
検索リテラシ育成のための検索結果可視化システム

理工学研究科 真 野 宗 和
システム理工学部教授 小 尻 智 子

1. はじめに

今や私たちの生活にウェブ上の膨大な情報は必要不可欠なものとなっており、それらの中から私たちが欲しい情報を探す手段として、Google などの検索エンジンもまた、必要不可欠なツールとなっている。検索エンジンでは欲しい情報を得られるような語句を検索クエリとして設定して検索結果を得る。検索結果は多くの場合、検索クエリと関連の強いものから弱いものまで、多数の情報が出力される。通常私たちはこれらをすべて見るということはず、上位のページをみるだけのことが多い。欲しい情報が検索結果の上位に存在しない場合、私たちは検索結果の内容と欲しい情報との差分を考慮して検索クエリを変更する。適切な検索クエリを設定できなければ、検索結果として欲しい情報が得ることができない。したがって、私たちが ICT を基盤とする情報化社会で生きていくためには適切な検索クエリを設定する能力、すなわち検索リテラシが求められる。

検索リテラシを対象とした研究には様々なものが存在する。Nguyen らは膨大な検索結果の中から欲しい情報を見つけるという活動に焦点をあて、ユーザの興味を考慮した検索結果の可視化システムを提案している¹⁾。これはユーザに検索結果の語句に対して興味のある語句の選択とその語句に対する重み付けを行わせ、そのうえでユーザの興味のある語句と関係のあるページを重みに比例した大きさで回転させることで、興味のあるページを見つけやすいように支援している。この方法は、検索結果に興味のある語句が含まれている場合は有効であるが、含まれていない場合は欲しい情報を得ることはできない。欲しい情報を得るには適切な検索クエリを設定しなければならない。システムによる検索クエリの設定に関する研究も存在する^{2, 3)}が、これらの研究では検索クエリの設定は可能となるが、システムが検索クエリを設定するため、ユーザ自身は適切な検索クエリが設定できるようにはならない。したがって、システムがない環境では情報の収集ができない可能性がある。

ユーザ自身による検索クエリの設定支援として、長畑らはユーザが欲しい情報を得るために試行錯誤した一連の検索結果に含まれる名詞の出現頻度の変化をもとに、有用な語や不要な語を推定し、可視化している⁴⁾。この手法では、出現頻度をもとに計算した名詞のスコア

をもとに、常に高いスコアを維持している語やスコアが大きく増減した語などにわけて名詞を表示することで、ユーザが検索クエリとして設定すべき語を特定することを支援している。Maらはユーザと同じ検索クエリを設定した人が次に設定した検索クエリがユーザにとって有用であるという考えをもとに、協調フィルタリングの概念を導入した検索クエリの推薦方法を提案している⁵⁾。このようにシステムが検索クエリの候補となる語句を提示する手法では、ユーザはより適切な検索クエリに変更することができるが、その検索クエリに変更する理由を理解することは出来ない。したがって、システムがなければ検索クエリを設定できない。

一方、検索リテラシの育成支援として、齋藤らは検索クエリの変更か検索結果のリンク先の閲覧かという行動の選択能力に焦点をあて、ユーザの検索活動のリフレクションを促す環境を構築している⁶⁾。この環境ではユーザの行った検索プロセスの提示と、プロセスを省察するための質問やガイドを与える。例えば、「これまでにどういうキーワードをどのように組み合わせましたか？」や「1つのページからどのくらいリンクをたどっていますか？」といった質問を与え、ユーザに自身の検索プロセスを説明させることにより、自身の活動の良し悪しを考えさせる。この研究では検索クエリを変えるべきかという判断能力は育成できるが、検索クエリの設定方法は対象とはしていない。検索クエリを変更する判断ができて、どのような検索クエリを設定すれば良いかが分からなければ欲しい情報を得ることはできない。そこで、本研究では検索リテラシのうち、検索クエリ設定能力の育成を対象とする。

野球のバッティングスイングを対象に能力の育成を支援している西山らによると、自身の活動の振り返りには、その活動によって得られる効果の可視化が有効である⁷⁾。検索クエリの効果は前の検索結果との変化に応じて決まる。したがって、検索クエリの変更に対する状態の変化を理解することは、状態に応じた検索クエリの設定に繋がる。そこで、本研究では検索結果の状態を可視化する手法を提案するとともに、検索クエリ変更前後の状態変化を比較できるシステムを提案することで、検索クエリの設定能力を育成する。

2. 検索クエリ設定能力育成のためのアプローチ

2.1 情報検索のプロセス

情報検索のプロセスを図1に示す。検索エンジンを用いた情報検索では、初めに欲しい情報が得られるような検索クエリを設定して検索する。次に、検索結果を閲覧し、得られた検索結果が欲しい情報か否かを判断する。この際、検索結果のリストに表示されるウェブページのタイトルや概要を基に判断することが多い。欲しい情報を得られなかった場合には、検索結果の内容に基づいて検索クエリを変更する。これを欲しい情

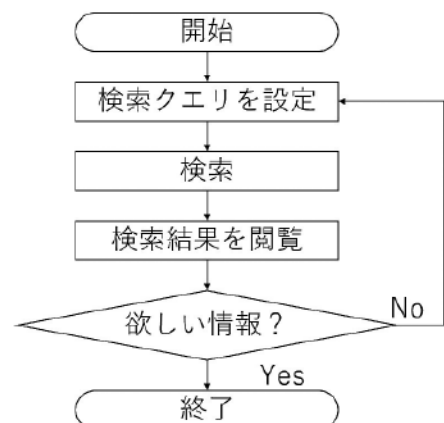


図1 情報検索プロセス

報が得られるまで繰り返す。

例えば、大阪市福島区の居酒屋の情報が欲しい場合に、検索クエリとして「福島、居酒屋」を設定したとする。この検索に対して、欲しい情報である大阪市福島区の居酒屋に加えて、要らない情報である福島市の居酒屋が出てきている場合、検索結果を大阪市に絞るために、「福島、居酒屋、大阪市」と検索クエリに「大阪市」を追加する。このプロセスにおいて、欲しい情報を得るには、検索結果の内容と欲しい情報との差分を考慮して検索クエリを修正することが重要である。

2.2 検索クエリ設定能力

検索クエリ設定能力とは、現在の検索結果の状態に対して、次の検索で期待する状態に変化させるために、検索クエリ中の語句を変更する能力である。検索クエリの設定方法は検索の意図に応じて異なる。Broder は検索クエリ設定の意図を、特定のウェブページへ辿り着くことを目的としたナビゲーションクエリ、特定の情報を得ることを目的としたインフォメーションクエリ、サービスを提供するウェブページを通してサービスを利用することを目的としたトランザクショナルクエリの 3 つに分類している⁸⁾。ナビゲーションクエリは、得られた検索結果が欲しい特定のウェブページの内容と近いかな否か、インフォメーションクエリは、欲しい情報に近いかな否かを判断し、次の検索クエリを設定する。一方、トランザクショナルクエリは、サービスを提供している特定のウェブページへ辿り着くことが目的か、特定のサービスを提供しているページへ辿り着くことが目的であると捉えると、ナビゲーションクエリか、インフォメーションクエリかに分類できる。このことから、検索クエリの設定には、検索結果が特定のページ、もしくは特定の情報をどれくらい含んでいるかという状態を正確に認識できる必要があるといえる。

検索結果の状態に基づいた検索クエリの設定は、検索結果に欲しい情報が含まれていない場合と、不要な情報が含まれている場合で異なる。検索クエリは、類似した話題が近くに来るように配置された話題空間において、焦点を当てる空間を指定する操作とみなすことができる。欲しい話題空間と検索クエリの指定した話題空間が異なっていない場合、欲しい情報を表す語句を追加するとともに、欲しくない情報を表す語句を削除する必要がある。図 2 (a) や図 2 (b) はこれらの語句の追加と削除を両方行う場合である。図 2 (c) は不要な語句はないが欲しい情報を表す語句が不足しているため、語句の追加のみを行う。図 2 (d) は検索クエリが不要な語句を含んでおり、語句の削除のみを行う。ただし、図 2 (d) は欲しい情報を含んでいるため、語句の削除を行わなくても欲しい情報を得ることは可能である。これらの操作を行うことで、検索クエリが表す話題空間を欲しい情報の空間に近づけることができる。

本研究では、得られた検索結果と欲しい情報の差分をイメージし、検索クエリに語句を追加するか削除するかを判断できる能力のことを検索クエリ設定能力という。もちろんこの時、具体的にどんな語句を追加し、削除するかも重要な能力であるが、本研究ではこの具体的な

語句の設定の支援は今後の課題とし、語句の追加、削除を判断する能力の育成を対象とする。

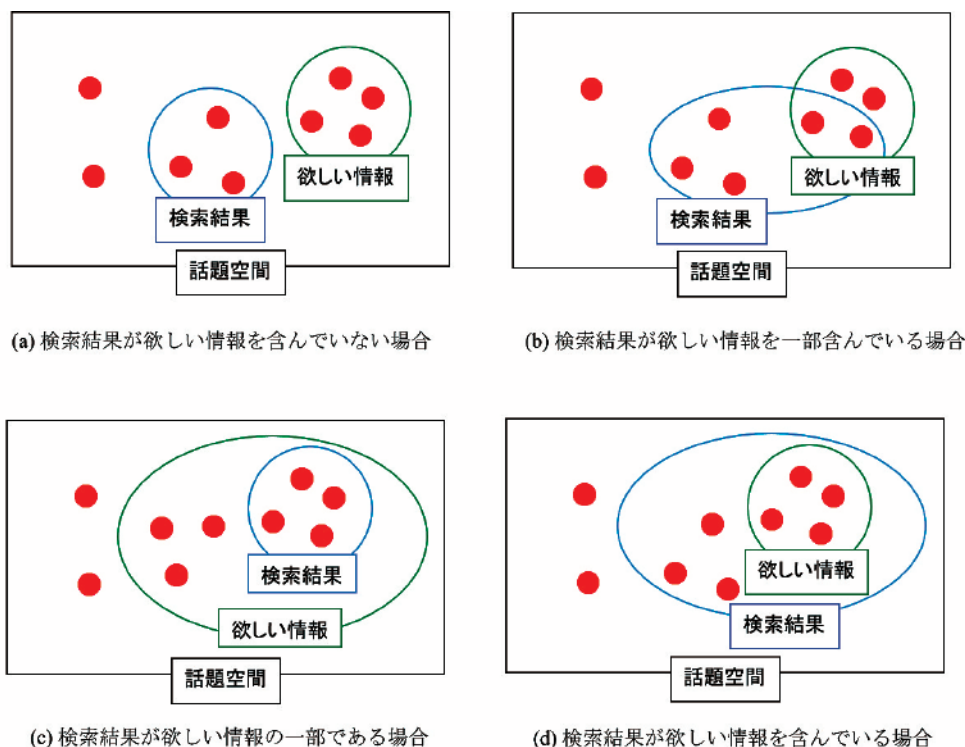


図2 検索結果と欲しい情報との関連

2.3 検索クエリ設定能力の育成支援システムの概要

検索クエリ設定能力の育成には、検索結果の状態を正確に認識できることに加え、検索クエリを変更した際に生じる検索結果の状態変化を理解することが必要である。そのため、タイトルに含まれる語句が欲しい情報か否かをユーザに判別させ、判別された語句の分布を可視化するシステムを構築する。一方、検索クエリがもたらす状態変化を理解するためには、検索クエリを変更したときに話題がどのように変化するかを判断できる必要がある。そこで、システムでは検索クエリの変更前後での語句の分布の比較を可能とし、検索クエリがもたらす語句の分布の変化の理解を促進する。

システムとユーザのインタラクションを図3に示す。ユーザが検索クエリを入力すると、システムは検索結果と、タイトルに含まれる語句を検索された話題として提示する。ユーザは提示された語句が欲しい情報か否かを判断し、システムに入力する。システムはユーザの語句の判別に基づいて現在の検索結果の状態を可視化する。ユーザは可視化された検索結果の状態を基に検索ク

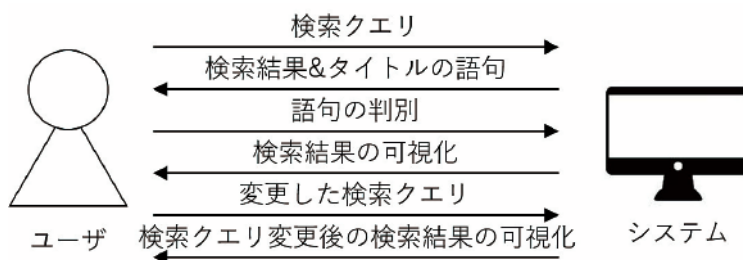


図3 システムとユーザのインタラクション

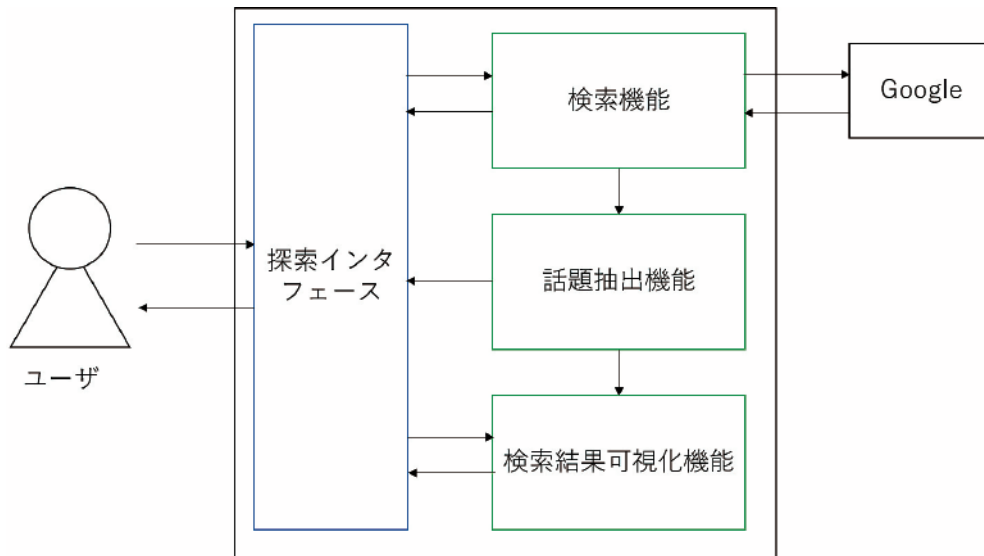


図4 システム構成図

エリを変更し、変更後の状態の可視化を得る。検索クエリ変更前後の状態変化を比較することで検索クエリの変更がもたらす状態変化を理解する。

このようなインタラクションを実現するシステムの構成を図4に示す。検索クエリが入力されると検索機能はその語句をGoogleに入力し、検索結果を得る。話題抽出機能は検索機能によって取得した検索結果のタイトルから話題を表す語句を抽出し、インタフェースを介してユーザーに提示する。これらの語句に対してユーザーから得られた欲しい情報か否かの判別データは、語句と一緒に検索結果可視化機能に送られる。検索結果可視化機能では、語句の出現頻度と語句同士の類似度を計算した後に、出現頻度、類似度、判別データに基づいて可視化する。

3. 話題抽出機能

話題抽出機能は検索されたページのタイトルから話題を表す語句を抽出する機能である。話題は名詞で表現されることが多いが、タイトルには話題と関係のない語句が含まれる可能性もあるため、本研究ではユーザーが設定した検索クエリの検索結果として得られる上位10件のタイトルの中から、品詞が名詞かつ出現回数が2回以上となる語句を出力とする。

名詞の抽出には形態素解析器を用いる。形態素解析は自然言語で書かれた文を言語上の最小単位である形態素に分割し、それぞれの形態素の品詞や変化などを判別する処理である。例えば、「福島（大阪）のウまい居酒屋20選～人気店から穴場まで～ -Retty」を形態素解析すると表1の結果が得られる。表1のうち、「福島、大阪、居酒屋、20、選、人気店、穴場、Retty」が名詞であるので、これらを話題を表す語句として抽出する。同じ語句が複数回抽出された場合には、語句の重複が無いように1つのみを出力とする。

表1 形態素解析結果（形態素：品詞）

福島：名詞，（：記号，大阪：名詞，）：記号，の：助詞，ウマイ：形容詞，
居酒屋：名詞，20：名詞，選：名詞，～：記号，人気店：名詞，から：助詞，
穴場：名詞，まで：助詞，-：記号，Retty：名詞

4. 検索結果可視化機能

検索結果可視化機能は、検索クエリで検索された結果の状態を表示するための機能である。検索結果の状態は、現在の検索クエリで抽出された話題に対してユーザの欲しい情報・不要な情報がどれほど含まれているか

が分かると良い。そこで、検索結果のタイトル中の語句の話題の分布を図5のように現在のクエリが表現している話題からの分布として表現する。半円の中心は現在の検索クエリが示す話題、中心からの距離はタイトル中の語句と検索

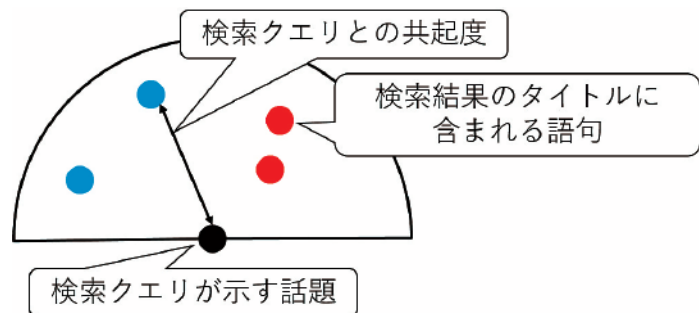


図5 検索結果の可視化

クエリの共起度を表現している。語句同士の距離は類似している語句が近くなるよう配置する。また、ユーザが欲しいと思った話題を赤、欲しくないと思った話題を青で区別することで、現在のクエリが欲しい話題をどれほど含んでいるかを瞬時に認識可能とする。

図5のような配置に対し、すべての語句間の類似度語を計算して2次元上に配置することは計算コストが高い。そこで、本研究では個々の語句の中心からの距離と、基準軸からの角度を以下のように定義する。なお、基準軸は中心から水平右方向への辺とする。

- 中心からの距離：検索結果の上位10ページのタイトルに含まれる語句の割合を共起度とする。式1に共起度の計算式を示す。また、式1で得られた共起度を用いて決定する半円の中心からの距離を式2に示す。例えば、検索結果として出力された上位10件のタイトルのうち6件のタイトルに検索クエリ中の語句aが含まれていたとする。このとき、タイトルと検索クエリ中の語句aの共起度は0.6となる。また、半円の半径が10とすると、語句aの半円の中心からの距離は4となる。

$$\text{共起度} = \frac{\text{語句を含んでいるタイトルの数}}{10} \quad (1)$$

$$\text{中心からの距離} = \text{半径} \times (1 - \text{共起度}) \quad (2)$$

- 基準軸からの角度：検索クエリと語句の概念の類似度をもとに決定する。語句の概念は、それぞれの語句を検索クエリとして検索した結果を示しているとみなすことがで

きる。そこで、現在の検索クエリ (C) と語句 (W) をそれぞれ検索クエリとして設定したときに出てくるタイトルに含まれる語句を要素とする概念ベクトル (V_c, V_w) を用意し、式 3 に示すそれらのコサイン類似度で基準軸からの角度を決定する。ただし、検索結果のタイトルはウェブページの管理者が任意に設定することができるため、必ずしも検索クエリの概念を表している語句が含まれているとは限らない。検索クエリの概念を表さない語句は複数回出現しないと考えられるため、語句の出現回数が 2 回以上の語句を用いて、 V_c と V_w を表現する。

$$\cos(\vec{V}_c, \vec{V}_{w_n}) = \frac{\vec{V}_c \cdot \vec{V}_{w_n}}{|\vec{V}_c| |\vec{V}_{w_n}|} \quad (3)$$

例えば、C の検索結果のタイトルに含まれる語句が (a, a, b, b, c, c, c, c)、W の結果に含まれる語句が (a, a, c, c, d, d, e) とする。このとき 2 回以上出現するのは a ~ d であるため、 V_c, V_w の要素は (a, b, c, d) となり、それぞれの出現回数を用いて V_c は (2, 2, 4, 0)、 V_w は (2, 0, 2, 2) となる。W の位置は V_c と V_w のコサイン類似度に対応する角度を、右の半径からの角度とする。ただし、出現回数を用いてコサイン類似度を算出しているため、コサイン類似度の値は 0 ~ 1 の値を取り、角度の範囲は 0 ~ 90 度となる。そこで、半円に対応させるためコサイン類似度より得られた角度を 2 倍にする。先ほどの V_c, V_w の角度は 45 度であるため、W はこれを 2 倍した 90 度の位置に配置される。

以上のようにして配置された語句の分布より、検索クエリを設定した場合の検索結果の状態の認識が可能となる。また、検索クエリ変更前後の語句の分布を提示することで、検索クエリ変更による効果を理解することが可能となる。例えば、大阪市福島区の居酒屋の情報が欲しい場合に、検索クエリとして「福島, 居酒屋」と設定して検索し、検索結果の話題を表す語句として「福島, 福島区, 福島県, 居酒屋, おすすめ, ランキング, 人気」が得られたとする。要る語句として「福島, 福島区, 居酒屋, おすすめ, ランキング, 人気」、要らない語句として「福島県」を選択した場合、語句の分布は図 6 (a) のようになる。検索結果中に福島県に関する情報が出てきており大阪に関する情報が出ていないと判断して、検索クエリを「福島, 居酒屋, 大阪」に変更したとする。この時の語句の分布が図 6 (b) のようにな

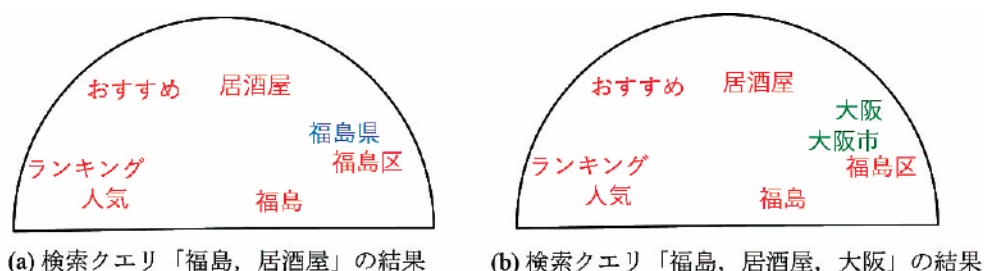


図 6 検索結果の可視化例

ったとすると、検索クエリに「大阪」を追加することで要らない語句と判別した「福島県」を除くことができ、要る語句と判別した「大阪」を多く含むことができるということが理解できる。

5. プロトタイプ・システム

HTML, CSS, Python, Node.js, JavaScript を用いて開発したプロトタイプ・システムのインタフェースを図7に示す。画面①のテキストボックスに検索クエリを入力し、検索ボタンを押すと画面①に検索結果と画面②に語句の一覧が表示される。画面②で表示された語句の一覧に対して必要な語句は赤色、不要な語句は青色、どちらでも良い場合はグレーのラジオボタンをチェックし、可視化ボタンを押すことで画面③に語句の分布が表示される。画面④は検索クエリを変更した時の結果を見ることができる。テキストボックスに変更した検索クエリを入力し、検索ボタンを押すことで語句の分布が表示される。この時、変更前の検索クエリで出てきた語句は画面②で設定した色で、新たに出てきた語句は緑色で表示される。

本研究では、Google からの検索結果を表示するために Custom Search Engine を用いた。また、検索結果の上位10件のタイトルの取得には Custom Search JSON API⁹⁾を、名詞の抽出には新語や固有表現に対応可能な形態素解析器である mecab-ipadic-NEologd¹⁰⁾を用いた。



図7 プロトタイプ・システムのインタフェース

6. 評価実験

6.1 概要

本研究で提案した検索クエリ設定能力の育成支援システムを評価するための実験を行った。実験ではシステムを利用することで設定する検索クエリを適切なものに変更できるようになるかを評価した。実験協力者は本学の学部生11名である。

まず、実験協力者の検索クエリ設定能力を測るため、検索課題と欲しい情報が得られない不十分な検索クエリを与え、欲しい情報をえるためにどのように検索クエリを変更するか、なぜそのように変更したかについて紙上で回答してもらった (Step 1)。その後、システムを用いて、Step 1 とは別の検索課題に取り組んでもらった (Step 2)。システム使用後、必要に応じて Step 1 で設定した検索クエリを変更してもらい、なぜその検索クエリに変更したかについても回答してもらった (Step 3)。最後に、システムに関するアンケートに回答してもらった (Step 4)。

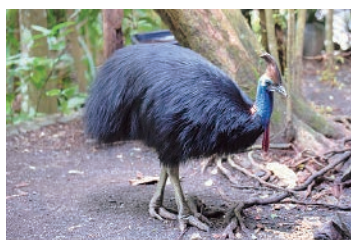
Step 1 では 3 つの検索課題を設定した。表 2 に設定した課題、不十分な検索クエリとして与えた語句、理想的なクエリに含まれているべき語句を示す。1 つ目の検索課題は日本で 3 番目と 5 番目に高い山に関する情報を問うものであり、「日本」「山」「順位」「高い」を検索クエリに含むべき話題とした。2 つ目の検索課題は手に入るゲームに関する情報を得るものであり、「人気」「ゲーム」「入手可能」を含むべき話題とした。3 つ目の検索課題は図 8 の赤線で囲まれた部分の名称を得る課題であり、「飛行機」「尾翼」「後ろ」を含むべき話題とした。Step 2 では、図 9 (a) の生き物の名称を調べる検索課題と、図 9 (b) の船が展示されている場所を調べる検索課題を与えた。

表 2 Step 1 の課題

ID	検索課題	不十分な検索クエリ	理想的なクエリ
1	日本で 3 番目と 5 番目に高い山	日本, 山	日本, 山, 順位, 高い
2	手に入るおもしろいゲーム	人気, ゲーム	人気, ゲーム, 入手可能
3	図 8 の赤枠の部分の名称	飛行機, 尾翼	飛行機, 尾翼, 後ろ



図 8 Step 1 の検索課題 3



(a) 検索課題4



(b) 検索課題5

図 9 Step 2 の検索課題 4, 5

Step 4 で使用したアンケートの内容を表 3 に示す。質問 1 はシステムを用いて語句の分布を比較することで、語句の分布がどのように変化するかをイメージできたかについて回答してもらう項目である。「はい」・「どちらかといえばはい」と答えた場合は持つことのできたイメージについて回答してもらい、そのイメージが正しいかを確認する。「どちらかといえはいえ」・「いいえ」と答えた場合はイメージを持つことができなかった理由を回答してもらい、原因を明らかにする。質問 2 はシステムが使いやすかったかを回答してもらう項目となっている。「はい」・「どちらかといえばはい」と答えた場合は使いやすいと感じた理由を、「どちらかといえはいえ」・「いいえ」と答えた場合は使いにくいと感じた理由を回答してもらう。

表 3 アンケート項目

質問 1	検索クエリを変更したときに語句の分布がどのように変化するかをイメージを持つことはできましたか？ (1 はい・ 2 どちらかといえばはい・ 3 どちらかといえはいえ・ 4 いいえから選択)
質問 2	システムは使いやすかったですか？ (1 はい・ 2 どちらかといえばはい・ 3 どちらかといえはいえ・ 4 いいえから選択)

6.2 結果と考察

Step 1 の全ての検索課題で検索クエリに含むべき話題を含んでいた人はすでに検索クエリ設定能力が備わっていることを示している。また、Step 2 でシステム使用時に検索結果の可視化機能を用いずに検索課題に取り組んだ人はシステムの効果を検証できない。そこで、それらを除いた 4 人（A ～ D）の結果で考察する。

Step 1 と Step 3 の結果をまとめたものを表 4 に示す。検索課題 1 では、Step 1 のシステム使用前に検索クエリを正しく設定できていなかった協力者 A, B, C の 3 名のうち、システム使用後に検索クエリを変更したのは協力者 A, C の 2 名だった。これらの 2 名とも、検索クエリを変更した結果、含むべき話題を含んだ検索クエリとなっていた。検索課題 2 は全員が Step 1 で正しく解答できていた。検索課題 3 では Step 1 で正解できていなかった協力者 B, D の 2 名のうち、検索クエリを変更したのは協力者 D のみだったが、含むべき話題を含んだ検索クエリとはならなかった。これらの結果をまとめると、システム使用前に正しい検索クエリを設定できていなかった 4 名のうち 3 名がシステムを使うことで検索クエリを変更し、2 名が含むべき話題を含んだ検索クエリを設定できた。このことは、本システムはすべての人の検索クエリ設定能力の育成に効果があるとは言えないものの、一定数の人には効果があることを示唆している。

Step 2 のシステムの利用時に、協力者 D は検索クエリを「飛ばない鳥」と設定し、語句として「飛ばない鳥, 下田逸郎, 鳥, LOVE SONGS, AND, 空, たち」が提示された。このとき、「飛ばない鳥, 鳥」を必要な語句、「下田逸郎, LOVE SONGS, 空」を不要な語句、「AND, たち」をどちらでも良いと選択し、語句の分布（図10(a)）を見た。その後、検索クエリを「飛ばない鳥類」に変更し、語句の分布を見た（図10(b)）。この時、図10(b)にヤン

バルクイナしか出ていないという理由で検索クエリを「飛べない鳥類 首 青い」に変更した。これは、Step 3 で与えた検索課題の生き物がヤンバルクイナではないということを理解しており、表示された語句の分布に欲しい情報が表示されていないと判断し、新たな特徴を追加したと考えられる。また、協力者 C は課題 1 の検索クエリ変更理由に「目的のいくつかの山に範囲を少し広げた」と回答している。このことから、視野を広げるために新しい語句を追加するという変更をしていることが分かり、検索クエリの設定能力の育成に効果があった可能性がある。

表 4 Step 1 と Step 3 での検索クエリの変更内容

課題	協力者	Step 1	Step 3	変更理由
1	A	日本, 山, ランキング	日本, 山, 高さ, ランキング	別なランキングが出てきそう
	B	日本, 高い山, 標高	変更なし	-
	C	日本で最も高い山	日本の高い山, 上位	目的のいくつかの山に範囲を少し広げた
	D	日本, 山, 標高, ランキング	変更なし	-
2	A	人気, 最新, ゲーム, ダウンロード可	変更なし	-
	B	人気, 最新, ゲーム, リリース済み	変更なし	-
	C	人気, ゲーム, ランキング, 売れ筋	変更なし	-
	D	人気, 最新, ゲーム, ダウンロード	新作, 人気, ゲーム	新しく作られているゲームだから新作を選んだ。人気ゲームで流行のゲームを調べようとした。
3	A	飛行機, 尾翼, 後ろ側	変更なし	-
	B	飛行機, 尾翼, 部位	変更なし	-
	C	飛行機, 尾翼付近, 名称	変更なし	-
	D	飛行機, 尾翼, しくみ	変飛行機, 尾翼, パーツ名	より多く情報が得られると思った。

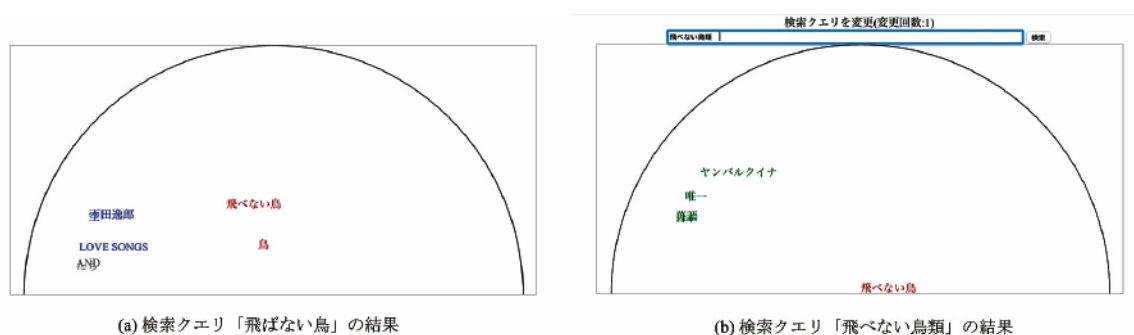


図10 協力者 D のシステム使用例

アンケートの結果を表 5 に示す。質問 1 より、検索クエリの変更に対する語句の分布をイメージできたと思われる協力者は A のみであった。イメージできなかった理由として、真ん中に寄らないという意見があった。これは語句の中心からの距離を検索結果の上位10件のタ

タイトルに占める語句の割合としているが、語句を中心に配置するには上位10件のタイトルすべてにその語句が含まれていなければならない。しかし、検索クエリに含む語句の数を増やした場合には、検索結果のタイトルは多様になり、その語句を含むタイトル数が減ることになる。その結果、中心に寄らなかったと考えられる。可視化の理想として、検索クエリに語句が追加された場合には、追加された語句に関する話題が中心に寄る方がよい。そのため、これらを反映するように改善する必要がある。検索クエリに具体的な語句を設定した場合には、タイトル中の語句より、概要の語句がヒットし上位に表示される場合があるため、概要を考慮して可視化することで改善されると思われる。

システムの使いやすさについては質問2の解答の通り、協力者Bより語句の分布中の文字が重なって読めないという意見があった。語句の位置は検索クエリとの概念の類似度のみに応じて決定しており、語句間の類似はみていない。そのため、検索クエリとの類似度が似ている語句がすべて同じような位置に表示される結果となった。また、表示の際、語句の文字すべてを表示しているため、必然的に文字同士が重なる割合が大きくなってしまっていた。可視化する際に語句をすべて表示せず、先頭の一文字だけとしてクリックしたら語句全体が見えるようにするなど、表示方法を工夫する必要がある。

表5 アンケートの結果

質問	協力者	回答	理由
1	A	2	出てくる回数や量によって近づいたりする。
	B	2	語句が変わることでどんな検索結果かが想像できたため
	C	4	システムをあまり使わなかった
	D	4	変更した結果自分の思い通りの語句の分布の結果が得られなかったから。 真ん中に寄らない
2	A	3	ほとんど普通に検索してしまった
	B	3	語句が散らばっていなかったため読めないことがあった
	C	4	画像検索を使いたかった。分布が文字ばかりで分からなくなった。
	D	2	検索する前に右下で調べられたから事前にクエリの設定を考えた。

7. おわりに

本稿では、検索クエリ変更前後の検索結果の状態を比較することで検索クエリ設定能力を育成することを目的としたプロトタイプ・システムを構築した。本システムは、検索結果の状態を把握するために、検索結果の話題を表す語句を提示し、語句の判別に基づいて検索結果の状態を可視化する機能を有する。また、検索クエリ変更時の効果を理解するために、検索クエリ変更前後における検索結果の状態の可視化を比較することができる環境を構築した。

システム使用前に正しい検索クエリを設定できていなかった4名のうち2名がシステム使用後に正しい検索クエリを設定できたことから検索クエリ設定能力の育成に効果がある可能性が示唆された。今後はさらに評価実験を実施し、本システムの有効性を検証していく必要がある。

ある。一方で、本実験の協力者は普段からスマートフォンなどを用いて日常的に検索をしている大学生だったため、ある程度検索クエリを設定できた可能性が高い。そのため、普段検索をしない年齢層を対象にした実験もする必要がある。

参考文献

- 1) T. N. Nguyen and J. Zhang, “A Novel Visualization Model for Web Search Results”, IEEE Transactions on Visualization and Computer Graphics, Vol. 12, No. 5, pp. 981-988 (2006)
- 2) 鈴木 泉、大里 有生、“Web 検索の認知モデルとその支援システム”、情報処理学会研究報告自然言語処理 (NL)、Vol. 2006, No. 1, pp. 55-60 (2006)
- 3) 吉田 大我、小山 聡、中村 聡史、田中 克己、“Web 検索結果におけるキーワード出現相関の可視化と対話的な質問変換”、電子情報通信学会第18回データ工学ワークショップ (2007)
- 4) 長畑 洋臣、太田 学、“検索結果の推移の可視化による検索支援”、電子情報通信学会技術研究報告、Vol. 108, No. 329, pp. 11-12 (2008)
- 5) C. Ma and B. Zhang, “A New Query Recommendation Method Supporting Exploratory Search Based on Search Goal Shift Graphs”, IEEE Transactions on Knowledge and Data Engineering, Vol. 30, No. 11, pp. 2024-2036 (2018)
- 6) 齋藤 ひとみ、三輪 和久、“Web 情報検索におけるリフレクションの支援”、人工知能学会論文誌、Vol. 19, No. 4, C, pp. 214-224 (2004)
- 7) 西山 武繁、諏訪 正樹、“身体運動時の姿勢変化の分節化によるスキル熟達支援”、身体知研究会、Vol. 1, No. 3, pp. 13-16 (2008)
- 8) A. Broder, “A Taxonomy of Web Search”, ACM SIGIR Forum, Vol. 36, No. 2, pp. 3-10 (2002)
- 9) Google, “Custom Search JSON API”, <https://developers.google.com/custom-search/v1/overview> (参照日2020/2/18)
- 10) 佐藤 敏紀、橋本 泰一、奥村 学、“単語分かち書き辞書 mecab-ipadic-NEologd の実装と情報検索における効果的な使用方法の検討”、言語処理学会第23回年次大会発表論文集、pp. 875-878 (2017)

