

## 新型コロナワクチンをめぐる人々の話題・関心の変化を分析 —— 1億超の大規模 Twitter データを読み解く ——

### 1. 発表者：

小林 亮太（東京大学大学院新領域創成科学研究科 複雑理工学専攻 准教授／  
東京大学数理・情報教育研究センター 准教授）

中山 悠理（東京大学大学院新領域創成科学研究科 複雑理工学専攻 修士課程 2 年生）

武富 有香（国立情報学研究所 情報学プリンシプル研究系 特任研究員）

須田 永遠（国立情報学研究所 情報学プリンシプル研究系 特任研究員）

### 2. 発表のポイント：

- ◆新型コロナウイルスワクチン接種期間中の 2021 年 1 月から 10 月までに Twitter で投稿された「ワクチン」を含む 1 億件以上の日本語の全ツイートデータを時系列的に分析しました。
- ◆2021 年 6 月から開始された職域接種を境に、人々の関心が社会的トピックから個人的事柄へと推移したことがわかりました。
- ◆本研究から、Twitter による個人的体験の共有がワクチン接種への安心感を醸成した可能性が示されました。今後、人々の気持ちや関心をおしはかる手段として、ソーシャルメディア分析が活用されることが期待されます。

### 3. 発表概要：

日本での新型コロナワクチンの接種は、ワクチンの安全性・有効性に対する不安やワクチン接種に関する政策への不満があったにもかかわらず、欧米諸国と比べて迅速に進みました。短期間で高い接種率に達する過程で、人々が何を考え、何に関心を抱いたかを知ることは公衆衛生上の重要な問題です。

今回、東京大学大学院新領域創成科学研究科の小林亮太准教授、中山悠理氏（修士課程 2 年生）、国立情報学研究所の武富有香研究員、須田永遠研究員らは、千葉商科大学 橋本隆子教授、国立情報学研究所の宇野毅明教授、喜連川優所長、東京大学 生産技術研究所 豊田正史教授、吉永直樹准教授、アントワープ大学（ベルギー） Luis EC Rocha 教授と共同研究を行い、2021 年 1 月から 10 月につぶやかれた新型コロナワクチンについての 1 億以上ものツイートを分析しました。これまで、1 つの国（日本）におけるワクチン接種期間の全ツイートデータを網羅的に分析した研究はありませんでした。

テキスト情報を考慮に入れた時系列分析を行った結果、2021 年 6 月の職域接種の開始を境に、ワクチン政策、ワクチンの有効性、関連ニュースなどの社会的トピックに関するツイートの割合が減り、接種の予定や報告、自身の副反応などの個人的事柄に関するツイートの割合が増えていることを発見しました。本研究から、Twitter による個人的な体験の共有がワクチン接種に対する安心感を醸成に寄与した可能性が示されました。

今後、公衆衛生などの政策決定を行う上で重要な人々の気持ちや興味・関心をおしはかる手段として、ソーシャルメディアデータ分析が活用されることが期待されます。

本研究成果は、2022 年 12 月 22 日（米国東部時間）に「Journal of Medical Internet Research」のオンライン版で公開されます。

本成果は、内閣府 Covid-19 AI・シミュレーションプロジェクトで実施し、科学技術振興機構（JST） 戦略的創造研究推進事業 個人型研究（さきがけ）研究領域：「数学と情報科学で解き明かす多様な対象の数理構造と活用」（研究総括：坂上貴之 京都大学大学院理学研究科 教授 JPMJPR1925）、科学研究費助成事業 基盤研究 A 19H01133、による支援を受けて行われました。

#### 4. 発表内容：

##### <研究の背景と経緯>

日本における新型コロナワクチンの接種開始は、欧米諸国と比べると2ヶ月以上も遅れました。しかしながら、国民の間にワクチンに対する多くの不安や不満があったにもかかわらず、日本におけるワクチン接種は世界の中でも速く進み、医療従事者への接種開始からわずか8ヶ月ほどでワクチン接種率は72%（世界229カ国中14位）に達しました。この接種期間中に、人々のワクチンについての興味・関心はどのように変化したのでしょうか。

ワクチンについての人々の考え方を調べることは、公衆衛生における重要な課題の1つであり、これまでは主にアンケートによって調査されてきました。しかし、同じ大規模なアンケートを繰り返し実施することには非常にコストがかかります。また、質問項目の設定によって結果が大きく変わりうることや、被験者が控えめに回答する傾向が多いことなど、アンケート調査には多くの問題点が指摘されています。

一方、スマートフォンの普及に伴い、ソーシャルメディアのデータを分析する研究も行われるようになってきました。日本ではTwitterの普及が進んでおり、20～30代のおよそ60%がTwitterユーザであると推計されています。本研究ではTwitterでつぶやかれた長期間にわたるツイートデータを時系列的に分析することで、ワクチンに関する人々の興味・関心の変化を調べました。

##### <研究の内容>

まず、本研究グループは2021年1～10月に投稿された「ワクチン」を含む全ての日本語ツイート（800万ユーザが投稿した1.1億ツイート）を収集しました。次に、機械学習技術（トピックモデル、注1）を用いてツイートを「ワクチン接種についての個人的感想」「ワクチンの有効性に関する議論」「接種会場からの実況」「政策に対する意見」「マスコミの報道に対する意見」などの15個のトピックに分類しました。機械学習を使うと膨大なツイートをトピックに自動的に分類できますが、トピックの意味を解釈することはできません。そこで、各トピックからランダムに抽出したツイート群について、データサイエンスの研究者（小林、中山）と文学の研究者（武富、須田）が協力して精査することによって15個のトピックの意味解釈を行い、4つの主要な話題（1. 個人的な事柄、2. ニュース、3. 政治、4. 陰謀論とユーモア）に整理しました。

次に、各テーマのツイート割合の時間変化を調べました（図1）。2021年1月当初は、個人的な事柄（約30%）、ニュース（約30%）、政治（約25%）、陰謀論とユーモア（約15%）というように4つのテーマにツイートは分散していました。しかし、6月以降には、ワクチン接種の予定、接種後の体調や副反応など個人的事柄に関するツイートが急増することがわかり、10月には個人的な事柄に関するツイートは「ワクチン」を含むツイート全体の約70%にまで増加しています。この結果は、6月以降にTwitterユーザの興味・関心が個人的事柄という話題に集中したことを示唆しています。また、陰謀論（注2）関連のキーワードが含まれるツイートの割合はわずか6%と小さく、その中には冗談をつぶやいているツイートも多かったため、

日本語 Twitter はワクチン関係のトピックに関しては陰謀論の温床になってないことが示唆されます。

最後に、6月以降に個人的事柄に関するツイートが急増した原因を調べるため、主要なイベント（A. 医療従事者の接種開始：2月、B. 高齢者の接種開始：4月、C. 職域接種開始：6月、D. 東京オリンピック：7月）がツイートされる話題に与える影響を分析しました。各話題のツイートの割合を時系列で計算し、中断時系列分析（注3）を用いてこれらのイベントが与える影響を分析した結果、C.職域接種、D.東京オリンピックの2つのイベントがツイートに与える影響が大きい（統計的に有意である）ことがわかりました。特に、職域接種が開始された2021年6月21日以降に個人的事柄に関するツイートの割合が急速に増加することが確認されました（図2）。この結果は、多くのTwitterユーザがワクチン接種を受けることが可能となった職域接種開始以降に、Twitterユーザの興味・関心が個人的事柄に集中したことを示唆しています。

#### <今後の展開>

本研究は、職域接種が始まるとともに、Twitter上の人々の関心がワクチン接種についての個人的な体験に集中し、接種後の体調や副反応に対する準備などワクチン接種を受ける上で重要な情報がTwitter上で共有されたことを示しています。その一方、Twitterなどソーシャルメディアの問題として、陰謀論や偽情報の拡散によるワクチン接種に与える負の影響が考えられます。今回の分析によると、日本のTwitterでは陰謀論のツイートはあまり見られませんでしたので、Twitterはワクチン接種を阻害する要因にはならなかったことが推測されます。以上のことから、Twitterが日本における迅速なワクチン接種率の向上の一因になったことが示唆されます。

今後は、本研究で行ったようなデータ解析技術を発展させることにより、公衆衛生などの政策決定を行う上で重要となる、人々の気持ちや興味・関心をおしはかる手段として、ソーシャルメディアデータ分析が活用されることが期待されます。

#### <研究プロジェクトについて>

本研究は、内閣府 Covid-19 AI・シミュレーションプロジェクトとして実施され、科学技術振興機構 戦略的創造研究推進事業 個人型研究（さきがけ）研究領域：「数学と情報科学で解き明かす多様な対象の数理構造と活用」（研究総括：坂上貴之 京都大学 大学院理学研究科 教授）、科学研究費助成事業 基盤研究A 19H01133、による支援を受けて行われました。また、本研究は、東京大学 小林亮太准教授、中山悠理氏、国立情報学研究所の武富有香研究員、須田永遠研究員に加えて、国立情報学研究所の宇野毅明教授、喜連川優所長、千葉商科大学の橋本隆子教授、東京大学の豊田正史教授、吉永直樹准教授、ゲント大学（ベルギー）Luis EC Rocha教授が加わった国際共同研究によって行われました。

#### 5. 発表雑誌：

雑誌名：「Journal of Medical Internet Research」（オンライン版：12月22日）

論文タイトル：“Evolution of the public opinion on COVID-19 vaccination in Japan: Large-scale Twitter data analysis”

著者：Ryota Kobayashi\*, Yuka Takedomi, Yuri Nakayama, Towa Suda, Takeaki Uno, Takako Hashimoto, Masashi Toyoda, Naoki Yoshinaga, Masaru Kitsuregawa, Luis E C Rocha.

DOI 番号 : 10.2196/41928

## 6. 注意事項 :

日本時間 12 月 23 日 (金) 午前 0 時 45 分 (米国東部時間 : 22 日 (木) 午前 10 時 45 分) 以前の公表は禁じられています。

## 7. 問い合わせ先 :

<研究に関すること>

東京大学大学院新領域創成科学研究科

准教授 小林 亮太 (コバヤシ リョウタ)

〒277-8561 千葉県柏市柏の葉 5-1-5 新領域基盤棟 5H4

Tel : 04-7136-4134

E-mail : r-koba@k.u-tokyo.ac.jp

<報道に関すること>

東京大学大学院新領域創成科学研究科 広報室

Tel : 04-7136-5450

E-mail : press@k.u-tokyo.ac.jp

情報・システム研究機構 国立情報学研究所

総務部企画課 広報チーム

Tel : 03-4212-2164

E-mail : media@nii.ac.jp

科学技術振興機構 広報課

〒102-8666 東京都千代田区四番町 5 番地 3

Tel : 03-5214-8404 / Fax : 03-5214-8432

E-mail : jstkoho@jst.go.jp

<JST の事業に関すること>

科学技術振興機構 戦略研究推進部 ICT グループ

前田 さち子 (マエダ サチコ)

〒102-0076 東京都千代田区五番町 7 K's 五番町

Tel : 03-3512-3526 / Fax : 03-3222-2066

E-mail : presto@jst.go.jp

## 8. 用語解説 :

(注 1) トピックモデル

トピックモデル (Latent Dirichlet allocation, LDA) は、ツイートの文章 (テキスト) に含まれる単語の情報 (単語分布) に基づき、多数のツイートをいくつかのトピックに分類する機械学習技術のことです。詳しい解説は Blei et al., Communications of the ACM, 2011 などを参照してください。トピックモデルは教師なし学習であるため、テキストデータだけからツイートをトピックに自動的に分類できます。しかし、この技術ではトピックの意味までは解釈でき

ません。そこで、本研究ではデータサイエンスと文学の研究者が協力して精査することによって分類されたトピックの意味解釈を行いました。

#### （注 2）陰謀論

陰謀論は、何らかの出来事（本研究では新型コロナのワクチン接種）に対して、明確な根拠なしに「邪悪で強力な組織や集団による陰謀によるものである」と断定したり、信じたりする主張のことです。ワクチンについての陰謀論の例としては、「新型コロナワクチンにはマイクロチップが入っていて、接種を受けると私達の行動が管理される」などがあります。ソーシャルメディア上で広まる陰謀論や偽情報が、ワクチンによる集団免疫獲得を阻害する要因の1つではないかと考えられています。

#### （注 3）中断時系列分析

中断時系列分析（Interrupted Time Series analysis）は、時系列（時間とともに変化する量のデータ）とあるイベントが起きた時刻から、そのイベントが時系列に与えた影響の大きさを見積もる統計的手法のことです。詳しい解説は Bernal et al., *International Journal of Epidemiology*, 2017 などを参照してください。本研究では、4つのイベント A. 医療従事者の接種開始、B. 高齢者の接種開始、C. 職域接種開始、D. 東京オリンピックを考え、これらのイベントがツイートの内容（各テーマの割合についての時系列）に与える影響について分析しました。

9. 添付資料：

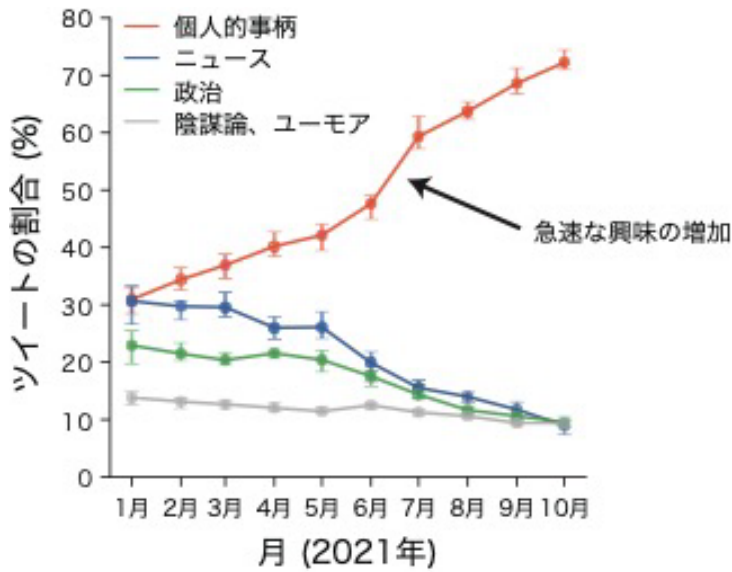


図1. ワクチン関連ツイートの主要4テーマ（1. 個人的事柄、2. ニュース、3. 政治、4. 陰謀論、ユーモア）の割合の月ごとの変化。6月以降、個人的事柄（赤）についてのツイートの割合が急上昇していることがわかります。

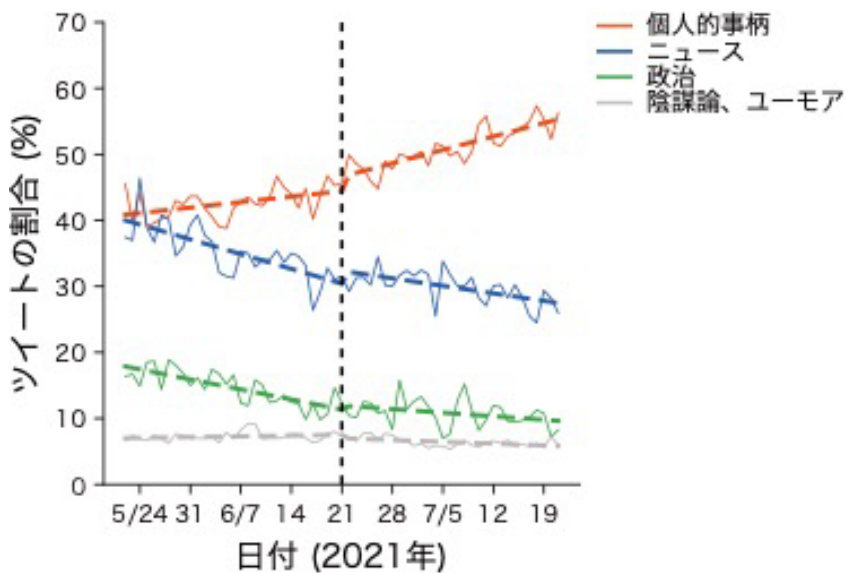


図2. 職域接種開始（2021年6月21日：縦の点線）前後におけるワクチン関連ツイートの主要4テーマの割合の日ごとの変化。