

## 大規模 Twitter データを用いた新型コロナワクチン接種に関する世論変遷の分析

武富有香<sup>1</sup> 中山悠理<sup>2</sup> 須田永遠<sup>1</sup> 宇野毅明<sup>1</sup>  
橋本隆子<sup>3</sup> 豊田正史<sup>4</sup> 吉永直樹<sup>4</sup> 喜連川優<sup>5</sup> Luis E C Rocha<sup>6</sup> 小林亮太<sup>2,7</sup>

国立情報学研究所 情報学プリンシプル研究系<sup>1</sup>  
東京大学大学院新領域創成科学研究科 複雑理工学専攻<sup>2</sup> 千葉商科大学 商経学部<sup>3</sup>  
東京大学大学院生産技術研究所<sup>4</sup> 国立情報学研究所 所長<sup>5</sup>  
Department of Economics, Ghent University<sup>6</sup> 東京大学 数理・情報教育研究センター<sup>7</sup>

## 1. はじめに

日本は先進国の中でも国民のワクチン信頼度が最も低い国のひとつとされる[1]。しかし日本の新型コロナワクチン接種は、欧米諸国と比べ接種開始が2ヶ月以上遅れたにも関わらず、医療従事者への接種開始からわずか8ヶ月ほどで接種率が72%（世界229カ国中14位）に達した。

日本での新型コロナワクチン接種が短期間で高い接種率に達した過程における過程で、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ある。これまで人々の関心のあり方は主にアンケートによって調査されてきた。しかし、大規模なアンケートを繰り返し実施することには大きなコストがかかり、人々の興味・関心の時間変化を追うことは難しい。一方、人々の関心のあり方を知るために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ており、ソーシャルメディア上の活動が現実の人々の興味や関心を反映することが示されている[2]。そこで本研究では Twitter でつぶやかれた長期間にわたる大規模データを分析することで、人々の関心や興味の推移を把握することを目的とした。さらに、重要な社会的イベントがツイートに与える影響についても分析を行った。

## 2. 方法

本研究では、2021年1月から10月までに800万ユーザが Twitter に投稿した「ワクチン」という語を含む日本語の全ツイート1.1億件を分

析した。このデータセットは NTT データから提供を受けた。まず、形態素解析を行ってツイートを単語に分割し、ストップワード（「これ」「する」など）を除去するなどの前処理を行った。次に、トピックモデル（Latent Dirichlet Allocation, LDA）[3]を用いてトピック分析を行った（詳しくは[4]を参照）。最後に中断時系列分析（Interrupted time series analysis）[5]を用いて、東京オリンピックの開催や職域接種開始などの重要な社会的イベントが世論に与える影響を調べた。

## 3. 結果

2021年1月から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（1.1億ツイート）のうち、引用ツイートやリツイートを除いたツイート（2,400万ツイート）に着目して分析した。データを収集した10ヶ月におけるツイート群の話題推移を調べるため、トピックモデル（LDA）を用いて、15個のトピックに分類を行った。各トピックからランダムに抽出したツイート群を実際に読み精査することにより意味解釈を行い、15のトピックに「接種会場からの実況」、「ワクチンの有効性」、「ワクチン政策に対する意見」などの名前をつけた。これらを4つの主要なテーマ（1. 個人的事柄、2. ニュース、3. 政治、4. 陰謀論とユーモア）に整理した。

つぎに、各テーマのツイート割合の時間変化を調べた（図1）。2021年1月当初は、個人的事柄（30%）、ニュース（30%）、政治（25%）、陰謀論とユーモア（15%）というように、4つのテーマにツイートは分散していた。しかし、6月以降には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し、10月には「ワクチン」を含むツイート全体の70%を占めるに至った。この結果は、6月以降に Twitter ユーザの興味・関心が個人的事柄に関する話題に集中したことを示唆する。また、陰謀論関連のキーワードが含まれるツイートの割合は全体の約6%にす

Evolution of the public opinion on COVID-19 vaccination in Japan

Yuka Takedomi<sup>1</sup>, Yuri Nakayama<sup>2</sup>, Towa Suda<sup>1</sup>, Takeaki Uno<sup>1</sup>, Takako Hashimoto<sup>3</sup>, Masashi Toyoda<sup>4</sup>, Naoki Yoshinaga<sup>4</sup>, Masaru Kitsuregawa<sup>5</sup>, Luis E C Rocha<sup>6</sup>, Ryota Kobayashi<sup>2</sup>  
<sup>1</sup>Principles of Informatics Research Division, National Institute of Informatics, <sup>2</sup>Graduate School of Frontier Sciences, The University of Tokyo, <sup>3</sup>Faculty of Commerce and Economics, Chiba University of Commerce, <sup>4</sup>Institute of Industrial Science, The University of Tokyo, <sup>5</sup>National Institute of Informatics, <sup>6</sup>Department of Economics, Ghent University, <sup>7</sup>Mathematics and Informatics Center, The University of Tokyo

ぎなかった。

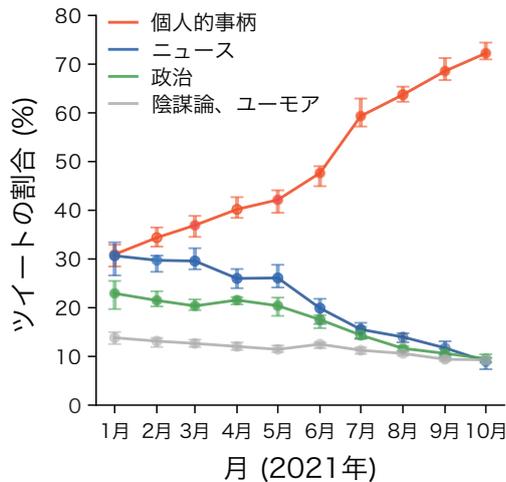


図1. 各テーマのツイート割合の時間変化

最後に、4つの社会的イベント (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中断時系列分析[5]を用いて分析した。この結果、C. 職域接種開始, D. 東京オリンピックの2つのイベントがツイートの話題に影響を与えることがわかった。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る。

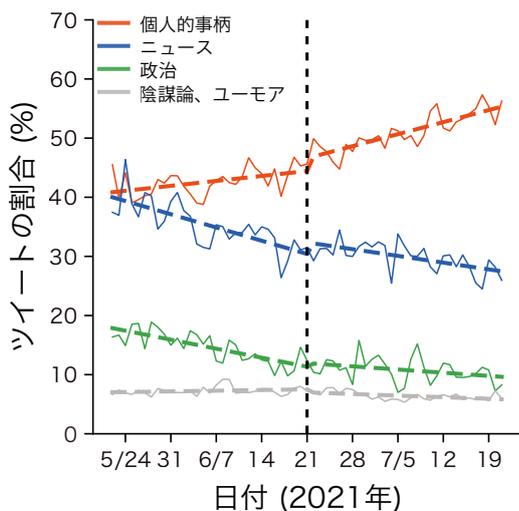


図2. 職域接種開始前後における各テーマのツイート割合の時間変化

#### 4. 結論

本研究では、2021年1月から10月までのワクチン接種期間中に、「ワクチン」という語を含む日本語の全ツイートを分析した。トピックモデルを用いてツイートを15のトピックに分類し、意味解釈を行うことにより4つの話題(個人的事柄, ニュース, 政治, 陰謀論とユーモア)に整理した。さらに中断時系列分析を行った結果、2021年6月の職域接種を境に、ワクチン政策, ワクチン関連ニュースなど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, 自身の副反応など個人的事柄に関するツイートの割合が増えたことを発見した。

#### 謝辞

本研究は、内閣府 Covid-19 AI・シミュレーションプロジェクトの一環として実施され、科学技術振興機構 (JST) JPMJPR1925, JPMJCR1401, 科学研究費助成事業 基盤研究(A) 19H0113, 21H04571, 基盤研究(B) 21H03559, 22H03695, 基盤研究(C) 18K11560, 22K12285, 二国間交流事業協同研究 JPJSBP1202022 01, 日本医療健康開発機構 JP21wm0525004 による支援を受けて行われました。

#### 参考文献

- [1] de Figueiredo A et al., (2020). "Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study.", *Lancet* 396: pp. 898-908.
- [2] Kwak H. et al., (2010). "What is Twitter, a social network or a news media?", *WWW '10*, pp.591-600.
- [3] Blei DM, Ng AY, Jordan MI. (2003). "Latent dirichlet allocation.", *J Mach Learn Res* 3: pp. 993-1022.
- [4] Kobayashi R, Takedomi Y, Nakayama Y, Suda T et. al., (2022). "Evolution of Public Opinion on COVID-19 Vaccination in Japan: Large-Scale Twitter Data Analysis.", *J Med Internet Res* 24: e41928.
- [5] Bernal JL, Cummins S, Gasparrini A. (2017). "Interrupted time series regression for the evaluation of public health interventions: a tutorial.", *Int J Epidemiol* 46: pp. 348-355.