

「デジタル空間における情報流通の健全性確保の在り方に関する検討会」(第10回)・  
ワーキンググループ(第4回)

1 日時 令和6年2月27日(火)17時00分～19時00分

2 場所 オンライン開催

3 出席者

(1) 構成員

宍戸座長、山本(龍)座長代理、生貝構成員、石井構成員、越前構成員、奥村構成員、  
落合構成員、クロサカ構成員、後藤構成員、澁谷構成員、曾我部構成員、田中構成員、  
増田構成員、水谷構成員、森構成員、安野構成員、山口構成員、山本(健)構成員、  
脇浜構成員

(2) オブザーバー団体

一般社団法人安心ネットづくり促進協議会、一般社団法人新経済連盟、一般社団法人セーフ  
ティーインターネット協会、一般社団法人ソーシャルメディア利用環境整備機構、一般社団法  
人テレコムサービス協会、一般社団法人電気通信事業者協会、一般社団法人日本インターネ  
ットプロバイダー協会、一般社団法人日本ケーブルテレビ連盟、一般社団法人日本新聞協会、  
日本放送協会、一般社団法人MyData Japan、一般財団法人マルチメディア振興センター

(3) オブザーバー省庁

内閣官房、内閣府、警察庁、消費者庁、デジタル庁、文部科学省、経済産業省

(4) 総務省

湯本大臣官房総括審議官、西泉大臣官房審議官、田邊情報通信政策課長、  
大澤情報流通振興課長、恩賀情報流通適正化推進室長、内藤情報流通適正化推進室課長補佐、  
上原情報流通適正化推進室専門職

(5) ヒアリング関係者

株式会社はてな 大西氏、中川氏

#### 4 議事

(1) 関係者からのヒアリング

(2) 令和6年能登半島地震におけるデジタル空間の偽・誤情報流通状況

【宍戸座長】 それでは、定刻でございますので、「デジタル空間における情報流通の健全性確保の在り方に関する検討会」の第10回会合並びにワーキンググループ第4回会合の合同会合を開催させていただきます。

本日も御多忙のところ、本会合に御出席をいただき、誠にありがとうございます。

議事に入る前に事務局より連絡事項の説明をお願いいたします。

【内藤補佐】 本日事務局を務めます総務省の内藤です。

まず、本日の会議は公開とさせていただきますので、その点、御了承ください。

次に、事務局よりウェブ会議による開催上の注意事項について案内いたします。本日の会議につきましては、構成員及び傍聴はウェブ会議システムにて実施させていただいております。本日の会合の傍聴につきましては、ウェブ会議システムによる音声及び資料投影のみでの傍聴とさせていただいております。事務局において傍聴者は発言ができない設定とさせていただきますので、音声設定を変更しないようお願いいたします。

本日の資料は、本体資料として、資料10-1-1から10-2までの4点を用意しております。万が一お手元に届いていない場合がございますら事務局までお申しつけください。また、傍聴の方につきましては、本検討会のホームページ上に資料が公開されておりますので、そちらから閲覧ください。

なお、ヒアリングシートの回答にはURLが記載されているものもありますので、御参加の皆様におかれましては、適宜アクセスしながら御確認ください。

なお、本日は、江間構成員が御欠席予定、落合構成員、曾我部構成員、山本健人構成員は会議途中から御出席予定、後藤構成員は会議途中で御退席予定と伺っております。

事務局からは以上です。

【宍戸座長】 ありがとうございます。それでは、まず、本日の議事について御説明を申し上げます。まず、関係者からのヒアリングといたしまして、本日、プラットフォーム事業者1社様からの御発表、それから質疑の時間を設けたいと思います。その後でございますけれども、澁谷構成員から御発表、そして質疑の時間を設ける、このように進めさせていただきたいと思います。

それでは、早速議事に入ります。本日は、株式会社はてな様からのヒアリングを予定してございます。はてな・大西様、中川様においでいただいておりますので、30分で御説明をお願いいたします。なお、5分前、1分前には事務局より御案内を申し上げますので、恐縮でございますが、時間厳守で御発表いただければと思います。

それでは、準備ができ次第、どうぞよろしくお願いいたします。

【はてな（大西氏）】 よろしくよろしくお願いいたします。

では、先ほど御紹介あずかりました株式会社はてなです。すいません、自己紹介のスライドを入れていませんでしたが、私ははてなの創業メンバーで、今は組織基盤開発本部長と取締役をやっております大西と申します。よろしくお願いいたします。

本日は、はてな、御存じない方もたくさんいらっしゃるかなと思いますので、私のほうから会社のどういったサービス事業やっているかという説明をさせていただきまして、その後、中川、特にユーザーサポートとかコミュニティーマネジメントをしてもらっている中川から今回の会の趣旨に沿った取組の内容について御説明させていただこうと思いますので、よろしくお願いいたします。

株式会社はてなです。2001年7月に創業して、今は東証グロース市場にあります。

所在地は京都と東京、私は今日は京都からリモートで接続させていただいております。

インターネット関連事業ということで、2001年から、今年24年目、プラットフォーム事業を中心にいろいろやってきましたということで、事業の説明をさせていただこうと思います。

こちらはサービスのミッションですけど、もともと人力検索はてなという、人が聞いて人が答えるQ&Aサービスというのを始めたというのが創業でして、これからユーザー投稿サービスを中心にインターネットサービスを開発・提供してまいりました。

今やっている事業としては、個人ユーザー向けのコンテンツプラットフォームサービス、そして企業向けのコンテンツマーケティング、テクノロジーソリューションといった3つの分野でサービス事業を展開しております。

簡単に順に御説明いたしますと、コンテンツプラットフォームサービスというのは、ブログであったり、はてなブックマークというソーシャルブックマークサービスですね、個人ユーザー向けのコンテンツプラットフォームサービスを提供しているという事業範囲です。

こちらに関して、今回のお話の趣旨にあるのかなと思っておりますので、後ほどこれらのサービスについてはもう少し説明してまいります。

コンテンツマーケティングとしては、そういったブログやブックマークというプラットフォームを使って企業さんのコンテンツマーケティングを支援しましょうということで、他社さんのオウンドメディアを一緒につくったりとか、そういった事業をやっていたり、あとは、個人ユーザー向けで培った技術やノウハウを生かしたサーバー監視とか、マンガビュ

ワー、共同開発といったソリューションも提供しております。

こちらにも簡単に説明しますと、サーバー監視サービス、Ma c k e r e lであるとか、デジタルマンガビューワ、G i g a v i e w e rであるとか、あとは、漫画投稿サービス、こちらはややプラットフォームに近いんですが、漫画の投稿サービスであるとか、あと小説の投稿サービスというのもやっております。こちらに関しては自社だけでやっているものではなくて、出版社さんと共同で開発してやっているとか、出版社さんからの御依頼を受けて受託開発しているみたいなサービスですので、今回は特に話題にはしないでおこうと思っておりますが、こういった事業もやっております。

あとは、任天堂さんと一緒にゲームに関連するサービスをつくったりもしておりますということで、改めてコンテンツプラットフォームサービスのほうに戻ってきますけども、提供開始からそれぞれ10年を超えるという個人ユーザー向けのコンテンツプラットフォームサービスをやっております。

まずは、はてなブログですね。もともとはてなダイアリーという前身となる日記サービスがあり、ブログブームと、インターネット上で日記的なコンテンツを書いていくという中で、ブログという言葉も2005年ぐらいからかなり認知されるようになってきて、はてなダイアリーもはてなブログという名前でもリニューアルスタートして、それも10年以上という形です。

「思いは言葉に。」という、そういったキャッチコピーも掲げておまして、特にしっかりと書いてもらおうと。あなたの思いをしっかりと書き残してくださいと、そういったことを大事にしているブログサービスになります。

ブログサービスというものに関しては、特段これ以上の説明は不要なのかなと思っておりますが、もし分からないこととかありましたら、後ほどでも聞いていただければと思います。

そして、はてなブックマークというのを、こちらは先ほどソーシャルブックマークと申し上げましたけども、ブックマークというのは、ブラウザのブックマーク、もう1回見たいページをどこかに保存しておくというような機能があると思いますけども、そちらをインターネット上に保存しましょうというのがソーシャルブックマークです。

どこがソーシャルなのかというと、ただ単に自分がブックマークしたものを後で見返せるという自分のためだけのツールではなくて、ほかの人がブックマークしたものが、それが可視化されるということによってソーシャル性を持つてくるというものになります。

こちらに関して、我々は、はてなブックマークには3つの機能がありますというふうにさせていただいている、保存と共有と発見ですね。ブックマークという行為自体は、先ほど申し上げたように、自分が見たサイト、インターネットのウェブページをどこかに保存しておこうと。で、もう1回見たいと。こちらはオンラインのウェブサービスですので、スマートフォンで保存してPCで見るとか、家で保存したものを会社や学校で見るとか、そういう時間やデバイスを選ばずに自分がはてなのIDでログインしたらどこでも同じものが見れるといった、まずそういった自分にとっての便利サービスというのと、あとは共有ですね。このページを読んだ人の感想が見たいと。はてなブックマークのブックマークページにはユーザーのコメントが表示されるんですね。

なので、実際見てみましょう。はてなブックマークの、例えばこれは、今は、「味付け塩だけ、放置1時間、調理5分。『塩豆腐と塩挽肉の炒めもの』」。これが今ちょっと人気でたくさんブックマークされています。282人がブックマークしている。はてなブックマークのトップページがこういった形で、今、最近たくさんブックマークされているというページが並んでおりますので、このページがまずインターネットの縮図のようになっていて、今話題のページというのが割とリアルタイムに近く見れますと。

先ほど言った共有ですね。このページをブックマークした人が、今1人増えて283人になりましたけど、この人たちが自分がブックマークする際にコメントをつけられるんですね。そのコメント、ほかの人がこのページをどういうふうにコメントしているかというのを見ることで、このページに対してほかの人はこういう意見を持っているんだとか、こういう感想を持っているんだみたいなことを知ることができると。今見たのは、レシピみたいな、割と問題のないコンテンツですけども、例えばうそや偽情報みたいなのがあった場合は、こういったブックマークコメントで、この情報って怪しいんじゃないのと指摘するとか、そういった効果もあるのかなと思っております。

というのと、そしてもう一つは発見ですね。今話題になっているページが知りたいと。はてなブックマークは、旬なページが、旬なニュースが見つかるメディアですということで、先ほど申し上げたように、はてなブックマークのトップページにはこういった今話題のコンテンツというのが出てきて、例えば戦争が始まりましたとか、芸能人がスキャンダルを起こしましたみたいなときって、インターネット上にたくさんのニュースサイトにそういった記事が載ると思うんですけども、そういったものがいち早く、たくさんの人が注目されたものがここに集まってきて、それに対してみんながどういう意見を持っているかみたいな。

例えばヤフーとかがやっているニュースとかにもコメントがついたりするかなと思うんですけど、どちらかというと、どういう人が注目しているかという注目度で集めたりとか、ほかの人のコメントを見比べることができるみたいなところがはてなブックマークの特徴的なところかなと思っております。

こういった保存、共有、発見を中心としたプラットフォームとなっております。

はてなとしては、そういった情報、はてなブログというのはブログサービスで、情報を発信するというサービスで、はてなブックマークは情報を拡散するという意味合いがあるかなと思っております。はてなブックマークに掲載されることで、ある意味この記事が目立つと。その目立ったことによってよりブックマークされてみたい、そういう人気の記事をより拡散して、はてなブックマークを例えばSNS、X、元ツイッターに連携される方とか、そういう機能も提供しておりますので、はてなブックマークで注目された記事がSNSを通じてより広まっていく。そういう人気の記事を拡散していくみたいな効果もあると考えておまして、それを我々としてはコンテンツプラットフォームビジネスにも活用いたしますし、また、そういった記事を、コンテンツをつくるという部分と拡散する部分と両方を担っていることから、これらの事業をコンテンツプラットフォームであると呼んで成長させていこうとしているという感じでございます。

以上、簡単ですが、はてなの会社と行っている事業について御説明させていただきました。この後は中川にバトンタッチさせていただこうと思います。

**【はてな（中川氏）】** よろしくお願ひします。改めて資料を共有させていただきます。

では、よろしくお願ひします。株式会社はてなでコミュニティーマネージャーという役職をやっておりますはてなの中川と申します。コミュニティーマネージャーは、はてなの運営するコンテンツプラットフォームを健全に運営する、あと、ユーザーさんにストレスなく使っていただくという、そういうところ、ユーザー体験をよくするという、この2つのところを担っている仕事です。よろしくお願ひいたします。

これまで、先ほど大西のほうからはてなのサービス、運営するサービスの概要を説明させていただきましたが、はてなブログとはてなブックマーク、あと、よく目にされるであろうサービスとしてははてな匿名ダイアリーというサービスがありまして、これは2006年からサービスをやっております。はてなでは、はてラボとって実験的なサービスを提供するプラットフォームがあるんですけども、その上で動いている匿名でコミュニケーションをするというサービスです。これも匿名ですけども、何でも書き捨てというわけじゃな

く、どちらかというと言み人知らずみたいな形で独自の文化を生み出そうというような、そういうような目的で運営しているサービスとなります。

いずれも、これまで多分発表されたほかの会社さんと比べてかなりユーザー数は小規模だと思っています。

小規模なサービスでの対策を共有する意義として、近年では個人レベルで分散型SNSのメトリクスを運用されている方も含め、小規模のサービスをつくられるスタートアップの方もおられますので、小規模なサービスをコンパクトな体制で健全に運営する方法というのにも需要があると思っていますし、また、そこも整備されないと安心して新しいサービスを生み出せないとも思いますので、弊社の事例をその一助としていただければと考えています。

まず、弊社における偽情報、フェイクニュースとかの対策への考え方なんですけれども、まず、先ほど大西が言いましたように、はてなブックマークなどでは、この情報おかしいんじゃないの、間違っているんじゃないのみたいなコメントが忌憚なく寄せられるような環境です。理想としては、ユーザーコミュニティの健全な集合知によって自然に正しい情報が目立って流通するというのが一番の理想かなと思っています。

ただ、実際には間違った情報が飛び交ってしまうということもありますので、そういう場合には、ユーザーからの通報、さらにそれをスタッフによる目視で公開停止などの措置をとるとというのが基本的な動きになります。それに加え、近年ではシステム面でフィルタできる要素も様々ありますので、そのような要素があれば取り入れていくというふうな体制で運営しております。

今資料のほうに、ヒアリングシートのほうに記載したんですけれども、このような業務に割かれる人間というのが大体5人ぐらいの体制で運営しておりまして、サービス規模の割には小規模で運営しているかなと考えております。

スタッフが目視判断を行う際の判断基準なんですけれども、まず情報が合っているか、正しいか、間違っているかというものに関する判断というのは現実的には困難であると考えております。

と申しますのは、ある情報を正しいか間違っているか判断するには一般的にはかなりコストとか知識とかが必要となります。もちろん荒唐無稽なものですとか、完全にうそみたいな誰もが偽情報とわかるものであればいいんですけれども、多くの場合はある程度のリアリティーがある情報であると思いますので、対応するスタッフがそのような情報



の精度を検証することはまず不可能であると考えています。

判断基準とできるような公正な情報源があるとも限りませんし、特に賛否のある意見とか党派性のある意見について、誤情報であると運営側が決めつけてサービスを削除してしまうと、サービスの中立性とか公正性みたいなところについて疑念を持たれてしまうということにもなりやすく、結果、そこでサービスへの信頼が損なわれたり、安心して使っていないことになりかねないです

なので、目視判断の際の基準として、より公正に納得感が得られる形として、投稿の目的自体がよくないとか、投稿によって人に迷惑がかかる、そういうような観点で判断基準を設けていくというのが、公開するガイドラインとしても、社内の内規としても設計しやすいと考えています。

はてなとして、投稿動機によって誤情報を分類しました。ヒアリングシートに記載したんですけども、実際の投稿管理業務において、誤情報をどのようなジャンルで分類しているかというところでは、やはりどういう内容であるかという分類ではなく、さきほども申し上げましたように、投稿の動機とか、投稿による実際に生じた影響で分類して対応方針を考えていることが多いので、そのような分類としています。

1つ目が、組織的な世論操作を目的に投稿されているものです。

2つ目が、営利宣伝とかインプレッションを稼ぐためにセンセーショナルな投稿している。

嫌がらせや誹謗中傷のためにデマを投稿している。

4番目が、発信者自身が間違っている。そういうこの4つの分類をしています。

もう一つ、AIの実験など、本人に悪意がない理由で意図せず適当な情報が流されてしまうみたいなのも近年ではあるのかなと思いますけれども、現時点ではまだ大きい影響はないと考えられますのでひとまずこのような分類でお話します。

1つ目、組織的な世論操作に関して、近年、世論を誘導するために大規模な誤情報による情報操作が行われているということがニュースになっておりますけれども、弊社のような小規模なサービスの場合、そもそも社会的な影響力がさほど大きくないということもあり、そのような操作を行っても世論を動かすよう流れをつくるのに対して必ずしも効率がいいとは言えないため、そのようなターゲットにはなりづらいと考えています。これは弊社に限らず、小規模なサービスがそのような操作のターゲットになるということ自体があまりないと思っております。

次、営利宣伝とかインプレッション稼ぎに関しては、これまで一般的にサービスにおいて行われているスパム対策ですとか過剰な最適化、ブラックSEOみたいなものですね、そういうものへの対策の延長線上で対応ができるかと思っております。

これは先ほど申し上げた通報から目視というプロセスに加えて、例えば投稿者の接続元情報が何か偽装しているとか、利用の状況が不自然であるとか、大量に同じところから投稿されているとか、そういうようなものに関してシステムである程度検知ができるので、目視に加えてシステムでの検知で対応しております。

次、嫌がらせとか誹謗中傷に類するものですね。これは従来より権利侵害情報や誹謗中傷情報対策を行っておりますので、その延長線上にあるものかと思えます。具体的にはプロバイダ責任制限法に基づく対応ですとか、不適切な投稿、嫌がらせ、迷惑行為などに対する通報対応で対応しております。

ここまであげた2種類の投稿行動に対応する際の考え方として、ルール作りや通報、監視体制なども大事なんですけれども、まず利用者が不適切な投稿で効率よくインプレッションを稼いだり安直にデマを投稿できない仕組みとか、サービス風土がそれを許さないみたいなサービス設計もとても大事であると考えております。

弊社の場合、まずサービス全域で健全な批判言論を歓迎するコミュニティーポリシーというものを明確に打ち出しています。なので、無責任な書き捨ては許容していません。何か強い表現でデマを発信すると必ず強く批判されるというような状態になっているのではないかと考えています。

次、閲覧数や反応数を単純にインプレッションに結びつけない配慮をしています。具体的には、数のランキングとか、あるいは注目度だけの何かバズっているものが見えるみたいな、数に基づく評価については慎重に行っていて、たとえば、はてなブックマークでは人気エントリーの選定については、建設的な内容であるとか、多様性があるトピックであるかといったブックマーク数以外の指標も設定したアルゴリズムとしております。

また、このようなアルゴリズムの設計方針に関するポリシーについては、はてなブックマークの場合、ガイドラインで公開しております、どのような情報が注目されやすく設計されているサービスかということは共有されているという状態になっています。

あとは、先ほど申し上げた匿名ダイアリーですね。匿名であることをいいことにデマの温床となるということはよくあることかと思うんですけれども、匿名ダイアリーにおいては、固有の人名とか法人名が特定できる記事については、本人からの依頼があった場合、無条件

で削除するなど、通常のサービスよりも厳しいルールで運営しております。これによってデマの温床となるような怪文書とか、そういうものは比較的投稿されづらい状態になっておりますし、サービスの現状としても、単に匿名性を悪用した不適切不謹慎な記事というよりも、表現自体は不穏当であってもあえて思考実験や、議論のたたき台として投稿されたであろう記事が目立っていると感じています。

最後、はてなブログですね。はてなブログもトップページは目視でのテーマ選定を基本にしています。よく見られているものが自動的に出てくるという仕組みではないですね。サービスとしては、あくまでダイアリーの延長線として、パーソナルな心情とか経験、自分の物語をつづるような記事を歓迎するポリシーを提示しておりますので、世論をあおるとか、何か危険性を訴えるとか、そういうようなもの、サービスの趣旨にそぐわない記事というのはそもそも選定されないので目立つ場所に出ないというふうな設計にしています。

最後になります。発信者自身が確信的に間違っているというのが実は一番対応が難しいと考えています。これはまだまだ課題があると思います。

今のところ、弊社によってできているのは、例えば反ワクチンの記事みたいなもの、医療とか健康などについての情報に関して、標準的な医療情報とは違う情報が掲載されているときには、このようなポップアップ表示が出るようにしています。これはユーザーが消せな

いのです。なので、この画像は、実際にあったブログなんですけど、割と平和な感じの生活ブログにワクチンは危険だから打ちちゃいけないよみたいなことを強めに訴えている方がおられるというパターンですね。サービスを悪用したいと思っていなくて、この方はみんなのためを思って投稿しているので、それを有無を言わせず削除するような強い対応をしてしまうと、やはり反発されてしまう。さらには、サービス自体が何らかの陰謀によって運営されているという形でかえって陰謀論の思い込みを強化させてしまうということが発生します。

なので、あくまで一般的な情報を提供して誘導を行うという工夫をしております。これは恐らく他社でも行われていることで、様々な形で代替情報への誘導というものがコロナ関係では行われていたと思っています。

そのほかにも、ユーザー間のコミュニケーションによって適切な理解に導かれるようなサービスづくりというのを目指していきたいと考えています。

ただ、課題としては、ワクチンに限らずほかの情報でも、そういう確信的な誤認があったりとか、代々言われ続けているような誤認というようなものがあって、そういうものに関し

ては、その方からすると常識であつてうそではないと思っているので、なかなか対応することは難しいです。

確信的な誤認に至る背景として、インターネット上の誤情報の影響を受けている場合であればリテラシー教育ですとか誤情報対策という今議論されているような領域だと思うんですけども、むしろリアルに周囲の人からの口コミですとか、デマが書かれている書籍とか、偏ったメディア、マスメディアからの誤学習というのがありますね。そのような場合は、インターネットにとどまらない社会的な取組がない限り、インターネットがいくら取り締まっても、「これは本に書いてあった」「知り合いがこう言っていた」みたいな形で、より陰謀論を強固にしてしまうということがありますので、ここも何らかの取組が必要なのではないかと考えています。

さらに、私は実際にエンドユーザー対応を行い、日頃から利用者と接しておりますが、対応の現場では、利用者本人が抱えるいろんな特性、問題が背景にあるのかなと考えられるような事例も多く見受けられ、そこにも課題感を抱いております。

具体的には、情報を適切に読解できなかったために結果的に誤情報を拡散してしまうような方です。例えば「Aはアレルギーがある方が食べると致命的なことになります」といような情報に関して「Aは毒、Aを食べると死にます」みたいな形で理解してしまわれたりとか「Xは殺人的行為です」みたいな暗喩的な表現が理解できなくて「Xで殺された人がいます」と誤解したり、「Xをしていた人を殺人犯だと誹謗している」と誤読したりとか、そのようなふうになってしまう、理解してしまう方がおられます。

さらには、精神的な問題で妄想に基づくデマや反社会的な主張をあえて書いてしまう方、認知に問題があるような高齢者や保護者の管理監督がない子供みたいな年齢的な問題。そのような物事の正誤判断自体が難しい特性とか能力傾向をお持ちの方というのも当然存在します。インターネットは利用者が広がっておりますので、そこにもバリアフリーが必要だと思っております。

現実的な対応としては、利用者のおかれた背景を問わず、投稿の内容が不適切であれば表示抑制を行ったり、ガイドラインの規約で定めるような限度を超えた段階で利用停止という対応をせざるを得ないんですが、やはり本来はそのような方も市民として排除するわけにはいかないですし、サービスの中には包含されるべきでもあるので、そこに社会的な何らかの支援が必要になってくるのではないかと考えております。

私からは以上となります。

【宍戸座長】 大西様、中川様、本当にありがとうございました。手際よく御説明いただいたので、質疑の時間もたっぷりございます。

ということで、ただいまのお二方のはてな様からの発表について御質問、御意見のある方はチャット欄で私にお知らせいただきたいと思います。いかがでございましょうか。

田中先生、お願いします。

【田中構成員】 名古屋工業大学の田中です。貴重な御発表どうもありがとうございました。

後半の中川様の御発表について、大変有意義な情報であると思います。1点だけちょっと御確認というか、質問させていただきたいんですけども、4つの投稿動機による誤情報の分類で、4つ目の確信的誤認というところが個人的には非常に重要なポイントの1つであると思いました。

それで、対応としては、陰謀論を強めないような形でポップアップを提示して、代替情報に導くというような取組をなさっているということと、あと、精神面での問題により妄想に基づくデマや反社会的な主張をあえて書いてしまうとか、認知的に問題がある高齢者、この辺りの方の対応が今後の課題であるというふうなお話でしたけれども、頻度的にはこういったパターンというのはどのくらいあるかというのが、大まかな目安で結構ですので、もし御認識があれば御教示いただけないでしょうか。

【はてな（中川氏）】 そうですね。特にブログなどに関して言いますと、やはり自身の妄想に基づく投稿をなさっている方というのは、割合というような形、数値という形では公開できないんですけども、定期的に一定のパトロールを行っている为数日に一度ぐらいには上がってきてしまうというような、そういうような状況です。

そういうものに関しては、やはり御本人のためにもあまり表に目立つ場所には出さない、出ないよう、弊社のほうで表示抑制の措置、ゾーニングみたいな措置がありますので、そのような対応をしております。

【田中構成員】 ありがとうございました。

【水谷構成員】 関西大学の水谷と申します。貴重な御報告ありがとうございました。どちらの方にお答えいただくかが私も分からないので、取りあえず質問をさせていただきますと、コンテンツモデレーションのはてなさんの対応で、人間のモデレーターさんが、先ほどの説明だと5名ぐらいという感じだったと思うんですけども、偽情報対策だとコスト的にかなり難しいというお話があったかと思います。

そこで対応として、AIとかアルゴリズムによる自動的な対応というのがやっぱり重要になってくるかなと思います。この点で、偽情報対策とはちょっと離れるかもしれませんが、現在ヤフーさんの建設的コメント順位づけモデルというAIが、はてなブックマークさんのコメントのところにも入っているというニュースを見た覚えがあるんですけども、これまでに運用を多分1年2年ぐらいされていると思うんですけども、データでなくても結構ですので、雑感として、実装してみてどうだったかという、感想でも結構ですので、その感触についてお話しいただけると幸いです。

以上です。

【はてな（中川氏）】 ありがとうございます。私のほうから話させていただきます。LINEヤフーさんの建設的コメント順位づけモデルに関しては、はてなブックマークの人気コメントのランクづけに活用させていただいております。はてなブックマークは一応コメントは全員の分が見れるんですけども、人気のコメントだけ別のタブで見れるようになっていきます。従来ははてなスターという評価、いいねの評価ですね、その数を基に人気コメントを選定していたんですけども、数だけですとどうしてもセンセーショナルなのが上に上がってしまうということもありまして、なので、ヤフーコメントの仕組みを応用させていただけるというお話がありました際、ぜひ使わせてくださいということでAPIを使わせてもらったんですけども、感触としては、かなりコメントの治安はよくなったと捉えております。

ただ、通報数という形の指標で見ますと、必ずしもコメントに対する通報が減ったということではなく、不適切なコメントと目されるようなコメントの投稿の数自体が減ったというわけではなく、具体的には見る方からするとやはり気に入らないコメントはたくさんあるというふうな状況が続いています。

ただ、目立つところによくないコメントが上がるという状況は従来に比べてとても少なくなつたというふうな実感を持っていますし、なので、よくなつたのではないかと思います。

【水谷構成員】 ありがとうございます。私もコンテンツ・モデレーションは、大きくに分類できると思っておりますし、よくないコメントとかを消すとか、見えなくするという方法もあるし、何で評価するかは別にして、よいコメントを目立つ位置に、見やすい位置に置くという方法があり、これを併合してやっていただくというのが実は偽情報対策にとっても非常に重要なのではないかなと考えております。特にAIに関しては、こうやってブラッ

トフォーム間で協力があるといいのではないかなと思っていましたので、いい機会で質問できてよかったと思います。ありがとうございました。助かりました。

【はてな（中川氏）】      ありがとうございます。はてなブックマークに関して言うと、何をもっていいコメントとするかという定義もガイドラインのほうに厚く記載しております、それを理解していただくユーザー様が増えることによってよりよい言論空間になるのかなと考えています。

【水谷構成員】      ありがとうございます。

【宍戸座長】      ありがとうございます。次に、奥村構成員、お願いします。

【奥村構成員】      奥村でございます。貴重な情報いろいろありがとうございました。お二人のどちらかに質問で、少し水谷構成員と質問がフォローアップするところもございませけれども、お許してください。

多分Xや何かに比べてはてなで使なされるユーザーの方は滞在時間が比較的長いということや少し考える時間もあるということで、非常にモデレートな空間が今実現できているということがよく理解できました。ありがとうございました。

伺いますが、よい投稿や何かを判別するのに建設的な情報を優先するというので、ガイドラインのほうには5つ、例えば事象をより理解しやすくするとか、記事の内容に関連した体験談とか、そういうものが挙げられているわけですが、これをデジタルのアルゴリズムで実現するというのにはなかなか御苦労が多いものと思われま。そうすると、大体どれぐらいの変数を使って、どのような形で今、モデレーションができていのか、それから、それがどれぐらいで要するにいい投稿を判別できるような機能を果たし得るのか、もしくはどれぐらい改善の余地として課題を感じていらっしゃるのかというような実感について、お答えできる範囲で伺えたらと思いま。ありがとうございます。

【はてな（中川氏）】      よろしくお。コンテンツ、特にはてなブックマークのコンテンツモデレーションの基準に関して言うと、ちょっと秘密要素があるんですね。はてなブックマークはやはり、サービスに記事に対する集客要素みたいなものがあって、ちょっと過剰に最適化されてしまうとよくないということがありまして、アルゴリズム情報に関して詳しいところだと、変数をどれぐらい使っているとか、どのような変数があるかみたいなものというのは、これは外部に出せないというふうなことになります。御承ください。

ただ、基準としては、ユーザー自身の行動、使い捨てのアドレスを使っていないかとか、そういう信用できるユーザーであるかみたいな行動情報というのと、コメントの表現自体

の情報、あと、コメントに対する他者からの評価という情報みたいなものがあって、複合的に考えているとお考えいただければと思います。

もちろん通報情報ですとか、他ユーザから多数非表示にされているとか、そういうような情報というのも見つつ、信頼の置けないユーザーの重みづけを図っていくというふうなつくりにはなっています。

【奥村構成員】 承知しました。秘密の部分が多いということもよく分かりました。そういうようなものが、多分世の中で誤情報とか、そういう間違いを誘発するような情報というのはかなりの分量になっていて、やっぱりそれを世の中でどういうふうに全体でモデレートしていくかということになると、やっぱりテクノロジーをどうやって使いこなすかというところの共有みたいなことが必要になると思いましたが、質問させていただきました。ありがとうございました。

【はてな（中川氏）】 そうですね。明らかに信頼が置けないユーザーというものは恐らくユーザー行動からある程度推測できる要素がある一と思います。それは恐らくこれまでスパム対策などに各社取り組まれてきた、その延長線上ではないかと私は考えています。

【奥村構成員】 ありがとうございました。

【宍戸座長】 ありがとうございます。それでは、生貝構成員、お願いします。

【生貝構成員】 一橋大学の生貝です。貴重な御説明ありがとうございました。はてなは20年来のユーザーでございます。

2点ほど、ざっくりとした御質問になるんですけども、1つは、中川様が御紹介いただいたうち、特にスライドの枚数でいうと25枚目、「サービス設計、コミュニティづくり」の部分、わけても健全な情報流通といったときに、やはりコミュニティの性質ですとか、コミュニティの育て方というのが非常に重要なんだろうなと思っているところがあり、まさにこの中でサービス全域で健全な批判言論を歓迎するポリシーを明確に打ち出す、無責任な書き捨てを許容しないとといったような、こういった全体的な方針を持っているといった時に、こういったまさにポリシーをコミュニティにある種浸透させていく上で重要なことというのがもしあれば。というのも、やはりガイドラインの中でもかなり詳しくその方針を記入している。しかし、読んでくれるユーザーだけではなかなかないだろうなといったときに、そういったコミュニティのルールというものを浸透させる上での努力というものがあればというのが1つと、それからもう一つ、コンテンツモデレーションを様々実施する中で、例えばですけども、Facebookなんかですと、第三者の委員会をつくっ



て、不服があったらそちらでも必要に応じて判断するといったような仕組みなんかも取り入れているSNSなんかもあったりする。

そうしたときに、そこまで当然大がかりなものでもなくとも、まさしくコンテンツモデレーションにユーザーコミュニティであるとか、あるいは外部の第三者の意見や見解というものを反映させていくようなお取組ですとか、そういったものがあれば教えていただきたいと思います。よろしく願いいたします。

【はてな（中川氏）】 ありがとうございます。まず、コミュニティの設計に関して、ポリシーを浸透させるには、恐らくサービスを運営するに当たって、例えば事業全体の収益スタイルとか、メディアとしてこういう情報が目立つとか、そういう全てに対して矛盾がないように設計するというのが大事なのではないかと考えています。

なので、いくらきれいごとを言っても、実際そのサービス内でベネフィットを得ているユーザーというのが、そのポリシーを違反したユーザーの方が目立っていたりとかお金がもうかっているという状態ですと、ユーザーとしても矛盾したメッセージになってしまいますので、いいことをした人にいい見返りがある。そのためにはいいコンテンツが目立つというのが一番ですよね。やはり投稿される方というのは露出が増えるとうれしい方が多いと思いますので。なので、まず、いいコンテンツがたくさん投稿されているというような状況を見せるというのが大事かと考えています。

次に細かいところの説明とか、ユーザーからの問合せに対する対応ですとか、そういうような末端まで同じ考えを基に動いていないとなかなか浸透はしないのではないかと考えますし、そういうことを愚直に繰り返していくうちにやはり理解をされていくかなとも思っています。

モデレーション、第三者の委員会を入れるかどうかという話ですね。これに関して言うと、個別の対応に対して第三者の意見を聞くというよりは、むしろポリシーを打ち出す段階で第三者に意見を伺うということはこれまでもやってまいりました。

例えばはてなブログですね。はてなブログはもともと商用利用を禁止していたんですけども、これはやはりアフィリエイトとかお金が絡むとよくない利用が増えてしまうのではないかと危惧もありまして、健全性を保つために商用利用は禁止というポリシーをしていたんですけど、ちょっとそれが実情にそぐわないということで、健全な商用利用のガイドラインをつくったんですね。

そのときには、日本アフィリエイト協議会さんにルール設計についてアドバイスをいた

だいたりとか、そういうことをしましたし、法的なアドバイスとか顧問弁護士の先生と相談したり、様々な有識者の方と相談しながらポリシーを決めていると、そういうふうな状態です。

【生員構成員】 ありがとうございます。サービスのインテグリティというか、全体としての矛盾のなさをちゃんとユーザーに伝えていくということの重要性含めて大変勉強になりました。ありがとうございます。

【はてな（大西氏）】 すいません、ちょっと私も補足させていただきます。はてなブックマークは性質上、はてながガイドラインを出したりしたときってユーザーさんの反応はかなりいただけるんですね。なので、そういうユーザーコミュニティから生の反応、もっとこうあるべきだみたいな話は随時いただきつつ、我々としてもそれをうのみにするというよりは参考にさせていただきながら、有識者会議とか、そういった形ではないんですけど、ユーザーさんの声のある程度見ながら、反映しながらやっているのかなと思っております。

【生員構成員】 ありがとうございます。特にフィードバックを受け付けやすいアーキテクチャという部分が非常に強いかと思います。

【宍戸座長】 よろしいですか。

それでは、澁谷先生、お願いします。

【澁谷構成員】 澁谷でございます。2点質問させていただければと思います。1点目はヒアリングの資料の11-3のところ、その他、日本国内におけるファクトチェックの推進の観点から行っている具体的な取組があれば教えてくださいという質問で、ユーザーによる自然発生的な対抗言論に期待しておられると書かれておりますが、差し支えない範囲で結構なんですけども、こちらは特に、対抗言論、言説みたいなものが、対抗するものに対して表示するようなことも含めて検討、つまり、アルゴリズム的な、あるいは表示、モデレーションも含めて何か検討されたり期待していらっしゃるのか、それとも、何かプラットフォーム側でモデレーション等を行うというよりも、そういう言説がもう少しユーザー側で発信されることを期待されているのかというところが少し気になりましたので、教えてください。ただければと思います。

【はてな（中川氏）】 対抗言論と申しますか、そのようなものを、特に対抗言論専門のプラットフォームみたいなものをつくるというのではなく、例えばはてなブログですと、世間にあふれるデマ情報みたいなものを検証する方が一定数おられて、またそのようなブログがよく読まれるものですから、はてなブックマークが集まり目立つところに表示される、そういう流れはありますね。

それとあと、はてなブックマークは、その構造上、インターネット上の様々な記事に皆さんの忌憚ない意見を投稿するというプラットフォームですので、やはりそこを見ていただくことによって、大まかにどのような情報がある程度妥当なのかとか、言われている情報はどれくらい信用性が置けるのかということはある程度浮き上がってくる構造になっているとは思っています。

【澁谷構成員】 なるほど、ありがとうございます。2点目続けてお伺いしたいんですけども、何か表示方法やコメントに対する御対応とか、そういった対応や変更を行った際にそれがどのような実際効果があったのかなどに関して、何か実際に御社の中で検証したり、あるいは第三者を入れて検証したりということはこれまでにあった、あるいはこれから何か検討されているということがございましたらぜひ教えていただければと思います。

【はてな（中川氏）】 ありがとうございます。まず、建設的コメント表示アルゴリズムに関しては、導入の際に自社独自のパラメータも加えて調整する必要がありまして、その結果の順位づけみたいなものをパラメータを変えて幾つか出して検証したりとか、そういうことを社内でもかなり大規模な目視でやりました。

ある程度効果が見込まれる。これはあまり定量的なものではなく「これだと快適に情報を見れるね、妥当だね」みたいな、そういうような定性的な判断にはなるんですけども、ある程度の妥当性があると見越されたパラメータでやっているというような状態です。

なので、一応社内での体制、社内での検証体制みたいなものは、少なくともはてなブックマークに関してはあるという状態です。

【澁谷構成員】 分かりました。ありがとうございます。

【宍戸座長】 ありがとうございます。それでは、後藤構成員、お願いします。

【後藤構成員】 後藤でございます。詳しい御説明ありがとうございました。

1つお聞きしたいのは、ユーザーの規模感と、例えば4つの誤情報分類の関係みたいなものがあつたかどうかです。今は登録数が1,000万を超えていると伺っていますが、多分最初はユーザー規模が小さいところから広がっていったと思います。もちろん時代も変わっているとは思いますが、例えば100万人を超えたところから組織的な世論操作が増えてきたとか、嫌がらせみたいなのは人数に関係なくあるとか、そういう概観でも結構なんですけど、ユーザーの規模感と偽情報、誤情報との関係みたいなもの、何かお気づきになったことがあれば教えていただきたいと思います。

【はてな（中川氏）】 大西さんどうですか。

【はてな（大西氏）】　そうですね。あまり規模感でというところはぱっとお答えできるよいお答えはないんですけれども、はてなブックマークというサービスが認知されて、例えば拡散の役に立つみたいなの、例えばニュースサイトとかにはてなブックマークのボタンを置いていただいて、そこからブックマークがつけられるみたいになって、メディアの方がブックマーク数というのを意識していただいたりとか、そういったことで、はてなブックマークで目立ちたいみたいなのが増えてきたみたいなのは、やっぱりユーザー数がある程度増えてきてサービスが人気になってみたいところで、そこはすごく実感したところではあります。

ただ、それを使った偽誤情報の拡散みたいなののがどの規模で感じられるようになったかみたいなのは、すいません、ちょっと我々としても把握しているわけではないので、申し訳ありません。

【はてな（中川氏）】　そうですね。述べましたとおり、まだ弊社ぐらいの規模ですと、そこまで組織的な誤情報の拡散のプラットフォームとして使われるということはあまりまだないのではないかと考えています。

【後藤構成員】　今後発展なされるとそういうリスクも高まってくるという、そういう認識でいらっしゃるでしょうか。

【はてな（中川氏）】　そうですね。サービスが大きくなり、社会的インフラみたいな色が増えてきますと、やはりそれを乗っ取ることによる強みが出てきてしまうと思います。弊社はまだそこまでは至っていないと思います。

【後藤構成員】　ありがとうございました。

【宍戸座長】　ありがとうございます。それでは、越前構成員、お願いします。

【越前構成員】　越前でございます。大変緻密に偽情報の対策を、人手を使って緻密に対応されていることに非常に感銘を受けました。どうもありがとうございます。

私からは、私は技術屋ですので、特に28ページですね。28ページの利用者が情報を適切に読解できず誤情報を拡散してしまうという箇所がございます。こういったものに対して、例えば利用者に何かしら拡散を抑止するような気づきを与えるような取組をすることで、何かしら偽情報の拡散を防ぐような手だてというのは検討されているのか。これを人手でやるというのは相当な手間だと思いますので、例えばそこに技術的なツールを入れてユーザーに提示するような、ユーザー自身が判断いただくと思うんですけど、そういったところの抑止する手段というのは必要性としていかがでしょうか。もし必要であれば、

こういった技術が望ましいかというような御意見があればぜひお願いいたします。

【はてな（中川氏）】　そうですね。これは完全に自分の私的な考えで会社の方針というわけではないのですが、例えばある会話、投稿に対して何らかのリプライがついているとか、元のコメントに対して何らかのコメントがついているというような対話に応じて、そこがかみ合っていないというのは比較的A Iで検出しやすいのではないかと考えていて、文脈の中でことここはつながりがないからここで誤読が発生しているのではないかとというのは検知しやすいのではないかと私は考えています。もちろん人工知能の専門家とかではないので的確ではない考えかもしれません。

対話の接続がおかしいというのを検知して提案できるようなことができると、そこで何か気づきというか、振り返って考えるみたいなことができているかもしれないですね。

【越前構成員】　ありがとうございます。現場での御知見というのは、我々研究者にとって非常に大事ですので、今後ともどうぞよろしくお願いします。我々もそういったことを日々考えながら検討していきたいと思えます。どうもありがとうございました。

【はてな（中川氏）】　最近、SNSとかで元記事を読まずにシェアをすると警告が出たりとかしますよね。元記事を読んでからシェアしてくださいみたいな。そのようなのは完全に全自動だと思いますので。

【越前構成員】　そうですね。フリクションみたいな機能はいろいろ実装されているところございますけど、もう少し高度にいろいろできて、誤った情報拡散をできるだけ技術的な処理で防げないかなというところもちよっと考えておりますので、今後ともどうぞよろしくをお願いいたします。ありがとうございました。

【はてな（中川氏）】　お願いします。

【宋戸座長】　ありがとうございます。それでは、増田構成員、お願いいたします。

【増田構成員】　増田と申します。すごく有益な情報ありがとうございました。

私からは組織的な世論操作がしにくいということの理由として、小さいサービスということが挙げられておまして、それからまた、安易にデマを投稿できない仕組み、そのサービス風土づくりということで、ガイドライン等、あるいははてなさんからの情報発信により、それが参加する方たちにかなり浸透しているのかなと受け取ったんですが、それもまたある程度小さいサービスだからというところが少し理由としてあるのかどうか。そして、やはり同じ思考、考え方、こういうところに参加する人の中でも目的がいろいろあると思うのですけれども、同じ目的に近い人たちがコアに集まっているような傾向があるのかどうか

というようなところをお伺いしたいと思います。

【はてな（中川氏）】 サービスの規模感とかターゲットがほぼほぼ日本国内、日本語圏だけみたいな、そういうことというのは、世論操作みたいなものをやるのに関して、やはり効率が悪いというのは、普通に考えてそうかなと思います。

サービスの規模が小さいというのとそれとターゲットがある程度限られているというのは、それは因果関係、卵と鶏みたいなものもありまして、ある程度長く腰を落ち着けてサービスに投稿したい方ですとか、自分の頭で考えてコメントを投稿したい方ですとか、そういう知的な興味がある方というのは必ずしも大きな層ではないんですね。

ですので、そこは、私のほうから言ってしまうのはどうかと思うんですけども、現状ではやはり限られているがゆえに、安全性も保たれているというふうなところかと思います。

もちろんサービスとしては成長を目指しておりますので、なので、成長してもそれが保たれるような方法というのはこれからちゃんと模索していかないといけないとも考えています。

【増田構成員】 ありがとうございます。

【宍戸座長】 ありがとうございます。構成員の皆様のほうからさらに何かお伺いしたいこととかございますでしょうか。

よろしいでしょうか。

水谷構成員。

【水谷構成員】 すいません。ちょっと時間があるなら追加でお伺いしたいことなんですけども、現時点では多分されてないかもしれませんが、今後の偽情報対策という意味も込めて、ファクトチェックの組織との連携みたいなものというのについては何か考えておられるということがあれば教えてください。

以上です。

【はてな（中川氏）】 そうですね。考えていないわけではありません、もちろん。ただ、現状、連携するのに当たって具体的な方法ですとか、ファクトチェックの団体自体の信頼性みたいなものがちゃんと検証されているのかとか、そういうところで若干まだ不安があるなというふうなのが今のところ。なので、サービス内での検討の俎上にはまだ上がっていないですね。

ただ、具体的にこれはとてもいいと思えるようなものがありましたら、ぜひぜひ積極的に考えたいとは思っておりますが、現状はそのような状態ではないという状況です。

【水谷構成員】 ありがとうございます。私も日本ファクトチェックセンターの運営委員を末席ながらやっております、ファクトチェック組織自体の信頼度を上げないという点は耳が痛い話ではありますが、今後頑張っていきたいと思っておりますので、よろしく願いいたします。

【はてな（中川氏）】 申し訳ありません。期待しております。よろしくお願いいたします。

【宍戸座長】 ありがとうございます。さらに森構成員、お願いします。

【森構成員】 ありがとうございます。時間がまだ少しあるということですので、お願いいたします。どうも御説明ありがとうございました。大変しっかりした対策を講じていらっしゃるというふうに伺っていて感じました。

透明性レポートも拝見いたしました。今後のことということなんですけども、今のところ現状では発信者情報開示請求とか、あと、スパム行為とか、そういうことでの対応の数値をお書きいただいていたけれども、今後のこととして、偽情報・誤情報対策で対応したことを数値化していただくというようなことは考えられますでしょうか。よろしくお願いいたします。

【はてな（中川氏）】 申し上げましたとおり、現場では偽情報か、偽情報でないかという軸で対応しておらず、スパムで偽情報も含むもの、あるいは発信者情報開示や苦情申立てデマを含む、偽情報による権利侵害を含むものみたいな形で対応数に内包されています。あまり偽情報というところにスコープを絞って対応するということがないので、定量的な数値を出すこと自体がかなり難しいのでないかと思っています。

今後、何らかのファクトチェックのシステムを入れるとか、そういうような偽情報にスコープを絞った対策を行うような段階になりましたら、もちろんそこで対応したものというのは定量評価ができると思うんですけど、現状では偽か、偽でないかというふうな軸では対応していないため、ちょっとその辺の数値を出せないかなと思います。

【森構成員】 なるほど、なるほど、ありがとうございます。先ほどワクチンは危険だという投稿についてのお話があったと思いますけれども、ああいうものは今はスパムで対応されているということでしょうか、それとも特に今は対応されていないということなんでしょうか。

【はてな（中川氏）】 それに関しては現状では特例的かつ限定的な対応なので、まだ。あれは確かに偽情報にスコープを絞った対応と言えますが、ただ、ちょっとまだ特例的な対応で、数値を出せるというふうな状態ではないです。

ああいうものに関して言うと、目立たないところで投稿されている分にはあまり問題がないんですけども、それが目立つところに拡散されてしまったりとか、とても読者が多いとか、そういう状況で特例的にそのような表示をして何らかの中立性を保つみたいに対応していて個別対応に近い状態です。なので、あまり定量的には計測していません。

**【森構成員】** 分かりました。ありがとうございました。

**【宍戸座長】** ありがとうございます。それでは、クロサカ構成員お願いします。

**【クロサカ構成員】** クロサカです。御説明ありがとうございました。お伺いしたいことは、例えばマスメディアのような権利者といいますか、情報、ニュースを持っているようなところから、例えばブックマークや匿名ダイアリー等の貴社サービス全般において、この扱いはよくないから対応してほしい、削除してほしいというような問合せとがどれくらい発生しているのか、あるいは、発生している場合、どのような対応されているのか、こういったことをまず一つ教えていただければと思います。

もう一つ、同じ質問を広告主、例えばこの広告がこんなところに配信されているので削除してほしい、というような問合せがあるかどうか、こういったところを教えていただければと思います。

**【はてな（中川氏）】** 結論から言うと、マスメディアから、メディアに対する何らかの言及が問題があるとして削除依頼を受けるということはほぼほぼゼロですね。広告主からも同様です。広告主からこのような記事に広告が掲載されているということで削除申立てを受けるみたいなこともゼロです。

マスメディアからの削除依頼で一番多いのは、掲載期間が終わったとか、誤報で差し替えたので削除してくれみたいなことがたまにあるぐらいだけで、利用者がメディアに対して何らか言及していることをもってメディアが削除依頼をしてくるということは基本的にはないです。

ただ、メディアで報道された人とかが後日削除を申し立ててくるということはあります。それはいわゆる忘れられる権利といったような文脈です。

ですので、メディアや広告主からの削除依頼というのはほぼほぼ受けたことがありません。

**【クロサカ構成員】** 分かりました。ありがとうございます。

**【宍戸座長】** 大体予定の時間でもございますが、ほかに何か御質問、御意見等ございますでしょうか。



よろしいでしょうか。

本日、こういう形でおいでいただいて、大変充実した取組について、非常に歯切れよくどうか、明快に御説明をいただき、また質疑応答にも丁寧に御対応いただき、座長といたしまして大変感謝を申し上げたいと思います。

先ほどいろいろファクトチェックとか、いろんなお話ございましたけれども、引き続き、総務省あるいは関係する民間の組織であったり、他企業を含めて、私自身が関わっているものもございますけれども、意見交換や連携をさせていただければと思いますし、また、この検討会に本日おいでになって、あるいは御覧になっていて、あれ、ここは違うんじゃないかとか、こうしたらいいのではないか、こういう協力だったらできるよということを思いつかれましたら、何なりと事務局ないし私まで引き続きお寄せいただければと思っております。本日は本当にありがとうございました。

【はてな（大西氏）】      ありがとうございました。

【はてな（中川氏）】      ありがとうございました。

【後藤構成員】      ありがとうございます。

【宍戸座長】      また、構成員の皆様も追加で、そういえばあれをお伺いしたかったとか、ユーザーとしてこうであるとか、何かはてな様への御質問とか御意見とかございましたら事務局までお寄せいただければと思っております。

それでは、アジェンダの2番目に移りたいと思います。これは令和6年能登半島地震におけるデジタル空間の偽誤情報の流通状況ということで、まさに現在進行形の課題について、理解を深め、御議論をいただきたいと思います。

まずは、澁谷構成員から御発表いただきたいということでお願いをしております。資料10-2に基づいてということでございますが、どうぞよろしくお願いいたします。

【澁谷構成員】      どうぞよろしくお願いいたします。それでは、私のほうから令和6年能登半島地震時の偽誤情報等の流通状況の速報値的な内容を少し御紹介差し上げて、この検討会での検討事項等について何か反映できることがあればということで話題提供としてさせていただきます。

1点留意点といたしましては、この発表資料に含まれる研究がまだ進行中のものもございまして、まだまだ進めている途中でございますので、確定版ではないというところだけは御留意いただければと思います。

説明資料のほうは少し長くなっておりますので、最初にエグゼクティブサマリーをつけ

ておりますが、本日いただいた時間のほうではエグゼクティブサマリーに書いている事項を中心にポイントだけ押さえてお話を申し上げるという形で進めていければと思います。

まず初めに何をしたかということは、簡単に申し上げますと、能登半島地震のときに偽誤情報等の流通はどうだったんだろうかという全体像を確認してみようということです。特にXの仕様変更等もございましたし、共有される情報なども変化しておりますので、その辺も含めてどうだったんだろうかというところを見たということでございます。

研究方法としては2つございます。両方とも日本語の偽誤情報等を対象としておりますが、1つは、偽誤情報と言われているものとしてどのようなものがあったのかというのをマニュアルで我々のほうで収集し、偽誤情報の類型やトピック、動機・意図について分類をしました。特に動機・意図に関する分類については、もちろん投稿を確認するだけでは100%言い切れるものではないので、あくまでコーダーによる推測による分類です。2つ目は、各プラットフォーム事業者で公開しているデータを用いて量的にどんなトレンドがあったのかというところを概観しました。

では、まず1つ目のほう、そもそも偽誤情報と呼ばれているものにどんなものがあったのかということですが、こちらはマスメディアやファクトチェックによる記事、あるいはそれ以外のブログ等も含めましたウェブ上の記事で偽誤情報等として言及されていたものを我々のほうで96投稿を収集しました。投稿内容の重複を除くと70ぐらいの投稿になっております。収集投稿一覧はスプレッドシートで公開してございますので、もし御関心がある方は見ていただければと思います。参考までに、この方法ではX以外の投稿というのは抽出はされませんでした。

スプレッドシートのほうにそれぞれの類型についてはラベルづけをしたものを載せておりますが、ここでは幾つかの事例について御紹介します。これらの事例はもともとニュース等でも取り上げられておりますので、既に御承知のことかとは思いますが、少しスキップしながら説明したいと思います。例えばこちらは救助要請の複製投稿、いわゆるコピペ投稿で、閲覧数稼ぎ、あるいは場合によっては善意によってコピペをして救助要請を拡散していたというものでございます。

あるいは、こちらは虚偽の救助要請や振込依頼です。金銭的動機に恐らく基づいているであろうと思われるような投稿はこれまでの災害等でも見られましたが、今回も引き続き見られたというものでございます。

それから次は、いわゆる詐称と呼ばれるタイプの偽誤情報で、公式アカウントのなりす

まし、イデオロギー的な動機によると見られる投稿がありました。あるいは個人になりすます詐称では、事故現場にいた、あるいは被災地にいたかのような投稿内容の発信がございました。

それから、こちらは外国人等に関する虚偽捏造投稿です。恐らくイデオロギー的な動機、あるいは、閲覧数稼ぎのためと思われるような外国人による犯罪に関する偽誤情報の投稿が今回も比較的多様に見られました。

そして、これもこれまでの災害でも見られておりますが、人工地震・原発等に関する陰謀論に関する投稿で、こちらもイデオロギー的な動機に基づくと思われる投稿が多様に見られました。

こちらは、恐らく欄閲覧数稼ぎのために、今回の災害とは全く異なる災害に関する、よりセンセーショナルな動画などを投稿するといった事例が多様に見られました。

ここまですと、これ我々のほうでマニュアルで収集類型化をしますと、こちらのマッピング図に示すように、非常に多様な偽誤情報等が今回も確認されましたが、特に金銭的なインセンティブの中でも、閲覧数稼ぎと思われる投稿と、イデオロギー的な動機と思われる投稿の2つの動機・意図に関しては、特に多様なトピック、そして多様な偽誤情報の類型に当てはまる投稿が見られたということが一つ特徴として挙げられると思います。

続いて、2つ目の調査研究として実際の投稿データの分析を通じてわかった今回の地震時のデジタル空間上の情報流通の特徴を御紹介申し上げます。

ここからはXについてフォーカスして申し上げます。Xに関しましては、1月1日から1月3日までのオリジナル日本語投稿で「地震」というキーワードを含む投稿データを分析したものを紹介いたします。こちらの図では、地震というキーワードを含む毎時投稿数を左Y軸に対応する緑の棒で示しています。本震の発災のときに非常に多くの投稿がなされています。

ただ、先ほどご紹介したマニュアルでの投稿収集の中で我々のほうで抽出したキーワードとして「#SOS」、「#助けて」、あるいは「人工地震」、「志賀原発」、それから「窃盗団」等を含む投稿ごとの投稿数を見ますと、特徴的なキーワードとしては、図の右Y軸に対応する青い実線で示す「人口地震」に関する投稿が、特に1日に見られたということが確認できます。Xのトレンドという仕様があり、Xのトレンドに人工地震が1月1日に上がっていたことによる影響もあると考えられます。

続きまして、地震関連投稿群で、投稿ユーザーがどんな言語を使用しているのかというの

を推定しました。投稿ユーザーのプロファイル文から推定ユーザー使用言語というのを推定しております。今回の分析対象のXのデータは日本語投稿でございますので、当たり前と言えば当たり前ですが、日本語使用者と推定されるユーザーによる投稿は地震というキーワードを含む投稿全体の8割を超えています。他方で、偽誤情報関連のキーワードを含んでいる投稿群での推定ユーザー使用言語を見てみます。この図では赤色棒が日本語で、グレー色の棒がそれ以外の言語を使用していると推定されるユーザー数になります。「#SOS」、「#助けて」を含む投稿群では、日本語使用者以外と推定されたユーザーによる投稿が7割ぐらいを占めているという状況が特徴として挙げられます。一方で、「人工地震」、「志賀原発」、あるいは「窃盗団」等に関する恐らくイデオロギー的な背景に基づくと思われる投稿群に関しては、日本語使用者と推定されるユーザーによる投稿がほとんどであったということが分かります。

それから、もう一つ今回の災害関連X投稿での特徴として、いわゆるコピペ投稿がかなり見られたと言われておりますので、そちらについても確認しました。これは、偽誤情報関連キーワード(#SOS, #助けて, 人工地震, 人工的地震, 地震兵器, 志賀原発, 志賀原子力発電所, 窃盗団, 強盗盗賊団)を含む投稿を対象に、投稿内容文の類似度を分析しました。類似度が任意で定めた閾値より高いものは複製投稿の可能性が高いとみなしたところ、少なくとも54パターンの複製投稿が3,938が抽出されました。

非常に興味深いのは、複製投稿の発信ユーザーを見てみますと、複製投稿をしているユーザーで日本語使用者以外と推定されるユーザーによる複製投稿が複製投稿の9割を占めるということがございます。最も複製されていた投稿は日本への地震のお見舞いメッセージで、投稿内容自体は偽誤情報とは言えない内容ですが、こちらの投稿のパターンが少なくとも1,700を超えて複製されています。

そのほか、30回以上複製された投稿と投稿数を示した図を左下に示しております、内容を見てみますと、被害状況や救助要請、お見舞いの内容がよく見られているということが分かります。

また、投稿内容とは無関係に「人工地震」というキーワードやハッシュタグを用いている複製投稿も複数見られました。推測にはなりますが、X上でのトレンドとして「人工地震」上がったことなどをを受けて、閲覧数を伸ばす目的などで、複製する投稿にもトレンドに上がっているキーワードを入れるというようなことも行われた可能性もあると思われれます。

複製投稿のオリジナル投稿は日本語ユーザーによる投稿だったのかということも気に

なります。収集時に削除されている投稿は分析対象外であり、オリジナルがX以外に存在する可能性などもあることからオリジナル投稿を確実に見出すことが難しく参考値にはなりますが、抽出された複製投稿のうち最も早いタイミングで発信された投稿をオリジナル投稿の可能性があるとみなして、それらの投稿の投稿ユーザーの推定使用言語を見ますと、日本語使用者以外と推定されるユーザーは5割にとどまります。

具体的にユーザーの推定使用言語ごとに発信内容が変わっていたのかということも概観しますと、日本語を使用しているユーザーに関しては、緊急地震速報が鳴ったとか、揺れへの恐怖心とか、被災地を気遣う投稿、実際に日本にいる経験に基づくと思われる内容、寄付・義援金、それから、正月でしたので、正月に関する挨拶などがよく見られます。

それに対して日本語ではない言語を使用していると思われるユーザーに関しましては、地震の被害のほか、インパクトのある別の動画、例えば北九州の火災など別のトピックについても投稿していました。

ここまでが投稿そのものについてお話ししましたが、もう一つ、今回の能登半島地震の特徴として、Xの「コミュニティノート」の実装以降、初めての国内の大規模な災害とも位置づけられると思いますので、こちらについて簡単に説明します。

Xの「コミュニティノート」というのは、誤解を招く可能性があるツイートに関しまして、Xのユーザーが背景情報等を追加することができるというものでございます。

コミュニティノートの投稿数のデータを見てみますと、「コミュニティノート」の機能開始から数えて能登半島地震発生直後の日本語のノートの作成数が最大値を記録しております。特に能登半島地震関連だけでも2,700件、羽田空港衝突関連が108件の「コミュニティノート」が作成されています。

ここで作成されたといっても、必ずしもX上で他のユーザーのフィードに表示されるには限りません。また、ユニークユーザーとしては782でした。

「避難」、「原発」、「強盗」、「人工地震」というキーワードごとに「コミュニティノート」の作成数を見てみますと、例えば、比較的人工地震に関する投稿が実際にも多かった時期に即応的にノートがつくられていたことが示唆されます。

以降のスライドのほうでは幾つか「コミュニティノート」の事例についても紹介しておりますので、御関心がありましたら見ていただければと思います。

先ほど、必ずつくったからといって表示されるには限りないと申し上げましたけども、どうやって表示されるのかということに関しましては、Xのほうでアルゴリズムやコードを

公開しております。他のユーザーからの評価と多様性に基づき評価するアルゴリズムになっております。

ただ、コミュニティノートの留意点といたしましては、比較的に関心が高いユーザーによるポストに関する「コミュニティノート」の評価が得やすかった可能性が、今回の地震時にも見られました。

これは実際に2021年のデータを用いた先行研究におきましても、「コミュニティノート」の課題として、様々なユーザーのボトムアップ的な貢献はあるものの、基本的には政治的なコンテンツへのノートの作成が好まれたり、さらにユーザーが自分とは異なる政治的見解を含む投稿に、より情報を追加するといった傾向があるということが示されておりました。偏りがあるということを確認する必要があります。今回の能登半島でもそういった傾向がやや見られる可能性があるということで、この辺りについては今後もう少し調査が必要と思っております。

先ほども申し上げましたが、能登半島が今回「コミュニティノート」の活用がほぼ初めてでしたが、能登半島発生直後に「コミュニティノート」の作成に初めて参加した人の数というのも能登半島地震のときに最大の数を記録しているというのも非常に特徴的な災害であったと考えております。

ここまではXでしたが、参考までにFacebookとInstagram、TikTok、YouTubeに関しても、こちら、取得できるデータが非常に限られておりますので、結論づけたことは言えないんですけども、幾つかポイントとなることをまとめておりますので、スライドを御確認いただければと思います。基本的にはXで特徴的に見られたような「#SOS」のようなキーワードを含む投稿の大幅な増加などは見られませんでした。これまでの災害と同じようにイデオロギー的な動機に基づくと思われる投稿は、Facebook、Instagram、それからTikTok、そしてYouTubeに関しても多様に見られたということは申し添えておきたいと思っております。

では、以上のことをまとめて今後の課題について申し上げたいと思っております。まとめといたしましては、これまでの災害でもそうでしたけれども、災害時の偽誤情報というのはトピックが多様であり尚且つ時系列で変化し、トピックが絶えず変化しますので臨機応変な対応が重要になると考えられます。

また、今回の災害ではXの仕様変更に伴い、閲覧数稼ぎの投稿が多様に見られたほか、閲覧数稼ぎが目的と見られる複製投稿も見られ、さらに日本語使用者以外と推定されるユー

ザーが全複製投稿の91.9%を占めるということが特徴的かなと思います。

また、Xのトレンドとして「人工地震」が取り上げられたということがどの程度情報流通のところで影響を与えたのかというところはもう少し慎重な解析が必要と思っています。

ただ一方で、人工地震を否定する投稿や注意喚起を呼びかける記事も大量にトレンドになったことで取り上げられております。

それから、「コミュニティノート」を見てみますと、一般ユーザーの参加によって災害時に特に即応的に必要なときに偽誤情報関連の修正情報の発信とか注意喚起で一定程度の役割を果たした可能性というのもあると思っています、この辺、もう少し慎重に見ていく必要があると思っています。

ただし、「コミュニティノート」の脆弱性とか限界というのは、海外の例えばウクライナ、それからガザなどの事象においてもそれぞれ指摘がされているところがございますので、例えば外部グループによる組織的な操作に弱いとか、「コミュニティノート」自身が偽誤情報の流通に寄与するとか、それ以外にも、ボトムアップ的なものなので即応的に対応できる一方、政治的なものに寄ってしまうなどのバイアスが指摘がされております。この辺りが日本語の「コミュニティノート」でどうなのか、また、偽誤情報対応のところはどう役に立っていくのかということも慎重な議論が必要と思っています。

最後に、Facebook、Instagram、YouTube、TikTokについては、データに限りがあるため結論づけることが難しいんですが、Xと比べてイデオロギー的な動機に基づく偽誤情報が多い可能性が示唆されております。

なお、Xの「コミュニティノート」データに関しては、XのAPIの利用コストが非常に高くなっている中、「コミュニティノート」に関してオープンデータで、あるいはオープンソースで整備されているというのは歓迎すべきと思っています。

その上で、オープンになっているデータの限界として、実際に「コミュニティノート」の作成対象となった潜在的な偽誤情報と結びつけて議論するようなことは難しいといったこと等がございますので、この辺はもう少しプラットフォーム側で取組に関してのデータや取り組みの公開があると、より効果検証が進むのではないかと考えております。

以上でまとめますと、今回、特に能登半島地震のところに限定してみたわけですが、やはりプラットフォーム事業者の多くでアルゴリズム・システムをどのように設計して管理しているのかについて、限られた透明性しか提供がされていないというのを改めて認識しております。

偽情報等のリスクを特定して対応するためのプラットフォームの取組においても、必要なデータに確実にアクセスできなければ第三者が精査するというのは事実上は不可能であるということを感じております。

データ流通の透明性の確保という観点から、研究機関、あるいは外部団体等の第三者が偽誤情報の実態やその対応策の効果検証を検証できるような形でデータの提供、アルゴリズムの開示が行われるというのが望ましいのではないかと考えております。

以上でございますが、最後にAppendixのほうにつけておきまして、少し参考までにとのことですけれども、Xで今、偽誤情報の話、マイナスのことばかり申し上げましたけれども、少しプラスのことも御紹介申し上げますと、Xの投稿の中で最も共有が多いサイトは圧倒的にヤフーのネット募金なんですね。なので、災害時にプラスの側面、恐らくネット募金、呼びかけとかしましたといったような投稿だと思うんですけども、よいネットワークの効果というのもやはりあるというのは少し最後に申し添えておきたいと思っております。

私のほうからは以上でございます。ありがとうございました。

**【宍戸座長】** 澁谷先生、大変興味深い御報告をありがとうございました。

それでは、ただいまの御発表につきまして、御質問、御意見のある方はチャット欄で私にお知らせをいただきたいと思っておりますが、いかがでございましょうか。

田中構成員、お願いします。

**【田中構成員】** 田中です。非常に貴重で詳細な分析の御報告どうもありがとうございました。大変参考になりました。

1点、人工地震がかなりキーワードとして上がってきていると思われましたので、ちょっとそれについて1点だけ御質問させていただきたいんですけども、スライド番号でいうと21なんですけど、ちょっとその前に1点だけ確認で、縦軸はポストであってユニークユーザーではないという理解でよろしいでしょうか。

**【澁谷構成員】** おっしゃるとおりで、投稿数でユニークユーザーではないです。

**【田中構成員】** ありがとうございます。先ほど考察の部分で、これはトレンドに上がったことが影響してこのような急激な上昇になったのではないかという仮説を御提示いただきましたけれども、例えば人工地震だとか地震雲のようなことはこれまでの災害時でもほとんど必ずと言っていいほど出てくるようなデマの典型例だと思います。例えばそれを逆に悪用というか、活用して、インプレッション稼ぎのためだけではないかもしれませんけれども、それを見越して、ハッシュタグをつけて、それでトレンド入り化させて、それからさ



らに増幅を狙うというようなことも仮説としては成り立つのでしょうかということを1つ。

例えばそれを検証するためにはほかにどういったデータが必要で、今はそういったものが入手しにくいとか、そういったものがあれば御教示いただけないでしょうか。

**【澁谷構成員】** 御質問ありがとうございます。今のおっしゃっていただいた仮説は十分にあり得ると思います。その仮説の検証に関して一番足りないデータは、Xのトレンドに関してどういうものがトレンドに上がっているのか上がっていたのかというデータを取得できないということです。トレンドか投稿数の伸びのどちらが先だったかは我々のほうでは今見るできない状態になっておりますが、検証が必要と考えております。

もう1点目の人工地震はこれまでの災害時にも結構あったということでございまして、おっしゃるとおりでございまして、非常に興味深いのは、Y o u T u b eの詳細分析というのはできないんですけども、少なくとも能登半島地震直後に投稿された動画はそこまで多くはないと考えられます。ただ、過去に投稿された人工地震に関する動画を能登半島発生時にXで引用して投稿される事例も見られました。人工地震を含む陰謀論に関する投稿は平常時も比較的コンスタントに投稿されています。平常時から積み重ねられてきた言説が、何かのきっかけで、より多くの人目に触れるような状態になるというのは例とも捉えることができると思います。この辺はおっしゃるように詳細な解析が必要と思っております。

質問の答えになっているでしょうか。

**【田中構成員】** 大変参考になりました。質問した1つの理由というのが、今後の情報流通の健全性を考えるに当たって、どちらの仮説がより予測できるのかみたいなのを検証すると、恐らく例えばトレンドをつくることができるということが支持されるとすると、今後どういうふうな対策をしていくのかの方向性がまたちょっと変わってくるかと思っておりますので、やはりそういった分析をするためにもデータへのアクセスというものが必要なんだろうなと感じました。どうもありがとうございました。

**【澁谷構成員】** ありがとうございます。

**【宍戸座長】** ありがとうございます。それでは、山口構成員、お願いします。

**【山口構成員】** 山口です。御発表いただきありがとうございます。これは本当大変、大変、大変すばらしい研究で、非常に興奮しながら聞かせていただきました。本当に面白くて、研究途中であるということで、最初に冒頭注意という話がありましたけども、昨日たまたまですけど、災害デマというテーマで朝日新聞に原稿を提出したんですけども、早速修正して、引用したいなと考えております。

私からは3つコメントがございまして、まず1つ目が、今回分析していただいた内容を見ますと、やっぱりXとほかのプラットフォームで結構違うなど。ということ踏まえると、お話しいただいたとおり、やはりインプレッション数稼ぎの意図というものが相当あるのではないかなと感じました。内容を理解せずにコピペしているということもあるでしょうし、日本語話者でない方が多いということで。また、実際災害に限らず、Xのトレンドを見ると、あらゆるXトレンドでコピペが大量にあるんですね。なので、全部そうなんだろうなと感じております。なので、今までとは偽誤情報のステージが変わってきたなと私は感じております。

こういった中でまず1つ目なんですけども、偽誤情報の例えば発信アカウントの収益化の停止とか、規約違反を理由に削除とかということが対策としてはあり得ると思うんですけど、あり得ると思いますし、Xも収益化のところを文章を読むと、例えばセンシティブな出来事は収益化しちゃいけませんと書いてあるんですね。つまりは悲惨な出来事、紛争、集団暴行とか駄目だと。恐らく災害も駄目なんですけど。あともう一つ、虚偽の主張、誤解を招く主張は駄目ですと書いてあるんですね。なので、本当は駄目なはずなんです。でも、恐らく収益化ができてしまっている。なので、全く対策が十分でないということがあると思います。

こういった中で、もちろんXに引き続き要請していくということもあるとは思いますが、ほかに、先生の視点から、政府ができること、あるいは技術的な解決でもいいんですけども、何か対策があり得るんじゃないかというところでお考えがあれば教えていただくと幸いです。これが1点目です。

2点目が、救助要請が今回多かったと思うんですね。救助要請が偽かどうかというのは極めて判断が困難なんじゃないかなと感じておまして、記事の中でも真偽不明の救助要請が云々という書き方が結構多かったと思うんですね。私もそういったものを安易に拡散してはいけませんよという話をこの間もテレビの取材でお話ししたんですけど、そのときにスタジオで、でもそれで実際に助けられた例もあるし、なかなか止まるというのも難しいんじゃないかみたいな話をされている方もいて、私もそう思うんですね。つまり、拡散しては駄目というのは簡単なんですけど、一方で、全員が全員拡散しないということを守ると、SNSの良さまで恐らく殺すことになってしまうというところで、偽の救助要請という、偽誤情報の中でも結構ピンポイント、けどもすごく重要な領域について、先生として何かこういう仕組みがあるといいんじゃないかというのがあれば教えていただくと幸いです。

最後、データ利用というキーワードが出てきておりますけども、Xの「コミュニティノート」のデータの課題とか、API有料化の課題とかというところを御記載いただいておりますけども、もちろん私も完全に同意するところなんですけども、そもそもほかのプラットフォームのデータが極めて収集しにくいという現実があると思います。先生の分析でも、ごく一部しか分析できないと御指摘いただいているとおりで、そもそもX以外のプラットフォームで分析が困難なんですよね。なので、やはりX、そしてX以外のプラットフォーム、全てのプラットフォーム、大手プラットフォームについて、データ活用に前向きになるということが大事なんじゃないかなと。こういった非常に有益な分析ができるということが何よりもその理由になっているんじゃないかなと感じました。3つ目は感想です。

私からは以上です。

**【澁谷構成員】** ありがとうございます。2点目のほうから御回答申し上げます。救助要請に関するものですが、今回の特徴としては、こういったいろんな方、困っている方をリスト化して投稿するようなもので、我々の方でコーディングしても、閲覧数稼ぎと思われるものと、もしかしたら善意なのかもしれないというものもあります。また、災害時はトレンドが常に変わりますので、タイミングよくどこまで対応できるかというのは課題と思います。

一方で、今回の救助要請に関する投稿の特徴として、恐らくコピーのものが多いということで、コピー投稿、複製投稿を防ぐようなシステムというのは一つ有効なのではないかと思えます。少なくとも大量の投稿で情報が錯乱し、救助要請が遅れてしまうということは少し減った可能性はあるのではないかと。複製投稿の増加の背景としては閲覧数インセンティブが大きいと思えますので、そういったインセンティブ設計も含めて取り組むというのは一つだと思っております。

1点目のほうは非常に難しいお話ですので、すぐに回答できるかは自信がないんですけども、例えばXのほうは、Xが変わってから、「コミュニティノート」による偽誤情報対応を重要視しているというふうに彼らの透明性レポートのほうにも書いておまして、なので、アルゴリズムでの制御にフォーカスしていて、これまでやっていたような、例えば人的リソースをしっかりと確保して、ポストをチェックして削除するといったリソースの割り方はかなり減っているようであります。ただアルゴリズム、「コミュニティノート」の機能だけに頼ることの弊害というのはいろんなところで指摘されていて、先ほども少し申し上げました脆弱性の観点から、そもそもどれだけ効果があるのかということに関して指摘がなされております。ファクトチェック団体ではない、あるいはジャーナリズム的なスキル

がある方ではない方の貢献というものについては、即応的な対応が可能などのメリットを活かしつつデメリットをどのように抑制するかは課題が多いと考えております。

以上のことから、やはり一つや少数の対応策での包括的な対応というのは非常に難しいんだらうなと思ってしまして、「コミュニティノート」や、それからファクトチェック団体、それから政府や有識者、あるいは専門家による発信や、あるいはもうちょっと別のモデレーション的なもの、そして人のチェックによる削除とか、いろいろな多面的な取組を推し進めるというのが1つの方向性とは感じております。

回答になっているか分かりませんが、以上です。

【山口構成員】 ありがとうございます。大変勉強になりました。

【宍戸座長】 ありがとうございます。それでは、安野構成員、お願いいたします。

【安野構成員】 中央大学の安野です。本当に貴重なデータで、まずデータを集めるだけでも本当に大変だったと思いますし、記録的な価値もすごく高い、素晴らしい御研究というふうに思いました。御報告どうもありがとうございました。

簡単な質問で恐縮なんですけど、私から2点ほどお尋ねしたいと思います。まず1点目は、誤情報、虚偽情報のタイプ分けとといいますか、カテゴリー分けをなさっていらっしゃるんですけども、特に例えばシェアとかリツイートをされやすい、要するに情報空間の中で共有されやすい誤情報、虚偽情報のタイプというものはあるのでしょうかというのがまず1点目のお尋ねになります。トレンドのデータなんか取りにくい状況で、分かりにくいこともあるかもしれませんが、もしかしたらこれから分析なさるかもしれないのですが、もし既に御存じでしたら教えてくださいということでございます。

次が、日本語と非日本語の投稿の分類というのも非常に面白いというか、重要だと思いついて、これは一部の虚偽情報あるいは誤情報が海外からアクセスされているということと理解していいんでしょうか、それとも国内で何らかの海外向けの混乱情報として出ているんでしょうか。これを2点ほどお尋ねしたいと思います。よろしく申し上げます。

【澁谷構成員】 御質問ありがとうございます。2点目のほうから御回答申し上げます。すいません、ちょっと説明が足りなかったかもしれませんが、基本的には今回分析した投稿は、投稿が日本語の投稿に絞っております。そういう意味では今回分析した投稿というのは基本的には日本語話者の方が見ていると思われるということで、恐らく日本国内で見られた可能性がある投稿がほとんどであると理解しております。

その上で、投稿自体は日本語なんですけれども、ユーザー自体は日本語話者ではない可能

性があるということから、推定ユーザー使用言語というのを推定しまして、これは投稿自体は日本語なんですけれども、投稿しているユーザーのプロフィール文を見てみると、別の言語を使っている方も結構いると。そうすると、その方は恐らく日本語話者ではなくて、プロフィールに使っている言語をふだん使っている可能性が高いだろうということで推定したものが推定ユーザー使用言語でございます。

ですので、可能性としては、恐らくふだんは別の言語を使っているんですけども、例えば閲覧数稼ぎのために日本語で書いているポストで何か拡散、閲覧数が伸びそうなものを、恐らく意味は分からないけれども、コピペをして、貼り付けて投稿しているというのが見られたというのが一つ特徴なのかなと思っております。

1点目のほうが、累計でどれくらい伸びたかということなんですけれども、残念ながら、どれくらいリツイートされたのか、あるいはどれくらい反応が絵荒れたのか、閲覧されたのかというデータは今回アクセスが難しいというところで、そこは一つ限界と考えております。

過去、アカデミック用のAPIが公開されていたときはそういったものもある程度分析が可能だったので、影響力をしっかりと解析する上では、アカデミック等の他の第三者機関に対してデータを開示するというのは今後もう少し必要なかと思えます。

ですので、どの類型がどれくらい伸びたかというのはちょっと申し上げることはできないんですが、肌感覚というか、実際に投稿群をざっと全部見て感じた中で、少し申し上げましたように、被害要請のリツイートが多い事例もありましたし、火災の映像とか、津波の映像など少しセンセーショナルと思われるような動画を含んでいるものはシェアが多かったような印象はございます。肌感覚でございますので、あくまで参考ということで申し上げます。

**【安野構成員】** どうもありがとうございました。すいません、無理なことをお尋ねしてしまいましたが、丁寧な御回答どうもありがとうございました。大変勉強になりました。

**【宍戸座長】** ありがとうございます。それでは、奥村構成員、お願いします。

**【奥村構成員】** ありがとうございます。とてもタイムリーな話題を提供していただきどうもありがとうございました。非常に興味深く伺いました。ありがとうございます。

「コミュニティノート」について伺おうと思ったんですが、詳細なコメントいただきましたので、別のことをお聞きしたいと思います。いろいろなツイートの中で、深刻な拡散とか、もしくは例えばそれで偽の通報があつて消防隊が出動してしまったとか、実際の被害につ

ながってしまうようなものと、それから一定、そういう人工地震の話が流布されているよとか、そういうインプレッション稼ぎの人が偽の被害者の人たちの情報を流しているよといった情報も流れていたというような部分で、多分災害対策をなさっている方はそのようなことを選別しながら行動なさっていたと思いますが、御覧になった中で、実際の例えば救助隊が出てしまって全く分からない場所に行ってしまったとかというような実際の被害につながったものがどれぐらいあったかとか、それから、かなりリポストなどがたくさんあって拡散力が非常に大きいようなインフルエンサーが何かそういうつぶやきをしてしまったものがどんと広がったようなものが、もしマニュアルで御覧になっていたと思いますから、90数件のうちに何かあったかどうかということ。

それから、もう1点は、津波や何かの別のときの情報というようなものは、災害があるときはつきものでして、落雷があると絶対に表れるインドのすごい落雷の画像というのがありますし、銃の乱射事件がありますとアメリカではサム・ハイドという人間が絶対にリツイートされるというようなことがあるわけなんですけど、そういう例えばファクトチェック機関や何かが入り込んでばつとそういうお話を潰すようなことがあると、恐らくこれはそこで途絶えたであろうなと思うような事例みたいなものがあつたのかということについて、ちょっと感触でお伺いできればと思います。データも頂けるといふことでしたので、それを後で検証させていただきたいと思います。

以上です。

**【澁谷構成員】** 御質問ありがとうございます。1点目の実際の被害のところは私たちのほうでも非常に興味を持っているところでして、偽誤情報が実際の実社会のところでもどれくらい影響があつたのかということはいろいろ検証が必要と思っています。例えば救助要請などは、本当に救助を必要とする人の救助が遅れるとつたことが考えられますが、今回の分析はそこまで対象としておりません。ただ、私たちの方でマニュアルで偽誤情報を集めるときに、元の記事を拝見しておりますので、各種主流メディア等でそういったものが報じられているものに関しては把握しているというところで申し上げますと、実際には被害に遭っていないのにX上に投稿されたので何度も救助隊の人が確認しに来て困つたというような事例とか、そういったものがございました。

2つ目のところで、ファクトチェックの介入があつたらもう少し止まつたのではないかとと思われる事例があつたかどうかということですけども、おっしゃるように、災害とは無関係のセンセーショナルなものが使われるというのはこれまでもあつたと思うんですが、

感覚的にはさらに増えている可能性もあると考えられますので、1つ1つをファクトチェック団体の方で対応しても、常にトピックが代わり大量の情報が流通する災害時に、即応的に対応できるかという点、同時並行で進める上では限界もあると思います。それ以外の特にイデオロギー的なものとか、それから詐称のものとかは、やはりファクトチェックのところ为抓手とした記事を出し発信するということで抑制されたものも恐らくあるでしょうし、そういったものが今後も役に立つというのは非常にあるんじゃないかと考えております。

「コミュニティノート」のところでも少し申し上げました。「コミュニティノート」は、即時的なところで少し役に立ったところがある一方、そこでカバーできないというところもあり、そこはまさにファクトチェック団体の方が有用になるところではないかと感じております。

**【奥村構成員】** ありがとうございます。「コミュニティノート」の最大の問題はやはりどなたがやっているかということについて甚だ不透明であるということと誰ができるかという基準みたいなものがやっぱり明らかじゃないというところだと思いますので、その部分のギャップが埋まるような工夫がこれからできるかどうかというところに注目しております。どうもありがとうございました。

**【宍戸座長】** ありがとうございます。お三方手が挙がっていますので、ここから申し訳ありませんが、お一方1問ずつ簡潔にお願いいたします。

それでは、山本健人構成員、お願いします。

**【山本（健）構成員】** 北九州市立大学の山本です。非常に興味深い御報告をありがとうございます。私は、偽誤情報の発信者の動機について非常に関心を持っています。今回、Xの仕様変更で金銭的インセンティブを理由とした発信が増えたということですが、こうした金銭的インセンティブを理由とする発信は今回から増えたのかというのがお聞きしたいことです。今後、インセンティブ構造に基づく対策を考える際に、発信者の動機が何なのか。イデオロギーを動機とするかと金銭が動機であるかによって、対策のつくり方も変わるのではないかと考えています。すいません、よろしくお願いします。

**【澁谷構成員】** ありがとうございます。閲覧数に関しては、少なくとも国内の大規模な災害時には恐らく初めてではないかと思うんですが、もちろん災害時以外の平常時において閲覧数稼ぎのインプレッション目的のものは継続的にありましたが、災害時はさらに情報が増えて、情報の氾濫というか、混乱みたいなところでの影響というのは今回非常に大きかった可能性があると考えております。

【宍戸座長】 ありがとうございます。落合構成員、お願いします。

【落合構成員】 ありがとうございます。私のほうも途中からではありますが、資料を拝見して、大変詳細で、特にXについてかなり突っ込んだ分析をされていることは本当にすばらしいなと思いました。

一つ、今回のこの分析をしていただいて、今後、プラットフォーマーを介した情報流通に関して、どういう情報を理解しているとよりよく比較分析、検討ができるのかという観点でお伺いしたいと思います。Xについては、情報を開示していたということで詳細に分析していただいたというところがあって、それでこういう形で情報が拡散しているということで、むしろある意味さらされてしまっている部分はあるとは思いますが、一方で対応としては、一つ、実際にはいい対応をほかのプラットフォーマーよりされている側面もあるように思い、研究者の先生などに解析していただくという意味では、よい部分があるのではないかなと思いました。だからこそこういう状況にあるという実態が議論できているように思いました。

そういう意味では、情報としてどういうものが分かると分析に役に立つのでしょうか。特に研究者であったり、もしくは、民間のファクトチェックであったり、分析しようとする団体が、どういった構造で偽情報が流れていたりだとか、これが特に危険だとかという点を察知したりだとかができるかです。そういった観点で見たときに、例えば御指摘いただいた中で、インセンティブが金銭面によってというか、閲覧数とある程度連動したような形の収入でという、金銭的な点のお話もあるのだと思います。

あとはまたトレンドに出たりする中で、その記事が見られたりするようになるというお話もあったかと思いますが。実際には投稿者自体と、さらに2次加工とまで言うべきなのかどうかはありますが、プラットフォームにおけるモデレーションと申しますか、何らかの媒介をするときに情報を編集していたり表示方法を変えたりしているという、この部分の情報が一般的には課題がありそうなのかなとも思いました。

ほかのプラットフォームもそうなのかなとも思いますが、ほかのプラットフォームについてはちょっと分からないところもあるということでもありました。こういった点について、情報として認識することが、まずは重要なのだと思いますが、先生のほうでこういった点に着目することでよさそうかという点であったり、分析するに当たってお気づきになっている点があればお伺いできればと思いました。

【澁谷構成員】 ありがとうございます。すぐに回答できるかは自信がないんですけど



も、おっしゃるようにXのほうはある程度データが開示されているからこそできたところもございますし、一方で、災害時、やっぱりXの利用者が一般的に非常に増えるという、増えるというか、投稿も増えるということで、Xを対象としたことは一定程度意味があると思っております。

その一方で例えばT i k T o kはどうだったのかとか、F a c e b o o kの一般のところはどうだったのかみたいなどころはもう少し丁寧に見なくてはいけなくて、そういう意味では、偽誤情報の流通自体はやっぱり完全になくすことはできないので、じゃあ、どういうふうに抑制したりリスクを減らすのかというところを考えたいと思います。やはりアルゴリズムの変更とかシステム仕様の変更によって、どういう仕様のときにどれくらいどういう効果があるのか、あるいはよくないことが起きたのかみたいなどころを丁寧に見いだすという意味では、そういったシステムの設計とかアルゴリズム、これはもちろん全てを開示することは難しいかもしれないですけど、こういったものを用いているのかというところを、データプラスアルファでそういったところを開示していくというのが一つ方向性としてはあるのかなと思っておりますが、まだまだほかにも必要なものはありそうで、もう少しぜひ先生方のお力を借りて検討していくべきところかなとは感じております。以上です。

**【宍戸座長】** ありがとうございます。まさにこの検討会で何が、どういうことができるかとかを、Xさんであったり、先ほど出てきたバイトダンスさん含めて、いろいろお話し合いながらやらせていただければと思います。

それでは、お待たせしました。最後、クロサカ構成員、お願いいたします。

**【クロサカ構成員】** クロサカです。まず御発表ありがとうございました。大変興味深い内容でした。御見解をぜひ教えていただければと思いますが、今回、この研究が偽誤情報の流通状況を調べられているので、今後の計画になるのかもしれませんが、実際に偽誤情報が流通したことに対して、能登半島の方、ないしは日本国民がどれくらい偽誤情報に接しているのか。つまり、実際情報の流通という情報に着目した流通状況と、実際にそれを人間がどれくらい受け止めているのか、受け止めなかったのかというようなことがあるかと思いません。

簡単ではないとは思いますが、何らか調査をする見立てであるとか、方法であるとかということがあるか、あるいはこの辺り、どういうふうに考えていけばいいかとお考えを先生に伺えればと思いますが、いかがでしょうか。

**【澁谷構成員】** 御質問ありがとうございます。非常に大事なところで、ぜひそこは取り

組みたいと思っております。具体的な方法については、正直なところ持ち合わせてないんですが、やはりおっしゃっていられたように、例えば人工地震とか、被災者の方が偽の情報を見て、実際に被害を被ったとか、そういったところ、一つ一つの実際に起こり得るものを洗い出して、それに関するデータを収集するというのは一つだと思いますし、よく言われることですが、いいねとかリツイートとかしなくてもやっぱりポストを見るだけでも影響を受けるというところはあると思いますので、閲覧がどのように行われていて、それによって閲覧した方々へどういう影響があったのかというところは、今回プラットフォームから提供されているデータのみですので、ユーザー側からの何らかのデータというのをしっかり収集しないといけないと思っております、その辺はやっぱり実験的な手法やユーザー調査等と組み合わせる必要があるんだろうなとは思っております。

回答になっているか分からないですけど、以上です。

**【クロサカ構成員】** ありがとうございます。

**【宍戸座長】** プラットフォーム研究会でも、ユーザー側にこういう偽情報見たことありますかというアンケートを前にとったこともありますので、うまく組み合わせられるといいなと思いました。

澁谷構成員、本当にありがとうございました。本当に貴重な御研究だと思いますし、研究が完成されることもそうですが、構成員の先生方も含めて、ぜひそれこそ拡散をお願いできればと思います。

**【澁谷構成員】** すいません。発表が長くなってしまって大変失礼いたしました。申し訳ありませんでした。

**【宍戸座長】** とんでもないです。今回、はてなさんの御発表もそうですし、澁谷先生の御発表もそうですけれども、これまでにない非常に有意義な場になっていると思います。改めて御礼を申し上げたいと思います。

それで、時間を超過しておりますが、若干私から申し上げたいこともございます。今のお話を伺っておりますが、プラットフォーム上で、言わばインプレッション、あるいは広告目当てで投稿を行う。その背景に、山本龍彦座長代理がよくおっしゃる、アテンションエコノミーが背景にある。そしてインプレッション稼ぎを目的に、質の低いコンテンツが発信また拡散されるということが情報空間における情報流通の健全性を考える上で一つ問題になっているということに、やはり踏み込んで考えようと思っております、投稿者に対して広告収益を分配する機能を持ったプラットフォームサービスの在り方について検討を深掘りしてい

く必要があろうかと思えます。

つきましては、議論のために、収益分配の基準など、サービスごとの機能について詳細を1回事務局に整理いただき、それをこの場で共有いただいた上で、さらに議論したり今後のヒアリングを進めるということとしたいと思えます。

ということで、今のような資料を準備いただくということを、会合が間が刻んであるので大変ではございますが、事務局のほうにお願いしてよろしいでしょうか。

【内藤補佐】 事務局でございます。宍戸座長の御指示について承知いたしました。

【宍戸座長】 ありがとうございます。それでは、御準備をいただきたいと思えますし、構成員の皆様で御知見のある方はこういうところを見とくんだよとか事務局にアドバイスをいただければと思えます。また、事務局からアドバイスを求められた構成員の方は、ぜひこれまた御協力をいただければと思えます。

それでは、最後に事務局より何か連絡事項ございますでしょうか。

【内藤補佐】 ありがとうございます。次回の第11回会合につきましては、3月5日火曜、午前10時からの開催を予定してございます。議事など詳細につきましては、別途事務局より御案内を差し上げます。

以上でございます。

【宍戸座長】 ありがとうございました。

それでは、以上をもちまして、「デジタル空間における情報流通の健全性確保の在り方に関する検討会」の第10回会合並びにワーキンググループ第4回会合の合同会合を閉会とさせていただきます。

本日も遅い時間まで、御参集、熱心に御議論いただきありがとうございました。これにて閉会といたします。