探索を経たレポート作成作業の概況

協力者	テーマ(上段:1回目, 下段:2回目)	作業環境	検索クエリ数	閲覧ページ数	ブックマーク数	作業時間(分)
A	ごみ問題	支援無し	12	35	20	91
	少子高齢化	支援有り	13	40	20	115
B	ごみ問題	支援無し	27	37	20	92
	少子高齢化	支援有り	25	29	23	97
C	少子高齢化	支援無し	34	41	20	72
	ごみ問題	支援有り	22	56	21	78
D	少子高齢化	支援無し	22	30	20	97
	ごみ問題	支援有り	33	37	20	112

有用ページ示唆は、全ての協力者に対して行われていた。実際状況を精査した結果、システムの示唆を受けることで、閲覧ページの「幅」が広がっている様子を確認できた。また、支援システムによるページ示唆が、有用ページ獲得に繋がっており、示唆支援の一定の効果が伺える。一方、探索者が支援システムの示唆を活用せずに作業を続け、有用ページを獲得できないケースも確認できた。この点について精査したところ、支援システムの示唆の基となる分析タイミングと、示唆を受けた時点での協力者の状況に差異が生じていた。例えば、協力者がある程度探索を完了した内容についてのページを示唆している等、探索目的の変化に対応できていない可能性が考えられる。これについては、タイミングに応じた示唆の絞り込みの検討が必要である。それぞれの協力者の詳細な履歴と照らし合わせると、協力者3名は、支援有りの場合に閲覧ページ数の増加が確認できる。支援システムからの示唆により、積極的にページを探索したことが推察できる。一方で、閲覧ページ数の増加が見られない協力者Bは、ブックマーク20件を集める前にレポートを殆ど作成し終えており、作業の趣旨からやや外れた状況が伺える。

探索成果物の整理フェーズにおける示唆は、実際状況を精査した結果、システムの示唆により、既知情報の探索が減少している様子を確認できた。支援システムからの示唆により、探索者が探索状況を把握した後に、探索成果物の整理フェーズ以前の作業に立ち返った更なる探索を実施することに繋がったことが伺える。示唆の活用状況を確認すると、レポート作成の際に本示唆機能を利用していた。概ね狙い通りの活用と言える。さらに詳細な状況を精査すると、協力者Aは、ある程度ブックマークを集めた後に示唆を活用してレポートを作成している様子が確認できた。一方、協力者B~Dは、レポート作成時に示唆を軽く確認する程度に留まって

いた。これは、探索と並行してレポート作成が行われていた状況から、半ば自然な流れと解釈できる。今回の実験では、作業期間・探索量が比較的小さいため、やむを得ない部分と言える。今後、より実証的な検証を重ねる必要がある。

これらの結果から、全ての協力者について、いずれかのフェーズで、行き詰まり軽減と模索経験促進を認めることができる。加えて、行き詰まりが軽減されつつも作業時間が単純減少に転じていないことを見て取ることができる。このことから、いずれかのフェーズでの行き詰まりが解消されたことで、他の作業に傾注している様子が伺える。余計なハードルを軽減し、より意義のある試行錯誤を行わせる本研究の趣旨に照らして、良好な結果と言える。

最後に、質問紙調査の結果から、支援システムによる示唆のうち、検索クエリ示唆、有用ページ示唆、成果物間の関係性示唆については、良好な評価が得られていることが確認できた。インタビューの結果からも、「示唆を受けることで、探索経緯を整理するきっかけができた」等のコメントが得られた。このことから、模索経験促進という本研究の趣旨と照らして、良好な感触と言える。一方、探索不足箇所示唆は、低めの評価となっており、インタビューでも「集めた情報はまとまりで確認できたが、新たに必要な情報は特に把握できなかった」などのコメントが得られた。今回の実験の制約上やむを得ない部分もあるが、探索の不足と探索すべきエリアの示唆方法などを更に検討することが重要と考えられる。このことから、探索状況の視覚化方法について検討の余地が伺える。

今回の実験は限定的なものであり、直ちに提案手法の有効性を確認するまでには至らないものの、今後に期待を持てる感触が得られた。今回得られた知見を基に、支援システムの改善を進め、更なる検証・知見集約を重ねることが重要と考える。

表2 各協力者に対する示唆の概況

協力者	検索クエリ示唆	有用ページ示唆	関係性示唆	探索不足箇所示唆
A	0	3	3	1
B	0	2	1	1
C	0	3	1	2
D	2	4	2	2

5. 主な発表論文等

〔雑誌論文〕(計3件)

渡辺剛硫, 中山祐貴, 大沼亮, 神長裕明, 喜久川功, 宮寺庸造, 中村勝一, “未熟者の模索経験促進のための Web 探索過程における手がかり示唆システムの開発,” 電子情報通信学会 技術研究報告, 査読無, Vol. 118, No. 510, pp. 47-52, 2019.

Hiroki Nakayama, Takeru Watanabe, Ryo Onuma, Hiroaki Kaminaga, Isao Kikukawa, Youzou Miyadera, Shoichi Nakamura, “Methods for Facilitating the Experiences of Meaningful Trial and Error in Complicated Web Exploration,” Proc. 2018 IEEE Conference on e-Learning, e-Management and e-Services, IEEE, 査読有, pp. 67-72, 2018.

渡辺剛硫, 中山祐貴, 大沼亮, 神長裕明, 喜久川功, 宮寺庸造, 中村勝一, “Web 探索未熟者の模索経験促進のための手がかり示唆,” 電子情報通信学会 技術研究報告, 査読無, Vol. 118, No. 131, pp. 13-18, 2018.

〔学会発表〕(計3件)

Yota Kawawa, Hiroki Nakayama, Ryo Onuma, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura, “Extracting the Springboard for Reflecting User's Recognition on the Obtained Pages to Future Search,” 17th International Conference on WWW/Internet, 2018年10月21日-23日(ブダペスト・ハンガリー).

渡辺剛硫, 中山祐貴, 大沼亮, 神長裕明, 宮寺庸造, 中村勝一, “未熟者の模索経験促進のための Web 探索における主要フェーズを見渡した手がかり示唆,” 電子情報通信学会 2018年総合大会, 2018年3月20日-23日, 東京電機大学(東京・日本).

Takeru Watanabe, Hiroki Nakayama, Ryo Onuma, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura, “Visualization of the Clues in Web Exploration for Enhancing Unskilled Learner's Meaningful Trial and Error,” 2017 International Conference on Computational Science and Computational Intelligence, 2017年12月14日-16日,(ラスベガス・アメリカ合衆国).

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。