

## 博士請求論文審査要旨

情報セキュリティ大学院大学  
情報セキュリティ研究科

論文題目 : 怒りが Disinformation の共有に及ぼす影響とその対策  
申請者 : 鈴木 悠  
審査委員会 : 主査 教授 後藤 厚宏  
副査 教授 稲葉 緑  
副査 教授 大久保 隆夫  
副査 教授 村上 康二郎

## I. 論文内容の要旨

本研究は、ソーシャルメディア上における Disinformation、すなわち悪意をもって作成された虚偽または誤解を招く情報の無計画な共有抑制を目的として、この共有を助長する怒りに働きかけるユーザ向けの介入策を提案するものである。

本論文は8つの章から構成される。

第1章では、本研究の背景と必要性、および貢献について述べている。近年発展したソーシャルメディアは速報性と拡散性に優れる一方、Disinformation の短時間での拡散も可能とし、社会や人々が混乱する事態を招いている。本研究は、感情的要因、特に怒りに着目した新たなユーザ向け介入策の提案を目的とする。社会の混乱を誘発する Disinformation には強い怒りを喚起させるものが多く、これがユーザの情報共有への心理を刺激する可能性が示されてきた。しかし、このような情報共有の抑制にあたり、既存の法律や情報の信憑性判断支援等の対策には限界がある。本研究の貢献は、Disinformation の共有メカニズムを踏まえて怒りにまかせた共有について一考させる介入策を提案したことで、社会の混乱防止に貢献することである。

第2章では、民主主義国家における既存の Disinformation 対策およびその研究を広く調査した。結果、表現の自由のため、投稿された Disinformation に対しては注意喚起や行動規範遵守に留まること、情報のファクトチェック判定が難しいことなどの課題が示された。一方、デバンキング・反論、プレバンキング、メディア情報リテラシー教育の主目的は真偽判断の促進であるが、怒りによる投稿共有にも適用し得る可能性が確認された。

第3章では先行研究より、感情、特に怒りが情報の信憑性を高く評価させることで、Disinformation の共有を促進する可能性を把握した。これに基づき、第2章で適用可能性を認識した真偽判断の促進策の効果が怒りによって減じられる課題を指摘した。この課題を回避するためには怒りに直接働きかける必要があり、情動調節の発想を導入したインターフェース上の介入策、すなわち、ナッジの有効性を予測した。

第4章では本研究が目的とする怒りによる Disinformation 共有の抑制策を提案するにあたり、次の3つの課題を解決する必要を述べた。第一に、怒りによる Disinformation の共有メカニズムを解明する。第二に、複数の種類が提案されている情動調節手法を導入したナッジをそれぞれ試作し、効果の高いものを選定する。第三に、選定したナッジが現在最も普及している対策であるメディアリテラシー教育よりも怒りによる Disinformation 共有抑制には有効であることを示す。

第5章では、第4章で挙げた第一の課題についての検証結果を報告した。調査の結果、先行研究で可能性の

提示に留まっていた、怒りによる Disinformation が共有されるメカニズムを確認した。怒りは情報の信憑性を高く評価させるほかに、信憑性判断を関与させずに Disinformation を共有させることが示された。このように二経路で共有が促されることが、怒りによる Disinformation の共有されやすさの背景にあると示唆された。

第6章では、第二の課題に関して、怒りによる Disinformation 共有に有効なナッジの選定評価について報告した。予備実験では情動調節の複数の手法を反映した情動調節ナッジをそれぞれ試作し、その中で「視点取得ナッジ」と「気晴らしナッジ」の2種の有効性が高い可能性が示された。大規模調査の手法で実施した本実験ではこの2種のナッジの効果を評価し、「気晴らしナッジ」の方が有効であるとの結果が得られた。

第7章では、第三の課題への取り組みを示した。怒りによる Disinformation 共有を抑制する効果について、第6章で最も有効であることが示された情動調節ナッジである気晴らしナッジと最も普及している対策の一つであるとメディアリテラシー教育との間で比較した。結果、教育を受けた参加者による投稿の共有には喚起された怒りが大きく関与することが示された。一方、気晴らしナッジを提供された参加者による共有では怒りの関与が低下し、代わりに信憑性判断の関与が高まった。ここから、気晴らしナッジが怒りによる Disinformation の共有を抑制する可能性が示唆された。

第8章では、本研究の貢献と限界を示した。貢献は、実験を通して、怒りが Disinformation の共有を促すメカニズムを明らかにし、そのメカニズムを踏まえて共有についてユーザに一考させる情動調節ナッジとして気晴らしナッジを提案したことにある。感情に直接的に働きかける対策として、真偽判断促進などの既存の対策を補完する可能性が期待される。また、プラットフォーム企業が気晴らしナッジを実装するにあたって解決すべき技術的課題について具体的に示した。限界は、情動調節ナッジの共有抑制効果について限定的な Disinformation を対象とした実験環境での評価にとどまっている点、強い怒りを喚起する真実の共有抑制や情動調節ナッジ表示を悪用した偽情報共有の悪化等、ナッジの悪影響の検証までは網羅していない点が挙げられる。

## II. 論文審査結果の要旨

本論文は、既存の Disinformation 対策の限界によるソーシャルメディアの無秩序化に歯止めをかけることを目指し、怒りに助長された Disinformation 共有のメカニズム、および、これを抑制する介入策として情動調節ナッジの有効性を実験的に明らかにした。これらの成果は、今後悪化が予想される情報戦に備えた国内の偽情報対策に関する議論や規制整備への活用が期待され、情報学への貢献が認められる。

なお、情動調節ナッジ実装の前提として必要な、社会やプラットフォームの理解をいかに得るかについての具体的な構想が明示されていれば、本論文はより説得力を持ったであろう。しかし、上記の点は本論文の博士（情報学）の論文としての意義を否定するものではない。

以上の理由から、本論文は、博士（情報学）の論文として合格と認められる。

## III. 審査経過

本審査委員会は、2025年7月26日に論文内容とこれに関連する事項について口述試問を行った。審査に当たっては、博士学位のディプロマ・ポリシーに基づいて総合的に評価し、申請者が学位取得にふさわしい知見を持つものと判断した。