

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とりまとめ（抜粋）

第7章 総合的な対策 — 制度的な対応

I 情報伝送 PF 事業者による偽・誤情報への対応の在り方

情報伝送 PF サービスは、デジタル空間における情報流通の主要な場となっており、その中で偽・誤情報が流通・拡散すること等により、

- ・ 人の生命、身体又は財産への影響
（例えば、健康被害、災害時の救命・救助活動や復旧・復興活動の妨害、詐欺被害、事業者への風評被害を含む営業妨害等）
- ・ 人格権やその他基本的人権への影響
（例えば、誹謗中傷、なりすましによる肖像権等の侵害、ヘイトスピーチ等）
- ・ 健全な民主主義の発達への影響
（例えば、集団分極化に伴う民主的政治過程への悪影響等）
- ・ その他の社会的混乱等の実空間への影響
（例えば、株価の下落、公共インフラの損壊、外交関係の悪化等）

が発生・増幅している。

こうした社会的影響については、その背景として、情報伝送 PF サービスの特徴と、当該特徴が一因となって情報流通空間としてのデジタル空間を支配するに至った経済モデルの存在が指摘されている¹。

すなわち、情報伝送 PF サービスには、

- ① 誰もが低コストで不特定の者に向けた情報発信を行うことができ（情報発信コストの低廉性）、
- ② 情報の流通・拡散を促進する「いいね」やリポスト等の機能を備えていたり（拡散促進機能の具備）、
- ③ 閲覧等受信側の利用者の興味・関心等に応じてコンテンツやデジタル広告の表示順位その他の表示方法を変更する機能を備えている（レコメンデーション機能・広告ターゲティング機能の具備）

といった特徴を有するものが存在し、特に、コンテンツに付随して表示されるデジタル広告に対して支払われる広告費を主たる運営基盤としている情報伝送 PF サービスの場合には、デジタル広告、ひいてはデジタル広告が付随するコンテンツの流通・拡散しやすさが収益に直結することから、上記①から③までに挙げた特徴をより強化するインセ

¹ 本検討会第6回会合（2024年1月19日）における山本龍彦座長代理による発表（資料6-2-2）を参照。

ンティブが存在すると言える。

その結果、情報伝送 PF サービスを中心とする現在のデジタル空間は、第 1 章 II 2. で述べたように、多様な一般利用者や広告主が実名・匿名で投稿・出稿する玉石混交のコンテンツやデジタル広告が、伝統メディアをはじめとするプロの書き手・送り手によるコンテンツと混じり合いながら溢れる情報過多の状況になっており、その状況がもたらす、一般に「アテンション・エコノミー」と呼ばれる経済モデルの下で、情報伝送 PF サービスは、上記①から③までに挙げた特徴を一層先鋭化させ、更なる情報過多の状況をデジタル空間において再生産する傾向がある。

さらに、放送、新聞等の伝統メディアが制作・編集・発信し、かつ自ら伝送したコンテンツが受信者の主な情報源であった従来の情報流通空間²と異なり、情報伝送 PF サービスを通じた情報流通³に関しては、個人の情報処理能力を遙かに超える量の情報が未整理の形で流通することにより、表現の自由や知る権利の前提となる思想同士の自由競争の場（思想の自由市場）が機能不全に陥る可能性も指摘されている⁴。

このように、情報伝送 PF サービスは、偽・誤情報の流通・拡散や、それによる社会的影響の発生・増幅をもたらすリスクを構造的に有していると言える。

このことを踏まえ、情報伝送 PF 事業者には、偽・誤情報の流通・拡散に関連して自社サービスやそのアーキテクチャ（サービスに組み込まれた拡散促進機能、レコメンデーション機能及び広告ターゲットティング機能に加え、それらを支えるアルゴリズムを含む。）がデジタル空間における情報流通の健全性、ひいては権利侵害、社会的混乱その他の実空間や個人の意思決定の自律性に与える影響・リスクを自ら適切に把握した上で、影響・リスクに応じたコンテンツモデレーションを実施し、またデジタル広告の質の確保や質の高いメディアへの広告配信に資する取組を実施するなど、情報流通の適正化に取り組む一定の責任を果たすことが期待されている。また、利用者の表現を預かる立場として、利用者の表現の自由の確保に取り組む一定の責任を果たすことも期待されている。特に、大規模事業者は、人々の日常的なコミュニケーション手段等として、国民生活や社会経

² NHK 受信料制度等検討委員会 次世代NHKに関する専門小委員会「次世代NHKに関する論点とりまとめ(第2次)報告書」(2023年3月31日)。

³ 人々の主な情報取得先が伝統メディアから情報伝送 PF サービスに移行していることについて、総務省「令和4年度情報通信メディアの利用時間と情報行動に関する調査報告書」参照。例えば、テキスト系メディアの平日1日の平均利用時間について、全年代において、テキスト系サイト利用(ブログやウェブサイトに加え、SNS等の利用も含む。)に費やす時間が最も長く、新聞閲読や書籍・雑誌・コミックに費やす時間を大きく上回っており、かつその傾向が10年以上継続している。また、個人が最も利用しているテキスト系ニュースサービスについては、「ポータルサイトによるニュース配信」と「ソーシャルメディアによるニュース配信」を合計した割合が一貫して増加する一方、「紙の新聞」の割合は一貫して減少している。

⁴ WG における議論では、デジタル空間における情報流通に関し、個人の情報処理能力を遙かに超える量の情報が未整理の形で流通していることの問題が顕在化していることなどから、思想の自由市場論の書かれざる前提(①市場に流通する情報は多ければ多いほどよい、②そのためには国家の介入は少ないほどよい、③情報の受け手は自律性を有し、情報の選別・判断能力を備えている)の妥当性が失われつつあるとの意見があった(WG 第 14 回会合(2024年4月12日)における曾我部主査代理による発表(資料 WG14-2)参照)。一方で、現時点で思想の自由市場の機能不全の危険性が高まっているのは、デジタル空間の中でも情報伝送 PF サービスや広告仲介 PF サービスといったプラットフォームサービスを通じた情報流通に限られていることから、そのような情報伝送主体を対象を限定した対応を検討すべきとの意見もあった(WG 第 16 回会合(2024年4月17日)における森構成員による発言参照)。

済活動等に広くかつ深く浸透しており⁵、我が国のデジタル空間における情報流通について公共的役割を果たしている。

現に、主要な情報伝送 PF 事業者は、自社サービスの利用規約等の中で、例えば次のような情報について、情報の削除、表示順位の低下等の一定の対応をする旨を定めている。

■「Yahoo!ニュースコメントポリシー⁶」

- 健康被害等をもたらす可能性のある偽情報であって、ファクトチェックにより反真実であることが明らかな投稿（明らかな偽情報）（「新型コロナウイルスのワクチンを接種すると、流産する。不妊になる。」、「ワクチン接種された実験用の動物が全て死亡した。」、「ワクチンを接種することでコロナウイルスに感染する。」など）

■Yahoo!知恵袋「利用のルール⁷」

- 明らかに事実と異なり社会的に混乱を招く恐れのある投稿（明らかな偽情報）（「(そのような事実がないにもかかわらず)昨日、〇〇(地名)で大地震があったけど、、、」、「コロナワクチン殺人計画は本当ですか?接種したネズミは3年以内に死亡したそうです。」、「トランプ大統領、コロナで亡くなったみたいですね。やはり突然の重症化、恐ろしいウイルスです」など）

■ファイナンス掲示板「ヘルプページ⁸」

- 明らかな偽情報として、架空の出来事のでっちあげや虚偽の内容の投稿（風説の流布に該当する場合違法行為となる可能性）（「〇〇会社が製造したワクチンを接種された実験用の動物が全て死亡した。」など）

■LINE オープンチャット「安心・安全ガイドライン⁹」

- 真偽不明の情報の拡散（新型コロナウイルス関連：「コロナワクチンによって不妊になる」、「コロナは人口削減のために人工的に作られた」など、災害関連：「能登半島地震は人工地震である」など）

■「LINE VOOM コミュニティガイドライン¹⁰」

- 誤情報の拡散：当社または第三者になりすまし行為や、フェイクニュースなど虚偽の情報、身体に影響を及ぼす食品、医療、医療薬などの虚偽の情報を意図的に発信・拡散・流布させる行為（新型コロナウイルス関連「コロナワクチンによって不妊になる」、「コロナは人口削減のために人工的に作られた」など、災害関連「能登半島地震は人工地震である」など）

■YouTube「誤った情報に関するポリシー¹¹」

- 特定の種類の誤解を招くコンテンツまたは虚偽が含まれるコンテンツで、重大な危害を及ぼす可能性のあるもの（現実の世界で危害を与える可能性がある特定の種類の誤った情報、技術的に操作された特定の種類のコンテンツ、民主的な手続きを妨害するコンテンツが含まれる）
- 国勢調査の妨害：国勢調査に関する時間、場所、方法、資格要件について参加者を誤解させることを目的としたコンテンツ、または国勢調査を著しく妨げる可能性のある虚偽の主張（国勢調査の参加方法に関して誤った手順を示す。回答者の在留資格が法執行機関に報告されるという誤った主張により、国勢調査への参加を妨げる行為。）

⁵ 総務省「令和4年度情報通信メディアの利用時間と情報行動に関する調査報告書」によると、例えば SNS について、X(旧 Twitter)の利用率は 20 代では 78.8%と高く、Instagram の利用率も全年代で一貫して増加している。また、平日・休日を問わず、インターネットの平均利用時間はテレビの平均視聴時間を超過しており、かつ、インターネットの利用項目別の利用時間では、休日の 10 代及び 20 代の「動画投稿・共有サービスを見る」、「ソーシャルメディアを見る・書く」の平均利用時間がいずれも 100 分を超過している。

⁶ <https://news.yahoo.co.jp/info/comment-policy>

⁷ <https://chiebukuro.yahoo.co.jp/topic/guide/rule/>

⁸ <https://support.yahoo-net.jp/ScFinance/s/article/H000011273>

⁹ <https://openchat-jp.line.me/other/guideline>

¹⁰ https://terms.line.me/line_voom_community_guideline?lang=ja&country=JP

¹¹ <https://support.google.com/youtube/answer/10834785?hl=ja>

- 改ざんされたコンテンツ：ユーザーの誤解を招くように技術的に操作または改ざんされ（前後関係を無視してクリップを切り抜く以上の操作が多い）、重大な危害を及ぼす可能性のあるコンテンツ（地政学的緊張を悪化させ、重大な危害を及ぼす可能性のある不正確に翻訳された動画の字幕。政府関係者の死を装うために技術的に操作（前後関係を無視してクリップを切り抜く以上の操作が多い）された動画。重大な危害を及ぼす可能性のある事件を捏造するために技術的に操作（前後関係を無視してクリップを切り抜く以上の操作が多い）された動画。）
- 虚偽のコンテンツ：過去の事象の古い映像を現在の事象のものであると虚偽の主張をすることで、重大な危害を及ぼす可能性のあるコンテンツ（実際は別の地域や事象に関するコンテンツを、特定の地域の人権侵害の記録として不正確に提示したコンテンツ。実際は数年前の映像であるにもかかわらず、そのコンテンツが現在の事象のものであるという虚偽の主張とともに、抗議活動への参加者に対する軍事的弾圧を示したコンテンツ。）

■Meta「偽情報」ポリシー¹²

- 実際の危害や暴力：弊社は、人々に対する差し迫った暴力または実際の危害のリスクに直接つながる可能性が高いと、専門家のパートナーが判断した偽情報および検証できない噂を削除します。偽情報とは、信頼できる第三者が虚偽であると判断する主張を含むコンテンツと定義されます。検証できない噂とは、専門家のパートナーによる情報元の特定が極めて困難であるか不可能な主張、信頼できる提供元がない主張、その内容を証明するための具体性が不十分な主張、またはその内容があまりに信じがたい、もしくは不合理で信用できない主張と定義されます。弊社は、一見無害に思われる偽情報でも、特定の文脈では、死や深刻な怪我といった実際の危害の高いリスクにつながる可能性がある暴力的脅威など、オフラインでの危害のリスクにつながる場合があることを認識しています。弊社は、こうした各地域の動向について専門知識を有する非政府組織(NGO)、非営利団体、人道支援組織、国際機関のグローバルネットワークと連携しています。社会的暴力のリスクが高まっている国では、どの虚偽の主張が差し迫った身体的危害のリスクに直接つながるかを理解するために、現地のパートナーと積極的に協力しています。その上で、弊社のプラットフォームでそのような主張をするコンテンツを特定し削除します。例えば現地の専門家と協議して、文脈から切り離して虚偽の主張をし、暴力行為、暴力の被害者や加害者、武器や軍事用具を描写するメディアを削除する可能性があります。）
- 有害な健康関連の偽情報：弊社では、公衆衛生と安全に対する差し迫った危害に直接的な害をもたらす可能性の高い健康誤情報を特定するために、主要な保健機関と協議しています。弊社が削除する有害な健康関連の偽情報には次のようなものがあります。ワクチンに関する偽情報（弊社は、主にワクチンに関する偽情報について、保健当局がその情報は虚偽であり、差し迫ったワクチン接種の拒否を促進する可能性が高いと判断した場合は、その情報を削除します。）、公衆衛生上の緊急事態の間の偽情報（偽情報については、公衆衛生当局が、その情報が虚偽であり、差し迫った身体的危害のリスクに直接つながる可能性が高いと判断した場合、公衆衛生上の緊急事態の間、弊社はその情報を削除します。これには、個人が有害な疾病に感染したり、それを拡散させたりするリスクや、関連するワクチンを拒否するリスクにつながるものが含まれます。弊社は、世界と地域の保険機関と連携して公衆衛生上の緊急事態を特定しています。現在対象となるものとしては、ウイルスに関する緊急事態宣言が引き続き公に発表されている国における、新型コロナウイルス感染症関連の特定の虚偽の主張などがあります。新型コロナウイルス感染症およびワクチンについて、弊社が許可していない偽情報の種類をまとめたルール一覧についてはこちらをクリックしてください。）、健康上の問題に対する有害な「奇跡的な治療法」の宣伝または擁護（これには、医療の観点から推奨される使用法が、重傷または死亡のリスクに直接つながる可能性が高く、正当な医療上の使用法がない治療が含まれます(例：漂白剤、消毒剤、黒色軟膏、苛性ソーダ)。)
- 投票者または国勢調査への干渉：選挙や国勢調査の健全性を促進する取り組みとして、弊社は、

¹² <https://transparency.meta.com/ja-jp/policies/community-standards/misinformation/>

このようなプロセスに人々が参加する能力を妨害するリスクに直接つながる可能性が高い偽情報を削除します。これには次のようなものが含まれます。投票や有権者登録、国勢調査への参加に関する日付、場所、時間、方法に関する偽情報、投票できる人、投票資格、投票の有効性、投票するために提供しなければならない情報や書類に関する偽情報、候補者が立候補するか否かに関する偽情報、国勢調査への参加資格や、国勢調査へ参加するために提供しなければならない情報や書類に関する偽情報、個人の国勢調査の情報が別の(国勢調査の実施を担当していない)政府機関と共有されると述べること(ただし該当する場合)など、政府の国勢調査への関与についての偽情報、米国移民関税執行局(ICE)が投票所に待機しているとの偽りのまたは根拠のない主張、投票プロセスに参加すると新型コロナウイルス感染症(または別の伝染病)に感染するとする明確な偽の主張、選挙当局によって検証された、米国の投票所の現状に関して投票ができなくなるような偽りの主張。弊社では、暴力の呼びかけ、違法な参加の促進、選挙への協調的干渉の呼びかけを対象とする追加ポリシーを設けており、これらは弊社のコミュニティ規定の他のセクションに記載されています。

- 加工されたメディア：メディアはさまざまな方法で編集することができます。多くの場合、このような変更は、芸術的な理由による切り取りもしくは短縮、または音楽の追加など、無害なものです。ただ、特に動画コンテンツの場合、中には加工の有無がはっきりせず誤解を与えてしまう可能性が生じることもあります。このようなコンテンツはすぐに拡散されやすく、専門家の助言によれば、加工されたメディアに関する誤った思い込みはさらなる議論で修正できないことが多々あるため、弊社ではこのようなコンテンツを削除します。次の一定の条件を満たす動画については、本ポリシーに基づき削除します。(1)明瞭さや画質・音質の調整にとどまらず、動画の被写体が実際には発言していない言葉を使ったと一般の利用者に誤解させるような編集または合成が、一般の利用者にはわからない形で行われている動画、および(2)ディープラーニング技術を含む人工知能または機械学習(例：人為的なディープフェイク)によって制作され、特定の動画に他のコンテンツを統合、結合、置換、重ね合わせるなどして動画が本物に見えるように作成されている動画。
- 偽・誤情報ポリシー等に違反していないにもかかわらず、プラットフォームの信頼性と完全性(authenticity and integrity)を損なう誤情報については、独立した第三者ファクトチェック・パートナーのネットワークと協力して、誤情報の配信を減らし、強力な警告ラベルを表示し、誤情報に出くわした人、それを共有しようとした人、またはすでに共有した人に通知するというアプローチをとっている。1つのファクトチェックに基づき、誤りであることを暴いたストーリーと重複するものを特定する類似性検出方法を発動し、特定されたポストに対して、フィード上の流通を減らし、警告ラベルを表示し、利用者に通知するという同じペナルティを適用することができる。

■TikTok「コミュニティガイドライン」¹³

- 意図にかかわらず、個人や社会に重大な危害を及ぼし得る不正確な、誤解を招く、または虚偽のコンテンツ(重大な危害には、身体的、心理的、または社会的危害、および物的損害が含まれる。商業上の損害や風評被害はこれには含まれず、単なる不正確な情報や作り話も対象外)
- 公共の安全に危険を及ぼしたり、危機または緊急事態についてパニックを引き起こし得たりする誤情報(以前に行われた攻撃の過去の映像を現在進行中であるかのように使用することや、特定の場所で基本的な生活必需品(食料や水など)が在庫切れを起こしていると誤った主張することなど)
- 誤った医療情報(ワクチンに関する誤解を招く発言、生命を脅かす疾患に対して適切な治療を受けることを思いとどまらせる不正確な医療上の助言、公衆衛生に危険を及ぼすその他の誤情報など)
- 確立した科学的コンセンサスを阻害する気候変動に関する誤情報(気候変動やその一因となっ

¹³ https://www.tiktok.com/community-guidelines/ja-jp/integrity-authenticity/?enter_method=left_navigation#1

ている要素の存在を否定することなど)

- ・ 暴力的またはヘイトに満ちた危険な陰謀論（暴力的な行動を呼びかける、過去の暴力行為と関連付ける、十分に立証されている暴力事件を否定する、保護属性を持つグループに対する偏見を引き起こすなど）
- ・ 個人を名指しして攻撃する、特定の陰謀論
- ・ 現実世界の出来事について人に誤解を与えるようなやり方で編集、接合、または合成された素材（動画や音声など）
- ・ 根拠がなく、特定の出来事や事態が「政府」や「秘密結社」などの秘密の集団や強力な集団によって引き起こされたとする一般的な陰謀論
- ・ 詳細がまだ明らかになっていない緊急事態や展開中の出来事に関連する未確認情報
- ・ ファクトチェックの審査中の、危険性が高い可能性のある誤情報

■ LinkedIn 「プロフェッショナルコミュニティポリシー¹⁴」

- ・ 虚偽または誤解を招く内容。事実として提示された特定の主張が、明らかに虚偽であるか実質的に誤解を招くものであり、害を及ぼす恐れがある場合は、これを削除。虚偽であるか実質的に誤解を招くものではあるものの、害を及ぼす恐れがないコンテンツの場合、投稿者のネットワーク外での配信は認められない。

■ LinkedIn ヘルプセンター「虚偽または誤解を招くコンテンツ¹⁵」

- ・ まもなく実施されるか終わったばかりの選挙の投票の時間、場所、手段、資格要件に関する虚偽の、または実質的に誤解を招く情報が含まれたコンテンツ
- ・ 被害を引き起こす恐れのある脅威、暴力、危険についての根拠のない主張など、緊急時パニックを誘発したり、安全対策を講じる意欲を失わせたりするような主張（例：「森林火災が発生している地域で略奪者が横行している」など）
- ・ 特定の場所における人権侵害または軍事紛争の証拠として提示された、実際には別の場所、出来事、期間の、不正確なコンテンツ
- ・ 実際の出来事を歪曲した、対象者、他の個人やグループ、または社会全体に害を及ぼす恐れのある、加工された画像や動画などの合成または操作されたメディア
- ・ 有害な治療法や奇跡の治療法を宣伝するコンテンツ、または専門的な医学的アドバイスを求めたり、聞き入れたりすることを妨げるコンテンツ
- ・ 地域の保健当局や世界保健機関（WHO）の医療ガイダンスと相反する主張や記述
- ・ COVID-19 の治療、予防、変異株、感染に関する医療上の誤情報（医療用、非医療用を問わず、マスクは身体に悪影響を及ぼす、イベルメクチンまたはヒドロキシクロロキンは COVID-19 の治療または予防に有効である、承認された COVID-19 のワクチンは、COVID-19 を含む感染症による死亡、不妊、流産、自閉症、または筋収縮を引き起こす恐れがある、承認された COVID-19 のワクチンは、追跡/監視装置を埋め込んだり、患者に磁気を発生させたりする、COVID-19 のワクチン接種者は、ワクチン未接種者に比べ、ウイルスを拡散させる可能性が高い、COVID-19 は、特定の世界的な指導者、公人、または世界各地の保健当局からの資金提供または支援を受けて開発された）
- ・ COVID-19 の有病率または重症度に関する医学的誤報（COVID-19 は存在せず、デマである、ウイルスとその変異株が根絶され、パンデミックは終息した、COVID-19 の症状、死亡率、感染率は季節性インフルエンザよりも深刻なものではない、COVID-19 に感染して、死亡または重症化した人はいない）

■ X 「プラットフォームの操作とスパムに関するポリシー¹⁶」

- ・ 情報を人為的に拡散または隠蔽したり、X のユーザー体験や、X によるプラットフォーム操作

¹⁴ <https://jp.linkedin.com/legal/professional-community-policies>

¹⁵ <https://www.linkedin.com/help/linkedin/answer/a1340752/>

¹⁶ <https://help.x.com/ja/rules-and-policies/platform-manipulation>

の防止策を操作または侵害する行為に関与したりする意図で、Xのサービスを利用することは禁じられている。

■ X「誤解を招くアイデンティティや虚偽のアイデンティティに関するポリシー¹⁷」

- ・ 他の利用者を欺く目的で、個人、集団、組織の身元情報を流用したり、自身の身元を詐称したりすることはできない。

■ X「合成または操作されたメディアに関するポリシー¹⁸」

- ・ 利用者を欺いたり、混乱させたりして、損害をもたらす可能性のある、合成または操作されたメディアや、文脈から切り離されたメディア（以下、「誤解を招くメディア」）を共有することは禁止されている。さらに、Xでは誤解を招くメディアが含まれるツイートにラベル付けを行い、利用者がツイートの信ぴょう性や文脈を把握するのを支援する場合がある。

■ X「市民活動の阻害に関するポリシー¹⁹」

- ・ 選挙期間中に、市民活動への参加を抑圧したり、市民活動に関する日時、場所、参加方法について誤解を招いたり、現実世界の暴力を引き起こしたりするおそれのあるコンテンツをポストまたは共有するなど、選挙またはその他の市民活動の操作や妨害を目的にXのサービスを利用することを禁じている

しかしながら、本検討会におけるプラットフォーム事業者ヒアリングの結果を踏まえると、デジタル空間における情報流通の適正化や利用者の表現の自由の確保に向けた情報伝送 PF 事業者による取組として、我が国国内における偽・誤情報の流通・拡散への対応状況（情報の削除等）を含む取組状況に関する透明性・アカウントビリティの確保は総じて不十分であり、取組状況そのものについても全体として十分とは言えない。事業者団体による偽・誤情報対策に関する自主的な行動規範の策定に関する議論が白紙に戻り中断されていることにも鑑みると、情報伝送 PF 事業者による自主的な取組のみには期待できない状況であり、新たな具体的な対応が必要である。

加えて、今後、生成 AI 等の新たな技術やサービスの進展・普及による偽・誤情報の爆発的増加・巧妙化も懸念されるほか、情報伝送 PF サービスに組み込まれたアルゴリズム等の影響（フィルターバブル、エコーチェンバー等）により、人々が多様な情報を受信できずに適切な判断を下すことが困難となり、インターネット上で集団分極化が進み、結果として社会経済の混乱や民主主義への悪影響をもたらす可能性が指摘される²⁰など、誰もが自律的に情報を発信し、情報を摂取できる場としてのインターネットの存立が脅かされつつある近年の状況に鑑みると、情報伝送 PF 事業者による取組を中心としたデジタル空間の情報流通に関して、健全性を確実かつ持続的に確保するためのガバナンスを確立することが急務である。

以上のようなデジタル空間における情報流通の現状や、情報伝送 PF 事業者に期待される役割・責務及びこれまでの取組状況等を踏まえると、制度整備も含め、情報伝送 PF 事業者に対して以下の具体的措置を求めることが適当である。

¹⁷ <https://help.x.com/ja/rules-and-policies/x-impersonation-and-deceptive-identities-policy>

¹⁸ <https://help.x.com/ja/rules-and-policies/manipulated-media>

¹⁹ <https://help.x.com/ja/rules-and-policies/election-integrity-policy>

²⁰ 鳥海不二夫＝山本龍彦「共同提言「健全なプラットフォームに向けて ver2.0－情報的健康を、実装へ」」KGRI Working Papers No. 1 (2023年5月) (本検討会第6回会合(2024年1月19日)配付資料6-2-3)参照。

1. 対応を検討すべき「偽・誤情報」の範囲に関する基本的な考え方

「偽・誤情報」²¹の中でも、情報伝送 PF 事業者において（制度整備による対応に限らず、自主的な対応も含む）何らかの対応を検討すべき範囲については、利用者の表現の自由をはじめとする様々な権利利益に配慮する観点から、前述の主要な情報伝送 PF 事業者における現状の利用規約等の内容を踏まえつつ、対象範囲の客観的な明確性を確保するとともに、必要かつ相当な対策が適正に講じられることを担保できるよう定められる必要がある。

具体的には、少なくとも、次の①及び②の要件をいずれも満たす情報は、原則として、何らかの対応を検討すべき「偽・誤情報」の範囲に含まれるものと考えることが適当である。

- ① 検証可能な誤りが含まれていること
- ② 次の各要素の有無・軽重に照らし、具体的な方策との関係で比例性が認められること²²
 - i. 当該情報そのものが有する権利侵害性その他の違法性や客観的な有害性（及びその明白性）
 - ii. 当該情報が流通・拡散することによる社会的影響の重大性（及びその明白性）
例）人の生命、身体又は財産に重大かつ明白な悪影響を与えるような情報
 - iii. ①の誤りが含まれていることについての検証の容易性（誤りが含まれていることの明白性）

①の要件は、対象範囲の客観的な明確性を確保する観点から要求されるものである。「内容」に誤りが含まれている情報のみならず、なりすましアカウントによる投稿など、発信者の「名義」に誤りが含まれる情報も、①の要件を満たし得る。

一方、誤りが含まれていることに関する発信者の認識（主観的意図）については、「偽情報」と「誤情報」とを画する要件にはなり得る²³ものの、情報伝送 PF 事業者において判別困難と考えられることから、対応を検討すべき「偽・誤情報」の範囲を画する要件とはしないことが適当である。

②の要件は、どのような「偽・誤情報」に対して、具体的にどのような対応（制度整備による対応に限らず、自主的な対応も含む何らかの具体的な方策）を実施すべきかが、

²¹ 前掲脚注 41 及び 42 のとおり、本章において特に留保を付さずに「偽・誤情報」の表現を用いる場合、その意味するところ（いわゆる広義の「偽・誤情報」）は、偽情報（誤りが含まれる情報のうち、発信者が事実でない事項を事実であると誤認・誤解させる意図を持って発信したもの）及び誤情報（誤りが含まれる情報のうち、発信者が事実でない事項を事実であると誤認・誤解させる意図を持たずに発信したもの）の総称である。

²² 様々な具体的な方策があるため、全体的な考え方としてはこのようなものとなるが、特定の方策との関係では対象となる情報の定義・内容が明確にされる必要がある。

²³ 例えば、EU の欧州民主主義行動計画 (European Democracy Action Plan) は、「有害な意図を持たずに共有されながらも、その効果は未だ有害である虚偽の、又は誤解を招くコンテンツ」を「誤情報」(misinformation)、「欺き、又は経済的若しくは政治的利得を確保する意図を持って拡散され、公共への損害を生じさせ得る虚偽の、又は誤解を招くコンテンツ」を「偽情報」(disinformation) とそれぞれ定義しており、この定義は「2022 年偽情報に関する強化された行動規範」(the Strengthened Code of Practice on Disinformation 2022) でも踏襲されている。

i. から iii. までの各要素の有無・軽重により異なり得ることを示している。

この点について、情報そのものに「権利侵害性その他の違法性」(i.)がある場合には、制度整備による対応を含む比較的広い範囲の具体的な方策との関係で②の要件に合致するものと評価し得る²⁴。

また、「権利侵害性その他の違法性」がない情報であっても、例えば、当該情報そのものが、又は当該情報が流通・拡散することにより、人の生命、身体又は財産に重大かつ明白な悪影響を与えるような情報²⁵については、情報伝送 PF 事業者において、少なくとも、これらの情報の流通・拡散に関連して自らのビジネスモデルがもたらす社会的影響を予測し、有効な軽減措置を実施する(II参照)といった方策(又はそれ以上の方策)を要する程度の「客観的な有害性」(i.)又は「社会的影響の重大性」(ii.)を備えている、すなわち②の要件に合致するものと評価し得る。

一方、これらの情報の具体的な範囲や、これらの情報以外のいかなる情報²⁶(又はその流通・拡散)について、いかなる具体的な方策との関係で、「客観的な有害性」や「社会的影響の重大性」が認められ得るかについては、今後、更なる検討が必要である。その際、特に制度整備による対応との関係で「客観的な有害性」や「影響の重大性」が認められる範囲については、表現の自由への過度の制約を避ける観点から、明確に定められることが重要である。

②の要件に関連して、「客観的な有害性」及び「社会的影響の重大性」がともに小さいなど一定の類型の情報については、対応を検討すべき「偽・誤情報」の範囲に含まれないものと考えることが適当である。なお、「一定の類型の情報」として具体的にどのようなものが考えられるかについては、今後、更なる検討が必要である^{27,28}。

以上のほか、必ずしも誤りは含まれていないが文脈上誤解を招く(ミスリーディングな)情報や、事実ではあるが人を害する意図を持って発信された悪意ある情報への対応の要否及び具体的な対応の在り方については、具体的なケースを想定しつつ、今後、更

²⁴ 本検討会における議論では、オブザーバー団体(一般社団法人日本民間放送連盟)から、違法アップロードや放送コンテンツの改ざんへの対応を情報伝送 PF 事業者に求める意見があった(本検討会第 21 回(WG 第 24 回)会合(2024 年 5 月 30 日)配付資料 21-1-3)。

²⁵ 例えば、感染症流行時に健康被害を生じさせ得る医学的に誤った治療法を推奨する情報、地震等の災害発生時における救命・救助活動の妨げとなる実在しない住所を摘示しての救助要請など。

²⁶ WG における議論では、例えば、社会的混乱を招くような情報(例:存在しない災害が存在するかのように見せかけた偽画像、政府機関の建物が外部からの攻撃に遭ったかのように見せかけた偽画像、国民生活に欠かせない公共インフラについて科学的根拠なく有害なものであると主張する偽・誤情報等)や、特定の集団に対する差別・偏見を招くような情報(例:地震等の災害発生時に外国人が犯罪を犯しているとする偽・誤情報等)への言及があった。これらの情報への対応の在り方についても、今後の検討対象となり得る。

²⁷ 「一定の類型の情報」として、例えばパロディ・風刺などが考えられる。この点、豪州の 2023 年通信法改正案(Combating Misinformation and Disinformation Bill)は、「真に娯楽、パロディ又は風刺の目的で作成されたコンテンツ」を適用除外としているところ、WG における議論では、我が国で対応すべき「偽・誤情報」の範囲を検討するに当たって、「真に」等の要件を求めることによる限定の要否が問題になり得るとの意見があった。

²⁸ このほか、パロディ・風刺などとは性質が異なるが、報道の自由(最大判昭和 44 年 11 月 26 日刑集第 23 卷第 11 号 1490 頁(博多駅テレビフィルム提出命令事件))に基づく自律的な努力が求められる報道機関等について、その自律性を尊重する観点から、伝統メディア等による誤報を上記「一定の類型の情報」に含めることも考えられる。関連して、EU の「2022 年偽情報に関する強化された行動規範」でも「誤報」(reporting error)が「偽情報」の定義から除外されているほか、豪州の 2023 年通信法改正案でも「専門的なニュースコンテンツ」(professional news content)が適用除外とされている。

なる検討が必要である。

2. 偽・誤情報の流通・拡散を抑止するための「コンテンツモデレーション」の類型

情報伝送 PF 事業者は、自らのビジネスモデルがもたらす社会的影響の軽減に向け、偽・誤情報の流通・拡散を抑止等するために、例えば次に挙げるような多様な類型によるコンテンツモデレーションを講じている²⁹。

(1) 偽・誤情報の流通・拡散を直接的に抑止するための措置

① 発信者に対する警告表示

投稿自体は可能だが、不適切な内容を投稿しようとしている、又は直近で投稿したことが判明している旨の警告を表示する措置

② 収益化の停止

広告を非表示にしたり、広告報酬の支払を停止することにより、収益化の機会を失わせる措置

③ 情報の可視性に直接の影響がないラベルの付与

本人確認を行っていない利用者の明示³⁰等、情報発信者の信頼性等を見分けるためのラベルを付与する措置

④ 情報の可視性に一部影響するラベルの付与

ファクトチェック結果の付与等、情報の信頼性等を見分けるためのラベルを付与する措置

⑤ 表示順位の低下

投稿された情報を、受信者側のおすすめ欄等の表示候補から外したり、上位に表示されないようにする措置

⑥ 情報の削除

投稿された情報の全部又は一部を削除する措置

⑦ サービス提供の停止・終了、アカウント停止・削除

サービスの一部から強制退会、又はその一部の利用を強制終了し、新規投稿等できないようにする措置や、アカウントの一時停止又は永久停止（削除）を実施する措置

(2) 信頼できる情報の流通促進を通じて間接的に偽・誤情報の拡散を抑止するための措置

⑧ いわゆるプロミネンス

信頼できる情報を、受信者側のおすすめ欄に表示したり、上位に表示されるようにする措置

²⁹ 本検討会におけるプラットフォーム事業者ヒアリングの結果によると、情報伝送 PF 事業者は、このほかに例えば、再取得したアカウントにおける投稿制限、他のサービスや機能へのアクセス制限、プロフィールの編集要請、第三者が閲覧又はアクセスできない状態化、閲覧を望まない第三者への非表示等のコンテンツモデレーションを実施している。

³⁰ 本人確認を行った利用者のみを積極的に明示することで他と区別する措置を含む。

これらのコンテンツモデレーションの有効性については、具体的な偽・誤情報の流通・拡散状況等により、類型ごとに差があり得ることに留意が必要である。

例えば、収益化の停止（②）は、発信者の認識（主観的意図）が経済的インセンティブ目当てである、いわゆる「インプレッション稼ぎ」の投稿による偽・誤情報の流通・拡散に対しては有効と考えられる一方、イデオロギーや政治的意図を持って発信された偽・誤情報の流通・拡散に対する効果は限定的である可能性がある。

また、コンテンツモデレーションの類型ごとに、その実施を促進等するための適切な方策は異なり得ることに留意が必要である。

例えば、収益化の停止（②）は、情報伝送 PF 事業者の広告収入が間接的に偽・誤情報の発信・拡散主体の収入源となることを抑止する上で一定の効果が見込まれる一方、情報自体の可視性には直接の影響がない措置であると考えられる。

また、情報の可視性に影響しないラベルの付与（③）は、情報受信者における情報発信者の信頼性等の判断を容易にする効果が期待される一方、情報自体の可視性には直接の影響がない措置であると考えられる。

一方、情報の可視性に一部影響するラベルの付与（④）は、当該情報そのものの信頼性を否定するラベルを強制的に付すこと等を通じ、情報の可視性に一定の影響を与える措置であるが、発信者が投稿した情報に別の情報が加えられる措置であって、情報を削除する措置とは性質が異なると考えられる。

表示順位の低下（⑤）は、情報の可視性への影響が大きくなり得る措置であると考えられる³¹が、発信者の投稿自体は削除されずに残る上、特にレコメンデーション機能を介して優先表示されていた情報の優先表示を停止することを通じた表示順位の低下については、通常通りの伝送がされるに過ぎないと捉えることもできる。

情報の削除（⑥）は、情報の全部又は一部を不可視とする措置であると考えられる。

アカウント停止・削除（⑦）は、既に流通する情報を不可視とするのみならず、将来的な表現行為の機会を与えない点において、事前抑制的な性格を有する措置であると考えられる。

このように、上記（1）（偽・誤情報の流通・拡散を直接的に抑止するための措置）で挙げた類型には、情報の可視性に直接の影響がないもの（①・②・③）や、情報の可視性に一部影響を与えるもの（④・⑤）、情報の全部又は一部を不可視とするもの（⑥・⑦）等、情報の可視性に与える影響の大小に差異があり得るところ、一般的には、情報の可視性に与える影響が小さいものほど、表現の自由への制約は小さくなると言える。このことも踏まえ、情報伝送 PF 事業者においては、対象とする偽・誤情報の範囲（上記1.）に照らし比例的な対応を検討することが適当である。

一方、WG における議論では、いわゆるプロミネンス（⑧）について、いかなる情報を「信頼できる」ものと考え³²、また具体的にいかなる場合に「受信可能性の向上」がなさ

³¹ 特に透明性の確保が不十分な状態で行われる表示順位の低下（表示順位が低下したこと及びその理由が発信者に通知されないなど）は、一般に「シャドウバン」(shadowban)と呼ばれ、米国の一部州ではこれを規制する動きも存在する（例えば、フロリダ州法 SB7072。ただし、合憲性を巡り係争中）。

³² なお、総務省「令和4年度情報通信メディアの利用時間と情報行動に関する調査報告書」によると、世の中の動きやできごと

れたと認められるべきかなど、その実施を促進等するための方策の設計に当たって特有の課題が存在するため、他の類型（①から⑦まで）とは異なる取扱いが必要との意見があった。具体的な促進等の方策の在り方について、今後、更なる検討が必要である³³。

以上の措置をはじめとして、上記1. で述べた「対応を検討すべき「偽・誤情報」の範囲に関する基本的な考え方」に含まれる偽・誤情報の流通・拡散を抑止するため、制度整備も含め、情報伝送 PF 事業者によるコンテンツモデレーションの実施を促進等するとともに、その確実な実施を図ることが適当である。

3. 偽・誤情報に対するコンテンツモデレーションの実効性確保に向けた方策

(1) 総論

情報伝送 PF 事業者によるコンテンツモデレーションの実効性を確保するための方策としては、例えば次のようなものが考えられる。

- | |
|---|
| <ul style="list-style-type: none">① コンテンツモデレーションに関する透明性の確保を通じた過不足ない実施の確保② コンテンツモデレーションに関する対応の迅速化を通じた実施の促進③ 収益化の停止、ラベルの付与等、情報の可視性に直接の影響がないコンテンツモデレーションを中心に、体制を整備して確実に実施④ 情報の削除、アカウントの停止・削除等、情報の可視性への影響が大きいコンテンツモデレーションについて、体制を整備して確実に実施⑤ 上記①から④までの組合せによる対応 |
|---|

これらの方策については、以下（2）から（7）までにおいて詳述するように、対象とする偽・誤情報の特性・性質（権利侵害性その他の違法性・有害性、流通することによる社会的影響の重大性、誤りの明白性）等に応じた対応とすることが適当である。

(2) コンテンツモデレーションに関する透明性の確保を通じた過不足ない実施の確保

上記（1）①の方策は、情報伝送 PF 事業者によるコンテンツモデレーションが日本国内において過不足なく実施されていることに関し、利用者を含む社会一般が確認し、情報伝送 PF サービスに対する信頼性を向上させるとともに、利用者において各サービスを比較・選択し、自律的な意思決定に基づいて情報の発信・受信を行うことを可能にするための方策である。この方策については、特定電気通信による情報の流通によって発生する権利侵害等への対処に関する法律（以下「情報流通プラットフォーム対処法」とい

について信頼できる情報を得るために最も利用するメディアについて、全年代では「テレビ」を挙げる人が 53.1%で最も多く、また、各メディアの信頼度については、各年代で「新聞」又は「テレビ」と回答する人の割合が高かった。特に災害時の情報収集手段については、「テレビ」を挙げる人が7割前後で最も多いとする調査結果も存在する(株式会社 NTT ドコモ「モバイル社会白書 2023 年版」)。

³³ 特に災害発生時等における情報伝送 PF サービスがもたらす社会的影響の軽減措置としてのプロミネンスの重要性について、II 2. 参照。

う。) ³⁴における透明化規律³⁵を参考としつつ、次の (i) から (iv) までの対応を中心に具体化を進めることが適当である。

- | |
|--|
| <ul style="list-style-type: none">(i) コンテンツモデレーションに関する基準 (どのような場合にどの種類のコンテンツモデレーションを行うか) や手続を事前に策定・公表(ii) コンテンツモデレーションの実施要否等の判断に関与する人員等の体制に関する情報を公表(iii) 上記 (i) の基準の運用状況を事後に公表(iv) コンテンツモデレーションを実施した場合に、その旨及び理由並びに不服申立ての方法を発信者に通知 |
|--|

その際、これらの方策の具体化に当たっては、公表・通知等すべき具体的事項の詳細や、対象とするコンテンツモデレーションの範囲 (特に、収益化の停止、ラベルの付与等、情報の可視性に直接の影響がないコンテンツモデレーションや、情報の可視性に一定の影響を与える表示順位の低下を対象に含めるか否か) 等について、今後、更なる検討が必要である。

この点については、外国に本社を置く情報伝送 PF 事業者 (以下「外資系」という。) を含む情報伝送 PF 事業者によるコンテンツモデレーションが日本国内において過不足なく適正に実施されていることに関し、利用者を含む社会一般が確認し、情報伝送 PF サービスに対する信頼性を向上させるという目的に照らして具体化することが適当である。

特に (ii) の具体化に当たっては、日本語や日本の社会・文化・法令を理解する者の配置状況に関する情報や、情報伝送 PF 事業者によるコンテンツモデレーションの実施要否等の判断プロセスに AI を含む自動的手段が用いられる場合における当該手段の実効性に関する情報 (例えばエラー率や、発信者からの不服申立てを受けて判断に変更を加えた件数・割合等) を公表することも含めて検討することが適当である。

その際には、外資系も含む情報伝送 PF 事業者から情報の提供を受けつつ、AI を含む自動的手段による判断プロセスや判断結果等の実態を、より詳細に把握・分析することも必要である。

(3) コンテンツモデレーションに関する対応の迅速化を通じた実施の促進

上記 (1) ②の方策は、情報伝送 PF 事業者によるコンテンツモデレーションの実施の有無及び内容に関する判断の自主性は維持しつつ、当該判断を含む対応そのものの確実な実施及び迅速化を図る方策である。この方策については、情報流通プラットフォーム

³⁴ 令和6年改正後の特定電気通信役務提供者の損害賠償責任の制限及び発信者情報の開示に関する法律 (以下、改正前の法律を「プロバイダ責任制限法」という。) をいう。

³⁵ 情報伝送 PF 事業者を含む特定電気通信役務提供者 (プロバイダ責任制限法第2条第3号) による情報の削除やアカウントの停止・削除 (送信防止措置) について、大規模な特定電気通信役務提供者に対し、削除等の基準の策定・公表 (運用状況の公表を含む。) や削除等した場合の発信者への通知といった措置を義務付けるもの (情報流通プラットフォーム対処法第 27 条から第 29 条まで)。

対処法における迅速化規律³⁶等³⁷を参考としつつ、次の(i)から(iv)までの対応を中心に具体化を進めることが適当である。

- (i) 外部からのコンテンツモデレーション申出・要請窓口を整備・公表
- (ii) 上記(i)の窓口を通じて申出・要請があった場合に、一定期間内にコンテンツモデレーションの実施の要否・内容を判断し、申請者に判断結果(及び不服申立ての方法)を通知
- (iii) コンテンツモデレーションの実施の要否・内容を判断するための体制(コンテンツモデレーションに関する不服申立てを受け付ける体制を含む。)を整備
- (iv) 一定の条件(例えば、行政機関等の特定の第三者からの申出・要請を受けて実施した場合等)の下で行ったコンテンツモデレーションにより発信者が被った損害について、情報伝送PF事業者を免責

その際、これらの方策の具体化に当たっては、濫用的な申出・要請から生じる情報伝送PF事業者の実務上の負担に配慮する観点から、対象とする偽・誤情報の特性・性質に応じ、いかなる主体からの申出・要請を契機としたコンテンツモデレーションの実施を促進すべきかについて、以下を基本的な方向性としつつ、今後、更なる検討が必要である³⁸。

① 他人の権利を侵害する違法な偽・誤情報

他人の権利を侵害する違法な偽・誤情報については、(既に情報流通プラットフォーム対処法に規律が置かれたように、) まずは自己の権利を侵害されたとする者(被侵害者)からの申出・要請³⁹を契機としたコンテンツモデレーションについて、上記(i)から(iv)

³⁶ 特定電気通信役務提供者による(偽・誤情報に限らない)他人の権利を侵害する情報の削除(送信を防止する措置)について、大規模な特定電気通信役務提供者に対し、被侵害者からの削除申出窓口・手続の整備・公表、被侵害者からの削除申出への対応体制の整備、被侵害者からの削除申出に対する判断・通知(原則、一定期間内)といった措置を義務付けるもの(情報流通プラットフォーム対処法第23条から第26条まで)。

³⁷ 例えば、特定電気通信役務提供者による(偽・誤情報に限らない)他人の権利を侵害する情報の削除(送信を防止する措置)について、プロバイダ責任制限法第3条第2項及び第4条、私事性的画像記録の提供等による被害の防止に関する法律(以下「リベンジポルノ防止法」という。)第4条、性をめぐる個人の尊厳が重んぜられる社会の形成に資するために性行為映像制作物への出演に係る被害の防止を図り及び出演者の救済に資するための出演契約等に関する特則等に関する法律(以下「AV新法」という。)第16条等が、被侵害者からの申出を契機とした情報の削除により発信者が被った損害について一定の条件の下で特定電気通信役務提供者を免責する旨を定めている。

³⁸ 特に行政法規に抵触する違法な偽・誤情報について、行政機関(及び当該行政機関の委託や認証を受けた機関)だけでなく、司法機関や地方自治体、さらにはその他一般私人からの申出・要請を契機としたコンテンツモデレーションの実施を促進することも検討すべきという意見があった。なお、EUデジタルサービス法(Digital Services Act。以下「DSA」という。)は、ホスティングサービスの提供者に対し、一般私人を含むあらゆる個人や団体から違法コンテンツと思われる情報が存在する旨の通報を受け付ける体制の整備を求める(第16条)と同時に、ホスティングサービスを含む仲介サービスの提供者に対し、司法機関や行政機関から違法コンテンツについて何らかの対応を求める措置命令を受けた場合に、不当に遅滞なく当該司法・行政機関へ対応結果を報告することを求めている(第9条)。

³⁹ 他方、WGにおける議論では、被侵害者以外の第三者(例えば、行政機関や地方自治体が設置する審査会等)からの申出・要請を契機としたコンテンツモデレーションの迅速化を通じた実施の促進のための方策の在り方についても検討が必要との意見があった。こうした方策の要否及び具体的な在り方については、濫用的な申出・要請の防止という観点も含め、今後、更なる検討が必要である。

までのような対応を中心に、その迅速化を通じた実施の促進のための方策の在り方について具体化を進めることが、濫用的な申出・要請のおそれも小さく適当である。

② 行政法規に抵触する違法な偽・誤情報

行政法規に抵触する違法な偽・誤情報については、対応の迅速化を通じた実施の促進を図ることとする場合、違法性の判断能力の観点から、当該行政法規を所管する行政機関（当該行政機関の委託や認証を受けた機関を含む。）からの申出・要請を契機としたコンテンツモデレーションについて、上記（i）から（iv）までのような対応を中心に具体化を進めることが基本的には適当である⁴⁰。なお、この点については、今後、具体的な行政法規⁴¹を洗い出しつつ、具体化を進めることが適当である。

ただし、この場合の対応については、前提として、行政機関による恣意的な申出・要請を防止し、透明性・アカウントビリティを確保するとともに、過度な申出・要請に対し発信者や情報伝送 PF 事業者を救済するため、次のような方策を併せて検討することが不可欠である⁴²。

- （ア）行政機関において、申出・申請に関する手続等（事後救済手段を含む。）を事前に策定・公表
- （イ）行政機関において、実際に行った申出・申請の状況を事後的に公表
- （ウ）申出・要請に応じて実施されたコンテンツモデレーションにより発信者が被った損害について、情報伝送 PF 事業者を免責
- （エ）コンテンツモデレーションを実施した情報伝送 PF 事業者において、行政機関の名称等の情報を発信者に通知

その際、このような行政機関による申出・申請に関する透明性・アカウントビリティを確保するための方策の具体化に当たっては、行政手続に関する既存の法的枠組み（例えば、行政指導指針の策定・公表義務⁴³、国家賠償法等）との関係について、今後、更な

⁴⁰ なお、特定電気通信役務提供者による医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律（以下「薬機法」という。）で禁止される医薬品等に関する虚偽・誇大広告や無承認医薬品等の広告である情報の削除（送信を防止する措置）については、既に同法第 72 条の6が、厚生労働大臣又は都道府県知事による要請を契機とした削除により発信者が被った損害について、一定の条件の下で特定電気通信役務提供者を免責する旨を定めている。

⁴¹ 保護法益の重要性（及びその表れとしての違反した場合の処分の重大性）、所管行政機関の専門性等が考慮要素となり得る。

⁴² みずほリサーチ&テクノロジーズ株式会社「令和5年度 国内外における偽・誤情報に関する意識調査 結果紹介」（本検討会第 18 回(WG 第 17 回)会合(2024 年5月9日)資料 18—2—1)によると、政府がデジタルプラットフォームサービス事業者(SNS、検索、動画投稿・共有など)に対してコンテンツモデレーションを要求するに当たって必要な原則を尋ねたアンケート調査の結果は、日本を除くすべての対象国(米国、英国、フランス、韓国、豪州)で「透明性」と回答した割合が最も高く、日本においても「法的根拠」(51.4%)、「正当な目的」(51.1%)に次ぐ3位(44.2%)であった。

⁴³ 同一の行政目的を実現するため一定の条件に該当する複数の者に対し行政指導をしようとするときは、行政機関は、あらかじめ、事案に応じ、行政指導指針を定め、かつ、行政上特別の支障がない限り、これを公表しなければならない(行政手続法第 36 条)。なお、薬機法第 72 条の5第2項に基づく厚生労働大臣又は都道府県知事の要請については、「インターネット上の無承認医薬品及び指定薬物等に係る広告監視指導について」(平成 26 年 12 月 17 日薬食監麻発 1217 第1号厚生労働省医薬食品局監視指導・麻薬対策課長通知)において、都道府県知事が特定電気通信役務提供者に対して送信防止措置の実施を要請する場合の手順・方式(事後に厚生労働省宛てに連絡することを含む。)を規定・公表している。

る検討が必要である。

③ 権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい偽・誤情報

権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい偽・誤情報は、上記①の偽・誤情報（他人の権利を侵害する違法な偽・誤情報）や上記②の偽・誤情報（行政法規に抵触する違法な偽・誤情報）とは異なり、違法性のない情報であることから、これに対するコンテンツモデレーションについては、IIで後述する「情報伝送 PF サービスが与える情報流通の健全性への影響の軽減に向けた方策」としての影響評価・軽減措置の確実な実施を求める枠組みの活用を含め、情報伝送 PF 事業者による取組を促す観点が重要である。

一方、こうした取組の実効性を補完する観点からも、以下の方向性を基本としつつ、上記（i）から（iv）までのような対策を含め、情報伝送 PF 事業者によるコンテンツモデレーションの迅速化を通じた実施の促進のための方策の在り方について、利用者の表現の自由の保護とのバランスを踏まえながら具体化を進めることが適当である。

（ア）情報の可視性への影響が大きいコンテンツモデレーション

上記のとおり、権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい偽・誤情報は、違法性のない情報であることから、第三者からの申出・要請を契機とした可視性への影響が大きいコンテンツモデレーション（情報の削除、アカウント停止・削除等）について上記（i）から（iv）までのような対策の実施を制度的に担保することは、そうした措置の実施により、違法性のない情報に関する利用者の表現の自由を実質的に制約するおそれがあるため、当該偽・誤情報の特性・性質（有害性や社会的影響の大小・明白性、誤りが含まれることの明白性）を考慮しつつ、引き続き慎重な検討が必要である。

一方、情報伝送 PF 事業者が自主的な判断により、こうした情報の流通・拡散を抑止するため、利用規約等に基づいて、情報の可視性への影響が大きいコンテンツモデレーションの措置を講ずることは妨げられるものではない。こうした取組を促す観点からは、IIで後述する「情報伝送 PF サービスが与える情報流通の健全性への影響の軽減に向けた方策」として情報伝送 PF 事業者による影響評価・軽減措置の確実な実施を求める枠組みを活用することが適当である。

（イ）情報の可視性に直接の影響がないコンテンツモデレーション等

上記のとおり、権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい偽・誤情報に対するコンテンツモデレーションについては、IIで後述する「情報伝送 PF サービスが与える情報流通の健全性への影響の軽減に向けた方策」としての影響評価・軽減措置の確実な実施を求める枠組みの活用を含め、情報伝送 PF 事業者による取組を促す観点が重要である。一方、こうした取組の実効性を補完する観点からも、特に情

報の可視性に直接の影響がないコンテンツモデレーション（収益化の停止、ラベルの付与等）を中心とした対応について、発信者や情報伝送 PF 事業者以外の特定の第三者（当該情報付近に広告を表示された広告主、ファクトチェック機関、行政機関等）から申出・要請があった場合における上記（i）から（iv）までのような対策を含め、**情報伝送 PF 事業者によるコンテンツモデレーションの迅速化を通じた実施の促進のための方策の在り方について、利用者の表現の自由の保護とのバランスを踏まえながら具体化を進めることが適当**である。

なお、当該偽・誤情報の特性・性質（有害性や社会的影響の大小・明白性、誤りが含まれることの明白性）に応じた適切な申出・要請主体の範囲や対象とするコンテンツモデレーションの範囲等の詳細については、今後、更なる検討が必要である。

（４）情報の可視性に直接の影響がないものを中心としたコンテンツモデレーションの確実な実施

上記（１）③の方策は、コンテンツモデレーションのうち、情報の可視性に直接の影響がない収益化の停止、ラベルの付与等を中心に、情報伝送 PF 事業者による確実な実施を担保するための方策である。

特に、権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい偽・誤情報に対するコンテンツモデレーションについては、上記（３）のとおり、IIで後述する「情報伝送 PF サービスが与える情報流通の健全性への影響の軽減に向けた方策」としての影響評価・軽減措置の確実な実施を求める枠組みの活用を含め、情報伝送 PF 事業者による取組を促す観点が重要である。一方、こうした取組の実効性を補完する観点からも、情報の可視性に直接の影響がない方策を中心としたコンテンツモデレーションを確実に実施する方策について、利用者の表現の自由の保護とのバランスを踏まえつつ、発信者や情報伝送 PF 事業者以外の特定の第三者（当該情報付近に広告を表示された広告主、ファクトチェック機関、行政機関等）からの申出・要請を契機としたコンテンツモデレーション（収益化の停止、ラベルの付与等）の実施も含め、具体化を進めることが適当である。

なお、WGにおける議論では、権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい偽・誤情報の一部について、脆弱な個人に対するレコメンデーションや広告ターゲティングの停止（これもコンテンツモデレーションの一類型に該当すると考えられる。）の確実な実施を担保することが適当とする意見もあった。こうした方策の適否については、情報伝送 PF サービスにおけるレコメンデーションや広告ターゲティングの実態を踏まえつつ、今後、更なる検討が必要である。

（５）情報の可視性への影響が大きいコンテンツモデレーションの確実な実施

上記（１）④の方策により、情報伝送 PF 事業者に対し、偽・誤情報の流通・拡散に対する対応として、情報の削除やアカウント停止・削除の確実な実施を罰則付きで義務付けて、その流通・拡散の抑止を制度的に担保することについては、情報伝送 PF 事業者による過度な削除やアカウントの停止・削除が行われることにより、利用者の表現の自由

を実質的に制約するおそれがあるため、対象とする偽・誤情報の特性・性質（権利侵害性その他の違法性・有害性、流通することによる社会的影響の重大性、誤りの明白性）を考慮しつつ、引き続き慎重な検討が必要である⁴⁴。

ただし、情報伝送 PF 事業者が利用者との契約に基づき、自主的に情報の削除やアカウント停止・削除の確実な実施を行うことは妨げられるものではない。

（６）違法性を有する偽・誤情報の発信を繰り返す発信者等への対応

明白な権利侵害性その他の違法性を有する偽・誤情報を繰り返し発信する者など、特に悪質な発信者に対する情報の削除やアカウント停止・削除を確実に実施する方策については、こうした対応の段階的な実施を担保することも含め、具体化を進めることが適当である。

なお、こうした方策を実施する具体的な要件等については、情報伝送 PF 事業者によって実施されているコンテンツモデレーションの実態等を踏まえつつ、今後、更なる検討が必要である。

（７）情報流通の態様に着目したコンテンツモデレーションの実施

令和6年能登半島地震では、X（旧 Twitter）の仕様変更（発信者への経済的インセンティブの付与開始）に伴い、閲覧数稼ぎが目的とみられる複製投稿（いわゆるコピー投稿）その他偽・誤情報を含む投稿が多数確認されたとの調査結果がある。このように、例えば、別の投稿を複製した投稿が高頻度で送信された場合等、送信された情報の内容そのものの真偽に着目するのではなく、情報流通の態様に着目したコンテンツモデレーションの実施の在り方についても、偽・誤情報の流通・拡散を抑止する観点も含め、具体化を進めることが適当である。

なお、具体的にどのような態様の情報流通を対象とするか等の詳細については、当該態様の情報流通によって発生又は増幅する影響及びリスクを特定しつつ、今後、更なる検討が必要である。

4. 偽・誤情報の発信を抑止するためのその他の方策

コンテンツモデレーションの実施を促進する以外に、偽・誤情報の発信を抑止するた

⁴⁴ なお、インターネット異性紹介事業を利用して児童を誘引する行為の規制等に関する法律第6条に定める禁止誘引行為（児童を異性交際の相手方となるように誘引すること等）に係るいわゆる出会い系サイトへの書き込みについては、既に同法第12条第1項が、当該出会い系サイトを運営する事業者に対し、当該禁止誘引行為が行われていることを知ったときに当該書き込みを削除（インターネットを利用して公衆が閲覧することができないようにするための措置）すべき義務を課している（ただし、義務違反に対する罰則は定められていない。）。また、青少年が安全に安心してインターネットを利用できる環境の整備等に関する法律第2条第3項に定める青少年有害情報（インターネットを利用して公衆の閲覧（視聴を含む。）に供されている情報であつて青少年の健全な成長を著しく阻害するもの）については、既に同法第21条が、情報伝送 PF 事業者を含む特定サーバー管理者（同法第2条第11項）に対し、その管理する特定サーバーを利用して他人により青少年有害情報の発信が行われたことを知ったとき又は自ら青少年有害情報の発信を行おうとするときに、当該青少年有害情報についてインターネットを利用して青少年による閲覧ができないようにするための措置（青少年閲覧防止措置）をとるべき努力義務を課している。

めの方策として、国内の既存の法制度の下では、発信者に対する刑罰による対応⁴⁵、民事法による対応⁴⁶、行政処分等による対応⁴⁷が講じ得ることとされている。

一方、情報伝送 PF 事業者を含む伝送主体においては、こうした発信者の法的責任を前提に、共同正犯（刑法第 60 条）や幫助犯（同法第 62 条）として処罰されたり、一定の条件の下で被侵害者から損害賠償請求や差止請求を受けたり、行政機関等から一定の処分・要請等⁴⁸を受けたりすることがあり得るところ、情報伝送 PF サービスがデジタル空間における情報流通の健全性に与える影響の大きさや、情報伝送 PF 事業者に期待される役割・責務等に照らすと、こうした既存の対応では実効性の観点から十分でない可能性が指摘されている⁴⁹。

このように、情報伝送過程で偽・誤情報の発信を抑止するための追加的な方策として、例えば次のようなものが考えられるが、いかなる方策が必要かつ適当か、また、その方策をどのように実現するかについては、情報伝送 PF 事業者による取組等の実態を踏まえつつ、その自主的な実施を促す方策も含め、今後、更なる検討が必要である。

- ① アカウント登録時やアカウント情報変更時等の本人確認の厳格化
- ② bot アカウントの抑止策の導入（アカウントの有料化等）
- ③ 特定のサービスアーキテクチャの採用（シェア、リポスト等の拡散機能の利用に複数のアクションを要求する等）

特に①については、発信者のトレーサビリティを確保することを通じ、偽・誤情報の発信に一定の抑止効果を期待できるものの、その効果は事後的な責任追及の可能性を前提とした間接的なものに過ぎないことを踏まえれば、偽・誤情報の発信を抑止するための方策としての実効性に疑義があること、匿名表現の自由への制約となり得ること等から、情報伝送 PF サービスにおけるアカウント登録時等の本人確認の実態を踏まえつつ、制度的な対応の要否について慎重な見極めが必要である⁵⁰。

⁴⁵ 例えば、名誉毀損（刑法第 230 条第 1 項）、侮辱（同法第 231 条）、信用毀損・偽計業務妨害（同法第 233 条）、詐欺（同法第 246 条）、私電磁的記録不正作出（同法第 161 条の 2 第 1 項）、著作権法違反（同法第 119 条第 1 項）、金融商品取引法上の風説の流布・偽計（同法第 158 条）、わいせつ電磁的記録等頒布（刑法第 175 条第 1 項）、児童ポルノの提供等（児童買春、児童ポルノに係る行為等の規制及び処罰並びに児童の保護等に関する法律第 7 条第 6 項）、公職の候補者へのなりすまし（公職選挙法第 235 条第 2 項）、公職の候補者に関する虚偽事項の公表（同条第 1 項）等の罪に当たるとして処罰されることがあり得る。

⁴⁶ 例えば、名誉権、名誉感情、営業上の利益、肖像権、パブリシティ権、プライバシー権、アイデンティティ権等の権利又は法律上保護される利益を侵害するとして、被侵害者から損害賠償請求や差止請求がなされることがあり得る。

⁴⁷ 例えば、不当景品類及び不当表示防止法に基づく表示規制（同法第 5 条）、医師に対する行政処分（医師法第 7 条）、弁護士に対する懲戒処分（弁護士法第 56 条）、条例に基づくヘイトスピーチ等を行った者の氏名等の公表等。

⁴⁸ 警察庁の委託を受けたインターネット・ホットラインセンター（IHC）による違法情報、自殺誘引等情報及び重要犯罪密接関連情報の削除要請、法務省人権擁護機関による人権侵害情報の削除要請等。

⁴⁹ この点、例えばドイツでは、裁判例上、情報伝送 PF 事業者において、他人の権利を侵害するコンテンツが流通していることについて通知を受けた場合、当該通知されたコンテンツそのもののみならず、通知から理解した権利侵害の内容に基づき対象となる他のコンテンツをも特定した上で、対応する義務を負う場合がある（この義務は、技術的・経済的観点から合理的に可能な手段を全て講じた場合に限り免責される）とされており、情報伝送 PF 事業者に対し、より積極的な義務を課す法制度が存在すると言える。

⁵⁰ 他方で、本人確認自体は任意としつつ、前述のとおり、本人確認を行ったアカウントのみを積極的に明示することで他と区別する措置については、コンテンツモデレーションの一類型として、その実効性確保に向けた方策を別途検討し得る。

II 情報伝送 PF サービスが与える情報流通の健全性への影響の軽減に向けた方策の在り方

情報伝送 PF サービスは、デジタル空間における情報流通の主要な場となっており、その中で偽・誤情報その他真偽が不確かな情報や、事実ではあるが人を害する悪意を持って発信された情報（以下「偽・誤情報等」と、総称する。）が流通・拡散することにより、人の生命、身体又は財産への影響、人格権やその他基本的人権への影響、健全な民主主義への影響その他の社会的混乱等の実空間への影響が発生・増幅し得る。

このような社会的影響については、権利侵害性その他の違法性を有する偽・誤情報等の流通・拡散によって生じ得ることはもちろん、それ以外の偽・誤情報等の流通・拡散によって生じる場合もあるところ、特に後者のような、権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい偽・誤情報等の流通・拡散による影響の軽減に向けた方策としては、表現の自由への制約となるようなコンテンツモデレーションを直接的に促進等することよりも、情報の可視性に直接の影響がないコンテンツモデレーションの実効性を確保しつつ（I 参照）、当該情報の有害性や社会的影響の重大性の軽重に応じ、情報伝送 PF 事業者の自主性・裁量を一定程度確保しながら、有効な影響軽減措置の検討及び実施を促していくことが必要かつ適当な場面が多いと考えられる⁵¹。

以上に加え、第1章や第5章で述べたデジタル空間における情報流通の現状や、情報伝送 PF 事業者に期待される役割・責務及びこれまでの取組状況等を踏まえると、制度整備も含め、情報伝送 PF 事業者に対して以下の具体的措置を求めることが適当である。

1. 情報伝送 PF 事業者による社会的影響の予測・軽減措置の実施

情報伝送 PF 事業者は、デジタル空間における情報流通に関する役割・責務等を踏まえ、自らが設計するサービスアーキテクチャ（サービスに組み込まれたアルゴリズムを含む。）や利用規約等を含むビジネスモデルがもたらす社会的影響の軽減に向け、将来にわたる社会的影響を事前に予測し、その結果を踏まえて、影響を軽減するための措置（サービスアーキテクチャの変更、利用規約等の変更、コンテンツモデレーションの方法・プロセスの変更、レコメンデーション機能の変更等を通じた措置）を検討・実施すること⁵²が適当である。

特に、WG における議論では、情報伝送 PF サービスのレコメンデーションシステムや広告ターゲティング技術を通じ、個人の特性や状況に応じた脆弱性に着目してコンテンツやデジタル広告の表示先・表示順位が決定されることにより、偽・誤情報等の流通・拡散等によるものをはじめとする社会的影響が深刻化するリスクを指摘する意見があっ

⁵¹ 他方で、権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい偽・誤情報等の中に、情報伝送 PF 事業者自身による影響軽減措置の検討及び実施よりも、むしろ可視性への影響が大きいものも含むコンテンツモデレーション等の直接的な促進等を図るべき領域が存在するかについては、今後、更なる検討が必要である。

⁵² 同様に、サービス等が外部的な権利・利益に対して将来与えるであろう影響等を事前に予測し、あらかじめ軽減措置の実施を求める既存の法的枠組みとして、例えば、金融商品取引法に基づく内部統制報告制度、行政手続における特定の個人を識別するための番号の利用等に関する法律に基づく特定個人情報保護評価、環境影響評価法に基づく環境影響評価、EU DSA に基づくシステミックリスクの評価、英国オンライン安全法(Online Safety Act 2023)に基づくリスク評価等が存在する。

たところ、こうしたリスクを含め、情報伝送 PF サービスのアーキテクチャに起因する社会的影響を軽減するための方策として、最低限必要な具体的措置の実施を直接的に促進等しつつ、それに加え、当該サービスアーキテクチャを自ら設計する情報伝送 PF 事業者自身に影響予測と軽減措置の確実な実施を求める枠組みの具体化を進めることが適当である⁵³。

その際、情報伝送 PF サービスの利用者及び情報伝送 PF 事業者自身の表現の自由への過度の制約を回避しつつ、影響予測に用いられる指標の客観性の確保を通じて情報伝送 PF 事業者による恣意的な影響予測を防ぎ、かつ、情報伝送 PF 事業者が実施する軽減措置の実効性を担保するため、次のような方策を中心に具体化を進めることが適当である。

- ① 政府による基本的な制度設計の下、民産学官のマルチステークホルダーが、その協議に基づき、影響予測の視点や軽減措置の実施項目を定めた実施指針を策定・公表
- ② 情報伝送 PF 事業者において、上記①の実施指針に従って影響予測を実施し、その結果やそれに基づき講じた措置の内容を公開又は行政機関その他の民産学官のマルチステークホルダーに報告
- ③ 民産学官のマルチステークホルダーが、その協議に基づき、上記②の影響予測の結果やそれに基づき講じられた措置の内容を検証・評価

上記①（実施指針の策定・公表）及び③（検証・評価）の場面における民産学官の役割分担の具体化に当たっては、政府（官）による大枠の制度設計の下、少なくとも表現の自由との関わりがある部分は民産学のステークホルダーが主として協議・決定を行い、機能不全が生じた場合に補完的に政府が関与するという、段階的・多層的な形を基本とすることが適当である。

この制度設計に当たっては、偽・誤情報等が流通・拡散することによる社会的影響（人の生命、身体又は財産への影響、人格権やその他基本的人権への影響、健全な民主主義への影響その他の社会的混乱等の実空間への影響）を予測し、第5章で述べた「情報流通の健全性」に関する基本理念⁵⁴に照らし必要かつ十分な軽減措置の実施を確保するという目的に沿ったものとすることが重要である。特に、流通・拡散する偽・誤情報等の内容や態様、及び技術・サービス等の外的要因を含むデジタル空間における情報流通を巡

⁵³ 他方で、情報伝送 PF サービスのレコメンデーションシステムや広告ターゲティング技術がデジタル空間における情報流通の健全性にもたらすリスクへの対応として、情報伝送 PF 事業者による影響予測及び軽減措置の実施に加えて、いかなる方策が必要かつ適当かについては、今後、更なる検討が必要である。なお、本検討会における議論では、オブザーバー団体（一般社団法人 MyDataJapan）から、

・ 広告ターゲティングやレコメンデーションに使われるデータを収集する際に、情報を取得するタッチポイントにおいて、誰が何の目的で収集し、どのようなリスクを伴うのかを分かりやすく生活者に対し通知し、取得の拒否などのコントロール性を確保することを義務付ける

・ 社会問題・政治に関する広告や投稿に関しては、そのターゲティングやレコメンデーションシステムで用いることのできるパーソナルデータを制限する

という提案があった（本検討会第19回（WG第19回）会合（2024年5月15日）配付資料19-2-2）。

⁵⁴ 表現の自由と知る権利の実質的保障及びこれらを通じた法の支配と民主主義の実現、安心かつ安全で信頼できる情報流通空間としてのデジタル空間の実現、国内外のマルチステークホルダーによる国際的かつ安定的で継続的な連携・協力。

る状況は絶えず変化しており、今後も変化し得るところ、そうした変化に応じ、予測すべき社会的影響や必要かつ十分な軽減措置の内容も継続的にアップデートが必要となることに留意が必要である。

一方、具体的な設計として、例えば、マルチステークホルダーによる実施指針の策定や検証・評価プロセスの具体的な在り方等については、表現の自由との関わりがあるために民産学のステークホルダーが主として協議・決定すべき部分と、政府が関与すべき部分をどのように切り分けるかを含め、今後、更なる検討が必要である。

また、上記②の方策の具体化に当たっては、公開又は報告される情報の内容・粒度はある程度統一され、各社の取組を比較できるものとするのが適当である。

さらに、上記③の方策の具体化に当たっては、マルチステークホルダーの検証・評価能力を確保するため、情報伝送 PF 事業者からマルチステークホルダー又はこれを構成する研究者・研究機関に対して関連する情報やデータ（例えば、投稿に関するメタデータや、サービスに組み込まれたアルゴリズムに関する情報等）が確実に提供されることを制度的に担保する仕組みを検討するのが適当である⁵⁵。その際、当該情報やデータに含まれる個人情報や機密情報の適正な取扱いも併せて担保されることが適当である⁵⁶。

上記①から③までの枠組みの具体的な設計については、これらの観点を踏まえつつ、今後、更なる検討が必要である⁵⁷。

以上に加え、情報伝送 PF 事業者による影響予測と軽減措置の確実な実施を担保する枠組みの具体化に当たっては、マルチステークホルダーが策定した実施指針や検証・評価の結果を情報伝送 PF サービスに確実に反映するための方策についても、併せて具体化を図ることが適当である。

こうした方策としては、例えば次のようなものが考えられるが、いかなる方策が必要かつ適当か、実効性を担保するためにどのような執行手段を整備するかについても、上記①から③までの枠組みの具体的な設計と並行して、今後、更なる検討が必要である。

- | |
|---|
| <ul style="list-style-type: none">(i) 実施指針を踏まえた行動規範を各情報伝送 PF 事業者が策定・公表(ii) マルチステークホルダーによる検証・評価の結果を公表 |
|---|

⁵⁵ このような、研究者・研究機関によるデータに基づく研究を通じ、情報伝送 PF がもたらす社会的影響(予測)に関するマルチステークホルダーの検証・評価能力を確保する仕組みについては、EU において、DSA に基づき超大規模オンラインプラットフォーム(VLOP)及び超大規模オンライン検索エンジン(VLOSE)に研究者等へのデータアクセス提供義務が課されていること(第40条)、さらに、こうしたデータアクセスを前提に、DSAの施行をサポートするための科学的及び技術的専門知識を提供し、オンラインプラットフォームや検索エンジンによって展開されるアルゴリズムシステムの影響に関する更なる研究を行うことを目的として2023年4月に設立された欧州アルゴリズム透明性センター(European Centre for Algorithmic Transparency)が、データアクセスを通じて実施される研究の知識ハブとして機能していることなどが参考になる。https://algorithmic-transparency.ec.europa.eu/index_en

⁵⁶ 個人情報や機密情報の適正な取扱いを担保するための具体的な方策については、対象となる情報やデータの受領者における情報管理体制の確認方法、情報漏えい等が生じた場合の責任の在り方等の論点が検討課題になり得ると考えられる。

⁵⁷ WG における議論では、独立した第三者による監査等を段階的に求める仕組みを導入することも含めて検討すべきとの意見があった。

- (iii) 検証・評価の結果を踏まえ、情報伝送 PF 事業者において影響軽減策を策定・公表
- (iv) 検証・評価の結果の中に具体的な影響軽減策を含めた上で、情報伝送 PF 事業者において当該影響軽減策を確実に実施

2. 特に災害発生時等における対応

災害発生時、感染症流行時、テロ発生時等、限られた時間の中で多くの人の間で適時に正確な情報の共有が求められる場面における情報収集・伝達手段としての情報伝送 PF サービスの存在感や公共的役割は高まっている。

こうした場面では、偽・誤情報等の流通・拡散による社会的影響が質的にも量的にも大きくなり得、また、特にコンテンツ（それに伴うデジタル広告を含む。）の閲覧数等に応じて発信者に経済的インセンティブを付与する仕組みを取り入れている情報伝送 PF サービスにおいて、経済的インセンティブ目当てのいわゆる「インプレッション稼ぎ」の投稿が増加するなど、情報流通に伴う社会的影響のリスクが高まると言える。

その中で、情報伝送 PF 事業者は、偽・誤情報等の流通・拡散による社会的影響を抑止するとともに、公共的役割として人々にとって必要な正確な情報を迅速かつ適時・確実に伝送すべく、平時から計画を立て、災害発生時等には当該計画に従って即応することが適当である。

特に、前記1. の影響予測と軽減措置の確実な実施について制度整備を含む具体化を進めるに当たり、情報伝送 PF 事業者が災害発生時等に備えて立案すべき計画の一部として、災害発生時等に自社のビジネスモデルがもたらす社会的影響を平時から予測し、有効な軽減措置をあらかじめ講じておくことが適当である。

この場合の軽減措置としては、例えば次のような措置が考えられるところであり、上記1. の制度設計や実施指針にどのように反映するかも含め、更なる具体化を進めることが適当である。

- ① 信頼できる情報源からの情報の伝送確保（プロミネンス）及びその基準の明確化
- ② 災害発生時等に特に適用されるコンテンツモデレーション（収益化の停止を含む。）に関する利用規約等の整備
- ③ 上記②の利用規約等を踏まえた適正な対応を実施するために必要な人員等の体制の整備とその状況の公表
- ④ 上記②の利用規約等の運用状況を事後に公表（平時における定期的な公表とは別途、より短期的な運用状況を公表）
- ⑤ 関係機関（行政機関、ファクトチェック機関、研究機関、偽・誤情報等付近に広告を表示された広告主等）との連絡窓口の明確化と、当該窓口を通じた迅速かつ緊密な連携・情報共有（偽・誤情報等の流通・拡散等による社会的影響の大小や軽減措置の有効性を検証するに足りるデータの提供を含む。）

ただし、表現の自由に対する過度の制約を避ける観点から、この場合の「災害発生時等」に該当するための要件や、始期・終期を誰がどのような手続で決定するのかについて

ては、明確に定められる必要があり、少なくとも始期・終期の要件、誰がどのような手続で決定するのか等については、マルチステークホルダーによる平時からの協議で決定することが適当である。

こうしたマルチステークホルダーによる協議・決定のプロセスの具体については、透明性を確保しつつ、具体的にどのように定めることが必要かつ適当かという観点から、今後、更なる検討が必要である。

以上のほか、災害発生時等における更なる対応（例えば、上記①から⑤までの影響軽減措置のうち一部の確実な実施等）については、個々の場面ごとに平時とは区別した追加的な対応が求められる具体的な理由を整理しつつ、今後、引き続き検討が必要である。

Ⅲ マルチステークホルダーによる連携・協力の枠組み整備の在り方

マルチステークホルダーによる連携・協力に関する具体的な方策の在り方については、デジタル空間における情報流通の健全性確保に向けたガバナンスの在り方等に関し、国内外の民産学官のマルチステークホルダーが相互に連携・協力しながら安定的かつ継続的に議論・検討する枠組みについて、マルチステークホルダーにより構成される協議会の設置に関する制度整備も含め、具体化を進めることが適当である。

1. 連携・協力の目的

上記のとおり、デジタル空間における情報流通の健全性確保に向けたガバナンスの在り方等について、国内外の民産学官のマルチステークホルダーが相互に連携・協力しながら安定的かつ継続的に議論・検討する枠組みを整備することで、そうした議論・検討及びその結果に基づく取組を推進することが重要であるところ、特に、Ⅱで述べた情報伝送 PF 事業者による影響予測及び軽減措置の実施に関する指針の策定等、デジタル空間における情報流通に関わる特定のステークホルダーに影響を与える重要な協議・決定については、当該ステークホルダーが参加する場において透明性を伴う形で行われることにより、民主的な正統性を確保することが適当である。

具体的には、表現の自由をはじめとする各ステークホルダーの権利利益に配慮しつつ、自主的な取組を推進する観点から、政府による大枠の制度設計の下、当該影響を受けるステークホルダーを含む民産学のステークホルダーが一次的に協議・決定を担い、機能不全が生じた場合に補完的に政府が関与するという、段階的・多層的な形を基本として、次のような目的で連携・協力することが適当である。

- ① デジタル空間における情報流通の健全性確保に向けて各ステークホルダーに求められる取組に関するガイドライン（情報伝送 PF 事業者による影響予測及び軽減措置の実施に関する指針を含む。）や行動規範を策定・推進
- ② 情報伝送 PF 事業者による社会的影響の予測結果や軽減措置の内容について検証・評価
- ③ 情報伝送 PF 事業者に対してコンテンツモデレーション等の対応を要請し、過不足がなかったか等を事後に検証
- ④ その他、デジタル空間における情報流通の健全性確保に向けた各ステークホルダーの役割・取組・協力関係等に関する協議を実施

以上と併せ、民産学のマルチステークホルダー（場合により官を含む。）の間で、次のような目的での連携・協力も具体化することが適当である。

- ⑤ 情報交換（例えば、偽・誤情報の流通・拡散の実態や傾向、対策に向けた取組（ベストプラクティス）、技術動向等の情報交換）

- ⑥ 情報の収集・分析・活用（例えば、ファクトチェック結果に基づき、偽・誤情報の発信源・拡散主体の情報を収集・分析し、関係者間で共有・活用⁵⁸）
- ⑦ ステークホルダー間のパートナーシップの確立（例えば、情報伝送 PF 事業者とファクトチェック機関とのパートナーシップにより、情報伝送 PF サービスにおいてファクトチェック結果を実効的に活用）
- ⑧ 政府に対する意見陳述・政策提言

2. 協議会の設置

上記 1. のようなマルチステークホルダーによる協議・決定については、その実効性を担保するため、マルチステークホルダーにより構成される協議会の設置に関する制度整備も含め、具体化を進めることが適当である。

3. 協議会の役割・権限等

マルチステークホルダーにより構成される協議会については、協議会がその構成員（連携・協力の主体）に対して情報提供、意見表明等の必要な協力を求めることができ、構成員は協議会の協議結果に基づき必要な取組を行うものとするなど、協議会に一定の役割・権限等を持たせる方向で具体化を進めることが適当である。

また、協議・決定の過程で交換される情報やデータに含まれる個人情報や機密情報の適正な取扱いを担保するため、協議会が構成員に対し情報提供を求めることができる旨を制度上担保したり、協議会の構成員等に制度上の守秘義務を課したりすることも含め、具体化を進めることが適当である。

なお、連携・協力の場に具体的にどのような主体が参加すべきかについては、上記 1. に挙げた目的ごとに個別的に判断することが適当である。

その際、我が国においては、主として情報発信に関わる個人や情報受信に関わる利用者・消費者その他の市民社会の利益代表となる主体の数が必ずしも多くないことを踏まえ、市民社会の利益をマルチステークホルダーによる協議・決定に反映させるための方策も重要である。また、連携・協力の目的によっては、制度的な枠組みを整備せず、関係者の自主的な取組としてアドホックな連携・協力を積み重ねることも考えられるところ、その場合の連携・協力関係の安定性・継続性を担保することも重要である。これらの具体的な論点について、今後、更なる検討が必要である。

⁵⁸ より具体的には、広告関係事業者・団体との情報共有により発信源に対する広告の出稿停止措置に活用したり、bot アカウントの検出に活用したりすることが考えられる。この点、サイバーセキュリティ対策に関する IoT 機器調査及び利用者への注意喚起の取組 (NOTICE)、インターネット・ホットラインセンター (IHC) や一般社団法人コンテンツ海外流通促進機構 (CODA) による広告関係団体への情報提供の取組等が参考になる。

IV 広告の質の確保を通じた情報流通の健全性確保の在り方

情報伝送 PF サービス上で、本人や組織の許可を得ずに当該本人等であるかのように加工・編集されたなりすまし型の「偽広告」や、その他偽・誤情報を含む広告をはじめ、違法・不当なデジタル広告が流通・拡散しており、閲覧者に財産上の被害をもたらしたり⁵⁹、なりすまされた者の社会的評価を下げたり⁶⁰といった社会的影響が発生・増幅している。

こうした影響の軽減に向けては、第5章で述べた役割・責務が期待される情報伝送 PF 事業者のみならず、広告仲介 PF 事業者も、広告主（及び広告代理店）と直接取引を行い、広告主（及び広告代理店）から直接デジタル広告の出稿を受ける限り、

- ・ アルゴリズムを用いた入札方式の運用型広告を広告主に提供し、デジタル広告の質の確保に向けた取組やその透明性・アカウントビリティの確保等を通じ、アテンション・エコノミーの下で情報流通の健全性確保に貢献することが期待されること
- ・ どのようなデジタル広告がどのようなメディアにどのような形で流通・掲載されるかについて、広告流通の入口の段階で一定のコントロールを及ぼす立場にいること
- ・ 流通させているデジタル広告の内容について合理的な注意義務を負うこと⁶¹

等の理由から、対応を検討することが適当である。

現に、主要な情報伝送 PF 事業者及び広告仲介 PF 事業者は、自社サービスの利用規約等の中で、例えば、「身元や提携関係、資格に関する重要な情報について、誤解を招く表現を使ったり、わかりにくくしたり、省略したりすること」（Google 広告）、「人々を広告へと誘導するために、公人・著名人のイメージや誤解を招くような手口を使うこと」（Instagram 広告）、「誤解を招くコンテンツおよび虚偽のコンテンツ」（TikTok 広告）、「欺瞞的マーケティング」（X（旧 Twitter）広告）等に該当するデジタル広告について、配信停止やアカウント停止の対象とする旨を定めているほか、例えば次のような取組を行っている。

- ・ Yahoo! JAPAN ID 及び LINE ID の取得に当たってはそれぞれ SMS 等による認証を実施。これらの ID は他サービスを利用するにあたって必要となること、投稿時においてはサービスごとに表示名の設定が可能であること等から、なりすましに関しては、アカウント開設時ではなく、アカウント開設後に個別サービスにおける禁止事項等への違反等があった場合に必要範囲でアカウント停止等の措置を実施。

⁵⁹ なりすまし型「偽広告」を入口にしたものを含む SNS 型投資詐欺（相手方が、主として SNS を用いて投資を勧め、投資名目で金銭等をだまし取る詐欺）の被害は、2023 年下半期の増加が顕著であり、同年1年間の被害額は合計約 277 億 9000 万円であった（本検討会第 17 回（WG 第 15 回）会合（2024 年4月 15 日）における警察庁発表（資料 17-3-2））。

⁶⁰ なりすまされた者が正規に出稿していたデジタル広告の掲載を情報伝送 PF 事業者に停止されるという影響も指摘されている。「堀江貴文氏「プラットフォーム規制を」なりすまし広告被害に 提訴は「前沢さんに任せる」（2024 年4月 11 日）産経ニュース（<https://www.sankei.com/article/20240411-JTWGKAKB2NF27FFTYBFU6ZP4NY/>）参照。

⁶¹ 広告媒体業務にも携わる新聞社及び同社に広告の仲介・取次ぎをする広告社について、新聞広告の内容の真実性に疑念を抱くべき特別の事情があって読者らに不測の損害を及ぼすおそれがあることを予見し、又は予見し得た場合に、真実性の調査確認をして虚偽広告を読者らに提供してはならない義務がある旨判示した裁判例として、最三小判平成元年9月 19 日集民第 157 号 601 頁。

- オフィシャルブロガーには与信審査、トップブロガーにおいても所定の確認を実施。
- 以下のような場合には、サービスの使用を許可せず、アカウントや他のエンティティ（ページ、グループ、イベントなど）を制限または無効。
 - スクリプトやその他の不正な手段でアカウントを作成または使用する場合、
 - 以前の削除を回避するためにアカウントやエンティティを作成または再利用する場合、
 - 他の人を誤解させたり欺いたりするために意図的に身元を偽るアカウントを作成または使用する場合
- 誤解を招く身元の偽装を評価する際には、次のような要素を考慮。
 - 名前や年齢などの身元情報の繰り返しまたは重大な変更
 - 誤解を招くような自己紹介内容や位置情報などプロフィール情報
 - スtock画像や盗まれた写真の使用
 - その他の関連するアカウント活動
- LINE 広告・Yahoo!広告について、アカウント開設にあたり、アカウント審査基準に基づく審査（「アカウント開設に関する基準」として、「アカウント申込時の登録情報から不正な広告出稿の懸念があると判断した場合は、アカウントを開設できません。」「開設後のアカウントに関する基準」として、「アカウントの登録情報から不正な広告出稿の懸念があると判断した場合」は「アカウントを停止します。」）と、アカウントを申込みいただいた広告主が本人かを確認するため、一部のお客様に対しお申込み時に「本人確認」を実施。
- 純広告については、掲載ガイドラインを含む弊社規約に同意頂いたことになる旨の申込確認書を締結しているため、全てのクリエイティブは広告主の責任の下で許諾されていることが前提。配信される全ての広告に対して、掲載開始前に弊社側で目視による審査を実施。
- 偽広告防止のための基準（掲載ガイドライン）としては、一般的に偽広告が多く見られる、健康食品やサプリメント、投資サービスといったカテゴリの広告に関しては別途レギュレーションを規定。
- 提携している DSP 事業者からの広告が配信。基本的には広告クリエイティブが入稿された際に審査が行われていると理解。基準については各事業者のポリシーに準拠されていますが、基本的には JIAA の定めるガイドラインに従っていると理解。
- Google の不実表示に関するポリシーでは、広告主による「広告主のビジネス、商品、サービスに関する情報について隠蔽または虚偽記載を行ってユーザーを欺く」広告や「広告のリンク先で『フィッシング』の手口を使ってユーザー情報を収集する」広告の運用は許可されない。
- 今年に入ってから、著名な人物、ブランド、組織の支持を受けている、またはそれらと提携関係にあると虚偽の宣伝をしている悪質な広告主のアカウントを素早く永久停止できるようにするポリシーの更新を発表。
- 新たに広告配信の制限ポリシーを施行。新たな広告主には「会社確認」を行う期間を設定。この期間中は、広告インプレッションが制限される可能性。（当面このポリシーは、広告キャンペーンでブランド名を対象にしている広告主にのみ適用。つまり、広告主とその広告で言及されているブランドの関係性が不明確な場合や、その広告が一般的な内容である場合に、このポリシーが適用。）
- 広告が Facebook または Instagram で表示される前に、広告基準に基づいてレビュー。これは、広告が表示される前に自動的に実施。また、個別のレビュアーが自動化技術を改善およびトレーニングするために働いており、場合によってはマニュアルで広告をレビュー。以前に承認された広告が再レビューの対象となることがあり、その理由は、利用者がその広告を非表示にしたり、ブロックしたり、報告したり、その他の否定的なフィードバックを行った場合等で、これらのフィードバックは、レビュー過程で見逃したものを示す手助け。既に表示されている広告が弊社のポリシーに違反している場合、当該広告は取り下げられ、広告の配信は停止。

- ・ 現在アクティブなすべての広告は広告ライブラリ⁶²で閲覧可能。社会問題、選挙、政治に関する広告は、アクティブでなくなっても 7 年間広告ライブラリで閲覧可能。利用者は、ポリシーに違反していると思われる広告を広告ライブラリから直接調査し報告可能。
- ・ 独立したサードパーティのファクトチェッカーによってファクトチェックされ、評価された広告を拒否。
- ・ 広告主が X 広告を使ってコンテンツを宣伝する場合、そのアカウントとコンテンツは承認プロセスの対象。このプロセスにより、X は広告主が広告ポリシーを遵守しているかを確認。

しかしながら、本検討会におけるプラットフォーム事業者ヒアリングの結果を踏まえると、アテンション・エコミーの下での情報流通の健全性確保に向けた情報伝送 PF 事業者及び広告仲介 PF 事業者による取組として、我が国国内におけるなりすましやデジタル広告の質の確保への対応状況を含む取組状況に関する透明性・アカウントビリティの確保は総じて不十分であり、取組状況そのものについても全体として十分とは言えない。

このようなデジタル空間における情報流通の現状や、情報伝送 PF 事業者及び広告仲介 PF 事業者に期待される役割・責務及び実際の取組状況等を踏まえると、制度整備も含め、広告主と直接取引を行い、広告主から直接広告出稿を受ける情報伝送 PF 事業者及び広告仲介 PF 事業者（以下「情報伝送 PF 事業者等」と、総称する。）に対して以下の具体的措置を求めることが適当である。

1. 対応を検討すべきインターネット上に流通する「違法・不当な広告」の範囲

情報伝送 PF 事業者等において対応を検討すべき「違法・不当な広告」の範囲については、 広告主の表現の自由をはじめとする様々な権利利益に配慮する観点から、前述の主要な情報伝送 PF 事業者等における現状の利用規約等の内容を踏まえつつ、対象範囲の客観的な明確性を確保するとともに、必要かつ相当な対策が適正に講じられることを担保できるよう定められる必要がある。

具体的には、少なくとも、次の i. 及び ii. の各要素の有無・軽重に照らし、具体的な方策との関係で比例性が認められるものは、対応を検討すべき「違法・不当な広告」の範囲に含まれる（すなわち、どのような「違法・不当な広告」に対して、具体的にどのような対応を実施すべきかは、次の i. 及び ii. の各要素の有無・軽重により異なり得る）ものと考えることが適当である。

- i. 当該広告そのものが有する権利侵害性その他の違法性や客観的な有害性（及びその明白性）
- ii. 当該広告が流通・拡散することによる社会的影響の重大性（及びその明白性）

なお、上記範囲の「違法・不当な広告」に対する具体的な方策を検討するに当たり、特に営利広告の流通については、

⁶² <https://www.facebook.com/ads/library/>

- ・ その他の表現に比して委縮効果を考慮する必要が小さいこと
- ・ 広告に由来する消費者被害も少なくなく実害が生じること

等の理由から、より広範な制限に服し得るとの考え方があることに留意が必要である⁶³。

2. デジタル広告の流通前の事前審査の在り方

情報伝送 PF 事業者等による広告の事前審査については、上記のような「違法・不当な広告」のインターネット上における流通・拡散を事前に抑止する観点から、まずは事前審査の確実な実施が担保されることが前提となるが、その上で、その実効性を向上させるための方策として、審査の実態も踏まえつつ⁶⁴、次のようなものを中心に具体化を進めることが適当である⁶⁵。

- ① 広告の事前審査基準の策定・公表等
- ② 広告審査体制の整備及び透明化
- ③ 広告主の本人確認の実施（及び確認した広告主に関する情報等の広告への付与）

その際、このうち①の方策（広告の事前審査基準の策定・公表等）の具体化に当たっては、策定・公表すべき基準の具体的内容（例えば、他人の肖像を使用した広告に関する肖像使用の許諾の有無の確認手続、広告関連法規に違反するなど特定の内容の広告の禁止等）について、広告主の表現の自由をはじめとする様々な権利利益に配慮しつつ、今後、更なる検討が必要である。

また、②の方策（広告審査体制の整備及び透明化）の具体化に当たっては、情報伝送 PF 事業者等による広告の事前審査プロセスの多くの部分に AI を含む自動的手段が用いられているという実態を踏まえ、日本語や日本の社会・文化・法令を理解する者を一定数配置することを確保した上でその配置状況を公表するほか、AI 等の自動的手段を利用する場合における当該手段の実効性向上に向けた措置を実施することや、その実効性に関する情報（例えばエラー率や、広告主からの不服申立てを受けて判断に変更を加えた件数・割合等）を公表することも含めて検討することが適当である。

さらに、③の方策（広告主の本人確認の実施）の具体化に当たっては、確認すべき具体的な事項及び確認方法（例えば、広告代理店を経由した広告出稿の場合において身元

⁶³ 他方で、政治広告の流通に対する制限については、表現の自由との関係で慎重な議論が必要である。

⁶⁴ 本検討会におけるプラットフォーム事業者ヒアリングでは、半数の国内事業者及び一部の国外事業者から、配信される広告に対する日本国内における事前審査について、アカウント審査基準による審査、一部に対するアカウント開設時の本人確認、掲載開始前の目視による審査、広告ポリシーへの準拠を確認するための広告主のアカウントとコンテンツに対する承認プロセス等について回答があり、うち一部の国内事業者からは、本人確認で否認されたアカウント数や、問題があった件数全体に対する本人確認で否認された件数の割合について回答があった。

⁶⁵ なお、本検討会及び WG における議論では、オブザーバー団体（一般社団法人 MyDataJapan、一般社団法人民間放送連盟）や外部有識者（藤代裕之 法政大学社会学部教授）から、デジタル広告に関するガバナンス確保に向けた情報伝送 PF 事業者等による対応が不十分であることの指摘や、そのことを踏まえ、事前審査も視野に広告審査（利用者データの収集・利活用に関するコントロールビリティの確保状況の審査を含む。）を強化することや審査基準の整備・公表を求める意見があった（WG 第 14 回会合（2024 年 4 月 12 日）配付資料 WG14-3、本検討会第 19 回（WG 第 19 回）会合（2024 年 5 月 15 日）配付資料 19-2-2、本検討会第 21 回（WG 第 24 回）会合（2024 年 5 月 30 日）配付資料 21-1-3）。

を確認すべき「本人」の範囲等)について、情報伝送 PF 事業者等による対応の実態、他領域における既存の制度等を踏まえ、かつ、本人確認の目的(違法・不当な広告の流通・拡散の事前抑止、広告主の事後的なトレーサビリティの確保等)に照らし合理的であるかという観点も含め、今後、更なる検討が必要である。

なお、広告の事前審査の実効性を向上させるための方策として、上記①から③までの方策以外に具体化すべき方策があるか否かについても、今後、更なる検討が必要である。

また、情報伝送 PF 事業者等による広告審査や広告掲載停止措置(後記3.)等の取組の具体化に当たっては、情報伝送 PF サービスや広告仲介 PF サービス上での広告の適正な伝送を通じた情報流通の健全性確保(受信者の認知領域等の BtoC 関連の利益や、質の高いメディアの持続可能性、民主主義等の社会的な利益の保護)という目的とは別の観点(BtoB 取引の透明性・公正性向上)から取引の透明性確保を図っている DPF 取引透明化法の運用状況にも留意が必要である。

3. 違法・不当な広告に対する事後的な掲載停止措置の実効性確保に向けた方策

(1) 総論

インターネット上で(事前審査をすり抜けて)流通する違法・不当な広告については、情報伝送 PF 事業者等が当該広告の掲載(や仲介)の可及的速やかな停止措置(アカウントの停止を含め、以下「広告掲載停止措置」という。)を実施することを促進等するための方策として、次のようなものを中心に具体化を進めることが適当である。

- | |
|--|
| <ul style="list-style-type: none">① 広告掲載停止措置に関する透明性の確保を通じた過不足ない実施の確保② 広告掲載停止措置に関する対応の迅速化を通じた実施の促進③ 広告掲載停止措置について、体制を整備して確実に実施④ 上記①から③までの組合せによる対応 |
|--|

これらの方策については、以下(2)から(5)までにおいて詳述するように、対象とする違法・不当な広告の特性・性質(権利侵害性その他の違法性・有害性、流通することによる社会的影響の重大性)等に応じた対応とすることが適当である。

(2) 広告掲載停止措置に関する透明性の確保を通じた過不足ない実施の確保について

①の方策(広告掲載停止措置に関する透明性の確保を通じた過不足ない実施の確保)については、次の(i)から(iv)までのような方策を中心に具体化を進めることが適当である。

- | |
|---|
| <ul style="list-style-type: none">(i) 広告掲載停止措置に関する基準や手続を事前に策定・公表(ii) 広告掲載停止措置の実施要否等の判断に関与する人員等の体制に関する情報を |
|---|

公表

- (iii) 上記 (i) の基準の運用状況を事後に公表
- (iv) 広告掲載停止措置を実施した場合に、その旨及び理由並びに不服申立ての方法を広告主に通知

その際、これらの方策の具体化に当たっては、公表・通知等すべき具体的事項の詳細（例えば、(i) に関し、広告掲載停止措置そのものに関する基準・手続のみならず、掲載する広告のターゲティングに用いる主なパラメータに関する情報を公表するなど）等について、今後、更なる検討が必要である。

この点については、**外資系も含む情報伝送 PF 事業者等による広告掲載停止措置が我が国国内において過不足なく適正に実施されていることに関し、広告主や広告閲覧者を含む社会一般が確認し、情報伝送 PF 事業者等が提供するサービスに対する信頼性を向上させるという目的に照らして具体化することが適当**である。

また、これら (i) から (iv) までの方策以外に、広告掲載停止措置に関する透明性の確保に向け、いかなる方策（例えば、過去に掲載（仲介）した広告に関し、ターゲティングの対象とされた集団の範囲等を含む情報をデータベース化し、公開するなど）が必要かつ適当かについては、その具体的な目的を整理しつつ、今後、更なる検討が必要である。

(3) 広告掲載停止措置に関する対応の迅速化を通じた実施の促進について

②の方策(広告掲載停止措置に関する対応の迅速化を通じた実施の促進)については、次の (i) から (iv) までのような対策を中心に具体化を進めることが適当である⁶⁶。

- (i) 外部からの広告掲載停止申出・要請窓口を整備・公表
- (ii) 上記 (i) の窓口を通じて申出・要請があった場合に、一定期間内に広告掲載停止措置の実施の要否・内容を判断し、申請者に判断結果（及び不服申立ての方法）を通知
- (iii) 広告掲載停止措置の実施の要否・内容を判断するための体制（広告掲載停止措置に関する不服申立てを受け付ける体制を含む。）を整備
- (iv) 一定の条件（例えば、行政機関等の特定の第三者からの申出・要請を受けて実施した場合等）の下で行った広告掲載停止措置により広告主が被った損害について、情報伝送 PF 事業者等を免責

その際、これらの方策の具体化に当たっては、濫用的な申出・要請から生じる情報伝送 PF 事業者等の実務上の負担に配慮する観点から、対象とする広告の特性・性質に応じ、いかなる主体からの申出・要請を契機とした広告掲載停止措置の実施を促進すべきかについて、今後、更なる検討が必要である。

⁶⁶ 行政機関等からの申出・要請に基づき実施した広告掲載停止措置については、広告主等への説明責任は、情報伝送 PF 事業者等ではなく、当該申出・要請を発出した行政機関等が負うという意見もあった。

例えば、他人の権利を侵害する違法な広告については、自己の権利を侵害されたとする者（被侵害者）からの申出・要請を契機とした広告掲載停止措置について上記（i）から（iv）までのような対応を中心に具体化を進めることが、濫用的な申出・要請のおそれも小さく適当である⁶⁷。

また、行政法規に抵触する違法な広告については、対応の迅速化を通じた実施の促進を図ることとする場合、違法性の判断能力の観点から、当該行政法規を所管する行政機関（当該行政機関の委託や認証を受けた機関を含む。）からの申出・要請を契機とした広告掲載停止措置について、上記（i）から（iv）までのような対応を中心に具体化を進めることが基本的には適当である。なお、この点については、今後、具体的な行政法規を洗い出しつつ、具体化を進めることが適当である。

ただし、この場合の対応については、前提として、行政機関による恣意的な申出・要請を防止し、透明性・アカウントビリティを確保するとともに、過度な申出・要請に対し広告主や情報伝送 PF 事業者等を救済するために、次のような方策を併せて検討することが不可欠である。

- | |
|--|
| <ul style="list-style-type: none">(ア) 行政機関において、申出・申請に関する手続等（事後救済手段を含む。）を事前に策定・公表(イ) 行政機関において、実際に行った申出・申請の状況を事後的に公表(ウ) 申出・要請に応じて実施された広告掲載停止措置により広告主が被った損害について、情報伝送 PF 事業者等を免責(エ) 広告掲載停止措置等を実施した情報伝送 PF 事業者等において、行政機関の名称等の情報を発信者に通知 |
|--|

その際、このような行政機関による申出・申請に関する透明性・アカウントビリティを確保するための方策の具体化に当たっては、行政手続に関する既存の法的枠組み（例えば、行政指導指針の策定・公表義務、国家賠償法等）との関係を含め、今後、更なる検討が必要である。

以上のほか、違法・不当な広告が掲載されたウェブページの管理者や、当該広告に隣接して掲載されたコンテンツの発信者、広告関係団体等からの申出・要請を契機とした広告掲載停止措置について、上記（i）から（iv）までのような対策を含め具体化を進めることが適当であるが、当該広告の特性・性質に応じた適切な申出・要請主体の範囲については、今後、更なる検討が必要である。

特に、ファクトチェック機関等の特定の第三者からの申出・要請を契機として、権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい広告に対して広告掲載停止措置を実施することに関し、上記（i）から（iv）までのような対策を制度的

⁶⁷ なお、前述の特定電気通信役務提供者による他人の権利を侵害する情報の削除（送信を防止する措置）に係る情報流通プラットフォーム対処法（旧プロバイダ責任制限法）、リベンジポルノ防止法、AV 新法等の規定は、広告かそれ以外のコンテンツかの区別なく適用される。

に担保することについては、権利侵害性その他の違法性がない広告に関する広告主の表現の自由への過度の制約を避ける観点から、当該広告の特性・性質（有害性や社会的影響の大小・明白性）を考慮しつつ、引き続き慎重な検討が必要である。

（４）広告掲載停止措置の確実な実施について

③の方策（広告掲載停止措置の確実な実施）については、少なくとも、他人の権利を侵害する広告又は行政法規に抵触する違法な広告の場合、営利広告の特徴（上記1.）を踏まえ、被侵害者又は行政機関からの申出・要請を契機とした広告掲載停止措置について、その確実な実施を担保する方策を含め、具体化を進めることが適当である。

一方、権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい広告については、権利侵害性その他の違法性がない広告に関する広告主の表現の自由への過度の制約を避ける観点から、当該広告の特性・性質（有害性や社会的影響の大小・明白性）を考慮しつつ、引き続き慎重な検討が必要である。

なお、WGにおける議論では、権利侵害性その他の違法性はないが有害性や社会的影響の重大性が大きい広告の一部について、脆弱な個人に対するターゲティングの停止の確実な実施を担保することが適当とする意見もあった。こうした方策の適否については、情報伝送 PF 事業者等が提供するサービスにおけるターゲティングの実態を踏まえつつ、今後、更なる検討が必要である。

（５）情報流通の態様に着目した広告掲載停止措置の実施について

以上のほか、出稿された広告の内容に着目するのではなく、情報流通の態様に着目した広告掲載停止措置の実施の在り方についても、違法・不当な広告の流通・拡散を抑止する観点も含め、具体化を進めることが適当である。

なお、具体的にどのような態様の情報流通を対象とするか等の詳細については、当該態様の情報流通によって発生又は増幅する影響及びリスクを特定しつつ、今後、更なる検討が必要である。

V 質の高いメディアへの広告配信に資する取組を通じた情報流通の健全性確保の在り方

広告仲介 PF 事業者が提供する運用型広告を通じ、偽・誤情報等をはじめ、違法なコンテンツや客観的に有害なコンテンツを掲載するオンラインメディアにデジタル広告が配信され、広告主が支払う広告費が偽・誤情報等をはじめとする違法なコンテンツや客観的に有害なコンテンツの発信・拡散主体の収入源となる一方、手間とコストを掛けて取材・編集・制作された質の高いコンテンツ等を発信するメディアの広告収入に影響が及んでいるとの指摘がある⁶⁸。

こうした状況は、広告主（企業のほか、行政機関や地方自治体を含む。）のブランド価値の毀損（ブランドセーフティ⁶⁹に関する問題）やアドフラウド⁷⁰による被害等につながるものであるが、意識の高い一部の広告主を除き、広告主においていまだ問題に気づいていない（問題とっていない）、又は問題に気づいていても対策できていない（対策の方法が分からない、対策する人がいない、対策にコストがかかけられない）という指摘が存在する⁷¹。

特に、広告主においてデジタル広告の出稿に直接関わる現場担当者は単価ダウンのみを追いかけがちであり、自らが所属する組織の評判悪化やブランドイメージの低下懸念に考えが及ばないことがあり得、また、対策コストの確保も現場担当者のみでは難しいと考えられるため、広告主の経営層レベルでの大所高所からの判断や後押し、組織内の関係部門の連携・協力の促進や対応リソースの確保が必要であり、経営層がコンプライアンス問題・リスクマネジメントとして認識し、主体的に取組を進めることが重要である⁷²。

このような観点から、**広告主やその経営陣、更には広告主から委託を受け、広告主と一体的にデジタル広告の出稿に関与する広告代理店による主体的な取組を促進するための方策を中心に、制度整備を含め、以下の具体的な措置を講じることが適当**である。

1. 質の高いメディアへの広告配信の確保に向けた広告主（及び広告代理店）による取組の促進方策

広告主やその経営陣及び広告代理店による主体的な取組（その前提としてのブランドセーフティやアドフラウド対策等に関する意識の醸成を含む。）を促進するための方策と

⁶⁸ 公正取引委員会「デジタル広告分野の取引実態に関する最終報告書」（2021年2月）。

⁶⁹ 広告掲載先の品質確保による広告主ブランドの安全性（ブランドを毀損する不適切なページやコンテンツに広告が表示されるリスクからの安全性）をいう。

⁷⁰ 自動化プログラム（bot）を利用したり、スパムコンテンツを大量に生成したりすることで、インプレッション（広告表示）やクリックを稼ぎ、不正に広告収入を得る悪質な手法をいう。

⁷¹ 本検討会第12回（WG第7回）会合（2024年3月15日）における公益社団法人日本アドバイザーズ協会（JAA）による発表（資料12-1）を参照。

⁷² 偽・誤情報等の発信・拡散によって他人の権利又は法律上保護される利益が侵害された場合、当該偽・誤情報等の発信・拡散に間接的に寄与した主体も共同不法行為責任（民法第719条）等の責任を負う可能性があり、また、当該主体が会社その他の法人であれば、その役員等は当該法人に対して任務懈怠責任（会社法第423条等）等の責任を負う可能性がある。なお、漫画の海賊版を掲載するウェブサイトへに広告を提供して広告料を支払っていた広告代理店について、漫画家への損害賠償責任を認めた裁判例として、知財高判令和4年6月29日裁判所ウェブサイト。

して、マルチステークホルダーによる連携・協力の下、広告主やその経営陣及び広告代理店に求められる取組に関するガイドライン、ガイドブック等を策定・公表することが適当である⁷³。

既に広告関係団体が策定・公表しているガイドライン等⁷⁴の内容を踏まえると、ガイドライン、ガイドブック等に盛り込むべき取組について、次のようなものを中心に具体化を進めることが適当である。

- | |
|--|
| <ol style="list-style-type: none">① ブロックリスト（掲載したくない配信先をリストアップ）② セーフリスト（掲載したい配信先をリストアップ）③ 媒体社（パブリッシャー）と広告主を限定したクローズな広告の取引市場である PMP（Private Marketplace）の活用④ アドベリフィケーションツールの導入⑤ 業界ガイドラインの遵守やそれについての第三者機関⁷⁵による認証の取得等の客観的な指標により、広告掲載品質の確保を適切に行っていることが担保された広告仲介 PF 事業者その他の広告関係事業者の利用⑥ 広告掲載品質の確保のための質の高いメディア（パブリッシャー）の真正性・信頼性確保技術の活用 |
|--|

その際、こうしたガイドライン、ガイドブック等を策定・公表するに当たっては、どのような内容とし、かつ、どのような手段で広告主企業やその経営陣に周知することが効果的かといった具体について、民間企業のみならず、行政機関や地方自治体におけるデジタル広告の出稿に係るガバナンス体制の効果的な方策としての有効性という観点も含め、今後、更なる検討が必要である。

なお、行政機関や地方自治体が自らの施策の周知等の目的でデジタル広告を出稿し、広告主となる場合、その広告費が公金から支出されること等も踏まえ、求められる取組について積極的に情報収集するとともに、デジタル広告の出稿に係るガバナンス体制を構築することが適当である。

2. 質の高いメディアへの広告配信の確保に向けた広告仲介 PF 事業者による取組の促進方策

以上のような広告主（及び広告代理店）を中心とした取組に加え、その実効性を高める観点からも、運用型広告の提供を通じてメディア運営者（パブリッシャー）との接点

⁷³ 同様に、企業の経営陣を対象に、デジタル空間における情報流通に関わる取組を促進する目的で策定されたガイドライン、ガイドブック等の例として、経済産業省＝独立行政法人情報処理推進機構（IPA）「サイバーセキュリティ経営ガイドライン Ver 3.0」（2023年3月）や、経済産業省＝総務省「DX 時代における企業のプライバシーガバナンスガイドブック ver1.3」（2023年4月）が挙げられる。

⁷⁴ 例えば、一般社団法人日本インタラクティブ広告協会（JIAA）「広告掲載先の品質確保に関するガイドライン（ブランドセーフティガイドライン）」（2019年4月）、公益社団法人日本アドバイザーズ協会（JAA）「デジタル広告の課題に対するアドバイザー宣言」（2019年11月）及び「社会問題化するデジタルメディア上の詐欺広告に対する緊急提言」（2024年5月）、一般社団法人日本広告業協会（JAAA）「インターネット広告における運用型広告取引ガイドライン」（2014年9月一部改訂）、「デジタルメディア上における詐欺広告について」（2024年5月）など。

⁷⁵ 例えば、一般社団法人デジタル広告品質認証機構（JICDAQ）など。

を持つことになる広告仲介 PF は、偽・誤情報等、違法なコンテンツや客観的に有害なコンテンツの発信・拡散主体への広告費流入抑止と質の高いメディアへの広告配信の確保に向けて対応することが適当である。

この点、パブリッシャーと直接取引を行う広告仲介 PF 事業者の中には、既に、偽・誤情報や違法アップロードコンテンツ等を掲載するメディアへの広告費の流入を抑止する観点から、例えば広告配信先メディアの事前審査（以下「メディア審査」という。）や、事後的に広告配信を停止する措置（以下「広告配信停止措置」という。）等を自主的に実施しているものが存在する⁷⁶。このような、広告仲介 PF 事業者による質の高いメディアへの広告配信の確保に向けた取組を促進するための方策について、メディア審査や広告配信停止措置の実態を踏まえつつ、次のようなものを中心に具体化を進めることが適当である。

- | |
|--|
| <ol style="list-style-type: none">① メディア審査や広告配信停止措置に関する基準の策定・公表と運用状況の公開② 偽・誤情報や違法アップロードコンテンツ等を掲載するメディアへの広告費流入抑止のためのメディア審査や広告配信停止措置の実施体制の整備及び透明化③ メディア審査時におけるパブリッシャーの本人確認の実施④ メディア審査を通じた偽・誤情報や違法アップロードコンテンツ等を掲載するメディアへの広告配信の拒否や広告配信停止措置を実施した場合に、その旨、理由及び不服申立て方法をパブリッシャーに通知⑤ 外部からの広告配信停止申出・要請窓口を整備・公表⑥ 上記⑤の窓口を通じて申出・要請があった場合に、一定期間内に広告配信停止措置の実施の要否を判断し、判断結果と不服申立て方法を申請者に通知⑦ 一定の条件の下で行った広告配信停止措置によりパブリッシャーが被った損害について、広告仲介 PF 事業者を免責 |
|--|

その際、特に①の方策（基準の策定・公表等）の具体化に当たっては、偽・誤情報等を掲載するウェブページや、広告費詐取を目的として作成されたウェブページ (Made For Advertisement : MFA) の取扱いを明記することも含めて検討することが適当である。

なお、広告仲介 PF 事業者によるメディア審査や広告配信停止措置等の取組の具体化に当たっては、広告仲介 PF サービス上での広告の適正な伝送を通じた情報流通の健全性確保（受信者の認知領域等の BtoC 関連の利益や、質の高いメディアの持続可能性、民主主義等の社会的な利益の保護）という目的とは別の観点（BtoB 取引の透明性・公正性向上）から取引の透明性確保を図っている DPF 取引透明化法の運用状況にも留意が必要である。

⁷⁶ 例えば、「誤解を招く表現」、「信頼性がなく有害な文言」、「欺瞞行為」、「操作されたメディア」、「自社コンテンツが乏しいサイト」等を含むメディアへの広告配信を停止したり、アカウントを停止したりする措置を実施しているものが存在する（Google AdSense、Yahoo!広告など）。

VI その他全体に共通する事項

1. 執行手段・プロセス

I から V までで検討した方策の実施を情報伝送 PF 事業者や広告仲介 PF 事業者に求める場合、その実効性を制度的に担保する方法については、執行手段・プロセス（例えば、報告徴収・勧告・命令、罰金、課徴金、民事的救済等）の選択及び具体的な設計を含め、比例性や外資系を含む事業者への抑止効果、自主的改善に向けたインセンティブ効果等を勘案しつつ検討を進めることが適当である。

2. 対象事業者の範囲

(1) 情報伝送 PF 事業者

I から V までで検討した方策の実施の対象となる情報伝送 PF 事業者については、利用者が投稿するコンテンツに含まれる偽・誤情報等の流通・拡散に対応するための方策（I から III まで）か、違法・不当な広告の流通・拡散に対応するための方策（IV）かにかかわらず、偽・誤情報等や違法・不当な広告の流通・拡散の拡がり・頻度や社会に与える影響の深刻度という観点から、利用者数や、サービスの目的・性質等を勘案し、一定の要件を満たす大規模なものにまずは方策の実施を求めることが適当である。

(2) 広告仲介 PF 事業者

IV 及び V で検討した方策の実施の対象となる広告仲介 PF 事業者については、違法・不当な広告の流通・拡散の頻度や社会に与える影響の深刻度（IV）、あるいは広告費の支払を通じて偽・誤情報等の流通・拡散に間接的に寄与するおそれ（V）の大きさ等の観点から、一定の要件を満たす大規模なものにまずは方策の実施を求めることが適当である。

この場合における「一定の要件」を画する指標として、例えば、IV で検討した方策であれば、広告主からの広告入稿数、最終的に掲載された広告が受信者によって閲覧された回数（インプレッション数）等、V で検討した方策であれば、広告配信先となり得る取引先メディアの数・取引高、当該メディア上の広告付きコンテンツが受信者によって閲覧された回数等が考えられるが、いかなるものが適当かについては、広告仲介 PF サービスの実態を踏まえつつ、今後、更なる検討が必要である。

3. 生成 AI を用いて生成される情報への対応

生成 AI を用いて生成される画像、動画、音声等の偽・誤情報の流通・拡散への対応⁷⁷

⁷⁷ なお、2024 年 4 月に総務省及び経済産業省がとりまとめた「AI 事業者ガイドライン」においては、偽情報等の対策に関して AI 開発者、AI 提供者及び AI 利用者の各主体が取り組む事項として、生成 AI によって内容が真実・公平であるかのように装った情報を誰でも作ることができるようになり、AI が生成した偽情報・誤情報・偏向情報が社会を不安定化・混乱させるリスクが高まっていることを認識した上で、必要な対策を講じることが挙げられている。

については、上記1.（情報伝送 PF による偽・誤情報への対応の在り方）や2.（情報伝送 PF が与える情報流通の健全性への影響の軽減に向けた方策の在り方）を踏まえたコンテンツモデレーションの実施等に加え、例えば、生成段階において AI 生成物であることをラベリングすること、情報伝送 PF サービスにおいて AI 生成物を検知したり、送信する際に AI 生成物をラベリングした上で送信したりすることが有効との指摘がある⁷⁸。生成 AI を用いて生成される偽・誤情報への対応に関する制度面からのアプローチについては、今後、技術の進展やサービスの普及の状況、国際的な議論動向等を踏まえつつ、必要な対応の検討を進めることが適当である。

⁷⁸ 本検討会第3回会合（2023年11月27日）における越前構成員による発表（資料3-2）及びWG第12回会合（2024年4月5日）における生員構成員による発表（資料WG12-2）を参照。