

Twitter における「コロナ」と「ワクチン」に関するツイート分析 Tweet analysis of "corona" and "vaccine" on Twitter

大谷 光一[†]
Koichi Otani

1. はじめに

近年インターネットの発展に伴い、膨大な情報が生み出されている。中でも **Social Networking Service** (以下 SNS と省略する) は幅広い年代で活用され、そこで生み出される情報も、著しく増加している。Twitter は、140 文字のテキストもしくは画像を投稿したり、フォローしているアカウントの投稿を確認したりすることができるサービスである。また、Twitter は自分がフォローしているユーザーだけでなく、他のユーザーが投稿したツイートを検索することができる。その言葉を含むツイートを調べることができることから「今、何が起きているのか」を知ることができる。その他にもユーザー同士で情報を交換し合う機能も実装しており、日本の Twitter のアクティブユーザーは 4,500 万人を超えている。

新型コロナウイルスに関する経緯について以下に記す。2019 年 12 月に中華人民共和国湖北省武漢市で「原因不明のウイルス性肺炎」として確認されて以降、世界的に感染が拡大し、世界保健機関 (WHO) が **coronavirus disease 19(COVID-19)** と命名し 2020 年 3 月 11 日にパンデミックを宣言した。日本では 2020 年 2 月大型クルーズ船の乗客で初めて感染者を確認し、3 月の小学校・中学校・高校などを臨時休校にするなど対策をとったものの、感染が拡大。同年 4 月 7 日に緊急事態宣言を発出し社会活動にも制限をかけた。その後感染拡大と沈静化を繰り返し現在に至る。

2020 年 3~5 月の第 1 波、7~8 月の第 2 波、11 月~2021 年 1 月の第 3 波、4 月~5 月の第 4 波、7 月~9 月の第 5 波と感染者の増減を繰り返してきた。現時点において、2022 年 1 月~4 月の第 6 波が最後とされるが予断を許さない。感染者の増加に対応すべく、医療機関では専用の外来を設置した他、専用の入院病床を確保するなど対応に追われた。企業活動においても、時差出勤を始めリモートワークの推奨、個人事業主を中心に売上減を補填する補助金事業などが展開された。

人々の健康だけでなく、社会も様相を一変させた。会話する際には、マスクを必ず着用し、2 メートル程度距離を確保する「ソーシャルディスタンス」を確保したり、透明の衝立を間に立てたりなどコミュニケーションの在り方も、新型コロナウイルスが登場する前後で一変した。

一方、令和 2 年版情報通信白書[1]では、「コロナウイルスは熱に弱く、26~27 度のお湯を飲むと殺菌効果がある」、「新型コロナにビタミン D が効く」などの誤情報が流布された。また、医学的な根拠のない感染予防効果を標榜する健康食品などのインターネット広告が出現したと報告されている。2020 年春から約 2 年が経過した時点において、デ

マは落ち着いてきたものの、YouTube などにおいて「新型コロナウイルス」や「新型コロナワクチン」に関する言動が不自由な状況が続いている。ワクチンに関して消極的な発言を行うことを「反ワクチン」と呼ばれ、議論が交わされている状況も存在する。

一方、新型コロナウイルス感染症の感染拡大を防ぐ効果として期待されているのが「新型コロナワクチン」である。2021 年 4 月より医療従事者・高齢者の順に接種が進み、2 回のワクチン接種が 2022 年 6 月時点において全人口に占める割合は 80.7%となっている。2022 年に入り、3 回目のワクチン接種が開始されているが、2022 年 6 月時点において 60.8%に留まっている。

本稿の調査対象時期にあたる 2022 年 3 月の直前に、ロシア連邦がウクライナに侵攻する事件が起きた。調査時期は侵攻直後に当たり、多くのツイートが行われた。

この論文は 2022 年 3 月から 4 月の期間に「コロナ」、「ワクチン」、「ウクライナ」に関するツイート収集を行い、ツイート数の変化・形態素解析による解析を行った。

ツイートの収集は、Twitter の API を利用して、収集するプログラムコードを自作し収集した。収集したテキストについて形態素解析を行った。形態素解析はオープンソースの形態素解析エンジンの中から MeCab を用いて品詞分解を行った。共起ネットワーク作成は KH CODER を使用し作成している。

2. Twitter について

Twitter は、先に述べた通り投稿以外にも様々な機能を持っている。本稿で登場する機能について説明する。

2.1 ツイート

Twitter では、テキストは最大で 280 文字 (すべて日本語の全角文字にした場合は 140 文字)、画像は最大で 4 件、GIF 画像と動画は最大で 1 件含めることができる。動画は最大 2 分 20 秒まで投稿することができる。Twitter には、自身が発信するツイートを公開・非公開にする機能を持っている。基本的には自身が発信するツイートは、Web サイトを閲覧できるすべての人に公開されることになる。非公開のツイートに関しては、フォロワーのみに公開されることになる。ツイートには URL を記載することも可能であり、ツイートから自社サイトへの誘導を行うことも可能である。

2.2 ツイート

リツイートとは、retweet (re: 返信の意味) から名付けられた Twitter の機能の一つ。他のユーザーのツイートを再びツイートすること。リツイート機能を使うと、そのツイートをフォロワー全員と素早く共有することができる。

[†] 武庫川女子大学共通教育部 Mukogawa Women's UNIV.

表 1 Twitter の機能一覧

| | |
|-------|--|
| ツイート | Twitterに投稿されるテキスト・画像・動画などのこと。 |
| リツイート | 誰かのツイートをフォロワーと共有すること。リツイートすることで、気に入ったニュースや面白い発見などをフォロワーに共有したり、タイムラインに拡散させたりすることができる。 |
| リプライ | タイムライン上にあるツイートに対して、個別に返信する機能。 |
| メンション | タイムライン上にあるツイートに対して、複数のアカウントに対して返信する機能。 |

2.3 リツイート

リツイートとは、retweet (re: 返信の意味) から名付けられた Twitter の機能の一つ。他のユーザーのツイートを再びツイートすること。リツイート機能を使うと、そのツイートをフォロワー全員と素早く共有することができる。

元のツイートには何回リツイートされたのか記録されている。企業が広告宣伝活動を行う際に、企業のアカウントが発信したツイートを消費者にリツイートさせることで、消費者をフォローしているアカウントに対しても情報伝播させることができる。企業側でも自身のツイートを誰がリツイートしたかは把握できるので、リツイートしたアカウントの中から便宜を図ることでキャンペーン活動を行うことができる。また、消費者が自社製品に関するツイートをリツイートすることで、消費者と自社ブランドの結びつきを強化することも可能である。

2.4 リプライ

リプライとは、他者のツイートに「返信」する機能である。リツイートは他者のツイートをもう一度情報発信するのに対して、リプライは発信したアカウントと自分のアカウント、そして両者をフォローしている人にしか表示されない特徴を持っている。ツイートごとに誰がリプライできるのか設定を行うことも可能であり、すべてのツイートがリプライできるわけではない。共有される範囲が狭いこともあり、個人の私信としての役割も持っている。

2.5 メンション

メンションとは、ツイートの文章に「@○○○」とアカウント名を記載して“タグ付け”するツイートのことをメンションという。リプライのツイートは、元ツイートアカウントと自分のアカウント両方フォローしているアカウントのタイムラインにしか表示されないが、メンションのツイートは自分をフォローしているアカウントのタイムラインに表示される違いがある。また、リプライは 1 対 1 の

情報交換になるが、メンションは複数のアカウント宛にツイートすることが可能となっている。

3. 既往研究と研究目的

3.1 既往研究

Twitter に関する既往研究は多岐にわたる。震災発生時におけるツイートの様子[2]や、リツイートの内容を解析したもの[3]が存在する。また、テキストの形態素解析によりユーザーの感情分析を試みるもの[4]などがある。その他では、リツイートを扱った研究もあり、先に挙げた震災時に起きたリツイート行動を解析したものやテレビ番組に対するリツイートを調べたもの[5]や、スポーツ大会に関するリツイートを調べた研究[6]などが存在する。

新型コロナウイルスのツイートを扱った事例では、情報拡散を調べたもの[7]、感情の変化を読み取ったもの、テキストより不安やストレスを読み取ったものなどがある。

3.2 研究目的・仮説

コロナ禍において「コロナ」に関するツイートと「ワクチン」を含むツイートにおいて、形態素解析を用いて情報発信の内容を抽出した上で、「コロナ」と「ワクチン」「ウクライナ」で発信されている情報量・情報伝播の内容を比較することで、それぞれの情報においてユーザー同士の交流に差異が生まれていることを確認することを目的とする。

■仮説 1: 「コロナ」「ワクチン」はリツイートの割合が「ウクライナ」より高い。

■仮説 2: 「コロナ」・「ワクチン」において、デマにあたる情報のツイートが存在する。

上記 2 つの仮説が成り立つか調査・分析を行った。

4. 分析

「コロナ」「ワクチン」「ウクライナ」の言葉を含むツイートを 2022 年 4 月に発信されたツイートを収集した。収集したツイートの発信数・発信形態・形態素解析を行った内容を以下に示す。

4.1 収集ツイートの全体数について

新型コロナウイルス感染者数の調査期間中の推移を図 1 に示す。2022 年 1 月～2 月における感染者数増加がピーク



図 1 新型コロナウイルス感染者数 (全国)

表 2 ツイートの収集状況

| 期間 | コロナ | ワクチン | ウクライナ |
|----|-----------|---------|---------|
| 3月 | 1,080,660 | 482,861 | 885,472 |
| 4月 | 753,775 | 723,695 | 749,999 |

表 5 「コロナ」ツイートの Web リンク解析

| 期間 | Webリンク | Yahoo | サプリメント |
|----|---------|--------------|-----------------|
| 3月 | 264,736 | 6,346 (2.4%) | 66,273 (25.0%) |
| 4月 | 482,676 | 9,770 (2.0%) | 271,394 (56.2%) |

表 3 ツイートの収集状況

| 3月 | 全体 | Webリンク | リツイート | メンション | リプライ |
|-------|-----------|-----------------|---------------|-----------------|-----------------|
| コロナ | 1,080,660 | 264,736 (24.5%) | 285 (0.03%) | 129,658 (12.0%) | 122,526 (11.3%) |
| ワクチン | 482,861 | 153,004 (31.7%) | 212 (0.04%) | 110,406 (22.9%) | 105,845 (21.9%) |
| ウクライナ | 885,472 | 545,551 (61.6%) | 1,173 (0.13%) | 207,409 (23.4%) | 165,595 (18.7%) |

表 4 ツイートの収集状況

| 4月 | 全体 | Webリンク | リツイート | メンション | リプライ |
|-------|---------|-----------------|-------------|-----------------|-----------------|
| コロナ | 753,775 | 482,676 (64.0%) | 303 (0.04%) | 121,801 (16.2%) | 116,470 (15.5%) |
| ワクチン | 723,694 | 236,484 (32.6%) | 296 (0.04%) | 178,319 (24.6%) | 172,006 (23.8%) |
| ウクライナ | 749,999 | 467,070 (62.3%) | 960 (0.13%) | 194,389 (25.9%) | 159,505 (21.3%) |

Yahoo! ニュースのツイートは3月・4月とも2.0%前後であった。一方、サプリメントの購入サイトへつなげる Web リンクは3月・4月とも大量に計測された。4月においては、「コロナ」を含むツイートの半数を占め

であり、その後微減していることが分かる。

「コロナ」・「ワクチン」・「ウクライナ」を含むツイートを収集し集計したものを表 2 に示す。「コロナ」と「ウクライナ」は3月の方が多く、逆に「ワクチン」は4月のツイート数が多いことがわかった。増減に着目すると、「コロナ」のツイート数の減少幅が大きいことがわかる。「ワクチン」のツイートはコロナとは逆に4月の発信数が増加している。「コロナ」・「ワクチン」のツイートは日々の新規感染者数との相関は見られない。

4.2 情報発信の分類

収集したツイートを「ツイート」・「リツイート」・「リプライ」・「メンション」に分類した。結果を表 3・表 4 に示す。外部サイトの URL を情報に含む「Web リンク」に分類したツイートをみると、「ワクチン」は3割程度、「ウクライナ」は6割程度で推移したことを確認した。一方、「コロナ」は3月と4月で投稿量に差がある。

ツイートを再投稿する「リツイート」を見ると「コロナ」と「ワクチン」は共に0.03~0.04%程度で Twitter を介した情報拡散はあまり行われていない。対照的に「ウクライナ」を含むツイートでは0.13%と「コロナ」・「ワクチン」に比べて多いことが確認された。

他のアカウント名を含む「メンション」では「ウクライナ」が最も多く、次いで「ワクチン」、最も少ないのが「コロナ」であった。

4.3 Web リンクのツイートについて

「コロナ」を含むツイートにおいて、Web リンクがあるツイートの内容について解析を行った。表 5 に結果を示す。

表 6 収集ツイートの Web リンク解析

| | Webリンク | Yahoo! ニュース |
|---------|---------|---------------|
| コロナ3月 | 264,736 | 6,346 (2.4%) |
| コロナ4月 | 482,676 | 9,770 (2.0%) |
| ワクチン3月 | 153,044 | 5,3 (3.5%) |
| ワクチン4月 | 236,184 | 7,789 (3.3%) |
| ウクライナ3月 | 545,551 | 40,281 (7.4%) |
| ウクライナ4月 | 467,070 | 30,463 (6.5%) |

ていたことがわかった。Yahoo! ニュースに着目すると「コロナ」・「ワクチン」よりも「ウクライナ」に関する Yahoo! ニュースに関するツイートがツイートに占める割合が高いことがわかった (表 6)。

4.4 リツイートについて

新型コロナウイルス感染症や新型コロナワクチンは、自身の健康や生活に大きくかかわる情報であるため、他のツイートよりも多くツイートしていると本稿の仮説 1 として設定した。しかし、実際のリツイートは0.04%程度であり、ウクライナと比べて少ないことが確認された。ウクライナに関するリツイートを見ると、戦場の様子を伝える報道に対する印象や意見を述べるツイートが確認された。

4.5 リプライについて

リプライについて内容を見ると、自分が得た情報をシェアする事例は確認されず、自身の体調や日常生活について意見交換する様子が確認された。一例を示すと『そうな

んだ。休みが合っ
て良かったね 🍀
ワクチンは3回目
かな?なんか、副
反応結構出てる人
多いから、せっか
くの休みが…副反
応出ないことを祈
ってるね 😊』。
本稿の仮説 2 に設
定した「コロナ」
や「ワクチン」を
含むツイートには
デマにあたる情報
が存在すると設定
したが、ユーザー
同士の情報交換や
リツイートにおいて
デマにあたる情報
が拡散されている
事例は確認され
なかった。

収集したツイ
ートについて Twitter

の API にそのツイートが何件リツイートされたか把握できる変数が用意されている。リツイート件数を元に情報の拡散具合を確認したが、現時点においてデマにあたる情報が拡散された様子は確認されていない。

5. 形態素解析

5.1 全ツイートの形態素解析について

Twitter から収集されたツイートに対して、MeCab を用い形態素解析を行った。上位 15 位の品詞について表 7 に示す。それぞれのツイートを構成する要素は、調査期間の2か月間大きな変化がなかったことが分かる。コロナにおいて「マスク」の投稿が増えたことなどが分かる。ウクライナのツイートにおいて「避難」という言葉が増加していることも確認された。

6. 考察

6.1 仮説 1 の検証

仮説 1: 「コロナ」「ワクチン」はリツイートの割合が「ウクライナ」より高い、という仮説を立てたが収集した結果リツイート率は 0.04%程度であり仮説は棄却された。一方「ウクライナ」に関するツイートは日本のニュースや海外のニュースを紹介する Web リンクが多いことがわかった。

6.2 仮説 2 の検証

仮説 2: 「コロナ」・「ワクチン」において、デマにあたる情報のツイートが存在する、について検証を行った。ニュース記事のリンクを示したうえで、このニュースは誤り

表 7 収集ツイートの形態素解析

| | コロナ3月 | コロナ4月 | ワクチン3月 | ワクチン4月 | ウクライナ3月 | ウクライナ4月 |
|-----|------------|------------|-----------|------------|------------|------------|
| 1位 | 感染74,628 | 思う47,755 | 接種199,452 | 接種271,020 | ロシア474,846 | ロシア372,095 |
| 2位 | ワクチン52,001 | ウイルス44,088 | 打つ112,259 | 打つ173,139 | 日本130,932 | 日本142,342 |
| 3位 | 思う45,494 | 行く35,111 | コロナ86,803 | コロナ131,496 | 戦争119,348 | 戦争107,370 |
| 4位 | ウイルス33,752 | 接種32,427 | 反応71,795 | 反応98,720 | 侵攻114,462 | 支援84,186 |
| 5位 | 接種28,353 | 言う32,372 | 今日46,718 | 今日62,714 | プーチン94,382 | 侵攻73,261 |
| 6位 | 言う27,622 | マスク27,313 | 出る42,169 | 思う61,972 | 思う86,187 | 思う71,348 |
| 7位 | 今23,561 | 出る26,315 | 思う40,985 | 出る58,742 | ニュース86,049 | 言う70,419 |
| 8位 | 今日23,051 | 今24,392 | 熱34,772 | 言う50,386 | 言う83,813 | ニュース66,386 |
| 9位 | 出る22,705 | 陽性23,205 | 痛い31,822 | 感染48,709 | 大統領70,031 | プーチン54,867 |
| 10位 | 見る18,158 | 今日22,812 | 言う31,767 | 熱48,002 | 支援59,456 | 見る44,110 |
| 11位 | 陽性17,808 | ニュース19,746 | 明日27,861 | 痛い45,025 | 見る52,653 | 避難42,651 |
| 12位 | ニュース16,736 | 日本19,619 | 腕25,571 | 明日40,896 | 世界51,729 | 政府42,109 |
| 13位 | マスク15,401 | 検査19,246 | 感染25,163 | 行く36,818 | 攻撃50,460 | 世界37,612 |
| 14位 | 日本15,261 | 見る18,771 | 行く24,405 | 腕33,746 | 今42,352 | 大統領37,612 |
| 15位 | 仕事14,669 | 仕事15,934 | 今22,421 | 今32,373 | 国民41,947 | 国36,871 |

であるという意見表明を行うツイートは確認できたものの、リツイートそのものが少なかった。特定のツイートを拡散させる活動は見受けられなかった。仮説 2 も棄却となった。

7. まとめ

今回 Twitter は 140 文字までの投稿となるため単独のツイートでデマ情報を発信したり、特定の情報を拡散したりする事象は見られなかった。今後情報伝播の様子を詳しく調べるために、

- ・テキストの感情分析を行い精査する
- ・リツイートの追跡することで情報拡散を可視化する

上記の2点の取組みを継続して行い、SNS による情報伝播の傾向からマーケティング・注意喚起の一助となる示唆を見出したいと考えている。

参考文献

- [1] 総務省, “情報通信白書令和 3 年版”,
 - [2] 梅島彩奈・宮部真衣・荒牧英治・瀧本明代, 「災害時 Twitter におけるデマとデマ訂正 RT の傾向」, 『2011 年度情報処理学会研究報告』, DBS-152, 1-6. (2011 年)
 - [3] 鳥海不二夫・篠田孝祐・栗原聡・榎剛史・風間和宏・野田五十樹, 「東日本大震災時におけるリツイートの分析」, 『2012 年度電子情報通信学会技術研究報告』, AI, 人工知能と知識処理, 19-24. (2012 年)
 - [4] 富田愛優, 玉岡雅之, Twitter データを用いた消費税増税に対する人々の感情の分析, 国民経済雑誌, 第 222 巻 5 号 (2020).
 - [5] 阿部倫之・米谷優也, 「テレビ番組の視聴リツイートを利用した中心ユーザの発見支援」 『2016 年度情報処理学会研究報告』情報システムと社会環境, (2016 年)
 - [6] 小林竜也・尾崎知伸, 「スポーツ大会におけるコンテンツに着目したリツイート行動の要因分析」 『2014 年度人工知能学会全国大会』, (2014 年)
- 鳥海不二夫 “COVID-19 下の情報拡散”, 第 11 回横幹連合コンファレンス, (2020).