

Web 検索クエリに関する子供向けサブトピックランキング手法の検討

岩田 麻佑[†] 原 隆浩[†] 西尾章治郎[†]

[†] 大阪大学大学院情報科学研究科マルチメディア工学専攻 〒 565-0871 大阪府吹田市山田丘 1-5

E-mail: †{iwata.mayu,hara,nishio}@ist.osaka-u.ac.jp

あらまし 子供が Web 検索を行うことが一般的となってきたが、子供は大人以上に Web 検索を苦手とするため、子供の特徴を考慮した支援が必要である。支援の方法として、クエリに関するサブトピック（検索意図を具体化する文字列）の提示が考えられる。しかし、一般に重要なサブトピックが、子供にとっても重要とは限らない。そこで筆者らの研究グループでは、これまでに子供目線で重要なサブトピックの特徴の調査を行った。本稿では、調査結果に基づき、子供向けサブトピックを優先的に提供するサブトピックランキング手法を提案する。提案手法では、サブトピックの抽出元ソースの種類や難易度などをもとに、サブトピックの子供向け度合いを算出する。評価実験の結果、提案手法の有効性を確認した。

キーワード Web 検索, 子供向け, サブトピック, ランキング

Mayu IWATA[†], Takahiro HARA[†], and Shojiro NISHIO[†]

[†] Department of Multimedia Engineering Graduate School of Information Science and Technology

1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

E-mail: †{iwata.mayu,hara,nishio}@ist.osaka-u.ac.jp

1. はじめに

インターネットの爆発的な普及により、子供がインターネットを利用し、Web 検索を行うことが一般的となってきた。2011 年の統計^(注1)によると、小学生のインターネット利用率は年々増加しており、小学校 6 年生では 9 割近くの子供が日常的にインターネットを利用し、インターネットで Web 検索する際には、学習や趣味に関する検索を行う子供が 5 割以上いることが報告されている。しかし、子供は、クエリの作成や検索結果の選択を困難とし、大人以上に Web 検索を苦手とするため [3]、子供の特徴や要求を考慮した検索方法や表示方法などの Web 検索の支援が必要であると考えられる。

Web 検索支援の有効な方法として、クエリに関するサブトピックを利用する研究が盛んに行われている [4]。クエリに関するサブトピックとは、入力されたクエリに対して“ユーザの意図を具体化する文字列”である。例えば、“ハリー・ポッター”というクエリについては、“映画”、“本”などがサブトピックの候補として考えられる。これらのサブトピックを利用したクエリサジェスション、検索結果のランキングの多様化などにより、ユーザの Web 検索を支援可能となる。子供にとっても、サブトピックを利用した支援は有効であると考えられるが、既存研

究のサブトピック抽出方法 [9] [4] において、検索エンジンの検索結果やクエリログから取得できる一般的に重要なサブトピックが、子供にとっても重要であるとは限らない。

子供にとっては、興味を持って楽しむことが学習に繋がり、重要であるという指摘があるように [1]、子供の Web 検索において、子供にとって興味をひくサブトピックを提示することが重要であると考えられる。さらに、子供にとって重要なサブトピックは、子供の検索目的を考慮すると、“趣味”や“学習”に役立つものであると考えられる^(注1)。例えば、クエリ“沖縄”については、“ツアー”、“レンタカー”というような旅行に関するサブトピックが一般的に重要と考えられるが、それらは子供にとって興味をひく内容ではない。一方、“食べ物”、“歴史”というようなサブトピックは、子供が授業で沖縄の文化や習慣について学んでいる際に興味をひく内容であり、重要であると考えられる。子供の興味をひくサブトピックを提示することで、そのサブトピックについて調べてみたいという子供の欲求を刺激し、Web を活用できるきっかけになると考えられる。もちろん、子供の興味をひくサブトピックが必ずしも Web 検索の際に役に立つとは限らないが、まず興味をひくものでなければ、中身を調べようという意欲が沸きにくい。つまり、Web 検索の際に子供にとって重要なサブトピックは、子供にとって興味をひくものである上で、Web 検索に実際に役に立つものである必要がある。そのため、子供の Web 検索支援のために重要なサ

(注1) : goo リサーチ: <http://research.goo.ne.jp/>

ブトピックを特定するには、子供目線での重要度と Web 検索時に役立つ重要度の両方の面で定量化しなければならない。

本研究では、学校の宿題などで Web 検索を行う機会が多いと考えられる小中学生を対象とし、まず、子供目線で興味をひくかどうかという観点でサブトピックの重要度を定量化することに着目する。子供の興味をひくという点では、まず内容を想像できないサブトピックは興味を持ちにくいと考えられるため、子供にとって分かりやすいものかどうかという点が一つの重要な要素となり得る。また、子供の好奇心を刺激するかどうかという点も重要な要素となり得る。さらに、内容について知ることが役立つと感じるかどうか、サブトピックへの興味を左右する重要な要素と考えられる。そこで本研究では、子供にとって分かりやすい (familiar)、調べてみたい (interesting)、役に立ちそう (useful) であるサブトピックを子供向けサブトピックとする。

以上のような子供の興味を引くサブトピックの特徴を明らかにするため、筆者らの研究グループでは、これまでに子供向けサブトピックの調査 [11] を行った。具体的には、一般向け検索エンジン、子供向け検索エンジン、Wikipedia という複数のソースから取得したサブトピックを、familiar, interesting, useful の 3 つの観点から子供にスコア付けしてもらうことで、子供向けサブトピックの特徴を調査した。調査の結果、以下のような点が子供向けサブトピックの重要度の定量化に重要であることが明らかになった。

- 子供向けサブトピックを簡単に取得できるのは検索エンジンの検索結果から得られる関連検索語であるが、重要なサブトピックを広くカバーするには、Wikipedia などの他のソースも効果的である。
- サブトピックの抽象度、難易度、検索意図、検索ヒット数といった要素を考慮することが、子供向けサブトピックを特定するための手がかりとなる。

本稿では、子供にとって重要なサブトピックを優先的に提供するため、文献 [11] の調査結果に基づき、子供向けサブトピックランキング手法の検討を行う。この手法では、調査結果に基づいて決定した、取得ソースの種類や難易度などのサブトピックの要素に関する素性をもとに、サブトピックの子供向け度合いを数値化し、ランキングを行う。これにより、子供が検索クエリを入力した際に、子供向けサブトピックをクエリサジェスションとして提示したり、子供向けサブトピックに基づいて検索結果のクラスタリングを行うことで、子供は自分の検索要求を具体的なクエリとして入力する手間なしに、必要とするサブトピックを容易に得ることができる。

以下では、まず 2 章で関連研究について述べ、3 章で子供向けサブトピックの調査について述べる。その後、4 節で子供向けサブトピックランキング手法について述べ、最後に 5 章で本稿のまとめを行う。

2. 関連研究

近年、子供の Web 検索プロセスを調査する研究が盛んに行われている。Bilal ら [3] は、大学院生と子供に、あるコンテン

ツを Web 検索により探すタスクを行ってもらい、大人と子供の間のタスクの成功率や検索行動の違いを調査している。また、Bilal [2] は、子供に、子供向け検索エンジンでタスクを行ってもらい、認識面、身体面、感情面での特徴についても調査している。調査の結果、子供には、検索失敗時の修正方法が分からない、タスク中に飽きてしまうというような特徴があり、Web 検索が苦手であることが示されている。Druin ら [5] は、子供に自宅で自由に Web 検索を行ってもらい調査により、スペリング、タイピング、クエリ作成、検索結果の解釈が子供にとって問題となることを明らかにしている。

これらの調査結果を考慮し、子供の Web 検索を支援するために、数多くの研究が行われている。文献 [8] [10] では、子供向けの検索結果のランキングを行うことで、子供の Web 検索支援を行っている。文献 [10] では、Web 検索時の子供の特徴を考慮して、Web ページ中の画像量や文章の難易度などに基づき Web ページをランキングする。文献 [8] では、子供向けサイトとのリンクの繋がりに基づいて Web ページの子供向け度合いを算出し、Web ページのランキングを決定する。さらに文献 [7] では、検索結果の表示方法に関する支援が行われている。具体的には、クエリログを利用して、クエリに関して適切なメディアタイプを決定し、子供向けのマルチメディア検索を実現するシステムについて述べている。また文献 [12] では、幼小者の生活様式オントロジーを構築し、子供の生活環境に密着した Web 検索を可能とするシステムを提案している。以上のように、様々な方法で子供の Web 検索支援を行う研究が行われているが、どのようなサブトピックを子供が重要視するのかは明らかになっておらず、子供向けサブトピックを優先的に提供する研究は行われていない。

一方、一般的な Web 検索を支援する方法として、サブトピックを利用する研究が多数行われている。これらの研究では、Web 検索エンジンから返される検索結果、外部ソース、クエリやクリックスルーのログなどを利用して、サブトピックを抽出する。文献 [4] では、クエリログ、アンカーテキスト、検索結果中のキーワードをもとにサブトピックを抽出し、検索結果の多様化を行っている。文献 [9] では、クエリを Wikipedia の概念にマッピングして利用することで、クエリの意図を推定し、カテゴリズを行っている。これらの手法で取得できるサブトピックは子供向けとは限らないため、子供にとって重要なサブトピックについて調査する必要がある。

3. 子供向けサブトピックの調査

本章では、筆者らが文献 [11] において実施した子供向けサブトピックの調査結果を紹介する。文献 [11] での調査に加え、新たなサブトピックの特徴の一つとして、検索エンジンにおけるヒット数と子供向けサブトピックの関係性についても調査した。

3.1 サブトピックのスコア付け

子供向けサブトピックの調査を行うため、クエリとサブトピックから成るデータセットを作成し、データセット中のサブトピックを子供に評価してもらうことで、サブトピックのスコアを決定した。

3.1.1 データセット

子供が検索する可能性の高いクエリで調査を行うため、Yahoo!きっずの2009年12月から2011年10月の週間検索キーワードランキング上位10位より選んだ“地球温暖化”，“聖徳太子”，“金環日食”，“沖縄”，“恐竜”，“トマト”，“ポケモン”，“ハリー・ポッター”の8個のクエリを想定した。

クエリには、学習を主な目的とするもの、趣味を主な目的とするもの、学習と趣味のどちらも目的とし得るものが存在し、何を目的とするかによって重要なサブトピックが異なると考えられる。そのため、“地球温暖化”，“聖徳太子”，“金環日食”については学習のみ(3ペア)，“沖縄”，“恐竜”，“トマト”については学習と趣味の2種類(6ペア)，“ポケモン”，“ハリー・ポッター”については趣味のみ(2ペア)の検索目的を想定した合計11種類の検索目的とクエリのペアについてサブトピックを抽出し、データセットとした。

サブトピックを抽出するには、検索エンジンの検索結果、Wikipediaなどの外部ソース、クエリやクリックスルーのログを利用する方法が代表的である。そこで、サブトピックの抽出には、子供がWeb検索の際に利用する一般向け検索エンジンや子供向け検索エンジンにおける関連検索語や検索結果、および、代表的な外部リソースとしてWikipediaを利用した。具体的には、以下の6種類のソースからサブトピックを抽出した。

- genRel : Yahoo! JAPANにおけるクエリの関連検索語
- genKey : Yahoo! JAPANにおけるクエリの検索結果上位100件のスニペットに出現するキーワード
- kidRel : Yahoo!きっずにおけるクエリの関連検索語
- kidKey : Yahoo!きっずにおけるクエリの検索結果上位100件のスニペットに出現するキーワード
- kidCat : Yahoo!きっずにおけるクエリの検索結果上位100件の登録カテゴリ
- wiki : クエリをタイトルとするWikipediaの記事内の項目タイトル

キーワードについてはキーワード抽出API^(注2)で取得した。このAPIを用いることで、検索結果のスニペットから特徴的な表現とその重要度を取得できる。

関連検索語から抽出できるサブトピックは、他のソースと比較すると少数であるが(Yahoo! JAPANで最大10個、Yahoo!きっずで最大3個)、ユーザが頻繁にクエリと一緒に検索する語であるため、重要性が高いと考えられる。

一方、検索結果中のキーワードからは、非常に多数のサブトピックを抽出することが可能であるが、ノイズも多く含まれると考えられる。Yahoo! JAPAN, Yahoo!きっずの両方を利用したのは、Yahoo! JAPANとYahoo!きっずでは検索結果に含まれるページが異なるためである。Yahoo!きっずの検索では、クエリに対応するおすすめサイトがあれば、そのサイトを検索結果上位に表示し、その後、Yahoo! JAPANと同様のページ検索結果のランキングを表示する。したがって、あらかじめ登

図1 サブトピック評価システム

録されていないページは表示されにくいいため、Wikipediaなどの子供向けでないと考えられるページは少なくなる。

また、Yahoo!きっずのおすすめサイトにはそれぞれカテゴリが登録されており、検索結果中に表示される。カテゴリはクエリの大まかな意図を示すと考えられるため、サブトピックを取得する一つのソースとして用いた。カテゴリは具体的ではないため、意図を大幅に絞ることはできないが、子供にとって知っている語が多いと考えられる。

Wikipediaについては、記事タイトルをクエリと見なしたとき、記事全体のキーワードは子供にとって難解である可能性が高いが、項目タイトルはそこまで難解でなく、クエリの内容を説明するサブトピックになり得ると考えられる。

複数ソースから取得したサブトピック中からクエリに該当する文字列は除外して統合し(例：“沖縄旅行”と“旅行”は“旅行”に統合)、複数のソースから同一のサブトピックを取得した場合も1つのサブトピックとして扱った。そして、Yahoo!キーワードAPIで返ってきた重要度のスコアが低いサブトピックは除外し、各クエリのサブトピックを100個とした。

3.1.2 手順

評価実験のために、Web上で簡潔な操作で評価を行えるシステム(図1)を構築した。被験者は、小1から中3の複数の学校に所属する子供49名とし、データセット内のサブトピックについて、3つの質問により評価を行ってもらった。被験者の内訳は小学校低学年が1名(男1名)、小学校中学年が11名(男5名、女6名)、小学校高学年が25名(男19名、女6名)、中学生が12名(男7名、女5名)であった。3つの質問で評価を行ったのは、内容を知っているのかどうか、検索する意欲を促進するかどうか、検索時に役立ちそうかどうかという多面的な要素で、子供の興味をひき、重要なサブトピックがどのようなものなのかを調べるためである。以下に実験の手順を示す。

(1) ランダムに表示されたサブトピック s について、以下の質問に‘はい’、‘どちらでもない’、‘いいえ’で回答してもらう。

- q に関する s について少しでも何か知っていますか?(familiar)
- *situation* のために、 q に関する s についてインター

(注2) : Yahoo!デベロッパー キーワード抽出 API:
<http://developer.yahoo.co.jp/webapi/jlp/keyphrase/v1/extract.html>

ネットで検索してみたいですか?(interesting)

- *situation* のために, *q* に関係する *s* は役に立ちそうですか?(useful)

(2) 各クエリで (2) を複数サブトピックで繰り返す.

(3) (1) から (3) を複数クエリで繰り返す.

検索目的 *situation* は, 学習を目的とする“学校の宿題”, 趣味を目的とする“友達と遊ぶ”などのあらかじめ定義した状況を利用する.

最初に, 被験者に大学に集まってもらい, 1 日で 100 個のサブトピックを評価してもらった. その後, 協力してもらえぬ被験者については, 自宅からも自由にサブトピックの評価を行ってもらった. その結果, 各サブトピックに対して, 少なくとも 4 人分の評価を取得した. 被験者間のサブトピックに対する評価の一致率を示す Fleiss's Kappa は, familiar で 0.234, interesting で 0.208, useful で 0.206 であり, サブトピックに対する評価は年齢による差が大きいと考えられるため, 一致率は低かった.

その後, 3 つの質問について評価してもらった結果に基づき, 1,100 の評価用サブトピックに対して, サブトピックごとのスコアを決定した. 具体的には, 評価の‘はい’を 1 点, ‘どちらでもない’を 0.5 点, ‘いいえ’を 0 点とし, 質問ごとに各サブトピックの合計点数を算出し, 評価した被験者数で平均した値をそのサブトピックのスコアとした. スコアが高いほど, 子供にとって高評価なサブトピックであることを示す.

3.2 調査結果

サブトピックのスコアの傾向を調査することで, 子供の興味をひくサブトピックを優先的に提供するために必要な要素について検討した. 具体的には, まず, データセットとして用意したクエリやサブトピックの抽出元ソースの種類が, 抽出できるサブトピックの種類に影響し, 子供によるサブトピックのスコアに影響すると考えられるため, クエリやソースの種類とサブトピックのスコアの関係について調査した (調査項目 1). これにより, 検索目的の違いによるサブトピックのスコアの違いや子供向けサブトピックの抽出に重要なソースを明らかにした. また, クエリやソースの種類を考慮しただけでは, 子供向けサブトピックとそれ以外を区別するためのサブトピックの定量化が難しいと考えられるため, 子供が重要視すると予想されるサブトピック自身の抽象度や検索意図などの特徴とサブトピックのスコアの関係について調査した (調査項目 2).

このような調査の結果, 子供向けサブトピックの抽出に重要であると考えられる以下の点が明らかになった.

(1) 学習を目的とするか趣味を目的とするかで重要なサブトピックが変化する.

趣味に関するクエリでは嗜好の影響が強く出る一方, 学習に関するクエリは, 子供にとって一見親しみにくくとも, 役に立ちそうなサブトピックが多数ある (3.2.1 項). また, 学習に関するクエリでは, 授業で学ぶようなサブテーマを示すサブトピック, 趣味に関するクエリでは, 動画像に関するサブトピックが高スコアであり, 重要な検索意図は異なる (3.2.2 項).

(2) 幅広く重要な意図をカバーするためには, 関連検索語

表 1 クエリの各検索目的の平均スコア

	学習	趣味
familiar	0.376	0.401
interesting	0.538	0.493
useful	0.541	0.586

表 2 各ソースの平均スコア

	wiki	genRel	genKey	kidRel	kidKey	kidCat
familiar	0.371	0.506	0.407	0.684	0.397	0.427
interesting	0.531	0.562	0.508	0.615	0.519	0.515
useful	0.537	0.565	0.513	0.668	0.518	0.515

表 3 各ソースから抽出したサブトピックの個数

	wiki	genRel	genKey	kidRel	kidKey	kidCat
274	106	378	28	507	190	

を中心に複数のソースを利用すべきである.

検索エンジンの関連検索語から抽出するサブトピックは高スコアであるが, 数が少なく, カバーできる意図が少ない (3.2.1 項, 3.2.2 項). 検索結果中のキーフレーズや Wikipedia などのソースを利用すると, 関連検索語だけでは抽出できないサブトピックも抽出することができる (3.2.2 項).

(3) サブトピックの抽象度, 難易度, 意図, 検索ヒット数といった要素は子供向けサブトピックを特定するための手がかりとなる (3.2.2 項).

複数の要素を組み合わせることで, 子供向けサブトピックの定量化を効果的に行えると考えられる.

以下では, これらの知見の詳細について順に述べる. なお, 上記において, 括弧内が各知見に関連する結果について述べる調査項目に対応している.

3.2.1 クエリやソースとスコアの関係

クエリとソースによるサブトピックのスコアの違いについての調査結果を述べる. クエリの各検索目的でのスコアの平均値を表 1 に, 各ソースでのスコアの平均値を表 2 に, 各ソースで抽出したサブトピックの個数を表 3 に示す. サブトピックのスコアに対して, 0.05 を有意水準とし, クエリとソースの 2 要因の分散分析を行ったところ, familiar についてはクエリ, ソース, interesting についてはクエリ, useful についてはクエリ, ソースの違いにより, 有意にスコアが変化していることが分かった. この結果より, クエリやソースの違いが子供にとってのサブトピックの評価と相関していることが分かる. さらに, 分散分析の後, 有意な要因について Bonferroni 法 [6] で多重比較を行い, 0.05 の有意水準で有意に平均値が異なる組合せを調査した.

クエリごとの違いについては, 多重比較の結果, 趣味に関するクエリと学習に関するクエリとの間で, スコアの平均値が有意に異なることが分かった. 具体的には, familiar については, 学校で学んだ内容かどうか依存する学習に関するクエリのスコアが, 趣味に関するクエリのスコアよりも低かった. interesting や useful については, 嗜好の影響が強く出るため, 興味のないクエリに関するサブトピックのスコアを低く評価する子供が多かった. そのため, 趣味に関するクエリのスコアは, 学習に関するクエリのスコアを下回った. この結果より, 学習を目的とした際には, 子供にとって, 親しみが薄くとも, 調べてみたい, 役に立ちそうと感じるサブトピックが多く存在すると考えられる.

ソースごとの違いについては、多重比較の結果、familiar については genRel と kidRel, useful については kidRel のみが他のソースのスコアの平均値を有意に上回った。これは、検索エンジンの関連検索語は実際にユーザが頻繁に検索しているキーワードの組合せであり、子供にとっても分かりやすいサブトピックが多く含まれるためと考えられる。genRel と kidRel を比較すると、Yahoo! JAPAN は一般向け検索エンジンであり、genRel の中には“沖縄ツーリスト”など子供にとって役に立たないと感じるサブトピックが多くあるため、genRel のスコアは kidRel のスコアを下回った。このように関連検索語を取得できる際には、積極的に子供向けサブトピックとして利用することが効果的であると考えられる。ただし、関連検索語から取得できるサブトピックは非常に少ないため、幅広いサブトピックを取得するためには、他のソースも利用すべきと考えられる。

他のソースの傾向としては、wiki では、難解な表現ではあるが、学習に役立つようなサブトピックが多いと考えられるため、familiar は低いが、interesting や useful は比較的高かった。つまり、interesting や useful を重視する際には、Wikipedia などのやや難解ではあるが整理された外部情報を利用することが効果的であると考えられる。genKey や kidKey からは、数多くのサブトピックを抽出できるが、サイト名や団体名などの難解な固有名詞を多く含むと考えられるため、スコアは高くなかった。そのため、多様なサブトピックを抽出するために検索エンジンを利用する際には、子供にとって不要なサブトピックを排除する必要がある。kidCat は、固有名詞や難解な表現は少ないと考えられるため、familiar は wiki や kidKey よりもやや高かった一方、抽象度の高いカテゴリは曖昧であり、interesting や useful というスコアは高くなかった。そのため、カテゴリについても、抽象度が高く役に立たないような不要なサブトピックを排除する必要がある。

以上の結果より、子供にとって重要なサブトピックを幅広く抽出するには、子供向け関連検索語を中心に、その他のソースも利用してサブトピック候補を抽出した後、何らかの特徴を考慮して、子供向けのサブトピックとしての重要性を決定する必要があると考えられる。

3.2.2 サブトピックの要素とスコアの関係

子供向けサブトピックとしての重要性を定量化するために、サブトピックのスコアに影響を与えると考えられるサブトピックの要素について検討する。

子供は語彙が少ないため、子供向け文章では漢字やアルファベットの割合や文字数が少ない。そのため、漢字数やアルファベット数が子供によるサブトピックの評価に影響すると考えられる。また、サブトピックの抽象度もサブトピックの評価に影響すると考えられる。つまり、具体的なサブトピックには難解なものが多くなるので、評価が低くなる可能性がある。

さらに、サブトピックが示す検索意図が重要となる。一般的に、検索を行う際には、どのような情報要求を達成したいかという何らかの検索意図が存在し、その検索意図に対応する有効なサブトピックが存在すると考えられる。つまり、提示するサブトピックは検索意図と対応すべきである。これをサブトピッ

表 4 サブトピックの要素とスコアの相関係数

	intent	difficulty	abstraction	kidHit	genHit
familiar	0.447	-0.284	0.294	0.286	0.252
interesting	0.370	0.005	0.133	0.171	0.170
useful	0.380	-0.019	0.197	0.231	0.202

クの示す検索意図とする。例えば、“沖縄”というクエリについては、“レンタカー”、“ホテル”というサブトピックは旅行プランに関する検索意図、“戦争”、“年表”というサブトピックは歴史に関する検索意図を示すと考えられる。サブトピック自体が分かりやすくとも、サブトピックの示す検索意図と子供にとっての情報要求が食い違えば、重要なサブトピックとはならないため、サブトピックの示す検索意図がどのようなものであるのかも重要であると考えられる。

さらに、サブトピックを示すキーワードを用いて、検索エンジンで検索した際に返ってくる検索ヒット数も、子供にとって重要なサブトピックなのかどうかを示す一つの要素になると考えられる。検索ヒット数の極端に少ないようなサブトピックは、一般的に利用されにくい不要なサブトピックである可能性がある。

以上の議論に基づいて、子供のサブトピックの評価に影響する要素として、本研究では以下の4点について調査する。

- difficulty: サブトピックの難易度。漢字とアルファベットの文字数。
- abstraction: サブトピックの抽象度 (1: 低~5: 高)。
- intent: サブトピックの示す検索意図。
- genHit (kidHit): Yahoo! JAPAN (Yahoo!きっず) においてサブトピックで検索して返ってきたページ数。

abstraction と intent については、3名の評価者で全サブトピックについてアノテーションを行った。具体的には、abstraction については3名の評価者が、5段階評価で各サブトピックの抽象度を付加し、3名の評価者に付加された抽象度の平均値をそのサブトピックの abstraction とした。intent については、3名の評価者が、各クエリのサブトピック 100 個を 10 個から 15 個の意図になるように分類し、3名の評価者に付加された意図のうち、大多数を占める意図をそのサブトピックの intent とした。3名の評価者の一致度を示す Fleiss's Kappa は、抽象度で 0.701, intent で 0.672 あり、中から高程度一致していた。

各要素とサブトピックのスコアの関係を調べるために、各要素の値と familiar, interesting, useful の相関係数を求めた (表 4)。difficulty は familiar と弱い負の相関、intent は familiar と弱い正の相関、genHit と kidHit は familiar と useful との間に弱い正の相関、intent については familiar, interesting, useful と弱から中程度の正の相関が見られた。これらの相関は単独では十分に大きいものであるとはいえないが、これらの要素を組み合わせることで、子供向けサブトピックの定量化の際に有効な素性になり得ると考えられる。特に意図については考慮する必要があると考えられる。

そこで、具体的に、どのような意図が子供にとって高評価なのか、3種類のクエリの例を表 5 に示す。各クエリで、高 (低) スコアな意図 3 つとそれぞれの意図の中で最高スコアのサブトピック、その抽出元のソースを括弧内に示す。

表 5 高スコア, 低スコアな意図とサブトピック例

検索目的	クエリ	スコア	意図の例 (サブトピック例: 取得元ソース)
学習	地球温暖化	高	現状 (現状: genRel, genKey), 乗り物 (乗りもの: kidCat), 影響 (気候変動: wiki, genKey)
		低	動画像 (アニメーション: kidKey), リンク集 (リンク集: kidCat), 議論 (議論: wiki)
学習	沖縄	高	歴史 (米軍基地: wiki), 地名 (那覇市: genKey), クイズ (クイズ: kidKey)
		低	レストラン (居酒屋: genKey), サイト (総合情報ポータルサイト: genKey), ニュース (ニュース: kidCat)
趣味	沖縄	高	スポット (美ら海水族館: genKey, kidKey), 気候 (気候: wiki, kidRel), 動画像 (地図: genKey)
		低	クイズ, データ (人口: wiki), 豆知識 (県名の由来: wiki)
趣味	ポケモン	高	カード (カードゲーム: genKey, kidKey), 動画像 (イラスト: genKey), 団体 (任天堂: kidKey)
		低	豆知識 (開発: wiki), スポット (ナガシマスパーランド: genKey, kidKey), 地名 (イスラム諸国: wiki)

学習に関するクエリに対しては, “動画像”, “リンク集” という情報の形式に関する意図が低スコアである一方, “現状”, “影響” というクエリに関するサブテーマを示すような一般名詞のサブトピックが高スコアであった. それに対し, 趣味に関するクエリに関しては, “豆知識”, “クイズ” などの知識に関する意図が低スコアである一方, 学習に関するクエリでは低スコアであった “動画像” が高スコアであった. この結果より, 検索目的が学習か趣味であるかによって重要な検索意図が変化すると考えられる.

また, 表 5 の取得元ソースに注目すると, 高スコアな意図を表現するサブトピックは幅広いソースから取得されたものである. この結果より, 子供にとって重要な意図を示すサブトピックを幅広く抽出するには, 特定のソースを利用するだけでは不十分であることが分かる.

4. 子供向けサブトピックランキング手法

上記の調査結果に基づき, 子供向けサブトピックランキング手法について検討する. この手法を用いることで, 多くのサブトピック候補の中から子供にとって familiar, interesting, useful とした観点で重要度の高いサブトピックを優先的に提示することが可能となる.

4.1 手法の設計

子供向けサブトピックランキング手法では, 複数のソースからサブトピック候補を抽出し, 抽出ソースの種類に加えて, クエリの検索目的, サブトピックの要素を考慮して, サブトピックのスコアを算出し, ランキングを行う.

4.1.1 素性

具体的には, 実験で得られた知見に基づき, 以下に示す素性のスコアを算出する. 括弧内は, 3.2 節に示すどの知見に対応する素性なのかを示している.

- クエリに関する素性 (知見 1)
 - クエリの検索目的: 学習もしくは趣味.
- 抽出ソースに関する素性 (知見 2)
 - genRel: Yahoo! JAPAN の関連検索語かどうか.
 - genKey: Yahoo! JAPAN の検索結果のスニペットに出現するキーワードかどうか.
 - kidRel: Yahoo!きっずの関連検索語かどうか.
 - kidKey: Yahoo!きっずの検索結果のスニペットに出現するキーワードかどうか.
 - kidCat: Yahoo!きっずの検索結果中の登録カテゴリかどうか.
 - wiki: Wikipedia の項目タイトルかどうか.
- サブトピックの要素に関する素性 (知見 3)

- difficulty: 難易度. 漢字やアルファベットの文字数.
- abstraction: 抽象度. 一般名詞か固有名詞か.
- intent: 意図. あらかじめ定義した教科カテゴリと属性.
- hit: Yahoo! JAPAN と Yahoo!きっずでの検索ヒット数.

サブトピックの要素に関する素性については, どのように各素性をスコア化するかを検討しなければならないが, 本研究では, 単純な方法でサブトピックの各要素を素性として用いたときに, どの程度効果があるのかを確認することに重点を置く.

クエリに関する素性については, 学習か趣味をクエリの検索目的として用いる. これは, 3.2 節で述べたように, 学習を目的とするか趣味を目的とするかで重要なサブトピックが変化するためである, 自動でクエリの検索目的を決定するには, 各目的において頻繁に検索される語を定義した辞書を利用する方法などが考えられるが, 今回は, データセット中で, クエリごとにあらかじめ定義された目的をそのまま利用した.

抽出ソースに関する素性については, サブトピックを抽出したソースの種類, つまり, それぞれのソースがサブトピックの抽出元であるかどうかを用いる. これは, 3.2 節で述べたように, 幅広く重要な意図をカバーするためには, 関連検索語を中心に複数のソースを利用すべきであるためである. ただし, 検索結果のスニペット中のキーワードである genKey, kidKey については, ソースにおけるそのサブトピックの重要度として, Yahoo!キーワード API で取得したキーワードの重要度を用いた.

サブトピックの要素に関する素性については, 3.2 節で述べたように, サブトピックの抽象度, 難易度, 意図, 検索ヒット数といった要素が子供向けサブトピックを特定するための手がかりとなることを考慮し, 以下のように決定する. difficulty については, サブトピックの文字列中に占める漢字とアルファベットの数を利用した. abstraction については, 3.2 節で 3 人の評価者が判定した抽象度に基づいて, あらかじめ定義した一般名詞か固有名詞かを用いた. 形態素解析ツールの品詞判定などを用いることで, 自動で一般名詞か固有名詞かを判定することも可能と考えられる. intent については, 自動で決定することは難しいため, 調査の結果, 子供にとってのサブトピックの重要度に影響を与え得ることが判明した「教科カテゴリ」と「属性」の 2 種類を素性として用いた. 教科カテゴリは, 教科ごとにあらかじめ作成した辞書に基づいて決定した. 具体的には, Yahoo!きっずのカテゴリである科学, 数学, 社会, 生活, 言語, 総合学習を用い, Yahoo!きっずの各カテゴリの登録サイトに出現する重要語を抽出して, 各カテゴリの辞書を作成した. そして, サブトピックがどのカテゴリの辞書に登録された重要語を含むかで該当カテゴリを決定した. どのカテゴリの辞書に

も含まれない場合は、その他のカテゴリとした。属性については、3.2節で3人の評価者が分類した組織、人名、地域、サイト、情報タイプから選択して用いた。形態素解析ツールの品詞判定などを用いて、これらの属性を自動で判定することも可能と考えられる。hitについては、Yahoo! JAPAN, Yahoo!きっずの2種類の検索エンジンにおいて、サブトピックを検索クエリとした際に返ってくるそれぞれの検索結果ページ数を用いた。

4.1.2 手順

前項で説明した素性に基づき、ユーザがクエリを発行すると、以下の手順で、子供向けにサブトピックのランキングを行う。これにより、ランキング上位ほど子供向けのサブトピックとなり、子供がクエリを入力してWeb検索を行う際に、容易に子供向けサブトピックを取得し、それらに対応する内容について調べることが可能になると考えられる。

- (1) ユーザがクエリを入力する。
- (2) クエリに対するサブトピックを以下に示す複数のソースから抽出する。

- Yahoo! JAPAN におけるクエリの関連検索語
- Yahoo! JAPAN におけるクエリの検索結果上位 100 件のスニペットに出現するキーワード
- Yahoo!きっずにおけるクエリの関連検索語
- Yahoo!きっずにおけるクエリの検索結果上位 100 件のスニペットに出現するキーワード
- Yahoo!きっずにおけるクエリの検索結果上位 100 件の登録カテゴリ
- クエリをタイトルとする Wikipedia の記事内の項目タイトル

(3) 4.1.1 項で述べた複数の素性に基づき、抽出した各サブトピックのスコアを算出する。具体的には、0 ~ 1 に正規化した各素性のスコアを Ranking SVM を用いて組み合わせ、各サブトピックのスコアとする。

- (4) スコアの降順にランキングする。

4.2 評価

4.2.1 データセット

データセットは 4.3.1 項で述べた子供 49 人に評価してもらった 11 種類のクエリに関するサブトピック 1,100 個とした。familiar, interesting, useful のそれぞれの観点で子供に評価してもらったスコアの降順に並べたサブトピックのランキングを、それぞれの観点における正解ランキングとした。

4.2.2 評価指標

ランキングの精度を測る代表的な指標として、MAP 値と NDCG 値を用いた。MAP (Mean Average Precision) は高スコアなサブトピックを網羅している割合を示す素性で、式 (1) より求める。

$$AP_q = \frac{1}{k} \left(\sum_k rel_i \right), MAP = \frac{1}{n} \left(\sum_n AP_i \right) \quad (1)$$

rel_i はランキング i 番目のサブトピックのデータセットで定義されたスコア、 AP_q はクエリ q に関する Average Precision (式 (4.1) より求める)、 k はクエリ数、 n は全サブトピック数である。

一方、クエリ q に対する NDCG は、スコアの高い順に並べた理想的な結果とのずれを表す指標で、式 (2) より求める。

$$NDCG_q = \frac{1}{IDCG_q} \left(rel_1 + \sum_{i=2}^l \frac{rel_i}{\log_2 i} \right) \quad (2)$$

rel_i はランキング i 番目のサブトピックのデータセットで定義されたスコア、 l はランキングの検索数であり、上位 1 件から 20 件で評価を行った。 $IDCG_q$ はクエリ q に対する $NDCG_q$ の理想値、つまりデータセットで定義された理想ランキングの $NDCG$ 値である。

4.2.3 評価結果

評価では、ソースを組み合わせる効果、および、サブトピックの要素を組み合わせる効果を調査するため、ソース単独を用いたもの (genRel, genKey, kidRel, kidKey, kidCat, wiki), ソース全てを用いたもの (Source), 全ての素性を用いたもの (All) の 8 種類のケース (手法) でランキングした結果を比較した。

素性の組合せは Ranking SVM を用いて決定した。具体的には、 SVM^{rank} (注3) を用いて実装を行い、カーネルは線形カーネルを用いた。データセット中の 1 種類のクエリをテストデータ、残りの 10 種類のクエリをトレーニングデータとし、Ranking SVM を用いてランキング関数を学習した。学習したランキング関数によって、テストデータのサブトピックのランキングを作成し、NDCG 値と AP 値を算出した。全てのクエリを順にテストデータとし、最終的に、全 11 種類のクエリの平均 NDCG 値と MAP 値を算出した。

各手法における MAP 値を表 6 に示す。有意水準 0.05 で Bonferroni 法で多重比較を行ったところ、MAP 値については、全ての質問に対して、ソース単独よりも Source と All が有意に高精度であった。具体的には、All については、familiar で 12.2% から 40.8%, interesting で 43.2% から 49.8%, useful で 21.6% から 50.8%, ソース単独よりも MAP 値が向上した。また、Source については、familiar で 8.9% から 37.5%, interesting で 41.9% から 49.5%, useful で 19.6% から 49.0%, ソース単独よりも MAP 値が向上した。これは、ソース単独では取得できるサブトピックに限界があるためである。Source と All の MAP 値については有意差は見られなかったが、All の方が安定して高精度であった。具体的には、All の方が平均としては、familiar で 3.3%, interesting で 1.3%, useful で 1.8%, Source よりも MAP 値が向上した。

各手法における NDCG 値を図 2 に示す。横軸は、式 (2) における上位何番目までのサブトピックを検索するかという検索数 l を示す。NDCG 値については、単独ソースの結果では、familiar や useful で、ランキング上位での genRel や kidRel は比較的高精度であり、familiar のみでは genKey も比較的高精度であった。しかし、ランキング 10 位を超えると途端に精度が下がったり、質問ごとに偏りがあつたりするため、単独ソースでの精度は不安定であるといえる。それに対して、Source や All は安定して高精度であり、さらに All は Source よりも高精

(注3) : SVM^{rank} : http://www.cs.cornell.edu/people/tj/svm_light

表 6 素性の各組合せによるランキングの MAP 値

	wiki	genRel	genKey	kidRel	kidKey	kidCat	Source	All
familiar	0.116	0.300	0.033	0.091	0.116	0.014	0.389	0.422
interesting	0.033	0.035	0.016	0.090	0.026	0.092	0.511	0.524
useful	0.176	0.314	0.043	0.090	0.106	0.022	0.512	0.530

図 2 素性の各組合せによるランキングの NDCG 値

度である。有意水準 0.05 で Bonferroni 法で多重比較を行ったところ、全ての質問で、ほぼ全ての点に対して、All の NDCG 値は単独ソースの NDCG 値だけでなく、Source の NDCG 値を有意に上回っていた。具体的には、All は、familiar で 0% から 12.4%，interesting で 3.8% から 8.2%，useful で 7.3% から 13.6%，Source の NDCG 値を上回っていた。これらの結果より、単独ソースよりもソースを複数用いる方が幅広く安定して高精度にサブトピックを取得できること、および、ソースだけでなく、サブトピックの要素を考慮することで、より有効なサブトピックを上位にランクできることを確認した。

5. おわりに

本研究では、子供にとって重要なサブトピックを子供に優先的に提供するため、先行研究の調査結果に基づき、子供向けサブトピックランキング手法について検討した。子供向けサブトピックの調査に関しては、小中学生 49 名に、一般向け検索エンジン、子供向け検索エンジン、Wikipedia という複数のソースから取得したサブトピックを familiar, interesting, useful の 3 つの観点からスコア付けしてもらうことで行った。その調査結果に基づいて、子供向けサブトピックランキング手法を提案した。子供向けサブトピックのランキングは、取得ソースの種類と難易度などのサブトピックの要素に関する素性をもとに、サブトピックの子供向け度合いを数値化することで行う。子供にスコア付けしてもらったデータセットを用いて、子供向けサブトピックランキング手法で順位付けしたサブトピックのランキングの精度を確認したところ、提案手法では、サブトピックの抽出元ソースの種類と、サブトピック自身の難易度、教科カテゴリ、検索ヒット数といったシンプルな素性を考慮するだけで、ソースを単独で用いたり、抽出元ソースの種類のみを考慮してサブトピックをランキングする場合よりも、子供向け度合いの高いサブトピックを多く上位にランクできたことを確認した。

今後は、より高精度な子供向けサブトピックランキングを実現するために、各素性を自動でスコア化する方法を改善する必要がある。また、現状で用いている素性である意図、抽象度、難易度、検索ヒット数というサブトピックの特徴の他にも有効な特徴がないのかをさらに検証する必要がある。さらに、今回

は子供目線での重要度のみを考慮していたので、子供にとって重要なサブトピックの中で、実際に Web 検索時に役立つサブトピックがどのようなものなのかを調査する必要がある。

6. 謝 辞

本研究は JSPS 科研費 (24187) の助成によるものである。ここに記して謝意を表す。

文 献

- [1] J. R. Anderson: *Cognitive Psychology and Its Implications*, Freeman & Co, 1980.
- [2] D. Bilal: Children's Use of the Yahoo!igans! Web Search Engine III. Cognitive and Physical Behaviors on Fully Self-Generated Search Tasks, *Journal of the American Society for Information Science and Technology*, Vol. 53, No. 13, pp. 1170–1183, 2002.
- [3] D. Bilal and J. Kirby: Differences and Similarities in Information Seeking: Children and Adults as Web Users, *Information Processing and Management*, Vol. 38, No. 5, pp. 649–670, 2002.
- [4] Z. Dou, S. Hu, K. Chen, R. Song, and J.-R. Wen: Multi-Dimensional Search Result Diversification, *Proc. WSDM 2011*, pp. 475–484 (Feb. 2011).
- [5] A. Druin, E. Foss, H. Hutchinson, E. Golub, and L. Hatley: Children's Roles using Keyword Search Interfaces at Home, *Proc. CHI 2010*, pp. 413–422, 2010.
- [6] C.W. Dunnett: New Tables for Multiple Comparisons with a Control, *Biometrics*, Vol. 20, No. 3, pp. 482–491, 1964.
- [7] K. Gyllstrom, M.-F. Moens: A Picture is Worth a Thousand Search Results: Finding Child-oriented Multimedia Results with CollAge, *Proc. SIGIR 2010*, pp. 731–732 (July 2010).
- [8] K. Gyllstrom, M.-F. Moens: Wisdom of the Ages: toward Delivering the Children's Web with the Link-based Agerank Algorithm, *Proc. CIKM 2010*, pp. 159–168 (Oct. 2010).
- [9] J. Hu, G. Wang, F. Lochovsky, J.-T. Sun, and Z. Chen: Understanding User's Query Intent with Wikipedia, *Proc. WWW 2009*, pp. 471–480 (Apr. 2009).
- [10] M. Iwata, Y. Arase, T. Hara, and S. Nishio: A Children-oriented Re-ranking Method for Web Search Engines, *Proc. WISE 2010*, pp. 225–239 (Dec. 2009).
- [11] 岩田麻佑, 原 隆浩, 西尾章治郎: 子供による Web 検索支援のための検索クエリに関する子供向けサブトピックの調査, 情報アクセスシンポジウム 2012, pp. 23–30, 2012.
- [12] M. Nakaoka, Y. Shirota, and K. Tanaka: Web Information Retrieval Using Ontology for Children based on Their Lifestyles, *Proc. ICDEW 2005*, p. 1260 (Apr. 2005).