

化粧品クチコミに特化した信憑性評価システムの実現とその評価

濱野花莉¹ 梶田美帆¹ 中村聡史¹

概要：肌質や化粧の仕上がりの好みは人によって異なることもあり、多くの消費者の需要に応じて多種多様な化粧品が開発されている。ここで、化粧品の選択を誤ると肌荒れや仕上がりの悪さに繋がるため、化粧品の購入に失敗したくないと考える消費者は多く、インターネット上のクチコミを参考にすることがある。しかし、化粧品のクチコミは実際の利用を確認できないなど信憑性において様々な問題を抱えている。そこで本研究では、これまでに検討した信憑性の評価軸を利用し、化粧品に特化したクチコミ信憑性評価システムを開発する。また、このシステムをユーザに使用してもらうことで、システムを利用した際に信憑性への意識の変化を調査し、システムの改善点について検討する。その結果、元々信憑性を意識していなかった実験参加者が信憑性の提示により信憑性を意識するようになることや信憑性提示の見方にも個人差が観察されたため、その個人差に対応可能なシステムにする必要があることが明らかになった。

キーワード：化粧品、クチコミ、コスメ、レビュー、レビュー分析、信憑性

1. はじめに

化粧は多くの人にとって日常的なものである。ポーラ文化研究所が2021年に行った調査[1]によると、女性の81.8%、男性の16.6%が化粧を行っており、スキンケアは女性の93.1%、男性の55.5%が行っている。多くの消費者の需要に応じて、化粧品メーカーは様々な特色をもった商品を研究開発している。例えばファンデーションの場合、シミや毛穴を隠すカバー力を重視したもの、皮脂や時間経過による崩れに強いもの、自然由来や肌に優しい成分にこだわったものなど、多種多様な商品が存在する。

ここで、化粧品は継続的に直接肌につけるものであるため、自身の肌に合わないものを使用すると肌荒れを引き起こすことがある。また値段も安くはないため、購入する化粧品の選択に失敗したくないと考える消費者は多い。しかし、株式会社M2が2019年に行った調査[2]の「ファンデーション選びで失敗したことはありますか?」という質問に対して、65.9%の人が「はい」と回答しているなど、化粧品選びは容易でないことがわかる。

化粧品の購入に失敗しないために、情報収集は重要である。化粧品に関する情報を得る方法のひとつとして、SNSやクチコミサイトなどでクチコミを閲覧することが挙げられ、我々の1547人を対象とした調査でもSNSを32.4%、クチコミサイトを47.6%の人が購入時の情報源として利用していた。こうしたSNSやクチコミサイトでは、多くのユーザの意見を手軽に閲覧できるという利点がある一方で、誰でも書き込むことができるため、質の低いクチコミやステルスマーケティングによるクチコミが存在するという問題もある。質の低いクチコミの具体例としては、専門知識をもたないユーザによる正しい知識に基づかないクチコミ、SNSで拡散されたいと考えるユーザによる誇張表現を含む

クチコミなどがある。こうした背景から、化粧品の購入に失敗しないために、SNSやクチコミサイトに投稿されたクチコミの信憑性をユーザ自身で見極める必要がある。しかし我々の研究[3]において行った1547名を対象としたアンケート調査では、「過去に口コミを参考にして失敗した経験があれば回答してください」という質問に対して多くの経験談が寄せられており、クチコミからの情報収集も容易でないことがわかる。

これまでクチコミの信憑性の判断を支援する研究[4]やシステムの開発[5]などが行われてきているが、これらは化粧品に特化したものではない。例えば化粧品は化粧下地の上にファンデーションを塗るなど複数の化粧品を同じ部位に併せて使うために、ファンデーションのクチコミであっても、その効果がファンデーションによるものなのか、併用した化粧下地によるものなのか判断ができないといった問題が起きる。また、化粧品は消費者の肌質や好みといった個人差によってクチコミの意見が大きく変わることで、投稿者が本当にその化粧品をつかっているのか、どの程度使っているのかなどクチコミだけでは判断できないことなどの特性がある。さらに、化粧品は試供品が配布されていることも多く、一度使用しただけの化粧品についてクチコミが投稿されることも珍しくない。

このように、化粧品のクチコミには特有の特性があるため、特に信憑性の判断を支援するシステムが必要である。しかしクチコミサイトやSNS上のクチコミが真実かどうかを完璧に判断することは現実的ではない。こうした問題をふまえ、我々は、システムによる客観的な信憑性評価に加え、信憑性評価が得意なユーザの評価基準を共有することで、信憑性の評価が苦手なユーザを支援するシステムの実現を目指し、化粧品のクチコミの信憑性評価軸の検討[3]やその評価軸について有用性の検証[6]を行ってきた。しかし、

¹ 明治大学
Meiji University

システムとしては実装していなかったため、実際に利用するとどうなのかを明らかにできておらず、またどういった情報が重要かも明らかにできていなかった。

そこで本稿では、これまで検討してきた信憑性評価軸を用いたシステムを実装し、システムを用いた評価実験を行う。そのために、まずシステムに使用するクチコミについて軸評価を行い、信憑性を推定した結果をユーザに提示するシステムを実装する。また、システムについての評価実験を行うことで、信憑性評価項目や信憑性の提示がどの程度役立つのかを調査し、ユーザからのフィードバックをもとにシステムの改善点を明らかにする。

2. 関連研究

2.1 化粧に関する研究

化粧は多くの人にとって日常的なものであるため、化粧支援に関する研究は数多く存在する。

高木ら[7]は、メイクアップテクニックの説明やユーザが自分の顔に行ったメイクアップに対するアドバイスを出力するシステムを提案している。このシステムでは、顔データを3次元で扱っており、正面からだけでなく横や斜めから見た顔のイメージも考慮できる。また神武ら[8]は、個人の好みを反映したメイクアップ支援のためのシステムを提案している。このシステムは、ユーザの顔画像と好みの顔画像を入力することで、ユーザの肌の色と好みを考慮したメイクアップを顔に施したシミュレーション画像を出力し、そのメイクアップを実現するための商品を推薦してくれる。Chongら[9]は、トランスジェンダーのための化粧支援システムを構築している。このシステムでは、顔写真と認識されたい性別を入力することで化粧のシミュレーションと他人から自分の望む性別として認識される可能性をスコアとして提示する。Kajitaら[10]は、リアルタイムでファンデーションの塗布状態を可視化するシステムの実現を目指し、スマートフォンのカメラで撮影した写真から機械学習を利用してファンデーション塗布画像と素肌画像を82.3%の精度で判別可能であることを明らかにした。このように化粧の支援に関する研究がこれまで多く行われてきた。本研究では化粧品使用時の支援ではなく、購入する化粧品を選ぶ段階からユーザの支援を行うことを目的としている。

2.2 情報の信憑性に関する研究

インターネットの普及により、ウェブ情報の信憑性を評価することは重要な課題となっている。

藤本ら[11]は、災害時など不測の事態においても臨機応変に信憑性の高い情報を収集することを目指し、Twitter上で信頼性の高いユーザを特定する手法を提案した。これは、過去の投稿と情報のリアルタイム性をもとにユーザを分類する手法であり、検証の結果、信頼性の高いユーザを特定できることを明らかにした。またMoranら[12]は、SNS上

のクチコミに着目し、クチコミの信憑性に関する文献の再検討を行ったうえで、オンライン環境におけるクチコミの信憑性を構成する要因をまとめたフレームワーク 4Cs (Community, Competence, Content, Consensus) を提案した。

このフレームワークは、クチコミの分析を支援し、クチコミをより有意義なものにできると考えられる。中島ら[13]は、ブロガーのマニア度を算出するマニア指向プログラミング方式を提案した。この手法を用いることで、情報を閲覧したユーザ自身がその信頼性を判断することが可能になると主張している。Huangら[14]は、誇張された情報や実証されてない情報を含む広告を問題視し、そういった広告を判定するために畳み込みニューラルネットワークを利用したシステムを提案し、高精度に判定できることを示した。山本ら[15][16]は、分析対象のウェブ情報とその関連情報をデータ対として表現し、support 関係を分析することで対象となる情報の信憑性を評価するモデルを提案した。この提案モデルの有効性を確認する評価実験では、適合率65%以上で自動判定が可能であることを明らかにした。また、正確なウェブ情報を収集する能力である「ウェブアクセスリテラシー」を測定する尺度を提案し、この尺度の有効性についても確認している。さらに、情報の信憑性を判断する際、ユーザはウェブページ作成者がどのような人物かということにあまり注意を払わないことも明らかにしている。

本研究は、こうしたインターネット上の情報の中でも化粧品のクチコミに特化して信憑性を判定していくものであるが、これらの手法と組み合わせることでさらに有用なシステムの実現につながると考えている。

2.3 クチコミの分析に関する研究

Leeら[17]は、Instagramにおけるインフルエンサーらのプロモーション投稿について、スポンサー情報の開示・投稿者の信頼性・ブランドの信頼性の観点からプロモーション効果への影響を調査した。その結果、インフルエンサー自身の信頼性はプロモーション効果に影響せず、ブランドの信頼性が影響することを明らかにした。また、Ohら[18]はネット通販のレビューについて、レビューの特徴と有用性の関係を調査した。その結果、レビューの長さ・投稿された時期・肌トラブルについての言及などが有用性に影響を及ぼすことを明らかにした。Rofiantoら[19]は、動画でのクチコミに着目し、開封動画において信頼性や有用性の形成とクチコミ採用率の関係について調査した。その結果、発信者の専門性や信頼度から形成される動画の信頼性は採用率に影響を与えるが、動画内での論証品質がもたらす動画の有用性は採用率に影響しないことを明らかにした。さらに、Mukherjeeら[4]は、Amazonのレビューアグループから偽のレビューアグループを検出することを目指し、ラベル付きデータセットの作成、偽のレビューアグループを判別するための指標や偽のレビューアグループを検出するためのアルゴリズムの提案を行った。この研究では、グルー

ブ単位で偽のレビュアーを検出することは、個々のレビューに注目して偽のレビューを検出するより容易であることも明らかにしている。

しかし、これらの研究は化粧品のクチコミの特性を考慮したものではないため、他のカテゴリとは異なる独自の特性をもつ化粧品のクチコミについて信憑性を評価する際には不十分である可能性がある。本研究では、化粧品のクチコミの特性を踏まえたうえで、信憑性を評価する基準を提案することを目指している。

2.4 マイクロタスクに関する研究

我々が目指すシステムではマイクロタスク的にクチコミの特徴をユーザに評価してもらうことを検討している。ここで、マイクロタスクを用いた研究は様々行われている。

Iwamotoら[20]は、クラウドソーシングにおけるワーカー確保という問題に着目し、床面にタスクを提示することで、通行人がタスクに容易にアクセスできる手法を提案し、データの品質を管理するための対策を検討した。その結果、通行人の観察結果に基づいた分類器によってデータの品質を向上させた。また、三輪ら[21]はマイクロタスクを行う機会の少なさに着目し、参加を促す手法として、音楽ゲームに埋め込むことを提案した。その結果、ゲームを邪魔することなくマイクロタスクへの回答を促すことを可能とした。一方、佐々木ら[22]はクライアントの要望に則したイラストを、マイクロタスクを用いて生成する手法を提案している。この手法では、描画能力を問わない複数人のワーカが少しずつ線を描画することで、一枚の写真からイラストの生成が可能であることを示した。また、三原ら[23]は、マンガのコマ領域の判定を目的として、画像認識とクラウドソーシングを組み合わせたマイクロタスクを実装した。これにより、マンガメタデータ作成にかかる手作業の工数を削減し、従来の計算機のみでのコマ抽出よりも精度を向上させることに成功した。

我々はユーザがクチコミを表示した際、マイクロタスク的に評価軸にまつわる簡単な質問を提示し、その結果を集約・蓄積することで信憑性評価を行うことを検討している。

3. 信憑性判断支援システムの実装

3.1 システム概要

我々が実現を目指す化粧品クチコミの信憑性判断支援システムは、クチコミにおいて他の商品との比較がなされているか、その商品の悪い点も書かれているかといったクチコミに対する各種評価軸に基づき行うものである。また、クチコミに対する軸評価はマイクロタスク化し、膨大なユーザの手で軸の有無などを評価してもらい、そのデータを集約・蓄積し、信憑性評価を行う予定である(図1)。

信憑性評価軸としてはこれまでの研究[3][6]で、「製品のデメリットについての記述の有無」「他の製品との比較の有

無」「それぞれの化粧品において重要な単語(カバー・肌質・発色など)の有無」「購入経緯についての記述の有無」「併用した製品についての記述の有無」「投稿者なりの使用方法(HOWTO)についての記述の有無」が挙げられていた。また、評価軸の検証用データセットを構築する際に得られた意見から「今後のことについての記述の有無」「どのような人におすすめかについての記述の有無」という軸も採用した。そこで、本稿でもこの軸を採用する。

本稿では、これまでに検討した化粧品クチコミの信憑性評価軸を用いた信憑性評価システムを、Vue.jsを用いたWebサービスとして実装した。実際に構築したシステムを図2に示す。画面左半分にクチコミが表示され、その右側に「信憑性評価項目」「信憑性ランク」として評価軸と評価結果(S・A・B・C)を提示する。信憑性評価項目は後述の軸評価の際当てはまると判断した評価者の人数に応じて色の濃さを変化させている(図3, 4)。信憑性評価項目を以下に示す。

- それぞれの化粧品において重要な単語の数
- 製品のデメリットについての記述の有無
- 他の製品との比較の有無
- 購入経緯についての記述の有無
- 併用した製品についての記述の有無
- 投稿者なりの使用方法(HOWTO)についての記述の有無
- 今後のことについての記述の有無
- どのような人におすすめかについての記述の有無

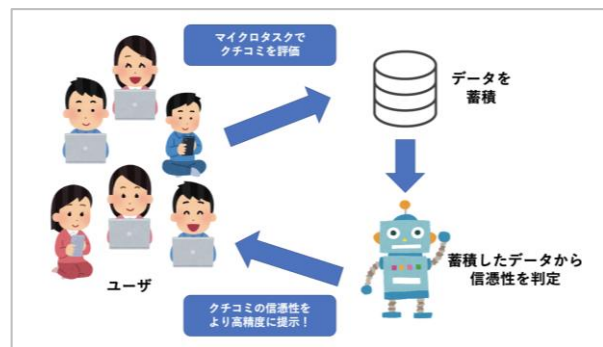


図1 システムでの評価軸利用イメージ図



図2 システム画面

画面右上の㊦マークをクリックすると信憑性評価項目に関する説明が表示される(図5)。また、化粧品の種類の絞り込み機能も実装しており、画面中央上部のラジオボタンを用いて選択した化粧品についてのクチコミのみを表示できる。

3.2 クチコミに対する軸評価

本稿では、過去の研究[3]で収集したファンデーション、化粧水、リップについてのクチコミ各100件ずつの計300件をシステムに使用する。そのため、これまで検討した信憑性評価軸をもとにこれらのクチコミに対して軸評価を行う。ここで、評価軸の検証[6]の際にはクチコミに対する同様の軸評価を13名に依頼していたが、この人数を減らして機械学習を行ったところ判定精度がほとんど変化しなかったため、本稿では著者を含む化粧をする大学生3名が行った。

軸評価は、クチコミに対してこれまで検討した評価軸の各項目に当てはまる場合1、当てはまらない場合0を入力してもらうものを300件分行った。3名全員の回答が一致したクチコミの割合を表1に示す。一致率ももっとも高かったのは「併用した製品についての記述の有無」で94.7%であった。また、もっとも低かったのは「投稿者なりの使用方法についての記述の有無」であり、71.0%であった。この一致率に関して、過去の研究[6]でのデータセット構築における軸評価の際も一致率ももっとも高かったのは「併用した製品についての記述の有無」で87.4%、もっとも低かったのは「投稿者なりの使用方法についての記述の有無」で51.8%となっており、今回の軸評価では全体的に一致率が向上していることが確認された。なお、実際にシステムとして運用する際には、今回の軸評価と過去の研究の両方で一致率ももっとも低かった「投稿者なりの使用方法についての記述の有無」を評価者が容易に判断できるような判断基準を設ける必要があると考えられる。

3.3 クチコミの信憑性スコア算出

3.2節で作成した軸評価データを用いてクチコミの信憑性スコアを算出する。スコア算出では「それぞれの化粧品において重要な単語の有無」の評価軸に関して、クチコミに含まれる単語数が8単語以上の場合1点、4~7単語の場合0.5点、3単語以下の場合0点とし、その他の7項目の評価軸に関しては3名の軸評価データの平均値をそのクチコミの評価値としたうえで、すべての評価値の和を信憑性スコアとした。システムではこの評価値が3.5点より大きい場合S、3.5~2点の場合A、2~1点の場合B、1点以下の場合Cとランク付けを行い、これをシステム上で提示する。

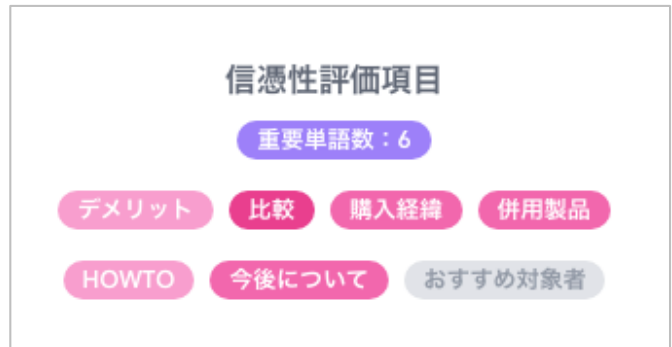


図3 信憑性評価項目 (信憑性ランク S の場合)



図4 信憑性評価項目 (信憑性ランク C の場合)

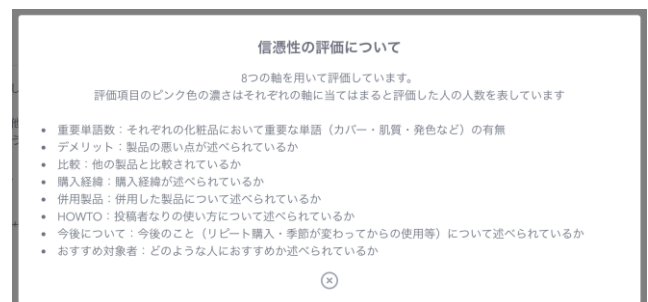


図5 信憑性評価項目に関する説明の表示

表1 3名の回答の一致率

評価軸	一致率
製品の悪い点についての記述の有無	75.3%
他の製品との比較の有無	87.7%
購入経緯についての記述の有無	79.7%
併用した製品についての記述の有無	93.7%
投稿者なりの使用方法についての記述の有無	71.0%
今後のことについての記述の有無	72.0%
どのような人におすすめの記述の有無	86.7%

4. システム評価実験

4.1 実験概要

信憑性評価軸やスコアを提示することの効果明らかにしつつシステムについてのフィードバックを得るため、評価実験を行った。本実験では化粧をする大学生 11 名（女性 10 名，回答しない 1 名）にシステムを利用してもらい、信憑性評価の提示に関する意見やシステムの使いやすさについてアンケートを行った。

実験ではまず、システムの利用前に事前アンケートを実施した。その後「ファンデーションが切れてしまったので、新しく購入するファンデーションを探す」といった状況を説明し、その状況に沿いつつシステムを用いてどのようなものを購入するか決めてもらった。同様の試行を化粧水とリップについても行い、最後に事後アンケートを行った。なお、順序効果を防ぐため、ファンデーション、化粧水、リップの各試行はこちらがランダムに指定した順番で行ってもらった。アンケート内容を表 2, 3 に示す。

4.2 事前アンケートの結果

事前アンケート Q1「化粧品のクチコミの閲覧頻度を教えてください」の結果を図 6 に示す。もっとも多かったのは時々閲覧する（2 週間に 1 度程度）と回答した人であり、閲覧したことはないと回答した 1 名を除く 10 名は化粧品

のクチコミを閲覧していることが明らかになった。

Q2「普段化粧品のクチコミを閲覧する際、どのようなものを参考にしますか」への回答では、自身と同じ肌質の人のクチコミや最新のクチコミなどが参考にするクチコミの特徴として挙げられるとともに、具体的なサービス名として Instagram や Twitter といった SNS や、EC サイトの Amazon, LIPS や @cosme などの化粧品専門のクチコミサイトが挙げられており、人によってどのようなものを参考にするかは異なることが確認された。

Q3「化粧品のクチコミの信憑性をどの程度気にしていますか」の結果を図 7（左）に示す。この結果より 11 名中 7 名が 5 段階中 4 以上を選択しており、信憑性を気にしていたが、3 名は 5 段階中 2 以下を選択しており、信憑性をほ

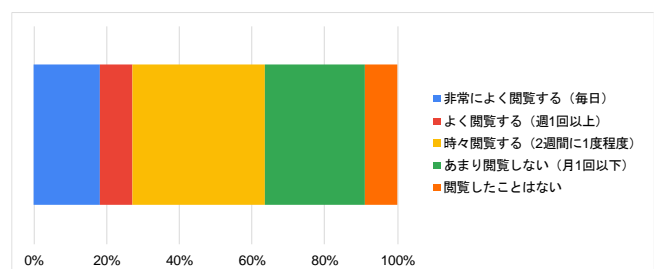


図 6 化粧品クチコミの閲覧頻度 (Q1)

表 2 事前アンケートの内容

Q1	化粧品のクチコミの閲覧頻度を教えてください	5 段階（非常によく閲覧する（毎日），よく閲覧する（週 1 回以上），時々閲覧する（2 週間に 1 度程度），あまり閲覧しない（月 1 回以下），閲覧したことはない）
Q2	普段化粧品のクチコミを閲覧する際、どのようなものを参考にしますか	自由記述
Q3	化粧品のクチコミの信憑性をどの程度気にしていますか	5 段階（5：非常に気にする～0：全く気にしない）
Q4	クチコミの信憑性を気にする場合、どのような点で信憑性を判断していますか	自由記述

表 3 事後アンケートの内容

Q5	ファンデーション・化粧水・リップについて、どのようなものを購入したいと思いましたか	自由記述
Q6	化粧品のクチコミの信憑性をどの程度気にしてシステムを利用しましたか	5 段階（5：非常に気にした～0：全く気にしなかった）
Q7	システムで表示される信憑性評価項目はどの程度役に立ちましたか	5 段階（5：非常に役に立った～0：全く役に立たなかった）
Q8	システムで表示される信憑性ランクはどの程度役に立ちましたか	5 段階（5：非常に役に立った～0：全く役に立たなかった）
Q9	システムで提示される信憑性についてどう感じましたか	自由記述
Q10	システムに対する意見・感想をお願いします	自由記述

とんど気にしていない人もいた。

Q4「クチコミの信憑性を気にする場合、どのような点で信憑性を判断していますか」の回答内容も人によって大きく異なっており、「実際の体験をもとに評価しているかどうかで判断している」「全てを褒めるのではなく、合わなかったものははっきり言っているという部分です」などクチコミ本文から判断している人もいれば、「評価の分布」「同じような意見を持っている人が複数人いるかどうか」など複数のクチコミから総合的に判断している人もいた。

4.3 事後アンケートの結果

Q5の「ファンデーション・化粧水・リップについて、どのようなものを購入したいと思いましたか」についての自由記述では、具体的なクチコミ内容を引用した記述が多かったが、自身と同じ肌質かつ信憑性ランクがSのものを参考にしたという意見もあった。

Q6「化粧品のクチコミの信憑性をどの程度気にしてシステムを利用しましたか」の結果を図7に示す。事前アンケートのQ3と比較すると、5段階中4または5を回答した人数は変化がなかったが、1を回答した人が事前アンケートでは2名いたのに対し、事後アンケートでは0名であった。この2名について回答を詳しくみると、2名とも事後アンケートでは5段階中4または5を回答しており、システムを利用することによって信憑性を意識するようになっていた。しかし、事前アンケートと事後アンケートの回答を比較すると事後アンケートの方が低い値を回答した実験参加者も2名存在していた。

信憑性評価項目と信憑性ランクがどの程度役に立ったかを比較するため、Q7「システムで表示される信憑性評価項目はどの程度役に立ちましたか」～Q8「システムで表示される信憑性ランクはどの程度役に立ちましたか」の結果を図8に示す。Q7では11名中6名が4以上を選択していたが、4名は2を選択していた。Q8では、8名が4以上を選択しており、信憑性評価項目に比べ信憑性ランクを参考にされる傾向があった。

Q9「システムで提示される信憑性についてどう感じましたか」では、「クチコミをととも信じてしまうため、信憑性

について人間の偏見なくシステムが判定してくれるのはとても信用できると感じる」といったシステムに対するポジティブな意見のほか、「信憑性が低いと記されているものは見ずにA以上のものを参考にしたが、AかSかはあまり気にしなかった」「信憑性ランクが低いクチコミでも自分の気になる項目について書かれていれば十分参考になったので、信憑性ランクが低いからといって読み飛ばすということはなかった」など、ユーザ各々の信憑性提示に関する利用方法についての意見も得られた。また、「短文中で重要単語数が0のレビューであっても信憑性評価に利用される8つの軸のうち複数個の内容が入ればランクが上がってしまう。基本的にレビューを読むときはランクが高いものを選択して読むのでそれが少々煩わしかった」といった意見もあり、信憑性スコアの算出方法については検討の必要性が示唆された。

Q10「システムに対する意見・感想」に対しては、「自分の重視する項目を選択して、自分に最適化された信憑性ランクが出てきたらいいと思った」「ポジティブな意見かネガティブな意見かが一目でわかるような欄があると良いなと思いました」「化粧品の種類によって多くの人が気にするポイント（例えばファンデなら「カバー力」「落ちやすさ」など）についての記述があるクチコミなのか、パッとみてわかると便利な気がした」などの意見が得られた。

5. 考察

事前アンケートQ3「化粧品のクチコミの信憑性をどの程度気にしていますか」と事後アンケートQ6「化粧品のクチコミの信憑性をどの程度気にしてシステムを利用しましたか」の比較から、普段信憑性を気にしていない人はシステムを利用することで信憑性を普段よりも意識できる可能性が示唆された。しかし、普段から信憑性を気にしている人の中にはシステムの信憑性評価提示によって信憑性を自身で疑わなくなってしまう人も存在する可能性がある。

また、多くの実験参加者が信憑性評価項目よりも信憑性ランクを参考にシステムを利用したことが明らかになった。

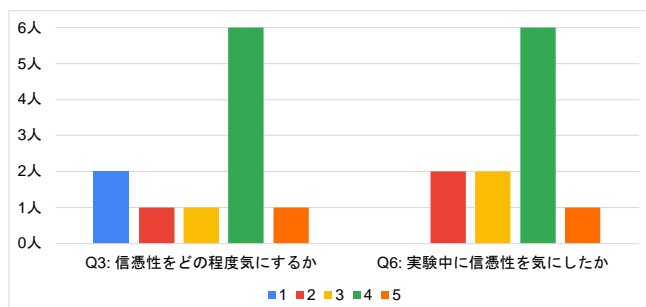


図7 信憑性に対する意識
 (事前アンケート Q3・事後アンケート Q6)

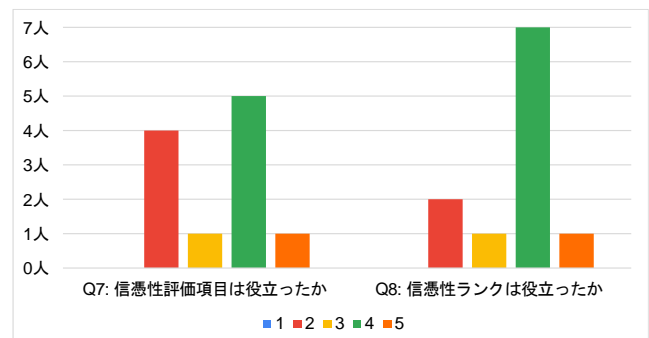


図8 事後アンケート結果 (Q7～Q8)

しかし、事後アンケート Q7「システムで表示される信憑性評価項目はどの程度役に立ちましたか」と Q8「システムで表示される信憑性ランクはどの程度役に立ちましたか」の回答を個別に見ると、Q7を Q8 より高い値で評価している参加者が 2 名確認できた。この 2 名について Q9「システムで提示される信憑性についてどう感じましたか」の回答をみると、信憑性評価項目の表示から自身が重視する記述が含まれるクチコミを探していることがわかった。さらに他の実験参加者の Q9「システムで提示される信憑性についてどう感じましたか」、Q10「システムに対する意見・感想をお願いします」への回答から、参加者ごとにクチコミの見方が異なり、信憑性ランクのみで閲覧するクチコミを決める人や、ランクが一定以上かつ自身が重視する信憑性評価項目を含むものを閲覧する人などが存在することが明らかになった。現在のシステムでは信憑性評価ランクの根拠として全信憑性評価項目を提示していたが、評価項目で検索する機能を実装することで様々なユーザの利用方法に対応可能なシステムにしていくことができる。また、評価項目を現在は色の濃さで軸評価の際に当てはまると回答した人数を表現しているが、評価項目やランクのソート機能を実装することでよりユーザが自身のニーズに合うクチコミを探しやすくなると考えられる。

また、Q10 で得られた意見から、重要単語数に関して数字のみでなく詳細も表示することでより便利なシステムにできると考えられる。ここで、現在の重要単語数はクチコミに含まれる重要単語の個数を信憑性評価軸として用いているが、システムに用いたクチコミの中には Twitter から収集した短いクチコミも存在しており、単語数のみでは文章量の差を考慮できていなかった。そのため、今後は文章量に対する重要単語の数にするなど文章量を考慮した軸を検討する必要がある。また、現在はスコア算出時のそれぞれの評価項目に対する配点を均等にしているが、信憑性判定精度を向上させるために重み付けを行うことも検討している。さらに、現状の評価軸は文章についてのもののみで構成されているため、投稿者や画像に関する軸などを検討することで、より高精度な推定が可能になると考えられる。

6. まとめ

本研究では、化粧品のクチコミがもつ独自の性質を踏まえたクチコミの信憑性評価システムの開発を行った。まずシステムの実装にむけ、300 件のクチコミに対し軸評価を行った。このデータを用いて信憑性スコアを算出し、スコアをもとに信憑性ランクとしてシステム上に提示した。このシステムを用いた評価実験の結果、信憑性評価項目に比べ信憑性評価スコアを多くの実験参加者が参考にしていることが明らかになった。また、信憑性評価項目を用いた検索機能を実装することで、評価項目の提示についてもより

有意義なものにできると示唆された。

今後は、評価実験で得られた意見をもとに検索機能の実現や文章量を考慮した重要単語軸の検討、信憑性スコア算出方法の改善を行う。また、新しい評価軸についての検討を進めることでシステムの精度向上を目指す。特に、これまでは文章のみに着目し、評価軸を検討・検証していたが、今後は投稿者や画像といった要素にも着目し、評価軸の追加を行うことで、システムをより良いものにしていく。

参考文献

- [1] “ウィズコロナ時代の美容行動「画面越しに見せたい顔」とは?” . <https://www.cosmetic-culture.po-holdings.co.jp/report/pdf/220623gamengoshi.pdf>, (参照 2022-07-12).
- [2] “メイク崩れ注意報発令！7 割近くの女性が「ファンデーション選びで失敗したことがある」と回答。気になる失敗エピソードとは?” . <https://www.cocondesoie.com/post/make-up-breakage-warning-issued>, (参照 2022-07-12)
- [3] 濱野花莉, 中村聡史. 化粧品に対するクチコミの信憑性判定に向けたクチコミ文章の調査. 第 13 回データ工学と情報マネジメントに関するフォーラム (DEIM2021), 2021, no.C14-3, pp.1-8.
- [4] Mukherjee, A. Liu, B. and Glance, N.. Spotting fake reviewer groups in consumer reviews. Proceedings of the 21st Annual Conference on World Wide Web. 2012, pp.191-200.
- [5] “サクラチェッカー” . <https://sakura-checker.jp/>, (参照 2022-07-12)
- [6] 濱野花莉, 伊藤理紗, 中村聡史. 文章に着目した化粧品クチコミの信憑性評価軸の検討. 情報処理学会 研究報告グループウェアとネットワークサービス (GN), 2022, vol.2022-GN-115, no.32, pp.1-8.
- [7] 高木佐恵子, 波川千晶, 吉本富士市. メイクアップ技術上達のためのアドバイスシステム. 芸術科学会論文誌. 2003, vol.2, no.4, pp.156-164.
- [8] 神武里奈, 星野准一. 好みの顔画像の色に基づくメイクアップ支援システム. 日本感性工学会論文誌. 2017, vol.16, no.3, pp.299-306.
- [9] Chong, T., Maudet, N., Harima, K., and Igarashi, T.. Exploring a Makeup Support System for Transgender Passing based on Automatic Gender Recognition. CHI'21. 2021.
- [10] Kajita, M., and Nakamura, S.. Basic Research on How to Apply Foundation Makeup Evenly on Your Own Face. 20th IFIP TC14 International Conference on Entertainment Computing (IFIP ICEC 2021), 2021, pp.402-410.
- [11] 藤本雄紀, 今井龍一, 中村健二, 田中成典, 有馬伸広, 荒川貴之. 信憑性の高い記事を投稿するマイクロブログユーザの特定手法に関する研究. 土木学会論文集. 2017, vol.73, no.2, pp.1_259-1.267.
- [12] Moran, G. and Muzellec, L.. eWOM credibility on social networking sites: A framework. Journal of Marketing Communications. 2014, vol.23, Issue 2, pp.149-161.
- [13] 中島伸介, 稲垣陽一, 草野奉章. 高信頼性情報の提示を目指した熟知度に基づくブログランキング方式の提案. 日本データベース学会論文誌. 2008, vol.7, no.1, pp.1-6.
- [14] Huang, H. Wen, Y. and Chen, H.. Detection of False Online Advertisements with DCNN. Proceedings of the 26th International Conference on World Wide Web (WWW'17). 2017, pp.795-796.
- [15] 山本祐輔, 田中克己. データ対間のサポート関係分析に基づく Web 情報の信憑性評価. 情報処理学会論文誌. 2010, vol.3, no.2, pp.61-79.

- [16] 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. 2019, vol.12, no.1, pp.24-37.
- [17] Lee, S. and Kim, E.. Influencer marketing on Instagram: How sponsorship disclosure, influencer credibility, and brand credibility impact the effectiveness of Instagram promotional post. *Journal of Global Fashion Marketing*. 2020, vol.11, no.3, pp.232-249.
- [18] Oh, Y.. Determinants of Online Review Helpfulness for Korean Skincare Products in Online Retailing. *Journal of Distribution Science*. 2020, vol.18, no.10, pp.65-75.
- [19] Rofianto, W., Kornelys, D., and Rifkhansyah, M.. Visual eWOM Usefulness and Credibility: The Antecedents and Consequences. *Jurnal Ilmu Manajemen & Ekonomika*. 2017, vol.9, no.2, pp.103-108.
- [20] Iwamoto, E., Matsubara, M., Ota, C., Nakamura, S., Terada, T., Kitagawa, H., and Morishima, A.. Passerby Crowdsourcing: Workers' Behavior and Data Quality Management. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 2018, vol.2, no.4, pp.1-20.
- [21] 三輪聡哉, 中村聡史. マイクロタスク埋め込み型音楽ゲームの提案. 情報処理学会 研究報告エンタテインメントコンピューティング (EC) . 2014, vol.2014-EC-34, Issue.2, pp.1-6.
- [22] 佐々木孝輔, 平田章, 井上智雄. マイクロタスクによる線画イラスト生成手法. 情報処理学会論文誌. 2016, vol.57, no.1, pp.260-269.
- [23] 三原鉄也, 石川夏樹, 豊田将平, 永森光晴, 杉本重雄. 画像認識とマイクロタスク型クラウドソーシングを組み合わせたマンガのコマ領域の判定. *人工知能学会全国大会論文集*. 2018, JSAI2018 巻, 第 32 回, p.4M105.