

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 26 日現在

機関番号：34416
研究種目：挑戦的萌芽研究
研究期間：2012～2014
課題番号：24653131
研究課題名(和文) ソーシャルメディアからの世論推定と民意抽出

研究課題名(英文) Social media and public opinion

研究代表者

安田 雪 (Yasuda, Yuki)

関西大学・社会学部・教授

研究者番号：00267379

交付決定額(研究期間全体)：(直接経費) 2,900,000円

研究成果の概要(和文)：本研究では、ソーシャルメディアからの民意推定を行うべく、Twitterデータの収集と分析システム(WTC)を構築した。参議院選等を含むの多様なテーマのtweet(つぶやき)のデータセットを収集しその特長を検討した。また、いわゆる「ネット上の炎上」を含む、ソーシャルイベントに関するデータから、ソーシャルイベントへのユーザーの反応の類型化を行ない、民意のパターン推定の理論枠組みを提案した。

研究成果の概要(英文)：We build a Web Tweet Crawler System (WTC) for collecting Tweets data. This project also examined how dissemination by Twitter occurred and how hubs and ordinary people contributed diffusion on social media. Examining various Tweets data, including demagogue data, election data, social-event data, we found opinion leaders, followers and ignorant influencers did affect people's recognition of social reality. Aggregation and converting micro tweets data macro, so called "big data" allowed us to monitor and sensor quasi public opinion. We propose categorization of social media users' reaction to various social events.

研究分野：社会学

キーワード：twitter 社会的ネットワーク つぶやき 民意 SNS ソーシャルメディア 社会的イベント

1. 研究開始当初の背景

2012年の研究開始当初は、いわゆる「ビッグデータ」時代の幕開け期であり、コンピュータのデータ処理能力の飛躍的發展にともない人間が扱いうるデータ量が膨大になるとともに、ソーシャルメディア上で発信されるデータも膨大に蓄積され、情報科学のみならず社会科学の研究対象となり始めた時期であった。

また、この時期はコンピュータとともに携帯電話によるソーシャルメディアの利用が普及し、通信速度の飛躍的な高速化にともない、市民によるソーシャルメディアの利活用が活発化しかかっていた。ソーシャルネットワークサービス(SNS)やミニブログなども含む、ソーシャルメディアの学術的定義は多様だが、なかでもTwitterは2011年の東日本大震災時に注目を集めたこともあり、Twitter上の発言(Tweet)の内容や量から、世論、民意、あるいは消費動向などを予測する技術やそのための基礎となる理論に対する、学術的及び商用的なニーズが顕在化し、情報科学の分野において国際的な競争が展開され始めていた。研究終了時点の現在と比べ、ソーシャルメディアからの民意推定や世論の考察において、Twitter上で発信される情報を対象とすることの妥当性、さらにはTwitterの発信力及び影響力についての社会的合意がえられていなかったことを、研究提案時の背景として記しておく。

2. 研究の目的

本研究の目的は、前述した背景をふまえ、研究期間においてソーシャルメディア上に膨大に蓄積されつつある膨大なデータを収集・分析し、それらの発信者の特性、集団的な民意の推定及びそのための実践的方法の検討を行うことである。

具体的には、(1)Tweetデータの学術的利用

のための収集・分析システムの構築、(2)特定のソーシャルイベントに関わるTweetデータの収集と収集したデータのデータベース化、(3)Twitterユーザーのソーシャルイベント時の行動特性調査(例えば、選挙前のTwitterユーザーの行動と意識に関するWeb調査)、(4)Twitterユーザーのフォロー追跡によるユーザープロフィールの特定と行動の関連検討、(5)多様なソーシャルイベントへのユーザーの反応の類型化、(6)以上を統合してのTwitter上の情報拡散と民意・世論の推定に関わる理論的検討を行うことである。

3. 研究の方法

研究統括である安田と分担者である鳥海はデータ収集、分析、理論的考察、実証研究を部分的に共有し、さらには本プロジェクト外の研究者との共同研究をも連関させつつ、本研究を展開した。

以下、時系列的に、本研究で実施した分析内容と方法を、研究結果も交えつつ概括する。

2012年度には、ソーシャルメディアからの民意推定を行うべく、Tweetデータの収集と分析体制を、鳥海が中心に構築し、年度内にTweetデータ収集システムの基本的体制を整えた。また、安田(2013)研究開始時に収集されていたソーシャルメディア上のデマ拡散事例のプロセスを再検討し、Twitterユーザーのハブの役割とともに、イグノラント・インフルエンサー(無知なる影響者)概念を精緻化した(安田, 2013)。本概念は、日本学術振興会によるプロジェクト「東日本大震災から学ぶ社会科学」、メディアとネットワーク班におけるソーシャルメディアの役割についての論考へ発展する出発点となった。

2013年度には、2012年度に鳥海が構築したシステムを活用して、2013年12月の参議院選挙直前に、政党名、当該選挙にあたり争点と考えられるキーワード群を決定し、それらをクエリーとしてTweetデータの収集を

行った。結果として、11,420,421 件の Tweet からなる「2013 年参議院議員選挙 Tweet データ」を収集した。この後、Twitter 社の API 制約条件を含むデータ収集上のメモリ及びハードディスク容量、回線速度の制約条件などの再検討と増加する Tweet データ収集のための新たなシステム構築の必要性を確認、対応策の検討を行った。

同時に、Tweet データの分析に際して注目すべき重要な特徴量は何かを明らかにすべく、収集したデータについて目視も含めて探索的分析を行った。一連の探索的分析の過程でもっとも注意すべき点として明らかになったのは、Twitter でアカウント名として取り扱われることの多い、ScreenName と、表示名すなわち Twitter 上のいわゆる「名前」である(Name)が同一ではないことである。つまり、ScreenName も Name もユーザー個人の一貫性を保証するものではないということである。

同一人物による複数 ScreenName の使用がありうること、さらには、同一の ScreenName であっても、それが時期によってはまったく異なるユーザー（個人）でありうる可能性があるのである。両者の差異は長期的に収集された Tweet データからの民意の推定には影響を及ぼしうる可能性である。これが Tweet データを用いて民意推定を行う際のミクロレベルでの複雑さを増す一つの要因であり、この点については、鳥海(2015)において指摘されている。

この知見をふまえ 2013 年度には、さらに、池田・安田(2013)による「参院選時のツイッター利用者調査」と部分的に共同した 1000 名を対象とした Web 調査を実施し、Twitter ユーザーの利用行動についての検討を行った。本研究の成果として注目したいのは、この調査の回答者は、対象者の選定時点で Twitter のアカウント名 (ScreenName)の回答提供に同意する者に限定されていた。個人

情報は提供されないものの、Twitter のアカウント名が回答に含まれているのである。「参院選時のツイッター利用者調査」は、これらの対象者について Web 調査により、ユーザーの参議院議員選挙時の実際の投票行動・支持政党をはじめ、社会的参加についての意識を尋ねており、ソーシャルメディア利用の実態及びユーザー意識を明らかにするものである。

さらにこの調査データを追跡すべく、2014 年度には、2013 年度の調査対象者のうち半数の 500 人を抽出して、個別にフォローすることで、そのプロフィール特性を事後的に関連づける試みを行った。(このデータは非公開データベースとして保存している。) Twitter ユーザーの公開されているプロフィールと行動とを対応させた研究を行うためである。

Twitter ユーザーのプロファイルからは、Twitter 利用の継続性、鍵の有無、凍結状態、性別、年齢、アカウント名、所在地、キーワード、URL、Twitter 参加日、Tweet 数、フォロー数、フォロー数が特定しうる。さらに、Twitter 利用者はその他にも、個人のアカウント数、課金の有無、ニュース閲覧頻度、Tweet の内容(日常のつぶやき、テレビを含むマスメディア内容について、URL のコピーアンドペーストの利用状態)、Twitter の連携機能の使用の有無、#の利用、トレンド検索の利用、メッセージの利用など、Twitter の機能の多様化にとまない、注目すべき特徴をもつ。

これらの一部は、アカウント名をフォローし続けることで短期的にはデータとして把握できるものもあるが、実質的には Twitter 社の協力なしに外部から直接アクセスしうる属性情報には大きな制約があり、発信者の個別情報との紐付けは困難であり、アカウント名のフォローだけでは実態との紐付けが困難であることが明らかになった。

さらに、「参院選時のツイッター利用者調査」の回答者のうち、都知事選に選挙権をもつ者を対象にした「東京都知事選挙に関する調査」(2014)も実施し、追跡調査を行っているが、ユーザー行動、Twitter 上のプロフィールと属性との対応関係は検討が継続中である。

最終年度の 2014 年には、鳥海らがソーシャルイベントに対するユーザーの反応の類型化を検討し、民意推定の一つの視座を切り拓いた。鳥海ら(2014)は、いわゆる「ネット上の炎上」を特殊なソーシャルイベントとしてとりあげ、複数のイベント分析を通じて、ソーシャルイベントへの反応を類型化した。この類型化により、Twitter ユーザーの多くが、炎上しやすい言説について対してどのような反応をするのかが具体的に明らかになった。この知見から、伝統的に民意あるいは世論と概念化されてきたものと、ソーシャルメディア上で観察される、ユーザーの多様な社会的イベント、とりわけ炎上現象の発火について理論的検討の糸口が開かれた。

4. 研究成果

研究期間の 2012 年～2014 年は、ソーシャルメディア全般を俯瞰すれば、技術的な大きな変化が観察された時期である。研究開始時点は、ソーシャルメディアを用いたビジネスの立ち上がり期でもあり、終了時点までの間に、日記による SNS サービスを主としていた mixi が相互交流機能からオンラインゲームビジネスへの完全な転換を図ったこと、ブログと Twitter のユーザーによる使い分けが進んだことなどが確認されている。さらに Tweet データの収集と分析が急速に商業化され普及するとともに、我々が 2011 年に挑戦的とした位置づけた Tweet データの収集とその分析からの民意推定は、統計物理学の研究者により、流行現象、議員や人気アイドル選挙結果の予測がなされるなどの展開が進

んだ。Tweet データの分析はその規模を拡大するとともに質的な変容を遂げたわけだが、これらについては、研究代表者と分担者が設立当初からその一翼を担い、本研究の成果についての議論も深めた、後述する、ソーシャルメディアワークショップでの議論が貢献したと考えている。

この間、個人間コミュニケーションが Twitter のメッセージ機能よりも不可視性の高い LINE へ吸収されているという変化も観察され、Twitter 上の個人間コミュニケーションの機能とその内容には、質的な変化が継続的に起こっていることを確認した。この変化が Twitter からの民意推定に及ぼす効果は、今後の検討課題である。さらに、Twitter に依存したソーシャルメディアからの民意推定は、常にデータ提供者である Twitter 社の API の仕様変更や、データ入手条件の厳格化への対応という制約を負い続けることが、研究期間中の度重なる Twitter 社のデータ利用に関わる仕様変更によって確認された。

一方で、Twitter をはじめ SNS により収集されるビッグデータ利用についての一般市民の利用認識や利用可能性についても社会的認識が変わりつつあることが明らかになったため、ソーシャルメディア上の個人情報の商用・学術的利用をめぐる研究倫理についてのサーベイを行い、倫理規範をあわせて検討した。その結果、個人によってソーシャルメディアに投稿された記事、個人と特定しうる可能性のある本研究の収集データは非公開とするものとした。

研究期間中に Tweet データの商用利用と民間企業による民意・世論及び浮動票の推定などが、研究提案時の予測よりも急速に進展した。だが、本研究では、現状として伝統的社会科学の分野における学術的な Tweet データの利活用研究が少ない現状をふまえ、鳥海が、学術目的及び一般市民の利用が可能な商用化ベースにのせない Twitter からのデー

タ収集システム(WTC)を構築した。このTweetデータの収集ならびに基礎的な解析方法を提案したことで、情報系研究者やプログラマーではない、社会科学研究者がそれぞれの視点でデータを分析、民意や世論推定を試みるための研究基盤をつくった。また、鳥海は、Tweetデータの商用利用の現状についてのサーベイも行った(鳥海、2015)。さらに、ソーシャルメディアデータ、とりわけTweetデータの学術的利活用法の現状については、安田が「組織科学」特集号(2015)にて検討を行い、統計物理学を含む最先端の研究事例を紹介し、組織論及び経営学研究への応用可能性を論じた。なお、当該特集号においては、本研究開始以前から継続的に安田・鳥海が参加し、ソーシャルメディアと民意についての議論を展開してソーシャルメディアワークショップ(SMWS)の報告内容が重点的にとりあげられている。

5. 主な発表論文等
(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 3 件)

鳥海不二夫 (2015) “Twitter 上のビッグデータ収集と分析” 「組織科学」48(4) pp.47-59. 査読無し

安田雪(2013) “ソーシャルメディア上の情報拡散の事例—東日本大震災時のデマの事例とハブの役割” 関西大学社会学部 紀要 45(1) pp.33-46. 査読無し

安田雪 (2012) “絆を作り関係を制する” 「Re」No.176, 建築保全センター. pp.7-13. 査読無し

〔学会発表〕(計 6 件)

鳥海不二夫、榊剛史(2014) 「社会的イベント発生時のソーシャルメディアにおける反応の分類」 Social Media Work Shop. 箱根 2014.10.05

池田謙一、安田雪(2014) 「東日本大震災後の情報環境・情報行動が 1 年半後の適

応に何をもたらすか」 Social Media Work Shop. 箱根 2014.10.04

松澤有、鳥海不二夫、大橋弘忠 (2014) “Why Majority Become Silent: Simulation Analysis with Vocal-Attitude Propagation Model” JWEIN2014. (電気通信大学) 2014.08.23

YASUDA, Yuki(2014) “Information Dissemination in Social Media: Hubs and Demagogues” International Network for Social Network Analysis: Sunbelt XXXIV (St.Pete, USA) 2014.02.20

安田雪(2013) 「ソーシャルキャピタルは絆かしがらみか」第 86 回日本社会学会大会(慶應義塾大学) 2013.10.12

安田雪(2012) 「東日本大震災時のデマツイートの拡散過程」第 85 回日本社会学会大会(札幌学院大学) 2012.11.04

〔図書〕(計 1 件)

池田謙一(編) 『震災から見えるメディアとネットワーク』第八章 池田謙一・安田雪・芝内康文 “必要な情報が届くために：情報環境と受けての対応関連性” (2015 年 7 月刊行予定) 東洋経済新報社.

〔産業財産権〕

出願状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況(計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
取得年月日：
国内外の別：

〔その他〕

ホームページ等

・安田雪 (2015) “特集「ビッグデータと新たな地平」に寄せて” 特集号「組織科学」48(4). pp.2-4.

6. 研究組織

(1) 研究代表者

安田 雪 (Yasuda, Yuki)
関西大学・社会学部・教授
研究者番号：00267379

(2) 研究分担者

鳥海 不二夫 (Toriumi, Fujio)
東京大学・工学(系)研究科(研究院)・
准教授
研究者番号：30377775