検索トピックに対する反証示唆がウェブ検索ユーザの情報精査態度に与える影響

Analysis on Effect of Disputed Topic Suggestion for User Behavior in Web Search

山本 祐輔
Yusuke Yamamoto

京都大学 経営管理研究部
Center for Research on Business and Administration, Kyoto University
yamamoto@gsm.kyoto-u.ac.jp, <http://hontolab.org/>

嶋田 敏
Satoshi Shimada

同上
Same as above
shimada@gsm.kyoto-u.ac.jp

keywords: web search, careful information seeking, information credibility, attitude change, behavior analysis.

Summary

During web search and browsing, people often accept misinformation due to their inattention to information credibility and biases. To obtain correct web information and support effective decision making, it is important to enhance searcher credibility assessment and develop algorithms to detect suspicious information. In this paper, we investigate how credibility alarms for web search results affect searcher behavior and decision making in information access systems. This study focuses on disputed topic suggestion as a credibility alarm approach. We conducted an online user study in which 92 participants performed a search task for health information. Through log analysis and user surveys, we confirmed the following. (1) Disputed topic suggestion in a search results list makes participants spend more time browsing pages than ordinary search conditions, thereby promoting careful information seeking. (2) Disputed topic suggestion during web browsing does not change participant behaviors but works as complementary information. This study contributes to system designs to enhance user engagement in critical and careful information seeking.

1. ま え が き

今日、ウェブ上で発信される情報は、社会、経済、政治、文化、日常生活など、社会のあらゆる方面にわたっている。今やウェブは我々にとって重要な情報基盤となっている。しかし、ウェブに存在するすべての情報が正しいとは限らない。最悪の場合、情報の信憑性を意識していなければ、誤った情報を鵜呑みにしてしまう可能性もある。Nakamura らは、検索エンジンユーザの 50%以上が検索エンジンが返す検索結果に含まれるウェブページはある程度信用できると考えていると報告している [Nakamura 07]。また、Lindgaard らの調査では、見た目が綺麗なウェブページを信用してしまう人が多いということが明らかになっている [Lindgaard 11]。

ウェブ情報は、公開前に第 3 者によって内容チェックを受けることは稀である。The Pew Research Centre の 2011-2012 年レポート^{*1}によると、米国のインターネットユーザの 70%が健康に関する情報を検索した経験があるとされている。しかし、Silence らの報告によると、医

療関係のウェブサイトのうち医療専門家によってチェックを受けているものは全体の 50%以下しかない [Silence 04]。このように、内容検証されていないウェブ情報が多いにもかかわらず、あまり注意せずにウェブ情報を受け入れてしまっているユーザは少なくない。それゆえ、信憑性の高い情報の取得を支援する情報システムの構築が重要となる。本稿では、ユーザの信憑性に対する意識を高め、情報精査を促進する情報システムのデザインについて検討を行う。

これまで、信憑性判断支援のアプローチとして、証拠情報の検索システム [Leong 12]、反証センテンスの提示システム [Ennals 10, Yamamoto 12]、信憑性判断に重要な指標に沿ったスコアの可視化システム [Schwarz 11, Yamamoto 11] など、様々なアイデアが提案されてきた。提案されてきたシステムは、ユーザが適切に使いこなすことができれば非常に有用である。しかし、人は事前に自身が抱いている信念や先入観を補強する情報を優先的に選んでしまうことが少なくない [Kahneman 03]。それゆえ、先に挙げたようなシステムが、ユーザの事前信念に沿うような補完情報を提示してしまった場合、ユーザによっては評価対象となっている情報の信憑性を都合よく解釈し、

*1 PewResearchCenter's Health Fact Sheet,
<http://www.pewinternet.org/fact-sheets/health-fact-sheet/>

図 1 検索結果リスト閲覧中およびウェブページ閲覧中における反証トピック提示の例.

信じ込んでしまうことも起こりえる。また、ユーザがそもそもウェブの情報に信用できると思い込んでいる、あるいは疑わしい情報が存在するとしても自分でそれを検知できると思い込んでいる場合、先に挙げたようなシステムに必要性を感じない可能性がある。これらを踏まえると、情報探索時の注意意識を高めることも重要な課題となる。注意深い情報探索、情報を精査することへの意識を喚起できれば、ユーザは批判的にウェブページを閲覧したり、複数の情報ソースを見比べるようになる。また、過去に提案されてきた支援システムも効果的に利用できるようになる。

本稿では、情報精査を促進する情報システムのデザインを検討するために、情報検索時における信憑性への注意喚起がユーザの検索行動に与える影響について分析を行う。説得コミュニケーションの分野では、態度変容の方法の1つとして恐怖アピールに関する研究が行われてきた [Rogers 83]。本稿では、恐怖アピールの一種として反証トピックの示唆に着目し、信憑性への注意喚起を行う。ここで、反証トピックとは、そのトピックが示す内容が真実かどうかにかかわらず、一部の人がウェブ上で疑義を唱えているトピックと定義する。例えば、あるウェブページで「朝バナナダイエットはブームになったが、実際は効果はない」と言及していた場合、「朝バナナダイエット」を反証トピックと見なす。本研究では、反証トピックを示唆するタイミングとして「ウェブ検索結果一覧ページの閲覧時」および「検索結果中のウェブページの閲覧時」の2種類に焦点を当てる（図1参照）。本研究では、反証トピックの示唆効果を調べるために、健康情報に関する検索タスクを行わせるオンラインユーザ実験を実施した。実験結果の分析を通じて、ウェブ検索中の反証トピックの示唆がユーザの検索行動や最終的な意思決定に与える影響を明らかにする。

本研究を通じて明らかになった知見を以下にまとめる：

- 反証トピックの示唆効果は提示するタイミングによって異なる。
- 検索結果一覧を走査している時に反証トピックを示

唆すると、検索結果の選択およびウェブページの閲覧にかかる時間が長くなる。

- ウェブページ閲覧時に反証トピックを示唆しても、検索行動はあまり変化しない。ウェブページ閲覧時に提示された反証トピックは、最終的な意思決定の際の補完情報として利用される。

2. 関連研究

2.1 ウェブ情報の正確性・信憑性の評価

疑わしいウェブ情報対策として考えられるアプローチの1つは、アルゴリズムによって情報の正確性や信憑性を評価することである。Galland らは、評価対象となっている情報の内容を支持する情報ソース、反対する情報ソースを集約することで情報の正確さを評価するアルゴリズムを提案している [Galland 10]。Dong らは、誤った事実や主張を含まないウェブページほど正確であると仮定し、ウェブページの正確さを評価するアルゴリズムを提案している [Dong 15]。これらの手法は、評価対象のドメインに関する事前知識が十分にある場合、有効に働く。

Wikipedia や Twitter、ウェブページなど、特定の情報タイプに焦点を絞った情報の信憑性評価手法も提案されている。Adler らは、編集されずに残っている文は信憑性が高いとの仮定のもと、Wikipedia に記載された文の信憑性の評価アルゴリズムを提案している [Adler 07]。Castillo らは、ニュース記事に対する Tweet（つぶやき）や記事の共有行動を解析することで、Twitter 上で伝播するニュース記事の信憑性を評価するアルゴリズムを提案している [Castillo 11]。これらの手法では、客観性や情報ソースの権威、情報の鮮度など、信憑性の認知に影響する要素に着目している [Metzger 03]。しかし、信憑性を感じる情報が常に正確であるとは限らない。信憑性評価アルゴリズムによって算出されたスコアは、ユーザがアルゴリズムが意味するところを理解して利用すれば有用であるが、不正確な情報を自動検知するためにスコア情報をそのまま使用することは危険である。

2.2 ウェブ情報の信憑性判断支援

ユーザによる信憑性判断を支援するシステムに関して、研究が少なからず行われている。Suh らは、Wikipedia 上の記事の編集履歴を可視化するシステム WIKIDASHBOARD を提案している [Suh 08]。Leong らはユーザがウェブページを閲覧中に疑わしい文の信憑性を検証できるよう、証拠情報の検索アルゴリズムを提案している [Leong 12]。Yamamoto らは、信憑性判断に有用とされるいくつかの観点に沿ってウェブページをスコアリング・可視化するプロトタイプシステムを提案している [Yamamoto 11]。Diakopoulos らは、ユーザの評価やコメントを集約・可視化することで閲覧中の動画コンテンツの信憑性判断を支援するシステムを提案している [Diakopoulos 10]。

検証すべき情報に気づかせるために、疑わしい情報への注意喚起に焦点をあてた研究もいくつか行われている。本稿でもこのアプローチを採用している。Ennals らは閲覧中のウェブページに信憑性が疑われる文が存在したときに、それをハイライトするシステム DISPUTE FINDER を提案している [Ennals 10]。Yamamoto らは、反証示唆に着目した新しいタイプのクエリ推薦手法を提案している [Yamamoto 12]。これらの研究では、提案システムの有用性に関する定性的な評価は行われているが、ユーザの態度や検索閲覧行動への影響に関する定量的な評価はほとんど行われていない。本論文では、疑わしい情報への注意喚起がユーザの検索行動や意思決定に与える影響を分析する。

2.3 ウェブ検索・閲覧におけるバイアス

ユーザが情報の信憑性に注意を払っていたとしても、勘違いや先入観によって誤った意思決定を行ってしまうことがしばしば起こりえる。これは認知バイアスとして知られている [Kahneman 03]。情報検索分野においても、いくつかの認知バイアスが知られている。Leong らは、特定のインターネットドメインに属しているウェブページなら信用できるとユーザが感じてしまう「ドメインバイアス」の存在を明らかにしている [Leong 12]。Clarke らは、商用のウェブ検索エンジンのクリックスルーログを解析した結果、可読性の高い検索結果ほどクリックされやすい傾向にあることを明らかにした [Clarke 07]。White らは、検索トピックに関してユーザが抱いている事前信念と検索行動の関係について分析を行い、検索トピックに関して事前に抱いている信念が弱い場合、検索によってその信念が変更される可能性が高いことを明らかにした [White 13]。また、事前信念が強い場合、検索やページ閲覧後も信念が変更される可能性は低いことも明らかにした。これらの研究は、誤った情報の取得を防ぐためには認知バイアスの考慮が重要であることを示唆している。

3. 方 法

反証トピックの示唆タイミングの違いが検索行動や意思決定に与える影響を調べるために、クラウドソーシングサービスを用いたオンライン実験を行った。実験では、用意した3種類の検索インタフェースのいずれかを用いて、ある健康問題に対する治療法・対処法をウェブから探すタスクを被験者に依頼した。本実験では、以下について明らかにする：

- 反証トピックの示唆によって、信憑性が不確かなトピックに対する注意が喚起されるか？
- ウェブ検索結果の精査を最も促進する反証トピックの示唆タイミングはいつか？
- 示唆された反証トピックは、最終的な意思決定にどのような影響を与えるか？

図2 (a) SERP-d UI および (b) Page-d UI の振る舞い（赤色でハイライトされた語は反証トピック）

本実験の主目的は、反証トピックの示唆が検索ユーザの注意深い情報探索を促進するか否かを確認することである。それゆえ、タスクの効率性や正確性については議論しない。

3.1 実験条件

本実験は、検索インタフェースを要因とする1要因3水準の実験計画で実施した。本実験では、検索ユーザインタフェース (UI) として、**Control UI**、**SERP-d UI**、**Page-d UI** の3種類を用意した。Control UI は、検索クエリに対して検索結果のリストが返されるシンプルな検索インタフェースである。各検索結果は一般的な検索エンジンと同様、タイトル、スニペット、URL で構成されている。SERP-d UI では、検索クエリが入力されると検索結果のリストが提示されるが、検索結果一覧ページ (SERP) 上でのみ反証トピックがハイライトされる。入力されたクエリに関して反証トピックが存在する場合、図2(a)のように、各検索結果のタイトルもしくはスニペットに含まれている反証トピックが赤色でハイライトされる。Page-d UI では、検索クエリに対して検索結果のリストが提示されるが、図2(b)のように、ウェブページを閲覧時にのみ反証トピックがハイライトされる。Page-d UI を使用する被験者は、検索結果をクリックしてウェブページを訪問するまで、どのトピックに反証が存在するかを知ることはできない。

3.2 実験環境

§1 タスクで用いる検索トピック

検索トピックとして、健康に関するトピックを9個用意した。トピックは、ある程度の話題性があり、民間療法や都市伝説など信憑性が怪しい治療法や予防法を解説

表 1 検索トピックと反証トピックの例

検索トピック	反証トピック例
癌	アバスタチン, BCG ^{*2} , MTX ^{*3} , 炭水化物, マクロビオティック, ホルモン療法, ビタミン C
ダイエット	カラオケ, バナナ, 水, 炭酸水, 半身浴, 低炭水化物, 豆腐
ニキビ	アセロラ, オリーブ, ステロイド, ヨーグルト, ワセリン, 水洗顔, レーザー
二日酔い	ターメリック, カフェイン, アクエリナス, サウナ, シジミ, ハチミツ, 梅干し
アトピー	BCG, うなぎ, チョコレート, ベビーオイル, マイタケ, パン, 抗ヒスタミン
うつ病	SNRI ^{*4} , SSRI ^{*5} , インデラル, スクリーニング, アルコール, 運動, 精神療法
便秘	こんにゃく, ごま, ブルーン, 下剤, ラクトバシラス, チンゲンサイ, 酵素ダイエット
花粉症	プロテイン, 喫煙, アルコール, マーガリン, セレスタミン, 魚卵, 漢方薬
喘息	β_2 刺激薬, アルコール, APAP ^{*6} , 喫煙, 砂糖, 牛乳, メブチン

するウェブページが多いものを選んだ。用意したトピックを表 1 に記す。

§2 反証トピック

本研究では、反証トピックを「あるウェブページで否定的なコメントがなされているトピック」として定義する。事前に用意した検索トピックに関連する反証トピックを収集するために、本研究では WISDOM X^{*7}を用いた。WISDOM X は、自然言語形式で質問文が入力されると、その質問に対する回答の候補一覧をウェブから自動的に抽出するシステムである [Akamine 09]。例えば「何がダイエットに効果的?」という質問文に対して「低炭水化物ダイエット」といった回答候補が返される。

表 1 に記した検索トピックに関する反証トピック候補を得るために、本研究では WISDOM X に対して「何が“検索トピック”に対して効果がない?」という質問をクエリとして入力し、得られた解リストから意味不明なものを除外したものを反証トピックとした。得られた反証トピックの例を表 1 の右欄に記す。

§3 検索タスクで表示するウェブページ

実験開始前に、Microsoft Bing Search API を用いて各タスクで表示するウェブページを収集した^{*8}。収集されたウェブページは (1) 反証トピックを含むウェブページ (*disputed pages*) と (2) 反証トピックを含まないウェブページ (*non-disputed pages*) の 2 種類である。(1) と (2) のウェブページは、各検索トピックごとにそれぞれ 100 件ずつ集められた。反証トピックを含むウェブページ (1) の収集には、検索トピックを q 、その反証トピック d とした

図 3 検索結果ランキングの調整方法。反証トピックを含むページは赤色のスロットに、含まないページは白色のスロットに割り当てた。事前に調査したクリック率が高いウェブページほど上位に表示されるように調整した。

時、クエリ “ $effective \wedge q \wedge d$ ” を用いた。反証トピックを含まないウェブページ (2) の収集には、クエリ “ $effective \wedge q$ ” を用いた。Bing API を用いて各種ウェブページを収集した後、得られたページの適合性を手動で判定し、適合と見なしたページのタイトル、スニペット、URL、HTML コードを保存した。

§4 検索結果ランキング

本実験では、どの検索インタフェースを用いても、各クエリに対してあらかじめ定義した検索結果リストが表示されるよう調整した。一般に、検索エンジンユーザの大半は、検索結果リストに表示された最初の 10 件しか見ないことが知られている [Silverstein 99]。しかし、今回の実験では、リストに表示される検索結果数は 100 とした。これは、被験者の中に注意深い被験者がいた場合、できる限り多くのウェブページを精査できるようにするためである。各検索結果は、タイトル、スニペット、URL から構成し、どの検索インタフェースを用いても、同一クエリに対しては同じ検索結果リストが表示されるよう設定した。なお、本実験ではタスク開始と同時にあらかじめ用意したクエリが自動的に投入されるようになっており、被験者がクエリを変更できないよう設定した。また、被験者の検索閲覧行動をモニタリングするために、検索結果中のウェブページに含まれるハイパーリンクをクリックしてもページ遷移しないよう設定した。

実験で提示される検索結果および検索結果のランキングについては、実験に先立って意図的に調整した。この操作の目的は以下の 2 つのケースを防ぐためである：

- (1) SERP-d UI を使用する被験者が反証トピックに気づかないケース
- (2) Page-d UI を使用する被験者が反証トピックを含むウェブページを訪問しないケース

^{*2} BCG: Bacillus Calmette Guerin

^{*3} MTX: Methotrexate

^{*4} SNRI: Serotonin & Norepinephrine Reuptake Inhibitors

^{*5} SSRI: Selective Serotonin Reuptake Inhibitors

^{*6} APAP: paracetamol

^{*7} WISDOM X: <http://wisdom-nict.jp> (in Japanese)

^{*8} Microsoft Bing Search API:

<https://datamarket.azure.com/dataset/bing/search>

ケース1への対応として、ランキングの上位ほど、タイトルもしくはスニペットに反証トピックを含む検索結果が高頻度に出現するよう検索結果リストを調整した。具体的な調整手順は以下の通りである。まず、図3に記したように、各クエリに対する検索結果リストを5つのグループに分割した(G1, G2, G3, G4, G5)。各グループには、タイトルもしくはスニペットに反証トピックを含む検索結果が80%, 65%, 50%, 35%, 20%含まれるようにした。反証トピックを含む検索結果の出現位置(スロット)、反証トピックを含まない検索結果のスロットは、各グループ内でランダムに配置した。

ケース2への対応として、反証トピックを含むウェブページと含まないウェブページのクリック確率を事前に計算し、クリック率の高い順にケース1の対応策で用意したスロットに各ウェブページを割り当てた(図3では、反証トピックを含むページは赤色に、含まないページは白色のスロットに対応)。クリック確率は、クラウドソーシングサービスの1つであるLancers.jp^{*9}を用いて事前に計算した。Lancers.jpのワーカーには、表1に挙げた検索トピックに関連するウェブページのリストを提示し、その中から設定した検索意図を満たすウェブページを探すタスクを依頼した。検索意図は本実験と同様、「ある健康トピックに対する治療法・対処法を探す」という設定にした。提示するウェブページは3.2節§3で述べたものと同一のものをを用いた。ウェブページリストのランキングは被験者毎にランダム化した。この予備タスクでは、反証トピックはハイライトされないよう設定した(Control UIと同じ条件)。タスク中、ワーカーのクリックスルーログを記録した。1つの検索トピックに少なくとも100人以上のワーカーが参加するようにした。最終的に、延べ1012人のワーカーが予備タスクに参加した。得られたクリックスルーログを用いて、リスト中のウェブページのクリック確率を算出した。

3.3 被 験 者

本実験には計121人の被験者が参加した。被験者はLancers.jpを通じて集められた。本実験には、ウェブページのクリック確率計算のための予備タスクに参加した者は含まれていない。どの被験者もウェブ検索・閲覧に慣れている者が選ばれた。被験者はタスクごとに27個のグループ(3UI×9検索トピック)のいずれかに割り当てられた。本実験への参加は最大9タスクに制限された。被験者には、各タスクの終了時に報酬として50円を支払った。

本実験では、121人の被験者から507件のタスクデータを得た。得られた結果のうち27件はタスクが完全に終了していない、180件はJavascriptの問題によって被験者の行動を完全にモニタリングできなかったため除外

した。最終的に、92人の被験者に関する310件のタスクデータを分析に用いることにした。

3.4 実 験 手 順

本オンライン実験は「被験者登録」、「タスクの説明」、「タスク」、「事後アンケート」の4ステップから成る。

本実験では、「健康に有効なものを調査する」タスクとしてLancers.jpを通じて被験者を募った。タスクの所要時間は5~10分として募集を行った。以上のタスク要件に合意した被験者にはLancers.jp上で仮契約を行い、著者が用意した実験用サイトに移動してもらった。

実験用サイトに移動したタイミングで、被験者には、タスク用の検索トピックと検索ユーザインタフェースをランダムに割り当てた。その後、被験者にはタスクの手順に関する短い導入説明を行った。さらに、SERP-d UIもしくはPage-d UIが割り当てられた被験者には、検索タスク中に反証トピックがハイライトされる旨を「水虫治療」を例に挙げ説明した。反証トピックについては、「あるトピックが対象となっている健康問題を解決・改善するのに有効でないと記述しているウェブページが存在した場合、そのトピックが赤色でハイライトされる」と説明した。タスクの手順に関する説明を行った後、被験者に以下のようなシナリオが提示された。

「あなたは健康に問題を抱えており、ダイエットをしようと思っています。あなたは今、効果的なダイエット方法について、ウェブで検索しようとしています。以下のリンクをクリックして検索を開始し、有効なダイエット方法を見つけ、報告してください」

シナリオを読んだ後、被験者は指定されたリンクをクリックし検索を開始した。検索ページに移動した直後、通常のウェブ検索エンジンと同様にウェブ検索結果が提示された。提示される検索結果の数は100件とした。被験者は時間無制限で検索結果リストおよびウェブページを閲覧し、シナリオに適合する有効な解を探した。納得できる解を発見次第、検索を終了し解を報告してもらった。検索の結果、満足する解が得られなかった場合、「解なし」と報告してもらった。タスク中、被験者のクリックスルーおよびウェブページの滞在時間はJavascriptを通じて記録された。

タスク終了後、SERP-d UIもしくはPage-d UIを使用した被験者には、各UIの有用性に関するアンケートに回答してもらった。有用性は5段階のリッカート尺度で評価してもらった(1:まったく有用でない~5:とても有用である)。また、どのような点が有用であったかについて、下記のリストから複数選択式で回答してもらった。

- 信憑性を意識した情報検索への意識づけ: 検索中、情報の信憑性を意識するようになる点
- 検証すべき情報の示唆: どのトピックを詳細に調べべきかに気づくことができる点

*9 Lancers.jp: <http://www.lancers.jp/>

表 2 検索インタフェースごとのページビューの平均と標準誤差 (SEM)

検索インタフェース	Control	Page-d	SERP-d
ページビュー (SEM)	5.24 (0.634)	5.55 (0.537)	4.64 (0.547)

- 疑わしい情報のフィルタリング: 無視すべき疑わしいトピックに気づくことができる点
- その他: その他
- なし: 有用な要素はない

4. 結 果

以下では、各種検索インタフェースを使用した際の被験者の行動分析と事後アンケートの結果を記す。なお、検索トピック間における被験者の検索行動の偏りの有無を確認するために、ベースラインである Control UI を用いた被験者のページビュー数、タスク所要時間、検索結果一覧ページにおける滞在時間の 3 指標について、検索トピックを要因とする 1 要因 9 水準の分散分析を行った。その結果、どの指標についても検索トピック間に統計的有意差は見られなかった (ページビュー: $F(8, 96) = 0.774, p = .627, \eta^2 = 6.05e-2$) (タスク所要時間: $F(8, 96) = 1.098, p = 0.372, \eta^2 = 8.38e-2$) (検索結果一覧ページにおける滞在時間: $F(8, 96) = 0.609, p = .768, \eta^2 = 4.83e-2$)。それゆえ、以後の分析では検索トピックは区別せずに 3 つの検索インタフェースの比較を行った。統計的検定には 3 つの検索インタフェースを要因とする 1 要因 3 水準の分散分析を用いた。ポストホック比較には、Tukey-Kramer テストを用いた。事後アンケートの解析には、ピアソンの χ^2 検定を用いた。

4.1 ページビュー

Page-d UI と SERP-d UI によって複数の情報ソースの確認が促進されるか否かを分析するために、クリックスルーデータを解析し、被験者が訪問したウェブページの数 (ページビュー) を分析した。分析結果を表 2 に記す。分散分析の結果、ページビューについては、Control UI, SERP-d UI, Page-d UI の間に統計的有意差は確認されなかった ($F(2, 307) = 0.638, p = .529, \eta^2 = 4.14e-3$)。

被験者が検索結果から訪問ページを選択する際に、Page-d UI と SERP-d UI がどのように影響したかを調べるために、1 タスクの中でクリックされた検索結果のうち、反証トピックを含む検索結果がクリックされた割合 (clickthrough ratio) を検索インタフェースごとに算出した。Control UI と Page-d UI については、検索結果一覧ページで反証トピックがハイライトされることはない。しかし、比較のために、SERP-d UI でハイライトされる検索結果が Page-d UI および Control UI でクリックされた割合を仮想的に算出した。その結果、3 種類の検索インタフェース間で

図 4 検索インタフェースごとに見る、タスク中にクリックされた反証トピックを含むページの割合 (図中の線はインタフェース間に統計的有意差が確認されたことを意味する)。

図 5 検索結果順位ごとに見る、反証トピックを含む/含まない検索結果のクリック確率の違い。

統計的有意差が確認された ($F(2, 307) = 19.8, p^{**} < .01, \eta^2 = 0.114$)。図 4 で示しているとおり、ポストホック解析でも Control UI や Page-d UI と比較して SERP-d UI を用いた際の clickthrough ratio は有意に小さかった (vs. Control; $49.7\% < 70.5\%, p^{**} < .01$) (vs. Page-d; $49.7\% < 78.9\%, p^{**} < .01$)。この結果は、検索結果一覧ページで反証トピックのハイライトすると、被験者は反証トピックを検索結果に含むページへの訪問を避けるようになることを示唆している。

[Joachims 07, Silverstein 99] など多くの研究で報告されているように、一般にウェブ検索において、ユーザは検索結果リストを上位からチェックする。それゆえ、検索結果上位のウェブページほどクリックされやすい。そこで、検索結果リストページ上での反証トピック提示が検索結果下位にあるウェブページの閲覧を促進するかを検証するために、検索結果のクリック確率を検索結果順位別に分析を行った。ここで、ある検索結果のクリック確率とは、全被験者のうち、その検索結果をクリックした被験者の割合を意味する。図 5 は、SERP-d UI を使用した被験者が反証トピックを含む検索結果をクリックした確率 (All on Control)、反証トピックを含まない検索結果をクリックした確率 (Non-disputed results on SERP-d)、Control UI を使用した被験者が検索結果をクリックした確率 (Disputed results on SERP-d) を順位ごとに記している。3.2 節 §4 で記したように、本実験ではどの検索 UI

を用いた場合でも、同一の検索結果が提示されるよう調整した。それゆえ、ある検索結果に対するクリック確率を検索 UI ごとに比較することができる。先に述べたように、通常、検索結果の下位にあるページほどクリック確率が低下する。

図 5 によると、SERP-d UI を用いた被験者は、検索順位の 1~40 位にかけては、反証トピックを含む検索結果よりも反証トピックを含まない検索結果をクリックすることが多かった。特に検索結果上位の 1~20 位にかけては、反証トピックを含まない検索結果のクリック確率が反証トピックを含む検索結果のクリック確率に比べて 10% 以上高かった。検索ユーザの多くは検索結果の上位しか閲覧しないという先行研究を踏まえると、この結果は比較的閲覧されやすい検索結果上位のグループでは、被験者は反証トピックを含まないウェブ検索結果をよりクリックする傾向にあることを示唆している。このことは、先の「1 タスク内における反証トピックを含むページのクリック割合」の結果と矛盾しない。さらに、図 5 によると、SERP-d UI で示唆された反証トピックを含まない検索結果に対するクリック確率は、Control UI におけるそれと比べ、検索結果の順位が下がるにつれて緩やかに減少していることが分かる。このことは、Control UI を用いた被験者とは異なり、SERP-d UI を用いた被験者は検索結果上位だけでなく、検索結果中位にあるウェブページも確認しようとしていることを示唆している。

一方、検索結果下位のグループでは、SERP-d UI を用いた場合の反証トピックを含んだ検索結果に対するクリック確率、含まない検索結果に対するクリック確率については、差がほとんどなかった。また、両クリック確率は、Control UI を用いたときのクリック確率とも差がなかった。これは Control UI のクリック確率が示しているように、そもそも検索結果リストの下位にある検索結果はほとんどクリックされることがないため、3 つのクリック確率の間にほとんど差が現れなかったと考えられる。

4.2 ページ滞在時間

Page-d UI と SERP-d UI によって注意深いウェブページ閲覧が促進されるかを調べるために、各種検索インタフェースを用いた場合の検索結果一覧ページの閲覧時間、ウェブページの閲覧時間を分析した。図 6 は、検索結果一覧ページにおける滞在時間、反証トピックを含むウェブページの総滞在時間、反証トピックを含まないウェブページの総滞在時間、全ウェブページの総滞在時間（タスク所要時間）を被験者ごとに平均を取ったものを示している。Control UI 使用時における反証トピックを含む/含まないウェブページの総滞在時間については、SERP-d UI もしくは Page-d UI で反証トピックがハイライトされる/されないウェブページに Control UI を用いた被験者が滞在した時間を算出し用いた。

タスク所要時間については、3 つの検索インタフェース

図 6 検索インタフェース別に見る、全ページ (All)、検索結果一覧ページ (SERP)、反証トピックを含むページ (Disputed)、反証トピックを含まないページ (Non-disputed) の総滞在時間の違い。図中の線は統計的有意差の度合いを意味する (*= $p < 0.1$, *= $p < .05$, **= $p < .01$)。

間で統計的な有意差が確認された ($F(2,307)=3.75, p^* < .05, \eta^2 = 2.38e-2$)。さらに図によると、SERP-d UI を用いた際のタスク所要時間は、Control UI を用いた際のそれに比べて有意に長かった ($317.8s > 214.7s; p^* < .05$)。また、Page-d UI を用いた際のタスク所要時間よりも長い傾向が見られた ($317.8s > 217.0s, p=.056, < 0.1$)。検索結果一覧ページにおける滞在時間に関しても、分散分析の結果、3 種類の検索インタフェース間に統計的な有意差が確認された ($F(2,307)=3.27, p^* < .05, \eta^2 = 2.09e-2$)。また、ポストホック解析の結果、SERP-d UI を用いた場合の滞在時間は Control UI および Page-d UI を用いた場合に比べて長い傾向があることが確認された (SERP-d UI vs. Control UI; $129.4s > 84.8s, p=.081$) (SERP-d UI vs. Page-d UI; $129.4s > 82.1s, p=.056$)。この結果から、SERP-d UI を用いた場合、他の検索インタフェースを用いる場合に比べ、より長い時間検索結果リストの走査に時間をかけるようになることが予想される。

反証トピックを含むウェブページの総滞在時間に関しては、3 種類の検索インタフェース間で統計的な有意差は確認されなかった (Control = $117.2s$, SERP-d = $104.3s$, Page-d = $111.5s$; $p = .961$; $\eta^2 = 2.60e-4$)。一方、反証トピックを含まないウェブページの総滞在時間に関しては、3 種類のインタフェース間で統計的有意差が確認された ($F(2,307)=13.0, p^{**} < .01, \eta^2 = 7.78e-2$)。ポストホック解析によると、SERP-d UI を用いた被験者が反証トピックを含まないページに滞在した総時間は、Control UI を用いた被験者、Page-d UI を用いた被験者のそれに比べて有意に長かった (SERP-d UI vs. Control UI; $81.6s > 18.2s, p^{**} < .01$) (SERP-d UI vs. Page-d UI; $81.6s > 23.4s, p^{**} < .01$)。検索結果一覧ページにおける滞在時間の分析結果と合わせると、SERP-d UI を用いた場合、Control UI と Page-d UI を用いたときと比べ、検索結果

一覧ページおよび反証トピックを含まないページでの総滞在時間が長くなることが予想される。

さらに詳細に分析するために、反証トピックを含む/含まないウェブページ 1 ページあたりの滞在時間を検索 UI ごとに比較した。実験では反証トピックを含む/含まないページを訪問しなかった被験者も存在する。そのため、この分析では、各検索インタフェースを用いた被験者が訪問した反証トピックを含む/含まないページの滞在時間をサンプルとみなし、検索インタフェース群ごとの比較を行った。その結果、反証トピックを含むウェブページ 1 ページあたりの滞在時間について、3 種類の検索インタフェース間で統計的な有意差が確認された ($F(2,1227) = 8.17, p^{**} < .01, \eta^2 = 1.31e-2$)。さらに、ポストホック解析によると、図 7 に示したように、SERP-d UI を用いた場合、Control UI もしくは Page-d UI を用いた場合に比べ、反証トピックを含むウェブページの 1 ページあたりの滞在時間が有意に長かった (SERP-d vs. Control; $44.8s > 29.8s, p^{**} < .01$) (SERP-d vs Page-d; $44.8s > 29.1s, p^{**} < .01$)。図 6 で示したように、反証トピックを含まないページの総滞在時間に比べて、反証トピックを含むウェブページの総滞在時間は短かった。しかし、図 7 が示しているように、SERP-d UI を用いた被験者は、いったん反証トピックを含むウェブページを訪問すると、Control UI もしくは Page-d UI を用いた被験者に比べてページをより注意深く閲覧していたことが予想される。

反証トピックを含まないウェブページ 1 ページあたりの滞在時間についても、3 種類の検索インタフェース間で統計的な有意差が確認された ($F(2,1227) = 5.49, p^{**} < .01, \eta^2 = 2.19e-2$)。さらに、ポストホック解析の結果では、SERP-d UI を用いた場合、Control UI を用いた場合に比べ、反証トピックを含むウェブページ 1 ページあたりの滞在時間が有意に長かった (SERP-d vs. Control; $41.9s > 22.6s, p^{**} < .01$)。また、SERP-d UI と Page-d UI との比較では、SERP-d UI を用いた場合のページ滞在時間の方が Page-d UI のそれよりも長い傾向が見られた (SERP-d vs Page-d; $41.9s > 29.5s, p < 0.1$)。この結果から、反証トピックを含まないページの閲覧に関しても、SERP-d UI を用いた場合、1 ページあたりの滞在時間が長くなることが予想される。

4.3 タスクに対する回答

SERP-d UI および Page-d UI が被験者の最終的な意思決定にどのような影響を与えたかを調査するために、被験者が報告した健康問題への対処法に関する 310 件の回答を分析した。本分析では、報告された回答と反証トピックを比較し、反証トピックリストに含まれるもの、含まれないものの 2 グループに人手で分類した。なお、分類の過程で意味が曖昧で分類できなかった 5 件の回答については、分析対象から除外した。

図 8 は、被験者の回答のうち反証トピックリストに含

図 7 反証トピックを含む/含まないウェブページ 1 ページあたりの滞在時間の比較。図中の線は統計的有意差の度合いを意味する ($\cdot = p < 0.1, * = p < .05, ** = p < .01$)。

図 8 反証トピックリストに含まれる (disputed answers) /含まれない回答 (non-disputed answers) の割合。

れるもの、含まれないものの割合を検索インタフェースごとに算出した結果を示している。図によると、Page-d UI および SERP-d UI を用いた被験者の回答の多くは Control UI を用いた被験者の回答に比べ、反証トピックリストに含まれないものが多かった ($69.5\%, 69.4\% > 56.9\%$)。ピアソンの χ^2 検定を行った結果、Page-d UI と Control UI の回答比率、SERP-d UI と Control UI の回答比率の間に有意差がある傾向が見られた (Page-d vs. Control; $\chi^2 = 3.57, p = 5.88e-2 < 0.1$, Cramer's $V = 1.31e-1$) (SERP-d vs. Control; $\chi^2 = 3.36, p = 6.67e-2 < 0.1$, Cramer's $V = 1.30e-1$)。これらの結果は、SERP-d UI と Page-d UI を用いた被験者は、疑わしいトピックを避けるように最終的な意思決定を下している可能性があることを示唆している。

4.4 反証トピック提示の有用性

Page-d UI と SERP-d UI に対する主観的な評価と利用方法について検討するために、事後アンケートの結果を分析した。表 3 は、効果的な健康促進方法を検索するのに、SERP-d UI と Page-d UI が被験者によってどの程度有用であると評価されたかを示している。SERP-d UI と Page-d UI の平均有用度はそれぞれ 4.02, 3.74 であった。

表 3 SERP-d UI および Page-d UI に対する有用度の平均値と標準誤差 (SEM)

UI condition	SERP-d	Page-d
Mean usefulness	4.02	3.74
(SEM)	(9.30e-2)	(1.01e-1)

図 9 各要素を有用と感じた被験者数の割合 (図中の線はインタフェース間に統計的に有意な傾向があったことを意味する)。

($F(1, 203) = 4.23, p^* < .05, \eta^2 = 2.04e-2$)。この結果から、反証トピックの提示の有用性は肯定的に捉えられていると考えられる。また、被験者は SERP-d UI の方が Page-d UI よりも効果的な健康促進方法を検索するのに有用と感じていることが示唆されている。

図 9 は、SERP-d UI および Page-d UI について、被験者が有用であると感じた要素をまとめたものである。SERP-d UI と Page-d UI について最も有用であると感じられた 2 大要素は「信憑性を意識した情報検索への意識付け (SERP-d = 51.5%; Page-d = 39.6%)」と「検証すべき情報の示唆 (SERP-d = 49.5%; Page-d = 43.4%)」であった。「信憑性を意識した情報検索への意識付け」を有用な要素と感じた被験者の数について、SERP-d と Page-d の間に統計的に有意な傾向が見られた ($\chi^2 = 2.92, p < 8.74e-2 < 0.1$, Cramer's $V = 1.19e-1$)。有用な要素として「疑わしい情報のフィルタリング」を挙げた被験者の数については、「信憑性を意識した情報検索への意識付け」と「検証すべき情報の示唆」ほど多くはなかった (SERP-d = 29.3%; Page-d = 34.0%)。Page-d を用いた被験者の 11.3%, SERP-d を用いた被験者の 4.04% は使用した検索インタフェースに有用な要素はなかったと回答した。

5. 考 察

92 名の被験者の行動分析と事後アンケート分析から、ウェブ検索時における反証トピック示唆について以下の知見が得られた。

- (1) SERP-d UI を利用した被験者は Control UI および Page-d UI を利用した被験者に比べ、タスクにより多くの時間を費やした。特に、検索結果一覧ページおよび反証トピックを含まないウェブページの閲覧

により多くの時間を費やした。

- (2) SERP-d UI を用いた被験者は、検索結果の一覧を走査する際、反証トピックを含まない検索結果をより頻繁にクリックする傾向が見られた。
- (3) SERP-d UI を用いた被験者は、反証トピックを含むページを訪問した場合、他の検索インタフェースを用いた被験者よりもページ閲覧により多くの時間を費やす傾向が見られた。
- (4) ページ滞在時間や全訪問ページ数における反証トピックを含むページの割合については、Page-d UI と Control UI の間で特に顕著な差は見られなかった。
- (5) SERP-d UI および Page-d UI を用いた被験者は Control UI を用いた被験者に比べ、反証トピックに関連する回答を避ける傾向が見られた。
- (6) Page-d UI と SERP-d UI は平均的には被験者から高評価を得た。

検索結果一覧ページでの反証トピックの示唆 (SERP-d UI) によって、検索ユーザはウェブページを閲覧する前にどのウェブページに第 3 者によって疑義が唱えられているトピックが含まれているかを知ることができる。実験から得られた知見を踏まえると、SERP-d UI は疑わしいトピックを含むウェブページのフィルタとしても機能するが、情報検索中にウェブ情報の信憑性への注意意識を高める効果もあると考えられる。図 4 で示したように、SERP-d UI を用いた被験者は、反証トピックを含む検索結果のクリックを避ける傾向にあった。また、SERP-d UI によって被験者のウェブ検索は注意深くなったと考えられる。これは以下の分析結果から推察される: (a) SERP-d UI を用いた被験者は他の検索インタフェースを使用した被験者に比べ、タスクの所要時間が長かったこと、(b) SERP-d UI を用いた被験者は検索結果一覧の走査により長い時間を費やしていたこと (図 6 参照)、(c) SERP-d UI を用いた被験者は反証トピックを含むページを訪問した場合、他の検索インタフェースを用いた被験者よりもそのページの閲覧により長い時間を費やしたこと (図 7 参照)。これらの分析結果から、SERP-d UI は被験者の検索プロセス全体に影響を与え、それによって被験者はウェブ情報の信憑性を吟味するようになったと考えられる。検索結果一覧ページにおける反証トピックの示唆は、注意深い情報探索を促進する可能性がある。

一方で、ウェブページ閲覧中の反証トピックの示唆 (Page-d UI) は、疑わしいトピックを含むページの回避や注意深い情報探索の促進ではなく、意思決定のための補完情報として利用されることが考えられる。Page-d UI が信憑性が疑わしいトピックを含んだページを回避するために活用されるならば、被験者は反証トピックをウェブページ上で確認したらすぐにページを離脱するため、Page-d UI を用いた被験者が反証トピックを含むページに滞在する時間は Control UI を用いた被験者のそれよりも短くなると考えられる。しかし実際には、Page-d UI を用いた際の反証

トピックが含まれるページにおける滞在時間は、Control UI を用いた際のそれと同じような値であった (図 7 参照). さらに、Page-d UI が SERP-d UI と同様に注意深い情報探索を促進するならば、Page-d UI を用いた被験者が反証トピックを含むページの閲覧やタスクそのものに費やした時間は Control UI を用いた被験者のそれよりも長くなることが予想される (図 6, 図 7 参照). しかし、実際の結果はそれとは異なっていた. 一方で、図 8 が示しているように、反証トピックリストに含まれないトピックをタスクの回答として報告した被験者は、Control UI よりも Page-d UI のほうが多かった. さらに、事後アンケート結果によると、被験者の中には Page-d UI は「疑わしい情報」をフィルタリングするのに有用であったと回答した者もいた (図 9 参照). これらの結果を総合すると、ウェブページ閲覧中にハイライトされた反証トピックは、最終的な回答の不確実性を下げるために利用されたと考えられる. すなわち、ウェブページ閲覧時の反証トピックの示唆は、閲覧中の情報が信用できるか否かを判断するための補完情報として機能したと考えられる.

実験結果から、同じ反証トピックを示唆するとしても示唆するタイミングによって、被験者の検索行動に与える影響が異なることが明らかになった. SERP-d UI および Page-d UI は疑わしいトピックに関連する意思決定を回避させる効果があった. しかし、SERP-d UI はその効果に加え、ウェブページを注意深くじっくりと閲覧させる効果があった. 一方、Page-d UI は注意深い情報探索を喚起させるほど強い影響力はなかった. 要約すると、ウェブページを訪問した後に反証トピックを示唆するよりも訪問する前に示唆したほうが、検索ユーザに注意深い情報探索を喚起できる可能性がある.

実験から得られた知見は、ウェブ検索時における情報の信憑性への意識を高める検索インタラクションのデザインに一定の示唆を与えるものである. しかし、今後さらに検討すべき課題がいくつかある. 第一に、反証トピックの示唆効果に個人差があることが挙げられる. 実験結果から、Page-d UI や Control UI を用いるよりも SERP-d UI を用いるほうが注意深いウェブ検索・閲覧が行われることを示したが、一部の被験者は SERP-d UI を用いても検索行動に変化が見られなかった. 表 4 は、SERP-d UI を用いた被験者を、反証トピックに関連する回答をした被験者とそうでない被験者に分け、各グループの検索行動の違いを分析した結果である. 表 4 によると、特に訪問ページ中、反証トピックがスニペットに現れるページを訪問した割合について、反証トピックに関連する回答をした被験者とそうでない被験者の間で統計的有意差が確認された ($F(1,97) = 25.7, p^{**} < .01$). 反証トピックに関連する回答をした被験者は、訪問したページのうち、反証トピックがスニペットに現れるページの割合が平均で 77.8%であった. この値は反証トピックに関連する回答を避けた被験者に比べて 40.3%も高かった. このこと

は、反証トピックに関連する回答をした被験者はそうでない被験者とは異なり、反証トピックが検索結果スニペットでハイライトされていても避けてページを訪問・閲覧していたことを意味する. 考えられうる原因の 1 つは、検索トピックに対する被験者の事前信念や関心の強さに起因する. White らは検索ユーザが検索トピックに対して事前に強い信念を持っている場合、そのトピックについて情報検索を行っても信念の強さが変わらないことを明らかにしている [White 13]. また、Liao らの報告によると、人は自分の意見や信念と対立する情報が存在することを知っていても、自分の意見を強化する情報を優先的に取得する傾向があることが明らかになっている [Liao 13]. これらの研究報告を踏まえると、検索結果一覧ページでの反証トピック示唆はすべての検索ユーザの行動や態度に影響を及ぼすとは限らない. 本研究で得られた検索行動データから、事前信念の有無や強度、それらと反証示唆効果との関係を分析することはできなかった. 注意深い情報探索を促進する方法を開発するためにも、今後は、ユーザ実験の開始前後に被験者へのインタビューやアンケート調査を行い、事前信念の有無や強度を調べることで、検索ユーザの事前信念と反証トピックの示唆効果の関係性について検討を行う必要がある.

第 2 の課題は、検索結果一覧ページで示唆する反証トピックの可視性についてである. 本研究では反証トピック示唆の効果を検証するために、反証トピックを含む検索結果がリストの上位に高確率で出現するよう検索結果リストを調整し、被験者が確実に反証トピックの存在に気づけるような人工的な実験環境を構築した. 実験の結果、検索結果上で反証トピックが存在することを示唆すると注意深い情報探索行動を促進することが示された. しかし、反証トピック示唆による注意喚起の効果は、反証トピックの示唆タイミングだけでなく示唆内容や示唆頻度 (被験者が反証情報を浴びる量) にも関係すると考えられる. Rogers らが提唱している防護動機理論 [Rogers 83] では、恐怖や不都合などの脅威をアピールすることによって態度変容を促す場合、アピールされる脅威の有害さや生起確率が態度変容の促進効果に影響を与えるとされている. 反証トピックの示唆は「それを含む情報を取得すると誤った情報を鵜呑みにしてしまう可能性を示唆している」という意味で一種の恐怖アピールとみなすと、反証トピックの示唆頻度は脅威の生起確率に対応する. 本研究で行った実験では反証の示唆頻度が高い条件であったため、被験者は脅威の生起確率が高いと感じ、それが注意喚起効果に大きく影響したことも考えられる. 今後は反証トピックの出現頻度や出現位置を変動させて実験することで、反証トピックの可視性と注意喚起効果との関係性を調査する必要がある.

第 3 の課題は、検索エンジンと調和する反証トピックの示唆・可視化方法である. 本研究では、検索結果上での反証トピックの示唆が注意深い情報探索を促進すること

表4 SERP-d UIを用いた被験者のうち、反証トピックに関連する回答をした被験者と関連しない回答をした被験者の検索行動の比較. (**) は、被験者間で有意水準 1%で統計的有意差があったことを意味する。

比較項目	反証トピックに関連する 回答をした被験者	反証トピックに関連しない 回答をした被験者
ページビュー	3.57	5.10
タスク所要時間	237.8s	353.1s
検索結果リストページ滞在時間	71.5s	154.5s
訪問ページ中、反証トピックがスニペットに現れるページの割合 (**)	77.8%	37.5%

を示した。実験では、被験者が反証トピックの存在に気づけるよう、検索結果リストの上位に必ず反証トピックが出現するように調整した。しかし、同様の反証トピックのハイライト機能を既存の検索エンジンに実装したとしても、特定のアルゴリズムによって生成される検索エンジンが返す検索結果（の上位）に常に反証トピックを含む検索結果が現れるとは限らない。それゆえ、仮に検索トピックに関して反証されているトピックが存在したとしても、それが検索ユーザの目に触れない可能性もある。注意深いウェブ検索を促すためには、検索結果中の反証トピックのハイライトだけでなく、[Yamamoto 12]のように検索エンジンに実装可能な形式で反証トピックの可視化方法、事前警告方法を検討する必要がある。

6. む す び

本論文では、反証トピックの示唆がウェブ検索行動および情報の信憑性判断に与える影響について、検索結果を走査中の反証トピックの示唆とウェブページ閲覧中の反証トピック示唆の2種類のケースに分けて分析を行った。反証トピックの示唆効果を分析するために、クラウドソーシングを用いたオンライン実験を行った。実験には92人の被験者が参加した。実験の結果、同じ反証トピックを示唆する場合でも、示唆するタイミングによって被験者の検索行動、意思決定に与える影響が異なることが明らかになった。検索結果リスト中で反証トピックを示唆すると、被験者は情報の信憑性を意識するようになり、反証が存在するトピックを含むウェブページの閲覧を避け、情報の精査に時間をかけるようになった。一方、ウェブページ閲覧中に反証トピックを示唆しても、被験者の検索行動は大きく変わらなかった。ページ閲覧中に反証トピックを示唆した場合、その情報は最終的な意思決定のための補完情報として使用される可能性が示された。これらを踏まえると、検索ユーザの意識を高め注意深い情報探索への向かわせるためには、ユーザがウェブページを訪問する前の事前警告が重要であると考えられる。

今後は、反証トピック示唆効果の個人差について明らかにする必要がある。また、ウェブから反証トピックを精度よく効率的に取得する手法の開発も課題としてあげられる。さらに、既存の検索エンジンに親和性の高い事前警告の手法や反証、矛盾を検証するための補完情報を

検索する手法についても研究開発が必要である。

正確な情報を取得し、意思決定を効果的に支援するには、計算機による信憑性評価システムの研究開発に加え、ユーザによる信憑性判断のプロセスを強化することが重要である。本論文が、注意深いウェブ情報探索の強化・促進に寄与できると信じている。

◇ 参 考 文 献 ◇

- [Adler 07] Adler, B. T. and Alfaro, de L.: A content-driven reputation system for the Wikipedia, in *Proceedings of the 16th International Conference on World Wide Web (WWW 2007)*, pp. 261–270 (2007)
- [Akamine 09] Akamine, S., Kawahara, D., Kato, Y., Nakagawa, T., Inui, K., Kurohashi, S., and Kidawara, Y.: WISDOM: A web information credibility analysis system, in *Proceedings of the ACL-IJCNLP Software Demonstrations (ACLDemos 2009)*, pp. 1–4 (2009)
- [Castillo 11] Castillo, C., Mendoza, M., and Poblete, B.: Information credibility on twitter, in *Proceedings of the 20th International Conference on World Wide Web (WWW 2011)*, pp. 675–684 (2011)
- [Clarke 07] Clarke, C. L. A., Agichtein, E., Dumais, S., and White, R. W.: The influence of caption features on clickthrough patterns in web search, in *Proceedings of the 30th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2007)*, pp. 135–142, ACM (2007)
- [Diakopoulos 10] Diakopoulos, N. and Essa, I.: Modulating video credibility via visualization of quality evaluations, in *Proceedings of the 4th Workshop on Information Credibility (WICOW 2010)*, pp. 75–82, ACM (2010)
- [Dong 15] Dong, X. L., Gabrilovich, E., Murphy, K., Dang, V., Horn, W., Lugaresi, C., Sun, S., and Zhang, W.: Knowledge-based trust: Estimating the trustworthiness of web sources, *PVLDB*, Vol. 8, No. 9, pp. 938–949 (2015)
- [Ennals 10] Ennals, R., Trushkowsky, B., and Agosta, J. M.: Highlighting disputed claims on the web, in *Proceedings of the 19th International Conference on World Wide Web (WWW 2010)*, pp. 341–350 (2010)
- [Galland 10] Galland, A., Abiteboul, S., Marian, A., and Senellart, P.: Corroborating information from disagreeing views, in *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining (WSDM 2010)*, pp. 131–140, ACM (2010)
- [Jeong 12] Jeong, S., Mishra, N., Sadikov, E., and Zhang, L.: Domain bias in web search, in *Proceedings of the 5th ACM International Conference on Web Search and Data Mining (WSDM 2012)*, pp. 413–422, ACM (2012)
- [Joachims 07] Joachims, T., Granka, L., Pan, B., Hembrooke, H., Radlinski, F., and Gay, G.: Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search, *ACM Transactions on Information Systems*, Vol. 25, No. 2 (2007)
- [Kahneman 03] Kahneman, D.: A perspective on judgment and choice: Mapping bounded rationality, *American Psychologist*, Vol. 58, No. 9, pp. 697–720 (2003)
- [Leong 12] Leong, C. W. and Cucerzan, S.: Supporting factual statements with evidence from the web, in *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM 2012)*, pp. 1153–1162, ACM (2012)
- [Liao 13] Liao, Q. V. and Fu, W.-T.: Beyond the filter bubble: Inter-

active effects of perceived threat and topic involvement on selective exposure to information, in *Proceedings of the 31st ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2013)*, pp. 2359–2368, ACM (2013)

- [Lindgaard 11] Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., and Noonan, P.: An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages, *ACM Transactions on Computer-Human Interaction*, Vol. 18, No. 1, pp. 1–30 (2011)
- [Metzger 03] Metzger, M., Flanagin, A., Eyal, K., Lemus, D., and McCann, R.: Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment, *Communication Yearbook*, Vol. 27, pp. 293–336 (2003)
- [Nakamura 07] Nakamura, S., Konishi, S., Jatowt, A., Ohshima, H., Kondo, H., Tezuka, T., Oyama, S., and Tanaka, K.: Trustworthiness analysis of web search results, in *Proceedings of the 11th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2007)*, pp. 38–49, Springer (2007)
- [Rogers 83] Rogers, R. W.: *Social Psychophysiology*, chapter Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation, pp. 153–176, Guilford Press (1983)
- [Schwarz 11] Schwarz, J. and Morris, M.: Augmenting web pages and search results to support credibility assessment, in *Proceedings of the 29th ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2011)*, pp. 1245–1254, ACM (2011)
- [Sillence 04] Sillence, E., Briggs, P., Fishwick, L., and Harris, P.: Trust and mistrust of online health sites, in *Proceedings of the 22nd ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2004)*, pp. 663–670, ACM (2004)
- [Silverstein 99] Silverstein, C., Marais, H., Henzinger, M., and Moricz, M.: Analysis of a very large web search engine query log, *SIGIR Forum*, Vol. 33, No. 1, pp. 6–12 (1999)
- [Suh 08] Suh, B., Chi, E. H., Kittur, A., and Pendleton, B. A.: Lifting the veil: Improving accountability and social transparency in Wikipedia with WikiDashBoard, in *Proceeding of the 26th ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2008)*, pp. 1037–1040, ACM (2008)
- [White 13] White, R. W.: Beliefs and biases in web search, in *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2013)*, pp. 3–12, ACM (2013)
- [Yamamoto 11] Yamamoto, Y. and Tanaka, K.: Enhancing credibility judgment of web search results, in *Proceedings of the 29th ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2011)*, pp. 1235–1244, ACM (2011)
- [Yamamoto 12] Yamamoto, Y.: Disputed sentence suggestion towards credibility-oriented web search, in *Proceedings of the 14th Asia-Pacific International Conference on Web Technologies and Applications (APWeb 2012)*, pp. 34–45, Springer (2012)

〔担当委員：奥 健太〕

2016 年 5 月 10 日 受理

嶋田 敏(正会員)

京都大学経営管理研究部特定助教。2015 年東京大学大学院工学系研究科博士後期課程修了。博士（工学）。サービスプロセスのモデル化、シミュレーション、対人サービスにおける人材の評価尺度に関する研究に従事。精密工学会、日本機械学会、観光情報学会、サービス学会等各会員。

著 者 紹 介

山本 祐輔(正会員)

京都大学経営管理研究部特定講師。2011 年京都大学大学院情報学研究科博士後期課程修了。博士（情報学）。情報検索、ウェブマイニング、情報の信憑性、態度変容のインタラクションに関する研究に従事。ACM、情報処理学会、人工知能学会、日本デザイン学会等各会員。