

網羅的検索のための Web 検索結果における拡張情報提示とその評価

山田 純平[†] 北山 大輔[†]

[†] 工学院大学 大学院工学研究科 情報学専攻 〒163-8677 東京都新宿区西新宿 1 丁目 2 4 - 2

E-mail: [†]tem19024@ns.kogakuin.ac.jp, ^{††}kitayama@cc.kogakuin.ac.jp

あらまし 検索エンジンではユーザに対して Web ページの選択を補助するためにタイトルやスニペットなどを提示している。しかしながら、旅行先の調査や、商品比較など、得たい知識を網羅的に習得するような検索においては、事前知識不足や、不適切なクエリ設定など検索戦略の失敗から Web ページを順々に閲覧してしまうことがある。この問題に対して、検索結果内の未訪問ページに対し、その内容を予想できるようなキーワードを提示し必要な情報が記載されていそうな Web ページを選択しやすくすることで解決する。我々は 2 種類の特徴語を検索結果に提示する拡張情報提示手法を提案する。1 つ目はその Web ページの内容を示す内容語、2 つ目は閲覧することで新たに得られそうな話題を表す既知話題語、未知話題語である。それぞれ、単語の分散表現を用いたスニペット文中の単語のクラスタリング、訪問済みページ中の単語のクラスタリングに基づいて抽出する。本稿では網羅的検索タスクにおける検索行動ログを収集し、提案する検索結果の拡張情報提示の効果を明らかにする。

キーワード 検索結果, Web 検索, 検索行動, 効率化, 可視化

1. はじめに

近年、インターネットや情報端末の普及に伴い、情報検索がより身近なものになった。中でも Web 検索は情報検索の代表例である。Web 検索では一つの検索クエリから多くの検索結果が提示される。その中でユーザは数多くの検索結果から必要な情報を効率よく閲覧したいと考えている。これまでにも、Web 検索をより便利にするための研究は数多くされてきた。しかしながら、未だ毎回効率良く Web 検索をすることは容易ではない。

検索エンジンでは入力されたクエリから生成される検索結果ページ (Search Engine Result Page, 以下 SERP とする) に提示された各ページを選択、閲覧することを繰り返すことで情報検索を実現している。既存の検索エンジンではユーザに対して Web ページの選択を補助するためにタイトルやスニペットなどを提示している。しかしながら、旅行先の調査や、商品比較など、得たい知識を網羅的に知りたい検索においては、事前知識不足など検索戦略の失敗から同じような内容のページを順々に閲覧してしまうことがある。

そこで我々はこの問題に対し、SERP 内の各スニペット文から 2 種類の特徴語を抽出し提示することで解決する。以前から我々は、SERP に提示される多くの検索結果から選びやすくなる検索支援手法として、特徴語を提示する拡張情報提示を検討している [1] [2]。この拡張情報を SERP に提示することで、ユーザは Web ページを開いて確認する前に内容を予想しやすくなり、自分の目的に合った Web ページの選択の助けになると考える。また、効率的に適切なページを閲覧することで、短時間に網羅的な探索が可能になると考える。本稿では、各検索結果の Web ページの内容を特に表す“内容語”と、ユーザがその Web ページを閲覧することで新たに得られそうな知識を表す“未知話題語”、“既知話題語”を提示する拡張情報提示について、網羅的検索における検索行動ログを収集し、その評価を行

う。被験者は提示されたタスクに従って網羅的検索を行い、その検索行動ログから拡張情報提示があった場合の効果として、どのようなユーザに対して有効か、どのような問題点があるのか、網羅的探索において有効かを分析する。

本研究の主な貢献は以下のとおりである。

- (1) ユーザの知識を可視化する拡張情報提示を新たに提案した。
- (2) 拡張情報提示の網羅的検索における効果として、実際のユーザの検索行動を用いて、検索時間が短くなることを示した。

本稿の構成は、以下の通りである。2. 節では、これまでの検索支援に関する研究と、本研究の位置づけについて述べる。3. 節では、本研究における拡張情報提示についての説明を述べる。4. 節では、実際の実験システムの概要と検索行動ログの収集について述べる。5. 節では、収集した検索行動ログとアンケート結果をもとに、拡張情報提示の評価について議論する。最後に 6. 節でまとめと今後の展望について述べる。

2. 関連研究

2.1 検索支援

これまで、Web 検索の支援手法について多くの研究がされてきた。検索エンジンの検索結果をリランキングに関する研究 [3] [4] [5] や検索クエリを推薦する研究 [6]、検索結果ページから閲覧ページへのスムーズな遷移を支援する研究 [7] などが挙げられる。

特に網羅的情報検索の支援としては、以下の研究が存在する。梅本らの ScentBar [8] は候補クエリの提示と未閲覧情報量を示すグラフを組み合わせることで網羅的検索を支援するアプローチを提案している。これは我々の目的と似ているが、我々は検索結果ページにスニペットを拡張することで検索結果の選択を支援することを目指している。早乙女ら [9] はウェブブラウザ

における既読コンテンツ検出とその表示手法について研究している。効率的な検索行動として、同じような内容の Web ページを見ないようにするために、既読コンテンツを検出することは有効である。しかしながら、検出した既読コンテンツを Web ページ中で提示していることから、我々が目指す検索結果の選択を支援することとは異なる。また、この論文では既読コンテンツをコサイン類似度によって検出しているが、我々は蓄積した訪問済みページ内でクラスタリングを行い、トピックを考慮している。

これらの手法と、我々の拡張情報提示は、支援を行う箇所が異なるため、併用することでそれぞれの効果を得ることが可能であると考えられる。

2.2 検索スニペット

次に、検索スニペットに関する研究について紹介する。Feildら[10]は検索結果スニペットと Web ページの遷移を支援する Clickable Snippet を提案している。スニペット文から気になる内容を見つけても、Web ページを開くとどこにその内容が書かれているか探してしまうことに対してページ遷移を支援している。拡張情報提示という観点では我々の研究と似ているが、非効率な検索になりがちな網羅的検索においては、ユーザ自身が知りたいことにフォーカスした動きになると考えられる。我々が提案する拡張情報提示では検索にユーザの意識とは別の区切りを与えることから、検索に夢中になり非効率な検索をしてしまう場合に有効ではないかと考える。

Ageev ら[11]は検索行動データに基づきスニペットのテキストの質を改善することに取り組んでいる。こうした研究と本研究は相性がいいと考える。より個人化されたスニペット文や Web ページを正確に表したスニペット文が提示されることで、スニペット文から含まれるトピックを単語として提示する我々の拡張情報提示の効果は向上すると考える。

一般に提示される検索結果やスニペット文自体を工夫する研究は存在しているが、検索結果に付加情報を提示し、Web ページ選択を支援する手法はあまりない。スニペットを拡張する手法は存在するが、スニペット文から拡張するような形で提示する工夫をしている研究は少ない。我々は、様々な検索支援がある中で検索エンジンを利用する以上、ユーザに対して表示される多くの検索結果から Web ページを選びやすくすることに着目し、その評価を行う。

2.3 検索行動ログの分析

検索支援の研究をするにあたって検索行動ログを正確により大規模に収集することが評価において重要である。また、検索行動ログを収集することによってユーザの検索行動から検索における問題点や、ユーザの属性から新しい支援アプローチを考えることができる。Kalyani ら[12]は 150 人を対象にユーザの知識獲得が認知レベルに応じてどのように変化するか分析した。結果として検索行動や知識獲得に影響が出ることが確認できた。我々は今回の実験を計画するにあたり、検索行動ログの収集、アンケートなどを考える過程で参考にした。

検索行動ログの分析は本研究において非常に重要である。検索行動の分析を行う先行研究は多数存在する[13][14][15][16][17]。

実験タスクの設計、検索行動の分類に関しては、これらの先行研究を参考にした。ユーザと検索タスクの満足度の関係性に着目し分析した研究なども存在する[18]。評価実験では、検索専門性と事前知識があるユーザには必要な情報だけをより判断しやすい付加情報を添えて提示し、逆に検索専門性と事前知識がないユーザには多くの検索結果を提示することが満足度向上につながると確認できた。本研究では検索クエリ 1 つにつき検索結果上位 10 件と提示される検索結果件数が固定されているが、事前知識と拡張情報提示の関連についても調査する。

3. 検索結果選択支援のための拡張情報提示

本研究で提案している拡張情報提示とは、SERP 上で提示される各検索結果に対して Web ページを選びやすくする特徴語を提示する。本手法ではスニペット文に含まれる特徴語から、各検索結果の Web ページの内容を示す“内容語”とユーザの訪問済みページを元に“未知話題語”と“既知話題語”を抽出する。本節では“内容語”の抽出手法、“未知話題語”と“既知話題語”の抽出手法について説明する。なお、すべての語は Word2Vec 等による単語の分散表現が得られているものとする。

3.1 特徴語抽出

はじめに、本稿で扱う特徴語について説明する。本稿ではスニペット文から特徴語を抽出する場合と、訪問済みページから特徴語を抽出する場合がある。スニペット文から特徴語を抽出する場合にはスニペット文とそのタイトルを用い、訪問済みページから特徴語を抽出する場合にはページ本文を用いる。取得したテキストに対して形態素解析を行い、一般名詞、固有名詞を抽出する。さらに、抽出された一般名詞、固有名詞に対して、Wikipedia の全文書をもとに idf 値を求め、閾値以上の idf 値を持つ語を特徴語としている。

3.2 Web ページの内容を表す語の抽出

ここでは、Web ページの内容を表す“内容語”を抽出する手法について説明する。本手法ではスニペット文から特徴語を取得する。高い精度を求める場合には Web ページの本文を直接解析できることが望ましいが、検索結果すべてにアクセスするのは処理時間として現実的ではないため、スニペット文から内容を表す特徴語を抽出することにした。なお、本稿では Google 検索で生成されるスニペットを用いた。詳しくは 4.1 節で述べる。

我々はスニペット文に含まれる単語集合に対しクラスタリングすることで生成されるクラスタが、それぞれ文章のおおまかなトピックを表すと考えた。また要素数が最大となるクラスタは、その文章の中心的なトピックを表すと考えた。内容語抽出手法の概要を図 1 に示す。

まず、スニペット文から抽出した特徴語集合に対し用意した単語ベクトルモデルを用いてコサイン距離で階層的クラスタリングを行う。このクラスタリングによって生成されるクラスタ群の中から要素数が最大のクラスタに属している特徴語を“内容語”とした。

3.3 既知/未知話題を表現する語の抽出

ここでは、ユーザの訪問済みページから抽出された特徴語を

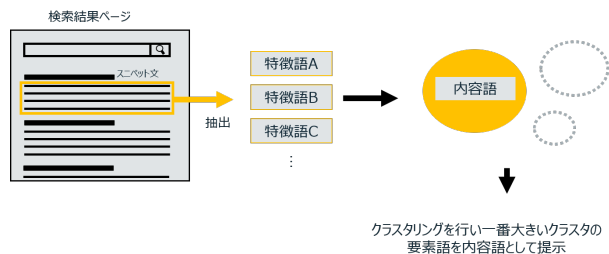


図 1 内容語抽出手法

Fig. 1 Content Words extraction method

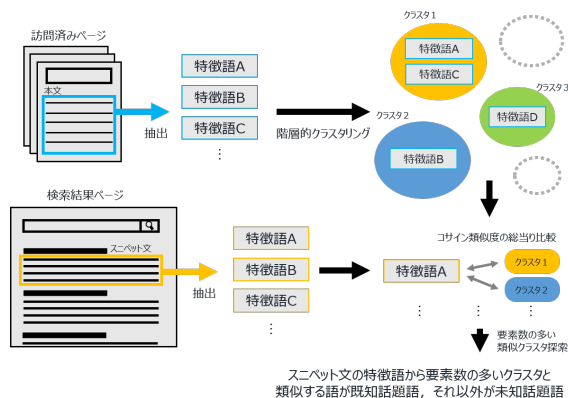


図 2 既知/未知話題語抽出手法

Fig. 2 Known/Unknown-topic Words extraction method

用いて、検索結果の各スニペット文から“未知話題語”，“既知話題語”を抽出する手法について説明する。

我々は訪問済みページの特徴語集合に対してクラスタリングした結果が、おおよかなユーザの未知，既知のトピック分布を表せると考えた。そして、各クラスターの要素数でトピックの既知度合いを表現できると考え、あるトピックについて閲覧していくとそのトピックを表すクラスターの要素数が増えていき、その要素数で既知/未知話題を表現できると考えた。既知/未知話題抽出手法の概要を図 2 に示す。

まず、抽出した特徴語についてコサイン距離による階層的クラスタリングを行い、その結果で得られたクラスター群を既読トピッククラスターとした。この既読トピッククラスターとスニペット文の各特徴語を総当たりで類似度を算出した。そして、要素数が多い既読トピッククラスターとコサイン類似度が高いスニペット文の特徴語を“既知話題語”とし、各スニペット文からの抽出した特徴語と“既知話題語”の差集合を“未知話題語”とした。ここでのクラスターと特徴語の類似度は、クラスタリングでも用いた単語ベクトルモデルを使って求めたクラスターの各単語と特徴語のコサイン類似度の平均値で算出している。

3.4 提示例

前述の方法で得られた“内容語”と“既知話題語”，“未知話題語”は SERP の各検索結果タイトルとスニペット文の間に提示する。実際の提示例を図 3 に示す。

拡張情報はクエッションマークで示す“未知話題語”の行とチェックマークで示す“既知話題語”の 2 行で構成される。スニペット文から抽出された特徴語の中で“内容語”は太字で表示さ

れている。

本手法では、訪問済みの検索結果にも“未知話題語”が表示されることがある。これはシステムのアルゴリズムを理解していないと混乱を招くと考えられることから、訪問済みの検索結果には拡張情報が表示されないようにした。

4. 評価実験

本稿では検索行動を分析し提案する拡張情報提示が、ユーザの検索行動にどのような影響を与えるのか評価を行う。そのために、検索行動ログを収集するための実験システムを構築した。これを用いてクラウドソーシングサービス“CrowdWorks”^(注1)で被験者を募り検索行動ログを収集した。被験者は延べ 360 人である。本節では構築した実験システム、本実験の検索タスクと検索達成の確認手法、収集した検索行動ログ、事後アンケートについて述べる。

4.1 実験システム

実験システムは、検索エンジン、キーワード収集機能、実験タスクによっては拡張情報の提示、クリックされたページ URL や入力された検索クエリなどを含む検索行動ログ記録の機能を持つ。拡張情報は SERP 上に提示され、SERP にアクセスした際と、SERP から閲覧ページにアクセスした際にログが記録される。PC ブラウザで Google Chrome 環境を想定している。初回の SERP では訪問済みページが存在しないため、各検索結果にはすべて未知話題語が提示される。

今回 SERP 上に表示される検索結果は Google Custom Search API を利用して取得した。また、検索クエリ 1 つに対して上位 10 件の検索結果を表示した。実験システムの検索ページには従来の検索エンジン同様の検索結果と検索クエリの入力フォーム、右上には実験を終了するボタン、閲覧ページでは左上に検索結果ページに戻るボタンを表示した。また、本稿では検索網羅性の観点からも調査を行うため、ユーザが検索中見つけた気になるキーワードを収集するための機能を持ったフォームを右のバーに表示した。実験ページを図 4 に示す。

閲覧ページは Web ページをそのままダイレクトするのではなく、ページの内容部分を表示した。ページの内容部分の抽出には Python の readability-lxml^(注2)というライブラリを用いて Summary を取り出す機能を使った。また、閲覧ページ上部に title タグから取得した Web ページのタイトルを表示した。

閲覧ページでは、既読クラスター生成のために訪問済みページで表示される内容部分を取得する。これらを基に既読トピッククラスターを生成する。ここで、訪問済みページ数が増えるとクラスタリングの対象となる訪問済みページから抽出され蓄積された特徴語が増え、処理に時間がかかる問題がある。そこで、Wikipedia の全文書をもとに idf 値を算出した idf 値が 6.7 以上かつ単語出現数と idf をかけ合わせた値の上位 500 語を採用した。また、要素数が少なすぎるクラスターについてはノイズであることが多いことから、要素数が 1 のクラスターは除外して

(注1) : <https://crowdworks.jp/>

(注2) : <https://github.com/buriy/python-readability>

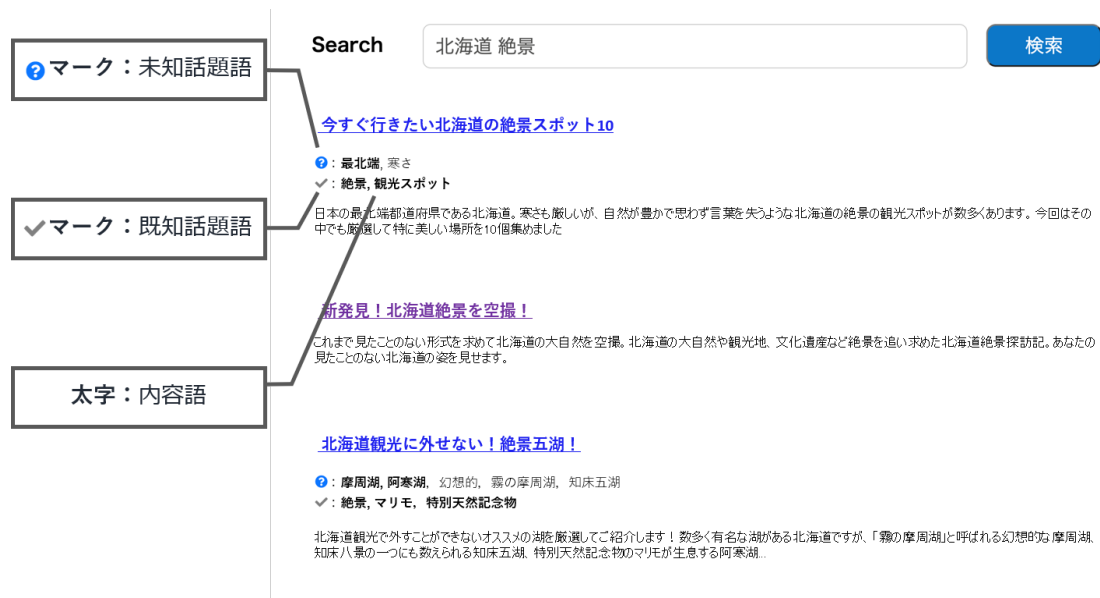


図 3 拡張情報の提示例
Fig. 3 Example of an Enhanced presentation

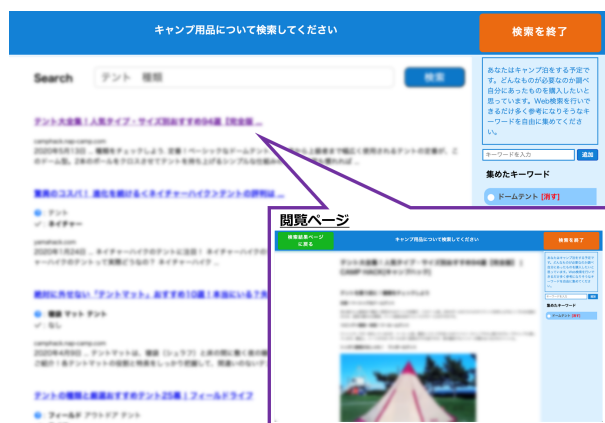


図 4 実験ページ
Fig. 4 Example of experiment pages

いる．今回要素数が 10 より多いクラスタを既知であるトピックを表すクラスタとして採用した．さらに，既知であるクラスタ内の各特徴語と判定するスニペット文の特徴語を総当りでコサイン類似度比較をし，その平均値が 0.3 より高いスニペット文の特徴語を既知話題語として採用した．未知話題語は，既知話題語以外のスニペット文の特徴語とした．本実験では，学習済みの単語の分散表現として，Wikipedia ダンプを形態素解析器である MeCab に辞書「mecab-ipadic-NEologd」を用いて分かち書きし，fasttext [19] で学習したものを用いた^(注3)．また，Wikipedia の全文書をもとに idf 値を算出し 6.7 以上の値を持つ単語に関して特徴語とした．内容語や既知話題語を得るための階層的クラスタリングは群平均法を用い，コサイン距離は 0.7 を採用した．

こうして生成され記録された既読トピッククラスタは閲覧ページから検索結果ページに戻るボタンをトリガとし，常に最

新の既知/未知話題語を SERP 上の拡張情報に反映している．

4.2 実験手法

被験者はクラウドソーシングサービスの仕事依頼画面から実験開始ページへアクセスし，実験を開始する．仕事依頼画面には実験の流れ，実験ページの説明，検索タスク，アンケート，拡張情報が提示されるタスクの場合には拡張情報についての説明図が提示されている．仕事依頼画面を見て実験に同意をした被験者は実験説明文と検索タスクを確認し，実験開始ページへアクセスする．実験開始ページからの実験の流れを以下に示す．

- (1) 実験開始ページにアクセスしユーザ ID を入力
- (2) 実験システム (図 4) を使って検索
- (3) 実験完了リンクをクリック
- (4) 事後アンケートに回答

4.3 検索タスク

本実験では，4 つの網羅的検索タスクを用意した．各検索タスクに対し，Standard, Enhanced, Limited の 3 手法を用意し，合計 12 種類のジョブをクラウドソーシングサービス上で募集した．Standard は記事の閲覧と検索ができるシンプルな検索インターフェース，Enhanced は Standard をベースに拡張情報提示がある検索インターフェース，Limited は Enhanced をベースに同じクエリで 4 ページ以上記事を読んだ場合に未知話題語を非表示にした検索インターフェースである．これら 3 手法は順にシステムがユーザの検索行動への介入度合いを強め，未知話題語の提示の効果と，クエリの変更という区切りを与えられるかを調査するために用意した．各タスクは 1 人あたり 1 回までの参加とし，同じ検索タスクは 1 回しか参加できないように実験システム側で制限を設けた．被験者は検索タスクを仕事依頼画面上で確認し，実験ページを用いてキーワードを収集しながら自由に検索を行い，検索できたと感じたら，検索行動を終了する．

(注3) : <https://qiita.com/Hironasan/items/513b9f93752eccc9e670>

4.4 収集した検索行動ログ

検索行動ログにはユーザ ID、ジョブの種類、URL、アクセスした時間、検索クエリが記録されている。今回は、各ジョブ 30 件ずつ検索行動ログを収集した。検索行動ログの他に、検索結果に使用した Google Custom Search API からのレスポンス、各閲覧ページを見た際に生成される既読トピッククラス、収集したキーワードなども記録されている。また、以降本稿で述べる検索行動ログとは、ユーザのジョブにおける検索行動の一連のログ集合を指す。

4.5 アンケート調査

本実験では検索行動の前後にそれぞれアンケートを実施した。検索行動前では、検索タスクを提示した直後に事前アンケートとして、その検索タスクに関して事前知識を持っているか質問した。検索行動後では、検索行動において新しい発見があったかどうか、満足に検索できたかについての質問を実施した。また、拡張情報ありのシステムを利用した被験者には拡張情報の印象評価についての質問を 4 項目実施した。被験者は各質問に対して、5 段階のリッカート尺度で回答した。

5. 拡張情報提示の評価

本節では収集した検索行動ログから拡張情報提示の評価を行う。拡張情報の有無で各ユーザの検索行動ログを比較することで、拡張情報の提示による影響を調査する。主に、入力クエリ数、検索時間、閲覧ページ数の観点からの検索行動の変化について述べる。他にユーザの主観を軸とした、アンケートの回答から事前知識の有無、新しい発見の有無、満足度という観点からの検索行動の変化についても述べる。

5.1 検索行動ログの選別

収集した検索行動ログには報酬目的で早く終わらせた検索行動や、明らかに他のユーザより異なる検索行動が存在する。このような、不適切な被験者や外れ値となる被験者の検索行動を含めて集計すると変化の傾向が見えづらくなってしまう。そのため、以下の条件に当てはまる検索行動ログを分析対象にした。その結果、分析に用いる検索行動ログ（延べ被験者数）は、200 件となった。

- ・ 閲覧ページが 2 ページより多い
- ・ 収集キーワードが 6 個より多く、80 個より少ない
- ・ 検索時間が 100 秒より長く、3000 秒より短い
- ・ 入力クエリ数が 2 つより多く、70 個より少ない

5.2 全体での比較

ここでは、収集した各ユーザの検索行動ログから入力クエリ数、検索時間、SERP 閲覧時間、Web ページ閲覧時間、閲覧ページ数、収集キーワード数を集計し、手法ごとに平均値を算出し比較を行う。

入力クエリ数は単純にユーザが検索クエリを入力した回数と、重複した入力クエリをカウントしなかった回数の両方を示す。この 2 つの違いは、入力クエリ数は違うクエリを挟んでもう一度 1 つ前のクエリに戻った場合にもカウントされてしまうが、重複なしの回数ではこうした場合カウントされないことである。検索時間は、最初のログが書き込まれた時間から最後のログが

表 1 全体での被験者一人あたりの平均

	Standard	Enhanced	Limited
入力クエリ数	12.0	12.2	12.8
入力クエリ数（重複なし）	9.9	11.4	10.5
検索時間（秒）	956.9	912.3	1014.8
SERP 閲覧時間（秒）	445.4	532.9	529.1
Web ページ閲覧時間（秒）	511.4	380.9	486.5
閲覧ページ数	12.9	12.8	14.9
エラーページ数	3.2	3.0	2.9
収集キーワード数	15.0	12.2	15.6
延べ被験者数	62	66	72
タスク 1	12	17	16
タスク 2	13	19	16
タスク 3	19	18	17
タスク 4	18	12	23

書き込まれた時間を示している。エラーページ数は、ユーザが見たいと思って Web ページの構造上、実験システムが解析できなかったページが存在するため、そのようなページにアクセスを試みた回数を示す。エラーページは閲覧ページとはカウントされず、多くのエラーページにアクセスしている場合検索時間に影響を及ぼす可能性があるため、参考として示した。収集キーワード数はユーザが網羅的検索を行ったかどうかの確認をするためのものであり、平均 12 から 15 個のキーワードが収集されており、ある程度検索タスクに対して、網羅的に情報を得たものと考えている。

表 1 に示すように、入力クエリ数は Standard, Enhanced, Limited の順に増加し、検索時間と閲覧ページ数は Enhanced の方が Standard, Limited よりも少ないことがわかる。このことから、全体的に Enhanced で提示される検索行動は短くなる傾向にあると考えられる。また、Enhanced は、ウェブページの閲覧時間よりも SERP の閲覧時間の方が長いことが確認できる。

Limited は、入力クエリ数や検索行動が増加するように見える。ウェブページの閲覧時間に対する SERP の閲覧時間は、Enhanced の場合ほどではないが、SERP の閲覧時間が若干長くなっている。Enhanced と Limited の効果を比較するためには、1 回のクエリで 4 ページ以上のページが閲覧された場合のみを分析する必要がある。詳細は 5.5 節で説明する。

5.3 入力クエリ数によるユーザ分類での比較

ここでは、全検索行動を入力クエリ数で k-means による非階層的クラスタリングを行い、3 つに対象ユーザを分類し分析を行う。分類されたユーザグループ内で手法ごとに平均値を算出し、拡張情報の評価を行う。また、検索タスクごとのユーザ数に大きな偏りが無いことから、検索タスクに依存した結果ではないことがわかる。また、どのクラスタにおいても、収集キーワード数に大きな差はなく、入力クエリ数に関わらず被験者は網羅的に情報を収集したと判断してタスクを終えていると考える。

表 2 の結果から、入力クエリ数が少ない Enhanced のユーザの検索行動は、短いことがわかる。また、入力クエリ数が少

表 2 入力クエリ数分類での被験者一人あたりの平均

	入力クエリ数が少ないユーザ群			中間ユーザ群			入力クエリ数が多いユーザ群		
	Standard	Enhanced	Limited	Standard	Enhanced	Limited	Standard	Enhanced	Limited
入力クエリ数	6.6	7.1	6.7	15.8	16.6	17.2	27.1	34.4	31.4
入力クエリ数（重複なし）	6.3	6.5	6.3	12.9	15.3	13.2	17.9	33.6	25.0
検索時間（秒）	879.0	725.2	787.4	1056.5	1077.7	1283.6	1036.5	1713.7	1353.6
SERP 閲覧時間（秒）	339.1	373.3	379.9	554.3	688.2	659.5	634.9	1158.1	915.3
Web ページ閲覧時間（秒）	539.7	352.5	408.4	502.4	391.0	625.3	400.8	565.6	437.4
閲覧ページ数	12.0	9.3	10.3	14.3	14.5	17.9	13.0	34.2	30.4
エラーページ数	2.7	2.2	2.3	3.9	3.2	3.3	3.2	6.8	4.5
収集キーワード数	14.1	11.7	15.3	15.0	13.3	16.4	19.4	11.8	14.1
延べ被験者数	34	40	40	21	21	25	7	5	7
タスク 1	6	10	9	3	5	6	3	2	1
タスク 2	8	13	12	4	5	4	1	1	0
タスク 3	9	11	6	9	6	6	1	1	5
タスク 4	11	6	13	5	5	9	2	1	1

表 3 事前知識の有無での被験者一人あたりの平均

	事前知識がないユーザ群			中間ユーザ群			事前知識があるユーザ群		
	Standard	Enhanced	Limited	Standard	Enhanced	Limited	Standard	Enhanced	Limited
事前知識	1.8	1.7	1.7	3.0	3.0	3.0	4.1	4.1	4.0
入力クエリ数	12.3	12.1	11.9	11.6	12.1	13.4	11.7	12.5	13.3
入力クエリ数（重複なし）	10.3	11.6	9.7	8.5	10.9	10.3	9.9	11.1	12.3
検索時間（秒）	995.6	979.6	1088.4	909.0	866.8	882.3	892.2	750.9	1057.1
SERP 閲覧時間（秒）	432.3	557.0	513.4	468.0	527.3	555.3	462.2	467.0	522.6
Web ページ閲覧時間（秒）	563.2	424.5	575.7	441.3	341.0	328.0	429.6	284.3	535.4
閲覧ページ数	12.9	12.9	14.4	12.1	13.9	13.2	13.7	11.8	18.0
エラーページ数	3.3	3.3	2.6	2.5	3.1	2.7	3.3	1.9	3.7
収集キーワード数	14.9	12.0	14.7	16.9	12.5	16.0	13.5	12.6	16.6
延べ被験者数	37	41	31	11	11	23	14	14	18

ないほど SERP の閲覧時間が短くなる傾向がある。

SERP と Web ページの閲覧時間を見ると、入力クエリが多いほど SERP の閲覧時間の比率が高くなることが確認できる。拡張情報の有無に問わずこの傾向が見られるが、Enhanced の場合は、検索時間が長くなるため顕著である。これは入力クエリ数が多いユーザに対して提示された拡張情報の効果が薄れ、検索結果から適切なページを選べなかったことが考えられる。一方で、Limited は、その中間の比率を示している。

入力クエリが多いユーザは、拡張情報の効果が薄れ、検索結果から適切なページを選択できなかった可能性がある。このことから、入力クエリ数が少ない場合には、拡張情報を利用して適切なページを選択することができ、効率的な検索動作をサポートできる可能性があることが示唆された。一方、提案手法における既知／未知のトピック単語は、トピック／サブトピックを考慮していない。そのため、親トピックが異なるサブトピックのキーワードを既知のトピック単語と誤解してしまう可能性がある。また、検索が長くなるにつれて、蓄積された訪問ページの上位 500 個の特徴語しか読み取れず、既知の話題語をうまく抽出できなかった可能性もある。

今後は、検索トピックの切り替わりや長い検索行動にも対応できるように改善していく必要がある。

5.4 事前知識での比較

ここでは、実施した事前アンケートの結果を元に事前知識の有無と拡張情報の効果を比較する。事前知識は全く知識がないを 1、とても知識があるを 5 としたときの集計結果を元に、1 または 2 と回答したユーザを事前知識のないユーザ群、3 と回答したユーザを中間ユーザ群、4 または 5 と回答したユーザを事前知識のあるユーザ群と分けた。事前知識の有無での検索行動の平均値を表 3 示す。

結果からは、全体的にどの層のユーザも Enhanced の検索時間が短くなっている。特に事前知識があるユーザに対してその効果が確認できる。事前知識がないユーザの検索行動では適切な検索クエリを入力できずに長くなってしまい、拡張情報の効果を十分に発揮できなかったことが考えられる。これについては、クエリ推薦などの検索支援技術と組み合わせることや、検索結果中に必要な情報がないことを瞬時に判断し見切りを促す仕組みを検討することでより効果が発揮できると考える。

5.5 拡張情報提示の比較

ここでは、2 種類の拡張情報提示手法、Enhanced と Limited を比較する。Limited は、Enhanced をベースに、同じ検索クエリで 4 つ以上の Web ページが閲覧された場合に、未知話題語を非表示にするものである。そこで、この条件に合う「同じ検索クエリで 4 つ以上の Web ページが閲覧した」検索行動に絞

表 4 閲覧ページ数（閲覧 4 ページ以上の検索行動ログ）

	Standard	Enhanced	Limited
閲覧ページ数	6.7	6.3	6.1
入力クエリ数	60	46	68

表 5 検索の満足度と発見度合いに関する回答

	Standard	Enhanced	Limited
満足度	3.9	3.8	3.7
発見度	4.3	3.9	4.1

り、各手法の平均閲覧ページ数を算出した。結果を表 4 に示す。

この結果を見ると、Standard, Enhanced, Limited の順に Web ページの閲覧回数が減少していることがわかる。これは、未知話題語を早期に排除することで、現在のクエリでの閲覧を終了させる効果があることを示唆している。また、Enhanced と Limited を比較した場合に、ユーザあたりの入力クエリ数が増加しており、Enhanced が適切に制御されていないと、検索行動を悪化させる可能性があることも確認できる。

今回の拡張情報は、補完集合で未知話題語を定義しているため、実際のユーザの閲覧状況を反映し、適切に削減することができていない。次の検索にスムーズに移行するためには適切な介入が必要である。表示される単語の精度を向上させることで、検索行動が改善される可能性がある。

5.6 ユーザの主観評価での比較

ここでは、被験者が回答したアンケートを元に、今回の実験での拡張情報の効果、印象を確認する。まず、被験者が検索終了後に回答した自身の検索行動を振り返るアンケートである。検索で新しい発見があったかどうか、満足に検索できたかどうかについて、1 をネガティブ、5 をポジティブな回答としたときの結果を表 5 に示す。

この結果から、検索と発見に対する満足度は平均的に概ね肯定的であり、2つの提案手法のうち満足度が高いのは Enhanced であることが確認できた。発見度合いは、Standard に次いで介入のレベルが高い Limited の方が高いことが確認できる。Limited は、これまで見てきたユーザの能動的な検索行動と合わせて、参加者自身が発見をしていることを示しており、網羅的に検索ができている可能性が高いと示唆される。

6. おわりに

本稿では、ユーザが多くの検索結果から閲覧ページを選ぶ際に、訪問済みページの内容と同じようなページなどを閲覧することで時間や労力を無駄にしまう問題に対し、閲覧判断基準となる“内容語”と既知度合いを表現する“未知話題語”、“既知話題語”で構成された拡張情報の評価を行った。クラウドソーシングサービスで被験者を募り、構築した実験システムで収集した計 200 件以上の検索行動を元に、網羅的な検索タスクにおける拡張情報提示の評価を行った。

収集した検索行動ログを解析した結果、提案した拡張情報提示が検索行動の検索時間の短縮に有効であることを確認した。一方で、SERP の閲覧に時間がかかり、検索結果から SERP を

選択できないユーザも確認できた。このようなユーザに対しては、すぐに切り上げて検索クエリを変更させるためのクエリ推薦や他の検索支援方法を組み合わせることが有効であると考えられる。

システム介入の度合いを変えた 2 つの拡張情報提示を比較した結果、未知話題語を早期に排除することで、現在のクエリでの閲覧を中止させる効果がある可能性を確認できた。未知話題語を強制的に非表示にすると、拡張情報提示としての検索行動が悪化するため、他の効果的な介入方法を検討する必要がある。

今後は、拡張情報提示における特徴語の精度を向上させるためのアプローチを検討し、ユーザに負担をかけずに、クエリの変更を促し閲覧を切り上げさせることができるような方法で提示することを計画している。また、今回トピック抽出するにあたって、クラスタリングを用いた類似単語集合ベースでの検討を行ったが、今後正確なトピックを抽出するためにトピックモデルを用いた手法を検討していく必要がある。

謝 辞

本研究の一部は、2020 年度科研費基盤研究 (C)(課題番号：18K11551) によるものです。ここに記して謝意を表すものとします。

文 献

- [1] 山田純平, 北山大輔. ユーザの未訪問ページ予測のための拡張スニペットによる検索支援手法. 第 24 回一般社団法人情報処理学会シンポジウム インタラクション 2020 論文集, pp. 1B-43, 2020.
- [2] 山田純平, 北山大輔. ユーザの習得度を表現する web 検索スニペットを用いた検索行動の調査. Technical Report 5, 工学院大学, 工学院大学, aug 2020.
- [3] 倉門浩二, 大石哲也, 長谷川隆三, 藤田博, 越村三幸. Wikipedia のリンク共起とカテゴリに基づくリランキング手法. 研究報告データベースシステム (DBS), Vol. 2010, No. 12, pp. 1-8, jul 2010.
- [4] Elif Aktolga and James Allan. Reranking search results for sparse queries. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management, CIKM '11*, pp. 173-182, New York, NY, USA, 2011. Association for Computing Machinery.
- [5] Yan Chen and Yan-Qing Zhang. A query substitution-search result refinement approach for long query web searches. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Volume 01, WI-IAT '09*, pp. 245-251, USA, 2009. IEEE Computer Society.
- [6] 大石哲也, 倉元俊介, 峯恒憲, 長谷川隆三, 藤田博, 越村三幸, 堀憲太郎. 関連単語抽出アルゴリズムを用いた web 検索クエリの生成. 情報処理学会研究報告, No. 56(2008-DBS-145), 2008.
- [7] Matthias Keller, Patrick Mühlischlegel, and Hannes Hartenstein. Search result presentation: Supporting post-search navigation by integration of taxonomy data. In *Proceedings of the 22nd International Conference on World Wide Web, WWW '13 Companion*, pp. 1269-1274, New York, NY, USA, 2013. Association for Computing Machinery.
- [8] 梅本和俊, 山本岳洋, 田中克己. 網羅性指向タスクにおける未閲覧情報量の提示. 人工知能学会論文誌, Vol. 32, No. 1, pp. 1-12, 2017.
- [9] 早乙女高大, 相田仁. ウェブブラウザにおける既読コンテンツの検出・表示手法の検討. 研究報告情報基礎とアクセス技術

(IFAT) , 第 2017-DC-104 巻, pp. 1—8, mar 2017.

- [10] Henry Feild, Ryen W. White, and Xin Fu. Supporting orientation during search result examination. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pp. 2999–3008, New York, NY, USA, 2013. Association for Computing Machinery.
- [11] Mikhail Ageev, Dmitry Lagun, and Eugene Agichtein. Improving search result summaries by using searcher behavior data. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '13, pp. 13–22, New York, NY, USA, 2013. Association for Computing Machinery.
- [12] Rishita Kalyani and Ujwal Gadiraju. Understanding user search behavior across varying cognitive levels. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media*, HT '19, pp. 123–132, New York, NY, USA, 2019. Association for Computing Machinery.
- [13] Ariyam Das, Chittaranjan Mandal, and Chris Reade. Determining the user intent behind web search queries by learning from past user interactions with search results. In *Proceedings of the 19th International Conference on Management of Data*, COMAD '13, pp. 135–138, Mumbai, Maharashtra, IND, 2013. Computer Society of India.
- [14] Kazuma Nagano, Yutaka Arakawa, and Keiich Yasumoto. Trackthink: A tool for tracking a thought process on web search. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, UbiComp '17, pp. 681–687, New York, NY, USA, 2017. Association for Computing Machinery.
- [15] Jin Young Kim, Mark Cramer, Jaime Teevan, and Dmitry Lagun. Understanding how people interact with web search results that change in real-time using implicit feedback. In *Proceedings of the 22nd ACM International Conference on Information Knowledge Management*, CIKM '13, pp. 2321–2326, New York, NY, USA, 2013. Association for Computing Machinery.
- [16] Jon Pearce, Shanton Chang, Basil Alzougool, Gregor Kennedy, Mary Ainley, and Susan Rodrigues. Search or explore: Do you know what you're looking for? In *Proceedings of the 23rd Australian Computer-Human Interaction Conference*, OzCHI '11, pp. 253–256, New York, NY, USA, 2011. Association for Computing Machinery.
- [17] Ahmed Hassan, Ryen W. White, Susan T. Dumais, and Yi-Min Wang. Struggling or exploring? disambiguating long search sessions. In *Proceedings of the 7th ACM International Conference on Web Search and Data Mining*, WSDM '14, pp. 53–62, New York, NY, USA, 2014. Association for Computing Machinery.
- [18] 梅本和俊, 山本岳洋, 田中克己. 検索専門性と事前知識に着目した検索行動とタスク満足度の関係性分析. 情報処理学会論文誌データベース (TOD) , Vol. 7, No. 4, pp. 13–28, 2014.
- [19] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, Vol. 5, pp. 135–146, 2017.
- [20] Mengyang Liu, Yiqun Liu, Jiaxin Mao, Cheng Luo, Min Zhang, and Shaoping Ma. "satisfaction with failure" or "unsatisfied success": Investigating the relationship between search success and user satisfaction. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, pp. 1533–1542, Republic and Canton of Geneva, Switzerland, 2018. International World Wide Web Conferences Steering Committee.