

ゲーミフィケーションを利用したウェブ情報の信頼性評価行動を促進するシステム

佐藤 祐吾^{1,a)} 松村 敦^{2,b)}

概要: ウェブ上の誤情報に惑わされないためには、ユーザ自身でウェブサイトの情報を確認し、信頼できるかどうか判断するウェブ情報の信頼性評価行動が重要である。そこで本研究は、ウェブ情報の信頼性評価行動をゲーミフィケーションで促進するシステムを構築し、その利用による効果を検証することを目的とする。提案システムでは、ウェブサイトの評価用フォームを作成し登録数に応じて経験値を与えたりランクを設けることでフィードバックを行う。本システムの有効性を検証するために評価実験を行った。その結果、ウェブを確認する意識は、情報学系学部の所属やゲーミフィケーションの有無による違いは見られなかったものの、1評価あたりで入力された評価項目数については、提案システムを使った非情報学系学部の実験参加者と、比較用システムを使った情報学系学部の実験参加者で有意に多かった。

1. 研究背景

現在、インターネットや情報通信端末が普及し、いち早く世の中の出来事や動きを把握したり、趣味や娯楽の情報を得るメディアとして多く利用されている [1]。また、ブログや SNS、動画投稿サイトなどのサービスによって、行政や公的機関、報道機関や会社、社会人や学生といった個人など、多様な背景を持つ人が情報を発信できるようになった。一方で、間違った情報も発信されやすくなり、デマやフェイクニュースといった誤情報の拡散による被害が社会問題となっている。

こうした問題への対策として、検索エンジンなどのサービス側が表示情報の操作・制限を行うアプローチと、ユーザ自身がウェブサイトの情報が信頼できるようにするアプローチの両方面が考えられる。

本研究では、最終的に閲覧するウェブサイトを選択し、そこで記載された情報を信頼するか判断する主体はユーザであることから、ユーザ側に焦点を当てたアプローチをとる。ここで、本研究では、ユーザ自身が発信者の情報や根拠、最終更新日などの項目を確認した上で、そのウェブサ

イトの情報が信頼できるかどうか判断するという行動を「ウェブ情報の信頼性評価行動」と定義する。

ユーザがウェブ情報の信頼性評価行動をとるにあたり、ウェブサイト毎に一つ一つの項目を確認して信頼性を評価することは面倒に感じられると考えられる。その要因として、より多くウェブサイトの情報を確認して評価するための目標が明確にはないことによって、行動意欲が削がれることが考えられる。そこで、娯楽であるゲームのデザインや構成要素を取り入れる「ゲーミフィケーション」を活用する。ゲーミフィケーションでは、ポイントを使った達成状況の可視化や、バッジや称号のような報酬の付与によるモチベーション強化を行うことで、ユーザを作業に惹きつけ、課題解決を促す。このゲーミフィケーションを利用することで、ウェブ情報の信頼性評価行動が面倒に感じられる要因を緩和することが期待できる。

そこで本研究では、ウェブ情報の信頼性評価行動にゲーミフィケーションを適用したシステムの利用によって、ウェブ情報の信頼性評価行動が促進されるか検証する。

2. 関連研究

2.1 ウェブ情報の信頼性に関する評価尺度

鷲見ら [2] は、「調べる目的で利用する情報源としての Web サイトの質に対する評定尺度」として「Web サイト評価インベントリ (WEI: Website Evaluation Inventory)」を作成した。WEI は既存の評価ツール 27 件を基に項目の選定・追加・削除が行われた、20 項目について 6 段階で評価する尺度である。山本ら [3] は、「検索エンジン

¹ 筑波大学情報学群知識情報・図書館学類
College of Knowledge and Library Sciences, School of Information, University of Tsukuba, Kasuga 1-2, Tsukuba, Ibaraki 305-8550, Japan

² 筑波大学図書館情報メディア系
Faculty of Library, Information and Media Science, University of Tsukuba

a) s1811494@s.tsukuba.ac.jp

b) matsumur@slis.tsukuba.ac.jp

等の情報アクセスシステムを上手く使いながら、情報を批判的に精査し、正確なウェブ情報を収集するための能力」を「ウェブアクセスリテラシー」として、その能力を測定するための「ウェブアクセスリテラシー尺度」を開発した。ウェブアクセスリテラシー尺度は7つの下位尺度で計59項目の質問項目があり、それぞれ5段階で回答する。

2.2 ウェブ情報の信頼性評価行動をゲーミフィケーションで促進するシステム

Huangら[6]はウェブサイトの信頼性を評価するクラウドソーシングのためのプラットフォームとして、CredibleWebを開発した。CredibleWebではウェブサイトを「全体的な信頼性スコア」、「真実性」、「公正さ」、「セキュリティ」、「ページデザイン」について5段階で評価する他、ウェブサイトについて発見された問題点を記載するテキストボックスが設けられている。また、他のユーザが作成したウェブサイトの評価に良いと感じられるものがあれば、投票することができる。導入されたゲーミフィケーションは「ポイント」、「達成度」、「リーダーボード」である。

Juzwinら[7]はウェブページを評価するプラットフォーム Reconcileを開発し、ゲーミフィケーションの導入によるウェブサイトの信頼性評価への影響を検証した。Reconcileではウェブサイトを「信頼性」、「外観」、「著者の専門性」、「著者の意図」、「情報の完全性」という観点から5段階で評価する。ゲーミフィケーションとしては、「クエスト」、「賭け」、「クイズ」を導入した。「クイズ」は真偽で答える問題を提示し、ユーザの知識や評価の質の向上を図った。

これらの研究では、ウェブ情報の評価項目が「ページのデザイン」や「著者の専門性」など、ウェブサイトで確認すべき項目が曖昧になっている。さらに、評価項目はほとんどが5件法で評価されるため、具体的にウェブサイトのどの箇所を確認して入力しているかが不明瞭である。

3. 提案システム

3.1 システム概要

本研究ではウェブ情報の信頼性評価行動を促すために、細分化された評価項目からなるウェブ情報の信頼性評価フォーム（以下、評価フォーム）と、ゲームメカニクスとしてレビュー・経験値・ランクを備えたシステムを提案する。提案手法のイメージを図1に示す。

ユーザは任意のウェブサイトに対して、評価フォームを入力しながら信頼性評価を行う。そして、入力した評価項目の数に応じて、経験値を獲得できる。保存された評価は、他のユーザから見ることができる。ユーザは、既に自分で評価したウェブサイトと同じものを対象とした他ユーザの評価について、点数をつけるレビューができる。レビューも評価フォームと同様、入力したレビュー項目数に応じて経験値を獲得できる。評価やレビューを行うことで、ユー

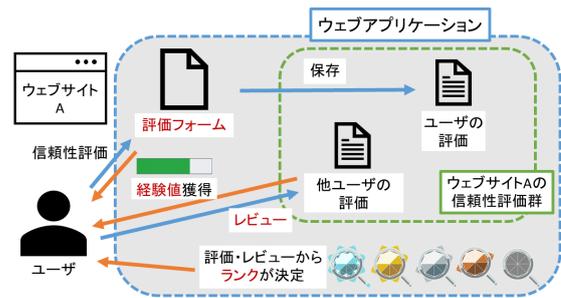


図1 提案手法のイメージ図

ザのランクが決定される。

3.2 評価フォーム

評価フォームは、WEIとウェブアクセスリテラシー尺度の質問項目のうち、ウェブサイトの内容に関する項目を利用して作成した。この評価フォームにより、主観や事前知識の影響を抑えたウェブ情報の信頼性評価行動を支援することが期待できる。表1作成した評価項目と入力形式を示す。

3.3 導入したゲームメカニクス

3.3.1 レビュー

ユーザの行った信頼性評価の内容に対するフィードバックを行うために、レビューを用いる。提案システムのレビューでは、他のユーザが記録したウェブ情報の信頼性評価に点数をつける。これにより、ユーザ自身がウェブ情報の信頼性評価行動を自省するきっかけを与える。またレビューをするには、レビュー対象と同じウェブサイトの評価していることを前提とした。この仕様により、レビューの整合性を担保し、乱雑なレビューの抑制を狙っている。また、他ユーザの名前、ランクを示すことで、他ユーザと自分を比較するリーダーボードとしての機能も備える。表2に作成したレビュー項目を示す。

3.3.2 経験値

ユーザの行動量を可視化するため、経験値を採用する。経験値は、評価フォームやレビューの入力を行うことで獲得できる。これによって、ウェブ情報の信頼性評価行動を定量的に可視化する。

本研究では、信頼に足る十分な情報を掲載したウェブサイトを集めることではなく、評価する行動を促すことに焦点を置いている。したがって、各経験値は対応する評価項目を入力すれば獲得でき、評価項目に入力した値によって得られる経験値は変化しない。

3.3.3 ランク

ランクは、システム上のユーザの地位を示す。ランクを昇格させることで、他のユーザがレビューにて閲覧した際のユーザ名の色とバッジ画像が固有のものに変化する。ラ

表 1 評価フォームの評価項目

項目	入力形式
発信者の名前	テキスト
発信者のプロフィールや資格、実績	テキスト
発信者の連絡先が明確に示されている	6 段階
発信者が信頼できる機関に所属しているか、信頼できる機関そのものであることが明確に示されている	6 段階
ウェブサイトに、公的機関からの推薦や信頼性の高い第三者からの認定が明確に示されている	6 段階
ウェブサイトに推薦や認定を与えている、公的機関や信頼性の高い第三者の名称	テキスト
掲載内容は、研究やデータによる事実に基づいていることが明確に示されている	6 段階
参照している研究やデータの情報	テキスト
発信者による根拠の少ない主張はほとんどない	6 段階
根拠が少ない主張部分	テキスト
誤字、脱字、スペルミスなど、間違いはほとんどない	6 段階
ある特定の製品/サービスの購入をすすめる掲載内容はほとんどない	6 段階
掲載内容のテーマが明確に示されている	6 段階
掲載内容の元になっている情報源のタイプ（政府発表、統計データ、新聞など）が、明確に示されている	6 段階
掲載内容に関連し、役立つ、他サイトへのリンクが十分なものである	6 段階
FAQ における回答内容が十分なものである	6 段階
最終更新日	テキスト
掲載内容の更新頻度が明確に示されている	6 段階
掲載内容の更新箇所が明確に示されている	6 段階
ウェブサイトの目的・掲載意図	テキスト
ウェブサイトの利用対象者	テキスト
ウェブサイト内を移動しやすくする工夫（サイトマップ、サイト内サーチエンジンなど）が分かりやすいところに置かれている	6 段階
ウェブサイトの利用方法を説明するヘルプ情報が十分なものである	6 段階
表示速度を落としていそうな画像/動画の掲載はほとんどない	6 段階
ウェブサイトの利用には、特定のブラウザ、ブラウザ以外のソフトが必要である	3 択
ウェブサイトの利用に必要な、特定のブラウザやソフトに関する説明が十分なものである	6 段階
ウェブサイトの情報の信頼度	5 段階
信頼度の判断理由	テキスト

表 2 レビュー項目

項目	入力形式
「ウェブサイトの情報の信頼度」が、他の評価項目の内容に基づいている。または、他の評価項目の内容に基づかない理由が「信頼度の判断理由」に記述されている	5 段階
ウェブサイトの内容と相違がない	5 段階
記入漏れがない	5 段階

図 2 評価フォーム

リンクを昇格させるためには、特定の条件を満たす評価やレビューを行う他、一定点数以上のレビューを受け取る必要がある。こうした仕様を設けることで、ユーザがより多くの評価項目を注意深く確認して評価することを狙っている。

これらのインセンティブや昇格条件は、システムを利用してウェブ情報の信頼性評価行動を行う際の目標やモチベーションとなることを狙っている。

3.4 システムのインターフェース

ブラウザで使用可能なウェブアプリケーションと補助ツールとしての Google Chrome の拡張機能を実装した。

3.4.1 評価フォーム

評価フォームのインターフェースを図 2 に示す。評価フォームはウェブサイトと見比べながら利用できるように、別ウィンドウで表示される。評価フォーム上部の「獲得 exp」は入力状況を示すプログレスバーとなっており、経験値が付与される評価項目に入力を行うほど、緑の面積が増加する。

評価を保存した後、同じウェブサイトを他のユーザが評価している場合はメッセージを表示し、レビューを促す。評価一覧のページでは、ユーザが評価してきたウェブサイトの評価内容を確認できる。

3.4.2 レビュー

レビューする評価を検索するツールとして、評価検索機能を実装した。評価検索機能では条件を指定して評価を絞り込みできる。絞り込み条件は、評価やレビューの有無、テーマを指定できる。またレビューに偏りが生じないように、誰もレビューしていない評価を優先的に表示した。

レビューのフォームでは、図 3 に示したように同じウェブサイトに対する自分の評価と相手の評価を並べて表示する。



図 3 レビュー



図 4 送ったレビュー



図 5 受け取ったレビュー

送ったレビューと受け取ったレビューは図 4, 5 のように一覧で見ることができる。受け取ったレビューの一覧ではレビューを返す機能を実装している。レビューを返す機能は、レビューを送ってきたユーザの評価に対してレビューを促す仕組みである。この機能は、レビューの促進やレビューがもらえないユーザへの補助を狙っている。

3.4.3 Google Chrome 拡張機能

実装した Google Chrome 拡張機能は、ページタイトルや URL といった基本情報の入力を補助するツールとして実装した。評価したいウェブサイトで使用すると、URL とページタイトルが既に入力された状態で、ウェブサイトの評価に取り組み始めることができる。

4. 評価実験

4.1 実験の概要

提案システムによってウェブ情報の信頼性評価行動が促進されたか検証するため、評価実験を行った。ユーザの行動への影響がゲーミフィケーションによるものかを確認するため、提案システムと、提案システムからゲーミフィケーションの要素を除去した比較用システムを用意し、実験参加者を割り振った。そして実験参加者自身の PC で、1 週間毎日システムを利用してもらった。

4.2 手続き

以下に実験の流れを示す。

- (1) 事前アンケート：Google フォームで実験参加者の属性、インターネットの利用状況、ウェブサイト閲覧に関するアンケートを行った。
- (2) システムの割り振り：事前アンケートの回答に基づき、実験参加者が偏らないように提案システムと比較用システムを使うグループに割り振った。
- (3) 操作説明：実験初日の 0 時に、各参加者に割り振られたシステムの操作説明書を、メールで送付した。
- (4) 実験システムの利用：実験期間中毎日 0 時に、検索テーマをメールで送付し、信頼性評価を行うよう指示した。毎日 30 分以上の作業時間をとり、必ず 1 つ以上のウェブサイトをシステム上で評価するよう指示した。
- (5) 事後アンケート：実験最終日の 0 時に、インターネットの利用状況、ウェブサイト閲覧に関する質問と、システムについてアンケートをメールで送付した。

4.3 事前アンケート

インターネットの利用状況は、インターネット利用時間、調べ物に使う時間、利用年数などである。ウェブサイト閲覧に関する質問項目では、ウェブ情報の信頼性評価行動において一つ一つの項目を確認する意識について質問する。質問項目を表 3 に示す。これらの項目は評価フォームと同様に、WEI[2]とウェブアクセスリテラシー尺度 [3] のうち、ウェブサイトの内容に関する項目を利用して作成した。Q-C1~Q-C22 の回答の選択肢は「必ず確認する」「確認する時の方が多い」「確認しない時の方が多い」「全く確認しない」の 4 段階とし、得られた回答を 1~4 の整数値に変換する。こうして得られた値の合計値を、ウェブ情報の信頼性評価行動において、一つ一つの情報を確認する意識(確認意識)とした。

表 3 ウェブサイト閲覧に関する質問

番号	質問項目
Q-C1	作成者／機関の名前が、明確に示されているか
Q-C2	作成者／機関の資格や実績が、明確に示されているか
Q-C3	作成者／機関の連絡先が、明確に示されているか
Q-C4	作成者／機関が、信頼できる機関に所属しているか、信頼できる機関そのものであることが、明確に示されているか
Q-C5	公的機関からの認定や信頼できる第三者からの推薦があるかどうか
Q-C6	掲載内容は、研究やデータによる事実に基づいていることが、明確に示されているか
Q-C7	作成者／機関による、根拠の少ない主張／意見はないか
Q-C8	誤字、脱字、スペルミスなど、間違いはないか
Q-C9	ある特定の製品／サービスの購入をすすめる掲載内容はないか
Q-C10	掲載内容のテーマが、明確に示されているか
Q-C11	掲載内容の元になっている情報源のタイプ（政府発表、統計、データ、新聞など）が、明確に示されているか
Q-C12	掲載内容に関連し、役立つ、他のサイトへのリンクが、十分なものであるか
Q-C13	FAQ（ウェブサイト内の「よくある質問」、掲載内容についてウェブサイトを見た人から寄せられた疑問・質問、そのサイトの使い方についての質問など）における、回答内容が、十分なものであるか
Q-C14	最終更新日が、明確に示されているか
Q-C15	掲載内容の更新頻度が、明確に示されているか
Q-C16	掲載内容の更新箇所が、明確に示されているか
Q-C17	ウェブサイトの目的が、明確に示されているか
Q-C18	ウェブサイトの利用対象者が、明確に示されているか
Q-C19	ウェブサイト内を移動しやすくする工夫（サイトマップ、サイト内サーチエンジンなど）が、わかりやすい場所におかれているか
Q-C20	ウェブサイトの利用方法を説明するヘルプ情報の内容が、十分なものであるか
Q-C21	表示速度を落としていそうな画像／動画の掲載はないか
Q-C22	特定のブラウザ、ブラウザ以外のソフトが必要な場合、それに関する説明が、十分なものであるか

4.4 検索テーマ

検索テーマの内容は、意見や情報の信頼性が分かれやすいと考えられるトピックとして、先行研究 [8] の評価用トピックと、2021 年 10 月 31 日に行われた第 49 回衆議院選挙の公約 [9] を参考にして作成した。提示した検索テーマの一例を表 4 に示す。

4.5 実験参加者

実験では Google Chrome をメインブラウザとして利用している大学生・大学院生 20 人を対象とした。提案システムを利用するグループを G グループ、後述する比較用システムを利用するグループを N グループとした。以降、ユーザはグループ名 G, N と番号で表す。なお、途中からデータ収集が十分できなくなった参加者 2 名を実験対象か

表 4 検索テーマ (例)

提示日	テーマ文
1 日目	SNS で、選択的夫婦別姓制度の賛否に関するネット記事が推薦されました。そこで、改めて夫婦別姓制度の議論について確認することにしました。
2 日目	新型コロナウイルス感染症の対策として、ワクチンの 3 回目の接種が日本で予定されているというテレビの報道をみかけましたが、番組によって言っていることが異なりました。そこで、ワクチンの 3 回目接種についての現状や必要性・安全性を調べることにしました。

表 5 実験前の確認意識（ウェブサイト閲覧に関する回答の合計点）

G01	G02	G05	G06	G07	G08	G09	G10	平均
46	49	55	55	59	60	62	75	56.30
N01	N02	N03	N05	N06	N10	平均		
43	45	48	52	53	64	51.88		

ら除外した。

4.6 実験期間

提案システムではレビューによるユーザ間フィードバックがあるため、全参加者で実験期間を統一した。実験は 2021 年 12 月 7 日から 12 月 13 日までの 7 日間で行った。

4.7 比較用システム

グループ N の参加者に割り当てた比較用システムでは、ゲーミフィケーションの要素としてレビュー・経験値・ランクを除去し、評価フォームと、自身が作成した評価の一覧ページのみアクセスできるよう実装した。

5. 結果

本実験では「毎日必ず 1 つ以上のウェブサイトを経験システム上で評価する」という条件を課していたが、これを満たさない実験参加者が複数存在したため、注意喚起のために、メールにて呼びかけを行った。このような、実験参加者 (G03,G04,N04,N09) は呼びかけの影響を受けていると判断し、以降の結果では省いている。

5.1 実験前後の確認意識と実験中のシステム利用状況

表 5 に事前アンケートの「ウェブサイト閲覧に関する質問」の回答結果を示す。これは実験前の確認意識を表す。

表 6 に、グループ G・グループ N の評価フォームから入力・保存されたウェブサイトの評価数の合計を示す。

実験期間中にグループ G の実験参加者が受け取ったレビュー数を、表 7 に示す。実験期間中にグループ G の実験参加者がレビューした数を、表 8 に示す。

表 9 に、グループ G の実験参加者の実験終了時点でのリンクを示す。

表 6 評価数の合計

G01	G02	G05	G06	G07	G08	G09	G10	合計
15	17	8	7	7	10	10	16	113
N01	N02	N03	N05	N06	N10			合計
19	14	12	15	14	11			101

表 7 受け取ったレビュー数 (グループ G)

G01	G02	G05	G06	G07	G08	G09	G10	合計
3	9	1	3	1	3	1	3	30

表 8 レビューした数 (グループ G)

G01	G02	G05	G06	G07	G08	G09	G10	合計
4	8	1	1	3	2	0	3	30

表 9 グループ G の最終ランク

実験参加者	ランク
G01	シルバー
G02	ダイヤモンド
G05	ブロンズ
G06	ブロンズ
G07	シルバー
G08	ブロンズ
G09	原石
G10	シルバー

表 10 実験後の確認意識 (ウェブサイト閲覧に関する回答の合計点)

G01	G02	G05	G06	G07	G08	G09	G10	平均
66	88	87	47	53	77	56	60	63.30
N01	N02	N03	N05	N06	N10			平均
77	44	47	69	59	68			61.50

表 10 に事後アンケートの「ウェブサイト閲覧に関する質問」の回答結果を示す。これは実験後の確認意識を表す。

5.2 分析結果

5.2.1 学部とゲーミフィケーションの有無による確認意識の変化

実験の前後で確認意識に変化があったかを確認するため、事前アンケート・事後アンケートの回答から得られた確認意識について差をとり、確認意識の変化量とした。

そのうえで、学部やゲーミフィケーションによる確認意識の変化量への影響を確認するため、学部 (情報学系学部・非情報学系学部) と割り振られたシステムのグループ (グループ G・グループ N) を独立変数、確認意識の変化量を従属変数とした 2×2 の分散分析を行った。表 11 に、各

表 11 学部とグループによる確認意識の変化量の平均

学部 グループ	情報学系学部		非情報学系学部	
	G	N	G	N
変化量	10.00	-1.00	8.25	15.25
	(19.11)	(.00)	(24.68)	(13.75)

上段：平均値, 下段：標準偏差

表 12 学部とグループによる入力された評価項目数の平均

学部 グループ	情報学系学部		非情報学系学部	
	G	N	G	N
評価項目数	25.38	26.62	26.50	25.75
	(1.50)	(.57)	(1.00)	(1.09)

上段：平均値, 下段：標準偏差

要因での確認意識の変化量の平均を示す。

分散分析では、学部とグループについて有意な交互作用が見られなかった ($F(1, 10) = .74, p = .41$)。学部とグループそれぞれの主効果も見られなかった ($F(1, 10) = .48, p = .50$; $F(1, 10) = .04, p = .85$)。

5.2.2 学部とゲーミフィケーションの有無による入力された評価項目数の変化

学部やゲーミフィケーションによる、1 評価あたりで入力されたの評価項目数への影響を確認するため、学部 (情報学系学部・非情報学系学部) と割り振られたシステムのグループ (グループ G・グループ N) を独立変数、入力された評価項目数を従属変数とした 2×2 の分散分析を行った。

表 12 に各要因での評価項目数の平均を示す。分散分析では、学部とグループについて有意な交互作用が見られた ($F(1, 171) = 31.78, p < .001$)。そこで、単純主効果の検定を行った。その結果、非情報学系学部におけるゲーミフィケーションの有無の単純主効果が有意であり ($F(1, 171) = 12.35, p < .001$)、グループ G のほうがグループ N より入力された評価項目数が大きくなった。また、情報学系学部におけるゲーミフィケーションの有無の単純主効果が有意であり ($F(1, 171) = 24.24, p < .001$)、グループ N のほうがグループ G より入力された評価項目数が大きくなった。グループ G においては、学部の単純主効果が有意であり ($F(1, 171) = 22.56, p < .001$)、非情報学系学部のほうが情報学系学部より入力された評価項目数が大きくなった。また、グループ N における学部の単純主効果が有意であり ($F(1, 171) = 10.95, p < .01$)、情報学系学部のほうが非情報学系学部より入力された評価項目数が大きくなった。

5.2.3 確認意識とゲーミフィケーションの有無による確認意識の変化

事前アンケートで得た確認意識について、上位群と下位群に分けて分析を行った。この区分を、実験参加以前の確認意識とした。

表 13 確認意識とグループによる確認意識の変化量の平均

確認意識 グループ	上位群		下位群	
	G	N	G	N
変化量	-3.60	9.00	30.33	10.67
	(12.10)	(7.00)	(9.61)	(20.21)

上段：平均値，下段：標準偏差

表 14 確認意識とグループによる入力された評価項目数の平均

確認意識 グループ	上位群		下位群	
	G	N	G	N
評価項目数	26.28	25.70	25.65	26.29
	(1.16)	(1.31)	(1.51)	(.63)

上段：平均値，下段：標準偏差

実験以前の確認意識やゲーミフィケーションによる確認意識の変化量への影響を確認するため、実験以前の確認意識（上位群・下位群）と割り振られたシステムのグループ（グループ G・グループ N）を独立変数、確認意識の変化量を従属変数とした 2×2 の分散分析を行った。

表 13 に、各要因での確認意識の変化量の平均を示す。分散分析で、確認意識とグループについて有意な交互作用が見られた ($F(1, 10) = 5.15, p < .05$)。そこで、単純主効果の検定を行った。その結果、グループ G における確認意識の単純主効果が有意であり ($F(1, 10) = 12.82, p < .01$)、下位群のほうが上位群より確認意識の変化量が大きくなった。

5.2.4 確認意識とゲーミフィケーションの有無による入力された評価項目数の変化

実験以前の確認意識やゲーミフィケーションによる、1 評価あたりで入力された評価項目数への影響を確認するため、実験以前の確認意識（上位群・下位群）と割り振られたシステムのグループ（グループ G・グループ N）を独立変数、入力された評価項目数を従属変数とした 2×2 の分散分析を行った。

表 14 に、各要因での評価項目数の平均を示す。分散分析では、確認意識とグループについて有意な交互作用が見られた ($F(1, 171) = 11.54, p < .01$)。そこで、単純主効果の検定を行った。その結果、上位群におけるゲーミフィケーションの有無の単純主効果が有意であり ($F(1, 171) = 5.35, p < .05$)、グループ G のほうがグループ N より入力された評価項目数が高かった。また、下位群におけるゲーミフィケーションの有無の単純主効果も有意であり ($F(1, 171) = 6.19, p < .05$)、グループ N のほうがグループ G より入力された評価項目数が高かった。グループ G においては、確認意識の単純主効果が有意であり ($F(1, 171) = 6.32, p < .05$)、上位群のほうが下位群より入力された評価項目数が高かった。また、グループ N における確認意識の単純主効果も有意であり ($F(1, 171) = 5.26, p < .05$)、下位群のほうが上位群より入

力された評価項目数が高かった。

5.2.5 ランクと確認意識の変化量の相関

ランクを昇格するには、次のランクが高レベルであるほど、評価フォームへの入力量やレビューの獲得数などの条件が厳しくなる。したがって、ランクの高いユーザはより多くシステムを利用しており、確認意識に変化があったと考えられる。そこで、グループ G について、ランクと確認意識の変化量に相関関係が見られるか確認した。

無相関検定の結果から、ランクと確認意識の変化量に有意な相関は見られなかった ($r = .45, N = 8, p = .87$)。

5.2.6 受け取ったレビュー数と評価数、確認意識の変化量の相関

レビューの獲得がランクの昇格条件に含まれていることで、より多くの評価を行ってレビューされやすくしようとすることが考えられる。そこでグループ G について、受け取ったレビュー数と評価数に相関関係が見られるか確認した。またレビューは、ウェブ情報の信頼性評価行動を自省するきっかけを与える役割を持つ。そこでグループ G について、受け取ったレビュー数と確認意識の変化量についても相関関係が見られるか確認した。

無相関検定の結果から、受け取ったレビュー数と評価数に有意な相関は見られなかった ($r = .69, N = 8, p = .97$)。また、受け取ったレビュー数と確認意識の変化量に有意な相関は見られなかった ($r = .52, N = 8, p = .91$)。

6. 考察

6.1 システム利用前後の意識変化

5.2.1 で行った分析結果から、実験参加者の所属学部やゲーミフィケーションによって、システム利用の前後で確認意識に変化があったとは言えなかった。また 5.2.5 で示したように、提案システムにおいて、ランクと確認意識の変化量には統計的に有意な相関が見られなかった。一方、記述回答に、評価フォームの内容が参考になったことやウェブサイトを注意深く確認するようになったことが、ゲーミフィケーションの有無を問わず言及されていた。これらの結果から、システムの利用前後で確認意識に変化がなかったものの、評価フォームの利用を通して、自身の確認意識の不足に気づく参加者がいたことが分かった。確認意識への影響が見られなかったことの原因としては、実験期間が 1 週間という短期間であったことが考えられる。

5.2.3 で行った分析結果からは、提案システムで、実験参加以前の確認意識が高い実験参加者よりも、低い実験参加者の方が、実験前後で確認意識が有意に向上することが示唆された。これは、実験後に確認意識が向上する余地にも影響を受けていたと推測される。はじめから確認意識が高かった人がより意識を高めたかを検証する方法を検討しなければならない。

6.2 システム利用状況の変化

6.2.1 評価フォームの利用

5.2.2, 5.2.4 では1 ウェブサイト評価あたりの入力された評価項目数について、分析を行った。その結果、提案システムを使った実験参加者のうち、非情報学系学部の実験参加者と実験参加以前の確認意識が高い実験参加者で有意に多くなることが示された。また比較用システムを使った実験参加者では、情報学系学部の実験参加者と実験参加以前の確認意識が低い実験参加者で有意に多くなることが示された。

しかしながら、これらが起こった要因の特定は難しい。想定される要因として、情報学系学部の参加者では、実験参加以前からウェブ情報の信頼性評価行動や関連分野に関心があったことなどが挙げられる。また、実験以前の確認意識の低い参加者にとっては、評価フォームのみを備えた比較用システムの方が評価項目への入力に集中できたことも考えられる。反対に、評価フォームの入力画面とは別で現れるレビューやランクといったゲーミフィケーション要素がシステムの利用方法を複雑化させており、情報学系学部の実験参加者や実験以前の確認意識が低い実験参加者にとって、評価項目の入力を阻害する要因になったとも考えられる。

6.2.2 レビュー

5.2.6 で行った分析結果から、レビューと評価数や確認意識には相関関係が見られなかった。したがって、レビューに応じて、ウェブ情報の信頼性評価行動の増加やウェブサイトの情報を確認する意識の向上があったとは言えない。

これらの結果が得られた原因として、レビューが十分に利用されなかったことが考えられる。表7で示したように実験期間中に行われたレビューは30件であり、113件あった評価に対して多くレビューされたとは言えない。提案システムではレビューを促すために、レビューした相手にレビューを返させるボタンの表示などの工夫を施したが、レビューを促すための仕様も十分に機能していなかったと考えられる。

7. 結論

本研究では、ウェブ情報の信頼性評価行動を促進するためにゲーミフィケーションを適用したシステムの提案と実装を行い、その効果を検証した。実験の結果、確認意識に関しては、情報学系学部の所属やゲーミフィケーションの有無による変化は見られなかった。1 評価あたりで入力された評価項目数については、提案システムを使った非情報学系学部の実験参加者と、比較用システムを使った情報学系学部の実験参加者で有意に多いことが示された。評価数については、ゲーミフィケーションや所属学部、確認意識による差が見られなかった。また、レビューによってウェブ情報の信頼性評価行動が増えたことや、ウェブサイトの

情報を確認する意識が変わったことは確認できなかった。

今後の課題として、より長期的な実験や、ウェブサイトの評価数・検索テーマについて条件を設けない自然な利用での実験を実施し、提案システムによる行動への影響を確認する必要がある。

参考文献

- [1] 総務省. "令和3年版 情報通信白書: メディアとしてのインターネットの位置づけ". 総務省. 2021-07-30. <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r03/html/nd242550.html>, (参照 2021-12-13).
- [2] 鷲見克典, 四谷あさみ. 調べる目的で利用する情報源としての Web サイトに対する評価尺度の作成と信頼性および妥当性の検討. 情報処理学会論文誌. 2004, vol. 45, no. 3, p. 1032-1040, ISSN 1882-7764. <http://id.nii.ac.jp/1001/00010966/>, (参照 2021-10-05).
- [3] 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 (TOD). 2019-01-16, vol. 12, no. 1, p. 23-37, ISSN 1882-7799. <http://id.nii.ac.jp/1001/00193701/>, (参照 2021-06-03).
- [4] 一ノ瀬智浩, 上野秀剛. 特集, ゲームインタフェースとエンタテインメントコンピューティング: ゲーミフィケーションを構成する要素の違いと作業効率の評価. ヒューマンインタフェース学会論文誌. 2016, vol. 18, no. 2, p. 65-76. https://doi.org/10.11184/his.18.2_65, (参照 2021-07-29).
- [5] 倉本到, 片山拓馬, 渋谷雄, 辻野嘉宏. 懐優館: 作業意欲を持続的に維持向上させる EELF に基づく主観的比較型エンタテインメントシステム. 情報処理学会論文誌. 2009, vol. 50, no. 12, p. 2807-2818, ISSN 1882-7764, <http://id.nii.ac.jp/1001/00067513/>, (参照 2022-01-01).
- [6] Zhicong Huang, Alexandra Olteanu, Karl Aberer. "CredibleWeb: a platform for web credibility evaluation". CHI EA '13: CHI '13 Extended Abstracts on Human Factors in Computing Systems. Paris, France, 2013-04-27/05-02. Association for Computing Machinery, 2013, p. 1887-1892. <https://doi.org/10.1145/2468356.2468694>, (参照 2021-10-08).
- [7] Marta Juźwin, Paulina Adamska, Maria Rafalak, Bartłomiej Balcerzak, Michał Kąkol, Adam Wierzbicki. "Threats of Using Gamification for Motivating Web Page Quality Evaluation". MIDI '14: Proceedings of the 2014 Multimedia, Interaction, Design and Innovation International Conference on Multimedia, Interaction, Design and Innovation. Warsaw, Poland, 2014-06-24/25. Association for Computing Machinery, 2014, p. 1-8. <https://doi.org/10.1145/2643572.2643587>, (参照 2021-10-08).
- [8] 川田拓也, 赤峯亨, 河原大輔, 加藤義清, 乾健太郎, 黒橋禎夫, 木俣豊. "情報分析システム WISDOM のユーザ評価とその分析". 言語処理学会第17回年次大会発表論文集. 豊橋技術科学大学, 2011-03-07/10, 言語処理学会. 言語処理学会, 2010, p. 45-48. https://www.anlp.jp/proceedings/annual_meeting/2011/pdf_dir/B1-4.pdf, (参照 2022-01-02).
- [9] "各党の公約 政策別・政党別 衆議院選挙 2021". NHK 選挙 WEB. 2021-10-21. <https://www.nhk.or.jp/senkyo/database/shugiin/2021/pledge/>, (参照 2022-01-02).