

令和 3 年 6 月 29 日現在

機関番号：11601

研究種目：若手研究

研究期間：2019～2020

課題番号：19K14309

研究課題名(和文)探索遂行能力育成のための既有ページ注目行為に着目した段階的の手がかり示唆手法

研究課題名(英文)Methods for Strategically Presenting the Clues of Fostering Search Skills based on Analysis on Novices' Behaviors of Page Attention

研究代表者

大沼 亮 (Onuma, Ryo)

福島大学・共生システム理工学類・特任助教

研究者番号：60829729

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：PBLや研究活動等において試行錯誤性の高いWeb探索を行う機会が増加し、その能力育成の重要性が増している。この種の探索を上手に行うためには、探索結果と経緯を理解した上で、次の探索に繋げる必要があるが、未熟者には容易ではない。能力習得のためには模索経験の積み重ねが重要だが、複数の要素が関係する難しさゆえに、十分に行えない実状がある。本研究では、未熟者の探索進捗状況に対する認識や探索意図が、既有ページに対する注目行為に半ば無意識的に体现される現象に着目し、これを活かした段階的な手がかり示唆による模索経験促進方法を提案した。これにより、実際の経験を促しつつ能力育成を図る新たな支援の可能性を示した。

研究成果の学術的意義や社会的意義

重要性が指摘されながらも、これまで困難であった模索経験の積み重ねによるWeb探索能力の育成を可能とする教育支援充実の意義を有する。

また、これまで見過ごされてきた「半ば無意識的な既有ページに対する注目行為」を活かして、未熟者の認識や探索意図を反映し得る形で模索経験促進に挑戦する本研究は、教育工学研究の新たな可能性を提示する意義を有する。

研究成果の概要(英文)：Opportunities to perform Web explorations involving trial and error as a part of intellectual activities such as PBL have increased and cultivation of its skills is increasingly needed. In order to successfully conduct such exploration, it is necessary to understand the results of searches along with their circumstances and utilize them to later explorations. Although the stacking of exploration experience is important to acquire the skills, it is difficult for unskilled people to accumulate sufficient experiences since multiple factors such as expression of queries, examination of acquired pages and change of search intentions are involved. This research focus on a fact that novices' behaviors of page selection during looking back their exploration often express unconsciously their search intentions and recognition against search progress. Utilizing this fact, methods for promoting the experience of exploration with trial and error by suggesting the clues in steps were proposed.

研究分野：教育工学

キーワード：探索能力育成 ページ注目行為 探索意図 手がかり示唆 情報視覚化

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

PBL や研究活動における調査など、試行錯誤性の高い Web 探索を行う機会が増加し、その能力育成の重要性が増している。この種の探索を上手く行うためには、探索結果と経緯を理解した上で、次の探索に繋げる必要があるが、未熟者には容易ではない。能力習得のためには模索経験の積み重ねが重要だが、クエリ表現や獲得ページの吟味、自身の探索意図の変化など、複数の要素が関係する難しさゆえに、十分に行えない実状がある。

これに対して、Web ページの推薦や分類管理など、Web 探索支援に関する研究が多数報告されているが、殆どが探索効率化に主眼を置くもので、探索能力育成としては有効とは言えない。一部、能力育成に焦点を当てた研究も存在するが、模索的探索の成否を左右する要素の内、いずれかのみを対象としており、「一連の作業について実際の経験を集める」支援は実現されていない。ゆえに、これら介在要素を考慮した形で、模索的探索の実際の経験を促す能力育成方法が必要と言える。

2. 研究の目的

本研究では、未熟者の探索進捗状況に対する認識や探索意図が、既有ページに対する注目行為に半ば無意識的に体现される現象に着目し、これを活かした段階的な手がかり示唆による模索経験促進方法の実現を目指す。

3. 研究の方法

まず、探索履歴・動向の分析に基づいて、未熟者が探索成果・意図の整理のために注目すべきページの候補を抽出する手法を開発する。次に、探索者の着目に応じて手がかり(検索クエリ候補、有望な未知ページ)を抽出する手法を開発する。さらに、これらの手がかりを探索者に示唆し、経験状況に応じて段階的に減らしていく能力育成戦略を開発する。その上で、探索者の状況に応じて手がかりを示唆するシステムを開発する。これにより、探索における余計なハードルを軽減し、より重要な模索経験を促進することを目指す。

4. 研究成果

4.1 模索的探索における手がかり生成手法

4.1.1 試行錯誤を伴う探索におけるスプリングボード候補の抽出

本研究では、獲得ページの中で、探索者が注目するものを「探索者の現状認識・要求を体现し、更なる探索の踏み台になる存在(スプリングボード)」と捉え、それを活用した支援方法の開発を試みる。本研究では、典型的なスプリングボードとして以下の3つに注目する。

- ・探索者の求める情報を概ねカバーしているページ
- ・探索者の求める情報の一部を詳述しているページ
- ・複数の話題を含んでいるページ

スプリングボードに相当するページを見つけること自体が、探索者にとって容易な作業ではない。そこで本研究では、探索成果物の管理等を容易に行える環境を提供し、そこでユーザが生成する既有ページのクラスタの中からスプリングボードになり得るページを抽出する。

「求める情報を概ねカバーしているページ」を抽出するために、まず、各クラスタの主要な特徴語を抽出する。次に、クラスタ中の各ページの特徴語とその重要度を抽出する。その上で、クラスタを構成するページの中で、クラスタを代表する主要特徴語を網羅する度合いの高いものを選出する。このようなページは、所属するクラスタの内容全てをカバーするとは限らないものの、探索者にとって、そのクラスタを概ね表す存在となり得るものと考えられる。

「求める情報の一部を詳述しているページ」を抽出するために、各クラスタの内容を表す主要特徴語、および、クラスタを構成する個々のページの特徴語をそれぞれ抽出する。その上で、1つのクラスタを構成するページについて、テキスト量(文字数)、特徴語の有する重要度を分析し、「テキスト量が一定以上であり、より重要度の高い特徴語を含む度合いの高いページ」を選出する。このようなページは、ある情報について詳細を述べている可能性が高いものと考えられる。つまり、「求める情報を概ねカバーしているページ」で重視する網羅性は求めずに、独自性と一定程度の情報量が伺えるページをスプリングボード候補とする。このタイプのスプリングボード選出に際しては、探索者にとっての認識し易さを意識し、各ページの獲得時に使用された検索クエリに、ある程度重みをおく。

「複数の話題を含むページ」を抽出するために、クラスタを構成する個々のページの特徴語をそれぞれ抽出し、一定以上の重要度を有する特徴語を含むページのみを抽出する。次に、抽出したページ(クラスタを構成するページの一部)を、内容の類似性に基づいて分類する。つまり、仮のサブクラスタを抽出する。その上で、抽出した各サブクラスタを代表する特徴語を改めて診断し、他のクラスタの特徴語と一定以上の差を有するものを抽出する。このようなサブクラスタが存在する場合、当該親クラスタは、複数の話題を含んでいると考えられる。このとき、最初に抽出した「一定以上の特徴語を含むページ」の中で、各サブクラスタを代表する特徴語を含む度

合が高いものをスプリングボード候補とする。この際、獲得時に使用されたクエリを含む度合が高いものを優先する。これは、スプリングボードを選択する際の探索者にとっての分かり易さに配慮する狙いである。

なお、これらのスプリングボードは、全てのクラスタについて常に存在するとは限らず、該当するもの（種別）が存在する場合に提示する。

4.1.2 探索者が注目するページに応じた手がかり抽出

(1) クエリ候補の生成

探索者による新たな情報獲得に対する手助けとして、探索者がスプリングボードとして選択したページに応じて検索クエリ候補を生成する。

「求める情報を概ねカバーしているページ」に注目した場合、探索者はそのページの関連情報・詳細情報を得たいという意図があると考えられる。そのようなページの獲得に向けて、注目したページの主要な特徴語と、獲得時に探索者が使用した検索クエリから、新たなクエリ候補を複数パターン生成する。

「求める情報の一部を詳述しているページ」に注目した場合、探索者は「注目したページに関連しつつ、異なる情報」を得たいという意図があると考えられる。ゆえに、似て非なるページの抽出に向けて、獲得時に使用した検索クエリは用いず、注目したページの主要な特徴語から新たなクエリ候補を複数パターン生成する。

「複数の話題を含むページ」に注目した場合、個々の話題について更に関連情報や詳細情報を得たいという意図があると考えられる。ゆえに、当該スプリングボード抽出時の仮のサブクラスタごとに、「求める情報を概ねカバーしているページ」に注目した場合と同様に検索クエリ候補を生成する。

(2) 有望な未知ページの提示

探索未熟者にとっては、クエリの選択自体が難しさを伴う。本研究では、生成したクエリ候補を使用して、注目したページに応じた有望な未知ページを抽出し、探索者に提示する。これにより、探索意図の自覚度合が低い段階でも、それを反映した模索を現実的に行わせ、次第に意図を自覚させることを狙う。

注目したページに応じた有望な未知ページを抽出するために、まず、生成したクエリを用いて検索要求を準備し、検索エンジンから結果（ページ集合）を取得する。得られたページ集合について、注目したページとの特徴語の類似度、スニペットの類似度などを分析し、有望なページを抽出する。

なお、似て非なるページの抽出に際しては、生成した検索クエリを用いて得られたページ集合の内、探索者が注目したページの主要な特徴語を含み、なおかつ、「一定以上の重要度の特徴語で、注目したページには含まれないもの」を含むページを選出する。

4.1.3 経験状況に応じた段階的な手掛かり示唆

本研究では、試行錯誤を伴う探索を遂行する能力を育成するために、有望ページなどの手がかりを段階的に削減し、探索者による模索や意図の自覚を促す。

未熟者の経験状況として、「ページを広く収集することができていない状況」、「自分の探索意図を自覚できていない状況」、「獲得した情報を上手く整理できない状況」に着目する。これらの段階は必ずしも順番通りに進むとは限らず、前の段階に戻ることや段階を飛ばしてスキルアップすることも考えられる。これらを考慮した上で、未熟者の探索能力レベルを推定する。本研究では、検索の頻度、閲覧ページ数、ブックマーク数等から経験状況を推定する。

まず、検索頻度が低く、閲覧ページ、ブックマークが少ない場合、ページを広く収集することができていない状況と推定する。このとき、新たなページの獲得により探索を前進させることが模索経験の促進に繋がると考えられる。ゆえに、探索の手がかりとして、有望ページを提示する。

次に、探索頻度が高く、閲覧ページ数が増加し、ブックマーク数が少ない場合、自分の探索意図を自覚できないまま探索を行っている状況と推定する。このとき、何らかのページに注目させることで、不確かな意図を表出・確認させることが重要と考えられる。ゆえに、探索の手がかりとして、スプリングボードを提示する。注目するスプリングボードによって、示唆される有望未知ページが変化するため、その反復を通して探索意図の自覚を促す。

また、検索頻度が低下し、閲覧ページ数とブックマーク数が多く、既有ページを再閲覧する頻度が高い場合、ある程度情報を集めることができていないものの、獲得した情報を上手く整理できない状況と推定する。このとき、新たな探索よりも、ひとまず獲得済の情報の整理を促すことが重要と考えられる。ゆえに、ページ獲得時に使用したクエリの共通性、および、ページの内容類似性に基づいて獲得済ページをクラスタリングし、探索者に提示する。つまり、客観的な分類の例を示すことで、獲得してきた情報の概況理解を促す。

4.2 支援システム

提案する支援の実現のために必要な機能等の検討に基づいて、支援システムのプロトタイプを開発した。本システムの構成を図1に示す。支援システムには、履歴データ取得・蓄積、手がかりの示唆、ユーザの探索状況推定、手がかりの段階的削減等の各機能を具備する。これらの機能を実現するために、本システムは、探索状況推定部、閲覧候補ページ抽出部、UI部などから構成され、主に Python を用いて実装した。また、本システムで扱う各種情報を蓄積・管理するデータベースは、MySQL を用いて構築した。

支援システムのインタフェースを図2に示す。本システムは、ユーザによる Web 探索を監視し、検索クエリや閲覧したページ、ブックマーク等を収集・蓄積する。あわせて、探索成果物の管理・観察を容易に行い得る環境を提供する。具体的には、探索者自身による探索成果整理の下地として、これまでに獲得したページをクラスタリングし、提示する。

本システムは、探索者が任意のクラスタ中のページから注目するもの(スプリングボード)を選択することで、注目ページに応じて、手がかりとなる情報を生成・提示する。このように、既有情報の整理・観察のための環境を与えつつ、探索状況に対する認識、既有ページに対する評価を「クラスタ(探索意図中の話題に相当する)ごとに、注目するページを選択する」という簡易な行為によって現実的に行わせる。それに応じて提示される新たなページの吟味を含めた一連の行為を反復的に行うことで、探索者の現状認識を積極的に反映した探索の遂行を支援する。

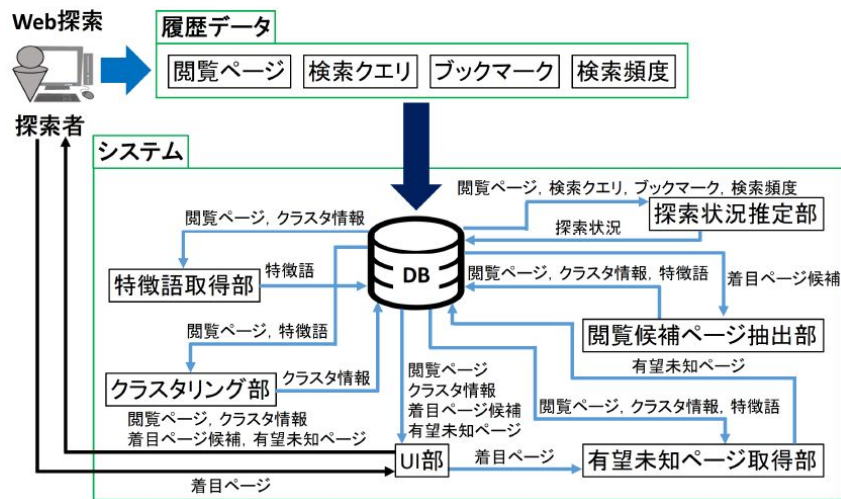


図1 システム構成

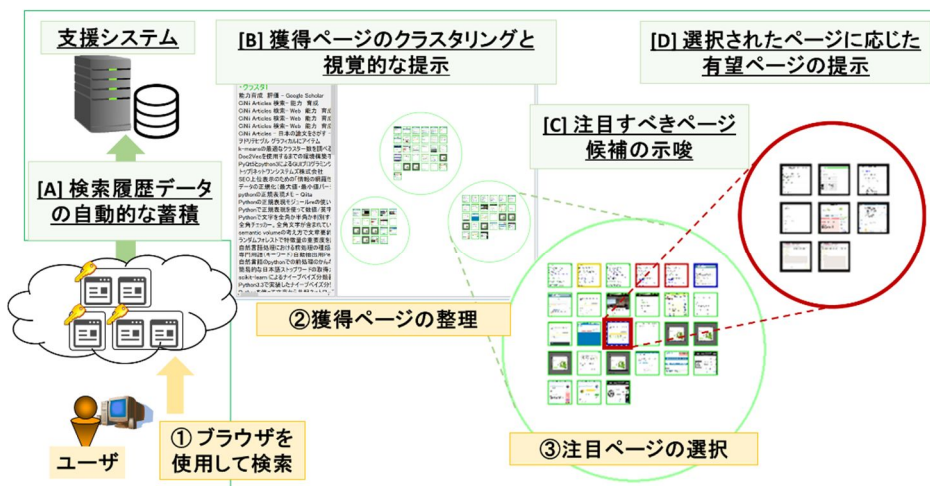


図2 支援システムのインタフェース

4.3 実験と考察

4.3.1 実験概要

スプリングボード抽出手法と有望未知ページ抽出手法の妥当性、および、手がかり示唆の有効性の検証を目的として実験を行った。

実験協力者(探索者側)に、2つのテーマについて Web 探索と、その結果を踏まえたレポート作成を実施していただいた。実験は探索フェーズ、探索成果整理フェーズ、再探索フェーズの3つの段階で行われた。まず、探索フェーズでは、与えられたテーマに沿って、「概要(調べた内容についての要約)」、「背景(現状に至るまでの経緯や原因、事例等)」、「影響(特に、自分に影

響がありそうなもの)」について、ブックマークが各 10 個以上になるまで探索していただいた。次に、探索成果整理フェーズでは、実験協力者に探索成果を整理させる目的で、獲得したページを分類していただいた。その後、探索テーマごとにレポートを作成していただいた。再探索フェーズでは、作成したレポートを実験協力者（教員側）が精査し、そこに記載された事柄の中からサブテーマごとに「この部分について、問題点の明確化、対応方法の把握を目指して更に深く調査せよ」という指示を与えた。この指示を踏まえて、実験協力者（探索者側）に、サブテーマごとに新たなブックマーク 10 件を目安として、探索を実施していただいた。その上で、探索の結果を、最初に提出したレポートに追記・修正する形でまとめていただいた。これらの作業を、システムによる支援の有無が異なる環境で各 1 回、それぞれ別の探索テーマで行った。

実験終了後に、実験協力者（探索者側）が獲得したページを 1 つ 1 つ精査し、スプリングボードに相当するものを選出した。これを正解データとして、支援システムによるスプリングボード抽出結果との一致状況を確認し、スプリングボード抽出手法の精度を求めた。また、実験協力者（探索者側）に支援システムが抽出した有望未知ページ候補を 1 つ 1 つ精査し、実際に有用ページと言えるか否かを確認していただいた。さらに、探索履歴データを丁寧に観察し、支援システムによる手がかり示唆の効果を検証した。

4.3.2 結果と考察

スプリングボード抽出の結果を表 1 に示す。一人の実験協力者については、良好な精度を得られているが、もう一方の協力者については、適合性が低めであった。本研究が目標とする経験促進のために「考えさせるきっかけ」の意味が大きいため、必ずしも高い適合性が必須ではないが、精度向上のために、更なる検討が必要と言える。ただし、いずれの協力者も適合性・再現性のバランスが取れた結果が得られており、その点は今後に期待が持てる。

有望未知ページ抽出の結果を表 2 に示す。いずれの協力者についても良好な精度が得られた。ただし、Web 探索の規模が大きい場合などについて、更に検証が必要と考えられる。

表 1 スプリングボード抽出手法の妥当性評価

	適合率	再現率	F 値
協力者 A	0.31	0.36	0.33
協力者 B	0.54	0.64	0.58

表 2 有望未知ページ抽出手法の妥当性評価

	適合率	支援システムが抽出した有望ページの総数
協力者 A	0.86	50
協力者 B	0.67	54

表 3 協力者 A の作業時間

	探索フェーズ	成果整理フェーズ	再探索フェーズ
支援有	0:19	1:27	2:16
支援無	0:17	0:54	0:58

表 4 協力者 B の作業時間

	探索フェーズ	成果整理フェーズ	再探索フェーズ
支援有	0:18	1:37	1:25
支援無	0:54	1:56	0:43

協力者 A、協力者 B が作業に要した時間をそれぞれ表 3、表 4 に示す。協力者によって差はあるものの、双方の協力者ともにシステムによる支援有の場合の方が、作業時間が長くなっていた。このことから、手がかり示唆、および、探索成果物の視覚的な管理が考察を促し、より深い探索に繋がったことが伺える。

また、ページ分類作業の際、協力者には各ページ群にその内容を表すタイトルを付けていただいた。それらをシステムによる支援の有無で比較したところ、支援無の場合は「概要」、「影響」といった抽象的なタイトルが多かったのに対し、支援有の場合は、テーマの内容により踏み込んだ具体的なタイトルが多い傾向が伺えた。このことから、探索成果物の視覚的な管理環境が、協力者の作業に影響を与え、より具体的なページ内容の把握に繋がったものと考えられる。

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、今後を期待する感触が得られた。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。

5. 主な発表論文等

〔雑誌論文〕 計5件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 Ryo Onuma, Hiroki Nakayama, Yota Kawawa, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura,	4. 巻 -
2. 論文標題 Initial Evaluation of System for Visualizing Novices' Unaware Search Intentions by Analyzing Behaviors of Article Selection	5. 発行年 2020年
3. 雑誌名 Proc. 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering	6. 最初と最後の頁 695-700
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TALE48869.2020.9368364	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 川和耀太, 大沼亮, 中山祐貴, 神長裕明, 宮寺庸造, 中村勝一	4. 巻 119(468)
2. 論文標題 Web探索能力育成のための未熟者の潜在的な意図を反映した手がかり示唆システム	5. 発行年 2020年
3. 雑誌名 電子情報通信学会 技術研究報告	6. 最初と最後の頁 103-108
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Ryo Onuma, Yota Kawawa, Hiroki Nakayama, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura	4. 巻 -
2. 論文標題 A Strategy for Cultivating Practical Exploration Skills by Suggesting the Clues based on the Analysis of Novices' Unaware Intentions	5. 発行年 2019年
3. 雑誌名 Proc. 18th International Conference on Information Technology Based Higher Education and Training	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ITHE46829.2019.8937357	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Yota Kawawa, Ryo Onuma, Hiroki Nakayama, Hiroaki Kaminaga, Youzou Miyadera, Shoichi Nakamura	4. 巻 -
2. 論文標題 A System for Cultivating Exploration Skills by Presenting Clues Based on the Analysis of Page Selection Behaviors	5. 発行年 2019年
3. 雑誌名 Proc. the 2019 2nd International Conference on Computational Intelligence and Intelligent Systems	6. 最初と最後の頁 134-139
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3372422.3372449	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 川和耀太, 大沼亮, 中山祐貴, 神長裕明, 宮寺庸造, 中村勝一	4. 巻 119(105)
2. 論文標題 記事選択行為に潜在する未熟者の認識・意図に応じた有望ページ抽出	5. 発行年 2019年
3. 雑誌名 電子情報通信学会 技術研究報告	6. 最初と最後の頁 35-40
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計1件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 大沼亮, 川和耀太, 中山祐貴, 神長裕明, 宮寺庸造, 中村勝一
2. 発表標題 ページ選択行動の分析に基づく探索遂行能力育成のための手がかり示唆
3. 学会等名 第44回教育システム情報学会全国大会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------