



Title	X (旧 : Twitter) のバズツイートの変化から見るプラットフォーム事業者の施策がもたらすELSI
Author(s)	若林, 魁人
Citation	ELSI NOTE. 2023, 33, p. 1-40
Version Type	VoR
URL	<a href="https://doi.org/10.18910/93093">https://doi.org/10.18910/93093</a>
rights	
Note	

*The University of Osaka Institutional Knowledge Archive : OUKA*

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka



# X (旧：Twitter) の バズツイートの変化から見る プラットフォーム事業者の 施策がもたらすELSI

---

## Author

---

若林 魁人

大阪大学 社会技術共創研究センター 特任研究員 (2023年11月現在)

---

本稿で取り上げる X のサービス動向は 2023 年 10 月 20 日時点の内容であり、その後もさまざまな施策や仕様変更が行われている。それぞれの情報については必ず脚注に出典を明記するので、引用する際にはその時点でのソースや状況を確認することを推奨する。

## 目次

はじめに.....	3
<b>1. Elon Musk 氏による買収後の X のサービス動向.....</b>	<b>5</b>
1.1. 閲覧数 (View Count) の可視化.....	5
1.2. 自殺防止機能の一時的な削除.....	7
1.3. 「あなたへのおすすめ (For You) 」タイムラインの強制表示.....	7
1.4. レコメンデーションアルゴリズムの公開.....	9
1.5. API の価格と仕様の変更.....	11
1.6. Community Note 機能の実装.....	13
1.7. 閲覧数に応じた広告収益分配プログラムの導入.....	14
1.8. 生体認証情報と雇用情報の収集.....	15
1.9. 外部サイトの記事の見出し情報削除.....	17
1.10. ブロック機能の将来的な廃止.....	18
<b>2. 現在のユーザーの行動に関する主成分分析.....</b>	<b>19</b>
2.1. 論点と分析目的.....	19
2.2. 方法.....	21
2.3. 結果.....	30
<b>3. X のサービス動向がもたらす可能性のある ELSI.....</b>	<b>34</b>
3.1. 煽動的バズツイートの拡散起点の多様化による分断の加速.....	34
3.2. 収益獲得や発信活動を目的とするユーザーの行動統制.....	36
3.3. プラットフォームの社会的影響や責任の不透明化.....	36
<b>4. おわりに.....</b>	<b>38</b>

## はじめに

ソーシャルメディアは人々の情報行動にとって重要な存在となって久しく、それと同時にさまざまな問題ももたらしている。例えば自分と似た興味関心をもつユーザーをフォローしてゆく中で自分と似た意見だけに触れてしまうエコーチェンバーや、表示最適化アルゴリズムによって「見たい情報」だけが表示されることで異なる考え方や価値観から隔離されてしまうフィルターバブル<sup>1</sup>によってもたらされる分断・情報の偏向が長らく問題視されている。これは倫理的・法的・社会的課題（ELSI）の観点でも重要な論点であり、例えば米国ではソーシャルメディア事業者にレコメンデーションアルゴリズムの説明やユーザーによるレコメンド機能の利用に対する選択の自由を義務付けるフィルターバブル透明性法案が2019年より議論されている（1.3節）。

特に日本においてはX(旧:Twitter)がソーシャルメディアの中でも人気が高い。Twitterは2006年に米国で創業したTwitter社がリリースしたソーシャルネットワーキングサービスである。日本には2008年に上陸しており、日本人のユーザー数は米国に次いで2番目の約6,700万人にのぼる<sup>2</sup>。またTwitter上でのエコーチェンバーやフィルターバブルについても、偽情報拡散<sup>3</sup>や政治的意見の対立<sup>4</sup>などさまざまな文脈で研究が行われてきた。

2022年10月27日、起業家のElon Musk氏はTwitter社（現在はX社へ統合され消滅<sup>5</sup>）を440億ドルで買収し、幹部らを解雇してCEOに着任した<sup>6</sup>。その後、2023年7月末にはサービス名称がTwitterからXへ変更された。Elon Musk氏は買収に際して、目的は金銭的利益ではなく自由な言論が可能なプラットフォームの構築であること<sup>7</sup>、アルゴリズムの透明性の向上やスパムボットの排除などのために様々な施策を行ってゆくこと<sup>6</sup>を述べた。実際にElon Musk氏の着任以降、さまざまな施策が行われている。しかし、それらの施策は必ずしも同氏の述べた目的の達成

---

<sup>1</sup> 総務省, “令和元年版 情報通信白書のポイント”. <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/html/nd114210.html>

<sup>2</sup> “Leading countries based on number of X (formerly Twitter) users as of January 2023”. Statista. <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/> (2023.10.20 閲覧)

<sup>3</sup> Grinberg, Nir, et al. “Fake news on Twitter during the 2016 US presidential election.” *Science* 363.6425 (2019): 374-378. <https://doi.org/10.1126/science.aau2706>

<sup>4</sup> Baumann, Fabian, et al. “Modeling echo chambers and polarization dynamics in social networks.” *Physical Review Letters* 124.4 (2020): 048301. <https://doi.org/10.1103/PhysRevLett.124.048301>

<sup>5</sup> Low de Wei and Marika Katanuma. “Twitter company ‘no longer exists’; it’s now part of Musk’s X Corp”. *Los Angeles Times*. (2022.10.27). <https://www.latimes.com/business/story/2023-04-11/elon-musk-merges-twitter-x-corp>

<sup>6</sup> Kate Conger and Lauren Hirsch. “Elon Musk Completes \$44 Billion Deal to Own Twitter”. *The New York Times*. (2022.10.27). <https://www.nytimes.com/2022/10/27/technology/elon-musk-twitter-deal-complete.html>

<sup>7</sup> Mia Sato. “Buying Twitter ‘is not a way to make money,’ says Musk in TED interview”. *The Verge*. (2022.4.15). <https://www.theverge.com/2022/4/14/23025343/elon-musk-twitter-takeover-ted-talk-quote-stock-buyout>

につながっておらず、逆に既存の問題の増幅や新しい ELSI をもたらすことにつながる可能性がある。例えば Elon Musk 氏の就任直前後のツイート<sup>8</sup>を分析し、同氏の施策によってヘイトスピーチと bot が増えた可能性を指摘した先行研究<sup>9</sup>なども存在する。

そこで本稿では Elon Musk 氏による X 買収後に行われた施策を ELSI の観点から概観した上で、ユーザーのツイートに見られる人々の情報行動の変化について分析する。それによって現在の X のサービス動向がもたらす可能性のある ELSI、プラットフォーム事業者の社会的責任について考察する。

なお、本稿についていくつかの注釈をあらかじめ付記する。

- ・ 本稿では古典的統計手法から機械学習を用いた手法までを含むいくつかのデータ分析を行っているが、いずれも後述する X API に課せられた制限下で可能な範囲での簡易な分析であり、分析プロトコルや結果に十分な妥当性は担保されていない。限られたデータ資源による局所的なスナップショットの記録、および X API の制限がソーシャルメディア研究にもたらす影響そのものへの考察に焦点を当てることに留意されたい。
- ・ X API のドキュメントに記された仕様、分析過程で明らかとなった仕様についても書き残すが、これについても 2023 年 10 月 20 日時点で確認されたものであり、今後仕様が変更される可能性がある。
- ・ 参考にした文献等は全て脚注に記した。ウェブサイトについては 2023 年 10 月 20 日時点でアクセスを確認しており、今後アクセスできなくなる可能性がある。

---

<sup>8</sup> Twitter から X へのサービス名称変更に伴ってユーザーの投稿および再投稿の機能の名称は「ツイート」「リツイート」から「ポスト」「リポスト」へ変更されたが、本稿の一部で「バズツイート」など以前の機能名称に由来するトピックを扱う都合上、簡単のため本稿では「ツイート」「リツイート」の名称を用いる。

<sup>9</sup> Hickey, Daniel, et al. "Auditing elon musk's impact on hate speech and bots." Proceedings of the International AAAI Conference on Web and Social Media. Vol. 17. 2023. <https://doi.org/10.1609/icwsm.v17i1.22222>

# 1. Elon Musk 氏による買収後の X のサービス動向

本章では Elon Musk 氏による買収が行われた 2022 年 10 月 27 日から 2023 年 10 月 20 日までの、機能追加や規約変更など X のサービス動向のうち ELSI の観点から主要なトピックとその論点を概観する。なお、社名やサービス名はいずれもその時点での名称に合わせて記述する。



図 1 本稿で取り上げる X のサービス動向の時系列

## 1.1. 閲覧数 (View Count) の可視化

2022 年 12 月中旬から下旬<sup>10</sup>にかけて、ツイートの閲覧数を第三者からも確認できるように表示する機能が実装された<sup>11</sup>。Elon Musk 氏はこの機能実装に際して、各ユーザーのツイートはいいね・リツイートの数値から受ける印象以上に閲覧されていることに言及し、プラットフォーム全体のメディアとしての価値を可視化することが目的であったことを述べている (図 2)<sup>12</sup>。

<sup>10</sup> なお 2.2 節の分析では、具体的に 2022 年 12 月 16 日以降のツイートから X API を通して閲覧数の取得が可能であった。

<sup>11</sup> Mitchell Clark. “Twitter is now showing everyone how many views your tweets get”. The Verge. (2022.12.23). <https://www.theverge.com/2022/12/21/23522064/twitter-view-count-roll-out-personal-info>

<sup>12</sup> <https://x.com/elonmusk/status/1605957811162054656> (2023.10.20 閲覧)



図 2 Elon Musk 氏の閲覧数表示についての言及と、閲覧数表示の例（赤枠）

この機能実装について、例えば動画共有サービス niconico の代表を務める栗田穰崇氏は「表示回数がオープンになることでフォロワー数が多くてもあまり見られていないアカウントが可視化されてしまう」と言及した<sup>13</sup>。ソーシャルメディア上で他者からの反応が数値化され続けることが自己承認やコミュニケーションへの悪影響、いわゆる「ネット依存」を引き起こす一因となることが指摘されている<sup>14</sup>。そこで2019年にはFacebook社（現在のMeta社）が運営するInstagramとFacebookでの「いいね」の数値の非表示機能の実装<sup>15</sup>、2021年には動画共有サービスYouTubeでの「低評価」のカウントの非表示<sup>17</sup>など、いくつかの主要なソーシャルメディアは自身のプラットフォーム上での数値的評価が人々にもたらすネガティブな影響に対応するための施策を行っている。

なお2023年10月6日にElon Musk氏は、将来的にツイートの閲覧数カウントのみを前面に表示し、いいねやリツイートなどのユーザーインタラクション数はツイート情報をタップした際のみ表示されるよう変更することを述べた<sup>18</sup>。同氏はこの変更の目的を「ツイートを読みやすく

<sup>13</sup> <https://x.com/sigekun/status/160609415477231360> (2023.10.20 閲覧)

<sup>14</sup> Jamie Leventhal. "How removing 'likes' from Instagram could affect our mental health". PBS NewsHour. (2019.11.25). <https://www.pbs.org/newshour/science/how-removing-likes-from-instagram-could-affect-our-mental-health>

<sup>15</sup> Catherine Thorbecke. "Facebook tests hiding likes to see if it will change how people engage with posts". abc NEWS. (2019.9.28). <https://abcnews.go.com/Business/facebook-tests-hiding-likes-change-people-engage-posts/story?id=65901080>

<sup>16</sup> Madeline Fitzgerald. "Instagram Starts Test To Hide Number of Likes Posts Receive for Users in 7 Countries". TIME. (2019.7.18). <https://time.com/5629705/instagram-removing-likes-test/>

<sup>17</sup> Mitchell Clark. "YouTube gives dislikes the thumbs-down, hides public counts. The Verge. (2021.11.11). <https://www.theverge.com/2021/11/10/22773299/youtube-dislike-button-hide-public-count-numbers-small-creator-protection>

<sup>18</sup> <https://x.com/xDaily/status/1710154397513388444> (2023.10.20 閲覧)

するため」と説明している。

## 1.2. 自殺防止機能の一時的な削除

2022年12月20日前後に、人々の自殺防止をはじめとする支援機能を提供する#ThereIsHelpが削除されていたことが発覚した<sup>19</sup>。#ThereIsHelpは、メンタルヘルス、HIV、ワクチン、児童の性的搾取、新型コロナウイルス感染症、ジェンダーに基づく暴力、自然災害、表現の自由など、ユーザーの安全リスクが示唆される検索が行われた際に、その内容に応じた支援団体の連絡先などが表示される機能である。

機能の削除へのアナウンスはなく、また関係者からはこの削除はElon Musk氏の指示によるものであったことが述べられた。その後、ユーザーや消費者安全団体などからの批判を受けたのち、同月25日に機能を復旧した。Twitter社のHead of Trust and Safety（信頼性と安全性に関する責任者）であるElla Irwin氏は機能を削除していたことを認めつつ、機能改善の作業中に一時的に削除しただけであると述べた<sup>20</sup>。

なおElon Musk氏がTwitter社への就任にあたって実施した人員削減によって、同月12日にはヘイトスピーチ・自殺・児童搾取などの問題に取り組む同社の委員会であるTrust and Safety Councilは解散を命じられていた<sup>21</sup>。

## 1.3. 「あなたへのおすすめ（For You）」タイムラインの強制表示

Xにはユーザーがフォローしているユーザーのツイートを表示する「Following」とアルゴリズムにおすすめされたツイートによる「For You」の2種類のタイムラインが実装されている。2023年1月11日に、サービス起動時には「For You」タイムラインがデフォルトで表示され、「Following」タイムラインを表示するにはその都度スワイプ操作を行わせるようにUIが変更された<sup>22</sup>。従来は2種類のどちらのタイムラインを表示するかをユーザーが選択することが

---

<sup>19</sup> “Elon Musk orders removal of Twitter suicide prevention feature, sources say”. The Japan Times. (2022.12.24).

<https://www.japantimes.co.jp/news/2022/12/24/business/twitter-elon-musk-suicide-prevention/>

<sup>20</sup> Kenneth Li. “Twitter restores suicide prevention feature”. The Japan Times. (2022.12.25).

<https://www.japantimes.co.jp/news/2022/12/24/business/twitter-elon-musk-suicide-prevention/>

<sup>21</sup> Matt O’Brien and Barbara Ortutay. “Musk’s Twitter disbands its Trust and Safety advisory group”. Associated Press News. (2022.12.13).

<https://apnews.com/article/elon-musk-twitter-inc-technology-business-a9b795e8050de12319b82b5dd7118cd7>

<sup>22</sup> Mitchell Clark. “Twitter defaults to a For You page now, just like TikTok”. The Verge. (2023.1.11).

でき、選択しなかったタイムラインは表示されなかった。またおすすめ表示のレコメンドアルゴリズム自体についても、2022年12月には「フォローしていないユーザーのツイートを積極的に表示する」などの修正が予告されており<sup>23</sup>、ユーザーからは自身がフォローしていないユーザーのツイートを強制的に見せられることへの批判も上がった<sup>24</sup>。ユーザーからの一連の批判を受け、同月25日に Elon Musk 氏は「アプリ起動時には、ユーザーが最後に選択したタイムラインを表示するよう修正する」という旨をツイートした<sup>25</sup>。その後のアップデートによって「For You」タイムラインのデフォルト表示はされなくなった。ただしサービス名称を X に変更後には再びおすすめ表示をデフォルト表示するよう修正され<sup>26</sup>、本稿執筆時点でも「For You」タイムラインがデフォルトで表示されている。

Twitter 社は Elon Musk 氏による買収以前の 2022 年 3 月 11 日にも、おすすめ表示タイムラインをデフォルト表示する機能を一時的に実装し、同月 14 日にはユーザーからの批判を受けて取り下げた<sup>27</sup>。米国では同時期に、企業がユーザーデータを用いた表示最適化を強制せず、レコメンド機能のオンオフを可能にすることを義務付けるフィルターバブル透明性法案 (Filter Bubble Transparency Act) が超党派の議員グループから提出される<sup>28</sup>など、不透明なアルゴリズムにより人々の情報行動が制御されることへの関心が一部で高まっていた。

---

<https://www.theverge.com/2023/1/10/23549368/twitter-for-you-page-tiktok-following-tab>

<sup>23</sup> Taylor Hatmaker. “Twitter is going to show you more tweets from people you don’t follow”. TechCrunch. (2022.12.1).

<https://techcrunch.com/2022/11/30/twitter-recommended-tweets/>

<sup>24</sup> Billy Perrigo. “Why Your Twitter Feed Is Suddenly Full of People You Don’t Follow”. TIME. (2023.2.1).

<https://time.com/6251833/twitter-for-you-page-tiktok/>

<sup>25</sup> <https://x.com/elonmusk/status/1616594332907372544> (2023.10.20 閲覧)

<sup>26</sup> Sarah Perez. “Twitter alternative T2 gets its own For You feed, but won’t force it as the default”. TechCrunch. (2023.9.7).

<https://techcrunch.com/2023/09/06/twitter-alternative-t2-gets-its-own-for-you-feed-but-wont-force-it-as-the-default/>

<sup>27</sup> Sarah Perez. “Twitter’s latest update makes it easier to escape its recommendation algorithm (Update: Twitter removed it!)”.

TechCrunch. (2022.3.11). <https://techcrunch.com/2022/03/10/twitters-latest-update-makes-it-easier-to-escape-its-recommendation-algorithm/>

<sup>28</sup> 本法案は John Thune 上院議員を中心に 2019 年に初めて提出されたものであり、2021 年 6 月に同議員らが主導して再度の法案提出、それに付随する形で 11 月に超党派のグループが法案を提出した。2023 年 7 月には米国上院通商科学運輸委員会で可決され、国会での可決を今後目指す。ここで可決された法案は、一定以上の規模のインターネットプラットフォーム事業者に対して、アルゴリズムの使用状況の開示と表示最適化のオンオフの切り替えの権利の提供を義務付け、その要件を満たさない事業を違法として連邦取引委員会 (FTC) が民事罰を求める権限を与えるものである。

参考：

Igor Bonifacic. “Bipartisan bill seeks to curb recommendation algorithms”. TechCrunch. (2021.11.10).

<https://techcrunch.com/2021/11/09/bipartisan-bill-seeks-to-curb-recommendation-algorithms/>

John Thune. “Thune’s Big Tech Algorithm Transparency Bill Unanimously Approved by Commerce Committee”. (2023.7.27).

<https://www.thune.senate.gov/public/index.cfm/2023/7/thune-s-big-tech-algorithm-transparency-bill-unanimously-approved-by-commerce-committee>

レコメンデーション機能によってソーシャルメディア上の分断が強化されるリスクについては多くの先行研究で検討されており、例えば Geschke et al. (2019) はレコメンデーション機能が個人レベルでは既存の態度、行動、アイデンティティを強化して安心感を与える方向に作用し、社会レベルでは他者との態度の違いを増大させながらもコミュニケーションを切断することで社会の断片化や二極化に作用することを示した<sup>29</sup>。Noordeh et al. (2020) は、レコメンデーションアルゴリズムに晒され続けることでユーザーが触れるコンテンツの多様性は損なわれてゆき、またそうしてフィルターバブルに陥ったユーザーは、個人レベルで（例えば「多様な情報に触れるよう心がける」といった形で）システム上での振る舞いを変えたとしてもレコメンダされるコンテンツの多様性を取り戻すことはできないことを指摘した<sup>30</sup>。これらの指摘に同意するならば、おすすめされたコンテンツを強制的に「ひとまず視界に入れさせる」機能を持ったプラットフォームは、選ばなくとも自分の気に入るコンテンツがおすすめされることによる無意識な安心感と引き換えに他者との分断を促進する可能性がある（この可能性を2章より検討する）。

また、同じ2023年1月11日には有料サブスクリプションサービス Twitter Blue（現在の X Premium）が日本でも提供開始された<sup>31</sup>。サブスクリプションサービスへ登録したユーザーのツイートがレコメンデーションアルゴリズムに優遇されることも公開されており<sup>32</sup>、例えばツイートのリプライ欄などを開くと X Premium へ登録した認証済みユーザーが優先して表示されることが実際に確認できる。

#### 1.4. レコメンデーションアルゴリズムの公開

2023年3月31日、Twitter社は「For You」タイムラインに表示するツイートを選択するレコメンデーションアルゴリズムのソースコードを GitHub に公開した<sup>33,34</sup>。Twitter社はこの目的を

---

<sup>29</sup> Geschke, Daniel, Jan Lorenz, and Peter Holtz. "The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers." *British Journal of Social Psychology* 58.1 (2019): 129-149.

<https://doi.org/10.1111/bjso.12286>

<sup>30</sup> Noordeh, Emil, et al. "Echo chambers in collaborative filtering based recommendation systems." arXiv preprint arXiv:2011.03890 (2020).

<https://doi.org/10.48550/arXiv.2011.03890>

<sup>31</sup> Reece Rogers. "日本でも開始、Twitterの有料サービス「Twitter Blue」とは？". WIRED. (2023.1.11).

<https://wired.jp/article/what-is-twitter-blue/>

<sup>32</sup> <https://help.twitter.com/en/using-x/x-premium> (2023.10.20 閲覧)

<sup>33</sup> "Twitter's Recommendation Algorithm". Twitter. (2023.3.31).

[https://blog.twitter.com/engineering/en\\_us/topics/open-source/2023/twitter-recommendation-algorithm](https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm)

<sup>34</sup> <https://github.com/twitter/the-algorithm#for-you-timeline> (2023.10.20 閲覧)

プラットフォームの透明性を担保するためであると述べた<sup>35</sup>。なお、ユーザーの安全とプライバシー保護のために、悪意あるユーザーによるアルゴリズムの逆用につながりうるコード、アルゴリズムに関連するトレーニングデータなどは非公開としている。

例えば 1.3 節で述べたフィルターバブル透明性法案など、レコメンデーション機能の透明性の担保の重要性については数多くの議論が行われており、Elon Musk 氏も Twitter 社の買収前からその重要性への同意を述べている<sup>6</sup>。一方で「ソースコードの公開」がレコメンデーション機能の透明性の担保につながるかについてはいくつかの指摘が存在する。

機械学習に基づくシステムはその開発者であっても入出力関係を理解、決定論的に制御できるものでないため、プラットフォームの社会的影響や責任を明確化するにはソースコードの読解のみならず、実際の入出力に関する検証も同時に行われる必要がある。それにもかかわらず、レコメンデーションアルゴリズムの公開と同日より実施された Twitter API の大幅な値上げや機能制限を伴う仕様変更 (1.5 節) によって、ツイートのデータを用いた分析の多くには制約が課された。Mozilla の上級政策研究員である Brandi Geurkink 氏は、公開されたアルゴリズム内には同時期に社会的関心の高まっていたウクライナ侵略に言及したツイートを人々に表示させないように重み付けを行うと思われる箇所が存在したものの、API の利用制限によってそれを実証することが困難となっていたことを指摘して「本当に Twitter のアルゴリズムに説明責任をもたらしたかったのであれば、アルゴリズムの公開に加えて精査も可能にできたはずだ」と批判した<sup>36</sup>。

加えて公開されたアルゴリズム自体も、プラットフォーム事業者の利益、あるいは (Twitter 社自身も述べているように<sup>35</sup>) アルゴリズムの悪意あるハックの防止など、様々な意図と裁量によって公開範囲が制限されうる可能性がある。例えばアルゴリズム公開直後には Elon Musk 氏のツイートをピックアップして追跡する機能が存在していることが見つかったが<sup>37</sup>、同氏はそのような機能が含まれていることを知らなかったと主張し、該当箇所は GitHub 上のソースコードから削除された。一方で 2023 年 2 月中旬には、同氏は自身のツイートが積極的にレコメンダされるようにアルゴリズムを修正することをエンジニアへ命じたという Twitter 社員からのリーク情

---

<sup>35</sup> “A new era of transparency for Twitter”. Twitter. (2023.3.31).

[https://blog.twitter.com/en\\_us/topics/company/2023/a-new-era-of-transparency-for-twitter](https://blog.twitter.com/en_us/topics/company/2023/a-new-era-of-transparency-for-twitter)

<sup>36</sup> Brandi Geurkink. “Twitter’s Open Source Algorithm Is a Red Herring”. WIRED. (2023.4.7).

<https://www.wired.com/story/twitters-open-source-algorithm-is-a-red-herring/>

<sup>37</sup> Matt Binder. “Twitter’s algorithm specifically tracks how Elon Musk’s tweets are doing”. TechCrunch. (2023.3.31).

<https://mashable.com/article/twitter-releases-algorithm-showing-it-tracks-elon-musk-tweets>

報が報じられており<sup>38</sup>、また同時期から Elon Musk 氏のツイートがおすすめタイムラインに数多く表示されるようになったというユーザーの声も上がっていた<sup>39</sup>。

また、レコメンデーションアルゴリズムによる重み付けパラメータを公開することはユーザーの振る舞いから多様性を奪うという視点も存在する。例えば英国のデータ倫理イノベーションセンター（Centre for Data Ethics and Innovation; CDEI）によるレコメンデーションアルゴリズムの音楽産業への影響の調査報告書<sup>40</sup>では、アルゴリズムを公開することで、クリエイターがアルゴリズムにレコメンデされやすいように振る舞いを最適化してしまうことによる画一化、システムによる支配がもたらされるリスクが指摘された。実際に X のサービス上でも、公開されたソースコードから「どのようなツイートやユーザーがレコメンデされやすいか」を様々なユーザーが読解を試みており<sup>41</sup>、例えば「いいねやリツイートの数よりもリプライの数の方がツイートの評価につながる重み付けが大きい」と判明すると、インフルエンサーや動画投稿者などが自身のフォロワーにリプライでのインタラクションを呼びかけるなど、アルゴリズムにレコメンデされるための行動変容も見られる。

## 1.5. API の価格と仕様の変更

2023 年 2 月 2 日、Twitter 社は Twitter API（現在は X API）<sup>42</sup>の無料アクセスのサポートの終了と新しい有料プランの提供を発表し<sup>43</sup>、2023 年 3 月 31 日より新プランの提供が実施された。この変更では実質的な大幅な値上げと、利用可能な機能や機能の呼び出し回数への制約が課せられている。新プランでは、Free（無料）、Basic（月額 100 ドル）、Pro（月額 5000 ドル）、Enterprise（個別問い合わせ、後述）の 4 つのプランが用意されている（図 3）<sup>44</sup>。例えばビッグデータ解

---

<sup>38</sup> Zoe Schiffer and Casey Newton. “Yes, Elon Musk created a special system for showing you all his tweets first”. Platformer (2023.2.15).

<https://www.platformer.news/p/yes-elon-musk-created-a-special-system>

<sup>39</sup> Emma Roth. “Twitter is just showing everyone all of Elon Musk’s tweets now”. The Verge. (2023.2.14).

<https://www.theverge.com/2023/2/13/23598514/twitter-algorithm-elon-musk-tweets>

<sup>40</sup> “The impact of recommendation algorithms on the UK’s music industry”. Centre for Data Ethics and Innovation. (2023.2.9).

<https://www.gov.uk/government/publications/research-into-the-impact-of-streaming-services-algorithms-on-music-consumption/the-impact-of-recommendation-algorithms-on-the-uks-music-industry>

<sup>41</sup> Steven Tey. “How the Twitter Algorithm works in 2023”. Twitter. (2023.3.31). <https://steventey.com/blog/twitter-algorithm>

<sup>42</sup> API（アプリケーションプログラミングインタフェース）はシステム同士を接続するためのインターフェースであり、多くの場合は外部から各種機能を利用するための操作方法をサービス・システム提供者が公式に提供するものである。X API（Twitter API）の場合は、自動ツイートや外部サービスからのツイート、X のユーザーアカウントを用いた外部サービスのユーザー認証、ツイートやユーザーデータの大規模取得などを行うための機能が提供されている。

<sup>43</sup> <https://x.com/XDevelopers/status/1621026986784337922> (2023.10.20 閲覧)

<sup>44</sup> <https://developer.twitter.com/en/products/twitter-api> (2023.10.20 閲覧)

析やトレンド分析などに用いられるツイートの検索・詳細情報取得では、それまでは無料で月間200万ツイートまでのデータ取得が可能であったが<sup>45</sup>、新プランでは Pro プランであっても取得数の上限は月間100万ツイートとなっている。

**Find the right access for you**

Free	Basic	Pro	Enterprise
For write-only use cases and testing the X API	For hobbyists or prototypes	For startups scaling their business	For businesses and scaled commercial projects
<ul style="list-style-type: none"> <li>Rate limited access to v2 post posting and media upload endpoints</li> <li>1,500 Posts per month - posting limit at the app level</li> <li>1 app ID</li> <li>Login with X</li> <li>Free</li> </ul>	<ul style="list-style-type: none"> <li>Rate limited access to suite of v2 endpoints</li> <li>3,000 Posts per month - posting limit at the user level</li> <li>50,000 Posts per month - posting limit at the app level</li> <li>10,000 Posts per month - read-limit rate cap</li> <li>2 app IDs</li> <li>Login with X</li> <li>\$100 per month</li> </ul>	<ul style="list-style-type: none"> <li>Rate-limited access to suite of v2 endpoints, including search and filtered stream</li> <li>1,000,000 Posts per month - GET at the app level</li> <li>300,000 Posts per month - posting limit at the app level</li> <li>3 app IDs</li> <li>Login with X</li> <li>\$5,000 per month</li> </ul>	<ul style="list-style-type: none"> <li>Commercial-level access that meets your and your customer's specific needs</li> <li>Managed services by a dedicated account team</li> <li>Complete streams: replay, engagement metrics, backfill, and more features</li> <li>Monthly subscription tiers</li> </ul>
<a href="#">Get started</a>	<a href="#">Subscribe now</a>	<a href="#">Subscribe now</a>	<a href="#">Apply now</a>

図 3 X Developer ページに記載されている X API のアクセス権プラン一覧

この変更によって、地震情報を発信するアカウントが災害時の運用に問題が生じる可能性をアナウンスする<sup>46</sup>、X との連携を伴うサービスの多くが有料化や値上げを行う<sup>47</sup>など、多くのサービス提供者への影響が生じた。またサービス提供者や事業者のみならず、偽情報拡散やエコーチェンバー形成などに関するソーシャルメディア研究も影響を受けている。Twitter 社からの Enterprise プランの営業を受け取った研究者らによると、同プランは最も安価なパッケージでも月額42,000ドルで5,000万件、上位のパッケージでは月額21万ドル2億件のツイートへのアクセスが提供される<sup>48</sup>。この営業を受け取った研究者らは、データ取得に月額42,000ドルを出資できる研究者はほとんど存在しないこと、仮にそれが可能であったとしても、Pro プラン以下はもちろん、Enterprise プランであっても取得可能な情報量はこれまで行われてきたソーシャルメディア研究

<sup>45</sup> Daigo Wakabayashi. “【2022年】TwitterAPI v2 の仕様まとめ・セットアップ方法”. Zenn. (2022.4.2).

[https://zenn.dev/mamushi/articles/twitter\\_api\\_v2\\_setup](https://zenn.dev/mamushi/articles/twitter_api_v2_setup)

<sup>46</sup> [https://x.com/earthquake\\_jp/status/1623494931666046977](https://x.com/earthquake_jp/status/1623494931666046977) (2023.10.20 閲覧)

<sup>47</sup> 例えば、“Updates to Twitter access”. IFTTT. <https://ifttt.com/explore/updates-to-twitter-2023>

<sup>48</sup> Daisuke Takimoto. “Twitter API の法人向け利用料が「月額500万円超から」になり、“誰も使えない代物”になるという決定的な証拠”. WIRED.jp. (2023.3.14). <https://wired.jp/article/twitter-data-api-prices-out-nearly-everyone/>

には不十分であることを指摘した。加えて、2021年1月26日以降は学術研究用のAPI利用プランも提供されていた<sup>49</sup>が、現在はこのプランでのAPIアクセスは不可能となっている。ユーザーからの意見をを受けて2023年3月30日にはTwitter社より学術研究用のAPI利用プランについても検討する旨のツイートがあった<sup>50</sup>が、以降アナウンス等は行われていない。

さらに2023年6月には、ユーザーのフォロー・フォロワーのリストを取得する機能が事前通知なく廃止された<sup>51</sup>。Twitter社はユーザーからの問い合わせや批判に回答することはなく、機能廃止の4日後に変更ログの更新のみを行った。変更ログによると同機能は(先述した月額最低42,000ドルの)Enterpriseプラン以外で廃止されたことが報告されている<sup>52</sup>。この影響によりフォロー管理サービスやサードパーティ製のクライアントアプリがサービス終了することとなった<sup>53</sup>。

なお、これらのX APIの機能制限は本稿で実施した分析にも影響を与えている。そこで、2章では本稿での分析プロトコルと合わせて、具体的にX APIを介して取得できる情報やプランごとの制約、廃止された機能、それらがソーシャルメディア研究の先行研究群で提案されてきた分析プロトコルに与える影響についても合わせて紹介する。

## 1.6. Community Note 機能の実装

2022年12月11日より、誤解を招く可能性があるツイートにユーザー同士で背景情報を追加できるCommunity Note機能が一般公開された<sup>54</sup>。同機能は日本のユーザー向けには2023年7月6日より一般公開、参加可能となった<sup>55</sup>。これはElon Musk氏の買収前にBirdwatch機能として構想されていた<sup>56</sup>ものであり、同氏による買収直後に名称がCommunity Noteへと変更された<sup>57</sup>。

---

<sup>49</sup> Larry Dignan. “Twitter、新APIの「学術研究」トラックで公開ツイート全履歴への無料アクセスなど提供”. CNET Japan. (2021.1.27).

<https://japan.cnet.com/article/35165639/>

<sup>50</sup> <https://x.com/XDevelopers/status/1641222788911624192> (2023.10.20 閲覧)

<sup>51</sup> 松浦立樹. “Twitter API、また急に仕様変更 有料機能「フォロワー／フォロワー取得」を突然削除 事前告知なし、開発者から批判相次ぐ”. ITmedia NEWS. (2023.6.28). <https://www.itmedia.co.jp/news/articles/2306/28/news175.html>

<sup>52</sup> <https://developer.twitter.com/en/updates/changelog> (2023.10.20 閲覧)

<sup>53</sup> Matt Binder. “Twitter’s API keeps breaking, even for developers paying \$42,000”. Mashable. (2023.6.29).

<https://mashable.com/article/twitter-api-elon-musk-developer-issues-apps>

<sup>54</sup> <https://x.com/CommunityNotes/status/1601753552476438528> (2023.10.20 閲覧)

<sup>55</sup> <https://x.com/CommunityNotes/status/1676767470961532928> (2023.10.20 閲覧)

<sup>56</sup> Sarah Perez. “Twitter is building ‘Birdwatch,’ a system to fight misinformation by adding more context to tweets”. TechCrunch.. (2020.10.3).

<https://tcrn.ch/3jqbUR7>

<sup>57</sup> [https://blog.twitter.com/en\\_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation](https://blog.twitter.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation) (2023.10.20 閲覧)

Community Note 機能は、協力者として登録したユーザーが任意のツイートに対してメモを追加し、その内容に他の協力者ユーザーから一定数の「役に立つ」という評価が集まった場合のみ一般公開される機能である。また、メモの内容はプライバシーポリシーに違反しない限り X 社は編集・変更できず、また同社の視点を表すものではないとしている<sup>58</sup>。

Elon Musk 氏の就任以降、Trust and Safety Council（1.2 節）をはじめ、X 社からは情報の信頼性や安全性を担保する多くのチームが解雇された<sup>59</sup>。同時に Elon Musk 氏は Community Note 機能について「Twitter の精度を向上させる画期的なツール」であると考えている旨をツイートしていた<sup>60</sup>。一方で米国での機能公開時には、選挙の公正性など保守派かリベラル派かの立場が意見に反映されるテーマへのメモは協力者ユーザー自身の立場が評価に反映されるために中立的な評価が集まりづらく、その結果ほとんどのメモは公開に至っていないこと、意見を対立させるトピックの信頼性向上にはつながっていない可能性があることが指摘された<sup>61</sup>。2023 年 10 月に WIRED が行った調査では、Donald Trump Jr 氏のツイートに添付された動画が偽情報であるというメモの内容が誤りであったことを明らかにし、同機能自体が偽情報拡散を促す可能性があること<sup>62</sup>も指摘された。加えて、協力者ユーザー間の評価によって未公開となったメモの履歴や協力者ユーザーからのヒアリングからは、X のプラットフォーム上と同様に Community Note 協力者ユーザーの相互評価プロセスでも政治的立場の対立による分断および偏った情報拡散が起きていること<sup>63</sup>も指摘された。

## 1.7. 閲覧数に応じた広告収益分配プログラムの導入

2023 年 7 月 14 日、ユーザーは自分のツイートに表示される広告の閲覧数に応じて収益の分配を受けられるプログラムが導入された<sup>64</sup>。このプログラムには、有料の X Premium サブスクリプ

---

<sup>58</sup> <https://help.twitter.com/en/using-x/community-notes>

<sup>59</sup> Vittoria Elliott and Chris Stokel-Walker. “Twitter’s Moderation System Is in Tatters”. Wired. (2022.11.17). <https://www.wired.com/story/twitters-moderation-system-is-in-tatters/>

<sup>60</sup> <https://x.com/elonmusk/status/1590773586616545280> (2023.10.20 閲覧)

<sup>61</sup> Eric Fan, Rachael Dottle and Kurt Wagner. “Twitter’s Fact-Checking System Has a Major Blind Spot: Anything Divisive”. Bloomberg. (2022.12.19). <https://www.bloomberg.com/graphics/2022-twitter-birdwatch-community-notes-misinformation-politics/?leadSource=uverify%20wall>

<sup>62</sup> David Gilbert. “A Graphic Hamas Video Donald Trump Jr. Shared on X Is Actually Real, Research Confirms”. Wired. (2023.10.11). <https://www.wired.com/story/x-community-notes-failures/>

<sup>63</sup> Vittoria Elliott and David Gilbert. “Elon Musk’s Main Tool for Fighting Disinformation on X Is Making the Problem Worse, Insiders Claim”. Wired. (2023.10.17). <https://www.wired.com/story/x-community-notes-disinformation/>

<sup>64</sup> <https://x.com/X/status/1679572360695824384> (2023.10.20 閲覧)

ションに登録している認証済みユーザーのうち、過去 3 か月以内のツイートで累計 500 万回以上の閲覧回数を獲得しており、500 人以上のフォロワーを持つユーザーが参加できる<sup>65</sup>。

このプログラムの実装により、ユーザーからは「収益目的で閲覧回数を稼ぐことを目的とした、バズツイート<sup>66</sup>や著名人のツイートへの認証済みユーザーのリプライ、自動リプライ bot が増えた」という声も散見される<sup>67</sup>。広告収益プログラムの導入によるこのようなユーザー行動の変化について 2 章から検討する。

## 1.8. 生体認証情報と雇用情報の収集

2023 年 9 月 29 日に発効された X のプライバシーポリシーでは、ユーザーの生体認証情報、職歴や学歴など雇用にかかわる個人情報を収集する旨が新たに追加された（図 4）<sup>68</sup>。これは安全性の担保、セキュリティ、身元確認を目的とした収集であると記載されており、BBC の問い合わせに対しては「X Premium ユーザーを対象に、顔写真と国民識別番号を紐づけるオプションを提供する予定である」と述べている<sup>69</sup>。雇用情報に関しては、ユーザーへ仕事をマッチングするためのレコメンド広告表示、雇用主が人材を探すための情報提供が目的であると記載されている。なお X 社は 2023 年 8 月 26 日に、企業向けの採用支援機能である X Hiring のベータ版を公開している<sup>70</sup>。ただし、現時点では生体認証情報や雇用情報の取得は利用者の同意が必要なものとなっているため直ちにユーザーの自由意志を侵害するものではないことも指摘されている<sup>69</sup>。

X 社は 2023 年 7 月に、同社のデータ収集がイリノイ州の生体認証情報プライバシー法 (Biometric Information Privacy Act) に違反しているとして集団訴訟の被告となった<sup>71</sup>。この訴訟では、同社がプラットフォームにアップロードされた写真から生体識別子を収集・保存していたにもかかわらずそれをユーザーに適切に伝えていないことが論点とされた。

---

<sup>65</sup> “Ads Revenue Sharing”. <https://help.twitter.com/en/using-x/creator-ads-revenue-sharing>

<sup>66</sup> 多くのユーザーの注目が集まり、急激に拡散されたツイート。

<sup>67</sup> 例えば、お笑いタレントの有吉弘行氏は自身のツイートに閲覧数稼ぎ目的の bot によるリプライが現れていることに言及している。  
<https://x.com/ariyoshihiroiki/status/1702292551729066452> (2023.10.20 閲覧)

<sup>68</sup> <https://twitter.com/en/privacy#x-privacy-1> (2023.10.20 閲覧)

<sup>69</sup> Chris Vallance. “X, formerly Twitter, to collect biometric and employment data”. BBC News. (2023.9.1).  
<https://www.bbc.com/news/technology-66679922>

<sup>70</sup> <https://x.com/XHiring/status/1695154429828247798> (2023.10.20 閲覧)

<sup>71</sup> Emma Roth. “X wants permission to start collecting your biometric data and employment history”. The Verge. (2023.8.31).  
<https://www.theverge.com/2023/8/31/23853618/x-privacy-policy-update-biometrics-job-history>

雇用情報の取得についても、ワシントン大学のデータサイエンス実践の専門家である Liberty Vittert 教授は、職歴や学歴と紐づけられたツイートやリツイート、フォロワーのリストなどが雇用主に悪用、不当な意思決定への利用が行われるリスクを指摘した<sup>69</sup>。国内でも近年、就職活動を行う人材の SNS アカウントを特定するサービスが広まっている<sup>72</sup>。

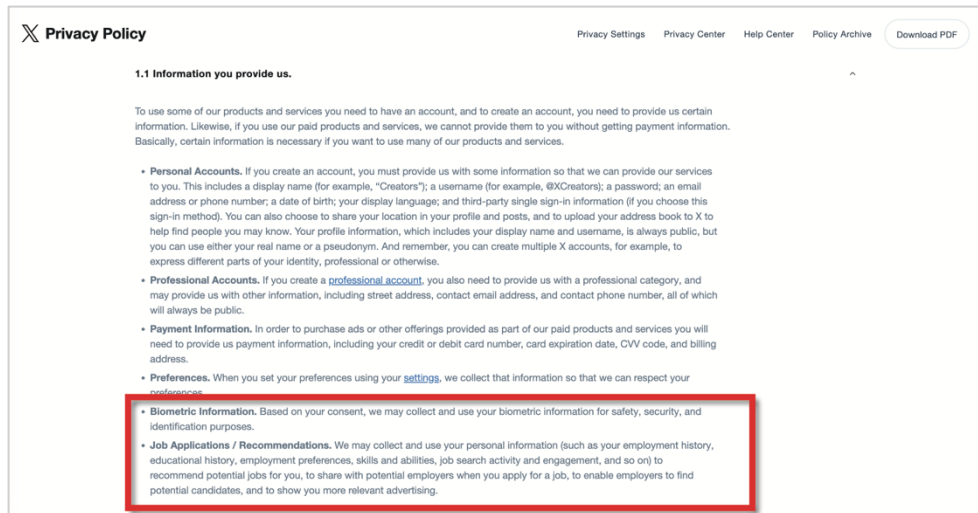


図 4 2023 年 9 月 29 日に発効された X のプライバシーポリシー

<sup>72</sup> 市原研吾. “「取り違えないといえるか」「社会の萎縮心配」裏アカ調査に識者は”. 朝日新聞デジタル. (2021.9.28). <https://www.asahi.com/articles/ASP9S1SWDP9RPTIL01M.html> (2023.10.20 閲覧)

## 1.9. 外部サイトの記事の見出し情報削除

2023年10月5日、外部サイトの記事へのリンクを掲載するツイートから見出しの表示を削除する仕様変更が実施された<sup>73</sup>。従来は外部サイトの記事のリンクをツイートした際にはその記事タイトルがカード状に自動的に埋め込まれていたが、変更後の仕様では、トップ画像のみが表示されて左下に小さく透かしてサイト名が表示される（図5）<sup>74</sup>。



図5 仕様変更後（左）と仕様変更前（右）の記事表示

この変更によってユーザーは、ツイートされた記事が実際にはどのような記事であるかがツイートからは判断できなくなった。仕様変更後に、例えばこの仕様変更に関するニュース記事に「Elon Musk氏が逮捕された」という偽の見出しを添えて投稿するユーザー<sup>75</sup>などが現れた。

Elon Musk氏は2023年8月に、ジャーナリストはX上に直接記事を公開するよう求めるツイート<sup>76</sup>を行っており、ニュースメディアサイトは競合相手であるとしている。同月にはNew York TimesやReutersなど特定のWebサイトのリンクへXからアクセスする際には5秒の遅延を加えていたことが判明し、それが報じられた後に遅延が修正された<sup>77</sup>。そのため、この仕様変更は、写真付き投稿と記事リンク付き投稿の区別をつきにくくすることで外部の報道機関などへのアク

<sup>73</sup> Siladitya Ray. “Borderline Useless’: X Removes Headlines On News Posts As Critics Say Move Changes Site’s Functionality”. Forbes. (2023.10.5). <https://www.forbes.com/sites/siladityaray/2023/10/05/x-hides-headlines-from-posts-after-musk-claims-it-will-greatly-improve-the-aesthetics/?sh=6e7d100b4c11>

<sup>74</sup> なお例に用いたツイートの表示はいずれも本稿執筆時点のものであり、外部URLに対して何らかの判定が行われたページに新しい仕様が適用されていると思われる。

<sup>75</sup> [https://x.com/Esqueer\\_/status/1709778449651384684](https://x.com/Esqueer_/status/1709778449651384684) (2023.10.20 閲覧)

<sup>76</sup> <https://x.com/elonmusk/status/1693756240260698550> (2023.10.20 閲覧)

<sup>77</sup> Jeremy B Merrill and Drew Herwell. “Borderline Useless’: X Removes Headlines On News Posts As Critics Say Move Changes Site’s Functionality”. The Washington Post. (2023.8.15). <https://www.washingtonpost.com/technology/2023/08/15/twitter-x-links-delayed/>

セスを阻害することを意図した可能性も指摘されている<sup>73</sup>。

## 1.10. ブロック機能の将来的な廃止

2023年8月1日、Elon Musk氏はブロック機能<sup>78</sup>を将来的に廃止する旨をツイートした<sup>79</sup>。ブロック機能はソーシャルメディア上での嫌がらせやストーカー行為などからユーザーを守る機能でもある。そのためApple社のApp StoreやGoogle社のGoogle Playなどアプリケーション配信プラットフォームはソーシャルメディアアプリへのブロック機能の実装を義務付けており、この変更が行われるとXのアプリケーションは各ストアの規約違反となる可能性も指摘された<sup>80</sup>。

Elon Musk氏が自身のツイートが積極的に Recommend されるよう Twitter 社のエンジニアに命じた際（1.4節）には、同氏のツイートの閲覧数が低くなったのはおすすめ表示によって同氏のツイートを大量に表示したことで多くのユーザーからブロックされたことも一因であったとエンジニアが指摘したこともリークされており<sup>81</sup>、ブロック機能の廃止は自身のツイートの閲覧数増加が目的である可能性も示唆されている。また Elon Musk氏は2023年9月28日にドイツの極右政党であるAfD（Alternative für Deutschland）を支持するツイートを引用リツイートで支持・拡散し、ドイツ外務省などからの批判を受けて反論を行うなどの言動を行っている<sup>82</sup>。Xをはじめソーシャルメディア上ではその政治的主張の是非が議論されているが、その主張がいかなる内容のものであったとしても、自身のツイートをユーザーに表示させるためにレコメンデーションアルゴリズムやサービス機能を変更できるステークホルダーが特定の立場に基づく発信を行なって強制的に表示すること自体が分断を増長する可能性もある。

---

<sup>78</sup> ユーザーが、特定のユーザーのツイートとのインタラクションを防ぐ機能。

<sup>79</sup> <https://x.com/elonmusk/status/1692558414105186796>（2023.10.20 閲覧）

<sup>80</sup> Faarea Masud. “Twitter, now X, to remove blocking feature – Musk”. BBC News (2023.8.18). <https://www.bbc.com/news/business-66550959>

<sup>81</sup> Georgia Wells, Alexa Corse and Kirsten Grind. “Inside Musk’s Twitter Transformation: Impulsive Decisions, Favors for Friends”.

THE WALL STREET JOURNAL. (2023.9.). <https://www.wsj.com/tech/elon-musk-twitter-x-transform-9627a8d5>

<sup>82</sup> Sophie Tanno. “Germany hits back at Elon Musk after he wades in on migrant debate”. CNN. (2023.9.30).

<https://edition.cnn.com/2023/09/30/europe/elon-musk-germany-migrant-debate-intl/index.html>

## 2. 現在のユーザーの行動に関する主成分分析

### 2.1. 論点と分析目的

#### 2.1.1. ネガティブ感情を誘発する発信が“バズリ”やすくなった可能性の検討

Xをはじめとするソーシャルメディア上では、中立的な投稿よりも感情的な投稿の方が拡散されやすく、またその感情は伝染してゆくことが長らく論じられている<sup>8384</sup>。そのため情報発信者がその発信内容で怒りや反感など激しい感情を煽ることは、インプレッション獲得や閲覧継続時間の増加、あるいは共感の獲得などの利益につながりうる。例えば2017年にFacebookは投稿へのリアクションを「いいね」ボタンのみでなくさまざまな感情を表すアイコンで行う機能を導入したが、当初はいいねボタンで評価された記事よりも人々が怒りを感じた投稿が重視される評価アルゴリズムが導入されていたことが報告されている<sup>85</sup>。ここでは、怒りを増幅させるアルゴリズムが悪用や誇大広告の増加といったリスクを持つ懸念から、2020年には怒りの評価の重み付けが引き下げられた。

このような背景を前提に、1章で整理したXのサービス動向のうち、特にレコメンデーションアルゴリズムに関する施策(1.3-1.4節)がネガティブ感情を誘発するツイートの拡散される機会を増加させ、また収益化プログラムの施行(1.6節)によってユーザーも意図的にそのような発信を行う利益が高まった可能性を検討する。

Milli et al. (2023) は「For You」タブの強制表示が施行された直後の2023年2月に、ランダムサンプリングされたユーザーの「For You」タブに表示されたコンテンツを収集・分析し、レコメンデーションアルゴリズムは強い感情、特に怒りを伴うツイートをレコメンデしている可能性を指摘した<sup>8687</sup>。加えて、レコメンデーションアルゴリズムはユーザーに見たいものを表示する以

---

<sup>83</sup> Stieglitz, Stefan, and Linh Dang-Xuan. "Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior." *Journal of management information systems* 29.4 (2013): 217-248. <https://doi.org/10.2753/MIS0742-1222290408>

<sup>84</sup> Ferrara, Emilio, and Zeyao Yang. "Measuring emotional contagion in social media." *PloS one* 10.11 (2015): e0142390. <https://doi.org/10.1371/journal.pone.0142390>

<sup>85</sup> Lexi Lonas. "Facebook formula gave anger five times weight of likes, documents show". THE HILL. (2021.10.26). <https://thehill.com/policy/technology/578548-facebook-formula-gave-anger-five-times-weight-of-likes-documents-show/>

<sup>86</sup> Milli, Smitha et al. "Twitter's Algorithm: Amplifying Anger, Animosity, and Affective Polarization." *ArXiv abs/2305.16941* (2023): n. pag. <https://doi.org/10.48550/arXiv.2305.16941>

<sup>87</sup> Mary-Ann Russon. "Elon Musk's Twitter algorithm changes are 'amplifying anger and animosity', say researchers". *The London Evening Standard*. (2023.5.30). <https://www.standard.co.uk/tech/elon-musk-twitter-algorithm-cyberbullying-discrimination-cornell-uc-berkeley-b1084490.html>

上は「そのような感情的なツイートを好んでいるユーザーがいるというだけだ」という主張に反して、ユーザーはアルゴリズムによって選択された感情的・政治的なツイートを「好んで」いなかったことも指摘している<sup>86</sup>。また、あくまで筆者の印象ではあるが、日本のユーザーにおいても、および日を追うごとに、X上で「対立を煽るかのような攻撃的・煽動的なツイートがおすすめ表示されることが多く疲れる」という声が見受けられるように思われる。

そこで本稿では、このようなXの「システムの変容」を受けて、ユーザーはどのようなツイートへ反応し、リツイートのような情報拡散行動やリプライのようなコミュニケーションはどのように変容したのか、といった「人々の変容」に着目する。例えば、本来はユーザーの気に入ってなかったネガティブツイートを強制的にユーザーの視界に入れさせ続けることで、そのようなツイートが次第に受容されやすく、“バズリ”やよくなっているなどのユーザー行動の変化が起きているならば、そのプラットフォームには倫理的・社会的な問題が指摘される可能性がある。

特にバズツイートへのリプライについては、X社の動向がもたらした可能性のある行動変容として、閲覧数を稼ぐことを目的としたリプライの増加が指摘されている（1.6節）。閲覧数に基づく収益、あるいは自身の発信するコンテンツの宣伝といった目的でリプライを行うアカウント（本稿ではビジネスアカウントと呼ぶ）が実際に増加しているのであれば、現在のレコメンデーションアルゴリズムではリプライの数が重視されるよう設計されていること、ビジネスアカウントもリプライで煽動的・攻撃的なツイートを発している可能性があることから、バズツイートへのリプライ全体に占めるビジネスアカウントの比率やその特徴にも着目する価値があると考えられる。

以上より本稿では、おすすめタブの強制表示が施行される直前と現在のバズツイートおよびバズツイートへのリプライを収集し、テキストに含まれる感情とユーザー同士のインタラクションの関係性を比較・分析する。ただし実際に行われたのは2.1.2節に示すX APIの仕様変更による制約下で可能な限りのデータ収集と分析であり、厳密に設計された分析でなく現在のソーシャルメディア空間に関する局所的なスナップショットの記録としての意味合いが強いことを付記する。

### 2.1.2. X APIの仕様変更がもたらすソーシャルメディア研究への影響の抽出

1.5節に概説したように、2023年3月31日よりX APIのプランと仕様に変更されたことでソーシャルメディア研究に用いられていたデータの取得に大幅な制約が加わっている。本稿もその制約を受けてデータ分析の設計として不十分な箇所が複数存在するが、その「X APIの制約によってもたらされた分析の不十分さ」自体にも着目する。

X APIの各プランの費用と取得可能な情報の対応を表1にまとめる。なお、データ取得に関する機能は例えば「Enterprise以外のプランでは直近7日以前のツイートは検索による一括取得ができない」など、件数以外にもプランごとにさまざまな制限が存在する（2.2節）。加えて本稿で

は議論しないが、データ取得には月間の件数上限のみでなく 15 分ごと・24 時間ごとのレートリミットも課せられており、上位のプランほど短期間に多くのデータ取得が可能である<sup>88</sup>。

	Free	Basic	Pro	Enterprise
費用/月	無料	100 ドル	5,000 ドル	42,000 ドル~
ツイート取得可能件数/月	不可	1 万件	100 万件	5,000 万件~
全期間でのデータ検索		×		○

表 1 各プランの費用とツイート情報の取得可能件数、機能

本稿では 1 ヶ月分の Basic プランを用いた。そのため 2.2 節では、合計 10,000 件のツイート情報と機能制限下で実施可能な分析が検討されていること、Enterprise プラン以外で使用できなくなった機能は多くの場面で求められるにもかかわらず Enterprise プランはおそらくどのような研究者にも利用できないこと（1.5 節）を前提として、本稿や先行研究の手法に X API の仕様変更や制限が与える影響を整理する。これによりプラットフォームの社会的影響の洞察が困難になった点を具体化する。同時に、現在の X API の仕様の範疇で可能な分析の検討、および分析過程で明らかとなった X API の仕様についても記録する。

## 2.2. 方法

本稿では、おすすめタブの強制表示が施行される直前と現在のバズツイート、およびバズツイートへのリプライの特徴を比較・分析する。まず、対象となるツイートおよびその詳細情報を収集する。次に、テキストの感情分析・リプライのユーザーコミュニティ分析を行い、各ツイートに特徴量を与える。最後に、それらの特徴量でベクトル化されたツイート群に主成分分析を行う。

各小段落の最後では、本稿の分析設計や先行研究の提案手法に X API の仕様変更や制限が与える影響を整理する。なお、X API の仕様は執筆時点のドキュメントや検証結果に基づいていること、

<sup>88</sup> <https://developer.twitter.com/ja/docs/twitter-api/rate-limits> (2023.10.20 閲覧)

データ取得数の制限に合わせて多くのデータ分析手法が簡易化されていることを再度注記する。

### 2.2.1. データセット

2023年9月25日から27日の3日間（期間1）、おすすめタイムラインの強制表示（1.3節）が施行される直前である2023年1月4日から10日の1週間（期間2）の、日本語で記述されたバズツイートのURLを収集した。本稿ではいいね数が10,000以上のツイートをバズツイートとして扱った。また、本稿ではテキストの特徴に着目するため、イラスト<sup>89</sup>や動画などメディアの含まれるバズツイートは対象外とした。対象となったバズツイートは期間1では253件、期間2では101件だった。なお、このツイート収集に関してはX APIでなくブラウザ版Xの検索機能を用いた手作業で行われており、両期間のデータ数の差異は検索機能のフィルタにもたらされたものである（機能制限の影響セクションを参照）。

このように取得されたツイート群について、X APIを用いて2023年10月1日時点の以下の詳細情報を取得した。

ツイートのテキスト・いいね数・リツイート数・引用リツイート数・閲覧数・リプライ数・リプライの許可範囲・投稿日時

ツイートをを行ったユーザーのフォロワー数・アカウント説明文・ユーザー名・ユーザーID<sup>90</sup>

次に、期間1のツイートへのリプライをX APIの検索機能を用いて取得した。まず、253件から著名人によるコンテンツ告知や日常的な内容に関するツイート<sup>91</sup>、リプライの許可範囲に制限を加えているツイートを目視で除いた。残った191件からツイートをランダムに取り出し、最新順に100件を上限としてリプライを取得した。これをX APIの月間データ取得数制限に至るまで繰り返し、バズツイートを投稿したユーザー自身のリプライを取り除いた。

これにより、131件のバズツイートに関する9,494件のリプライを取得した。含まれるユーザー

<sup>89</sup> なお、このフィルタリングを行わない場合、バズツイートのうちおよそ半数はイラスト作品に関するツイートであった。

<sup>90</sup> ユーザー登録時に自動的に設定される識別番号。2016年以前に登録したユーザーは登録順に割り振られた10桁の数値、それ以降に登録したユーザーには18桁の乱数が割り当てられている。

<sup>91</sup> このようなツイートは2つの期間で特徴が変わらず、またリプライ数が多くスパムや定型文の挨拶が大半を占めていた（2.3節に示す結果も参照）。そのため本稿ではデータ取得数の削減のために対象外とした。

数は9,164名であった。これらのリプライについても、X APIを用いてバズツイートと同様の詳細情報を取得した。

■ ツイート検索・収集の機能制限、ブラックボックス化されたフィルタ

ツイートの大規模な収集はさまざまな調査や研究・調査で行われており、その収集にも API は用いられてきた。X API を用いたツイートの情報取得は、特定のツイートを URL やツイート ID で直接指定する方法と、指定した条件に該当するツイートを検索して情報をリスト化する方法の2つがある。後者は大規模なデータ収集に不可欠である一方、利用プラン改定以降 Enterprise プラン以外に多くの制約が課せられている。

まず、X API の検索機能はさらに直近7日までのツイートの検索 (Recent Search) とサービス開始後全てのツイートの検索 (Full-archive search) に分けられており、Full-archive search は Enterprise プランでのみ利用できる。そのため、例えばトレンドやキーワードなど社会現象の事後的調査や評価、長期的な追跡などの分析<sup>92</sup>の実施が困難となった。

次に、X API やブラウザ・アプリ版 X のツイート検索機能では、キーワード以外にも検索条件を指定する検索クエリが提供されている。以下は期間2のツイート収集に用いた「いいね数10000以上、引用リツイート・メディア・リプライを除く、日本語」を指定する例である。

```
min_faves:10000 lang:ja -filter:media -filter:quote -filter:replies
since:2023-1-4 until:2023-1-10
```

しかし X API でツイート検索を行う場合、Enterprise 以外で使用できる検索クエリは“基本的な演算子 (the basic set of operators)”に限定されると記載されている。例えば本稿のデータ収集では、いいね数のフィルタリング (min\_faves:10000) を検索条件に加えると機能呼び出すことができなかった。つまり、Enterprise 以外のプランではいいね数・リツイート数などの条件設定を行うことができない。

これらの理由により、本稿ではブラウザ版 X の検索機能を用いて手作業で収集できる件数・

---

<sup>92</sup> Hino, Airo, and Robert A. Fahey. "Representing the Twittersphere: Archiving a representative sample of Twitter data under resource constraints." *International journal of information management* 48 (2019): 175-184. <https://doi.org/10.1016/j.ijinfomgt.2019.01.019>

なお、ここでも Twitter API (2019年当時) の価格や検索機能のフィルタがツイートのアーカイブ化・事後調査を困難にしている問題が提起されているが、この後2021年に学術研究プランが導入されることで当時は問題が緩和された。

条件のバズツイートを直接収集・指定し、また直近1週間の範囲に含まれる期間1のバズツイートへのリプライのみをX APIで収集した<sup>93</sup>。

また Enterprise プランでの API アクセス以外での検索は、Elon Musk 氏による買収以前の Twitter API Standard v1.1 の頃より「完全性ではなく関連性に焦点を当てており、一部のツイートやユーザーが検索結果に表示されない可能性がある」「完全な検索は Premium<sup>94</sup>もしくは Enterprise プランが必要」と述べられている<sup>95</sup>。そのため本稿ではブラウザ版 X の検索機能を用いてバズツイートを収集したが、条件に当てはまる全てのツイートを表示できなかった。例えば期間2では検索条件をいいね数5000以上のツイートと指定した場合にのみ表示されるいいね数10000以上のツイートなども確認され、期間2のバズツイートの多くは今回の検索条件では表示されなかったことが示唆された。このような、ブラックボックス化されたフィルタリングに基づくデータにしかアクセスできないことはソーシャルメディア研究に不可視のバイアスを与える問題も指摘されている<sup>92</sup>。

## 2.2.2. ツイートテキストの感情分析

各ツイートが攻撃的・煽動的な内容を含む可能性を特徴量化するために、収集したツイート群のテキストにポジティブ／ネガティブの感情極性分析（ポジネガ判定）を行った。分析には Python 用のオープンソースライブラリとして広く用いられる asari<sup>96</sup>を用いた。これにより、それぞれのツイートに0から1の範囲の値で表現されるネガティブスコアを与えた。この値は0に近いほどテキストがポジティブ、1に近いほどネガティブである確率が高いことを意味する。

### ■ ソーシャルメディアの感情分析とファインチューニング

ポジネガ判定に限らず、ソーシャルメディア上での喜びや怒りなど感情を分類する感情分析は、社会情勢の把握、プロダクトに対するユーザーの評判の分析など、さまざまな目的による

---

<sup>93</sup> 特定のツイートへのリプライは to:(ユーザー名) conversation\_id:(ツイート ID)の検索クエリで指定することが可能であり、この検索は Basic プランでも実行することができた。

<sup>94</sup> 当時存在した月額149ドルのプラン。

参考：Larry Dignan. “Twitter、有料の「Premium APIs」を発表、ベータ版をリリース”. CNET Japan. (2017.11.15).

<https://japan.cnet.com/article/35110434/>

<sup>95</sup> <https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/overview>

<sup>96</sup> <https://github.com/Hironsan/asari>

調査への応用が進んでいる<sup>9798</sup>。特にネット上での攻撃的発信が問題視されていることからその検出を目的とした研究も行われており、また独自の文章表現やスラングなども存在するため、感情分析や語句検知ライブラリのファインチューニングが求められることも多い<sup>99100</sup>。本稿でも攻撃的・煽動的な表現の評価を目的として `pymlask`<sup>101</sup>を用いた ML-ASK<sup>102</sup>による感情分析も試行したが、既存の日本語辞書とのパターンマッチングでは分類できないテキストも多かった。一方でファインチューニングを行うにはそれだけで大規模なデータ数が必要となり、例えば牧元ら（2022）による Twitter の攻撃的表現検出アルゴリズムの検討では、攻撃的表現を行うユーザーを同定してツイートをサンプリングする一連のプロセスで、少なくとも 3 万件のツイートを取得している<sup>99</sup>。そのため、ソーシャルメディアにおける感情分析や攻撃語検出などにも、X API のデータ取得数の制約が影響を与える可能性がある。

なお本稿の分析では結果的に、ネガティブスコアの高いテキストは定性的にはおおよそ批判的・攻撃的な内容が含まれるものであった。

### 2.2.3. プロフィール文のテキスト分類

リプライを行ったユーザーからビジネスアカウントをラベリングするために、自然言語処理によるテキスト分類を簡易化してプロフィール文の分類を行った。

まずトレーニングデータとして、データセットからビジネスアカウントとそうでないアカウントをそれぞれ 100 名ずつ目視でラベリングした。ここでビジネスアカウントの判断は、X Premium の認証ユーザーバッジを表示しており、かつ「役立ち情報」「コンサル」「有料記事」などそのユーザー自身のコンテンツ発信に関する特徴語を含む・フォロワー数 1000 人以上・bot やスパムアカウントでなく手動で運営していると思わしきアカウントであることを基準とした。

---

<sup>97</sup> 鳥海不二夫, 榊剛史, and 吉田光男. "ソーシャルメディアを用いた新型コロナ禍における感情変化の分析." 人工知能学会論文誌 35.4 (2020): F-K45\_1. <https://doi.org/10.1527/tjsai.F-K45>

<sup>98</sup> 安達由洋, et al. "感情語辞書を用いた日本語文の感情分析." 可視化情報学会誌 41.161 (2021): 21-27. [https://doi.org/10.3154/jvs.41.161\\_21](https://doi.org/10.3154/jvs.41.161_21)

<sup>99</sup> 牧元大悟 and 徳永健伸. "SNS 上の攻撃的表現の検出と位置特定" 言語処理学会 第 28 回年次大会 発表論文集 (2022). [https://www.anlp.jp/proceedings/annual\\_meeting/2022/pdf\\_dir/B8-4.pdf](https://www.anlp.jp/proceedings/annual_meeting/2022/pdf_dir/B8-4.pdf)

<sup>100</sup> 吉田基信, et al. "BERT を用いた SNS 上における攻撃的文章訂正システム." 第 84 回全国大会講演論文集 2022.1 (2022): 725-726. <http://id.nii.ac.jp/1001/00220992/>

<sup>101</sup> <https://github.com/ikegami-yukino/pymlask>

<sup>102</sup> Ptaszynski, Michal, et al. "A system for affect analysis of utterances in Japanese supported with web mining." Journal of Japan Society for Fuzzy Theory and Intelligent Informatics 21.2 (2009): 194-213. <https://doi.org/10.3156/jsoft.21.194>

次に、トレーニングデータのプロフィール文に MeCab<sup>103</sup>による形態素解析を行ったのち Word2Vec モデル<sup>104</sup>を使用してベクトル表現に変換し、この特徴量ベクトルについてのサポートベクターマシン (Support Vector Machine; SVM)<sup>105</sup>による分類器の学習・作成を行った。その後、同様に特徴量ベクトル化したテストデータのプロフィール文がビジネスアカウントであるかを分類器によってラベリングした。最後に、今回は分類器自体の十分な検証は行っていないため、定性的に誤ってビジネスアカウントと判別されたと思われるラベルを目視で修正した。

その結果、9,164 名のユーザーのうち 252 名がビジネスアカウントとラベリングされた。

#### ■ 認証済みユーザーの判別

X API のドキュメントには、ユーザーの認証情報を取得する機能について記載されている (図 6)<sup>106</sup>。しかし本稿の分析過程で試行したところ、いずれの種類の認証バッジ<sup>107</sup>が表示されたユーザー情報においても、非認証ユーザーを意味する false 値が取得されることが分かった。

アカウントの認証情報は、例えばコロナ禍での偽情報拡散の経路における個人ユーザーと認証済みユーザーの信頼性の比較や影響評価<sup>108</sup>など、ソーシャルメディア上で公的機関が果たす役割や責任を評価・洞察する際にも着目されることがある。

verified	boolean	Indicates if this user is a verified Twitter User.	Indicates whether or not this Twitter user has a verified account. A verified account lets people know that an account of public interest is authentic.
		<code>"verified": true</code>	

図 6 X API のドキュメントに記載された User 認証情報パラメータの説明

<sup>103</sup> <https://taku910.github.io/mecab/>

<sup>104</sup> Mikolov, Tomas, et al. "Distributed representations of words and phrases and their compositionality." *Advances in neural information processing systems* 26 (2013). <https://doi.org/10.48550/arXiv.1310.4546>

<sup>105</sup> Vapnik, Vladimir. *The nature of statistical learning theory*. Springer science & business media, (1999).

<sup>106</sup> <https://developer.twitter.com/en/docs/twitter-api/data-dictionary/object-model/user>

<sup>107</sup> <https://help.twitter.com/ja/rules-and-policies/profile-labels>

<sup>108</sup> Kouzy, Ramez, et al. "Coronavirus goes viral: quantifying the COVID-19 misinformation epidemic on Twitter." *Cureus* 12.3 (2020). <https://doi.org/10.7759/cureus.7255>

#### 2.2.4. ユーザーコミュニティの分析

ソーシャルメディア上ではフィルターバブルによるユーザーコミュニティの分断が発生していることが長らく指摘されている<sup>109</sup>が、新しく登場したと考えられるビジネスアカウントはトピックやコミュニティを問わずバズツイートにリプライを行うという独特な振る舞いをしていると考えられる。そこで、ネットワーククラスタリングの手法である Louvain 法<sup>110</sup>を用いてユーザーコミュニティを抽出し、各ユーザーが「どのくらい多様なコミュニティに所属しているか」を特徴量化した。

Louvain 法はノード（ここではユーザー）同士のエッジ（接続）の強さ・それにより構成されるコミュニティを検出するグラフクラスタリング手法であり、ソーシャルメディア研究の文脈では主にフォロー関係・リツイート・リプライによるユーザー間の接続が分析される。本稿ではリプライをエッジとして、2023年9月のバズツイートおよびそのツイートへのリプライを投稿したユーザーの間の関係性に基づくネットワークを構成した。その後、各ユーザーのエゴネットワーク（そのユーザーと直接つながったユーザーコミュニティからなるサブ・ネットワーク）にどの程度多様なコミュニティが存在するかを、情報エントロピーを用いて評価した<sup>111</sup>。情報エントロピーの値は大きいほどそのユーザーが多様なコミュニティに所属していることを示し、期間内に同じ回数バズツイートへのリプライを行ったユーザーであっても、同じコミュニティのトピックにのみリプライを行ったユーザーの値は小さく、トピックやコミュニティを問わずリプライを行っていたユーザーの値は大きくなる変数である。

#### ■ フォロー・フォロワー取得機能の削除、コミュニティ分析のデータ数

例えばリツイートは必ずしもそのツイートへの同意を意味しない<sup>97</sup>など、ユーザー間のインタラクションはアクションごとに異なる意味を持つ。そのためグラフクラスタリングによるコミュニティ分析では、フォロー・フォロワーの関係性に記述されるコミュニティ構造と論争的トピックの拡散経路を重ね合わせて比較するといったアプローチを行う先行研究も複数

---

<sup>109</sup> Garrett, R. Kelly. "Echo chambers online?: Politically motivated selective exposure among Internet news users." *Journal of computer-mediated communication* 14.2 (2009): 265-285. <https://doi.org/10.1111/j.1083-6101.2009.01440.x>

<sup>110</sup> Blondel, Vincent D., et al. "Fast unfolding of communities in large networks." *Journal of statistical mechanics: theory and experiment* 2008.10 (2008): P10008. <https://doi.org/10.48550/arXiv.0803.0476>

<sup>111</sup> 鳥海不二夫, and 榎剛史. "バースト現象におけるトピック分析." *情報処理学会論文誌* 58.6 (2017): 1287-1299. <http://id.nii.ac.jp/1001/00182259/>

ここでは、特定トピックがどの程度多様なコミュニティのユーザーに拡散されたかを情報エントロピーで評価する手法が提案されている。

存在する<sup>112,113</sup>。しかし 1.5 節で述べたように、2023 年 6 月に X API から各ユーザーのフォロー・フォロワー情報の取得機能が通知なく削除されて以降、このようなフォロー関係によるコミュニティ分析の手法にも Enterprise プランが求められるようになった。

また本来であれば Louvain 法などのグラフクラスタリング手法では、より広範かつ普遍的にユーザー同士の関係性を記述するために、単一のバズツイートとリプライのような局所的な関係性でなく、それぞれのユーザー自身の日常的なツイートと反応なども含めた行動を網羅的に取得されることが主流である。例えば上記の Cossard et al. (2020) の先行研究では 4,000 万件以上のツイートのデータが用いられている。しかし、その規模のデータ収集を行うためにも Enterprise プランが必要となる。

### 2.2.5. 主成分分析

バズツイートやそのリプライの性質を洞察するために、2.2.1 節から 2.2.4 節で得た特徴量群のベクトルについての主成分分析を行った。分析は「期間 1 のバズツイート」「期間 1 のバズツイートへのリプライ」「期間 2 のバズツイート」のデータ群それぞれに対して独立に行われた。

事前処理として「期間 1 のバズツイート」「期間 2 のバズツイート」を投稿したユーザーのフォロワー数分布の対数変換 (図 7)、いいね数が外れ値<sup>114</sup>にあたるツイートの除去を行った。主成

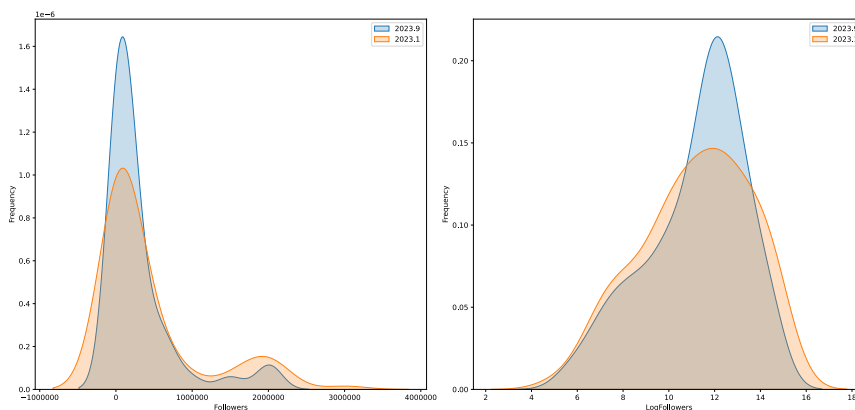


図 7 対数変換前 (左) と変換後 (右) の、バズツイート投稿者のフォロワー数の分布の概形

<sup>112</sup> Surian, Didi, et al. "Characterizing Twitter discussions about HPV vaccines using topic modeling and community detection." *Journal of medical Internet research* 18.8 (2016): e232. <https://doi.org/10.2196/jmir.6045>

<sup>113</sup> Cossard, Alessandro, et al. "Falling into the echo chamber: The italian vaccination debate on twitter." *Proceedings of the International AAAI conference on web and social media*. Vol. 14. 2020. <https://doi.org/10.1609/icwsm.v14i1.7285>

<sup>114</sup> 本稿では、データの分布を正規分布であるとした平均±2標準偏差の範囲外の値を外れ値とした。

分析は外れ値の影響を強く受ける<sup>115</sup>ことから、フォロワー数が100万人を超える著名人や特異的に社会的関心の集まったトピックのみによって主成分が定まることを避けるためにこれらの処理が行われた。

これらの処理を経て「期間1のバズツイート」「期間2のバズツイート」は7次元、「期間1のバズツイートへのリプライ」は8次元の、以下の特徴量を持つ特徴量ベクトルとして表現された。このツイート群に対して主成分分析を行い、それぞれ第1～7主成分（「期間1のバズツイートへのリプライ」では第8主成分まで）の寄与率と各特徴量の重み係数を導出した。

閲覧数・いいね数・リツイート数・引用リツイート数・リプライ数・フォロワー数・  
ネガティブスコア・エントロピー（「期間1のバズツイートへのリプライ」のみ）

---

<sup>115</sup> 主成分分析は特徴量空間内で分散が最大化される超平面を導出する手法であることから、例えば極端な外れ値が1つあった場合に「その外れ値1つ」と「それ以外のデータ群」を結ぶ超平面が引かれてしまうため。

## 2.3. 結果

### 2.3.1. 2023年1月と9月のバズツイートの主成分

図 8 に、期間 1（2023 年 9 月のバズツイート）の第 1～第 3 主成分空間での分布を示す。期間 1 のデータ群では、第 3 主成分までで分布の 68.0% が説明された。第 1 主成分には「閲覧数・いいね数・リツイート数・引用リツイート数は基本的に連動する」という（ある意味で自明な）特徴が示された。第 2 主成分は、フォロワー数とリプライの数が連動する方向の分散を示した。具体的に図 8 (A) は第 2 主成分軸上で特徴付けられる、フォロワー数とリプライ数が多いバズツイート群を示し、テキスト内容を確認すると著名な芸能人・動画配信者などの告知・日常的な挨拶ツイートなどが該当した<sup>116</sup>。第 3 主成分は「ツイート内容のネガティブスコアとリツイート数が連動する」という方向の分散を示した。加えてこの「ネガティブスコア・リツイート数」と「閲覧数・いいね数・フォロワー数」が、やや逆方向に連動していた。図 8 (B) は第 3 主成分軸上

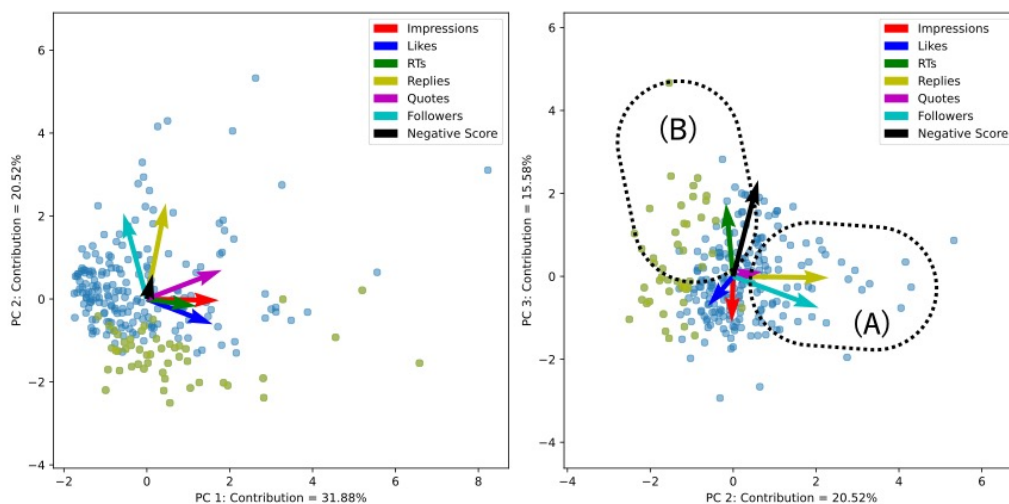


図 8 2023 年 9 月のバズツイートの第 1～第 3 主成分空間へのプロット。

黄色のプロットはフォロワー数が 10,000 人以下のユーザーによるツイートを示す。矢印の向きと長さは 7 次元空間から 2 次元の主成分空間に転写された各特徴量の基底ベクトルの向きと重みを示す。

<sup>116</sup> 例：[https://x.com/nino\\_honmono\\_j/status/1706428824685281496](https://x.com/nino_honmono_j/status/1706428824685281496) (2023.10.20 閲覧)

で特徴付けられる、リツイート数が多いネガティブツイートの中でもフォロワー数が相対的に少ないユーザーのバズツイートを示す。具体的なテキスト内容を確認すると、社会的話題に関する特定の立場からの批判が多数であった。その多くは次節 図 10 に示す政治的主張の対立に基づくものであったが、それに限らず例えば職場や家庭でのジェンダーバイアスなどに関する話題もあった。

図 9 には期間 2（2023 年 1 月のバズツイート）の第 1～第 3 主成分空間での分布を示す。期間 2 のデータ群では、第 3 主成分までで分布の 73.4 %が説明された。第 1 主成分、第 2 主成分の「閲覧数・いいね数・リツイート数・引用リツイート数は連動する」「フォロワー数が多い著名人の、リプライが多数ついたバズツイート群が存在する」という特徴は期間 1 と共通した。第 3 主成分は期間 1 と同様にネガティブスコアを主軸とした主成分軸だが、期間 2 ではネガティブスコアと他の特徴量のベクトルがほとんど直交した（図 9 右）。つまり、期間 1 と異なりネガティブスコアは他の特徴量とは連動せず、第 3 主成分は「バズツイートにはポジティブなテキストのものもネガティブなものもある」ということだけを意味した。また、期間 2 のネガティブツイートにも、政治的主張の対立に基づく批判が多く見られた。

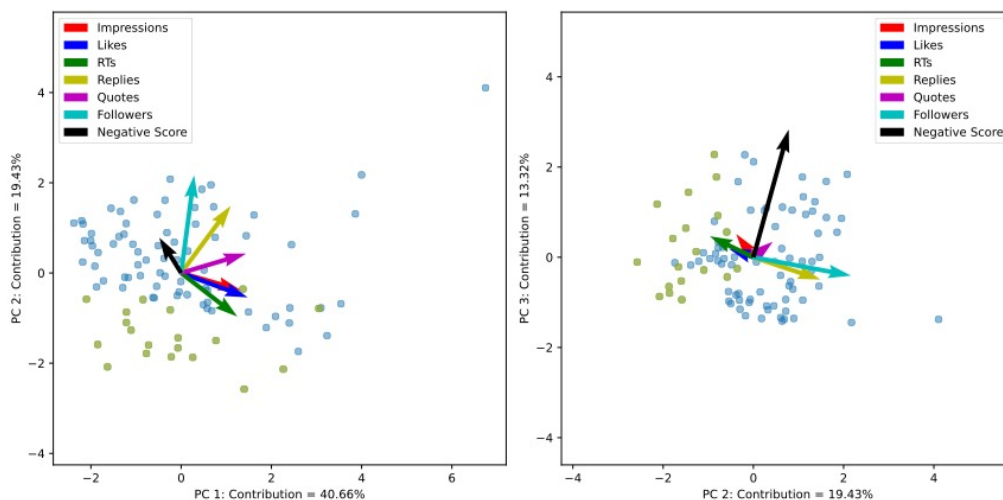


図 9 2023 年 1 月のバズツイートの第 1～第 3 主成分空間へのプロット。

黄色のプロットはフォロワー数が 10,000 人以下のユーザーによるツイートを示す。矢印の向きと長さは 7 次元空間から 2 次元の主成分空間に転写された各特徴量の基底ベクトルの向きと重みを示す。

### 2.3.2. バズツイートへのリプライ

図 10 に、期間1のバズツイート 131 件とそれらに対する 9,494 件のリプライを行ったユーザーのネットワーク、および目視で確認した各コミュニティのトピックの概要を簡易的に図示した。うち 237 名のユーザーが複数のバズツイートへリプライを行っており、同じ回数のリプライをしたユーザー同士でも同じコミュニティ内で多くリプライをしているユーザー、様々なコミュニティにリプライをしているユーザーが存在した。また、政治的立場の異なるコミュニティの分断に加えて、半数以上のユーザーが政治的話題を扱うコミュニティ群自体との接続が弱いことが記述された。

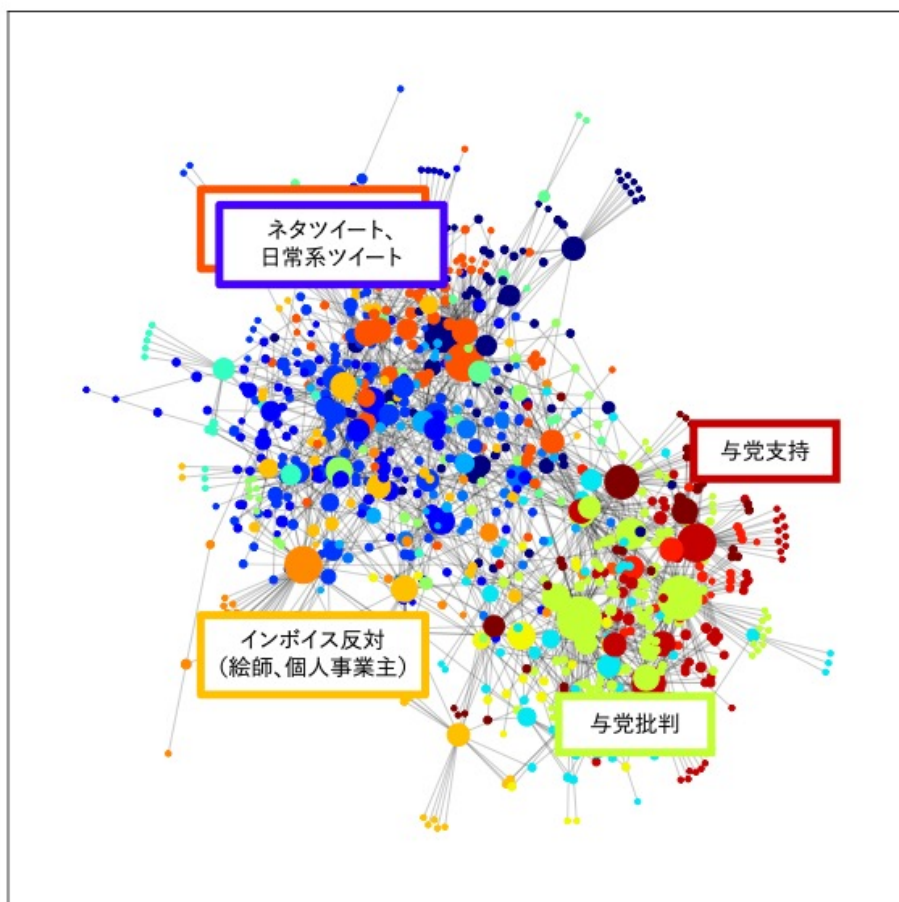


図 10 2023 年 9 月のバズツイートとリプライにより記述されたネットワークの模式図

図 11 には期間1のバズツイートへのリプライの第1～第3主成分空間での分布を示す。ここでは第3主成分までで分布の 58.4 %が説明された。第1主成分にはバズツイートと同様に「閲覧数・いいね数・リツイート数は連動する」という、話題性の多寡を示す方向の分散が示された。

ただしバズツイートの場合と異なり、引用リツイート数はこれらの特徴量との連動が弱い、つまりリプライの注目度には引用リツイート数はあまり寄与しないことが明らかとなった。第2主成分は、フォロワー数・エントロピーの高さ・テキストのポジティブさが連動する方向の分散を示した。第3主成分はネガティブスコアと引用リツイート数が連動する方向の分散を示した。

またビジネスアカウントの多くが、第2主成分軸上で特徴付けられる「エントロピーが高い・フォロワー数が多い・ネガティブスコアが低い」というリプライ群に位置付けられた(図 11(C))。該当するリプライを確認すると、この群は相対的に閲覧数が多いこと、テキスト内容には(例えば「なるほど」というような)バズツイートの内容への共感を表明するリプライが多く見られることがわかった。また、第2主成分軸の逆側に位置する「エントロピーが低くネガティブスコアの高いリプライ」の多くは、図 10 の政治的立場に基づくコミュニティの中で複数のリプライを行ったユーザーだった。

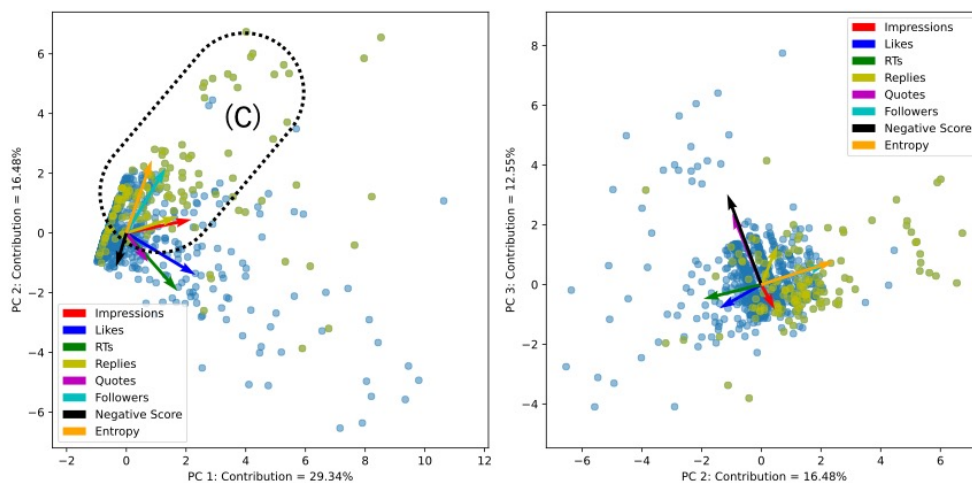


図 11 2023年9月のバズツイートに対するリプライの第1～第3主成分空間へのプロット。黄色のプロットはビジネスアカウントとラベリングされたユーザーを示す。矢印はそれぞれの主成分空間に転写された特徴量の基底ベクトルを示す。矢印の向きと長さは8次元空間から2次元の主成分空間に転写された各特徴量の基底ベクトルの向きと重みを示す。

## 3. X のサービス動向がもたらす可能性のある ELSI

### 3.1. 煽動的バズツイートの拡散起点の多様化による分断の加速

おすすめタブの強制表示実装前と現在のバズツイートいずれにも、社会的話題への政治的立場に基づく批判をネガティブな感情とともに発信するツイートが見られた。2つの期間でのこのようなツイートの違いとして、現在はバズツイートにネガティブ感情が伴うこととリツイート数が増えることが連動する傾向のバリエーションが見られた（図 8）。また、そのような「同程度のいいね数や閲覧数のバズツイートと比べてリツイート数が多い社会的批判」の投稿者の多くはフォロワーが（相対的に）少ないユーザーであり、おすすめタブ強制表示の実装前は、フォロワー数が 10,000 人以下のユーザーによるそのようなバズツイートはあまり見られなかった。

加えて、2.3.2 節の結果でリプライを通して記述されたネットワークでは、政治的立場の異なるコミュニティ同士の分断に加えて、半数以上のユーザーが政治的話題自体から距離を置いていることが記述された（図 10）。メディア環境の発展が触れる情報の選択の自由をもたらしたことでエンターテイメント志向の強い人々が政治的話題そのものから離れていること<sup>117</sup>は以前から指摘されている。レコメンド機能はユーザーが受け入れる情報の多様化にはつながらず、ソーシャルメディア上の分断は引き続き存在していると考えられる。

以上より、上述のようなネガティブ感情のバズツイートは、従来は著名人や「特定立場のコミュニティのインフルエンサー」などによる発信が主流であったが、現在は相対的に、ユーザー自身の影響力にかかわらず“バズ”機会が増加した可能性が考えられる。そのような拡散機会の増加がネガティブツイートのみで見られた<sup>118</sup>ことは、Milli et al. (2023)<sup>86</sup>の「おすすめタブには、ユーザーの好みを問わず怒りの感情を誘発するツイートがレコメンドされやすい」というレコメンドの特性に関する先行研究、ソーシャルメディア上では中立的な投稿よりも怒りなどの感情的な投稿の方が拡散されやすいという人々の情報行動の特徴（2.1.1 節）にも一致する。総じて、おすすめタブによって以前より多様なユーザーのネガティブ感情がレコメンドされるようになったことで、フォロワー数や影響力を問わず多様なユーザーの声が人々のネガティブな感情を誘発してリツイートされる機会、ひいてはバズツイートとなる機会が増加した可能性が示唆される。

---

<sup>117</sup> 小林哲郎, "ソーシャルメディアと分断化する社会的リアリティ < 特集 > Twitter とソーシャルメディア," 人工知能 27.1 (2012): 51-58. [https://doi.org/10.11517/jjsai.27.1\\_51](https://doi.org/10.11517/jjsai.27.1_51)

<sup>118</sup> もしポジティブツイートでも同程度に、フォロワーが少ないユーザーのツイートがリツイートされてバズる機会が増加しているならば、ネガティブスコアは他の特徴量と連動しない（視覚的には図 8 のネガティブスコアのベクトルが他の特徴量と直交する）。

おすすめタブの強制表示は、当初はユーザーからは否定的な声上がり（1.3節）、また一定数のユーザーはそこに表示されるネガティブな感情や対立を煽るようなテキストにも忌避感を持っている（2.1.1節）。しかし、それにもかかわらずおすすめタブによってネガティブな感情の拡散機会が増加したという考察が正しいならば、強制的にレコメンドされたツイートを「ひとまず視界に入れさせる」という X の施策は、自覚的・無自覚的を問わずユーザーに徐々に「フィルターバブルの居心地の良さ」を植え付けている可能性が示唆される。ソーシャルメディア上での分断がもたらす偽情報拡散や特定立場の過激派の増加などについて、ユーザー自身が客観的な情報を得よう心がけるなどのいわゆるメディアリテラシーは問題の根幹ではなく、その情報の真偽や主観性を「わかっていて」として自身に居心地の良さを与えてくれる構造に誰もが飲み込まれうることが問題であるという見方もある。だからこそ、例えば TikTok ではユーザーの「想定通り」ではない新規のコンテンツを表示する倫理的な義務があるのではないかという指摘<sup>119</sup>や、米国での表示最適化のオンオフ切り替え機能提供を義務付けるフィルターバブル透明性法案の提出（1.3節）など、レコメンド機能を提供するプラットフォーム事業者の倫理的・法的・社会的な責任についての議論も行われている。X の施策が人々に感情的なツイートを拡散する心地良さを広めているという解釈を正しいとするならば、それはソーシャルメディア上の分断をさらに加速させるという点で問題がある。

ただし、ソーシャルメディア上でのトレンド拡散の性質はそのトピックにも影響される。期間 1 のバズツイートにはインボイス制度に関する話題が含まれており、インボイス制度への反対運動に関する話題を通してエンターテイメント志向のクリエイターコミュニティと与党批判の立場のコミュニティにゆるやかな接続が記述された（図 10）。これは話題自体がクリエイターコミュニティに当事者性が高く、政治的話題に普段触れないユーザーにも拡散されやすかったことも影響した可能性がある。加えて本稿では直近 7 日以前のデータには大幅なフィルタが加えられているおり、データ件数にも 2 期間で大きな差がある。そのため、おすすめタブの強制表示などによって批判的バズツイート全般が拡散されやすくなったかについては慎重に検討する必要がある。もちろん、それは長期間のツイート、バズツイート以外のツイートも収集・分析することで議論できるが、2 章で述べた X API の制限により、そのようなデータ収集は本稿では（あるいは多くの研究者にとって）困難となっている。

---

<sup>119</sup> Golden, Paige, and David Danks. "Ethical obligations to provide novelty." Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society. 2021. <https://doi.org/10.1145/3461702.3462555>

### 3.2. 収益獲得や発信活動を目的とするユーザーの行動統制

2.3.2 節の結果からは、閲覧数獲得を目的としてバズツイートへリプライを行うアカウントは確かに存在しており、それらのアカウントはコミュニティを横断してトピックを問わずリプライを行うこと、リプライを行うユーザーの中では相対的に多くのフォロワーを抱えていること、bot やスパムアカウントでなく「人間らしい<sup>120</sup>」リプライを行うアカウントも多数存在していることが確認された。一方、当初の「ビジネスアカウントもリアクション獲得を目的として煽動的・攻撃的なツイートをする傾向があるのではないか」という想像と反して、むしろビジネスアカウントはポジティブなテキストによる「元のバズツイートへの共感」をリプライする傾向が強いことがわかった。

もちろん、攻撃的なコミュニケーションよりも共感的なコミュニケーションが行われることは歓迎されることである。一方で、このように収益を動機としてリプライを行うようになったユーザーは必ずしも真の共感を示しているとは限らない。1.4 節でも指摘されたように、収益や発信といった目的を持つユーザーはプラットフォームに「気に入られる」行動をせざるを得ないこと、それが場合によってはユーザーの行動の偏り、プラットフォーム事業者による統制・制御をもたらす可能性があることは留意する必要がある。

ただし、このようなリプライを行うアカウントが収益化プログラムの施行によって「新しく」発生あるいは増加したことを正当な統計的手続きに基づいて指摘するには当然、現在のリプライのみの分析ではなく収益化プログラム施行前との比較が必要である。そしてそれが本稿で行われなかったのは、2 章に述べた X API の制限によって過去のデータ取得が困難となったためである。

### 3.3. プラットフォームの社会的影響や責任の不透明化

2.2 節の各段落で整理したように、多くの研究者が現実的に利用できる X API のプランでは、扱えるデータ量の不足の問題はもちろん、事後分析や長期的な比較分析を行うための過去のデータ、情報拡散経路の分析に寄与するフォロワー・フォロワー関係のデータにアクセスできなくなった。

ソーシャルメディア上での人々の振る舞いや情報拡散経路、特に機械学習に基づくレコメンデーション機能の入出力関係やそれがもたらす影響は、システムの設計者であってもそのアルゴリズムのみから理解することは困難である（1.4 節）。そのため、実際のデータを用いてそれらを分

---

<sup>120</sup> バズツイートへのリプライを行う bot は、他のユーザーのリプライや引用リツイートの転載や、生成 AI による自動生成文などをリプライする傾向がある。

析・洞察するソーシャルメディア研究の取り組みは、ある意味で「公益性の観点でも適切なアルゴリズム、プラットフォームであるか」を評価・監査する役割がある。ここまでで指摘したようにレコメンドアルゴリズムの社会的責任についても倫理的・法的・社会的観点からの注目は高まっており、ユーザーがフィルターバブルに閉じ込められることを可能な限り回避し、ソーシャルメディア空間での分断をいかに緩和してゆくかがプラットフォーム事業者に問われている。

Hickey et al. (2023) は Elon Musk 氏の就任直前後の 2022 年 10 月 1 日から 11 月 29 日のツイートを（当時は利用可能であった学術研究用 API で）分析し、同氏の施策によってヘイトスピーチと bot が増えた可能性を指摘した<sup>9</sup>。加えて、Elon Musk 氏が COVID-19 の誤情報規制ポリシーの削除<sup>121</sup>、ヘイトスピーチ・自殺・児童搾取などの問題に取り組む同社の委員会である Trust and Safety Council の解散<sup>21</sup>などを行ったにもかかわらず、2022 年 11 月 23 日に「買収前よりもヘイトスピーチが低いレベルにまで低下した」と主張したこと、そのヘイトスピーチの調査方法については明らかにしなかったこと<sup>122</sup>も言及された。

今回の分析結果でも、現在の X はユーザーのネガティブな感情を誘発させ、分断をますます加速させる可能性があることが示唆された。しかしデータ取得の量や期間の制限により、その分析プロトコルや結果には十分な妥当性を担保できなかった。ここで、もし例えば Hickey et al. (2023) の先行研究と同様に、X 社が調査方法を明かさずに「X 上でフィルターバブルは改善されている」と主張した場合、少なくともその妥当性の検証や比較は行えず、そのまま X 社にとっての「居心地の良いフィルターバブル」で X 社の主張を受け入れてもらうことも可能かもしれない。

以上より、X 社が X API に課した制限は、人々の情報行動への影響や社会的責任を明確化させず「倫理的に適切なプラットフォームであるか自体を問わせない」という倫理的問題を引き起こしていると考えられる。

---

<sup>121</sup> Donie O'Sullivan. "Twitter is no longer enforcing its Covid misinformation policy". Associated Press News. (2022.11.29). <https://edition.cnn.com/2022/11/29/tech/twitter-covid-misinformation-policy/index.html>

<sup>122</sup> Sheera Frenkel and Kate Conger. "Hate Speech's Rise on Twitter Is Unprecedented, Researchers Find". The New York Times. (2022.12.2). <https://www.nytimes.com/2022/12/02/technology/twitter-hate-speech.html>

## 4. おわりに

本稿ではX社のサービス動向をELSIの観点から概観し、特にレコメンド機能や収益化プログラムがユーザーの行動にどのような影響をもたらすかに着目した。現在のXでは批判的な感情を誘発させられる機会がこれまで以上に増えた可能性があり、もしそれが正しければフィルターバブルのさらなる加速につながるリスクがある。加えて、これまでのソーシャルメディア研究の方法によってプラットフォームの社会的影響を評価すること自体がX APIの制限によって困難となっており、プラットフォーム事業者の社会的責任が不透明になってゆく可能性がある。

産学官のステークホルダーは、現代社会で影響力を持つソーシャルメディアの在り方を問い、事業者の営利と社会の公益性が共存可能な監査体制や法整備などを推進してゆくことが望まれる。最後に、ユーザー個人としての我々がいかにソーシャルメディアと付き合うことができるかについて、本稿で述べたXのサービスがもたらしうるELSIを踏まえて私見を述べる。

強制的におすすめタブを見せられることで誰もがフィルターバブルの「居心地の良さ」に飲み込まれうるという指摘（3.1節）が同意されるとしても、その課題解決を「ソーシャルメディアから距離を置けばいい」とユーザーの自己責任に押しつけることは容易い。しかし、例えば基地周辺などに自ら住む人々の意思決定は、転居による経済的・精神的負担、あるいは地域コミュニティへの愛着などの多層的な価値判断と制約の中で折り合いを付けて選ばれたものであり、それを「自らの意思で住んでいるのに不満を述べるのはおかしい」と批判することの持つ“自己責任論の暴力性”も生活史研究の中では指摘される<sup>123</sup>。つまり、ソーシャルメディアという生活圏においても、それぞれの事情によって望む・望まないに関わらず「ソーシャルメディアを辞められない」という状況も存在する。例えば1.1節で述べたFacebook社の「いいね」の数値の非表示機能実装時には、インフルエンサーとして生計を立てるユーザーのセルフプロモーションへの影響も同時に指摘された<sup>16</sup>。Xにおいても、例えばフリーランスとして活動しているイラストレーターの人々の「Xは生活のための仕事の基盤になっているので、アルゴリズムの都合でインプレッション数が下がったり、Xの都合で振り回されたりして困っているが辞めることもできない」といったツイートを目にすることもある。

また、仮に個人がソーシャルメディアの利用を辞めたとしても、ソーシャルメディア上の価値観で育った他者と現実世界で向き合う場面も今後ますます増えてゆく。筆者においても、ソーシャ

---

<sup>123</sup> 岸政彦、石岡丈昇、丸岡里美。“質的社会調査の方法”. 有斐閣ストゥディア. p.235-236. (2016).

ルメディアの影響で広まったと時折語られる「論破ゲーム」の価値観に根差した言動<sup>124</sup>を行う児童に科学教育の場で向き合った経験がある。

総じて、ソーシャルメディアの影響からはユーザー自らの行動で逃れることはできず、また個人がそのマクロな影響と対抗することも困難である。しかしそれでも「自分や他者はソーシャルメディアのシステムに何を思わされているのか・何をさせられているのか」を自覚的に捉えることは、例えば気づかないうちに感情を誘発されて攻撃や分断に加担してしまうことを防ぐ一助にはなると考える。「自ら主体的に、客観的な情報や多様な意見に触れましょう」という標語的なメディアリテラシーの啓蒙よりは、それを分かっていたとしてもできないかもしれない自己や他者を一旦認めることが、“リテラシーの無い異質な批判対象”同士の分断を再解釈する、少なくとも無自覚に攻撃し合うことに歯止めをかける役に立つのではないだろうか。生活者としての我々は、身を置いている構造をメタに捉えることから、ソーシャルメディアに飲み込まれず付き合ってゆく方法を考えてゆけるかもしれない。

---

<sup>124</sup> 岡田有花, “ひろゆき氏『『はい論破』言ったことない』『それってあなたの感想ですよねは1回きり』なぜ誤解広がる?”, Yahoo!ニュース エキスパート, (2023.9.10), <https://news.yahoo.co.jp/expert/articles/a7a22cf8faa4211c882d5d319169ac05cd8398af> (2023.10.20 閲覧)

**ELSI NOTE No. 33**

令和5年11月8日

**X (旧：Twitter) のバズツイートの変化から見るプラットフォーム事業者の施策がもたらす ELSI**

若林 魁人

大阪大学 社会技術共創研究センター 特任研究員 (2023年11月現在)

**Examination of the ELSI of X (former Twitter) platform providers focused on buzz tweets**

Kaito Wakabayashi

Research Center on Ethical, Legal and Social Issues, Osaka University



大阪大学 社会技術共創研究センター  
Research Center on Ethical, Legal and Social Issues

〒565-0871 大阪府吹田市山田丘 2-8  
大阪大学吹田キャンパステクノアライアンス C 棟 6 階  
TEL 06-6105-6084  
<https://elsi.osaka-u.ac.jp>

