

# 健康情報検索における信憑性判断と意見の形成に関する調査

浜島聡一郎<sup>†</sup> 山本 岳洋<sup>††</sup> 山本 祐輔<sup>†††</sup> 大島 裕明<sup>†,††</sup>

<sup>†</sup> 兵庫県立大学 応用情報科学研究科 〒650-0047 兵庫県神戸市中央区港島南町 7-1-28

<sup>††</sup> 兵庫県立大学 情報科学研究科 〒651-2197 兵庫県神戸市西区学園西町 8-2-1

<sup>†††</sup> 静岡大学 情報学部 〒432-8011 静岡県浜松市中区城北 3-5-1

E-mail: <sup>†</sup>{aa20w509,ohshima}@ai.u-hyogo.ac.jp, <sup>††</sup>t.yamamoto@sis.u-hyogo.ac.jp,

<sup>†††</sup>yamamoto@inf.shizuoka.ac.jp

**あらまし** 本研究では、ある健康情報について複数の情報源から情報が得られたとき、それらをどのように集約して自身の意見を形成しているのかについて調査する。また、ウェブページに対する信憑性判断が意見形成にどのような影響を与えているのかについても調査する。これらを明らかにするために、ユーザ実験を行った。実験では、健康情報を調べる2つの検索タスクを行った。そして、その検索タスクで調べた健康情報について、ユーザはどのように意見を形成したのかを調査した。また、ユーザが検索タスク中に閲覧したウェブページに対する信憑性判断をどのように判断しているのかについて調査した。加えて、ユーザにインタビューを行い意見形成と信憑性判断について調査した。その結果、ユーザは専門性、権威性、客観性のある情報かどうかで信憑性判断をしていた。意見形成では、多数決型、単一ページ確信型、事前信念先行型の少なくとも3つのパターンがあることがわかった。

**キーワード** 情報検索, 健康情報, 信憑性

## 1 はじめに

現在、ウェブ検索エンジンを使用し風邪の治療法やがんの症状などの健康情報を調べることは容易である[27]。また、ユーザはソーシャルメディアを用いて健康情報を探している[24]。

ウェブ上では誰もが情報を発信することができる。そのため、健康情報の質については様々な論文で議論されている[5][21]。正確ではない健康情報を信頼した場合に、健康被害にあう可能性が考えられる。ウェブ上の健康情報の質を判断する基準は様々な研究がされている[20]。例えば、著者、サイト運営者、日付、参考文献が判断基準としている JAMA ベンチマークがあげられる[18]。他にも様々なものが提案されている[2][1]。

ユーザはウェブページの情報に対しての信憑性判断をしている。信憑性とは社会心理学において、「情報発信者が、十分な専門度と信頼度を持っていると、受け手が感じる度合」とされている[10][29]。そして、信憑性は信頼性と専門性の2つから成り立っている。信頼性とは、「情報発信者を信じてよい、うそをついていないだろうと、受け手が感じる度合」である。専門性とは、「その情報が正しくあるために十分な知識を情報発信者が持っている、受け手が感じる度合」のことである。ユーザはウェブページの情報に対して信憑性を判断しながら検索を行っている。このような判断のことを本研究では、信憑性判断と呼んでいる。

ユーザは複数のウェブページを閲覧した後に、調べた健康情報についての自身の意見を形成する。その際に、信憑性があると判断したウェブページの情報から、調べた健康情報に対する自身の意見を形成することがある。意見を形成する際に例えば、病院、公的機関、企業といった異なる企業が調べた健康情

報について同様の内容を言及しているからといった理由が考えられる。

このような、ユーザが複数の情報源から得られた情報をどのように集約し意見を形成しているのかについて明らかにしている研究は少ない。本研究では、ある健康情報について複数の情報源から情報が得られたときに、どのように自身の意見を形成するのかについて明らかにする。また、意見を形成する際にウェブページに対する信憑性判断が影響していることが考えられる。そこで、ウェブページに対する信憑性判断が意見の形成に与えている影響についても調査する。意見の形成について明らかにすることで、どのような要因が意見の形成に影響を与えているのかについての知見が得られる。得られた知見をもとに、慎重な検索を促すような仕組みづくりの提案が可能であると考えられる。

本研究でのリサーチクエストンは以下の通りである。

**RQ** ある健康情報について複数の情報源から情報が得られたとき、どのように自身の意見を形成しているのか。

より具体的には、以下の2つのリサーチクエストンを明らかにすることに取り組む。

**RQ1** 複数の情報源から自身の意見を形成する際に、どのように情報を集約して意見を形成しているのか。

**RQ2** 個々の情報源や、複数の情報源に対する信憑性判断が意見の形成時にどのように用いられているか。

これらのリサーチクエストンを明らかにするために、ユーザ実験を行った。ユーザ実験では、著者らが用意したある症状に対する治療法に関する健康情報を調べる検索タスクを実施する。そして、検索タスクで調べた健康情報についての意見とその意見を形成した理由を調査する。また、ウェブページに対する信憑性判断についても調査する。さらに、意見の形成とウェ

ブページに対する信憑性判断についてインタビュー調査を行う。

このように、意見の形成とウェブページに対する信憑性判断について調査し、分析をすることで意見の形成や意見の形成に影響を与える信憑性判断について明らかにすることができると考えられる。

## 2 関連研究

### 2.1 健康情報を探すユーザ行動

Peterson らは、ユーザの医薬品に関する情報を探す際の検索行動を調査した。多くのユーザの情報を探すスキルは様々であった。しかし、医薬品の情報を探す情報源としてインターネットをユーザは重要視していた。ユーザは情報が正確かどうか考えるときにはウェブアドレスを参考にしていることを明らかにした [15]。

Feufel らは、ユーザが検索に慣れているかどうかで検索行動が変化するのかについて調査を行っている。検索に慣れている若年層と検索に不慣れな高齢者で対象として実験を行っている。高齢者は関係のない情報に気を取られていることを明らかにした [7]。

Zhang らは、検索エンジンを利用して健康情報を獲得するユーザが多いこと、また、ウェブ上で情報を探す際には、ユーザは上位 5 件のウェブページの中でページの適合性や質を考慮していることを明らかにした [23]。

Chen らは、ヘルスリテラシーが低いユーザと高いユーザの情報源に対する信頼度について調査を行った。調査の結果から 70% 近いユーザがオンラインで情報を獲得していた。最も信頼できない情報源として、半数以上のユーザが製薬会社以外の企業と法人があげられていた [3]。

Chi らは、健康情報が複雑で緊急性があるトピックが健康情報を探す行動にどのような影響があるのか調査した。緊急性のあるトピックにおいては医療従事者が情報提供しているウェブページを訪問し、自分に役に立つものとして採用していた。緊急性のないトピックでは、一般的なユーザが提供しているウェブページを採用しているという違いがみられた [4]。

### 2.2 ウェブ上の健康情報に対するユーザの考え方の調査

ウェブ上の健康情報に対する認識やオンラインの健康情報を求める場面について様々な研究が行われている。Hesse らは、ユーザは医者に行きたいと考えているが実際に最初に訪れるのはインターネットで健康情報に触れていることを明らかにした。また、どのような属性のユーザがインターネット上の健康情報を信頼しやすいかについても明らかにした [9]。

Song らは、アメリカ、韓国、香港でのオンラインの健康情報に対する認識の違いについて調査を行った。結果として、韓国、香港ではブログや SNS などの経験をベースとした記事を信じている傾向がみられた [19]。

Zulman らは、高齢者のインターネットに対する不信感とオンラインの健康情報を探す行動にどのような影響を与えるのかについて調査を行っている。調査の結果から、インターネットの

健康情報に対する信頼度が高いユーザであれば実際にインターネットで得た情報をきっかけに自身の行動を変化させたり、オンラインの健康情報について医者に相談していることが報告されている [25]。

### 2.3 オンラインの健康情報の信頼性

Sbaffi らは、ユーザはウェブページのレイアウト、ウェブページのコンテンツの読みやすさ、ウェブページの運営者など様々な要素を基準にして、ある健康情報が信頼できるかどうかについての判断を行っていることを明らかにした [17]。

Eysenbach らは、インターネット上での健康情報を探す際のユーザの検索や健康情報に対する評価について調査を行っている。調査の結果から、ユーザは情報の出典やウェブサイトのデザインや著者の写真の有無などをもとに情報の信頼性を評価している [6]。

## 3 健康情報を調べる検索タスク

本節では、ユーザ実験の手順について説明する。ユーザ実験で行ったタスク前アンケート、検索タスク、タスク後アンケート、インタビュー、実験後アンケートについて詳細に述べる。

### 3.1 実験手順

リサーチクエスションを明らかにするため、ユーザ実験を行った。なお、ユーザ実験を行うにあたり、兵庫県立大学応用情報科学研究科研究倫理委員会の承認を得た（承認番号:UHGSAL-2021-04）。

まず、実験についての説明をユーザに行った。その際に、実験中に収集するデータや収集したデータを学術目的に利用することも説明した。そして、実験に同意したユーザには同意書に署名してもらった。その後の実験の流れを図 1 に示す。著者らが用意した検索システムに慣れるために訓練タスクを行った。訓練タスクでは、「慢性腎臓病成人患者に対する運動トレーニング」<sup>1</sup> というトピックについての検索を行った。その後、ユーザにタスク前アンケートを回答してもらった。タスク前アンケート後に、ユーザは 2 種類の検索タスクに取り組んだ。検索タスク後に、ユーザにタスク後アンケートを回答してもらった。なお、2 種類の検索タスクを行う順序は、ユーザ間でバランスをとった。2 つの検索タスク実施後、休憩を挟んだのち、それぞれの検索タスクに関するインタビュー調査を行った。また、それぞれの検索タスクに関するインタビューの間に、休憩を挟んだ。インタビュー調査終了後、ユーザに、デモグラフィック属性、学歴、検索の専門性 [22]、eヘルスリテラシー [14] に関する実験後アンケートに回答してもらい、実験を終了した。検索専門性のアンケートは、山本らの研究を参考にして作成した [28]。eヘルスリテラシーのアンケートは、光武らの研究を参考にして作成した [26]。

### 3.2 実験参加者

著者らが所属する大学の大学院生および学部生を対象に実験

1 : <https://www.cochrane.org/ja/CD003236>

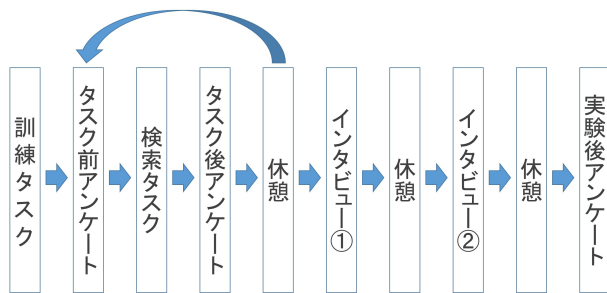


図 1 ユーザ実験の流れ

参加者を募集し 14 名の実験参加者が集まった。2021 年 11 月 18 日から 2021 年 12 月 16 日にかけて、ユーザ実験を行った。実験の開始から終了までは約 2 時間であった。実験を終了した 14 名の実験参加者には、1,940 円を支払った。14 名の実験参加者（男性 1 名、女性 13 名）の全てのデータを分析に用いた。どのユーザのデータであるか把握するために、各実験参加者に ID1 から ID14 を割り当てた。

### 3.3 検索タスクにおけるトピックの選定

本節では、検索タスクで用いたトピックの選び方と選んだトピックについて述べる。

トピックは Cochrane<sup>2</sup>に掲載されている「高齢者の認知症予防に対する魚油の効果」というようなある症状に対する治療法に関するトピックを対象とした。特に、Cochrane の補完代替医療<sup>3</sup>から選んだ。そして、症状に対する治療法の効果があるとされているものを 10 件、症状に対する治療法の効果がないとされているものを 10 件、計 20 件を選んだ。選んだ 20 件の中から、ユーザのトピックに対する事前の考えが偏っていないトピックを選ぶためにアンケート調査を行った。事前の考えが偏っていないトピックを選んだ理由は、ユーザが健康情報を検索する際に事前の考えが偏っていると検索行動に違いがみられないと考えたからである。

トピックに対する事前の考えの偏りを調査するアンケートは、クラウドソーシングサービスである Lancers<sup>4</sup>を使用した。アンケート調査では 100 人の実験参加者を集めた。実験参加者には 100 円の報酬を支払った。アンケートでは、「不眠症に対する音楽は効果的だと思いますか」という質問文と症状と治療法についての説明を付け加えてユーザに提示した。回答は 5 段階リッカート尺度を用いた（1：そう思わない、5：そう思う）。ユーザの回答を分析し、平均が 3 に近く、そう思うとそう思わないという回答の割合が半数以下であるトピックを対象とした。対象としたトピックの中から検索タスクで扱うトピックを選んだ。検索タスクで扱うトピックは表 1 に記載している、「高齢者での認知症予防に対する魚油」と「がんの予防と緑茶」を選んだ。以後、「高齢者での認知症予防に対する魚油」という健康情報を調べる検索タスクは「魚油タスク」とする。また、「がんの

表 1 実験で用いた検索タスク。Cochrane ID はコクランの記事の ID である。

タスク名	治療法	症状	Cochrane ID
魚油タスク	魚油	認知症予防	CD005379
緑茶タスク	緑茶	がん予防	CD005004

予防と緑茶」という健康情報を調べる検索タスクは「緑茶タスク」とする。

### 3.4 タスク前アンケート

検索タスクを行う前に、調べる症状についての興味関心や事前知識があるのか、ユーザの検索タスクに対する事前の考えが検索行動に影響を与えると考えた。これら 3 つの項目を調査するタスク前アンケートを行った。

魚油タスクにおけるタスク前アンケートでは、まず症状に対する興味関心について質問した。興味関心の有無を問う質問では、「あなたは認知症の予防についてどの程度興味関心がありますか」と質問した。質問の回答には、5 段階のリッカート尺度を用いた（1：全く関心がない、5：かなり関心がある）。事前知識を問う質問では、「認知症の予防についてどの程度知識があると思いますか」と質問した。質問の回答には 5 段階リッカート尺度を用いた（1：まったく知識がない、5：かなり知識がある）。最後に、ユーザの事前の考えを問う質問として「魚油は認知症の予防に効果的だと思いますか。」と質問した。質問の回答には、5 段階リッカート尺度を用いた（1：そう思わない、5：そう思う）。緑茶タスクでも同様の質問をした。

### 3.5 検索タスク

検索タスクでの詳細について述べる。検索タスクに取り組むにあたり、ユーザには以下に示すタスク文を読んでもらい、その後検索タスクに取り組んでもらった。タスク文では、Pothiratanachaikul らの研究を参考に治療法と症状についての簡易的な説明文と健康情報を調べる状況についての説明を行った [16]。検索タスクに対しての背景を知ってもらうために、このような説明を行った。

例えば、魚油タスクにおける、治療法の説明には、日本大百科全書<sup>5</sup>の内容を参考に作成した。また、症状の説明には、MSD マニュアル<sup>6</sup>を参考に作成した。ユーザに以下のタスク文を提示した。

魚油とは、イワシやサバなどから得られる脂肪油。  
認知症とは、記憶、思考、判断、学習能力などの精神機能が、ゆっくりと進行性に低下していく病気です。  
あなたは現在高齢者で、認知症にならないように気をつけています。そのような状況のなか、魚油が認知症の予防に効果があるらしいという噂を耳にしました。  
そこで、このタスクでは、ウェブ検索を用いて、魚油が認知症の予防に効果があるかどうかを自由に調べて

2 : <https://www.cochrane.org/ja/evidence>

3 : <https://cam.cochrane.org/cochrane-reviews-related-complementary-medicine>

4 : <https://www.lancers.jp/>

5 : <https://kotobank.jp/word/魚油-53502>

6 : <https://www.msdmanuals.com/ja-jp/ホーム/09-脳、脊髄、末梢神経の病気/せん妄と認知症/認知症>



図 2 実験で用いた検索システム

ください。

検索タスクに取り組む時間は最大 15 分とした。検索タスクにおいては、著者らが用意した検索システムを使用してもらった。ユーザが使用した検索システムの画面は図 2 になる。検索システムでは Microsoft Bing Web Search API を用いて、入力した検索クエリに対応する検索結果を表示した。

### 3.6 タスク後アンケート

本節では、ユーザが意見を形成した理由とウェブページに対する信憑性判断を調査するために行った、タスク後アンケートについて述べる。

#### 3.6.1 トピックに対するユーザの考え

ユーザは、調べたトピックについてどのように考えたかを選択肢の中から選んだ。そして、その考えになった理由について自由記述に記入してもらった。

まず、検索タスクについてどのような考えになったかを問う質問では、「症状に対する治療法の効果はあると思いましたか」と質問した。回答は「症状に対する治療法の効果はある」、「症状に対する治療法の効果はない」、「今回の検索ではわからなかった」の 3 つの選択肢の中から選んでもらった。そして、その選択肢を選んだ理由を自由記述欄に記入してもらった。自由記述欄には「複数の組織で言っていたため」という文を例示した。

検索したトピックについての意見を調査した後に、検索したトピックに対しての意見が反対の意見にならなかった理由についても質問した。ある症状に対する治療法の効果があると考えたユーザには「なぜ、ある症状に対する治療法の効果はないとなかったのかその理由を教えてください」と質問した。自由記述欄には「複数の組織で言っていなかったため」という文を例示した。例示した理由としては、意見を形成した理由をどのように記述してほしいかを伝えるようにしたからである。

#### 3.6.2 閲覧したウェブページの評価

ユーザがウェブページの信憑性判断をどのように行っているのか調査するために、検索タスク中に閲覧した全てのウェブページを評価してもらった。ユーザに質問した項目は 6 つある。質問項目については、表 2 になる。これらの質問項目でどのよ

うに回答してもらったか述べる。Q1 の回答には、5 段階リッカート尺度を用いた（1：まったく見ていない，5：かなり見ていた）。この設問を設けた理由は、ユーザがまったく見ていないと回答したウェブページを今回の分析するデータから除くためである。Q2 の回答には、「治療法は症状に効果的である」、「治療法は症状に効果的でない」、「治療法は症状に効果的であるか不明である」、「関係のない内容だった」の 4 つの選択肢からユーザに選んでもらった。Q3 の回答には、「症状と治療法の両方に言及している」、「症状についてのみ言及している」、「治療法についてのみ言及している」、「関係のない内容であった」の 4 つの選択肢からユーザに選んでもらった。Q4 の回答には、7 段階リッカート尺度を用いた（1：まったく信頼できない，7：非常に信頼できる）。Q5 の回答には、自由記述で回答してもらった。自由記述欄には「病院が情報提供していたため」という文を例示した。例示した理由としては、どのように記述してほしいかを伝えるようにしたからである。Q6 の回答には、7 段階リッカート尺度を用いた（1：非常に参考にならなかった，7：非常に参考になった）。

### 3.7 インタビュー

ウェブページに対する信憑性判断や意見の形成についてより詳細に知るために、半構造化インタビューを行った。インタビューではあらかじめ 6 つの質問を準備した。また、3.6.1 節のユーザが意見を形成した理由と 3.6.2 節の Q5 でウェブページが信頼できるかどうかを評価した理由を参考に質問内容を変化させた。インタビューの内容は全て録音した。

最初に、ウェブページが信頼できるかどうかを評価した理由にみられる評価基準について質問した。ウェブページが信頼できるかどうかを評価した理由に、「国立の研究センターの記事であったから」と回答していたユーザには、「なぜ国立の研究センターという組織が信頼できる理由として考えたのですか」と質問した。

2 つ目に、ウェブページが信頼できるかどうかを評価した理由に著者、サイト運営者、日付、引用のいずれかを評価していないユーザに質問した。ウェブページが信頼できるかどうかを評価した理由に著者がわかるから信頼できると書いていた場合には「なぜ、サイト運営者、日付、引用などの基準が信頼できる理由が含まれていないのですか」と評価した理由にない評価基準について質問した。

3 つ目に、ウェブページが信頼できるかどうかを評価した理由と意見を形成した理由を比較し、ウェブページが信頼できるかどうかを評価した理由のみにみられた基準について質問した。ウェブページが信頼できるかどうかを評価した理由に著者が明らかであったとしたと記述したユーザが、意見を形成した理由に著者が明らかであることを評価していない場合には、以下のような質問をした。「ウェブページを信頼するかどうかを評価した理由に著者をあげられているが、調べたトピックについての意見を形成した理由に著者が明らかであるという記述が見られない理由はなにかありますか」と質問した。

4 つ目に、ユーザが意見を形成する際に、どのような点を評

表 2 ユーザが閲覧したウェブページに対する質問項目

ID	質問項目
Q1	閲覧したウェブページをどの程度見ていましたか
Q2	閲覧したウェブページでは治療法と症状に言及していましたか
Q3	閲覧したウェブページの内容は以下のどれにあてはまっていますか
Q4	閲覧したウェブページをどの程度信頼できると思いましたか
Q5	Q4 の項目でなぜそのような選択をしましたか、理由を教えてください
Q6	閲覧したウェブページはどの程度参考になりましたか

価しているのかについて質問した。「厚労省が紹介しているなどの点で信用できそうな研究所が効果的だと述べていたから」という理由を自由記述欄に書いたユーザには、「調べたトピックについての意見を形成する際の理由に厚労省や研究所という組織をあげていますが、組織のどのようなところを評価していますか。」と質問した。

5 つ目に、ユーザが意見を形成するときに、どのようなことを考えているのかを知るための質問をした。ユーザに「今回のように、調べている情報に様々な意見があるとき、この意見が正しいのではないかと思うときにどのような考え方をしていますか」と質問した。

最後に、ユーザに意見を形成する際に影響を与えているものについて知るための質問をした。例えば、ユーザに「意見を形成するときに、影響を及ぼすものはなにかありますか」と質問した。

それぞれの質問において、ユーザからの回答に応じて意見の形成について明らかにするために、追加で質問を行うことがあった。

## 4 意見の形成と信憑性判断についての分析

ユーザ実験で収集したデータや録音したインタビューの内容などを分析した結果について述べる。

### 4.1 タスク前アンケートの分析

ユーザの検索タスクに対する事前の考えを調査したタスク前アンケートの結果を分析した。ユーザの事前の考えの項目で、魚油タスクでは平均 2.57 で分散 0.57 であった。緑茶タスクでは、平均 2.71 で分散 0.68 であった。両方のタスクでの平均が 3 に近いことから、著者らが意図した事前の考えに偏りが無いトピックを選べた。

### 4.2 ユーザ実験で収集した自由記述に対するラベル付け

信憑性判断をどのように行っているのかを分析するためにラベル付けを行った。ラベル付けを行ったデータは 3 つある。1 つ目は、3.6.1 節でユーザが意見を形成した理由を回答した文である。このデータのことを、本研究では意見形成データとしている。2 つ目は、3.6.1 節でユーザが反対の意見を形成しなかった理由を回答した文である。3 つ目は、3.6.2 節の Q5 でウェブページが信頼できるかどうかを評価した文である。このデータのことを、本研究ではページ評価データとしている。1 つ目のデータと 2 つ目のデータでは、ラベル付けを行う文に句

点がある場合、1 文ずつ評価した。

3 つの先行研究をもとに著者らがラベルを作成した [20] [13] [11]。表 3 に作成したラベルと定義を示す。ラベル付けについては、著者と同じ大学に所属する 2 人の計 3 人で個別にラベル付けを行った。評価する文にラベルの定義を満たす記述があるかないかの 2 値で評価した。複数のラベルを付与できる場合には複数のラベルを付与した。最終的なラベルは多数決で決定した。1 人でもいずれかのラベルを付与していた場合には議論を行い決定した。また、いずれのラベルの定義を満たしていない場合は、ラベルを付与しなかった。3 人の評価者のページ評価データの各ラベルごとの Fleiss [8] の Kappa 係数を求めた。その他を除くラベルの Kappa 係数を求めた。魚油タスクでの Kappa 係数の最小は専門性で 0.37 であった。最大は鮮度で 0.91 であった。平均は 0.70 であった。緑茶タスクでの Kappa 係数の最小は専門性で 0.33 であった。最大は人気、鮮度、両面提示で 1.00 であった。平均は 0.74 であった。

### 4.3 意見の形成についての分析

本研究の RQ1 である、複数の情報源から自身の意見を形成する際に、どのように情報を集約して意見を形成しているのかについての解答を明らかにする。

著者らがインタビューでの発話を分類し、意見の形成について分析した。分析の結果、少なくとも以下の 3 つのパターンがみられた。

- 多数決型
- 単一ページ確信型
- 事前信念先行型

#### 4.3.1 多数決型

閲覧したウェブページで主張している意見の分布を参考にする「多数決型」について述べる。インタビューで以下のような発話を得られた。

“信頼できるようなサイトでも効果があると言っている。そうではないサイトでも効果があると言っていたから。” (ID1)

このように、多数決型ユーザは閲覧したウェブページで主張している意見の分布を重要視していた。多数決型ユーザは閲覧したウェブページでどのような主張が多くみられたかを印象でとらえていた。

表 3 自由記述に付与したラベル

ラベル	定義
専門性	著者は情報を適切に扱うことができる十分な知識を持っているかどうか
権威性	情報を提供している組織が社会的に信用されているかどうか
客観性	情報が個人の考えや商業的な意図がない事実に基づいているかどうか
人気	情報が人気の情報であるかどうか
鮮度	新しい情報であるかどうか
詳細性	情報が詳細な情報であるかどうか
網羅性	情報が幅広い情報であるかどうか
両面提示	情報がメリットとデメリット両方を述べているかどうか
デザイン性	インターフェースの外観が視覚的に好みであるかどうか
了解性	内容がわかりやすいかどうか
断定の有無	治療法の効果の有無が明確に述べられているかどうか
その他	上記のラベルに当てはまらないものが基準によって、評価している

表 4 意見形成した理由の分析結果

	魚油タスク	緑茶タスク
専門性	2	2
権威性	3	3
客観性	7	9
人気	0	1
鮮度	0	0
詳細性	1	1
網羅性	1	1
両面提示	2	7
デザイン性	0	2
了解性	0	0
断言の有無	3	4
その他	6	6

#### 4.3.2 単一ページ確信型

最も信頼できるウェブページの情報をもとに意見を形成する「単一ページ確信型」について述べる。この意見の形成を行っていたユーザは、インタビューで以下のような発言が得られた。

“最終的には1個（1個のウェブページ）だけですが、ほかにも論文内容もちょっとだけは見たんですけどかなり1個（1個のウェブページ）です。”（ID9）

このように、閲覧したウェブページの中から最も信頼できるウェブページはどれであるか考え、最も信頼できるウェブページの内容から意見を形成していた。

#### 4.3.3 事前信念先行型

最後に、検索タスクに対する事前の考えを重視する事前信念先行型について述べる。インタビューで、以下のような発言が得られた。

“私は結構自分の主観に引っ張られやすいタイプなので、書いてあること、とりあえず、否定的なものから信じてしまう傾向がなんとなくあるとおもいます。”（ID11）

以上のように事前の信念に基づいて検索を行っていることが伺える。事前の信念に影響を受けて検索を行うことは、

Pothirattanachaikul らが明らかにしており同様の行動がみられた [16]。

以上の分析より RQ1 の解答は、ユーザが情報を集約する際に、多数決型、単一ページ確信型、事前信念先行型の少なくとも3つのパターンが存在している。

#### 4.4 意見の形成に用いられる信憑性判断の分析

ユーザが意見を形成する際に、どのような信憑性判断をしているのか分析する。まず、タスクごとの分析を行うために、4.2 節の意見形成データを魚油タスクと緑茶タスクのデータにわけた。そして、各タスクのデータから14名のユーザのデータに分割し、28個のデータを作成した。この28個のデータに対して、表3のラベルが存在しているかどうかを評価した。評価方法は、データ中の各ラベルに注目しラベルが1つ以上存在している場合には1とし、ラベルが1つも存在していない場合には0とした。各ユーザのデータに存在しているラベルの種類を評価したデータを作成した。このデータを用いて、各タスクのラベルごとの合計を算出した。その分析結果を表4に示す。表4より両方のタスクにおいて、客観性を多くのユーザが評価していた。実際に意見を形成した理由の一部に、「証拠が不十分であったため」という記述や「実験結果が掲載されている」という記述がみられた。

#### 4.5 ウェブページに対する信憑性判断の分析

ウェブページに対する信憑性判断の分析方法と分析結果について述べる。4.2 節のページ評価データで分析を行った。ページ評価データは、「国立がんセンターの記事だったから」といった文を含むデータである。

まず、ユーザがウェブページを信頼するかどうかを記述した理由の全てを対象に分析した。タスクごとの分析を行うために、ページ評価データを魚油タスクと緑茶タスクのデータにわけた。最初に、各タスクでラベルの定義を満たしていると評価された数を分析した。各タスクのデータでラベルごとに、定義を満たしていると評価された数の合計を算出した。2つ目の分析では、ラベルの定義を満たしていると評価された値があるユーザの数を分析した。4.4 節の分析と同様の分析をページ評価データで行った。



表 5 ウェブページに対する信憑性判断の分析結果. ( ) 内は実験参加者 14 人のうち該当のラベルを評価していた人数を表す.  $P$  は評価されたページ数を表す.

ページ数	魚油タスク			緑茶タスク		
	全ウェブページ	信頼できる	信頼できない	全ウェブページ	信頼できる	信頼できない
	$P = 93$	$P = 67$	$P = 11$	$P = 93$	$P = 60$	$P = 20$
専門性	35 (11)	29 (10)	2 (2)	23 (10)	18 (9)	4 (4)
権威性	39 (13)	35 (13)	2 (2)	40 (12)	31 (12)	7 (7)
客観性	26 (12)	20 (11)	3 (2)	33 (13)	19 (11)	9 (6)
人気	2 (1)	2 (1)	0 (0)	3 (2)	3 (2)	0 (0)
鮮度	4 (4)	2 (2)	1 (1)	1 (1)	1 (1)	0 (0)
詳細性	4 (3)	4 (3)	0 (0)	15 (9)	9 (7)	3 (3)
網羅性	3 (3)	3 (3)	0 (0)	5 (4)	3 (2)	2 (2)
両面提示	2 (2)	0 (0)	1 (1)	5 (4)	5 (4)	0 (0)
デザイン性	10 (8)	4 (3)	2 (2)	11 (7)	3 (2)	7 (6)
了解性	5 (4)	3 (3)	1 (1)	4 (3)	3 (2)	0 (0)
断言の有無	7 (4)	5 (3)	1 (1)	3 (3)	1 (1)	1 (1)
その他	16 (9)	5 (4)	4 (3)	10 (6)	1 (1)	3 (2)

分析の結果が表 5 になる. 表 5 の全ウェブページの項目に注目すると, 両方のタスクで, 専門性, 権威性, 客観性を多くのユーザが評価していた. 以上から, ユーザは信憑性がある情報かどうか判断する際に, 著者の知識, 社会的に信頼できる組織であること, 提供している情報に実験といった証拠があることを重要視していることがわかった.

次に, 表 5 にあるユーザが信頼できると判断したウェブページのみ記述された理由を調査した. 3.6.2 節の Q4 で回答でユーザがやや信頼できる, かなり信頼できる, 非常に信頼できると評価したウェブページを信頼できるウェブページとした. ページ評価データから, 信頼できるウェブページのデータを対象とし, 分析した.

分析では, 4.5 節と同様の分析を行いラベルの数とユーザの数を算出した. 分析結果は, 表 5 になる. 表 5 より, 両方のタスクで多くのユーザが, 信頼できるウェブページと判断するときには, 専門性, 権威性, 客観性を評価していることがわかった.

最後に, 表 5 にあるユーザが信頼できないと判断したウェブページのみ記述された理由を調査した. 3.6.2 節の Q4 でユーザがやや信頼できない, かなり信頼できない, 全く信頼できないと評価したウェブページを信頼できないウェブページとした. ページ評価データから, 信頼できないウェブページのデータを対象とし, 分析した.

分析では, 4.5 節と同様の分析を行いラベルの数とユーザの数を算出した. 表 5 の信頼できないウェブページのユーザ単位での分析をみると, 緑茶タスクで半数のユーザが権威性に注目している. ユーザは, 信頼しないウェブページを判断するのに権威性を用いていることが示唆される.

#### 4.6 ウェブページに対する信憑性判断と意見の形成との関係

インタビューでウェブページに対する信憑性判断と意見の形成との関係について調査した. 以下のような発話を得られた.

“論文で明確に予防できてるっていう結果が出たら,

信頼すると思います.” (ID10)

実験参加者の 14 人中 12 人が, 意見を形成する際に, 信憑性があると判断した情報を重要視していた.

以上から RQ2 の解答は, ユーザは意見を形成するときに用いられている信憑性判断は, 重要視する情報か否かを見分けている.

## 5 議 論

多数決型, 単一ページ確信型では, 信頼できるウェブページが意見の形成に影響を与えている. ウェブページが正確な情報を提供しているかどうかをユーザ自身は見極めなければならない. しかし, ユーザが獲得したい情報を得られたと考えていても, 正確な情報ではないことが明らかになっている [12].

また, ユーザの検索能力によって閲覧することのできるウェブページが限定されてしまうことが考えられる. 例えば, 主張が異なる複数のウェブページを閲覧することのできない検索クエリで検索をすることがある. この場合に, ユーザは同様の主張をしているウェブページの内容をもとに意見を形成することがある. そして, 異なる主張をしているウェブページの内容に目を触れることなく偏った情報をもとに意見を形成してしまう可能性がある. 意見を形成する以前にどのようなウェブページをユーザが獲得しているのかも, 重要である.

このように, ユーザの意見の形成に至るまでに, 検索クエリ, 検索結果, 閲覧したウェブページなど様々な要因が影響している. ユーザに慎重な検索を促すには, 検索から意見の形成に至るまでを支援するシステムの開発が望まれる.

## 6 ま と め

本研究では, ウェブ検索エンジンを用いて健康情報検索を行う際に, どのような情報に信憑性を感じるのか調査した. ユーザに健康情報を調べてもらう検索タスクを行った. まず, ウェブページに対する信憑性判断について調査した. 分析結果から,

多くのユーザがウェブページに対する信憑性判断に権威性、客観性、専門性で評価していることがわかった。次に、ユーザが健康情報を取得したのちに、どのように意見を形成しているのか調査した。多数決型、単一ページ確信型、事前信念先行型の少なくとも3つのパターンのいずれかで情報を集約し意見を形成していることがわかった。また、ユーザが意見を形成するときに、複数の情報源から重要視する情報か否かを見分けるのに信憑性判断を用いていた。

## 謝 辞

本研究は JSPS 科研費 JP21H03775, JP21H03774, JP18H03494, JP21H03554, JP18H03244 の助成を受けたものです。ここに記して謝意を表します。

## 文 献

- [1] C. Boyer, M. Selby, J.-R. Scherrer, and R. D. Appel. The health on the net code of conduct for medical and health websites. *Computers in Biology and Medicine*, 28(5):603–610, 1998.
- [2] D. Charnock, S. Shepperd, G. Needham, and R. Gann. DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology & Community Health*, 53(2):105–111, 1999.
- [3] X. Chen, J. L. Hay, E. A. Waters, M. T. Kiviniemi, C. Biddle, E. Schofield, Y. Li, K. Kaphingst, and H. Orom. Health literacy and use and trust in health information. *Journal of Health Communication*, 23(8):724–734, 2018.
- [4] Y. Chi, D. He, S. Han, and J. Jiang. What sources to rely on: Laypeople's source selection in online health information seeking. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*, pp. 233–236, 2018.
- [5] J. Y. Cuan-Baltazar, M. J. Muñoz-Perez, C. Robledo-Vega, M. F. Pérez-Zepeda, and E. Soto-Vega. Misinformation of covid-19 on the internet: Infodemiology study. *Journal of Medical Internet Research Public Health and Surveillance*, 6(2):e18444, 2020.
- [6] G. Eysenbach and C. Köhler. How do consumers search for and appraise health information on the world wide web? qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ*, 324(7337):573–577, 2002.
- [7] M. A. Feufel and F. Stahl. What do web-use skill differences imply for online health information searches? *Journal of Medical Internet Research*, 14(3):e87, 2012.
- [8] J. L. Fleiss. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382, 1971.
- [9] B. W. Hesse, D. E. Nelson, G. L. Kreps, R. T. Croyle, N. K. Arora, B. K. Rimer, and K. Viswanath. Trust and sources of health information: The impact of the internet and its implications for health care providers: Findings from the first health information national trends survey. *Archives of Internal Medicine*, 165(22):2618–2624, 2005.
- [10] C. I. Hovland, I. L. Janis, and H. H. Kelly. *Communication and Persuasion*. Yale University Press, New Haven, CT, 1953. (辻正三, 今井省吾 (訳). コミュニケーションと説得, 誠信書房, 1960.)
- [11] J. Kapoun. Teaching undergrads web evaluation: A guide for library instruction. *College & Research Libraries News*, 59(7):522–523, 1998.
- [12] M. Liu, Y. Liu, J. Mao, C. Luo, M. Zhang, and S. 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*, pp. 1533–1542, 2018.
- [13] M. J. Metzger, A. J. Flanagin, and L. Zwarun. College student web use, perceptions of information credibility, and verification behavior. *Computers & Education*, 41(3):271–290, 2003.
- [14] C. D. Norman and H. A. Skinner. eHEALS: the eHealth literacy scale. *Journal of Medical Internet Research*, 8(4):e27, 2006.
- [15] G. Peterson, P. Aslani, and K. A. Williams. How do consumers search for and appraise information on medicines on the internet? a qualitative study using focus groups. *Journal of Medical Internet Research*, 5(4):e33, 2003.
- [16] S. Pothirattanachaikul, T. Yamamoto, Y. Yamamoto, and M. Yoshikawa. Analyzing the effects of document's opinion and credibility on search behaviors and belief dynamics. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pp. 1653–1662, 2019.
- [17] L. Sbaffi and J. Rowley. Trust and credibility in web-based health information: A review and agenda for future research. *Journal of Medical Internet Research*, 19(6):e218, 2017.
- [18] W. M. Silberg, G. D. Lundberg, and R. A. Musacchio. Assessing, controlling, and assuring the quality of medical information on the internet. *JAMA*, 277(15):1244–1245, 1997.
- [19] H. Song, K. Omori, J. Kim, K. E. Tenzeck, J. Morey Hawkins, W.-Y. Lin, Y.-C. Kim, and J.-Y. Jung. Trusting social media as a source of health information: Online surveys comparing the united states, korea, and hong kong. *Journal of Medical Internet Research*, 18(5):e25, 2016.
- [20] Y. Sun, Y. Zhang, J. Gwizdka, and C. B. Trace. Consumer evaluation of the quality of online health information: Systematic literature review of relevant criteria and indicators. *Journal of Medical Internet Research*, 21(5):e12522, 2019.
- [21] H. Tatsumi, H. Mitani, Y. Haruki, and Y. Ogushi. Internet medical usage in japan: Current situation and issues. *Journal of Medical Internet Research*, 3(1):e12, 2001.
- [22] R. W. White and D. Morris. Investigating the querying and browsing behavior of advanced search engine users. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 255–262, 2007.
- [23] Y. Zhang. Consumer health information searching process in real life settings. In *Proceedings of the American Society for Information Science and Technology*, Vol. 49, pp. 1–10, 2013.
- [24] Y. Zhao and J. Zhang. Consumer health information seeking in social media: A literature review. *Health Information & Libraries Journal*, 34(4):268–283, 2017.
- [25] D. M. Zulman, M. Kirch, K. Zheng, and L. C. An. Trust in the internet as a health resource among older adults: Analysis of data from a nationally representative survey. *Journal of Medical Internet Research*, 13(1):e19, 2011.
- [26] 光武誠吾, 柴田愛, 石井香織, 岡崎勘造, 岡浩一朗. eHealth literacy scale(eHEALS) 日本語版の開発. 日本公衛誌, 58(5):361–371, 2011.
- [27] 厚生労働省. 平成 26 年版厚生労働白書 健康長寿社会の実現に向けて ～健康・予防元年～第 1 部第 2 章「健康をめぐる状況と意識」. 平成 26 年度厚生労働白書, 2014.
- [28] 山本祐輔, 山本岳洋, 大島裕明, 川上浩司. ウェブアクセスリテラシー尺度の開発. 情報処理学会論文誌 データベース, 12(1):24–37, 2019.
- [29] 大島裕明, 山本祐輔, 山家雄介, 高橋良平, ヤトフトアダム, 中村聡史, 田中克己. Web 情報の信憑性 (<特集>情報の信頼性). 情報の科学と技術, 61(1):2–7, 2011.