

Twitterにおける「コロナ」「ワクチン」を含むツイートにみられるWebサイト誘導について

武庫川女子大学 共通教育部 助手

大谷 光一

要約

Twitterにおいて2022年3月から4月にかけて、「コロナ」と「ワクチン」を含むツイートを収集し、ツイートから外部Webサイトへの誘導が行われているのかを分析した。関心が高いワードに対して、自社製品のWebサイトへの誘導が積極的に行われていると仮定し、ツイートの内容やリンク先を把握した。外部サイトへの誘導を図るツイートのリツイートが多いのではないかと仮説を立てたが、棄却された。「コロナ」を含むツイートで免疫力を向上させるサプリに関するツイートが一定数の割合が見られた一方で、「ワクチン」を含むツイートでは自社製品への誘導を図るツイートはごく少数であった。「いいね」の反応も確認したが自社サイトへの誘導のツイートの「いいね」は低い値にとどまった。

キーワード

Twitter, 形態素解析, 情報伝播, 新型コロナウイルス感染症, ワクチン

I. はじめに

新型コロナウイルス感染症が流行し始めて2年以上が経過した。マスクの着用やワクチン接種や行動制限が変わるなど、生活様式が変わった。経緯を振り返ると、2019年12月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 disease19（COVID-19）と命名し2020年3月11日にパンデミック宣言した。日本では2020年2月大型クルーズ船の乗客で初めて感染を確認し、3月の小学校・中学校・高等学校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日に非常事態宣言を発出し社会活動に制限をかけた。その後、感染拡大と沈静化を繰り返し現在に至る。2021年5月以降順次新型コロナワクチン接種が始まり、重症化する感染者は減少したものの、感染者数を激減させるまでには至っていない。2022年3回目のワクチン接種（ブースター接種）が始まる一方、ウイルスに罹患しても重症化する事例が少ない若年層（12歳から18歳）にも接種の対象が拡大されてきた。コロナワクチンを接種するべきか否か、コロナワ

クチン接種による副作用について、テレビ・新聞をはじめとするマスメディアは大きく取り扱わず、TwitterやYouTubeなどソーシャルメディアネットワークサービス（SNS）で意見交換されている。本稿では、2022年3月から4月にかけて「コロナ」と「ワクチン」という、関心度が高いと思われる言葉を含むツイートを投稿することで、自社のECサイトやWebサイトへの誘導する広報活動の一端を担えるのか分析および考察を行った。

II. 既往研究と仮説

I. 既往研究

Twitterにおけ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^[1]やリツイートの内容を解析したものの^[2]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^[3]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りした研究^[4]がある。新型コロナウイルスを扱った事

例では、情報拡散を調べたもの^[5]、感情の変化を読み取ったもの、不安やストレスを読み取ったものなどがある。

2. 仮説

2022年3月から4月にかけて「コロナ」と「ワクチン」を含むツイートを収集し、自社のECサイトやWebサイトへの誘導が見られるのか調査を行った。ツイート内容について形態素解析を行い、どのような意見交換が行われ、情報伝播がされているのか分析を行った。それぞれのツイートにおいて、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Webリンクへの誘導が行われているのか検証を行った。本稿では、以下の仮説を立て検証している。

- 仮説1: 自社のECサイトやWebサイトへの誘導を行うツイートが「コロナ」や「ワクチン」を含むツイート共に存在する。
- 仮説2: 自社サイトへの誘導を行うツイートに対する反応(リツイートやいいねなど)が一定数存在する。

III. 分析手法

I. ツイート収集及び形態素解析

ツイートの収集はTwitter社が提供しているAPIを利用した。Pythonでツイート収集し、MongoDBにドキュメントとして保存するコードを自作した。MongoDBに格納したテキストをMeCabで形態素に分割し、分析を行った。使用した辞書はmecab-ipadic-neologdである。

IV. 分析

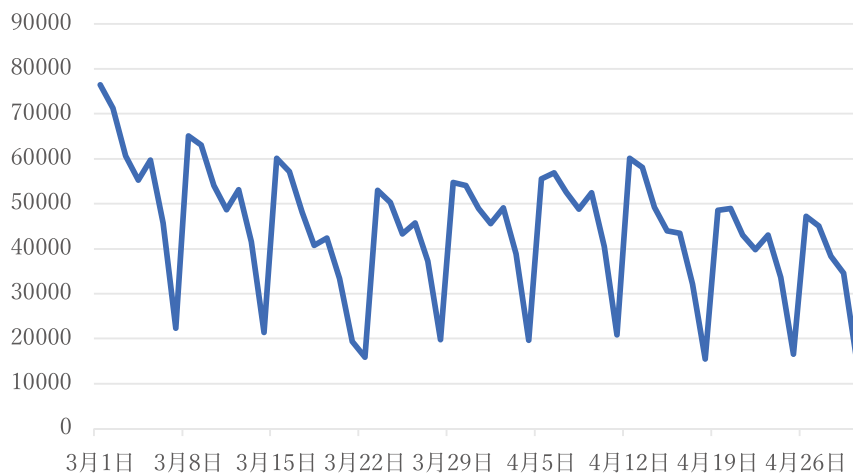
I. 収集ツイートの全体数について

2022年3月から4月の期間でツイートされた数量を表1に示す。合わせて、日本全国の新型コロナウイルス感染者数の推移を図1に示す。新型コロナウイルス感染症の第6波では2022年1月から2月がピークを迎えた。その後感染者数は図1が示す通り微減で推移した。「コロナ」は3月のツイートが多く、「ワクチン」は4月のツイート数が多い。増減に着目すると「コロナ」のツイート数が大きく減少している。

表—I ツイートの収集状況

期間	コロナ	ウイルス
3月	1,080,660	482,861
4月	753,775	723,695

図—I 新型コロナウイルス感染症の感染者数推移 (全国)



2. ツイートの分類について

収集したツイートを「ツイート（投稿）」「リツイート（再投稿）」「メンション（ユーザーのタグづけ）」「リプライ（返信）」「Webリンク」に分類を行った（表2・表3）。Webリンクとは、「【新型コロナ速報】小学校でクラスター発生 岡山県で新たに343人感染 岡山市で164人【岡山】(OHK岡山放送) #Yahoo ニュース <https://t.co/BEZSjhNSrO>」などのツイートで、テキストの中に「<https://>」とWebリンクが記載されているものを指す。実際には外部Webサイトの他に画像や動画を表示させるためのリンクも存在する。

分類ごとに見ると、情報拡散を示すリツイートは「コロナ」・「ワクチン」とも0.04%前後で少ないことがわかる。テキストの中に「@」を含むメンション¹⁾を見ると、「ワクチン」に比べて、4月における「コロナ」の増加割合が大きい。リプライに関しても、「コロナ」の増加割合が大きい。4月の「コロナ」のツイートにおいてWebリンクが64%を占めている。「コロナ」を含むツイート全体は減少しているにも関わらず、Webリンクに関するツイートは逆に増加していることがわかる。

3. ツイートの形態素解析

収集したツイートのテキストについてMeCabを用いて形態素解析²⁾を行った。上位15位の品詞について表4に示す。それぞれのツイートを形成する要素は、調査期間の2ヶ月間で大きな変化がなかったことがわかる。「コロナ」に関するツイートを見ると、同じ言葉が3月に比べて4月のツイートで増加している。ツイート数は減少しているが「マスク」の順位が12位から6位となっている。「ワクチン」のツイートで比較すると、ツイート数が増加しているが、構成している要素は大きな変化を見せていないことがわかる。

4. Webリンクのツイート解析

「コロナ」・「ワクチン」を含むツイートにおいて、Webリンクを含むツイートを表2と表3に示した。Webリンクを含むツイートを見ると、ビタミン・ミネラル・サプリメントなどを販売する企業の商品購入の際のクーポンコードを記載するツイートが多数を占めることがわかった。また、「ワクチン」のツイートを観察すると、ワクチン接種による副反応を記述したツイートが見られる。内容として発熱・腕の痛みや、痛みの程度をつぶやくものが多い。また、生理との重複により副反応が強くなるツイートが見られた。ワクチン接種により生理周期が乱れたり、生理時期ではないのに出血が起きたりするなど自分の身体に現れた異変をつぶやいている。コロナワクチン接種による「不妊」の可能性を心配するツイートも確認された。Webリンクにおける、「不妊」「生理」「免疫」「企業A³⁾」を含むツイート集計を表5に示す。「コロナ」を含むツイートにおいて、企業Aを含むツイートが75%を占めた。4月においても企業名を含む投稿が増加したが、全体の投稿量がそれ以上に増加しており、全体を占める割合は低下した。「企業A」と「ワクチン」を含むツイートはごく少数であるが存在することを確認した。

また、不妊・生理・免疫についても3月と4月のツイートを比較すると、投稿量が減少している。Webリンクを設けたツイートで他の事業体として観測されたのが、市区町村であった。日々のコロナの感染状況や、ワクチンの接種予約の案内など積極的に情報発信を行っていた。企業で「コロナ」や「ワクチン」を含むツイートを行い自社のWebサイトへ誘導する事例はごく少数であったため、本稿の分析からは割愛した。

表—2 ツイートの収集状況（3月）

3月	ツイート (全体)	リツイート	メンション	リプライ	Webリンク
コロナ	1,080,660	285 (0.03%)	129,658 (12.0%)	122,526 (11.3%)	264,736 (24.5%)
ワクチン	482,861	212 (0.04%)	110,406 (22.9%)	105,845 (21.9%)	153,004 (31.7%)

表—3 ツイートの収集状況 (4月)

4月	ツイート (全体)	リツイート	メンション	リプライ	Webリンク
コロナ	753,775	303 (0.04%)	121,801 (16.2%)	116,470 (15.5%)	482,676 (64.0%)
ワクチン	723,694	296 (0.04%)	178,319 (24.6%)	172,319 (24.6%)	236,484 (32.6%)

表—4 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
1位	感染 147,329	思う 47,755	接種 199,452	接種 271,020
2位	思う 93,026	ウイルス 44,088	打つ 112,259	打つ 173,139
3位	行く 75,432	行く 35,111	コロナ 86,803	コロナ 131,496
4位	ウイルス 68,371	接種 32,427	反応 71,795	反応 98,720
5位	言う 57,455	言う 32,372	今日 46,718	今日 62,714
6位	接種 50,631	マスク 27,313	出る 42,169	思う 61,972
7位	今 48,445	出る 26,315	思う 40,985	出る 58,742
8位	今日 45,030	今 24,392	熱 34,772	言う 50,386
9位	出る 43,196	陽性 23,205	痛い 31,822	感染 48,709
10位	見る 38,988	今日 22,812	言う 31,767	熱 48,002
11位	ニュース 34,674	ニュース 19,746	明日 27,861	痛い 45,025
12位	マスク 34,595	日本 19,619	腕 25,571	明日 40,896
13位	日本 33,804	検査 19,246	感染 25,163	行く 36,818
14位	陽性 29,963	見る 18,771	行く 24,405	腕 33,746
15位	仕事 29,454	仕事 15,934	今 22,421	今 32,373

表—5 Webリンクを含むツイートの形態素集計

	コロナ3月	コロナ4月	ワクチン3月	ワクチン4月
不妊	144 (0.05%)	100 (0.02%)	347 (0.23%)	281 (0.12%)
生理	287 (0.11%)	80 (0.02%)	1,116 (0.73%)	277 (0.12%)
免疫	56,720 (21.4%)	55,004 (11.4%)	6,060 (4.0%)	6,205 (2.6%)
企業A	199,344 (75.3%)	268,633 (55.7%)	58 (0.04%)	82 (0.03%)
Yahoo	12,377 (4.7%)	9,770 (2.0%)	5,316 (3.5%)	7,789 (3.3%)
ニュース	16,239 (6.1%)	7,114 (1.5%)	2,710 (1.8%)	4,237 (1.8%)
NHK	8,099 (3.1%)	3,932 (0.8%)	1,205 (0.8%)	3,007 (1.3%)
新聞	15,929 (6.0%)	8,681 (1.8%)	2,477 (1.6%)	4,359 (1.8%)

VI. 最後に

今回 Webリンクを含むツイートを分析する中で、企業 A のように「企業名」と「コロナ」の言葉を含み、キャンペーンコードを記載する事例は確認されたが、企業が投稿したツイートをリツイートするように誘導するような事例は確認されなかった。Webリンクを含むツイートにはYahoo!ニュースやNHKのWebサイト、新聞社のWebサイトのリンクをテキストに含ませる報道系のツイートが一定数存在した。Webリンクのあるツイートで「いいね」が多い投稿を確認すると、芸能人が新型コロナウイルス感染症から回復した投稿などが見られた。

「コロナ」や「ワクチン」においても、リツイートやいいねの反応が多いツイートは存在する。「台湾と日本は良き友人として、常に相手のことを心から思いやってきました。コロナ禍でも互いを思う気持ちに変わりはありません。近い将来、必ずまた会いましょう! <https://t.co/RrrcxVXqdL>」蔡英文 (President of the Republic of China (Taiwan)) のツイートは3月11日に投稿され、現在3,139件のリツイートと16,000件を超える「いいね」が送られている。

調査対象期間においては、18歳未満の子供にコロナワクチンを接種させるか、3回目のブースター接種を行うか意見を述べるツイートが見られた。医学的見地からではなく、自身の体験からコロナワクチン接種により、生理周期が乱れたり、生理痛が重くなったりという投稿が少数ではあるが見られた。また、「ワクチン」のツイートにおいては、ワクチン接種反対の意見を表明するツイートも存在した。今後も継続してどのような情報交換や情報伝播が起こっているのか、解析を継続したいと考えている。

注

- 1) 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。
- 2) 形態素解析では特定の企業が行なったツイートの影響を排除するため、テキスト内で使われた形態素を集計から排除した。同様にWebリンクが多いため、「https://」などについて集計から排除した。

- 3) 企業 A は iHerb を指す。アメリカに本社があり世界 185 ヶ国以上で利用されるサプリメントを中心とした EC ストア。コードを入力することで割引となるキャンペーンを行なっている。

引用文献

- [1] 梅島彩奈・宮部真衣・荒牧英治・瀧本明代 (2011年) 「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.
- [2] 鳥海不二夫・篠田孝祐・栗原聡・榊剛史・風間和宏・野田五十樹 (2012年), 「東日本大震災時におけるツイートの分析」, 『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.
- [3] 富田愛優, 玉岡雅之, (2020年) 「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.
- [4] 阿部倫之・米谷優也 (2016年) 「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, IS-136, 1-6.
- [5] 小林竜也・尾崎知伸 (2014年) 「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』, 1-2.