

# 「プラットフォームサービスに関する研究会」における 偽情報に関する検討

---

2024年3月4日

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  
ワーキンググループ事務局

# 設置（2018年10月）から最終報告書（2020年2月）までの議論①

- 「プラットフォームサービスに関する研究会」（座長：宍戸常寿 東京大学大学院法学政治学研究科教授）は、プラットフォーム事業者が大量の利用者情報を活用してサービスを提供していることを踏まえ、利用者情報の適切な取扱いの確保の在り方等について検討するため、2018年10月から開催（2024年1月に終了）。
- 2018年12月に公表された検討アジェンダでは、プラットフォームサービスの**社会基盤としての役割**や**自由な情報の流通の促進への貢献**に留意しつつ、「**フェイクニュースや偽情報への対応**」を検討する旨明記。

## 検討アジェンダ（2018年12月）

### IoTの普及・進展に伴うプラットフォームサービスの拡大に係る検討を行う背景（P2）

(…）本「プラットフォームサービスに関する研究会」においては、今後の本格的な「データ主導社会」への移行を見据え、電気通信事業及びプラットフォームサービスを取り巻くプライバシー保護（通信の秘密の保護を含む。）やトラストサービス・オンラインニュース配信事業などに関し、今後の市場動向や諸外国のプラットフォームサービスに係る政策動向を踏まえ、現状の課題・論点を抽出し、次の点に留意しつつ、これらへの対応の在り方について検討することとする。

- プラットフォームサービスは**社会基盤として重要な役割**を果たしており、**ユーザの視点から、適正な運営を確保する必要**があること
- 一方で、プラットフォームサービスは**自由な情報の流通の促進に大きな貢献**を果たしていること
- 自由なビジネス環境の実現を通じた**イノベーションの促進と適正なユーザの利益保護のための枠組みとのバランス**を確保すること

## 第3. その他

### 1. オンライン上のフェイクニュースや偽情報への対応（P15）

（EUの政策手段）

フェイクニュースや偽情報の問題は、ボットや閲覧者の嗜好に合わせた広告表示技術を通じた拡散により、年々深刻化しており、我が国においてもこれらに対する対策が課題となっている。

フェイクニュースや偽情報の対応に関し、EU では、2018年4月に欧州委員会が公表した政策文書の中で、今後の活動計画として、同年7月までにプラットフォーム事業者などが参照できる行動規範（Code of Practice）の策定を求め、その後、行動規範の効果測定などの検証を年内に行い、その検証結果が不十分な場合には規制の導入を含む更なる行動を採ることとしている。

（検討事項）

そこで、**EUにおける政策手段を参照**しつつ、プラットフォームサービスにおける**利用者利益の保護の観点**から、次の点に留意して**フェイクニュースや偽情報への対応**について検討する。

- ユーザリテラシーの向上及びその支援方策
- 情報内容の自浄メカニズム
- 表現の自由への配慮

# 設置（2018年10月）から最終報告書（2020年2月）までの議論②

- その後の検討を経て2020年2月に公表された最終報告書では、**偽情報への対応の在り方の基本的な方向性**として**民間による自主的な取組を尊重**するものとされた一方、自主的スキームが達成されない場合は**行政からの一定の関与も視野に入れて検討**を行うことが適当と提言。

## 最終報告書（2020年2月）

### 第2章 フェイクニュースや偽情報への対応

#### 第6節 我が国におけるフェイクニュースや偽情報への対応の在り方

##### 1. 自主的スキームの尊重（P35）

**偽情報への対応の在り方の基本的な方向性**としては、法律による規制、自主的取組とその実効性を支える法的規律を組み合わせた共同規制的な枠組みの構築、業界に対して行動規範の策定を求める等による自主的な対応、個社ごとの自主的な対応など、様々なレベルの対応が考えられる。

この点、第3節における記載のとおり、選挙への影響等により早くから問題が顕在化していた米国や欧州においては、**プラットフォーム事業者による自主的な対応を中心に取組が進められてきた**ところである。一方、一部の国々、具体的には、ドイツ・フランス・シンガポール等の国々では、偽情報の削除義務等を定めた法律が制定されているが、**表現の自由への萎縮効果への懸念**や、**偽情報の該当性判断の困難性による実効性の欠如**及び**恣意的運用への懸念**など、多くの課題や批判の声もある。

したがって、我が国における偽情報への対応の在り方の基本的な方向性としては、まずは**プラットフォーム事業者を始めとする民間部門における関係者による自主的な取組を基本とした対策**を進めていくことが適当である。

政府は、これらの民間による自主的な取組を尊重し、その取組状況を注視していくことが適当である。特に、**プラットフォーム事業者による情報の削除等の対応など、個別のコンテンツの内容判断に関わるものについては、表現の自由の確保などの観点から、政府の介入は極めて慎重であるべき**である。

他方で、仮に民間による自主的スキームが達成されない場合、あるいは、**将来的に偽情報の拡散等の問題に対して効果がないと認められる場合**には、**透明性・アカウントビリティの確保方策**に関して、プラットフォーム事業者に対する**行動規範の策定**や**対応状況の報告・公表**など、**行政からの一定の関与も視野に入れて検討を行うことが適当**である。

### 4. プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保

#### （1）プラットフォーム事業者による偽情報の削除等の適切な対応（P38）

偽情報への対応として、偽情報の生成・拡散を直接的に防ぐ観点から、問題となる様々な種類の情報やアカウント（ボットアカウント）の削除等を行うことが考えられるが、これらの対応については、各プラットフォームサービスの性質や考え方に応じたポリシーに沿って、プラットフォーム事業者自身が自律的に判断して対処することが期待される。

この点、これらの個別のコンテンツの内容判断に関わる対応については、表現の自由の確保の重要性に鑑み、過剰な削除などの問題が生じないようにする観点から、政府の介入は極めて慎重であるべきである。

#### （2）プラットフォーム事業者による透明性・アカウントビリティの確保（P38）

利用者が安心・信頼してプラットフォームサービスを利用することができるよう、上記（1）で記載したプラットフォーム事業者による自律的な情報の削除等の対応に加えて、それらの取組が適切に行われていることが利用者や社会全体に対して明らかにされることが望ましい。また、利用者の表現の自由を確保する観点から、プラットフォーム事業者によって過剰な削除や不当なアカウント停止等の行き過ぎた対応が行われていないかという点についても明らかにされることが望ましい。

以上の観点から、プラットフォームサービスの提供に当たって、いわばプラットフォーム事業者によるガバナンスが適切に機能しているか否かについて、利用者や社会全体が把握することができるようにすることが重要であることから、プラットフォーム事業者は、自らの取組の透明性やアカウントビリティを確保する方策についても、上記の削除等の対応と同時に積極的に取り組むことが適当である。

〔略〕

以上のとおり、プラットフォーム事業者が自主的に取組を実施し、それらの取組に関する透明性及びアカウントビリティの確保を図るとともに、プラットフォーム事業者自身による対応状況等の公開・説明やフォーラムの開催等を通じて、国民（利用者）やメディア等に対して取組の効果や課題などが明らかになることで社会全体としてのモニタリング機能が果たされ、それらの反応を踏まえてプラットフォーム事業者による更なる取組が進められていく、というサイクルが回っていくことが期待される。

また、政府は、こうしたプラットフォーム事業者による透明性やアカウントビリティの確保方策の状況について、フォーラム等の場を通じて随時適切に把握することが適当であり、仮にこれらの自主的スキームが達成されない場合、あるいは将来的に偽情報の拡散等の問題に対して効果がないと認められる場合には、透明性・アカウントビリティの確保方策に関して、プラットフォーム事業者に対する行動規範の策定や対応状況の報告・公表など、行政からの一定の関与も視野に入れて検討を行うことが適当である。

# 緊急提言（2020年8月）から第二次とりまとめ（2022年8月）までの議論①

- 最終報告書とりまとめ後、偽情報の問題についてプラットフォーム事業者の取組が十分かどうか、官民の取組が適切に進められているかどうか等について、**プラットフォーム事業者からのヒアリング等を通じてモニタリングと検証評価を実施**。2021年3～5月及び2022年3月の**モニタリング・検証評価**の結果を中間とりまとめ及び第二次とりまとめとして公表。
- 特に**第二次とりまとめ**では、**透明性・アカウントビリティ確保**をはじめ、**プラットフォーム事業者に期待される役割、情報流通の適正化やユーザの表現の自由の確保についての一定の責任**が改めて確認。
- その上で、プラットフォーム事業者による**偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は限定的**とのモニタリング・検証評価の結果（⇒スライド14）とともに、**行政からの一定の関与を具体的に検討**することの必要性が提言。

## 第二次とりまとめ（2022年8月）

### 第1部 誹謗中傷や偽情報を含む違法・有害情報への対応について

#### 第3章 今後の取組の方向性

##### 3 透明性・アカウントビリティ確保の重要性、枠組みの必要性とその方向性等

###### （1）透明性・アカウントビリティ確保の重要性（P92）

プラットフォーム事業者のサービス上では、**多くのユーザによる自由な情報の発信・受信を可能**としている一方で、**誹謗中傷や偽情報といった違法・有害情報も多く流通**している。プラットフォーム事業者は、**問題となる投稿の削除やそのような投稿を行ったアカウントの凍結・停止、投稿に注意喚起を促すラベルの付与、表示順位の低下等**といった、**ポリシーにあらかじめ定められた違法・有害情報などの流通を抑止するために講じる措置を実施**するなど、**情報流通の適正化について一定の責任を果たすことが期待される**。一方で、プラットフォーム事業者は、ユーザの表現を預かる立場でもあり、**ユーザの表現の自由の確保について一定の責任を果たすことが期待される**。

また、プラットフォーム事業者には、**自社サービスの特性にあわせて誹謗中傷等の情報がユーザやユーザ以外の者に与えるリスクを分析・評価**した上で、**文化的、社会的、政治的背景を踏まえた、投稿やアカウントに対する措置の実施に係るポリシーの設定、その適切な運用**、その運用に必要な体制の構築をはじめとする**リソースの確保**及び自社サービス上の投稿に係る発信者情報開示などの**法的手続への適切な協力**などが期待される。さらに、投稿やアカウントに対する措置については、**削除以外の手法による対応（例：投稿に対するラベルの付与や表示順位の低下、投稿時の警告表示等）**も含め、**事業者の自律的な創意工夫による対応が行われることが望ましい**。

# 緊急提言（2020年8月）から第二次とりまとめ（2022年8月）までの議論②

## （1）透明性・アカウントビリティ確保の重要性（続き）

当研究会におけるヒアリングを行ったプラットフォーム事業者においては、誹謗中傷や偽情報を含む違法・有害情報への措置を講じる必要性が認識されており、あらかじめ対応方針や基準となるポリシーを自主的に設定し、投稿やアカウントに対する措置が行われている。こうした措置については、**ポリシーが適時適切に定められるとともにポリシーに基づく措置の対象となる投稿やアカウントに対して確実に措置が行われることが望ましい一方で、行き過ぎた措置や恣意的な措置といった不適切な運用によってユーザの表現の自由が損なわれることがないよう、過不足なく実施される必要がある。**

**違法・有害情報への対応が適切に行われるとともにユーザの表現の自由に対する過度な制約とならないよう、過不足なく行われるためには、ポリシーの設定状況やその運用状況、対応の結果や異議申立ての機会の確保状況といった項目に関する透明性・アカウントビリティを確保し、「言論空間のガバナンスに対するガバナンス」、すなわち、プロセスの透明性を確保することが必要である。**

大規模なプラットフォームサービスの提供者は、そのサービスの提供により**情報流通について公共的役割を果たしている**と考えられることから、**当該サービスのユーザ及びユーザ以外の者に対して、その透明性・アカウントビリティが確保されることが必要**である。

また、プラットフォーム事業者における透明性・アカウントビリティの確保に当たっては、サービス上における、例えば、誹謗中傷の発生件数等の**流通実態やその抑制のための対策とその効果に関する総量的な数値等の把握という全体的な傾向に関する観点と、個別具体の誹謗中傷等の違法・有害情報に対する権利回復のための裁判手続への対応や、申請にもかかわらず十分に措置が行われないと考えられるケースや自身の投稿について行き過ぎた措置が行われたと考えられるケースが発生した場合の反論や異議申立ての機会の確保という個別具体の観点の両面から、ユーザ及びユーザ以外の者からの透明性・アカウントビリティの確保が必要**である。

こうした、プラットフォーム事業者による**投稿の削除やアカウントの停止等の措置に関する透明性・アカウントビリティを確保**することは、当該サービスの**ユーザ及びユーザ以外の者による客観的な根拠に基づく批評を可能**にし、こうした批評がプラットフォームサービスの運営にフィードバックされることを通じて、**投稿の削除やアカウントの停止等の措置の運用の改善**につながることを期待される。

さらに、プラットフォーム事業者による、透明性・アカウントビリティが確保されることは、より多くのユーザに最新の技術やサービスを柔軟に取り入れたサービス提供を目指すプラットフォーム事業者にとって、**サービスの設計や運営上の創意工夫に対するユーザからの信頼性の向上**につながるものと考えられ、このことは、プラットフォーム事業者にとっても経済的合理性を有する取組になるだけでなく、**ユーザが最新のサービスの利益を享受しながら、リスクを理解した上で、安全・安心にサービスを利用することが可能な環境の確保**につながると考えられる。

# 緊急提言（2020年8月）から第二次とりまとめ（2022年8月）までの議論③

## （3）偽情報への対応に係る透明性・アカウントビリティの確保について（P96）

偽情報は、その外延や個別の情報が偽情報であるか、また、その流通による我が国における影響について、モニタリングからも十分明らかではなく、我が国における実態が未だ明らかではない。そのため、プラットフォーム事業者は、引き続き、実態を把握しその結果を公表することが求められるとともに、プラットフォーム事業者に限らないメディアも含めた産学官民の社会全体で検討する環境が整備されることが必要である。

一方、既にプラットフォーム事業者では、偽情報等の不適切な情報への措置を講じる必要性が認識されており、あらかじめ対応方針や基準となるポリシーを自主的に設定し、投稿の削除やアカウントの停止等の措置を行っている。こうした措置については、措置の対象とされるべき情報に対して措置が確実に実行されることが望ましい一方で、行き過ぎた措置や恣意的な運用によってユーザの表現の自由を損なうことがないよう、過不足なく実施される必要がある。そのため、プラットフォーム事業者は、偽情報についても、我が国において生じている自らのサービス上の偽情報の問題について適切に実態把握とリスク評価を行った上で、そのリスクに応じて偽情報への対応を行うとともに、その透明性の確保を進めていくことが求められる。

しかしながら、プラットフォーム事業者による我が国における偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は限定的であり、偽情報に対して適切なリスク評価や低減措置が行われているか十分に明らかではなかった。

以上を踏まえ、プラットフォーム事業者において、違法・有害情報となり得る偽情報への対応については、本章1の違法・有害情報対策の方針も踏まえて、透明性・アカウントビリティが確保されることが必要である。

総務省は、違法・有害情報となる偽情報に関するプラットフォーム事業者の取組状況について、前述の違法・有害情報対策に関する記載内容を踏まえて、偽情報への対応に関する透明性・アカウントビリティの確保に向けて、行動規範の策定及び遵守の求めや法的枠組みの導入等の行政からの一定の関与を具体的に検討することが必要である。

なお、行政は、引き続きプラットフォーム事業者等による自主的な削除等の対応を促進することとし、プラットフォーム事業者等に対して削除義務を課すことや、個別のコンテンツを削除しなかったことに対して罰則等を設ける法的規制を導入することは、誹謗中傷の場合と比べても、極めて慎重な検討を要する。

また、総務省は、これらの取組に関するモニタリングと検証評価を継続的に行っていくことが必要である。この際、プラットフォーム事業者に対して具体的にどのような対応や情報公開を求めることにより、偽情報への適切な対応が図られているかどうかを評価することが可能かについて、引き続き検討が必要である。

## （4）その他プラットフォーム事業者に期待されること（P97）

上述の検討のほかに、インターネット上の言論空間には、違法情報でも有害情報でもない情報であっても、プラットフォーム事業者によるレコメンデーションによってユーザが無意識のうちにフィルターバブルに閉じ込められることやエコーチェンバー効果が発生することで、攻撃的な傾向への誘導やフェイクニュースの拡散、社会的分断を生じ得るとの指摘がある。

これを踏まえ、デジタル・シティズンシップの考え方も参考に、ユーザ自らが望ましいと判断する情報環境を選択するための環境整備、すなわち、個人がテクノロジーを通じて自身が触れる情報の自律的なコントロールを可能にするための環境整備が重要である。大規模なプラットフォームサービスが情報流通について公共的役割を果たしていることから、ユーザ自身が望ましいと判断する情報環境を選択するための環境整備について、利用者とコミュニケーションを図りながら、一定の役割を果たすことが期待される。



# 第三次とりまとめ（2024年2月）までの議論①

- 第二次とりまとめの後に行われた検討（2023年2～4月のモニタリング・検証評価を含む）の結果を踏まえ、2024年2月に**第三次とりまとめ**を公表。
- モニタリングの結果として、プラットフォーム事業者による我が国における**偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は未だ限定的**（一部事業者については**後退**）と提言（⇒スライド17）。
- また、透明性・アカウントビリティの確保として、**モニタリングの継続的な実施が適当**であることが提言。

## 第三次とりまとめ（2024年2月）

### 第2部 偽情報への対策について

#### 第1章 現状と課題

##### 1 プラットフォーム事業者による対応のモニタリング結果

###### (2) モニタリングの結果

###### ア 総論（P25）

Twitterを除く全ての事業者において、我が国における偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は、2022年3月28日に実施したプラットフォームサービス第34回におけるヒアリング（以下「前回ヒアリング」という。）と比較して、**一部で進展が見られるもののほぼ同等であり、未だ限定的**である。特に、Twitterからは、**口頭での発表が行われたものの、ヒアリングシート及び説明資料の提出がなく、透明性・アカウントビリティ確保の取組について後退があった。**

#### 第2章 今後の方向性（結論）

##### 1 モニタリング（P48）

本研究会では、プラットフォーム事業者による偽情報への対応の実施状況についてモニタリングを行ってきた。プラットフォーム事業者による投稿の削除やアカウントの停止等の措置に関する**透明性・アカウントビリティを確保**することは、当該サービスの**ユーザ及びユーザ以外の者による客観的な根拠に基づく批評を可能**にし、こうした批評がプラットフォームサービスの運営にフィードバックされることを通じて、**投稿の削除やアカウントの停止等の措置の運用の改善**につながることを期待される。したがって、こうした**モニタリングの取組については、継続的に実施していくことが適当**である。

なお、前述のとおり、2023年のモニタリングにおいて、Twitter（現X）からは、再三の求めにもかかわらずヒアリングシート及び発表資料が提出されなかった。**任意とはいえ、資料が提出されなかったことは遺憾**である。

# 第三次とりまとめ（2024年2月）までの議論②

- インターネット上の偽情報の生成・拡散やプラットフォーム利用者の情報に対する選択的接触の問題について、**アテンション・エコノミーを構造的要因とする場合**も含め、「**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**」における議論の深化に期待と提言。
- また、プラットフォーム事業者において、API開放等により、研究者が情報空間に関する実データを入手しやすい環境が整備されることに期待と提言。

## 2 今後の更なる検討（P48）

本研究会では、派生的論点として、フィルターバブル、エコーチェンバー及び分極化等の現象についても検討した。こうした現象が偽情報の拡散に寄与しているか否かは、計算社会科学等の学術分野における更なる研究が期待されるところであり、本研究会において結論づけることができるものではない。もっとも、第44回会合及び第46回会合において有識者から指摘された点を踏まえると、**レコメンデーションに関するアルゴリズムの公開やリテラシー教育等の方法により、利用者が情報に対して選択的接触を行っていることを、当該利用者に対して認知させることが重要**である。

加えて、近時は、**生成 AI やメタバース等の新たな技術・サービスの出現によりデジタル空間が更に拡大・深化**している。このような中、**インターネット上の偽情報の生成・拡散やプラットフォーム利用者の情報に対する選択的接触の問題**は、**アテンション・エコノミーを構造的要因とする場合**を含め、プラットフォーム事業者だけでなく、生成AI事業者、仮想空間関係事業者、通信・放送事業者、利用者等の多様なステークホルダーが連携・協力して対応すべき、デジタル空間における情報流通の健全性に関わる課題の一つと言える。

総務省は、生成AI等による巧妙な偽情報の生成や拡散に伴う社会的な影響の深刻化を含む、デジタル空間における情報流通を巡る新たな課題と多様化するステークホルダーによる対応等の現状を分析し、デジタル空間における情報流通の健全性確保に向けた今後の対応方針と具体的な方策について検討するため、2023年11月より「**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**」（座長：**宍戸常寿 東京大学大学院法学政治学研究科教授**）を開催し、検討を継続している。**今後、インターネット上の偽情報の生成・拡散やプラットフォーム利用者の情報に対する選択的接触の問題については、以上述べた観点を踏まえ、当該検討会において議論を深化させていくことが期待される。**

なお、フィルターバブル、エコーチェンバー及び分極化等の現象を分析する研究を含むデジタル空間における情報流通に関する研究においては、**実データの入手が重要**であると考えられる。プラットフォーム事業者においては、**APIの開放等により、研究者が情報空間に関する実データを入手しやすい環境を整備することが期待される。**

## 1. 前提(サービス概要)

- ① 対象サービス名・サービス分類
- ② 加入者数、月間アクティブユーザ数又は書き込み数

## 2. 「我が国における実態の把握」関係

- ① 偽情報等の発生・拡散状況を把握できる体制 分析・調査の有無
- ② 日本における偽情報等の発生・拡散状況、結果公表
- ③ 研究者への情報提供、利用条件
- ④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等)
- ⑤ 不正な申告や削除要請への対策の方法・仕組み、対応件数

## 3. 「多様なステークホルダーによる協力関係の構築」関係

- ① 産学官民の協力関係の構築
- ② 具体的な役割
- ③ 議論を踏まえた取組

## 4. 「プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保」関係

## (1) 偽情報等に関するポリシー

- ① (i)禁止行為 (ii)削除等の対応

## (2) 削除等の対応

- ① 偽情報等に関する申告や削除要請の件数
- ② (i)偽情報等に関する申告や削除要請に対する削除件数、(ii)アカウントの停止数
- ③ 偽情報等に関する主体的な削除件数(AIを用いた自動検知機能の活用等)
- ④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等)
- ⑤ 削除以外の取組 (i)警告表示 (ii)表示順位の低下 (iii)その他の取組内容
- ⑥ 不正な申告や削除要請への対策の方法・仕組み、対応件数

## (3) 削除要請や苦情に関する受付態勢・プロセス

- ① 一般ユーザからの申告・削除要請への受付窓口・受付態勢、対応プロセス
- ② 対応決定時における通知の内容、理由の記載の程度
- ③ 一般ユーザからの申告や削除要請に対応する部署・チームの規模・人数
- ④ ③の部署・チームに関する日本国内の拠点の有無、日本における責任者の有無
- ⑤ 削除等への苦情や問い合わせに対する苦情受付態勢及び苦情処理プロセス

## (4) 透明性・アカウントビリティの確保

- ① コンテンツモデレーションのアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照
- ② 透明性レポート 日本語で閲覧可能か
- ③ 取組の効果分析

## 5. 「利用者情報を活用した情報配信への対応」関係

- ① 広告表示先の制限
- ② 広告出稿制限
- ③ ターゲティング技術の適用に関する規定
- ④ 広告のアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照
- ⑤ 出稿者の情報や資金源の公開
- ⑥ 広告とコンテンツの分離
- ⑦ その他の透明性・アカウントビリティ確保方策 ユーザへのツール提供

## 6. 「ファクトチェックの推進」関係

- ① ファクトチェック結果の表示 具体的な仕組み・基準
- ② ファクトチェックを容易にするツールの開発及び提供
- ③ ファクトチェックを実施する人材の育成
- ④ ファクトチェック機関との連携

## 7. 「ICTリテラシー向上の推進」関係

- ① 普及啓発の取組・投資
- ② 他のステークホルダーとの連携・協力・投資

## 8. 「研究開発の推進」関係

- ① AI技術に関する研究開発
- ② 「ディープフェイク」対策の研究開発

## 9. 「情報発信者側における信頼性確保方策の検討」関係

- ① 信頼性の高い情報の表示
- ② ニュースの選別・編集に関する透明性・アカウントビリティ確保方策
- ③ メディアとの連携体制構築 具体的検討・取組
- ④ 情報源のトレーサビリティ確保、なりすまし防止・認証

## 10. その他

- ① 意見・補足

- 全体的な傾向として、プラットフォーム事業者の偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は限定的。
- 他方、多様なステークホルダーによる協力関係の構築、ファクトチェック推進、ICT リテラシー向上に関しては、まだ十分とは言えないものの、我が国においても取組が進められつつある。

項目	
我が国における実態の把握(2. 関係)	・すべての事業者について、我が国における偽情報の実態把握及び結果の分析・公開は未実施
多様なステークホルダーによる協力関係の構築(3. 関係)	・「Disinformation対策フォーラム」や「Innovation Nippon」により一定の進捗がみられる
プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保(4. 関係)	<ul style="list-style-type: none"> <li>・ヤフー(CGMサービス)は偽情報を直接禁止するポリシーが設けられておらず、対応件数等も非公開</li> <li>・LINEは、虚偽の情報の流布を包括的に禁止している。対応件数は一部公開</li> <li>・Googleはコロナ関連偽情報に関してポリシーがあり、我が国における削除件数等について一部公開</li> <li>・Facebook・Twitterは、削除ポリシーの整備や削除結果の公表等がグローバルでは進められているが、我が国における取組状況やその効果については非公開</li> </ul>
利用者情報を活用した情報配信への対応(5. 関係)	<ul style="list-style-type: none"> <li>・いずれの事業者も、偽情報に関する広告や政治広告について、広告配信先の制限や、広告内容に関する何らかの制限を規定するなど、一定の対応は行われている</li> <li>・LINE及びTwitterでは政治広告禁止。Googleでは日本において選挙広告は禁止</li> <li>・政治広告についてのターゲティング技術の適用については明確になっていない</li> </ul>
ファクトチェックの推進(6. 関係)	<ul style="list-style-type: none"> <li>・ヤフーやgoogleなど一部のプラットフォーム事業者において、プラットフォーム上の情報へのファクトチェック結果の紐付けや、資金提供等の取組が開始されている</li> <li>・Facebookでは、グローバルな取組は進められているが、我が国において国際ファクトチェックネットワーク(IFCN)に加盟しているファクトチェッカーがいないため、ファクトチェッカーとの連携は行われていない</li> <li>・LINE、Twitterでは、取組は未実施(LINEは台湾では取組を実施)</li> </ul>
情報発信者側における信頼性確保の方策の検討(9. 関係)	<ul style="list-style-type: none"> <li>・「Disinformation対策フォーラム」において、既存メディアや有識者との情報共有や協議が進展</li> <li>・いずれの事業者も、新型コロナウイルス感染症関係の政府(内閣官房や厚労省等)の情報をサービス内で優先表示させる仕組み等を積極的に実施しているが、認知度は低い</li> <li>・Googleでは、ジャーナリズムの支援策を実施</li> </ul>
ICTリテラシー向上の推進(7. 関係)	<ul style="list-style-type: none"> <li>・Facebook・Google・LINEでは、我が国において偽情報の問題に対応したリテラシー教育に関する取組を実施</li> <li>・Twitterの取組は、偽情報対策に資する内容となっているか不明。ヤフーは今後実施予定</li> </ul>
研究開発の推進(8. 関係)	<ul style="list-style-type: none"> <li>・Facebook、Google、Twitterではグローバルな取組としてディープフェイク対策のための研究開発が行われている</li> <li>・ヤフー、LINEではディープフェイク対策の研究開発は行われていない</li> </ul>

# 【参考】事業者が「回答を控えた理由」及び「今後の対応方針」（2021年）

事業者名	回答を控えた理由及び今後の対応方針
ヤフー	<p>4. (2)削除等の対応 理由:「偽情報」という切り口から対応件数をカウントしているサービスはないため。 対応方針: 対処すべき投稿がどのようなものかを含め、偽情報への対応につき有識者の意見も聞きながら検討を進める。</p> <p>8. 研究開発の推進 理由: 現在は行っていないため。 対応方針: 今後の事情に応じて、研究の要否も検討していく。</p>
Facebook	<p>(誹謗中傷関連回答の再掲)</p> <ul style="list-style-type: none"> <li>・コミュニティ規定施行レポートにおいて、国別の内訳を公表していない理由:             <ol style="list-style-type: none"> <li>①悪意のある利用者が意図的にサービスを悪用する場合、VPNを使って発信元の国を隠すことが多いため、国レベルのデータは信頼性が低くなる</li> <li>②1つのコンテンツが複数の場所で見られる可能性があるため、コンテンツの数や事前対応率を国別に報告することは困難</li> <li>③国レベルのデータではサンプルサイズが小さくなり、データの精度や信頼性が低下する可能性がある</li> </ol> </li> <li>・データの報告義務を規定すると、企業が直面する様々な技術的制限、利用可能なリソース、法的小およびプライバシーに関する考慮事項(各国のデータ保護法を含む)を適切に考慮できない可能性がある。特定のデータセットを追跡し報告するための過度に規定的で義務的な要件は、企業側、そして規制当局の両方にとってコストがかかり負担となる可能性があり、そのプラットフォームの安全性を正確に反映していない可能性もある。その結果、イノベーションと成長が損なわれ、テクノロジーの圧倒的な活用をもたらす利益が制限されることになる</li> <li>・利用者の皆様のためにサービスを向上させ、利用者の皆様の安全性を確保するための投資を継続的にやっていく</li> </ul>
Google	<p>プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保(4. 関係)</p> <ul style="list-style-type: none"> <li>・Googleでは、コンテンツを「偽情報」として特別に分類をしておらず、Googleのプラットフォーム上での詐欺的または低品質のコンテンツを防止することを目的に、長年にわたり多数のコンテンツポリシーを設けており、例えば、スパム、なりすまし、意図的な情報操作に関するポリシーを定めている。このため、「偽情報」という分類でのデータを有していない</li> <li>・偽情報への対処を目的とした政策立案・実施のため、どのような透明性の確保が有益なのか等について、プラットフォームサービス研究会、有識者等の意見を考慮しつつ、検討を行っていく</li> </ul>
Twitter	<ul style="list-style-type: none"> <li>・アルゴリズムに関する透明性・アカウントビリティについて:「責任ある機械学習イニシアティブの開始」というブログ記事を公表 (<a href="https://blog.twitter.com/ja_jp/topics/company/2021/introducing-responsible-machine-learning-initiative">https://blog.twitter.com/ja_jp/topics/company/2021/introducing-responsible-machine-learning-initiative</a>)</li> </ul> <p>(以下、誹謗中傷関連回答の再掲)</p> <ul style="list-style-type: none"> <li>・日本における一般ユーザからの誹謗中傷や偽情報に関する申告・削除要請の件数ならびに削除件数については、今後順次公開する予定</li> <li>・なお、データの集計頻度や分類方法、公開範囲については社内で検討中</li> <li>・一般ユーザからの申告や削除要請に対応する部署の規模については、非公開。サービスごとに直面する課題や受け取る申告・要請は異なっており、求められる対応も異なるため、サービス間の担当者の体制や人数を比較するのは文脈上適切でなく、また、偏った比較となる</li> <li>・対応部署の拠点については安全性やセキュリティの観点から非公開</li> </ul>

※LINEはヒアリングにおいて日本における取組に関する全ての項目に回答済み

## 1. 前提(サービス概要)

- ① 対象サービス名・サービス分類
- ② 加入者数、月間アクティブユーザ数又は書き込み数

## 2. 「我が国における実態の把握」関係

- ① 偽情報等の発生・拡散状況を把握できる体制 分析・調査の有無
- ② グローバル・日本における偽情報等の発生・拡散状況、結果の分析・公表
- ③ 研究者への情報提供、利用条件
- ④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等)
- ⑤ 不正な申告や削除要請への対策の方法・仕組み、対応件数

## 3. 「多様なステークホルダーによる協力関係の構築」関係

- ① Disinformation対策フォーラムなどの産学官民の協力関係の構築
- ② 多様なステークホルダーによる協力関係の構築に向けての具体的な役割
- ③ 議論を踏まえた取組

## 4. 「プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保」関係

## (1) 偽情報等に関するポリシー

- ① (i)禁止行為 (ii)削除等の対応
- ② ポリシー等の見直し状況及び外部レビューの有無とそのタイミング

## (2) 削除要請や苦情に関する受付態勢・プロセス

- ① 一般ユーザからの申告・削除要請への受付窓口・受付態勢、対応プロセス
- ② 対応決定時における通知の内容、理由の記載の程度
- ③ 一般ユーザからの申告や削除要請に対応する部署・チームの規模・人数
- ④ ③の部署・チームに関する日本国内の拠点の有無、日本における責任者の有無
- ⑤ 削除等への苦情や問い合わせに対する苦情受付態勢及び苦情処理プロセス

## (3) 削除等の対応

- ① 偽情報等に関する申告や削除要請の件数
- ② (i)偽情報等に関する申告や削除要請に対する削除件数、(ii)アカウントの停止数
- ③ 偽情報等に関する主体的な削除件数(AIを用いた自動検知機能の活用等)
- ④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等)
- ⑤ 削除以外の取組 (i)警告表示 (ii)表示順位の低下 (iii)その他の取組内容
- ⑥ 不正な申告や削除要請への対策の方法・仕組み、対応件数

## (4) 透明性・アカウントビリティの確保

- ① コンテンツモデレーションのアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照
- ② 透明性レポート 日本語で閲覧可能か
- ③ 取組の効果分析

## 5. 「利用者情報を活用した情報配信への対応」関係

- ① 広告表示先の制限
- ② 広告出稿制限
- ③ ターゲティング技術の適用に関する規定
- ④ 広告のアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照
- ⑤ 出稿者の情報や資金源の公開
- ⑥ 広告とコンテンツの分離
- ⑦ その他の透明性・アカウントビリティ確保方策 ユーザへのツール提供

## 6. 「ファクトチェックの推進」関係

- ① ファクトチェック結果の表示 具体的な仕組み・基準
- ② ファクトチェックを容易にするツールの開発及び提供
- ③ ファクトチェックを実施する人材の育成
- ④ 日本におけるファクトチェック推進に向けた取組
- ⑤ ファクトチェック機関との連携

## 7. 「ICTリテラシー向上の推進」関係

- ① 普及啓発の取組・投資
- ② 他のステークホルダーとの連携・協力・投資

## 8. 「研究開発の推進」関係

- ① AI技術に関する研究開発
- ② 「ディープフェイク」対策の研究開発

## 9. 「情報発信者側における信頼性確保方策の検討」関係

- ① 信頼性の高い情報の表示
- ② ニュースの選別・編集に関する透明性・アカウントビリティ確保方策
- ③ メディアとの連携体制構築 具体的な検討・取組
- ④ 情報源のトレーサビリティ確保、なりすまし防止・認証

## 10. その他

- ① 意見・補足

- 全体的な傾向として、プラットフォーム事業者の偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は未だ限定的。
- 多様なステークホルダーによる協力関係の構築、特定のトピックに関する偽情報や誤解を招く情報の流布に関するポリシーの設定、ファクトチェック推進、ICTリテラシー向上に関しては、まだ十分とは言えないものの、我が国においても取組が進められつつある。

我が国における実態の把握(2. 関係)	・すべての事業者について、我が国における偽情報の実態把握及び結果の分析・公開は未実施
多様なステークホルダーによる協力関係の構築(3. 関係)	・「Disinformation対策フォーラム」などにより一定の進捗がみられる
プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保(4. 関係)	<ul style="list-style-type: none"> <li>・ヤフー(ニュース及び知恵袋)は偽情報を直接禁止するポリシーを設定。対応件数等は一部公開</li> <li>・LINEは、虚偽の情報の流布を包括的に禁止している。対応件数は一部公開</li> <li>・Googleはコロナ関連偽情報に関してポリシーがあり、我が国における削除件数等について一部公開</li> <li>・Meta及びTwitterは、削除ポリシーの整備や削除結果の公表等がグローバルでは進められており、我が国における削除件数等について、ヒアリングの事後に一部公開</li> </ul>
利用者情報を活用した情報配信への対応(5. 関係)	<ul style="list-style-type: none"> <li>・いずれの事業者も、偽情報に関する広告や政治広告について、広告配信先の制限や、広告内容に関する何らかの制限を規定するなど、一定の対応を実施</li> <li>・LINE及びTwitterでは政治広告禁止。Googleでは日本において選挙広告は禁止。政治広告についてのターゲティング技術の適用については、ヤフー及びGoogleにおいて、政治広告に関する行動ターゲティングを制限</li> </ul>
ファクトチェックの推進(6. 関係)	<ul style="list-style-type: none"> <li>・ヤフーやGoogle、LINEなど一部のプラットフォーム事業者において、ファクトチェック結果のまとめの掲出やプラットフォーム上の情報へのファクトチェック結果の紐付け、資金提供等の取組が開始されている</li> <li>・Metaでは、グローバルな取組は進められているが、我が国において国際ファクトチェックネットワーク(IFCN)に加盟しているファクトチェッカーがいないため、ファクトチェッカーとの連携は行われていない</li> <li>・Twitterでは、特定の投稿にユーザが文脈を追加できるBirdwatch機能を試行予定。</li> </ul>
情報発信者側における信頼性確保の方策の検討(9. 関係)	<ul style="list-style-type: none"> <li>・「Disinformation対策フォーラム」において、既存メディアや有識者との情報共有や協議が進展</li> <li>・いずれの事業者も、新型コロナウイルス感染症やウクライナ侵攻関係の情報について、信頼性のあるニュースメディアやファクトチェック機関、政府(内閣官房や厚労省等)の情報をサービス内で優先表示を実施</li> <li>・Googleでは、ジャーナリズムの支援策を実施</li> </ul>
ICTリテラシー向上の推進(7. 関係)	<ul style="list-style-type: none"> <li>・すべての事業者において、偽情報の問題に対応したリテラシー教育に関する取組を実施</li> <li>・ヤフーでは、不確かな情報に惑わされないための「Yahoo!ニュース健診」を提供。</li> </ul>
研究開発の推進(8. 関係)	<ul style="list-style-type: none"> <li>・Google、Meta、Twitterではグローバルな取組としてディープフェイク対策のための研究開発が行われている。ヤフー、LINEではディープフェイク対策の研究開発は行われていない</li> </ul>

# 【参考】事業者が「回答を控えた理由」及び「今後の対応方針」（2022年）

事業者名	回答を控えた理由及び今後の対応方針
ヤフー	<p>&lt;Yahoo!知恵袋&gt; 4. (3)削除等の対応(①偽情報等に関する申告や削除要請の件数、②(ii)アカウントの停止数) 理由: 申告理由について、「偽情報」という切り口で件数をカウントしていません。 対応方針: 対処すべき投稿がどのようなものかを含め、偽情報への対応につき有識者の意見も聞きながら検討を進める。</p> <p>&lt;Yahoo!ニュース&gt; 4. (3)削除等の対応(①偽情報等に関する申告や削除要請の件数～④不正な申告や削除要請への対策の方法・仕組み、対応件数) 理由: 2022年3月にYahoo!ニュースコメントポリシーを改定したばかりであり、件数等定量データ収集中であるため。 コメントポリシー&lt;<a href="https://news.yahoo.co.jp/info/comment-policy">https://news.yahoo.co.jp/info/comment-policy</a>&gt; コメントポリシーの改定&lt;<a href="https://news.yahoo.co.jp/newshack/information/comment-policy-20220322.html?mcclid=443a99f8a9b611ec9a2d846d1a109ec">https://news.yahoo.co.jp/newshack/information/comment-policy-20220322.html?mcclid=443a99f8a9b611ec9a2d846d1a109ec</a>&gt; 対応方針: 来年度以降、回答準備を進めていく。</p>
Google	<p>&lt;誹謗中傷&gt; *2-4 一般ユーザからの申告や削除要請に対する部署・チームの規模・人数 及び部署・チームに関する日本国内の拠点の有無、日本における責任者の有無 →日本の案件の審査と承認は、日本語話者だけではなく、さまざまな部署のグローバルなチームがそれぞれの専門性を持って対応しておりますため、日本の案件だけに従事する人間の数を正確にお示しすることができませんが、日本をカバーする人数は必要に応じて拡充していきます。 *3-3-2 (ii)主體的な削除以外の対応(非表示化やラベルの付与等)を行った件数 →このような取り組みの個別件数については、グローバルに開示しておりません。引き続きユーザーの安全を守るために、どのような仕組みが機能しているかについては、さまざまな形で開示しています。Google 検索については、How Search Works(Google 検索の仕組みを確認する)&lt;<a href="https://www.google.com/intl/ja/search/howsearchworks/">https://www.google.com/intl/ja/search/howsearchworks/</a>&gt;をご覧ください。YouTube については、How YouTube Works (YouTube の仕組みとは)&lt;<a href="https://www.youtube.com/howyoutubeworks/">https://www.youtube.com/howyoutubeworks/</a>&gt;をご覧ください。Google マップについては、Google のクチコミの仕組みについてのブログ&lt;<a href="https://npan.googleblog.com/2022/02/google.html">https://npan.googleblog.com/2022/02/google.html</a>&gt;をご覧ください。 *4-1 発信者情報開示請求を受けた件数 (i)裁判上の開示請求件数、(ii)裁判外の開示請求件数 および *4-2 開示を行った件数 (i)裁判上の開示件数、(ii)裁判外の開示件数 →政府からの情報開示クエリに対して、Google がどのような対応をとったかについて透明性を持たせることを主目的としている弊社の透明性レポートでは、その目的に沿って裁判所を含む政府機関(民事訴訟を除く)におけるユーザー情報の開示についてのデータのみ提供しております。</p> <p>&lt;フェイクニュース&gt; *4(2)-3 一般ユーザからの申告や削除要請に対応する部署・チームの規模・人数 →日本の案件の審査と承認は、日本語話者だけではなく、さまざまな部署のグローバルなチームがそれぞれの専門性を持って対応しておりますため、日本の案件だけに従事する人間の数を正確にお示しすることができませんが、日本をカバーする人数は必要に応じて拡充していきます。 *4(3)-2-ii →何をもち「日本」のアカウントとするかは、技術的な課題がございます。弊社は、継続して利用可能な情報を拡大しますが、正確さとデータの妥当性には最大限の注意を払って行きます。</p>
LINE	<p>偽情報・フェイクニュース対策のヒアリングにおける質問6-④「日本におけるファクトチェック推進に向けた取組」について、「回答を控えた理由」当初段階には「6-④」が無かったところ、他方「6-⑤」において特定非営利活動法人ファクトチェック・イニシアティブ(FI)に正会員として参加し、日本におけるファクトチェック推進に寄与している旨既往回答済みであったことから、凡そ同様の趣旨の回答となることによる重複を避けた方が良いものと判断させて頂いたため。</p> <p>「今後の対応方針」 政府や関係団体等と連携して、日本におけるファクトチェック推進に向けた検討を進める。</p>
Meta	<ul style="list-style-type: none"> <li>・誤情報は、弊社のコミュニティ規定上の他のタイプの言論とは異なり、該当する内容の包括的なリストが決められないことに留意する必要があります。</li> <li>・ポリシーで単に「誤情報」を禁止したとしても、弊社のサービスの利用者にとって有益な通知にはなりません。弊社は全ての情報を知り得るわけではないので、そのようなポリシーを行使することもできません。</li> <li>・このため、弊社のポリシーでは、誤情報をさまざまなカテゴリーに分類し、そのような言動を確認したときにどのように扱うかについて明確な指針を示すよう努めています。</li> <li>・また弊社は、新型コロナウイルス感染症に関連する誤情報については、日本における措置件数を具体的に示しています。</li> <li>・弊社は、今後も外部のステークホルダーと協力し、ポリシーの構築と実施、そして誤情報に対応するための投資を続けて参ります。</li> </ul>
Twitter	<ul style="list-style-type: none"> <li>・日本における誹謗中傷や偽情報に関するデータについては、2022年5月11日に提供いたしました。</li> <li>・削除要請等に対応する部署の規模や視点について、社員の安全面等を考慮し、会社の方針として個別具体的な部署の人数は(今回の当該部署に限らず)公表していません。また拠点についても、セキュリティの観点から公表していません。</li> </ul>



## 1. 前提(サービス概要)

- ① 対象サービス名・サービス分類
- ② 加入者数、月間アクティブユーザ数又は書き込み数

## 2. 「我が国における実態の把握」関係

- ① 偽情報等の発生・拡散状況を把握できる体制 分析・調査の有無
- ② グローバル・日本における偽情報等の発生・拡散状況、結果の分析・公表
- ③ 研究者への情報提供、利用条件
- ④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等)
- ⑤ 不正な申告や削除要請への対策の方法・仕組み、対応件数

## 3. 「多様なステークホルダーによる協力関係の構築」関係

- ① Disinformation対策フォーラムなどの産学官民の協力関係の構築
- ② 多様なステークホルダーによる協力関係の構築に向けての具体的な役割
- ③ 議論を踏まえた取組

## 4. 「プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保」関係

## (1) 偽情報等に関するポリシー

- ① (i)禁止行為 (ii)削除等の対応
- ② ポリシー等の見直し状況及び外部レビューの有無とそのタイミング

## (2) 削除要請や苦情に関する受付態勢・プロセス

- ① 一般ユーザからの申告・削除要請への受付窓口・受付態勢、対応プロセス
- ② 対応決定時における通知の内容、理由の記載の程度
- ③ 一般ユーザからの申告や削除要請に対応する部署・チームの規模・人数
- ④ ③の部署・チームに関する日本国内の視点の有無、日本における責任者の有無
- ⑤ 削除等への苦情や問い合わせに対する苦情受付態勢及び苦情処理プロセス

## (3) 削除等の対応

- ① 偽情報等に関する申告や削除要請の件数
- ② (i)偽情報等に関する申告や削除要請に対する削除件数、(ii)アカウントの停止数
- ③ 偽情報等に関する主体的な削除件数(AIを用いた自動検知機能の活用等)
- ④ ③についての削除の方法・仕組み(AIを用いた自動検知機能の活用等)
- ⑤ 削除以外の取組 (i)警告表示 (ii)表示順位の低下 (iii)その他の取組内容
- ⑥ 不正な申告や削除要請への対策の方法・仕組み、対応件数

## (4) 透明性・アカウントビリティの確保

- ① コンテンツモデレーションのアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照
- ② 透明性レポート 日本語で閲覧可能か
- ③ 取組の効果分析

## 5. 「利用者情報を活用した情報配信への対応」関係

- ① 広告表示先の制限
- ② 広告出稿制限
- ③ ターゲティング技術の適用に関する規定
- ④ 広告のアルゴリズムに関する透明性・アカウントビリティ確保方策、AI原則・ガイドライン等の参照
- ⑤ 出稿者の情報や資金源の公開
- ⑥ 広告とコンテンツの分離
- ⑦ その他の透明性・アカウントビリティ確保方策 ユーザへのツール提供

## 6. 「ファクトチェックの推進」関係

- ① ファクトチェック結果の表示 具体的な仕組み・基準
- ② ファクトチェックを容易にするツールの開発及び提供
- ③ ファクトチェックを実施する人材の育成
- ④ 日本におけるファクトチェック推進に向けた取組
- ⑤ ファクトチェック機関との連携

## 7. 「ICTリテラシー向上の推進」関係

- ① 普及啓発の取組・投資
- ② 他のステークホルダーとの連携・協力・投資

## 8. 「研究開発の推進」関係

- ① AI技術に関する研究開発
- ② 「ディープフェイク」対策の研究開発

## 9. 「情報発信者側における信頼性確保方策の検討」関係

- ① 信頼性の高い情報の表示
- ② ニュースの選別・編集に関する透明性・アカウントビリティ確保方策
- ③ メディアとの連携体制構築 具体的検討・取組
- ④ 情報源のトレーサビリティ確保、なりすまし防止・認証

## 10. その他

- ① 意見・補足

- プラットフォーム事業者の偽情報への対応については、一部で進展が見られるものの、取組状況及び透明性・アカウントビリティ確保の進展は限定的\*。

\*Twitterからは、研究会に出席し発表が行われたものの、ヒアリングシート及び説明資料の提出がなく、透明性・アカウントビリティ確保の取組について後退があった。

- 多様なステークホルダーによる協力関係の構築、特定のトピックに関する偽情報や誤解を招く情報の流布に関するポリシーの設定、ファクトチェック推進、ICTリテラシー向上に関しては、まだ十分とは言えないものの、我が国においても取組が進められつつある。

我が国における実態の把握(2. 関係)	・前回ヒアリング(2022年)に引き続き、すべての事業者において、我が国における偽情報の実態把握及び結果の分析・公開は未実施。
多様なステークホルダーによる協力関係の構築(3. 関係)	・「Disinformation対策フォーラム」での議論や同フォーラムの報告書を受けて創設された日本ファクトチェックセンターを通じた協力関係の構築、一般社団法人ソーシャルメディア利用環境整備機構(SMAJ)における偽情報対策に係る行動規範策定の議論など、一定の進展がみられる。
プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保(4. 関係)	・すでにプラットフォーム事業者では、偽情報等の不適切な情報への措置を講じる必要性が認識され、あらかじめ対応方針や基準となるポリシーを自主的に設定し、投稿の削除やアカウントの停止等を行っている。 ・プラットフォーム事業者による偽情報の削除等に関する透明性・アカウントビリティ確保の取組の進捗は、前回ヒアリングに引き続いて、限定的であるものの、新たに公表する項目があるなど一部進展もあった。
利用者情報を活用した情報配信への対応(5. 関係)	・前回ヒアリングに引き続き、すべての事業者において、偽情報に関する広告や政治広告について、広告配信先の制限や、広告内容に関する何らかの制限を規定するなど、一定の対応を実施。
ファクトチェックの推進(6. 関係)	・日本ファクトチェックセンターの設立にあたり、ヤフー及びGoogleが、一般社団法人セーフターインターネット協会への資金提供等を通じて寄与。日本で活動する複数のファクトチェック団体がIFCNに加盟。 ・LINE及びMetaにおいても、国際ファクトチェックネットワーク(IFCN)加盟のファクトチェッカーとの連携を推進。
ICTリテラシー向上の推進(7. 関係)	・前回ヒアリングに引き続き、すべての事業者において、偽情報の問題に対応したリテラシー教育に関する取組が実施されている。
研究開発の推進(8. 関係)	・すべての事業者においてAI技術に関する研究開発が行われている。また、Google及びMetaにおいては、「ディープフェイク」対策の研究開発も実施されている。
情報発信者側における信頼性確保方策の検討(9. 関係)	・前回ヒアリングに引き続き、いずれの事業者も、公共性の高い情報等をユーザの目に入りやすい位置に優先的に配置する等の取組を行っている。

# 【参考】 2021～2023年のモニタリング結果概要（まとめ）

## 【2021年】

- 全体的な傾向として、プラットフォーム事業者の偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は限定的。
- 他方、多様なステークホルダーによる協力関係の構築、ファクトチェック推進、ICT リテラシー向上に関しては、まだ十分とは言えないものの、我が国においても取組が進められつつある。

## 【2022年】

- 全体的な傾向として、プラットフォーム事業者の偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は未だ限定的。
- 多様なステークホルダーによる協力関係の構築、特定のトピックに関する偽情報や誤解を招く情報の流布に関するポリシーの設定、ファクトチェック推進、ICT リテラシー向上に関しては、まだ十分とは言えないものの、我が国においても取組が進められつつある。

## 【2023年】

- プラットフォーム事業者の偽情報への対応については、一部で進展が見られるものの、取組状況及び透明性・アカウントビリティ確保の進展は限定的\*。

\*Twitterからは、研究会に出席し発表が行われたものの、ヒアリングシート及び説明資料の提出がなく、透明性・アカウントビリティ確保の取組について後退があった。

- 多様なステークホルダーによる協力関係の構築、特定のトピックに関する偽情報や誤解を招く情報の流布に関するポリシーの設定、ファクトチェック推進、ICT リテラシー向上に関しては、まだ十分とは言えないものの、我が国においても取組が進められつつある。