

令和 6 年 6 月 6 日現在

機関番号：12601

研究種目：若手研究

研究期間：2021～2023

課題番号：21K13278

研究課題名（和文）実社会の社会経済的状況とフェイクニュースの関連性や影響に関する実証分析

研究課題名（英文）Empirical analysis on relationships between misinformation and disinformation in cyberspaces and socioeconomic situations

研究代表者

渋谷 遊野（Shibuya, Yuya）

東京大学・空間情報科学研究センター・准教授

研究者番号：20847917

交付決定額（研究期間全体）：（直接経費） 3,600,000円

研究成果の概要（和文）：デジタル空間上での偽誤情報などの人々を誤解させるような情報流通の実態を把握し、デジタル空間上から実社会へ、また実社会からデジタル空間への直接的・非直接的な関係性を明らかにすることが求められる。そこで本研究は、デジタル空間上の偽誤情報等の情報流通と実社会の関係性に着目し、情報提供者と情報消費者それぞれの情報流通で果たす役割や現状、課題を実証分析を通じて明らかにした。具体的には、COVID-19や令和六年能登半島地震時の国内での偽誤情報流通状況の多面的な解析や実証分析のほか、本研究期間中に利用が拡大した生成AIなどが与える偽誤情報流通の状況や課題に研究の範囲を広げて実証分析を行った。

研究成果の学術的意義や社会的意義

偽誤情報等流通の実証研究はデータの入手可能性から研究の緒についたばかりで、特に、国内データを用いた分析は限られていた。本研究は、偽誤情報等の生成・流通・拡散に関する理論を国内外の大規模データ解析を通じて多角的に実証的に検証した。国内の偽誤情報の対応を巡っては、プラットフォーム事業者ごとの取り組みに委ねられてきたが、欧米で取り組まれる制度的枠組みの整備の必要性に関する議論も国内で行われるようになってきていて、本研究の成果は政策検討の一資料となることが期待される。日本を含む各国で深刻化する偽誤情報等の社会経済的影響の国内での検討に向けた基礎的資料となることが期待される。

研究成果の概要（英文）：This research explored intertwined relations between the physical and digital spheres in the context of the dis/misinformation and propaganda in Japan. This study empirically analyzed the cases of misinformation distribution in Japan at the time of COVID-19 and the 2024 Noto Peninsula earthquake. Also, we conducted online experiments by expanding the scope of our research to misinformation distribution in the age of generative AI. While the current approach taken in Japan to address misinformation has been largely left to the efforts of platform operators, there are ongoing discussions about the need to establish an institutional framework. This research offers insights that can be used to inform policy decisions.

研究分野：社会情報学

キーワード：偽誤情報 偽情報 誤情報 フェイクニュース 生成AI 令和6年能登半島地震 COVID-19 プロパガンダ

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

### 1. 研究開始当初の背景

デジタル空間上での偽誤情報などの人々を誤解させるような情報流通の実態を把握し、デジタル空間から実社会へ、また実社会からデジタル空間への直接的・非直接的な関係性を明らかにする必要がある。偽誤情報はデジタル空間での正の外部性を低下させ、報道やメディア、政府等の情報に対する懐疑心や不信感を高めることや、資源の誤配分、実社会への混乱を引き起こす可能性が指摘されてきた。他方、日本では、欧米のように選挙時に大規模な偽誤情報キャンペーンが行われて民主主義が脅かされた経験がなく、偽誤情報の脅威を社会全体として「自分ごと」として捉えて対応する準備が十分に整っているとは言い難い。デジタル空間上での偽誤情報等の流通によって、消費者は正確な情報を得ることが困難となるだけでなく、情報取得のためのコストが増加する。一方で、消費者は自らの信念をサポートするような情報を選好することにより偽誤情報から効用を得ている場合もある。偽誤情報に関する研究では、偽誤情報の動機や影響、偽誤情報の流通状況の把握推進により得られる利益とコスト、必要な政策的対応などの探求が求められているものの、実証研究はデータの入手可能性から研究の緒についたばかりで (Kogan et al., 2018)、特に、国内データを用いた分析は少ない。そこで本研究では、偽誤情報や情報流通に関する理論を国内外の大規模データ解析を通じて実証的に検証し、デジタル空間上の偽誤情報との情報流通の実態と実社会の社会経済・政治的な状況との関連性の解明を目指す。国内の偽誤情報の対応を巡っては、プラットフォーム事業者ごとの取り組みに委ねられてきたが、欧米で取り込まれる制度的枠組みの整備の必要性に関する議論も国内で行われるようになってきていて、本研究の成果は政策検討の一資料となることが期待される。また、本研究は、偽誤情報の生産流通を既往研究で示されてきた理論枠組みの中で実証的に捉え直すとともに、社会経済状況との直接的・非直接的関係性を解明し、日本を含む各国で深刻化する偽誤情報等の社会経済的影響の国内での検討に向けた示唆を得ることが期待される。既往研究の多くはメディア間での自由競争と消費者需要による言説の極化が主な議論の対象となっているが、複雑に関連しあう実社会とデジタル空間の関係性を分解し、偽誤情報の実社会への波及効果やその逆の波及効果なども対象とする。

### 2. 研究の目的

本研究では既往研究で議論されてきた理論に基づき、国内外のデジタル空間言説等の大規模データ解析を通じて実証する。特に、どのような市場参加者が偽誤情報の生成・流通に貢献しているのか、デジタル空間と実社会間でどのような関係性やスピルオーバー効果がみられるかなどを明らかにすることを通じて、今後求められる政策的検討の基礎的資料を提示する。

### 3. 研究の方法

偽誤情報に関連する既往研究と、デジタル空間言説の既往研究を対象に系統的レビューを行い理論や研究手法などを整理する。その上で対象とする偽誤情報事例を抽出し、プラットフォーム事業者の公式 API 等からデータを取得し、データ解析を行う。また、情報の消費者側の行動を把握するため、大規模なオンライン実験を行う。

### 4. 研究成果

#### (1) デジタル空間の偽誤情報流通に関する既往研究の整理

本研究では、まず偽情報の定義を関連語と対比・概観しながら偽誤情報に関連する研究のレビューを行ったほか (澁谷 2022)、さまざまな分野で研究活動が活発化しているソーシャルメディアデータ分析の手法について系統的レビューを行った (Shibuya et al., 2022)。まず、偽誤情報に関連する用語として、偽情報や誤情報、暴力的な過激主義やヘイト・ハラスメントやその関係、それぞれにまつわる研究をまとめた (図 1 : 澁谷 2022)。続いて、系統的レビューでは、人々とソーシャルメディアとの相互作用であるソーシャルメディア・インタラクシオンに関する関連研究を対象にレビューし、関連研究分野における研究トピックや研究手法、分析対象データの特徴とその傾向を明らかにした上で、今後の研究課題や研究手法に関する考察を提示した (Shibuya et al., 2022)。

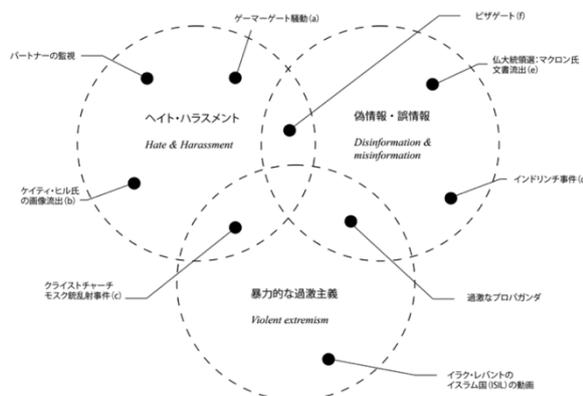


図 1 オンラインハラスメント・ヘイトに関する分類 (Thomas et al. (2021) Fig.1)

#### (2) COVID-19 禍の Facebook 上での偽誤情報の流通状況

研究蓄積の浅い国内の Facebook 上での COVID-19 関連偽情報等の生成・流通の概要を把握することを目的に、2020年1月から2021年5月までの18ヶ月分の日本語でのCOVID-19関連Facebook投稿を記述的に分析した(澁谷 2022)。その結果、国内外の関連研究で示されている他プラットフォームや他言語による偽情報の傾向と同じように、(1) 少数アカウントが偽情報生成・流通で中心的な役割を担っていて、主流メディアのアカウントと同等もしくはそれ以上の反応数を獲得するなど、COVID-19 関連の投稿としては最大級の反応を得ていることが明らかになった。また、(2) 偽情報の発信の動機としては、金銭的なインセンティブやイデオロギーに基づいた動機によると考えられるものがあること、(3) YouTube 等の外部情報源を利用しながら偽情報を流通させていること、(4) 偽情報の中心的なアカウントの周りには緩やかに繋がった大小様々なグループやアカウントが存在し協調的に偽情報の生成・流通に寄与していることが示唆された。偽情報等を展開するアカウントの中には、プラットフォーム事業者によるアカウントや投稿の削除への対策として複数アカウントを運営したり、既存コミュニティとつながることで、偽情報等を受容しやすい潜在偽情報消費者を獲得しているケースもみられた。こうしたことから、COVID-19 関連偽情報を巡ってはプラットフォーム事業者等の単一的なアカウントの削除等の対応のみによる効果は限られる可能性が示唆される。今後は国内の偽情報等の生成流通は、プラットフォーム事業者側の問題やリテラシーを中心とした個々の問題としてのみ捉えるだけではなく、社会経済システムの症状としても捉え直し、その背景にある社会・経済・政治な背景や偽情報需要の増減のメカニズムやエコシステムの解明を行うなど、多面的なアプローチも必要と考えられる。

### (3) COVID-19 禍の Twitter 上での専門家の役割の検討

まず、Twitter 上のユーザーのうち専門家が情報流通で果たす役割に焦点を当てた分析を行った(Nakazato et al., 2023)。ワクチンの誤報の拡散は、疾病発生のリスクを高めるなど、公衆衛生に深刻な結果をもたらす可能性がある。専門的な知識を持つ専門家は、ワクチンの誤報に対抗する上で重要な役割を果たすことができるにもかかわらず、COVID-19 初期段階において、医療専門家がソーシャルメディア上でどのようにコミュニケーションをとっていたのかについては、十分な理解が得られていない。そこで本研究では、日本語の大規模 Twitter データを用いて、2020年のワクチンに関する主要な議論に医療専門家がどのように参加していたかを検証した。具体的には、ソーシャルメディアのプロフィールから TF-IDF (Term Frequency - Inverse Document Frequency) アルゴリズムを用いて医療専門家ユーザーを特定し、ワクチンコミュニケーションにおける彼らの特徴的な行動を分析した。その結果、医療専門家はワクチン推進派である可能性が高く、政治的右派や中立派である可能性は低いことがわかった。また、これらのユーザーは、非医療専門家ユーザーと比較して、より多くのリプライを受け取り、より多くのフォロワーを持ち、非専門家と比較して、異なるタイプの情報源を利用する傾向があることがわかった。これらの結果は、ソーシャルメディア領域における医療専門職ユーザーのユニークな立場を示す。

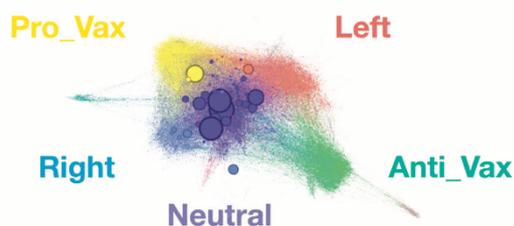


図 2 医療専門家ユーザーのフォロワー数を反映したリツイートクラスタのネットワーク

### (4) COVID-19 禍の Twitter 上での道徳基盤や有害性に関する分析

また、ワクチンに関する投稿の道徳基盤や有害性に関する解析を行った(中里他 2024)。具体的には、日本でのデジタル空間上の反ワクチンの言論流通状況の包括的な把握を目指し、COVID-19 拡大初期(2020年)の COVID-19 ワクチンに関するソーシャルメディアの投稿を対象に、ツイートにおける有害性の高さや COVID-19 関連のタイムライン、ユーザの投稿に表れるトピック、道徳基盤との関係性、情報の拡散状況について解析した。分析の結果、時系列やトピック、道徳基盤によって表出する有害性の程度が異なることが明らかになった。関連投稿の有害性が高い傾向のあるトピックについては、より積極的に正確な情報の発信を行うなどの対策が考えられる。また、ユーザごとの投稿数の分布については、有害性の低いグループでは、多数ユーザが少量ずつの投稿をしていることが示された。これは、少数のユーザーがより有害性の高いツイートを大量に投稿しているという予想とは反する結果であった。全体の投稿に対して分析を行

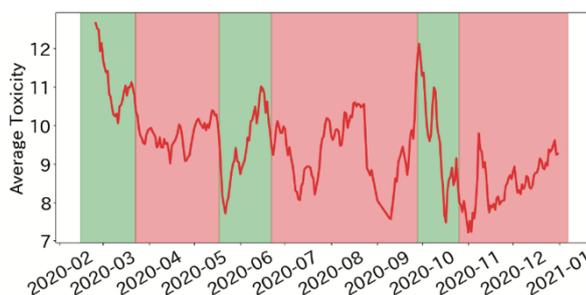


図 3 COVID-19 流行波ごとのツイートの有害性の時系列変化 (7日間の単純移動平均)

なっているため、企業等の広告アカウントや無関係な bot の投稿数が分析結果に反映されている可能性があり、今後の研究課題として、これらのユーザ属性を踏まえた分析も必要である。

#### (4) 令和 6 年能登半島地震時の偽誤情報の流通状況

令和 6 年能登半島地震発災後初期の偽誤情報等の流通状況を把握するため、偽誤情報の収集・類型化および情報流通の特徴のトレンドの分析を行った。また、アルゴリズムやプラットフォームに備わる機能の変更が情報流通にもたらす影響について、ソーシャルメディアプラットフォーム X に着目して調査した。質的分析の結果、偽誤情報の類型、意図・動機、トピックは多様であることが明らかになった。さらに、量的分析によって、流通する偽誤情報トピックの時系列的变化、日本語以外の言語を使用するアカウントによる複製投稿活動、外部情報源の利用など、発災後の X における情報流通の傾向が示唆された。複製投稿活動など、X の仕様変更による影響を精査する必要がある。さらに、近年のコミュニティノートによって投稿に背景情報を追加する取り組みは、能登半島発災後ユーザサイドからの一種のファクトチェックとして機能した可能性がある。本稿が明らかにした、公開されたコミュニティノートが対象とするポストのエンゲージメントが低い傾向は、コミュニティノートの公開と情報の信頼性評価との関連について示す。また、本稿における分析は、コミュニティノートによる情報付与の対象となるユーザの偏向や、トピックごとの評価の偏向についても示唆を与える。アルゴリズムや機能の変更が情報流通にもたらす影響については、プラットフォーム事業者の仕様変更により、情報流通のあり方が変化する可能性がある。

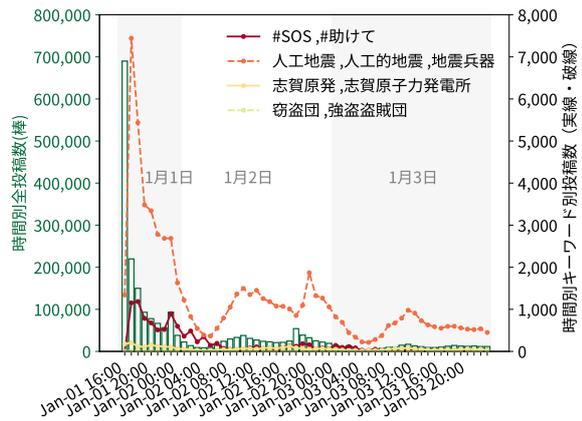


図 4 時間帯別の投稿数 (棒・緑色) とキーワード別の投稿数 (実線・破線)

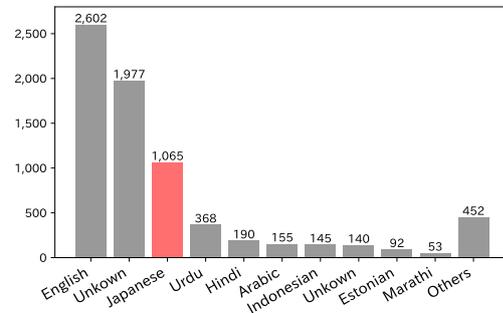


図 5 複製投稿として抽出された投稿群のうち、投稿者の推定使用言語別の投稿数

#### (5) デジタル空間での生成系 AI 偽誤情報消費者を対象としたオンライン実験

本研究を実施する 3 年間でもデジタル空間上の情報の種類や質は人工知能技術 (AI) の大衆化とともに大きな変化があった。特に生成 AI を活用した偽誤情報流通は 2021 年以降急速に広まっている。さらに、当初はテキストや画像がデジタル空間の情報流通の主流であったが、動画による情報流通の拡大が進んでいる。そこで、本研究では、生成系 AI 時代における偽誤情報対応策の検討を行うことを目的に、生成系 AI によって作成された動画を用いた大規模なオンライン実験を行った。AI によって生成されたコンテンツは、しばしば人間が作ったコンテンツと区別がつかないため、誤った情報やプロパガンダが広まることへの懸念が高まっている。これに対応するため、主要なプラットフォーム (Google, Meta, TikTok など) は、AI が生成したコンテンツについてユーザーに警告するポリシーを導入している。潜在的な課題のひとつは、AI に対する一般市民の認識が分かれているため、こうした警告に対する反応が異なる可能性である。AI が社会にもたらす潜在的な利益を評価する人もいれば、潜在的なリスクについて悲観的な人もいる。このような二極化した態度が、誤った情報に対する取り組みにどのような影響を与えるかは、まだ明らかになっていない。本研究では、AI に対する人々の態度が、AI が作成した投稿に対する彼らの認識にどのような影響を与えるかを調査した。実験では、参加者に、AI が生成した動画コンテンツの正確性についての判断に影響を与えた要因を自由回答形式で報告した。その結果、ほとんどの参加者は、AI が作成したコンテンツに対する警告メッセージが表示された場合でも、AI が作成した投稿を評価するために、既存の知識や信念に依存していることがわかった。興味深いことに、一部の参加者は、AI に対する肯定的または否定的な信念に基づいて、すべての動画の精度を肯定的または否定的に評価した。この発見は、生成 AI が作成したコンテンツに単純な警告を組み込むことは不十分である可能性があり、過小評価から無反応、過剰な反応まで、ユーザーにさまざまな影響を与えることを示唆している (Shibuya et al., 2024; Shibuya et al., Under Review)。

## 5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 2件/うちオープンアクセス 3件）

1. 著者名 澁谷遊野	4. 巻 10:3
2. 論文標題 FacebookにおけるCOVID-19関連偽情報の国内主要アクターとその特徴	5. 発行年 2022年
3. 雑誌名 社会情報学	6. 最初と最後の頁 13-30
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Yuya Shibuya, Chun-Ming Lai, Andrea Hamm, Soichiro Takagi, and Yoshihide Sekimoto	4. 巻 12:1
2. 論文標題 Do open data impact citizens' behavior? Assessing face mask panic buying behaviors during the COVID-19 pandemic	5. 発行年 2022年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 17607
掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/s41598-022-22471-y	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Yuya Shibuya, Andrea Hamm, and Teresa Cerratto Pargman	4. 巻 129
2. 論文標題 Mapping HCI research methods for studying social media interaction: A systematic literature review	5. 発行年 2022年
3. 雑誌名 Computers in Human Behavior	6. 最初と最後の頁 1-16
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.chb.2021.107131	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Yuya Shibuya, Tomoka Nakazato, Soichiro Takagi	4. 巻 -
2. 論文標題 How do people evaluate the accuracy of video posts when they know the videos were created by a Generative AI?	5. 発行年 2024年
3. 雑誌名 査読中	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計8件（うち招待講演 0件 / うち国際学会 7件）

1. 発表者名 Avishai Green, Yuya Shibuya, Tommaso Trillo, Saki Mizoroki, Rebecca Scharlach, Blake Hallinan, Tim Park, Effi Levi, Elad Segev, Shaul Shenhav, and Limor Shifman
2. 発表標題 How to Do Things with "Values": A Cross-Linguistic Analysis of the Meanings and Functions of a Core Concept on Twitter.
3. 学会等名 AoIR (国際学会)
4. 発表年 2022年

1. 発表者名 Etienne Gagnon, Kenneth McElwain, and Yuya Shibuya
2. 発表標題 Good News or Bad News? Political Social Media Messaging During the Tokyo Olympics
3. 学会等名 JPOSS (国際学会)
4. 発表年 2023年

1. 発表者名 Yuya Shibuya
2. 発表標題 Intertwined Relations between Communication in Cyberspace and Economic and Political Situations in the Physical World.
3. 学会等名 Early-Career Scholar Forum, UTokyo Center for Contemporary Japanese Studies (2022) (国際学会)
4. 発表年 2022年

1. 発表者名 Shibuya, Yuya, Hamm, Andrea, Raetzsch, Christoph
2. 発表標題 From Data to Discourse: How Communicating Civic Data Can Provide a Participatory Structure for Sustainable Cities and Communities
3. 学会等名 The International Sustainable Development Research Society (ISDRS) (国際学会)
4. 発表年 2021年

1. 発表者名 Tomoka Nakazato, Yuya Shibuya, Soichiro Takagi
2. 発表標題 Characterizing the Behavior of Healthcare Experts Towards COVID-19 Vaccine on Twitter
3. 学会等名 Proceedings of WI-IAT2023 (国際学会)
4. 発表年 2023年

1. 発表者名 Kai-Yi Tan, Chun-Ming Lai, Yin-Jie Foo, Yuya Shibuya
2. 発表標題 Russo-Ukrainian War on Facebook: A Temporal and Spatial Study Based on BERTopic modeling
3. 学会等名 International Conference on Security and Information Technologies with AI, Internet Computing and Big-data Applications (SITAIBA 2023) (国際学会)
4. 発表年 2023年

1. 発表者名 中里朋楓, 澁谷遊野, 大西正輝, 高木聡一郎
2. 発表標題 Twitter上のCOVID-19ワクチンに関するトピック・道徳基盤と有害性の関係性
3. 学会等名 2024年度 人工知能学会全国大会 (第38回)
4. 発表年 2024年

1. 発表者名 Yuya Shibuya, Tomoka Nakazato, Soichiro Takagi
2. 発表標題 AI Optimism, Pessimism, or Indifference? Challenges of Combating AI-Made Misinformation Under Mixed Perceptions of AI
3. 学会等名 Digital State Sponsored Disinformation and Propaganda: Challenges and Opportunities at the International AAAI Conference on Web and Social Media (国際学会)
4. 発表年 2024年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
イスラエル	エルサレム・ヘブライ大学			
ドイツ	ベルリン工科大学			
スウェーデン	ストックホルム大学			
その他の国・地域	Tunghai University			