

# 「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論点整理（案）

## 1. 情報伝送 PF による偽・誤情報への対応の在り方

- 情報伝送 PF 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）
  - ・ 個人の自律的な意思決定を含む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。
- 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 の特徴（①情報発信コストの低廉性、②拡散促進機能の具備、③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、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（いわゆる「アテンション・エコノミー」）の存在が指摘。
- 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と異なり、情報伝送 PF を通じた情報流通に関しては、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。
- これらを踏まえ、情報伝送 PF 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。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。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、国民生活や社会経済活動等に広くかつ深く浸透しており、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 は、自社サービスの利用規約等の中で、一定の範囲の偽・誤情報について、例えば、「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」(Yahoo!ニュースコメント欄)、「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」(YouTube)、「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」(Facebook、Instagram) 等に該当する情報について、情報の削除、表示順位の低下等の一定の対応をする旨を規定。
- 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 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も鑑みると、情報伝送 PF による自主的な取組も期待できない状況であり、新たな具体的な対応が必要。
- 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 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など、いわば「誰にも開かれた情報流通の場」としてのインターネットそのもの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 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。
- 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 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 に対して以下の具体的措置を求めることが適当ではないか。

## (1) 対応を検討すべき「偽・誤情報」の定義・範囲

- 情報伝送 PF において対応を検討すべき「偽・誤情報」の定義・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 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ところ、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、対応を検討すべき「偽・誤情報」の定義・範囲に含まれるものと考えることが適当ではないか。

- |   |
|---|
| <p>① 検証可能な誤りが含まれていること</p> <p>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</p> <p>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）</p> <p>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）</p> <p>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<br/>重大な社会的混乱を招くような情報</p> <p>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）</p> |
|---|

- ①の要件について、
- ✓ 「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、これを満たし得ると考えることが適当ではないか。
  - ✓ 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ものの、情報伝送 PF 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」の定義・範囲の要件とはしないことが適当ではないか。
  - ✓ 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について、対応を検討すべき「偽・誤情報」の定義・範囲に含まれるものと捉えるか否かについては、具体的なケースを想定しつつ、今後、更なる検討が必要ではないか。
- ②の要件について、
- ✓ 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、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のではないか。
  - ✓ 「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<sup>1</sup>

<sup>1</sup> 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

・ 重大な社会的混乱を招くような情報<sup>2</sup>

については、情報伝送PF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（後記2．参照）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のではないか。

- ✓ 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（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はないか。
- ✓ パロディ・風刺や伝統メディアによる誤報など、「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はないか。
- ✓ これらの他、対応を検討すべき「偽・誤情報」の範囲に典型的に含まれない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はないか。

## （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実施。

### ① 偽・誤情報の流通・拡散を直接的に抑止するための措置

#### (i) 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

#### (ii) 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

#### (iii) 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

#### (iv) 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

救助要請、投資詐欺を目的としたなりすまし型のいわゆる「偽広告」など。

<sup>2</sup> 例えば、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動画、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、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・誤情報など。

(v) 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

(vi) 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

(vii) 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、新規投稿等を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

(viii) 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

➤ 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はないか。

例) 収益化の停止 (ii) 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。

➤ 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はないか。

✓ 例えば、収益化の停止 (ii) 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。

✓ 情報の可視性に影響しないラベルの付与 (iii) 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。

✓ 情報の可視性に一部影響するラベルの付与 (iv) 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 (v) 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、人為的な操作を加えない原則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 (vi) 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 (vii) 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（i・ii・iii）や、情報の可視性に一部影響を与えるもの（iv・v）、情報の全部又は一部を不可視とするもの（vi・vii）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も言えることも踏まえ、情報伝送 PF においては、対象とする偽・誤情報の定義・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はないか。
- 一方、いわゆるプロミネンス（viii）については、いかなる情報を「信頼できる」ものと考え、また具体的にいかなる場合に「受信可能性の向上」がなさ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（i から vii まで）とは異なる取扱いが必要との指摘があるところ、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はないか。
- 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」の定義・範囲」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 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はないか。

### （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 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のではないか。

- |  |
|--|
| <ul style="list-style-type: none"><li>(i) 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</li><li>(ii) 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</li><li>(iii) 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</li><li>(iv) 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</li><li>(v) 上記（i）から（iv）までの組合せによる対応</li></ul> |
|--|

- これらの方策については、以下②から⑥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はないか。

## 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（i）の方策は、情報伝送 PF 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 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ところ、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」という。）における透明化規律を参考としつつ、次の（ア）から（エ）までの対応を中心に具体化を進めることが適当ではないか。

(ア) 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 (イ) 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 (ウ) 上記（ア）の基準の運用状況を事後に公表 (エ) 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
---

- 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はないか。この点については、外資系も含む情報伝送 PF 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 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はないか。
- 特に（イ）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 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はないか。その際には、外資系も含む情報伝送 PF 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はないか。

## 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（ii）の方策は、情報伝送 PF 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ところ、この方策については、情報流通プラットフォーム対処法における迅速化規律等を参考としつつ、次の（ア）から（エ）までの対応を中心に具体化を進めることが適当ではないか。

（ア）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  
（イ）上記（ア）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  
（ウ）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  
（エ）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 を免責

- 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 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はないか。

#### i. 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、）自己の権利を侵害されたとする者（被侵害者）からの申出・要請を契機としたコンテンツモデレーションについて上記（ア）から（エ）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はないか。

#### ii. 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（ア）から（エ）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はないか。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はないか。

- 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 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はないか。

- (a) 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
- (b) 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
- (c) 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 を免責
- (d) コンテンツモデレーションを実施した情報伝送 PF において、行政機関の名称等の情報を発信者に通知

- 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）との関係について、今後、更なる検討が必要ではないか。

### iii. 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（例えば、人の生命、身体又は財産に重大かつ明白な悪影響を及ぼす偽・誤情報、なりすまし型のいわゆる「偽広告」など）についても、発信者や情報伝送 PF 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、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、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について、以下の方向性を基本としつつ、上記（ア）から（エ）までのような対策を含め、情報伝送 PF によるコンテンツモデレーションの実施の在り方について具体化を進めることが適当ではないか。

#### (a) 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

- 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 i 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 ii 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、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（ア）から（エ）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はないか。

- 一方、情報伝送 PF 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ところ、こうした自主的な取組を促す観点からは、後記 2. の「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」として情報伝送 PF が自ら実施する影響評価・軽減措置の仕組みを活用することが適当ではないか。

(b) 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

- 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・誤情報に関し、発信者や情報伝送 PF 以外の特定の第三者からの申出・要請を契機としたコンテンツモデレーションのうち、特に情報の可視性に直接の影響がないもの（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は、上記（ア）から（エ）までのような対策を含め、情報伝送 PF による迅速なコンテンツモデレーションの実施を促進するための方策の在り方について具体化を進めることが適当ではないか。
- 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はないか。

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（iii）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 による確実な実施を担保するための方策であるところ、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・誤情報について、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は、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 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はないか。
- なお、本 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ところ、こうした方策の適否については、情報伝送 PF 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はないか。

## 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 (iv) の方策により、情報伝送 PF 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 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はないか<sup>3</sup>。一方、違法性を有する偽・誤情報を繰り返し発信する悪質な発信者への対応については、情報の削除やアカウント停止・削除も含む措置が必要ではないか。

## ⑥ 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ところ、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はないか。
- 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はないか。

### (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、民事法による対応、行政処分等による対応が講じ得ることとされている一方、情報伝送 PF 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 60 条）や幫助犯（同法 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・要請等を受けたりすることがあり得るところ、情報伝送 PF 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 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。

<sup>3</sup> ただし、情報伝送 PF 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。

- 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 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はないか。

- |  |
|--|
| <ul style="list-style-type: none"><li>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</li><li>② bot アカウントの抑止策の導入（アカウントの有料化等）</li><li>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）</li></ul> |
|--|

- 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 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はないか。

## 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

- 前記1. のとおり、情報伝送 PF 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、総称して「偽・誤情報等」という。）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、個人の自律的な意思決定を含む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（前記1. 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 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<sup>4</sup>。
- 以上に加え、前記1. のデジタル空間における情報流通の現状や、情報伝送 PF 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 に対して以下の具体的措置を求めることが適当ではないか。

### (1) 情報伝送 PF による社会的影響の予測・軽減措置の実施

- 情報伝送 PF 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・実施することが適当ではないか。
- 特に、情報伝送 PF 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・拡散等による社会的影響が深刻化するリスクが指摘されているところ、こうしたリスクを含め、情報伝送 PF のアーキテクチャに起因する社

<sup>4</sup> 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はないか。

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 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はないか<sup>5</sup>。

- その際、情報伝送 PF の利用者及び情報伝送 PF 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 による恣意的な影響予測を防ぎ、かつ、情報伝送 PF 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はないか。

- 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 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はないか。一方、上記の枠組みの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（特に、本検討会で検討されている「情報流通の健全性」に関する基本理念をどのように反映させるか）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はないか。

- 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ことが適当ではないか。

- 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 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

<sup>5</sup> 他方で、情報伝送 PF 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 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はないか。

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ことが適当ではないか。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はないか。

- 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はないか。
- 以上に加え、情報伝送 PF 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 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はないか。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はないか。

- |   |
|---|
| <ul style="list-style-type: none"><li>(i) 実施指針を踏まえた行動規範を各情報伝送 PF が策定・公表</li><li>(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</li><li>(iii) 検証・評価の結果を踏まえ、情報伝送 PF において影響軽減策を策定・公表</li><li>(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 において当該影響軽減策を確実に実施</li></ul> |
|---|

## (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中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 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 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 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はないか。

- 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 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はないか。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、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はないか。

- |   |
|---|
| <ul style="list-style-type: none"><li>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</li><li>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</li><li>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</li><li>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）</li><li>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。）</li></ul> |
|---|

- 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は、明確に定められる必要があり、少なくとも始期・終期についてはマルチステークホルダーの協議によって決定することが適当ではないか。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はないか。
- 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はないか。

### 3. 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

- デジタル空間の情報流通に関するガバナンスの確立が急務であるなどの前記1. で述べたデジタル空間における情報流通の現状等を踏まえれば、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はないか。

#### (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、前記2. で述べた情報伝送 PF 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はないか。
- 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はないか。

- 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 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 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 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はないか。

- 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）
- 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 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 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はないか。

## **(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はないか。
- 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はないか。
- 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はないか。
  - ✓ 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はないか。
  - ✓ 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はない

か。

- ✓ 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はないか。

## 4. 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 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、なりすまされた者の社会的評価を下げたりといった社会的影響が発生・増幅。
- こうした影響の軽減に向けては、前記 2. で述べた役割・責務が期待される情報伝送 PF のみならず、広告仲介 PF も、広告主と直接取引を行い、広告主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等の理由から、対応を検討することが適当。
- 現に、主要な情報伝送 PF 及び広告仲介 PF 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、広告の質の確保に向けた各種の取組を実施。
- 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・エコノ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 及び広告仲介 PF 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 及び広告仲介 PF 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 及び広告仲介 PF (以下、総称して「情報伝送 PF 等」という。) に対して以下の具体的措置を求めることが適当ではないか。

### (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 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 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ところ、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はないか。

- |  |
|--|
| 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、
- ・ 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はないか。

## (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 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、次のようなものを中心に具体化を進めることが適当ではないか。

- |  |
|--|
| 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） |

- 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、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はないか。

- 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 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はないか。

- 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を確認すべき「本人」の範囲等）について、情報伝送 PF 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はないか。
- 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はないか。

### (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 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はないか。

- (i) 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
- (ii) 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
- (iii) 広告掲載停止措置について、体制を整備して確実に実施
- (iv) 上記 (i) から (iii) 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はないか。

#### 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

- (i) 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 (ア) から (エ)

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はないか。

- (ア) 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
- (イ) 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
- (ウ) 上記（ア）の基準の運用状況を事後に公表
- (エ) 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

- 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、（ア）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はないか。この点については、外資系も含む情報伝送 PF 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 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はないか。
- また、これら（ア）から（エ）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はないか。

### 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

- （ii）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（ア）から（エ）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はないか。

- (ア) 外部からの広告掲載停止申出・要請窓口を整備・公表
- (イ) 上記（ア）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
- (ウ) 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
- (エ) 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 等を免責

- 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 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はないか。

- 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（ア）から（エ）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はないか。
- 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（ア）から（エ）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はないか。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はないか。
  - ✓ 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 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はないか。

- (a) 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
- (b) 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
- (c) 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 等を免責
- (d) 広告掲載停止措置等を実施した情報伝送 PF 等において、行政機関の名称等の情報を発信者に通知

- ✓ 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はないか。
- 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、ファクトチェック機関等からの申出・要請を契機とした広告掲載停止措置について、上記（ア）から（エ）までのような対策を含め具体化を進めることが適当ではないか。なお、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はないか。
  - ✓ 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（例えば、他人の権利を侵害するとまでは言えないなりすまし型の「偽広告」など）に対して広告掲載停止措置を実施することに関し、上記（ア）から（エ）までのような対策を制度的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はないか。

#### ④ 広告掲載停止措置の確実な実施について

- （iii）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はないか。
- 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はないか。
- なお、本 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ところ、こうした方策の適否については、情報伝送 PF 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はないか。

#### 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はないか。
- 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はないか。

## 5. 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 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、その他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存在。
- 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に関する問題）やアドフraud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、対策にコストがかけれない）という指摘が存在。
- 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。
- さらに、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。
- 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はないか。**

### (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raud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はないか。**

- 既に広告関係団体が策定・公表しているガイドライン等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はないか。

- |  |
|--|
| <ol style="list-style-type: none"><li>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）</li><li>② セーフリスト（掲載したい配信先をリストアップ）</li><li>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</li><li>④ アドベリフィケーションツールの導入</li><li>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 その他の広告関係事業者の利用</li><li>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</li></ol> |
|--|

- 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はないか。

## （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 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はないか。
- この点、パブリッシャーと直接取引を行う広告仲介 PF の中には、既に、偽・誤情報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ところ、このような、広告仲介 PF 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はないか。

- |  |
|--|
| <ol style="list-style-type: none"><li>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</li><li>② 偽・誤情報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</li></ol> |
|--|

- 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 を免責

- 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はないか。
- なお、広告仲介 PF 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 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、特定デジタルプラットフォームの透明性及び公正性の向上に関する法律により透明性確保が図られているところ、同法との連携の在り方を含め、今後、更なる検討が必要ではないか。

## 6. その他全体に共通する事項

### (1) 執行手段・プロセス

- 前記1. から5. までで検討した方策の実施を情報伝送 PF や広告仲介 PF 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、外資系を含む事業者への抑止効果等を勘案しつつ検討を進めることが適当ではないか。

### (2) 対象事業者の範囲

#### ① 情報伝送 PF

- 前記1. から4. までで検討した方策の実施の対象となる情報伝送 PF 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（前記1. から3. 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（前記4.）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はないか。

#### ② 広告仲介 PF

- 前記4. 及び5. で検討した方策の実施の対象となる広告仲介 PF 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（前記4.）、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（前記5.）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はないか。
- 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、前記4. 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、前記5. 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 の実態を踏まえつつ、今後、更なる検討が必要ではないか。

### (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したり、情報伝送 PF において送信する際に AI 生成物をラベリングした上で伝送したりすることが有効との指摘があるところ、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はないか。

以上