【オンラインセミナー】生成AI時代の選挙：民主主義を守るための対策

開催日時：2024年10月21日（月）14:00〜15:30
開催形式：オンライン
主催：国際大学グローバル・コミュニケーション・センター
協賛：日本マイクロソフト株式会社

プログラム（敬称略）：

- 講演①「生成AIで加速するデジタル空間の偽・誤情報に対する政策の現状と方向性」
  吉田 弘毅（総務省 情報流通行政局 情報流通振興課 企画官）
- 講演②「生成AIがもたらすwithフェイク2.0時代の民主主義」
  山口 真一（国際大学GLOCOM 准教授・主幹研究員）
- 講演③「2024年の選挙におけるAI とディープフェイク 〜選挙を守るための対策〜」
  井田 充彦（日本マイクロソフト株式会社 政策渉外ディレクター）
- パネルディスカッション「生成AI時代の選挙を守る」
  クロサカ タツヤ（オリジネーター・プロファイル(OP)技術研究組合事務局長/
  慶應義塾大学大学院政策・メディア研究科特任准教授）
  澁谷 遊野（東京大学大学院情報学環 准教授）
  古田 大輔（株式会社メディアコラボ 代表取締役/
  日本ファクトチェックセンター 編集長）
  井田 充彦（日本マイクロソフト株式会社 政策渉外ディレクター）
  山口 真一（国際大学GLOCOM 准教授・主幹研究員）

概要

日本における偽・誤情報の現状や生成AIの活用状況、そして選挙イヤーとなる2024年の国内外の偽・誤情報の状況を軸に、政府や企業、ファクトチェック団体、アカデミアによる取組と今後の課題を整理した。2024年時点で、生成AIの選挙への影響はほとんど見られないものの、技術発展による偽・誤情報生成や社会的分断等への影響が懸念されている。パネルディスカッションでは、今後に向けてマルチステークホルダによる連携やAIを含めた技術的対応の必要性のほか、日本語環境における偽・誤情報対応の透明性といった課題が提起された。

・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・

講演1「生成AIで加速するデジタル空間の偽・誤情報に対する政策の現状と方向性」（吉田弘毅）

インターネット空間では偽・誤情報が流通するという問題が顕在化している。この要因として、誤情報は事実よりも早く、深く、広範囲に拡散するという特性がある。インターネットにおける偽・誤情報に関して、イギリスの国民投票の際にも偽・誤情報が拡散したほか、2016年のアメリカ大統領選挙では、噂を信じた人が噂されたピザ店に実際に押し入ってしまったという事件も起こっている。また、この偽・誤情報がAIによってさらに加速している。マレーシアで政治関係の事件もあるほか、ウクライナの戦争においてゼレンスキー大統領が市民に投降を呼びかける偽画像も流布した。生成AIは非常にメリットも多いものだが、これによる偽誤情報拡散のリスクが増大している。G7構成国の全ての国が偽・誤情報、情報操作を生成AIに関するリスクとして認識しており、世界においてAIによって加速される偽・誤情報が懸念されている。「グローバルリスク報告書2024」でも、AIによる偽・誤情報の危険性を警告している。

特に今年は選挙イヤーで、インドネシア大統領選挙、台湾総統選挙から始まり、アメリカの大統領選挙まで多くの選挙が行われる。日本でも衆議院総選挙が行われると報道されている。このような中、偽・誤情報の流通・拡散のリスクは増大している。選挙における偽・誤情報やAIによる選挙での利用に関して、2月16日、ミュンヘン安全保障会議において「2024年選挙における AI の欺瞞的使用に対抗するための技術協定」がMicrosoftやGoogleはじめ、世界の20（10月現在27）社によって合意されている。これは民間企業がAI等の欺瞞的利用における対応に危機感を持って、取組を進めているものだ。リスク軽減にかかる技術開発やAI生成物の検知のポリシーのほか、多様な関係者が協力した取組、そして教育キャンペーンについて合意されており、企業におけるこのような取組は非常に好ましいと考えている。日本においてもAIの欺瞞的使用に関して、AIの開発事業者やSNS事業者がミュンヘン合意の趣旨を参考にした対応や、AI以外の偽・誤情報の流布に関しても、すでに利用規約等で対応を定められていると思うが、適切に対応していただけると非常に良いと考えている。

総務省は昨年11月から、偽・誤情報の流通に関して「デジタル空間における情報流通の健全性確保のあり方に関する検討会」を設置した。9月に取りまとめとして、デジタル空間における情報流通の健全性のための方策を公表している。基本的な考え方として、まずは社会全体で取り組もうということ。そして信頼性ある情報を流通促進することと、有害情報を抑制するという両輪の対応をすること。個人のレベルとシステム、サービスのアーキテクチャという技術の両面で対応しようということ。また、偽・誤情報の流通に関しては、デバンキング（偽・誤情報がすでに流通した段階での事後対応）、ファクトチェック等の取組と、

プレバンキング（偽・誤情報が流通する前に注意する）という取組が重要である等の考え方を示したあと、総合的な取組として6本の柱を書いている。1つは普及啓発、リテラシーの向上である。これは政府のみならず、企業、ファクトチェック団体、様々な市民社会と連携をして普及開発を進めていくべきという提言をもらっている。さらに人材の育成、社会全体のファクトチェックの普及、そして技術的対応や国際連携、制度的対応が必要だと述べている。普及啓発に関しては、これから総務省において大きな取組を進めようとしているが、技術研究開発においても支援をしようと思っている。制度的対応に関しては、制度の検討を一番に進めるとともに、情報伝送プラットフォームサービス、SNS等が与える情報流通の健全性の影響の軽減に関して様々な措置を検討するほか、偽・誤情報が流通するときの収益の元になる広告への取組について検討する。この取りまとめを受けて、10月10日から総務省において「デジタル空間における情報流通の諸課題への対処に関する検討会」を設置して、制度的検討や広告に関する諸課題を検討するという体制を整備している。

講演2「生成AIがもたらすwithフェイク2.0時代の民主主義」（山口真一）

2016年は偽・誤情報元年と言われている。米国大統領選挙では多くの偽・誤情報が拡散された。加えて、世論操作やロシアなどの外国からの介入も非常に大きな問題になった。その後、選挙のたびに偽・誤情報は大きな問題になっている。また、それだけではなく、新型コロナウイルスの感染拡大時にも様々な偽・誤情報が拡散し、WHOがその状況をインフォデミックとして警鐘を鳴らした。さらに、戦争や紛争でも様々な偽・誤情報がある。偽・誤動画を作って、それを世論操作に使うという動きも活発になってきている。

日本でも災害時のデマの投稿や感染症関連、政治的なものなど、多くの偽・誤情報が拡散している。例えば能登半島地震の時には、偽の寄付サイトや救助要請など、様々な偽・誤情報が拡散した。また、安倍元首相が台風の時に被災地を訪れて被災者を労っている写真について、これはスタジオで撮影されたものだと主張する人がSNSに現れたこともある。これは

かなり拡散された偽情報で、台湾でも拡散したため、台湾のファクトチェック組織がファクトチェックをしている。さらに、オーストラリアの戦略政策研究所のレポートによると、日本の原発事故に絡んで、中国寄りのSNSアカウントが中長期にわたって情報工作を続けていると指摘されている。すでに日本でもいろいろな偽・誤情報があり、時には海を渡ってしまうし、さらに国内外からの世論操作に偽・誤情報が活用されているという指摘もある。

私はGoogle Japanと長年実施しているInnovation Nipponというプロジェクトで、偽・誤情報の実証研究を行っている。今年発表した最新の研究では実際に拡散した偽・誤情報15件を使って、人々の真偽判断行動や拡散行動の分析をした。その結果、この偽・誤情報を見聞きした後に、その情報を誤っていると適切に判断できている人は加重平均を取ると、14.5%しかいないということがわかった。さらに、51.5%の人は正しいと信じており、かなり多くの人が騙されてしまっていると言える。しかも、10代も60代もほとんど結果が変わらなかった。つまり、偽・誤情報問題は老若男女問わず、あらゆる年代に関わってくる問題であると言える。しかも、偽・誤情報を見聞きした後で拡散するという行為の中で最も多かったのが、家族、友人、知人への直接の会話で直接伝えるというものだ。SNSで見た情報を誰かが食卓で家族に話す、その家族が友人に話す、その友人が今度はSNSに投稿する、こうしてSNSとリアルを行き来しながらどんどん広がっていくのがこの偽・誤情報である。しかも、この偽・誤情報は選挙結果を左右する可能性がある。私が以前行った実証実験では、政治家に不利な偽・誤情報を見ると、その政治家に対して弱い支持をしていた人ほど支持を下げやすいという結果が出ている。

こういった中で、さらに生成AIが偽・誤情報問題に深く関わってきている。つまり、ディープフェイクの大衆化が起こっているのである。誰もが自由に無料で生成AIサービスを使えるようになったために、偽動画や偽画像が投稿されるようになった。すでに日本でも災害とか政治関連のものが拡散されているが、米国大統領選挙では、ハリス候補に不利な偽動画が拡散されている。あるいはバングラデシュの選挙では投票日の前夜に、候補者が立候補をやめると急に言うという偽動画が拡散されるということが起こっている。2022年6月から2023年5月に、少なくとも16カ国で、生成AIが政治や社会問題に関する情報を歪曲するために使用された可能性が高いと指摘されている。注意しなければいけないのが、民主主義国家では人口の5%とか10%といった一部の人の意見を変えるだけで、選挙結果がガラっと変わってしまうことだ。この生成AIを使った偽画像や偽動画、偽音声を使って、誰もが簡単に世論工作できるようになったのが今の時代になる。生成AIによってフェイクがどんどん増加するという時代を、私は「ウィズフェイク2.0」ではなく、さらに時代が一段階変わったと言っている。

では、社会としてどう対処すればいいか。まず政府による制度的な対策が重要だ。しかし、あまりに厳しい規制を敷いてしまうと、表現の自由を損なってしまう。つまり、政府が気に食わないものを「これは偽情報だ」と認定して逮捕するということができてしまうようになる。ロシアやエジプトではすでにそうしたことが起こっている。

そこで、どういったことができるのか、6つのポイントを挙げたい。これは総務省の取りまとめの内容とかなり近い内容になる。ステークホルダー間の連携、国際協力の促進、あるいは現行法で対処できる問題に対する厳格かつ迅速な対処や、プラットフォーム事業者に透明性の確保を求めていくと同時に、適切なコミュニケーションと連携で対策を推進することと民間による対策の支援、最新の技術動向や情報空間の動向を踏まえた有効施策の検討、メディア情報リテラシー教育、AIリテラシー教育の推進、こういったものが欠かせないと考えている。ほかにもディープフェイクは人の目ではどうしても検証しきれないので、判定技術や来歴管理技術で、誰もがディープフェイクであることがすぐにわかる状態にするという技術的な対応も欠かせない。プラットフォーム事業者は、言論を司るプラットフォームとして責任が重い。日本ローカルの透明性の確保のほか、メディア企業やファクトチェック組織と連携して技術提供をする、あるいは研究者と連携した研究をするといったことも欠かせない。また、ファクトチェックの推進も重要だ。さらに、AIによる偽・誤情報のスクリーニングや効率的なファクトチェックのためのツールの開発、実装も大切だ。ファクトチェックは私の実証実験でも、少なくとも半数以上の人が、ファクトチェック結果を見ると考えを変えるという結果が出ているので、かなり重要であるといえる。

大事な点は、この偽・誤情報対策を考えた時に、結局、特効薬がないということだ。根絶は不可能だが、問題を改善していくことはできる。そのためにはステークホルダーが対等な立場で議論する会議体を作り、その中でベストプラクティスや技術を共有するといった連携を進めていくのが大事だ。また、偽・誤情報問題やAIの問題は国内で完結しないので、国際的な連携や情報共有、対策の実施が求められている。昨年のインターネットガバンスフォーラムでは総務省がプラクティスを発表しているが、国際的に連携していくことが欠かせないといえよう。

講演3「2024年の選挙におけるAIとディープフェイク ～選挙を守るための対策～」（井田 充彦）

2024年の選挙におけるAIとディープフェイク選挙を守るための対策として、AI生成コンテンツ、ミュンヘンの技術協定、そして選挙を守るためのマイクロソフトの取組みの3つの話題をお話する。

まず、AI生成コンテンツについてお話をしたい。ある調査によれば、人がAI画像を見分けることができる確率は60%程度という。急速に進化している技術だが、選挙の観点でも機会とリスクの両方をもたらしている。機会としては、選挙活動を強化するということで、生産性を向上させる、有権者との関わりの拡大に使うこともできる。リスクとしては民主プロセスへの干渉ということで、ディープフェイク、サイバー攻撃、有害なコンテンツといった悪用もあり得る。ディープフェイクは偽情報の拡散や詐欺、選挙操作などで使われ得ることになる。ディープフェイクの動画や画像は従来、口や目の動きが不自然とか、照明の当たり方がおかしい、話し方が不自然、指や手の数が多い・少ないといった特徴はあったが、AI技術の急速な発展によって、従来的な特徴を目で見抜くことはだんだん難しくなっていくであろうと思われる。

こうしたAI技術の急速な発展、そして今年、世界中で選挙が行われることから、AIの悪用から選挙を守ることが世界的な課題になっている。そこで、本年2月のミュンヘン安全保障会議において、テック企業が集まって自主的に技術協定を作成し、そこにコミットすることに合意した。選挙においてAIが欺瞞的に使用されるリスクに適切に対処するために、テクノロジ業界として目標を定めてコミットすることになっている。意図的かつ秘密裏に行われるコンテンツの生成及び配布が対象になっており、コンテンツの種類としては候補者、選挙管理人、主要な利害関係者の外見、声、行動などを偽ったり変えたりするようなAI生成の音声、画像、動画になる。また、他に投票に関して有権者に誤った情報を出すといった情報もスコープに入っている。現時点で27のテック企業が参加している。

内容として、3分野で8つのコミットメントがある。まずは「ディープフェイク作成への対応」で、AI生成コンテンツにプロヴェナンスや電子透かしを付加することによりコンテンツの信頼性を向上させる。また、コンテンツ作成ツールの安全アーキテクチャを強化する。例えばコンテンツを作成する時点でプロンプトに悪意のある画像を作成させようとした時に、プロンプトのレベルで止める、あるいは、悪意のある画像や動画を繰り返し作成するようなユーザのアカウントを停止することが含まれる。次の分野は「ディープフェイクの検出と対応」ということで、ディープフェイク拡散を検出して削除などの対応を行うことにな

る。また、テクノロジ業界全体で情報を共有して、ベストプラクティスを共有していく。もう1つは「透明性と対応力」ということで、一般社会へテック企業の対応などの透明性を提供すること、そして市民社会、学者、専門家の皆様方との連携、AI技術の悪用などに関する認知向上を目指すリテラシー教育も含まれている。

こうしたことを受けて、マイクロソフトでは選挙を守るために2つの観点から対応を取っている。1つはサイバー攻撃で、政党や候補者へのサイバー攻撃と選挙運動の妨害への対応である。もう1つはAIを活用した偽情報の作成、拡散への対応である。マイクロソフトは選挙を守るためのコミットメントを表明しているが、これはミュンヘンテック合意よりも前に弊社が自主的に作成してコミットしているものである。1つ目は、有権者が選挙に関する透明で信頼性の高い情報を得られること。2つ目として、候補者が、自分が発信したコンテンツであることを正確に主張できるということ。加えて、選挙中に自分に関する偽動画・偽画像が出回った時にプラットフォーマーに対して削除を申請し、迅速な審査をするプロセスにアクセスできるということ。3つ目として政治家や選挙管理委員会がサイバー攻撃から身を守り、選挙プロセスを確保するためのツールやサービスを利用できることで、サイバー攻撃への対応になる。

そして、AI生成コンテンツの悪用に対抗するために6つのアプローチを定めている。これは選挙に限らず、例えばAI生成のポルノ画像といった他の課題にも対応したものになる。1つ目は、ミュンヘンテック合意と同様の安全アーキテクチャを強化すること、次に耐久性のあるメディアのプロヴェナンスと電子透かしの活用すること。3つ目は、悪意あるコンテンツや行為からマイクロソフトのサービス、お客様を保護するということ。4つ目として、業界、政府、市民社会の方々との連携を図ること、そして5つ目として、テクノロジの悪用から人々を守るために法制度自体を近代化していく必要があるということ。これに対して、テック企業として政府と連携して制度設計に支援、貢献していくことが含まれる。最後に6つ目として一般市民の意識向上とリテラシー習得支援を行うこと。特効薬はないが、ステップ別でのそれぞれの対策があるので、その対策を組み合わせて使っていくことになる。

具体的な日本での取組みについては本年10月10日にブログで発表しているが、基本的には、今申し上げてきたことを選挙に近い形で提供している。「コンテンツ整合性ツール」は政党や候補者が自分のコンテンツに対し自分のものであるという認証情報を付加するもので、また、ユーザの方が弊社のウェブサイトにその画像をアップロードすると、コンテンツ認証情報が含まれている場合にはそれを確認することができる。また、政治家や候補者が弊社の消費者向けサービスで情報を見つけた場合に通報するウェブサイトを準備している。弊社のサービスを使っている選挙管理委員会や政党がサイバー攻撃を受けた時には、最優先で

のセキュリティサポートを提供する体制も整えている。また、検索エンジンのbingで投票方法などに関する信頼性の高い情報へのアクセスを支援している。

その他、国家による影響力工作を監視、分析、公表している。これは、市民社会への認知向上を支える取組となる。マイクロソフトにはマイクロソフト脅威分析センター（MTAC）というチームがあり、世界で30名規模で、日本語を含め15の言語の専門家が各国の影響力工作を監視している。1月の台湾の選挙でのAIの悪用事例を見ると、12月ぐらいからAIニュースキャスターやAI音声でのスキャンダルな動画ニュースが流されている。選挙当日にはAI音声で政治的な影響力の行使を狙った情報発信も行われた。台湾の選挙はAIの悪用が活発になったことが観察されていたが、影響は限定的だったと考えている。また、Microsoft Digital Defense Report 2024を今月15日に発表したが、台湾の事例以外にも中国、ロシア、イラン、北朝鮮の影響工作をフォローしている。例えばロシアが米国選挙をテーマにした複数のニュースwebサイトを立ち上げているが、これはソーシャルメディア上で偽アカウントによって拡散されている。そのニュースサイトに出てくるコンテンツはAIツールを使って作成されていると考えられる。また、他にも今年7月、中国関連の偽アカウントが米国の保守的な有権者を装って作成した欺瞞的なショートビデオを投稿して、それが150万回視聴されたといった事例も見られた。

こうしたAI生成コンテンツの悪用リスクに対処するために、弊社は米国政府向けの政策提言を行っているのだが、日本での対策を考えていく時にも参考になると思う。3分野あり、1つは「コンテンツの真正性の保護」である。デジタルエコシステムにおける信頼性を確保、構築するために、AI生成コンテンツであるという開示を促す、合成コンテンツにメディアプロヴェナンスの使用を要求するといった政策を提案している。2つ目は「欺瞞的なディープフェイクの検出と対応」で、 選挙における欺瞞的なAI、ディープフェイクに対応するための立法を求めている。児童ポルノ、あるいは非同意のポルノ画像に対抗するための法改正や立法も提案している。また、ディープフェイク詐欺に関する法を制定し、詐欺を行う者に対しての法執行を強化することも提案している。その他、AI生成コンテンツの悪用の被害者を調査し、その被害者を救済するための新しい官民連携を形成することや、それに対して政府からの資金提供も求めている。最後は「一般市民の意識向上と教育」で、政府がこの一般市民の意識向上のためのベストプラクティス集を毎年発表することや、合成メディアのプロヴェナンスの研究開発に資金提供することも提案している。こうしたことも日本で今後、皆様方と一緒に議論できればと思っている。

最後に強調したいことは、この問題は１つの会社で解決できるものではなく、適切な法制度の整備やファクトチェックの推進、リテラシーの向上、そして、関係各社の取組を連携し

て、より効果を持たせていくといったマルチステークホルダーの取組が非常に重要である。日本でこの部分は改善していく部分はあろうかと思う。今後、日本で適切なあり方に関して皆さんと一緒に議論していければと思う。

パネルディスカッション「生成AI時代の選挙を守る」

山口：　最初に自己紹介として、各パネリストから簡単にお話いただきたい。

クロサカ：　今日はオリジネータープロファイル技術研究組合（以降OP CIP）の事務局長として出させていただいている。私も関わっていたのだが、2016年の伊勢志摩サミットでは高松でデジタル大臣会合が行われ、この時にすでにAIとどのように向かい合うのかという議論があった。今、生成AIが大きく騒がれている一方で、 技術的な基礎は当時も現在も実は変わらず、ディープニューラルネットワークにある。そのさらに手前のマシンラーニングから新しいパラダイムが始まっているので、今のAIの時代は20年以上も続いている。今起きていることというのは、少なくともエンジニアリングの理解が深い方からすると、そういう可能性があり得るということが多かったかと思う。しかしながら、技術者が考える常識よりもはるかに高速であったり、多様であったり、複雑な形で社会に出現してきている。私は、OP CIPで複雑なデジタル空間の情報流通に対して、技術的な手段ないしはデジタルインフラという形で提供できないかと考えて、様々な方々と協力しながら取組を進めている。

澁谷：　東京大学大学院情報学環で教員をしている。偽・誤情報に関する研究に関して2つの観点から取り組んでいる。1つ目は、大規模なデジタル空間のデータを解析することで、偽・誤情報を含めた様々な情報がどのように流通し、その影響やリスクを把握するという研究である。2つ目は情報を受け取るユーザの側に着目し、大規模なオンライン実験をして、誤った情報を信じてしまったり、どう受け取ったりしているのかということに関して、色々な角度から研究をしている。最近の事例では能登半島地震の時の偽情報とか、SNS上での詐欺広告に関しての調査、そして生成AIによって作られた様々なコンテンツがどのように消費され、人々がどう受け取っているのかということに関する研究に取り組んでいる。

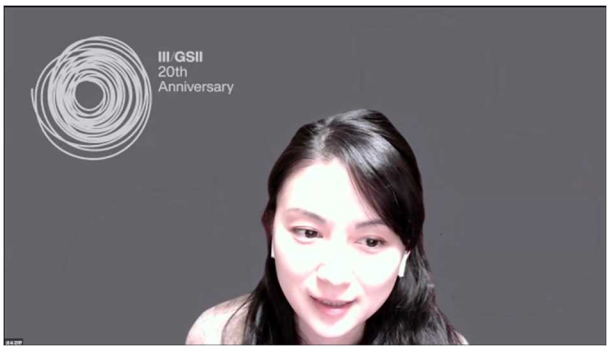
古田：　私は2022年9月から日本ファクトチェックセンターで編集長をしている。2016年、アメリカ大統領選の時に当時、日本版の編集長をしていたバズフィードのアメリカのチームが詳しく偽情報の現状を報道したことに衝撃を受け、Googleを経て、ファクトチェックセンターの設立に加わった。ファクトチェックや偽・誤情報対策の取組が日本は世界的に見ても非常に遅れている。世界的な団体である国際ファクトチェックネットワーク（IFCN）の認証を受ける団体は2022年、日本では我々が設立した段階まで0だった。民主主義国家においてG7では日本が最後で、G20で見ても非民主的な国家と同じようなレベルだった。今回の総選挙絡みの偽・誤情報に関して、すでに12本の検証記事と3本の解説記事を書いている(選挙終了段階で検証28本、解説5本)。短期的に見ると嘘の方が拡散しやすいという面があるが、我々が検証記事を出したものに関しては、検索結果が綺麗になっていくような中長期的な効果も見通した上でのファクトチェックに取り組んでいる。それと同時に、総合的な対策が必要であると考え、ファクトチェックで日々得た知見から具体的なメディアリテラシー教育も行っている。YouTube動画によるファクトチェックの講座や、情報リテラシー教育のための情報発信をしている。それに伴って、ファクトチェッカーの認定試験や講師を養成する講座、それ以外にツールの開発にも協力している。

山口：　テーマ1「生成AI時代における選挙、何が課題か」ということで、具体的にどういったところが問題かを改めてまずは整理したい。

古田：　2024年は世界的な選挙の年で、2023年の段階からIFCNを中心として、世界のファクトチェッカーの間でもこれは大変なことになる、特にAIが脅威であると議論されてきた。すでにいくつかの国では選挙が終わっており、アメリカも終盤戦に来ている。日々、情報交換をしているのだが、蓋を開けてみると、各国に共通して偽情報は大変だが、AIによる偽情報は思ったより少ないという状況だ。アメリカでNews Literacy Projectがレポートで700超の偽情報をダッシュボードにまとめているが、生成AIによるものは7%だった。私自身は実はこうなるのではないかと思っていた。AIは今のところ本物っぽいものを作れるようにはなったが、人の心を理解していないので、人が拡散したくなるようなデマを作るというところまでいっていないためだ。人が作ったほうが、拡散しやすくなるようなデマを作りやすい。面倒なプロンプトを何度も打ち込むより、人が作ったほうが早い。2024年においてはそんなに拡散していない。ただ、そのうちに技術も開発されて、人が拡散したくなるようなデマを拡散するためのプロンプトや技術が生み出されるはずなので、今から準備をしておかないといけない。いずれにしろ質の低い情報が増えることは、情報生態系全体の信頼性を落としていくことになる。現状で言えば、むしろこちらの方が課題としては大きいと思っている。

山口：　偽情報を流す側は目的がある。政治的動機や経済的動機が達成できれば、自分が作ろうとAIに作らせようとどっちでもいい。今はまだ精巧なものとか、あるいは拡散されやすいものがなかなか作れない。しかしながら、おそらく技術が進歩していくと、やがてみんなが拡散したくなる、そして見分けが全くつかないようなものが非常に簡単に作れてしまう。その、コストの閾値を超えた時に大量に発生すると理解している。

澁谷：　今の議論にもあったが、やはり情報がそもそもインターネット上にかなり増えている。インターネット上には基本的に正しいものも偽も両方あることを認識して使う必要があるし、インターネットを避けるわけにはいかない世の中で、どうしたらいいのかということだと思う。生成AIの観点でいうと、ディープフェイクや本物っぽいコンテンツ作成が簡単になるということはもちろん、文章の改善、自動翻訳といった細かなものも含めて日々の生活の中に入ってきて、誰でも簡単に作ることができるということが問題だ。私たちが能登半島地震の時の日本語による偽・誤情報やそれに関連する投稿を調べた際に、いわゆるコピペ投稿と言って、いいね稼ぎのために、他の人が言っていることやバズりそうなものをコピーして大量に投稿するという人がたくさんいた。この投稿者を調べたところ、推定ではあるが、日本語話者以外、おそらく海外にいる人がそうしたものを大量に作っていた。X上ではコピーの8割以上がそういった人から投稿されている。このことは、日本のコンテキストとか日本語といった前提知識を無視しても、何か心に響くものは簡単に作れてしまうことを示唆している。どんどん技術が発展しているので、ますます前提条件を無視したものが作られてくることを踏まえていかなければいけない。もう1点は、選挙に与えるこの影響を理解するために、拡散ベースで言うと少ないかもしれないが、投票行動とか人々の会話への影響に関して、データや研究はまだまだ蓄積がないと感じる。拡散数だけではない評価、特に若者はショート動画をよく見るので、若者に与える中長期的な影響に関して、もう少しじっくりと調査や研究、そしてデータの収集をしていかなければいけないのではないか。



山口：　澁谷先生の分析の日本語使用者以外の調査は非常に興味深く、衝撃的なデータだ。この背景にアテンションエコノミーというキーワードがあると思うが、経済的動機から偽情報を作るということにも生成AIは非常に使われている。選挙のような大きいイベントの場合、もちろん政治的な動機が多いが、同時に多くの人が注目しているのでお金稼ぎにも偽情報が扱われる。経済と政治、両方の面から考えていく必要がある。

井田：　生成AIの選挙への影響に関して、弊社社長が9月にアメリカの上院のインテリジェンスコミュニティで証言しているが、これまで世界の選挙においては、AIを活用して選挙に甚大な影響を与えたと思われる事例は見られない。おそらく、台湾であったように選挙当日に何らかの動画や音声が流されると、ファクトチェックをやっても間に合わないといったことも出てくるかと思われる。引き続き、できることは全部やっていくと良いと思う。また、アテンションエコノミーに関して、弊社ではLinkedInという関連会社がプロフェッショナルを対象としたSNSを運営しているが、LinkedInでは広告のレベニューシェアを行っておらず、金銭的動機に基づきアテンションを得るためのコンテンツの拡散を防いでいる。各社それぞれ自分のサービスに特化したところでのいろいろな対策を取っているので、これをうまく束ねて連携して、合わせて機能していくような仕組みを作っていければと思う。

山口：　今回のテーマ1「生成AI時代における選挙、何が課題か？」と、テーマ2「適切な対策とは何か？民主主義の将来に向けて何が必要か？」についても合わせてお話いただきたい。

クロサカ：　私は7月末からアメリカのジョージワシントン大学とバージニア工科大学の客員研究員を拝命して、ワシントンD.C.に引っ越した。このタイミングで日本にいる時よりも遥かに肌身で、大統領選は全く盛り上がってないことがわかる。断絶がひどいためだ。先日、ハリス対トランプのテレビ討論会があったが、両者が全く噛み合っていなかったことは皆さんもよくお分かりだと思う。ABCテレビが中継し、デビット・ミュアというキャスターが、トランプ氏の発言にファクトチェックをしていた。こうした取組で情報の適正な流通が実現するのではないかと思っていたのだが、トランプ支持者はABCこそ社会の敵、デビット・ミュアこそ我々の敵だと言わんばかりの態度で硬直化し、ハリス側は、あいつらは愚かだというスタンスを打ち出してしまっている。このように融和がないことが構造化される

と、人々は関心を失っていくという状況に入っている。先ほど、生成AIが今のところ選挙に悪影響を及ぼしているというケースは少ないという話があった。

敢えて議論を広げるためにこれを伏線でお話ししたが、AIシステムの範囲をどこまで想定するのか、ないしはその偽・誤情報の生成メカニズムをどの範囲で考えるのか。これによって、もしかすると評価が変わるかもしれない。あるパラメーターを勝手に投入して、生成AIに偽情報を発生させ、それを自動的にSNSに流して人間を混乱させる機械が、そこに人間が触れないものとして存在し、人間が右往左往するということ、こうしたことは、それほど今起きていないと思う。しかしながら、人間が特に偽情報を作るという時、ほとんどの場合、その人は検索エンジンやSNSで集めた情報を整理して作っているはずだし、検索エンジンもSNSも今、AI を使っている。そう考えると、生成AI全般で考えれば、十分我々の社会に一定のダメージを与えているのではないかという視点が出てくる。これが冒頭のAI、特にディープニューラルネットワーク全般でもある。

この「人間も要素として取り込まれたAIシステム」の仕組みは、とても厄介である。カーネマンのいうシステム1、システム2を引けば、システム1で脊髄反射してしまうのはAIではなく人間側であって、機械で構成される部分だけではなく、人間も含んだ様々な社会システムを1つの構造だと考えると、大きな意味でAIシステムを不安定にさせているのは人間なのではないか。こう解釈した方が偽・誤情報の理解は早いのではないか。

我々人間はシステム1で動いてしまっているが、偽・誤情報との対峙にはシステム2を安定的に動かすということが重要になる。この辺りがこのあとのテーマに繋がってくる話だと思うのだが、残念ながら、個人の中にシステム2を動かそうと考えるインセンティブ、または動かさないことのディスインセンティブも十分に持ちきれない。すなわちリテラシーを高めることだけでは解決できない。

さらに、セキュリティの一般的な考え方としても、何らか悪意を持っている人たちは必ず脆弱な部分を突いてくる。民主主義における選挙のシステムは非常に脆弱だが、実際これにしか我々は依拠することができない。社会の中には隙だらけというものがいっぱいある。そう考えると、インセンティブ、ディスインセンティブが個人の側で確立されていない脆弱なところが突かれる。こういう時に、社会システムとしてどう向かい合うのかということが重要になるのだが、それを規律として定める法律の概念が全く追いついていない。これはスピードの問題もあれば、そもそも社会にとっての権利・利益侵害の特定が難しいという背景もある。だからこそ、既存法の活用や自主規制、共同規制が非常に重要になる。これが、オリジネーター・プロファイル（OP）の話でもある。

OPは偽・誤情報を発出することにインセンティブがない、つまり普通に情報を出して健全な状態を作りたいと思っている人たちが、自分たちの考え方を世の中により分かりやすく実現するということを、技術としてサポートしたいと考えている。これは来歴管理だけでは

なく、誰がこの情報を出しているか、あるいはこの情報について誰が責任を持っているのかということをエンドユーザーが検証できるという技術だ。OPを使っていくことで、我々はシステム1に支配され、間違ったアテンションエコノミーの中の道具になってしまうのではなく、我々の主権や尊厳を取り戻すことができるのではないか。人間自身が自分の欲しいもの、正しいと思う世界を作り直すために、こういった規範をシステムとして実現していくといったことが必要になっているのではないかと考えている。

山口：　冒頭の、大統領選が盛り上がってないということに、衝撃を受けた。しかもその背景に分断がある。社会の分断が進みすぎると議論も、合意も、妥協もできないので、逆に皆が関心を失っていくという点は非常に興味深い。日本もそうなってしまわないよう、その辺りの対策は非常に重要だと感じた。続けて井田さんにお聞きしたいことが3つある。1つが、日本におけるマイクロソフトとしての対策、並びに将来的に必要だと考えていること、次に、ご質問として、マイクロソフトのように技術を持つ企業による対策について、対策者の信頼をどう判断したら良いか。もう1つ、国がプラットフォーム事業者に対し指導や注意喚起を行うケースがある、と認識しているが、事業者側の対応が甘い印象がある。この3点についてお聞かせいただければと思う。

井田：　先ほどの分断の話で、マイクロソフト共有分析センター（MTAC）が中国の動きを見ているという話をした。MTACの分析によれば、中国共産党系と思われる偽アカウントが、わざと米国の世論を分断するような質問を米国の人になりすまして投稿し始めて、米国民の反応を探っているのではないかといった動きも見られた。このように、正面から偽情報を流すだけではなくて、分断を煽るという、あまり気づかれにくい形で諸外国の影響といったものはあるのかもしれない。

質問の回答だが、日本と同じような対策はEU議会の選挙でも、イギリス、フランス、インド、台湾でもやっているので、基本的にはグローバルでやっていることを日本でもやっているということになる。日本において、もしローカライズが足りないような点があれば、来年の夏に参議院選挙もあるので改善できるところは改善していきたい。また、対策者の信頼については、一般社会に対して透明性を提供することが基本になると思う。例えばコンテン

ツであれば、どのような削除要請があって何件削除したのか、異議申し立てはどれぐらいあったのかといった情報をしっかりと開示していく。その上で、明らかにおかしな判断と思われるような事例があれば、問題提起いただき、関係者と議論していく。この積み上げによってしか信頼を得ていけないのだろうと思う。3点目として、コンプライアンスは非常に最大限に重視している。基本的に、まずご指導いただくことがないのが一番だが、ご指導いただいた場合にはしっかりと対策を取り、取った対策に対してはしっかりと説明していきたい。

山口： 古田さんに、ファクトチェック組織としてどんな対策をし、これからしようとしているかということ、あるいはそれ以外にも社会としてこんなことが必要ではないかということをお聞きしたい。もう1つ、「米国では、政治家がインタビューなどで虚偽の発言をした場合、それを瞬時に指摘する機能が発達しつつある。日本では、こうしたAIの開発は行われているか」という質問で、AIに限らず、虚偽と判断する基準とか中立性についてお聞かせいただきたい。

古田： ファクトチェック団体として、ファクトチェックだけをしていても絶対にこの問題は解決しない。偽情報対策としての柱の1つとしてのファクトチェックはもちろん続けていくし、強化していく。そして、色々と便利なツールができてきているので、そういったものも活用していく。これが2つ目の質問にも続くのだが、人の発言についてAIでこれが間違っている、当たっているということはまだほとんどできていない。例えば、AI活用で世界的に有名なイギリスのフルファクトという団体の編集長が6月の国際会議で、AIだけでファクトチェックが終わった事例はほとんどない、AIがこれは怪しいかもとサポートし、それを元に人間が検証しないと終わらないという話をしていたが、その通りだと思う。現在、AI活用で画像や動画、AIで生成されていると指摘するようなツールは、どんどん発達している。しかし、例えば生成AIである確率90%と言われても、ファクトチェックで、生成AIで作られたと断言はできない。10%本当じゃないかと言ってくる人たちもいるので、活用の仕方が難しい。ある程度のリテラシーがないと活用できないということになるという難しさがある。ツールは我々として絶対に必要なのだが、ツールさえ開発すれば終わりというわけでもないというのが、まさに我々のようなファクトチェック団体の存在が非常に重要というポイントでもある。

2つ目の日本での政治家の発言のチェックに関してだが、まず、なぜアメリカでは政治家のファクトチェックがあれだけ盛んになったのかというと、端的に言うと、トランプ氏の発言に間違いが多いからだ。他の国を見てもあんな事例はほとんどない。例えば、台湾の総統選の候補者討論会のライブファクトチェックでは、23か所チェックして、7か所は正確という判定だ。残りの16か所も一部間違えたというレベルだ。日本の場合、党首討論会に9人出てきて、1人1人の発言が短かすぎて、そもそもチェックする要素に乏しい。そういう各国

の特徴も理解した上で、偽情報とかファクトチェックの問題に取り組まないといけないと思う。

井田：　おそらくファクトチェックが難しい要因の1つとして、AIが作ったから必ず嘘とも限らないし、人間が書いたから正しいとも限らないわけで、これを技術だけで対処していくというのはかなり難しいと思う。とはいえ、膨大な情報の中から限られたファクトチェックリソースをどこに当てていくのかという判断のためには、検知の段階で、テクノロジーでかなりサポートできるのではないかと考えられるため、そういった使い方でうまく連携するべきだろう。あるいは、ファクトチェック団体の方から、こうした情報提供があるともっとファクトチェックがやりやすいといった話をいただいて、どうすればより良い連携ができるのかということを相談していければいいと思う。

山口：　ファクトチェックをしている人と話をしていると、何をチェックすれば良いかという点が一番大変だという話も伺う。そこに技術を活用できる余地があると感じている。

澁谷：　本日、皆さんがおっしゃっていたが、多面的なアプローチで、誰か単独では解決できないということを改めて強調させていただきたい。それは技術的なところだけではなくて、ユーザ側の話やデザインの話もある。さらに言うと、やはり既存のアクターのネットワークだけでなく、ネットワークそのものをオープンに広げていくこともますます重要になってくる。1人1人がこの民主的な参加を諦めないような環境を、どうやって私たちがデザインしていけるのか。インターネットを諦めない、民主的な参加を諦めない、色々な人々が自分事として参加できるような扉をどのように開けるのか。例えば私だったら研究者で閉じないとか、あるいは他のところと連携するだけではなくて、新しくそういったものに関心を持っている方とか、あるいは若者とか学生、いろいろな方に参加の扉を開くような仕組みも考えていかなければいけないと感じている。もう1点、データ解析などの観点から研究をしていると、やはりDSAのある欧州や米国と比べて、デジタルプラットフォーマーによるデータの開示は、特に日本に関して非常に限られていることを改めて指摘させていただきたい。今、日本で、あるいは日本語で、日本に関連するところでどのような情報がどのくらい流れていて、どのようなリスクがあるのかという全体像をなかなか把握できないという課題がある。さらに、プラットフォームの事業者の取組に関しても、他の第三者機関が検証できるようにオープンにしていくことが、今後も必要になるのではないかと思う。

山口：　やはり細かいプラットフォーム上の動きというデータはなかなか活用ができていないという現状、とりわけ日本ではできていないと感じている。日本語は特殊な言語で、特殊な言語のコミュニティができているので、そこに特化した分析はしなければいけないのだが、英語圏の分析機能が多く、なかなか日本語ではなされていないという課題感がある。そして、透明性もとても重要だ。とりわけ、どうしても外資系のプラットフォーム事業者が強いマーケットなので、日本でどんな対応をして、それがどんな効果があったかといった透明

性をもっと持っていただけると、日本国内の状況の改善には大きく繋がると感じている。パネリストの皆様、本当にありがとうございました。多角的な視点から議論が盛り上がったと思う。

以　上