

# 探索型検索を用いた学習ナビゲーションツールの検討

村上貴彦<sup>†</sup> 横山誠<sup>‡</sup> 中村太戯留<sup>‡</sup> 上林憲行<sup>‡</sup>

東京工科大学大学院<sup>†</sup> 東京工科大学<sup>‡</sup>

## 1. 目的

わが国では近年、自立的・能動的に学習を行い、変化し続ける社会において自らの力で問題の解決を図れる能力を求めており、受動的な教育の場を見直し自立的・能動的な学習態度を育成していくことが求められている[1]. それは、発生した未知となる課題において、教師や教科書に頼ることのできない状況で適切に情報の収集・取捨選択と学習および実践を行える能力が必要である. このような状況では、最新かつ大量の情報へ容易にアクセスが可能な Web 検索が重要な役割をはたす. 本研究では、上記の能力の育成を目指す支援ツールの提案と実証的検討を行った.

## 2. 方法

### 2.1 関連研究

Marchionini[2]は、学習者が未知の課題に対して、Web 検索を行いながら探索の目標を明確化していく Exploratory Search (以下、探索型検索)を支援する研究の必要性を主張している. 三輪ら[3]は、学生のレポート学習および旅行のプランという課題において探索型検索の行動を観察した結果、探索型検索のスキルは「検索技能」「学習テーマごとのウェブ情報源の知識」、「探索型検索の経験的知識」、「学習テーマの事前知識」などの要因から、学習テーマごとに個別に定義されることを指摘している. また、中島ら[4]は、高度な検索サービスの登場により、学習者の検索技能よりも目的知識に到達できるキーワードが創出できる領域知識の獲得が重要になったと指摘している. すなわち、ある分野の探索型検索による学習の経験が、別の分野の学習では直接的に寄与しづらいため、探索型検索では、学習者の経験によらず分野ごとに支援が必要であることが考えられる.

三輪ら[3]は、探索的検索が充足された場合、通常の参照型検索 (Lookup) に移行すること、参照型検索時に気になるワードを発見した場合に探索的検索に移行することを指摘している. White[5]は、探索的閲覧 (Exploratory Browsing) と絞り込み検索 (Focused Searching)の二つの検索プロセスによる、検索空間の拡大と絞り込みの繰り返しによって探索型検索が行われるモデルを提唱している. また、探索的閲覧では学習者の個人的興味や感情が大きく関係していると指摘している. すなわち、探索的閲覧と絞り込み検索双方に対して別のアプローチによる支援と、探索的閲覧では特に感情に対しての支援も必要であると考えられる.

Web による探索型検索の学習では、膨大な Web コンテンツの中から、学習者自らが判断して教材を選択する必要がある. その際のプロセス的問題として、学習者が教材を決定するには、その Web コンテンツを実際に閲覧し、判断する必要がある. 学習者の知識レベルに合わない難易度のテキストの場合、学習者のモチベーションに影響を与えてしまう. 解決案としては、学習者に合わせてテキストの難易度を推定し推薦する仕組みが考えられる. 河村[6]は、日本語能力試験出題基準とテキストを照合することを検討して

いる. 佐藤[7]は、日本語書き言葉均衡コーパスを用い、対象テキストを難易度順に並べたテキストと照合することで、相対的難易度を 9 段階に測定する仕組みを示した.

### 2.2 研究領域

#### 2.2.1 探索型検索モデル

White[5]による探索型検索モデルより、探索型検索による課題解決のタスクを定義する. ①課題の発見. ②キーワードの発見. ③クエリ創出・検索. ④Web 教材の選択・学習. ⑤学習が充足するまで②へ、充足したら⑥へ. ⑥課題解決方法のためのクエリ創出. ⑦解決方法の絞り込み. ⑧解決法決定まで⑥へ、決定したら⑨へ、方針転換や気になるキーワードがある場合は②へ. ⑨課題解決の実践.

②~⑤が探索的閲覧フェーズで、⑥~⑧が絞り込み検索フェーズである. 図 1 にその構成図を示す.

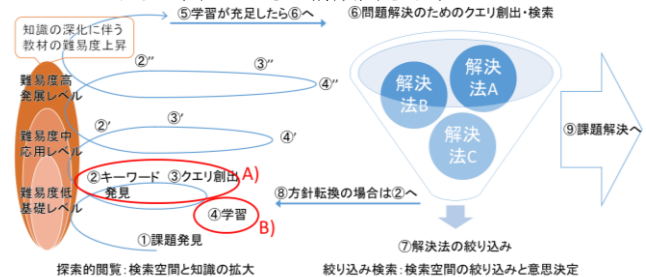


図 1: 探索型検索モデル

#### 2.2.2 研究領域

本研究では、中島ら[4]により、探索型検索における領域知識の獲得の重要性から、②~④の探索的閲覧に着目した. また、White[5]の探索型検索の支援に必要な特徴から、下記についてサポートを行う仕組みを構築、実験した.

- A) 「②キーワード発見」と「③クエリ創出」のタスクにおいて、クエリ作成・修正と洞察・意思決定のための可視化について支援する (図 1 の A)
- B) 「④Web 教材の選択・学習」のタスクにおいて学習・理解の支援について支援する (図 1 の B)

### 2.3 A)キーワード発見・クエリ創出支援

領域知識の獲得の重要性から[4][5]、「③キーワード発見」と「④クエリ創出」を支援する、Wikipedia から単語を抽出しネットワーク図として提示するプロトタイプを作成した. 単語抽出には堀ら[8]の単語抽出アルゴリズムが要件に近いのでこれを用いた. これは、入力したキーワードに対応する Wikipedia 記事内の文章を形態素解析して名詞を抽出し、得た単語と入力されたキーワードとの距離の近く、出現回数が多い程関連性が強いとするものである. 本研究では、関連性を確実に取得する為、取得する単語は文章内のハイパーリンクされた単語に限定した.

#### 2.3.1 A)キーワード発見・クエリ創出支援のプロトタイプ

実験に使用したプロトタイプの利用画面を以下に示す [図 2]. 画面左、放射状に表示されたグラフのうち、オレンジ色のノード (中央下) が入力したキーワード、周辺の繋がっているノードは関連するキーワードを示す. 周辺ノードには番号が振られており、若いほど関連が強い. 一度に表示する関連キーワードは上位 10 件とした. 画面右側で

<sup>†</sup>“A study of learning navigation tool based on exploratory search.”

<sup>†</sup>TAKAHIKO MURAKAMI, <sup>‡</sup>MAKOTO YOKOYAMA,

<sup>‡</sup>TAGIRU NAKAMURA, <sup>‡</sup>NORIYUKI KAMIBAYASHI

<sup>†</sup>Tokyo university of technology Graduate School,

<sup>‡</sup>Tokyo university of technology

関連キーワードの選択と検索を行う。図では例として、「インフォームド・コンセント」と入力し、取得した「パターンナリズム」を選択し関連キーワードを表示し、更に関連するキーワードを10件表示する。



図 2: A)プロトタイプ利用画面

2.3.2 実験方法

28名の大学2年生に、実際に履修中の演習で求められる2つの技能を、Web上の動画を検索し学ぶという課題を課した。プロトタイプは演習用サイトにアップロードした後、入手方法を紹介し、誰でも利用できる状況にした。課題は1)「実験におけるインフォームド・コンセントはどのようにして行いますか?」と2)「感情体験を観察する方法にはどのようなものがありますか?」の2つであった。課題終了後、検索履歴と動画視聴履歴を収集した。プロトタイプの利用有無、所要時間のアンケート調査を行った。なお、プロトタイプを利用するかどうかの判断は教育上の配慮から学生に選択させた。

2.4 B)学習・理解の支援

White[5]は、学習・理解の支援のために「概要を含んだページを返す」「わかりやすいページを返す」「他のユーザの検索履歴を利用する」などを指摘している。このうち「わかりやすいページを返す」という点に着目した。

2.4.1 方法

佐藤[7]のテキスト難易度測定の手法を用い、東京都23区の自治体サイトの出産・育児のテーマの母親向けWebページの難易度を測定した。また、Wikipediaの「インフォームド・コンセント」の記事における、記事のあるキーワード29件について同様に分析し、比較した。測定結果は小学校1年生レベルを難易度1とし、高校3年生レベルを難易度12、それ以降を難易度13としている。

3. 結果

3.1 A)キーワード発見・クエリ創出支援

プロトタイプ使用者(提案手法)は8名、未使用者(対照条件)は20名であった。プロトタイプの使用有無で単位時間あたりの平均獲得異なり検索キーワード数に有意な差が見られた[表1]( $t(26)=0.0168, p<.05$ ,片側検定)。これは、一人ひとりの検索履歴から重複するキーワードを省き、残ったキーワード数の合計を各実験参加者が回答した所要時間で割り、その平均をプロトタイプの利用有無で比較した。なお、提案手法では関連するキーワードを100個以上提案しており、対照条件より考案可能な検索キーワード数が減る可能性は、極めて小さいと推測されるため、片側検定を実施した。図表の数値は小数点第二位を四捨五入した。その他、平均検索回数、平均動画視聴本数、平均所要時間にも違いが見られた[表1]。

表 1: プロトタイプ使用有無による違い

	提案手法	対照条件	割合
平均検索回数	38.6	19.1	約 2.0 倍
平均動画視聴本数	33.5	20.0	約 1.7 倍
平均所要時間	2.3	3.0	約 2 割減
平均獲得異なり検索キーワード数	12.8	5.2	約 2.5 倍

3.2 B)学習・理解の支援

東京都23区中21区について測定することができ、合計4214ページのテキスト難易度を測定した。そのうち測定不能ページが13ページあった。表2に難易度別のページ数を示す。全体での平均難易度は約8.5(中学2年~3年)で全体の91%のページが難易度8~9であった。

表 2: 23区難易度別ページ数

難易度	1	2	3	4	5	6	7	8	9	10	11	12	13
ページ数	0	1	1	6	14	67	118	1557	2292	120	18	4	3

また、Wikipediaの「インフォームド・コンセント」記事のリンクキーワード29件の内訳は表3に示す。全体の平均は約11.8(高校3年生レベル)であった。

表 3: Wikipedia 難易度別ページ数

難易度	9	10	11	12	13
ページ数	1	5	3	9	11

4. 考察

4.1 A キーワード発見・クエリ創出支援

単位時間あたりの平均獲得異なり検索キーワード数に有意な差が見られたことから、プロトタイプの利用が学習者の探索型学習における探索的閲覧フェーズの検索空間と知識の拡大を促した可能性があることが示唆された。

4.2 B 学習・理解の支援

東京都の21区行政ページにおける、出産・育児のテーマの母親向けWebページでは、中学2年生~3年生のテキスト難易度が一般的であることが示唆された。一般的な母親の学歴や初めて出産・育児をすることを考慮して記述することが推奨されていると考えられた。Wikipediaのリンクキーワードの記事においては約69%が高校3年~大学・一般レベルであった。Wikipediaの目的として、オーソライズされた専門性の高い記事が推奨されているからと考えられた。このように、Webサイトの目的や対象ユーザの違いによって、テキスト難易度に一定の偏りがあることが示唆された。したがって、探索検索の最初の段階では専門性の低い、難易度の低いページを教材として推薦し、学習が進むにつれ、深化・専門化された高度なページにシフトして推薦する必要があると考えられた。

4.3 総合考察

本研究では、自らの力で問題解決ができる、自律的・能動的な学習が行える人材の育成を目的とし、探索型検索によるWebを活用した学習の支援について実証的検討を行い、A)探索型学習における探索的閲覧の支援が探索型検索に必要な検索空間と知識の拡大を促す可能性とB)Webページにはサイトごとにテキスト難易度の偏りがあり、学習進度別の教材を推薦する支援の必要性が示唆された。

参考文献

[1] 中央教育審議会 答申 新たな未来を築くための大学教育の質的転換に向けて~生涯学び続け、主体的に考える力を育成する大学へ~, 9p, (2014).  
 [2] Marchionini, G.: EXPLORATORY SEARCH:FROM FINDING TO UNDERSTANDING, COMMUNICATIONS OF THE ACM, Vol. 49, No. 4, pp.41-46 (2006).  
 [3] 三輪 眞木子:Web上のexploratory searchの特徴:発話プロトコルと事後インタビュー分析結果より, 情報処理学会研究報告. 情報学基礎研究会報告 2009-FI-96(2), pp.1-8, (2009).  
 [4] 中島悠:検索経験と領域知識のWWW情報検索行動に与える影響, ヒューマンインターフェース学会, pp.1-10, (2006).  
 [5] White, R. and Roth, R.: Exploratory Search:Beyond the Query-Response Paradigm, Morgan & ClaypoolPublishers (2009).  
 [6] 川村よし子: 語彙チェッカーを用いた読解テキストの分析, 早稲田大学日本語研究教育センター講座日本語教育, 第34分冊, pp. 1-22 (1998).  
 [7] 佐藤理史: 均衡コーパスを規範とするテキスト難易度測定情報処理学会論文誌 Vol. 52 No. 4 pp.1777-1789, (2011).  
 [8] 堀 憲太郎: Wikipediaからの拡張クエリ生成によるWeb検索とその評価,人工知能学会 pp1-8,(2008).