2023. 10. No.22

GLOCOM Discussion Paper Series 23-001

Center for Global Communications, International University of Japan

日本におけるコミュニティノートの受容
―成功か失敗か、偏りはあるのか

田中辰雄
横浜商科大学 教授 / 国際大学 GLOCOM 主幹研究員

GLOCOM
国際大学グローバル・コミュニケーション・センター
https://www.glocom.ac.jp/

2023 年 10 月 13 日発行（No.22, 23-001）
発行人　松山良一
編集長　山口真一
編集委員　渡辺智暁、豊福晋平　櫻井美穂子　小林奈穂　菊地映輝　逢坂裕起子
編集　安藤久美子　武田友希
発行所　国際大学グローバル・コミュニケーション・センター
〒106-0032　東京都港区六本木 6-15-21　ハークス六本木ビル 2 階
Tel：03-5411-6677　FAX：03-5412-7111
URL：https://www.glocom.ac.jp/

本論文は著者の見解に基づくものであり、国際大学グローバル・コミュニケーション・センターとしての公式見解を示すものではありません。

GLOCOM Discussion Paper Series 23-001
2023. 10.

日本におけるコミュニティノートの受容
—成功か失敗か、偏りはあるのか

田中辰雄（横浜商科大学 教授／国際大学 GLOCOM 主幹研究員）

要旨

コミュニティノートは、誤情報の修正を、専門的なファクトチェック機関ではなく、無名の人の多数の人の力で行う仕組みである（Crowd-sourced misinformation management）。日本で導入されて 1 か月後の時点でのユーザへのサーベイでは、この仕組みはユーザの支持を得ている。「役にたつ」、「誤情報を知ることができる」、「コミュニティノートでツイッターは良くなった」と思う人がそう思わない人より圧倒的に多いからである。リベラル論客の中にはコミュニティノートはリベラル側のツイートにばかり付くと批判する向きがある。確かにその傾向はあるが、その原因としては保守側が積極的にノートを使って書き込みをしている可能性だけでなく、そもそもリベラルのツイートが異議申し立て型なのでノートが付きやすい可能性があり、原因ははっきりしない。いずれにせよサーベイ回答者をリベラル的な人だけに限っても、コミュニティノートの導入でツイッターは良くなったと答える人が多いため、肯定的な評価は動かない。このように肯定的な評価が優勢な理由は、ノートを書く投稿者たちが、ノートには意見を書くのではなく重要な関連事実だけを書くという原則をいまのところ守っているからと考えられる。

キーワード

ツイッター、X、フェイクニュース、ファクトチェック、SNS、クラウド

ツイッターにコミュニティノートが追加された。コミュニティノートとは誤情報を修正するための工夫であり、その特徴は専門家や運営者が行うのではなく、ツイッターユーザ自らが行う点にある。

誤情報の修正は、これまではファクトチェック機関や SNS のプラットフォーマーがラベルを貼るなどして対応してきた。しかし、そのような専門家による修正には、人材や費用面の制約からカバーできる情報が少ないという問題点があり、さらにファクトチェック機関とプラットフォーマー自体が信用されなくなってしまうと効果が失われる。[1]

このような限界を受けて登場してきたのが無名の一般ユーザ（群衆 crowd）による誤情報修正の仕組み（Crowd-sourced misinformation management）である。ツイッター社によるコミュニティノート（導入時の名前は Birdwatch）はその代表的なもので、2021 年 1 月から試験がはじまった。日本では 2 年遅れて試験導入が行われ、2023 年 7 月から本格稼働した。

この機能には賛否両論があり、役立っているという評もあれば偏りがあるという批判も見られる。今回、ユーザへのアンケート調査でこの機能の実態と評価を調べたので報告する。結論から言うと、多少の偏りはあるものの、コミュニティノートをおおむね肯定的に評価する人が多い。ここまでのところ日本ではコミュニティノートは人々の支持を得ており、成功しているようである。

１．コミュニティノートとは

コミュニティノートについて最初に簡単に説明する。コミュニティノートとは、ツイートに対して他のユーザがつける注釈である。図 1 で「閲覧したユーザが背景情報を追加しました」とあるのがそれである。

図 1 (a) は「マイナンバーカード本人希望で廃止、4 割近くが自主返納」というニュースについたノートである。このノートは、4 割というのは本人希望で廃止した 250 件のうち自主返納だった分が 4 割ということであり、マイナンバー保有者 8 千万人の 4 割が自主返納したわけではないことを指摘している。ノートの趣旨は、ニュース記事の見出しがミスリードになる恐れがあることを指摘することである。

[1] アメリカでは共和党支持者の 70%はファクトチェック機関が偏っていると考えており（Walker and Gottfried, 2019）、それら機関がニュースに信用できないというラベルをつけても共和党支持者はその影響をあまり受けないとされる（Straub and Spradling, 2022）。

図 1 (a)

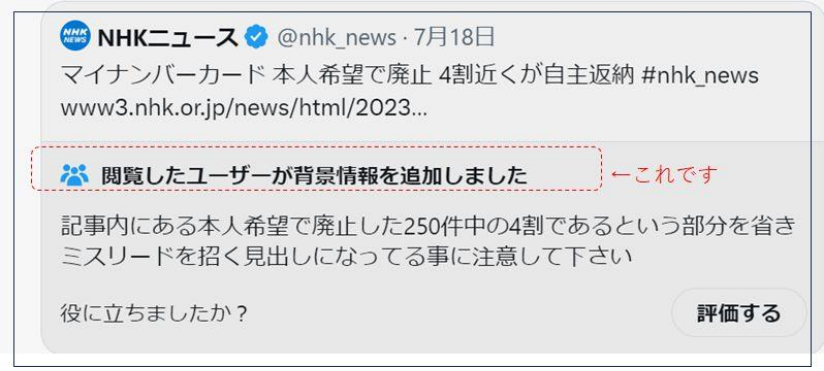
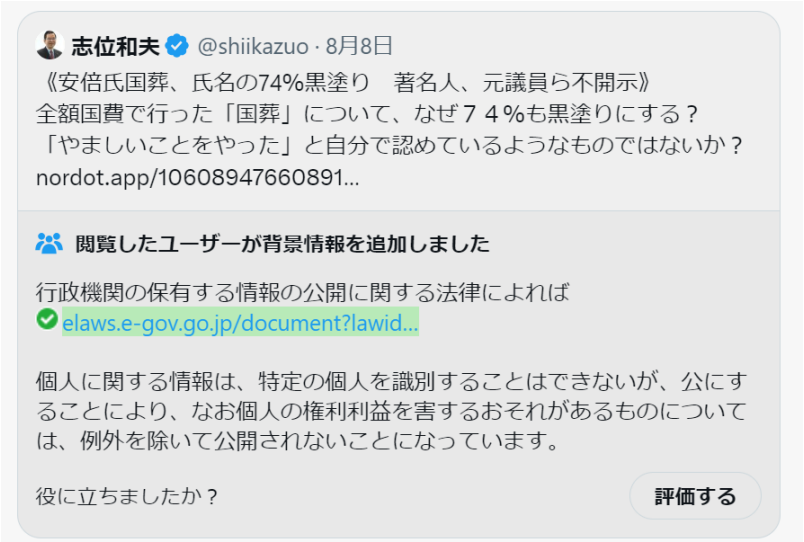


図 1 (b) は、安倍元首相の葬儀の参列者名簿の情報開示を求めたところ、74%が黒塗りだったことを批判するツイートに付いたノートである。ノートは、名簿の黒塗りは情報公開法で個人が特定できる情報は開示してはならないとされているからで、法律に沿った措置であるとしている。

図 1 (b)



この二つの例でわかるように、コミュニティノートとはファクトチェックとはやや趣が異なる。あからさまなフェイクニュースをフェイクとする例は少なく、多くはミスリードを正したり、重要な判断材料を提供したりするものである。特に判断材料を提供するタイプ、すなわち、このツイートでは A を主張しているが、それに関連して触れられていない B という重要な事実がありますよ、と指摘することが多い。

この二つのなかでは図 1 (b) がそれで、国葬の参列者名簿の黒塗りの是非を考えるための判断材料を提示している。直接に黒塗りの是非を述べているわけではない。このノートを見ても、なお葬儀出席者の黒塗りは問題であり公開すべきと主張することは十分可能であ

る。しかし、このノートがあることで、安倍元首相の威光を恐れ、あるいはどこかからの圧力があって黒塗りが行われたわけではなく、情報公開法の法令にしたがった措置であることはわかる。ノートがない場合、行政機関が政治家の圧力に屈した、あるいは忖度したという誤解に基づく不信が生まれる恐れがある。黒塗りに反対するなら情報公開法を覆しても公開すべきと主張する覚悟が必要であり、このことを伝えるのがコミュニティノートの目的である。

コミュニティノートはどうやって付けられるのだろうか。これには二つの段階がある[2]。まず、コミュニティノートを書く「投稿者（contributor）」が存在する。投稿者は過去にツイッターの規則に違反しておらず、ツイッターアカウント作成から 6 か月以上たっていれば、電話番号認証をして登録すれば誰でもなることができる。

投稿者になると、任意のツイートにノートを書くことができる。ただし、書かれたノートはすぐには公開されず、投稿者同士の評価にさらされる。ここで役にたった（helpful）という評価が一定のレベルに達すると一般に公開され、誰でも見られるようになる。ここまでが第一段階である。公開されたノートは、一般のユーザからの評価にさらされる。ここで役にたたなかった(not helpful)という否定的評価が多いとノートが非公開に戻ることがありうる。一定期間たつとノートは固定される。これが第二段階である。

いずれの段階でも、公開されるか非公開になるかは人々の評価で決まるが、単純な多数決ではない。ガイドラインによれば、ユーザを似た者同士でグループに分類しておき、多くの異なるグループから役にたったとされたノートが公開される。たとえば政治的に右あるいは左にかなり偏ったノートがあったとすると、特定のグループの人にしか評価されないため、公開されない。実際、公開前のノートの中には、事実の指摘ではなく単に自分の意見を述べているものがあるが、この場合は特定の人にしか評価されないため、公開されない。これは、コミュニティノートは意見を述べるのではなく、議論の材料を付加するためだからである。

コミュニティノートは強力である。元のツイートと一体化されるため、ブロックすることもリプライ制限することもできない。従来からもノートのような追加的な情報はリプ欄に書かれることがあったが、多くのリプライの中に埋もれてしまい人々の目には届かなかった。コミュニティノートは元のツイートと一体化されて表示されるため、埋もれることなくいやでも閲覧者の目に入る。この強力さのためコミュニティノートは物議をかもすことになった。

---

[2] コミュニティノートの動作の仕組みの解説は Wojcik et. al. (2022) が比較的詳しい。ウエブ上の記事としては、GIGAZIN, 2023/8/18, 「X（旧 Twitter）のコミュニティノートの表示・非表示を決めるアルゴリズムはどんな仕組みになっているのか」 https://gigazine.net/news/20230818-x-twitter-community-note-algorithm/

２．コミュニティノートはユーザに支持されているか

コミュニティノートは誤情報を防ぐという目的を達成しているのだろうか。これについては日本に先立って 2 年前に試験導入されたアメリカで、ツイートを多数収集し、コミュニティノートの導入前と後を DID や RD などの手法で比較する調査が多く行われている。ただし、結果は割れており評価は定まっていない。

ポジティブな報告としては、コミュニティノートの導入で、誤情報のリツイート数といいね数が減ったという報告がある（たとえば Wojcik et al. (2022)）。極端に感情的なツイートが減り、ツイートの品質の改善が見られたという調査もある（Borwankar et al. (2022)）。しかし、ネガティブな報告もあり、同じ DID で分析しながら誤情報のリツイート数もいいね数もあまり変わっていないという調査結果がある（Chuai et.al. (2023)）。さらにコミュニティノートは結局自分と政治的に反対の陣営にモノ申す手段になっているという指摘もある（Allen (2022)）。

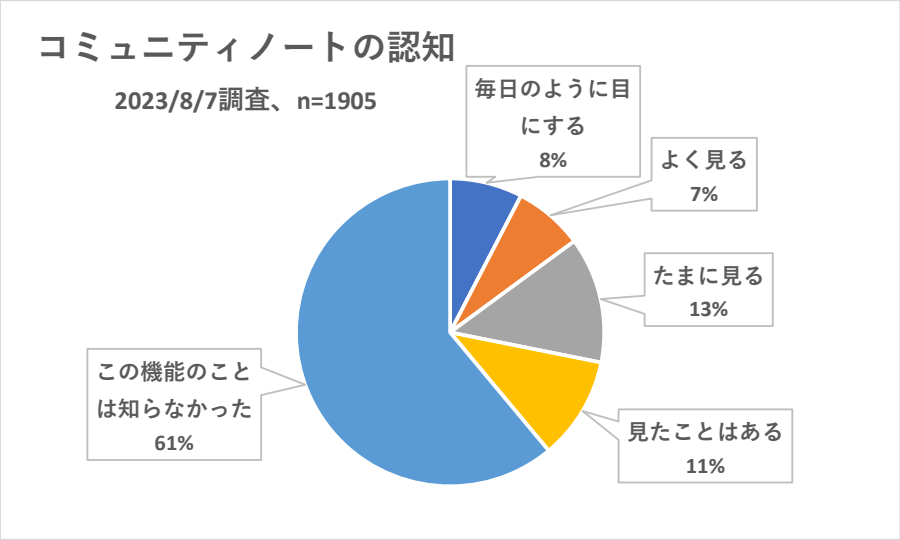
また、そもそもファクトチェック機関のような専門家による修正と、コミュニティノートのような多くの無名の人による修正のどちらが良いかを実験調査した研究もあるがこれも結果が一定しない。無名の人でも十分な数がそろえばジャーナリストと遜色ない判定ができたという報告もあれば（Resnick et al. (2021)）、素人の判定は専門家の判断には及ばなかったという調査結果もある（Godel et al. (2021)）。コミュニティノートのような無名の人による集合的判断という仕組みの評価は定まっておらず、これからの課題である。

日本での導入はまだ始まったばかりであり現時点での評価は難しい。ただ、アンケート調査でユーザがコミュニティノートをどう評価しているかは調べることはできるので、これを試みよう。主観的な評価になるが、新システムの導入直後は人々が起きた“変化”をもっともよく認識できる時期である。時間がたってしまうと記憶が薄れ、起きた変化の解像度が下がっていく。したがって、この時期に調査をして記録を残しておくことには意味があるだろう。

調査日は 2023 年 8 月 7 日、ウエブ調査会社 Freeasy 社の 20 歳～69 歳までのモニターから、年齢別男女別で均等割り当てした 5000 人のネットユーザを集めた。そのなかから現在ツイッターを使っている 1905 人に対して行った。コミュニティノートの本格導入が 7 月に入ってからなので、導入からほぼ 1 か月たったころである。

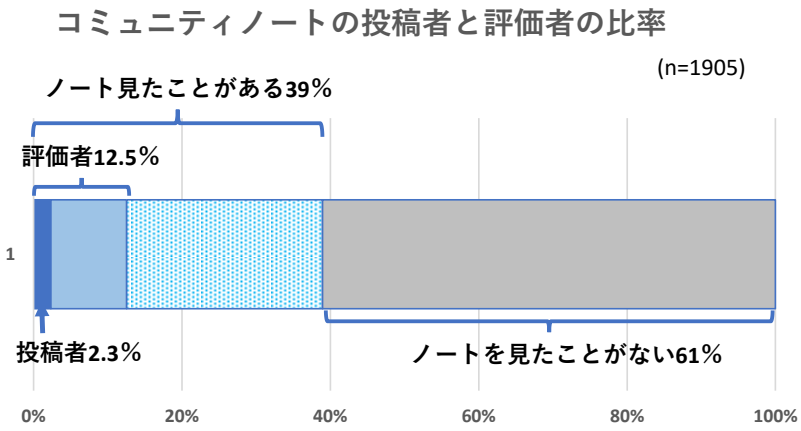
まずコミュニティノートの認知を尋ねた。コミュニティノートをどれくらい目にするかを聞いた結果が図 2 である。コミュニティノートを見たことのある人は 742 人で、今回のツイッターユーザ 1905 人の 39%であり、ほぼ 4 割である。少ないようにも思えるが、コミュニティノートが本格稼働してから 1 か月であること、関心対象が芸能・スポーツ・エンタメで、時事問題にまったく関心がないとコミュニティノートには出会わないことから考えると、4 割というのは妥当な数値であろう。

図２



このうちノートを評価したことがある人は 239 人で全体の 12.5%であった。つまり全ツイッターユーザのなかでノートが役立ったかどうかを評価したことがあるのは 1 割強である。さらに投稿者、すなわちノートを書いたことがある人は 44 人であった。これは全体の 2.3%である。まとめると図 3 のようになる。全ツイッターユーザのうちコミュニティノートを見たことがあるのは 39%で、ノートを評価したことがあるのは 12.5%、そしてノートを書いて投稿したことがある人が 2.3%である。逆に言うと、全ユーザのうち、2.3%の人がノートを書き、12.5%の人がそれを評価し、40%の人がそれを目にしていることになる。[3]

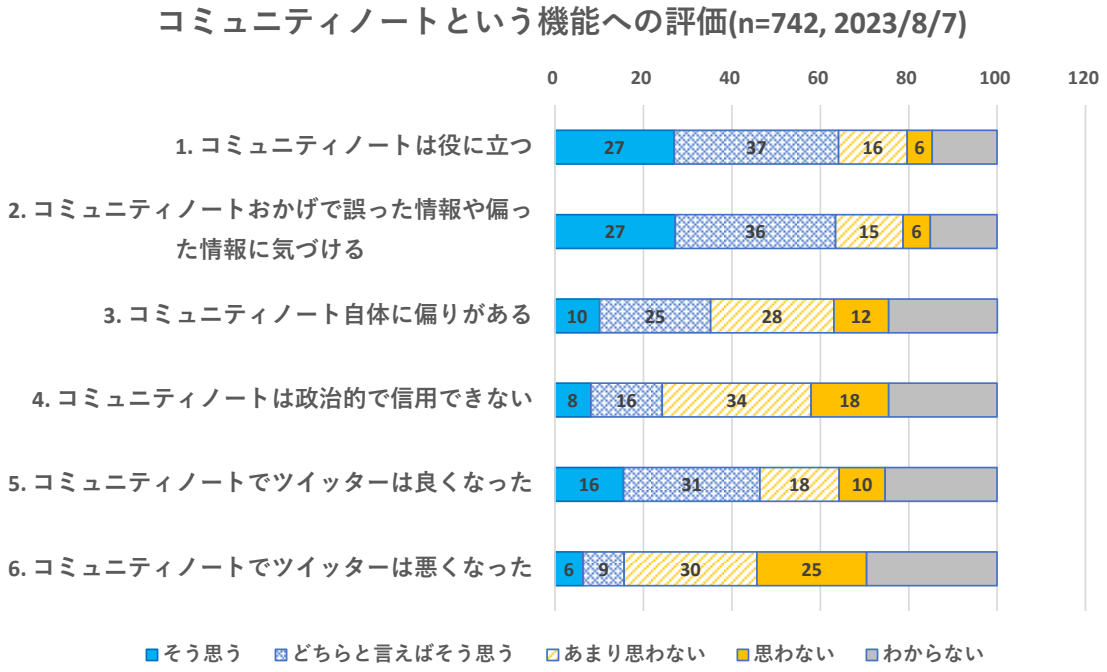
図３



[3] なお、投稿者は必ず評価もするため同時に評価者でもある。すなわち、12.5%の評価者には投稿者が含まれる。ノートを投稿せずに評価だけをする”純”評価者は引き算をして 10%程度である（=12.5－2.3）

コミュニティノートという仕組みを人々は支持しているだろうか。これを知るためにコミュニティノートについて肯定的と否定的な見解を取り混ぜて 6 つ用意し、それにそう思うか思わないかを 4 段階で答えてもらった。対象はコミュニティノートを見たことがあると答えた 4 割の人（742 人）で、図 4 がその結果である。左の青系統の二つが「思う」、「どちらかをいえばそう思う」で、右側の黄色系統の二つが「あまり思わない」、「思わない」、である。

図 4



最初の 1 と 2 はコミュニティノートが役にたつかどうか、そしてノートのおかげで誤情報や偏った情報に気づけるかどうかを聞いたものである。「そう思う」と「どちらかといえばそう思う」を合わせるとそれぞれ 64%、63%の人がそう思うと答えており、「思わない」、「あまり思わない」の 22%、21%を圧倒している。誤情報や偏った情報に気づくようにすることがコミュニティノートの目的であり、この目的は達成されている。

3 と 4 は否定的な見解である。コミュニティノート自体に偏りがあるかどうかを尋ねると、偏りがあると答えた人が 35%、無いと答えた人が 40%で、偏りがないと答える人の方が少しだけ多かった。否定的見解の否定度合いを一歩進めて、コミュニティノートは政治的で信用できないという見解については、「そう思う」は 24%だけで、「思わない」が 52%に達した。否定的な見解は少数派である。

最後に 5、6 でまとめとして、コミュニティノートの登場でツイッターは良くなったか悪

くなったかを聞いた。良くなったと答える人が 47%、悪くなったと答える人が 15%である。両者の差は圧倒的に大きく、ツイッターユーザはコミュニティノートを肯定的に評価しているとみてよい。

ユーザの間でコミュニティノートという機能への評価が高いのはアメリカでも同じである。Wojcik et al. (2021) の調査では、コミュニティノート（当時の名前は Birdwatch）が役立つと答えた人は民主党支持者では 79%で圧倒的であり（役だたないは 11%）、共和総支持者でも役立つと答えた人が 49%（役立たないは 29%）で多数派であった（Wojcik et al. (2021), Figure 4)。冒頭に述べたように誤情報の修正機能については研究者の見解が分かれているのにユーザの評価は高い。奇妙に思えるかもしれないが、見解が分かれていると言っても、コミュニティノートの導入で状況がかえって悪くなったという報告はない。変わらないかあるいは改善したかである。そうであるなら、人々の評価が肯定的であることと、これまでの検証結果は矛盾しない。

コミュニティノートをどのような人が肯定的に評価しているかを見るために、図 4 の評価をひとつの指数にまとめたうえで、いくつかの個人属性に回帰してみた（Appendix 参照）。教育水準、性別、掲示板・炎上への参加経験については有意でなく、フォロワー数は有意なものもあるが一貫した傾向はなかった。年齢は若い人の方がコミュニティノートを支持する傾向があるものの、係数が小さく影響は限定的である。影響を及ぼす属性変数はこれといったものが見当たらない。属性の影響が少ないということは、コミュニティノートへの支持は特定層ではなく広い範囲の人に広がっているという事を示唆する。

ただし、一つ例外があり、政治傾向の影響がはっきりと認められ、保守の人ほどノートを支持し、リベラルの人はあまり支持しない。次節ではこの偏りを検討する。

３．コミュニティノートに保守バイアスはあるか？

コミュニティノートにはリベラル側から批判がある。これまでにツイッター上でコミュニティノートという機能自体への不満・批判を述べた人は、津田大介氏、蓮舫氏、阿部岳氏などリベラル側の論客ばかりである。ノートが付く相手も、きっこ氏、ラサール石井氏、望月衣塑子氏などリベラル論客のツイートであることが多く、保守論客にたくさんのコミュニティノートが付いた例はあまり見られない。したがって、一部にはコミュニティノートはリベラルへの攻撃、あるいは政権の擁護に使われているという批判が見られる。

なぜそのような批判が起こるのだろうか。論理的な可能性としてはコミュニティノートのアルゴリズムが保守側に有利になるように仕組まれているという説明が考えられるが、これはありそうにない。コミュニティノートのアルゴリズムは公開されており、かつこの機能はアメリカで 2 年以上前から稼働しているからである。アメリカのコミュニティノートについてはすでに多くの先行研究があるので、アルゴリズムにそのような意図的な偏り

があればすでに発見されているだろう。

また、ツイッターユーザにはそもそもかなり強い保守（いわゆるネトウヨ）の人が多いからだという説も出そうであるが、これも事実に反する。この調査で当初あつめた 5000 人のモニターのなかにはツイッターユーザではない人が 3000 人近くいる。かれらの政治傾向（測定方法は後述）とツイッターユーザの政治傾向を比較するとほとんど同じであった。[4] ツイッターユーザの政治傾向が特に保守的という偏りはない。

アルゴリズムにもユーザにも偏りがないとすれば、コミュニティノートが運用される過程で内生的に偏りが出ていると考えるほかない。そのメカニズムは何だろうか。二つの可能性が考えられる。

第一は、何らかの理由で保守側がリベラル側よりも活発にノートを使っているという可能性である。理由はともかくとして、保守の人がコミュニティノートの導入で活性化し、意欲的にノート付与に励んでいるとすれば、リベラルのツイートにばかりノートが付くことの説明がつく。

なぜ保守側の方がコミュニティノートに対して活性化したかはわからない。ただ、保守側は、昔から新聞テレビなどのマスメディアはリベラルに偏向していると批判する傾向がある。それゆえ、保守論客はネットの上こそ自分たちの言論活動の場と心得るところがあり、ネット黎明期からネット上での情報発信に熱心であった（私見ではこれがいわゆるネトウヨの起源である）。そうだとするとコミュニティノートのような機能が備わったとき、真っ先に反応して利用し始めたのが保守側だというのはあり得ない話ではない。この説明を、コミュニティノートにおける「保守活性化説」と呼ぶことにしよう。煽りを込めた揶揄表現を使えば、いわばコミュニティノートにおける「ネトウヨ跳梁跋扈説」である。

第二の可能性として、リベラル側のツイートが、そもそもノートが付きやすいツイートだからという可能性が考えられる。これはリベラル側が意図的にデマを流していると言っているのではない。リベラルとは、人権や平等など普遍的理想をかかげて現状を批判し、変えていこうという思想的立場である。理想を掲げて現実を批判するいわば異議申立人であるため、特定の側面だけを訴えることになりやすい。問題点 A の告発を行おうとする弁護士は、相手にも B、C などいろいろ事情があるだろうと慮りはしない。異議申立人は様々の側面を考慮したりはせず、批判すべき A 点だけに限って舌鋒鋭く議論を展開するのが常である。しかし、他の側面に触れなければ、コミュニティノートからすれば格好の注釈対象であり、かくして、ほかにも B 点、C 点がありますと指摘するノートが付くことになる。

特に日本の場合、政権をずっと保守が担当しているため、異議申立人が常にリベラル論客になり、結果としてノートがリベラル論客に集まってしまう。アメリカのように政権がリベラルになることがあると、保守側が異議申立人になるため、保守側のツイートにもノ

[4] ツイッターのユーザの保守リベラル度（後述）は 3.01、非ユーザでは 3.00 でほとんど差がない。つまり、ツイッターユーザの保守・リベラルの度合いは国民全体と同じである。

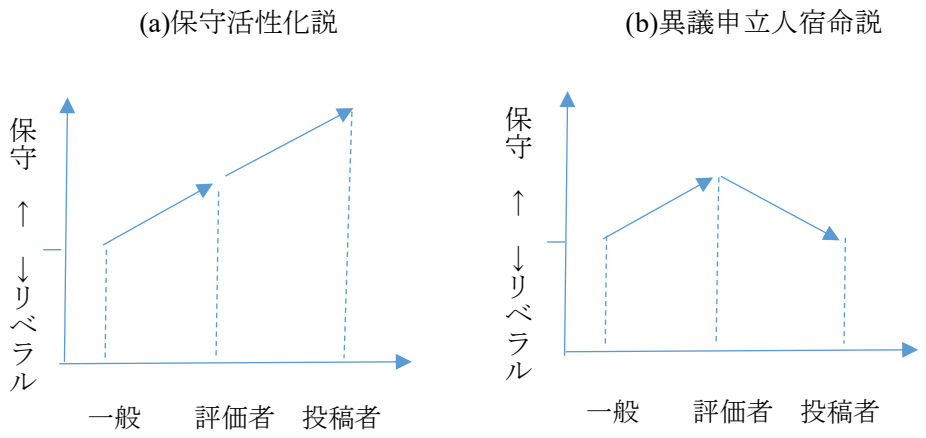
ートがつく。実際、アメリカでのコミュニティノートのテストは 2021 年にはじまったが、当時は民主党政権下で、共和党のトランプが野に下っていたため、トランプ支持者の虚実混ぜてのツイートにノートが付いていた。日本では政権がめったに代わらないため、異議申立人は常にリベラルであり、したがってノートが付くのはリベラル側のツイートに偏ることになる。この説明は異議申立人である限りノートが付くのは避けがたいということなので、「異議申立人宿命説」とでも呼ぶのがよいだろう。煽りを込めて揶揄表現をすれば、これは「リベラル自業自得説」である

保守活性化説と異議申立人宿命説のどちらの仮説が妥当だろうか。これを調べるために、コミュニティノートの活動に参加する人の政治思想の傾向を調べてみよう。ノートの活動に参加する人は投稿者と評価者とに分けられる。投稿者はノートを書く人であり、評価者は公開されたノートに評価を行う人である。もし保守活性化説が正しいなら、評価者は一般ユーザより保守的な傾向があるだろう。そして、ノート作成に時間と労力を投入する投稿者は、それだけの時間を割いて活動する情熱があるのだから、さらに保守思想が強いと予想される。

逆に異議申立人宿命説が正しいなら、ノートを書く投稿者は政治的に中立で特に思想的偏りはなくてよい。ただ、中立な人が書いてもノートは異議申立人であるリベラル系のツイートに付くことが多くなるので、評価する一般ユーザの方は保守傾向の人が多くなるだろう。リベラル系ツイートにノートが付いたとき、喜びいさんでそのとおりと拍手をして評価するのは保守の人と考えられるからである。

まとめると、保守活性化説が正しければ、保守度合いは、一般人＜評価者＜投稿者となる。一方、異議申立人宿命説が正しければ、一般人＝投稿者＜評価者となる。図示すると、左から一般人、評価者、投稿者と並べて保守リベラル度を測ったとき、保守活性化説が正しければ図 5 (a) のように右上がりになり、異議申立人宿命説が正しければ図 5 (b) のように山形になる。これで検証ができる。

図 5



検証のためには保守リベラル度を測る指標が必要である。図 6 はそのための問いである。政治的に意見の分かれそうな問いを 10 問用意し、これに賛成か反対か 5 段階で聞いて 1 点から 5 点までの点をつけ、その平均値をとる（ただし値が大きくなるほど保守度合いが増す方に方向をそろえる）[5]。このようにして得られた保守リベラル度は保守度を表す指数になる。この指数は筆者が他の調査でも用いてきたもので、一貫した結果が得られることがわかっている（田中、浜屋 (2019)）。

図６　保守リベラルの度合いの測定

項目

1 憲法9条を改正する
2 社会保障支出をもっと増やすべきだ
3 夫婦別姓を選べるようにする
4 経済成長と環境保護では環境保護を優先したい
5 原発は直ちに廃止する
6 国民全体の利益と個人の利益では国民全体の利益の方を優先すべきだ
7 政府が職と収入をある程度保障すべきだ
8 学校では子供に愛国心を教えるべきだ
9 中国の領海侵犯は軍事力を使っても排除すべきだ
10 男女の特性を生かした性別役割分業には良い面もある

選択肢

1 賛成
2 やや賛成
3 どちらでもない
4 やや反対
5 反対
6 わからない

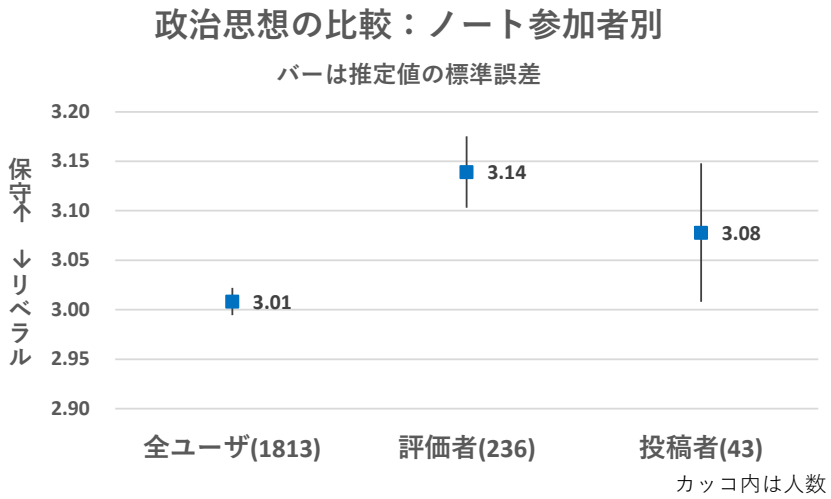
このようにして求められた保守リベラル度の平均値を、全ユーザ、評価者、投稿者の 3 者別に計算した結果が図 7 である。一番左が全ユーザの場合で、保守リベラル度は 3.01 である。この指数の最小値は 1、最大値は 5 なので、うまい具合にちょうど真ん中になる。

その隣が評価者に限ったときで、保守リベラル度は 3.14 に上昇する。保守活性化説、異議申立人宿命説のどちらの仮説でも評価者は保守が多いので、予想通りである。全ユーザの保守リベラル度 3.01 との差は統計的にも有意であり、コミュニティノートでノートに評

[5]　選択肢の番号をそのまま点数とするが、1、6、8、9、10 は保守の主張なので、逆転させて賛成の時に 5 点、反対との時に 1 点とする。

価を与えているのは保守層が多いと言ってよい。

図 7



仮説の検証にとって重要なのは、ノートを書く投稿者である。数が少ないため推定誤差が大きいが、図でわかるように投稿者の保守度合いは 3.08 で、評価者の 3.14 よりも低かった。

これは保守活性化説とは相いれない。保守勢力がコミュニティノートを歓迎して熱心に使おうとしているなら、時間と労力を使ってノートを書く投稿者にこそ熱心な保守がいるはずであるが、そうなっていないからである。投稿者がやや中立に近いということはむしろ異議申立人宿命説に適合的である。ただ、中立に近いといっても、全体平均 3.01 までは戻り切っていないため、異議申立人宿命説が支持されたとも言いにくい。つまり、どちらか一つの仮説だけでは説明できない中間的な結果が得られたことになる。

どう考えるべきか。この二つの仮説は排反ではないため、両者がともにある程度働いたと考えるとこの結果を解釈することができる。まず、保守が活性化して、保守寄りの人が評価者と投稿者に参加する。このとき労力を要する投稿者が多めに増える。投稿者が虚心坦懐にノートを付けても、リベラル側が異議申立人であるためリベラル側のツイートにノートが多く付く。その結果、これに拍手を送る形で役にたったと評価する保守がさらに増える。そして評価者の保守度が投稿者の保守度を上回るというストーリーである。投稿者のサンプル数が少ないため、あまり強い主張はできないが、暫定的な結論としては、二つの仮説はどちらもある程度働いたと考えるのが妥当である。どちらがどれくらいの大きさであるのかは、現データではわからず、さらなる調査を必要とする。

メカニズムはともかく、ノートが付くのがリベラルツイートに偏りがちであることが事実だとして、これが良いことか悪いことかは意見の分かれるところだろう。ノートが付くのがリベラル論客のツイートに偏ったとしても、ノートで重要な判断材料が提供されるの

だから良いことであるという意見もあろう。しかし、ノートは政権批判の出鼻をくじき、結果として政権擁護になるから良くないという意見もあり得る。本稿ではこの是非までは議論しない

ただ、節を閉じるにあたり、これらの偏りを踏まえてもなおコミュニティノートへの支持が強いことを指摘しておきたい。このことを示すために、回答者をリベラルの人に限ってみる。

図 8 がその結果である。6 つの評価項目のうち 4 つだけを取り上げた。各項目の最上段は全ユーザの場合で、2 番目は保守リベラルの軸で 50%ずつ半分に分けたときのリベラル側の人だけに限ったときである。3 番目はさらにリベラル度を強めて、リベラル側 25%の人、つまり強いリベラルの人に限った場合である。

図８　保護利活用仲介機構の詳細例

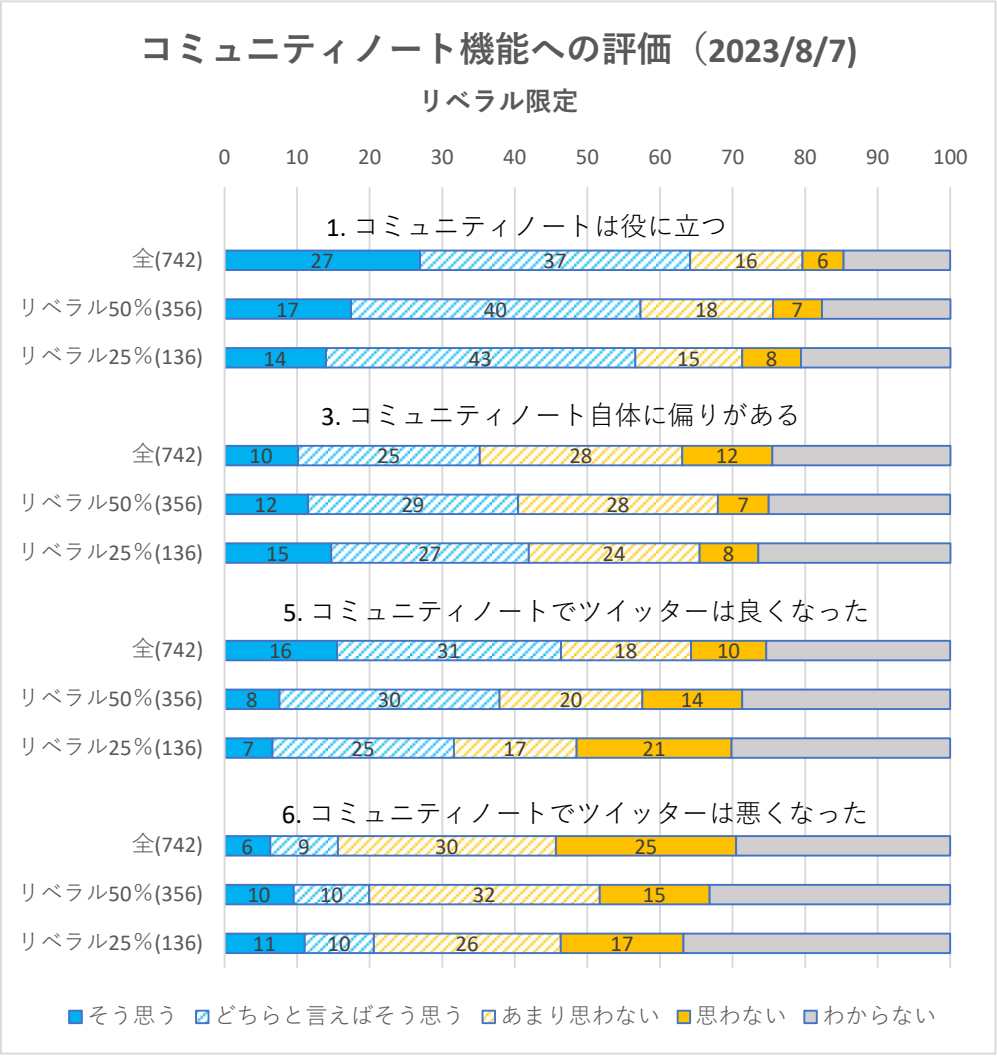


図 8 の (1) はコミュニティノートが役にたつかどうかの問いへの答えである。役にたつと答えた人は全体では 64%（=27+37）あったが、リベラルの人に尋ねると 57%に、さらにリベラル上位 25%に限ると 54%に低下する。リベラル側に寄るとノートの評価は低くなる。

しかし、それでも「そう思わない」25%、23%の 2 倍に達しており、全体としては役にたつという評価の方が多い。特にリベラルの上位 25%という強いリベラルの人のなかで、コミュニティノートは役にたつと答えた人が「そうは思わない」と答えた人の 2 倍いるというのは驚異的である。

ただし、リベラル側はコミュニティノート自体に偏りがあるとも感じている。図 8 (2) を見ると、リベラル側上位 25%では、コミュニティノートに偏りがあると思う人は 42%で、思わない人 32%より多くなっている。リベラル側にとってコミュニティノートに偏りがあることに不満があることは間違いないだろう

しかし、それでも、コミュニティノートの登場でツイッターが良くなったか悪くなったかを尋ねると、よくなったという答えが圧倒的である。図 8 の (3)、(4) を見ると、リベラル側上位 25%に限ってさえ、コミュニティノートの登場でツイッターが良くなったと思う人が 32%、悪くなったと思う人が 21%であり、良くなったと思っている人の方が多い。

この節をまとめれば、偏りはあるものの、コミュニティノートは役にたっており、この機能の追加でツイッターは良い方向に変化したというのが人々の評価だといってよいであろう。

４．投稿者のプロファイリング

コミュニティノートの働き方は投稿者しだいである。コミュニティノートの投稿者は自らノートを書き、相互にノートを評価してそのノートが公開されるかどうかを決めている。

評価者は公開されたノートを再び非公開にするかどうかに関わるだけであり、影響はそれほど大きくない。コミュニティノートの仕組みの中核を担うのは投稿者であり、投稿者が党派性を帯びたり、組織的に行動したりするとコミュニティノートの性格は歪んでしまう。そこで最後に投稿者のプロファイリングをしておく。ただし、投稿者の数は少なく、今回 5000 人をサーベイして見つかったのは 44 人というレアさなので、暫定的な結果になることをお断りしておく。

対象をツイッターユーザとし、投稿者であることを被説明変数としてロジット回帰する。表 1 がその結果である。この係数は限界効果で、投稿者になる確率がどれくらい上がるかを示している。値は小さいが、投稿者自体がツイッターユーザの中で 2%程度とそもそもレアなので、それを考えれば妥当な値である。統計的に有意なものは星印がついている。

表１　コミュニティノート投稿者のプロファイル
（ロジット回帰での限界効果）

VARIABLES	投稿者=1
＜政治傾向>	
1. 保守リベラル指数	-0.0012
	-(0.44)
＜ネットでの活動＞	
2. フォロワー数1～500	-0.003
	-(0.88)
3. フォロワー数501～1000	0.0223
	(1.44)
4. フォロワー数1001～3000	0.0106
	(0.94)
5. フォロワー数3001～	0.0677*
	(1.84)
6. 過去1年に掲示板書き込んだ	0.0548**
	(2.20)
7. 過去1年に炎上に書き込んだ	0.028
	(1.34)
＜デモグラフィック変数＞	
8. 大卒	0.0015
	(0.47)
9. 大学院卒	0.006
	(0.65)
10. 年齢（才）	0.0027**
	(2.60)
11. 年齢2乗（才）	-4E-05**
	-(2.88)
12. 性別（女性＝１）	-0.0058*
	-(1.66)
Observations	1,813
Log likelihood	-144.56048

t-statistics in parentheses
** p<0.05, * p<0.1

順に変数を見ていく。一番上の保守リベラル度は有意ではない。図 7 で投稿者の保守リベラル度を見た時、3.08 とわずかに保守寄りであったが、他の変数を制御すると政治思想の影響は消えてしまい、むしろマイナスにすらなっている。保守的あるいはリベラル的な人が投稿者になるという傾向はないことになる（なお、これは異議申立人宿命説を支持する材料である。）

2 から 5 まではフォロワー数を見たもので、3000 人以上が有意である。フォロワーが 3000 人以上いるということは、ツイッターである程度は情報発信を行っているということ

を示唆する。同じことは 6 の掲示板でも言えて、過去 1 年間にネット上の掲示板に書き込みをしたことがある人が有意に投稿者になっている。ここからもノートの投稿者とは情報発信の経験のある人と推測できる。

8、9 は教育水準を尋ねた。ノートを書くにはある程度の知識が必要であるので用意した変数であるが、有意ではなかった。ノートを書くかどうかに学歴は関係していない。

10、11 は年齢の要因で、一次項が正で二次項が負でいずれも有意である。興味深いことに年齢効果は逆 U 字型で最大値があることになる。ロジット回帰の元の係数に戻って最大となる年齢を求めると 36 歳であり、投稿者の中心は 30 代半ばである。

12 は性別で、女性が 1 となるダミーなので、係数が負ということは、男性が多いということを意味する。実際、44 人のうち、男性は 32 人、女性が 12 人であり、3 対 1 の割合で男性が多い。

まとめると、コミュニティノートで中核を担う人たち、すなわちみずからノートを書き、さらに公開するかどうかを決める投稿者は次のような人たちである。

> 「男性が 4 分の 3 で、年齢は 30 代半ば、フォロワーが 3000 人以上いて、
> 掲示板に書き込むなど情報発信に熱心な人たちである。学歴は問わず、
> 保守リベラルの政治思想に特にこれといった傾向はない」

実際に投稿者を見てみるとノートを書く分野に特徴がある（投稿者は匿名であるが、コードネームのような名前がシステム側から与えられるので、同じ投稿者のノートはまとめて読むことができる）。ある人はコロナ関係のツイートに熱心にノートを付けており、医療関係者であると推察できる。またある人は行政事務や手続きについて詳しいノートを書いており、おそらく行政事務の経験者である。それ以外にも、写真に詳しい人、軍事関係に詳しい人などがおり、それぞれ得意分野でノートを書いていることがうかがえる。

公開前の段階では、ノートとしてふさわしいかどうかについての議論が行われる。「このノートは事実の指摘ではなく意見なので不要である」、とか、「ありえない記述であるが、風刺であることは明らかなのでノートは不要である」、あるいは「ツイート内容と直接関連しない」などの書き込みがなされる。コミュニティノートの趣旨は重要な関連事実の指摘であり、意見の表明ではない。誤情報を正すことはあっても賛否を議論することが目的ではない。現状ではこのルールは投稿者間でよく守られているようである。コミュニティノートが党派性や組織的行為に振り回されず、一定の秩序を保っているのは、このルールが守られているからと考えられる。

この観察の傍証を得るために、投稿者に追加調査を行い、ノートを書く動機を尋ねた。44 人の投稿者の中で返答があったのは 37 名であった。人数が少なすぎて統計的な意義は乏しいが、最初の探索的研究として、また将来の研究のヒントとして結果を記しておく。

追加調査の実施日は 1 週間後の 8 月 14 日である。用意した問いと選択肢は次のとおりで

ある。

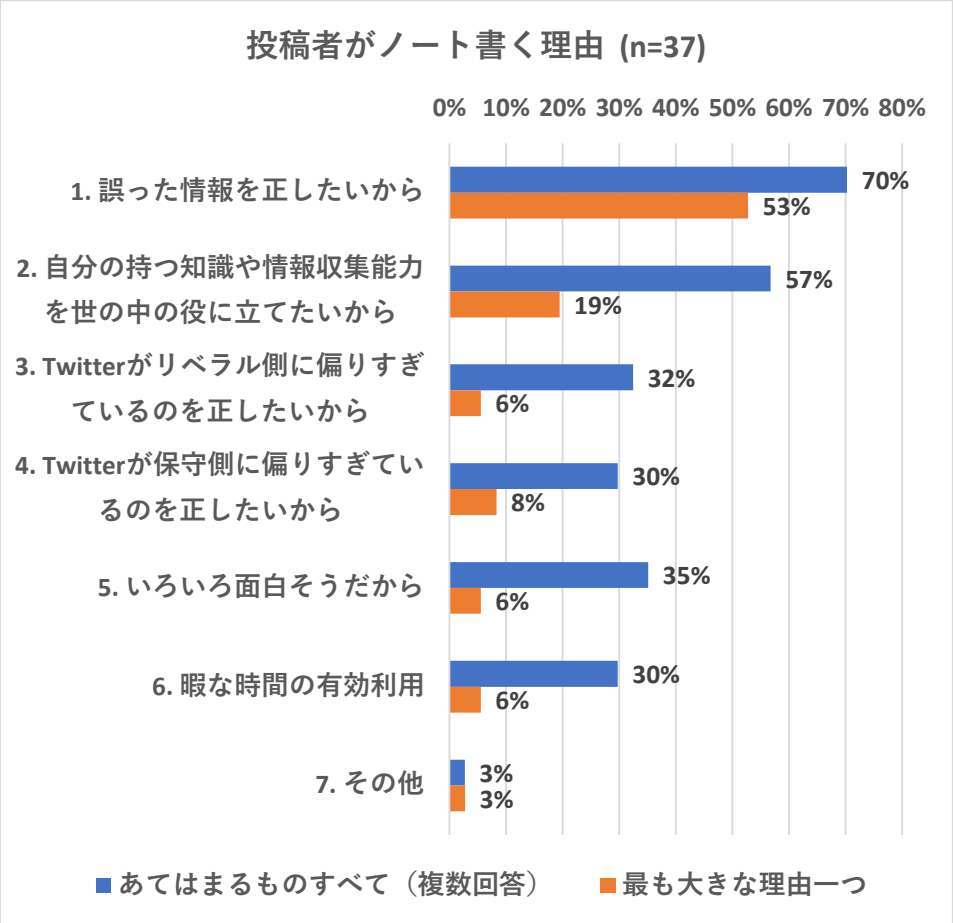
問：あなたはなぜコミュニティノートに書くようになったのでしょうか
あてはまるものをすべて、そのうち最も大きな理由を一つ選んでください。
1. 誤った情報を正したいから
2. 自分の持つ知識や情報収集能力を世の中の役に立てたいから
3. Twitter がリベラル側に偏りすぎているのを正したいから
4. Twitter が保守側に偏りすぎているのを正したいから
5. いろいろ面白そうだから
6. 暇な時間の有効利用
7. その他

1 はコミュニティノートの本来の趣旨である誤情報の修正である。2 は個人として持てる力を社会に役立てたいという動機である。この二つはコミュニティノートへの参加理由としては正当なものであろう。

3 と 4 は政治的な理由である。3 はツイッターのリベラル偏向を正すため、4 は逆に保守偏向を正すためという理由でノートを始めたとしている。回答者が 3 を選べば保守論客、4 を選べばリベラル論客であろう。この選択肢の表現はあざといので選ぶ人は少ないと想定されるが、それでもこれを選ぶ人がどれくらいいるかを見るためと、もうひとつ、3 と 4 に差があるかどうかを見るために置いた。特に 3 と 4 に差があるかどうかが重要で、保守活性化説に従えば 4 よりも 3 が多いはずである。5、6、7 は、それ以外の理由である。

結果は図 9 にまとめた。青いバーがあてはまる理由を全て上げた複数回答のときで合計は 100%を超える。赤いバーはもっとも重要な理由をひとつだけとした単数回答のときの答えで合計は 100%である。

図 9



結果を見るとまず、1 と 2 の「誤った情報を正したい」、「自分の力を役立てたい」が群を抜いて多い。複数回答では、誤った情報を正したいが 7 割、自分の力を役立てたいが 6 割弱である。単数回答でも 5 割と 2 割であり、二つ合わせて 7 割（72%）にも達する。政治的な動機である 3 と 4 は単数回答では合わせても 14%にしかならない。コミュニティノートの投稿者は、少なくとも主観的な自覚としては、ノートの本来の趣旨に沿って参加していることがわかる。

無論、政治的な理由も皆無ではなく、複数回答では 3 割の人が、保守への偏向あるいはリベラルへの偏向を正したいから、を選んでいる。ただし、調べてみるとこのうち半分強の人は 3 と 4 の両方を選んでおり、保守だけあるいはリベラルだけの偏りを正そうとしているわけではない。[6] さらに注目すべき点は保守とリベラルがともに 3 割で差がないこと

[6] 37 人のうち、複数回答で 3（4）を選んだのは 12 人（11 人）であるが、そのうち 7 人は 3 と 4 の両方を選んでいる。保守とリベラルの両方の偏りを正したいと思うことは可能であり矛盾ではないだろう。ただ、その場合、彼は政治的にはリベラルとも保守とも言い難く、強いて言えば中立と考えるべきである。

である。すなわち、投稿者に政治的動機があるとしてもそれは拮抗しており、どちらかに偏っているわけではない。これは、表 1 での投稿者へのロジット回帰で、保守リベラル度が有意にならなかったことと符合する結果であり、異議申立人宿命説を支持する結果である。

全体として、ノートを書く理由として政治的な動機の占める割合は小さい。コミュニティノートは誤情報を正すことが目的であり、意見を述べる政治活動の場ではない。ノートを書く動機を見るかぎり、このルールは守られてきたように見える。ツイッターユーザの間にコミュニティノートへの支持が強いのはこのルールが維持されているからと考えられる。

逆に言えば、このルールが崩れる時、コミュニティノートは混乱し、最終的には失敗するだろう。投稿者の中に政治的動機で戦略的に行動する大きな集団が現れ、ノートを操作するようになれば事態は変わる。当然対抗する集団が現れ、投稿者の間で闘争が起こればコミュニティノートは荒らしに荒らされる政争の具となり、人々からの支持は失われる。現状はまだ始まったばかりであり、投稿者の数は少ない。インターネットの歴史は、参加者が少ないうちはうまくいっていても、拡大するにつれて荒れ始めるのが常であり、コミュニティノートもそうならないかどうかはこれから見極める必要がある。

無論、運営もこのことは十分気を付けており、随所に荒らし対策が組み込まれている。投稿者になるには電話番号認証が必要で、ツイッターへの登録から 6 か月以上たった人だけである。1 日で書けるノート数に上限があり、問題行動を繰り返すとポイントが下がってノートを書けなくなる。このように荒らし対策は多岐にわたる。

荒らし対策のうち最大の物は、すでに述べたようにノートを公開するには同じ考えの人から支持されてもダメで、考えの異なる広い範囲の人からの支持がいるという仕様である。これにより過激派やあるいはカルト的な団体が組織的に投稿者になってノートを意図的に操作することを防いでいる。今のところこれは成功していると思われるが、これが持続するかどうかはこれからわかることであろう。

５．要約と課題

本稿は探索的研究である。ゆえに、得られた知見を簡単に箇条書きで要約し、そのうえで今後の課題を述べることにする。

（1） コミュニティノートはツイッターユーザの 2%程度の人が書き、12%程度の人が評価をくだし、40%くらいの人が見ている。

（2） コミュニティノートはツイッターユーザに支持されている。「役にたつ」、「誤情報を知ることができる」、「コミュニティノートでツイッターは良く

なった」と思う人がそう思わない人より圧倒的に多いからである。

（3） コミュニティノートはリベラル側のツイートに付くという批判がある。このような偏りが生じる理由としては、保守が積極的にコミュニティノートを使っていること、ならびにそもそもリベラルのツイートが異議申し立て型なのでノートが付きやすいという事情が考えられる。

（4） ただ、このような偏りがあっても、コミュニティノートへの支持は厚い。ユーザをリベラルの人に限っても支持する人のほう方が多いからである。

（5） ノートを書く投稿者は男性が 4 分の 3 で、年齢は 30 代半ば、フォロワーが 3000 人以上いて、掲示板に書き込むなど情報発信に熱心な人たちである。現状では彼らは政治的意見を述べるのではなく事実の指摘を行うというノートのルールを守っており、これが人々の支持を集めている理由と考えられる。

以上を踏まえてこれからの課題を考える。誤情報の修正はこれまではファクトチェック機関やプラットフォーマーが担ってきたが、これには限界があり、これを補完する方法として登場したのが、無名の人たちによる誤情報修正システム（Crowd-sourced misinformation management）であった。ツイッターのコミュニティノートはその代表的な仕組みであり、これまでのところうまく機能しているようである。しかし、課題もある。

最大の課題は、すでに述べたように、投稿者数が増えるにつれて荒れてくることはないかという点である。繰り返し述べるように、荒らしを防いでいるのは、異なる意見の人からの賛同がないと公開されないというコミュニティノートの仕組みである。これは優れたアイデアと考えられるが、まだ運用がはじまったばかりで評価は確定していない。様々の調査研究が待たれるところである。

調査研究方法としてはアルゴリズムの検討など様々な方法が考えられるが、本稿の延長として考えられるのは、投稿者を多数集め、その行動を詳細に追跡することである。これはサンプル数を大幅に増やせば達成できる。現状では 44 名しか集められなかったために言えることは限られるが、もし数百人の投稿者を集めれば、ノートを書く頻度、対象ツイートの種類、公開される確率などがわかり、それをその人の属性と組み合わせればノート生成・公開の過程が分析できる。これは次にコミュニティノートを調査する人の課題として期待したい。

文献

Allen, Jennifer, Cameron Martel, and David Rand, 2022, “Birds of a feather don’t fact-check each other: Partisanship and the evaluation of news in Twitter’s Birdwatch crowdsourced fact-checking program.” https://dl.acm.org/doi/10.1145/3491102.3502040

Borwankar, Sameer, Jinyang Zheng, Karthik Kannan, 2022, ”Democratization of Misinformation Monitoring: The Impact of Twitter’s Birdwatch Program.” https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4236756

Chuai, Yuwei, Haoye Tian, Nicolas Pröllochs, Gabriele Lenzini, 2023, “The Roll-Out of Community Notes Did Not Reduce Engagement With Misinformation on Twitter.” https://arxiv.org/abs/2307.07960

Godel, William, Zeve Sanderson, Kevin Aslett, Jonathan Nagler, Richard Bonneau, Nathaniel Persily, and Joshua A. Tucker, 2021, ”Moderating with the mob: Evaluating the efficacy of real-time crowdsourced fact-checking.” Journal of Online Trust and Safety 1, 1 (2021). https://doi.org/10.54501/jots.v1i1.15

Straub, Jeremy and Matthew Spradling, 2022, ”Americans’ Perspectives on Online Media Warning Labels.” Behavioral Sciences 12, 3 (2022), 59. https://doi.org/10.3390/bs12030059

Resnick, Paul, Aljohara Alfayez, Jane Im, and Eric Gilbert, 2021, ”Informed Crowds Can Effectively Identify Misinformation.” arXiv preprint arXiv:2108.07898 (2021). https://arxiv.org/abs/2108.07898v1

Walker, Mason and Jeffrey Gottfried, 2019, ”Republicans far more likely than Democrats to say fact-checkers tend to favor one side.” https://www.pewresearch.org/fact-tank/2019/06/27/

Wojcik Stefan, S. Hilgard, N. Judd, D. Mocanu, S. Ragain, F. Hunzaker, K. Coleman, J. Baxter, 2022, “Birdwatch: Crowd Wisdom and Bridging Algorithms can Inform Understanding and Reduce the Spread of Misinformation.” https://arxiv.org/abs/2210.15723

田中辰雄・浜屋敏、2019、『ネットは社会を分断しない』、KADOKAWA

Appendix

被説明変数「コミュニティノートへの（肯定的）評価値」は、図 4 のコミュニティノートへの 6 個の項目の評価から作成した。「そう思う」～「思わない」までの 4 段階に 1 点から 4 点まで点数をつけて、6 個の回答の平均値をとった値である。点数を振るとき、値が大きいほど肯定的になるように方向をそろえる。下の表の推定値は標準化回帰係数である。

結果を見ると、フォロワー数には有意なものがあるが、でこぼこがあり一貫した傾向とは言えない。年齢が有意であるが、標準化回帰が－0.0582 で値が小さい。最も影響が大きいのは保守リベラルの政治傾向で、係数も大きく有意性も高い。係数は正の値なので、保守の人ほどコミュニティノートに肯定的である。

VARIABLES	(1) コミュニティノートへの（肯定的）評価値
＜政治傾向>	
保守リベラル指数（保守度）	0.2359**
	(8.46)
＜ネットでの活動＞	
フォロワー数1～500	0.0701**
	(2.52)
フォロワー数501～1000	0.0344
	(1.25)
フォロワー数1001～3000	0.0773**
	(2.80)
フォロワー数3001～	0.0348
	(1.28)
過去1年に掲示板書き込んだ	0.0385
	(1.37)
過去1年に炎上に書き込んだ	-0.0275
	-(0.98)
＜デモグラフィック変数＞	
大卒	0.0105
	(0.39)
大学院卒	-0.0207
	-(0.76)
年齢（才）	-0.0582**
	-(2.12)
性別（女性＝１）	-0.0103
	-(0.36)
Constant	1.8111
	(14.49)
Observations	1,325
R-squared	0.077

t-statistics in parentheses
** p<0.05, * p<0.1