

コロナ禍における Twitter を介した 情報伝播について

武庫川女子大学 共通教育部 助手

大谷 光一

要約

コロナ禍において Twitter でどのような情報伝播がされているのか、2020年12月から翌年5月末までのツイートを収集し解析を行った。投稿を拡散するリツイートは0.1%以下であり、webサイトのリンクを含むツイートが4割を占めた。アカウント同士のやりとりであるリプライは、常時2割を超える割合であった。ツイートの文面を形態素解析したところ、リツイート・リプライでは使用される形態素に変化が見られた。情報交換に当たるリプライでは、「落ち着いたら、会おう」などのやりとりが確認された。1月以降「ワクチン」に関する情報が確認され、4月には、「オリンピック」に関するツイートが確認された。ユーザは Twitter を利用しながら、ニュースなどの外部サイトに関する情報をつぶやいたり、「おはよう」や「ありがとう」など日常の一コマをつぶやいたりしていたことがわかった。

キーワード

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

I. はじめに

近年インターネットの発展に伴い、多くの情報が生み出されている。中でもソーシャルメディアサービス（以下 SNS とする）により生み出される情報は、著しく増加している。SNS の一つである 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。Twitter には情報を共有する機能（リツイート）が存在する。自身が気になる情報をリツイートすることで、自分をフォローしているアカウントに対して情報共有ができる。また、ハッシュタグを用いることで、その時間帯に投稿されたツイートを検索することもできる。また投稿されたツイートに返信を行ったり、複数のアカウント向けに投稿したりすることもできる。情報を拡散したり収集したりできる機能を Twitter は持っている。

コロナ禍において、日々の感染者数の情報や緊急事態宣言に伴う生活の変化等、さまざまな情報発信がされている。日々の情報取得についても、SNS から得る行動が増加している昨今、緊急事態宣言下での通勤電車の混み具合や居酒屋などの飲食店での営業状態など何気ない日常を切り取ったものや、感染対策や周りの感染状況など様々な情報

が存在する。ウイルスの特徴や感染対策など、新しい情報を得ようと SNS で「コロナ」に関する検索や情報発信が行われている。

II. 既往研究と仮説

(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や、テキストの形態素解析を行い、ユーザの特性や分類を試みるもの^{[1] [2] [3]}などがある。リツイートを扱った研究では、震災時に起きたリツイート行動を解析したもの^{[4] [5]}が多い。しかし、平常時のリツイートを扱った研究は少ない。ニュースを引用したリツイートに関するニュースバリューを考察するものやテレビ番組の放送中におけるリツイートやスポーツ大会の増減を調査したもの^{[6] [7]}がある。

(2) 調査手法

本稿において、Twitter を対象として、「コロナ」に関する情報がどのように拡散されているのか Twitter API を用い分析を行った。コロナに関するツイートを2020年12月から

2021年5月まで収集し、各月ごとに解析を行い、情報拡散に差異が見られるか調べた。各月ごとのツイート・リツイート・リプライの解析も行い、「コロナ」に関する投稿が時系列毎にどのような変化が見られるのかを考察する。また、投稿内容を解析するため形態素解析も行い、どういう文脈で情報が伝播しているのか解析を行った。

III. Twitterの機能について

Twitterは先に述べたとおり、投稿以外にも様々な機能がある。本稿で登場する機能について以下に説明する(表—1)。

IV. 分析手法

(1) ツイート収集

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。

(2) 自然言語処理

得られたツイートをテキストファイルに変換し、形態素解析を行った。形態素解析はMeCabを用いて、入手したテキストデータを意味ある最小限の単位に分解した。名詞・形容詞・動詞をテキストから抽出し処理を行った。共起ネットワーク図の作成については、KHCODERを使用した。

V. 仮説

新型コロナウイルスに関して収集した全ツイートをリツイート、リプライなどに分類を行ったうえで、時期ごとの投稿数や

投稿内容について以下の仮説を立てた。

①感染者数の増減によりツイート数が増減する②通常のツイートとリプライでは投稿傾向が異なる③調査期間において、コロナに関するツイート内容が変化する。

VI. 分析

(1) ツイート集計について

「コロナ」に関するツイートを2020年12月から2021年5月末まで収集を行った。2回目の緊急事態宣言が発出される2021年1月8日から3月21日までと、3回目の緊急事態宣言が4月25日から実施されており、2回の緊急事態宣言を含む期間となっている。2020年12月から2021年5月末までの新型コロナウイルス感染症の陽性者発表の数値を図-1に示す。

12月から5月における日本国内・東京都・大阪府の新型コロナウイルス感染症の感染者数と収集したツイート数を表-2に示す。感染者数とツイート数の相関は見られないことが分かる。感染者数が減少した2月は1月に比べて、ツイート数も同じく減少するものの、感染者数の減少程ツイート数は減少をみせていない。3月後半において新規感染者数が増加に転じたが、同様にツイート数も増加したことが分かる。

続いて、リツイート・リプライ・webリンク・ハッシュタグ・Yahoo!ニュースに関するツイートの推移を表-3に表す。リツイートとハッシュタグは調査期間において、全体のツイートに対する占める割合の変化が見られなかった。リツイートの変化を見ると2月が全体に対して最も割合が高く0.097%、5月は0.065%であった。リプライは2月のツイートの32.6%を占めていた。期間中リプライの割合が最も低かったのは5月

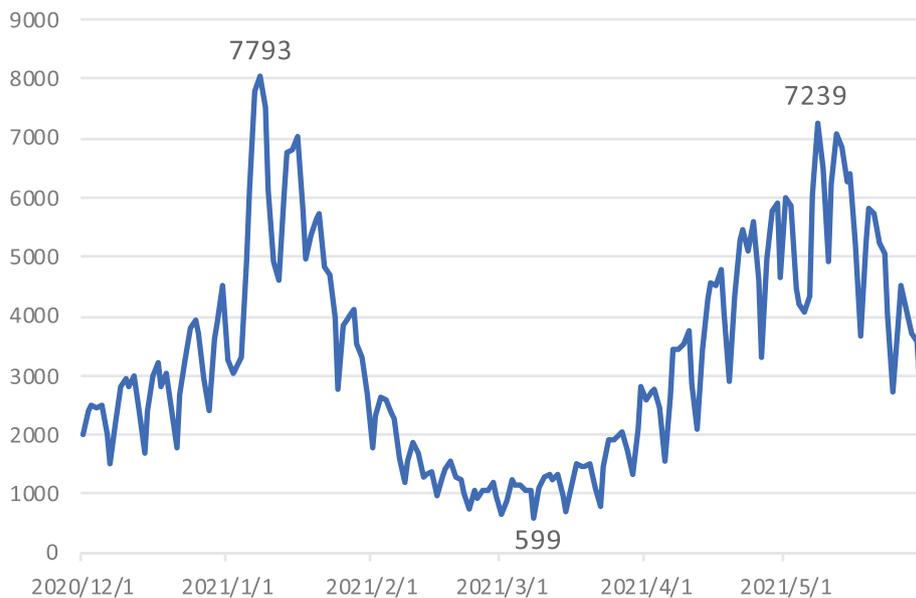
表—1 Twitter機能一覧

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ーに共有したり、タイムラインに拡散したりできる
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

で23.2%だった。調査期間を通して全体のツイートの2割以上がアカウント同士の情報交換であったことがわかった。2月はリプライ・ハッシュタグ・Yahoo!ニュースのツイートが占める割合が、他の月と比べ高い事が解析により分かった。

ツイートの中には、webサイトの外部リンクを含むものも存在する。調査期間において12月から3月までは全体の約44%を占めていたが、4月は41.6%、5月は28.8%まで減少している。

図一| 日本国内における新型コロナウイルス感染症感染者数推移（厚生労働省オープンデータより）



表一| 新型コロナウイルス感染症感染者数と収集したツイート数

	日本国内	東京都	大阪府	収集ツイート数
2020年12月	86,541	19,369	9,766	722,181
2021年1月	154,700	40,367	13,721	882,652
2021年2月	41,838	10,097	3,365	345,897
2021年3月	42,300	9,310	5,080	414,856
2021年4月	117,498	18,090	28,978	566,734
2021年5月	153,683	21,879	18,527	574,237

表二| 種類別ツイート集計

	ツイート数	リツイート	リプライ	webリンク	ハッシュタグ	Yahoo ニュース
2020年12月	722,181	676	190,057	320,561	100,060	6,179
2021年1月	882,652	654	229,406	383,024	126,399	21,711
2021年2月	345,897	336	112,617	151,864	54,134	9,070
2021年3月	414,856	323	114,752	184,443	60,807	8,944
2021年4月	566,734	404	149,011	235,811	73,309	9,838
2021年5月	574,237	374	133,335	165,444	75,511	13,686

(2) ツイート全体の形態素解析について

6ヶ月間のツイートについて、形態素解析を行った。「感染」「新型」「思う」「人」「ウイルス」などが上位を占めている。緊急事態宣言が発出された1月では「出る」という言葉が41,058件確認された他、1月以降は「ワクチン」という言葉が上位15位の中に含まれるようになった。5月では「接種」が43,174件確認されるなど社会の変化がツイートにも反映されていることが分かる。「東京」は12月に32,664件、4月に26,924件確認された。いずれも緊急事態宣言が発出される前のタイミングと重なっている。「大阪」も4月に32,619件確認されており、3回目の緊急事態宣言発のタイミングと重なる。各月とも10位前後に「ニュース」が確認される。Yahoo!ニュースを含めてニュース記事の外部リンクを含むツイートが一定数存在している。5月における「ニュース」の形態素は24位(21,239件)と減少していることがわかる。

(3) リツイートの形態素解析について

調査期間におけるリツイートについても、同様に形態素解析を行った。調査期間において一部を除いて同じ要素で構成されていることが読み取れる(表-5)。上位には「思う」

「人」「感染」などが占めた。自分の考えを発信している様子がわかる。各月における形態素を見ると、12月には「テレワーク」という言葉が入り、5月には「五輪」「オリンピック」が計測された。1月の時点から「ワクチン」が登場していることも特徴と言える。リプライに比べて、形態素の変化が乏しいことが今回の解析で判明した。

(4) リプライにおける形態素解析について

調査期間におけるリプライ(他のアカウントが投稿しているツイートに対して返信すること、返信先は投稿した本人のみ)のツイートについても形態素解析を行った。

全体のツイートでは見られない「落ち着く」「ありがとう」「おはよう」という形態素が測定された。「落ち着いたら、会いたいね」というやり取りや「気をつけてね」という内容が見受けられる。「ワクチン」が全体と同様4月以降上位に位置していることがわかる。「頑張る」が12月から2月にかけて上位に位置した。3月以降も測定されるも、3月は17位(4,685)、4月は19位(5,864)、5月は21位(5,116)であった。またリツイートと同様に「思う」という言葉が上位に位置している。

表-4 「コロナ」全ツイートにおける形態素集計

	12月	1月	2月	3月	4月	5月
1位	感染 153,292	感染 174,514	新型 54,100	新型 69,375	感染 110,264	感染 97,061
2位	新型 123,043	新型 138,385	感染 53,879	感染 67,457	新型 85,795	新型 86,667
3位	思う 73,776	思う 97,059	思う 35,331	思う 40,415	思う 61,009	ワクチン 77,321
4位	人 66,445	人 70,223	人 30,761	ウイルス 37,506	人 53,525	思う 57,101
5位	ウイルス 61,366	ウイルス 70,223	ウイルス 29,529	行く 36,668	行く 52,191	人 52,156
6位	行く 56,633	行く 69,893	ワクチン 28,767	人 34,502	ウイルス 46,839	ウイルス 46,078
7位	対策 54,991	言う 67,793	行く 28,729	ワクチン 33,874	ワクチン 46,131	接種 43,174
8位	言う 51,026	対策 62,069	対策 23,473	対策 32,024	対策 45,717	新規 42,172
9位	今年 44,778	ニュース 52,503	言う 22,174	言う 24,484	言う 42,438	対策 41,767
10位	ニュース 42,884	今 49,943	ニュース 18,843	言う 24,475	今 32,807	行く 41,754
11位	今 35,111	ワクチン 45,998	今 18,790	ニュース 24,041	大阪 32,619	言う 40,915
12位	日本 33,728	日本 43,743	日本 17,091	今 24,423	日本 29,798	進む 39,827
13位	東京 32,664	出る 43,407	見る 14,794	質問 18,680	ニュース 28,752	使える 38,425
14位	見る 31,805	見る 41,058	マスク 14,503	日本 18,671	東京 26,924	日本 18,671
15位	出る 31,629	今日 36,807	前 13,814	見る 18,019	見る 26,602	見る 18,019

表—5 「コロナ」リツイートにおける形態素集計

	12月	1月	2月	3月	4月	5月
1位	思う 79	思う 128	思う 55	思う 42	思う 79	人 71
2位	感染 75	人 100	人 52	人 37	感染 75	感染 56
3位	人 71	感染 90	新型 37	新型 33	人 58	思う 56
4位	新型 59	言う 79	行く 36	ワクチン 31	新型 52	新型 47
5位	言う 52	新型 78	言う 32	行く 30	ワクチン 43	ワクチン 43
6位	行く 47	自分 51	見る 29	見る 26	行く 41	今 32
7位	ウィルス 39	今 50	ワクチン 27	ウィルス 25	今 34	見る 31
8位	今 39	見る 49	感染 27	今 24	言う 32	東京 31
9位	今年 39	対策 47	ウィルス 24	言う 22	見る 31	ウィルス 30
10位	見る 35	行く 46	今 18	接種 18	考える 28	対策 30
11位	テレワーク 32	ウィルス 42	自分 17	感染 17	対策 27	日本 30
12位	問題 31	前 41	出る 17	前 16	大阪 27	言う 29
13位	対策 30	日本 41	対策 17	マスク 15	出る 26	五輪 25
14位	前 28	ワクチン 35	日本 17	日本 14	ウィルス 25	前 25
15位	自分 25	怖い 31	医療 15	関係 13	接種 25	オリンピック 24

表—6 「コロナ」リプライにおける形態素集計

	12月	1月	2月	3月	4月	5月
1位	思う 21,958	思う 28,016	行く 13,693	行く 15,454	思う 18,653	思う 17,120
2位	行く 19,469	行く 23,667	思う 13,564	思う 13,771	行く 18,011	行く 14,209
3位	感染 16,105	感染 19,145	落ち着く 10,310	落ち着く 10,284	感染 12,586	感染 10,644
4位	気 15,150	落ち着く 18,527	ありがとう 7,839	笑 8,289	人 11,239	ワクチン 10,596
5位	落ち着く 13,884	人 16,620	笑 7,410	ありがとう 8,164	落ち着く 11,225	人 10,537
6位	ありがとう 13,160	ありがとう 15,384	人 7,203	人 7,141	言う 10,265	落ち着く 9,873
7位	人 12,962	気 15,283	感染 6,577	今 6,804	ありがとう 9,866	言う 9,816
8位	今年 12,570	言う 15,009	今 6,551	言う 6,647	今 9,305	ありがとう 9,483
9位	言う 11,837	今 13,968	早い 6,138	気 6,120	笑 9,209	今 8,305
10位	笑 10,695	笑 13,081	気 6,009	感染 6,005	気 8,922	笑 7,672
11位	おはよう 9,885	早い 12,193	会う 4,897	早い 5,907	対策 8,166	おはよう 7,278
12位	今 9,823	対策 10,762	ワクチン 4,877	対策 5,644	早い 7,603	早い 7,125
13位	対策 9,801	おはよう 10,377	見る 4,615	行ける 5,048	おはよう 7,109	対策 6,962
14位	早い 9,070	頑張る 9,889	対策 4,572	前 5,048	見る 6,727	気 6,747
15位	頑張る 8,510	今日 9,708	頑張る 4,441	見る 5,019	ワクチン 6,368	日本 6,352

VII. ワクチンの変遷について

ここまで収集したツイートの形態素について解析を進めてきた。社会ではウィルス供給が本格化し、コロナウィルスワクチンを接種するかどうかで、意見が分かれる状態となってい

る。ワクチンの開発から承認までのプロセスが、他のワクチンと異なること、承認されるまでの治験期間が短いことが背景に存在する。積極的に「ワクチン」を打たない理由の一つに、「ワクチンを接種することで不妊になる」という情報が

影響を与えている事例がある。そこで今回収集したツイートから、「ウイルス」「ワクチン」「不妊」に関する形態素の集計を行なった（表-7）。「ウイルス」に関しては、調査期間において、全ツイートに対して8～9%を推移していた。「ワクチン」と「不妊」に関しては2月と5月が全ツイートに対しての割合が高いことがわかった。

VIII. 共起ネットワーク図について

収集したツイートから共起ネットワークを作成した。解析した形態素個々の関係性について計算を行い図示したものが共起ネットワーク図である。出現パターンの特徴が似てい

る形態素を分類し、強い共起関係を線で結び、出現数が多いほど大きい線で描画される。

形態素が分類されることで、どのような文脈で語られているかを理解することができる。2020年12月の全ツイート（722,181件）と2021年5月の全ツイート（574,237件）の共起ネットワーク図を示す（図-2・図-3）。12月と5月を比較すると、「ウイルス」「感染」「新型」につながる語群を見ると12月ではNHKやYAHOOが含まれる。12月の時点において、「東京」はこの語群に属していたが、5月の時点では、「五輪」「オリンピック」の語群に移っている事が分かった。5月の共起ネットワーク図を見ると、「ワクチン」「接種」に関する語群が形成されていることがわかる。

表-7 「ウイルス」「ワクチン」「不妊」における形態素集計

	12月	1月	2月	3月	4月	5月
ウイルス	61,366	70,223	29,529	37,506	46,839	46,078
ワクチン	22,477	45,998	33,874	33,874	46,131	77,321
不妊	163	275	193	169	155	411

図-2 2020年12月のツイートに関する共起ネットワーク図

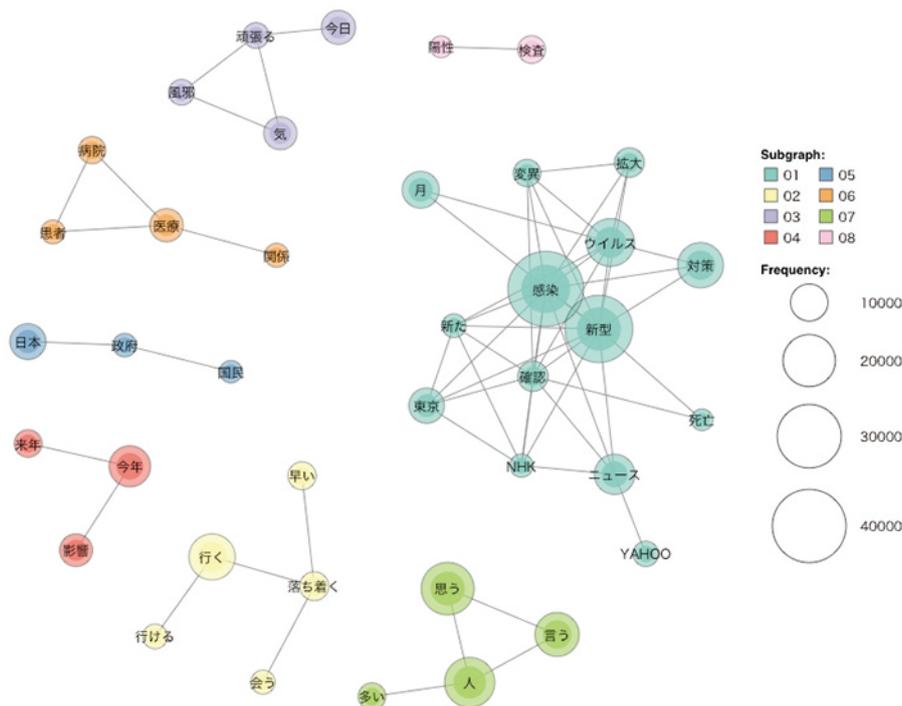
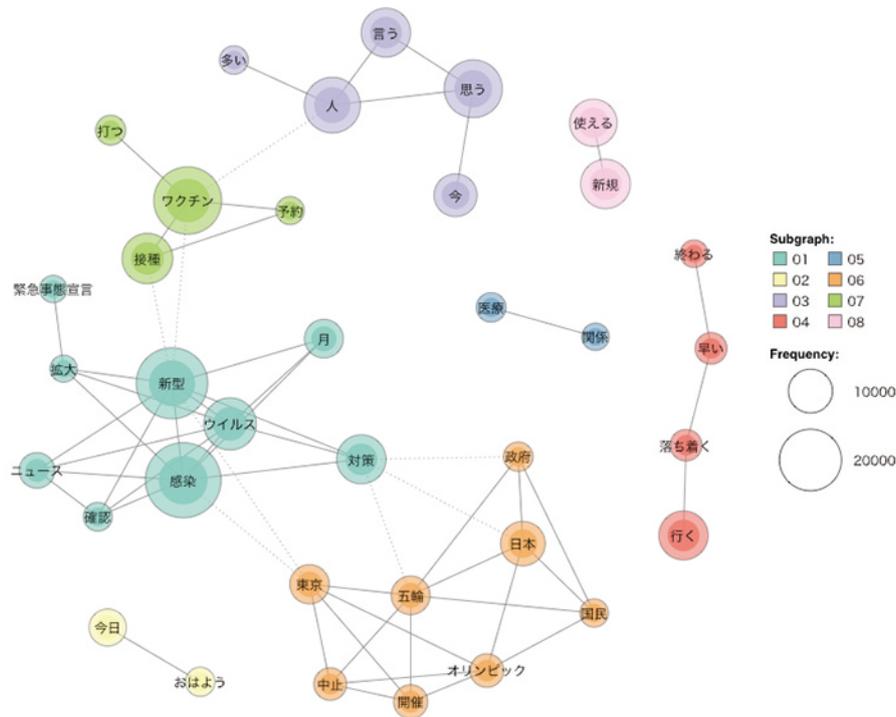


図-3 2021年5月のツイートに関する共起ネットワーク図



IX. 考察

「コロナ」を含むツイートの解析について仮説ごとに検証する。

(1) 感染者数の増減によりツイート数が増減する

新規感染者の増減により増加・減少に転じるものの、その変化の割合は近似値をとるものではなかった。相関関係が強いのであれば、新規感染者数が少なかった3月8日前後でツイート数は減少しなければならないが結果はそうではなかった。図-1が示すように新規感染者数が減少していた2月においてリツイート・リプライなどが活発に行われた。

(2) 通常のツイートとリプライでは投稿傾向が異なる

今回の解析において、全ツイートの形態素とリプライで使用された形態素が異なることが分かった。リプライとは、他のアカウントの投稿に対して返信する行為であり、アカウント同士の交流という側面が大きいことが分かった。アカウント同士での情報交換は少なく、「ありがとう」や「おはよう」、「落

ち着いたら、会おう」というつぶやきが行われていた。調査期間の間、その傾向は変わらなかった。

(3) 調査期間において、コロナに関するツイート内容が変化するか

調査期間全体において、「感染」「思う」「人」「ウイルス」を含むツイートが交わされていた。医療機関の状態を心配し、『コロナが終息したら、オフラインで会いたい、旅行やイベントに参加したい』と6ヶ月間変わることなく、ツイートされていたことが確認できた。1月から「ワクチン」という言葉が交わされ、2月に入ると「接種」できるかツイートに乗せて発信する様子が確認できた。4月に入ると「オリンピック」に関するツイートが散見され、「不妊」を含むツイートは2月と5月に多く観測されたことが分かった。4月以降webサイトのリンクが減少するなどの変化も確認された。

X. 最後に

「コロナ」に関するツイートを観察すると、他人の投稿をリツイート（情報拡散）する行為はあまり見られない一方、webサイトのリンクを含むツイートを行い、外部サイトの記事を発信することで情報の共有がなされたことがこの解析から見えてきた。アカウント同士の情報交換であるリプライでは、「おはよう」といった何気ない「つぶやき」をかわしながら日々生活している姿が浮かび上がってきた。4月以降webサイトのリンクのツイートが減少しているが、一部商品販売を含めたツイートが散見された。「コロナに打ち勝つために、健康食品を摂ろう」というような内容であった。当初マーケティングの意味合いで、消費者心理を突くツイートが一定程度存在するのではないかと想定していたが、4月以降の一部のツイートでしか確認できなかった。大量のツイートのテキスト精査を今後進める中で、見つかるかもしれないが、あったとしても少数に留まると考えられる。本稿では、収集したツイート全体からの俯瞰した内容であった。今後は、「不妊」やwebサイトリンクなど、焦点を絞った「コロナ」に関するツイートからどのような情報交換・情報交流が起きているのか、継続的に解析したいと考えている。

引用文献

- [1] 島海不二夫・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.
- [2] 田中淳史・田島敬史(2010年)「twitterのツイートに関する分類手法の提案」、『2010年度情報処理学会』, A5-4.
- [3] 岩橋瑠伊・矢吹 太郎(2018年)「SNSにおいてフェイクニュースを拡散するユーザーの特徴抽出」、『2018年度情報処理学会全国大会』, 113-114.
- [4] 山本雅人・小笠原寛弥・鈴木育男・古川正志(2012年)「東日本大震災における情報伝播ネットワーク」、『情報処理学会 情報処理』53巻11号, 1184-1191.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 工藤悠太・小林亜樹(2012年)「リツイート率の算出による価値あるツイートの発見手法」、『情報処理学会第74回全国大会』, 1-671-1-672