

デジタル空間における情報流通の諸課題への 対処に関する検討会 制度 WG 中間取りまとめ（案）

～新たな違法・有害情報対策のための羅針盤（仮称）～

2025.6

諸課題検 制度 WG

目次

I	検討の背景	3
II	検討の視点 ～レイヤーごとの整理～	6
III	事業者ヒアリング	8
IV	対応の方向性	11
1	情報レイヤーに関する検討①：権利侵害情報への対応の在り方	11
1-1	総論	11
1-2	被害者以外の第三者からの申出への対応	13
2	情報レイヤーに関する検討②：その他違法情報への対応の在り方	16
2-1	行政機関からの通報への迅速な対応	16
2-1-1	窓口整備などの体制整備	17
2-1-2	行政機関から法令違反情報に関する削除要請があった場合の免責	23
2-2	適格性のある第三者からの通報への対応	26
3	情報レイヤーに関する検討③：有害情報への対応の在り方	31
4	サービス設計レイヤーに関する検討：違法情報・有害情報共通の対応の在り方	34
4-1	違法・有害情報が流通するリスクとリスクへの対応の在り方	36
4-1-1	リスク評価・軽減措置に関する対応の在り方	36
4-1-2	経済的インセンティブに起因するリスクへの対応の在り方	47
4-1-3	レコメンダシステムに関する対応の在り方	51
4-2	適切な情報表示の在り方	56
4-2-1	プロミネンス（優先表示）に関する対応の在り方	57
4-2-2	AI 生成物のラベル付与に関する対応の在り方	58
4-3	利用者の確認に関する対応の在り方（アカウント開設時の本人確認等）	62
4-4	小括	65
5	その他	65
5-1	マルチステークホルダーによる協議会	65
5-2	執行当局の執行権限・体制の強化	66
V	結論	70

※ 本取りまとめにおいて、事業者名は、以下のとおり記載するものとする。

正式名称	記載名
Google LLC	Google
LINE ヤフー株式会社	LINE ヤフー
Meta Platforms, Inc.	Meta
TikTok Pte. Ltd.	TikTok
X Corp.	X

I 検討の背景

- ・ デジタル空間は、多様なコミュニケーションや、情報発信、情報収集を可能とし、人々の日常生活や社会経済活動を飛躍的に発展させ現在では、社会生活や経済活動に必要不可欠な場となっている。近時は、スマートフォン等の普及とともに、SNS 等のソーシャルメディアや動画共有サービスの利用が急速に拡大し、個人による情報発信がより容易かつ身近になった。
- ・ その一方で、デジタル空間において、誹謗中傷をはじめとする違法・有害情報の流通・拡散は依然深刻な状況である。また、生成 AI 等の新しい技術やサービスの進展及びデジタル広告の流通に伴う新たなリスクなど、デジタル空間における情報流通に伴う様々な諸課題が生じている。
- ・ 2023 年 11 月から開催された「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¹」（座長：宍戸常寿 東京大学院法学政治学研究科教授。以下「健全性検討会」という。）では、デジタル空間における情報流通に伴う様々な諸課題に対し、制度的対応を含む提言がなされた²（2024 年 9 月）。
- ・ 上記提言後も、デジタル空間における情報流通を巡る環境は変化してきており、違法・有害情報の流通を背景とした新たな社会問題（SNS 等における犯罪実行者の募集情報など）も発生している。
- ・ このような中、我が国においては、「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の一部を改正する法律」（令和 6 年法律第 25 号。以下「令和 6 年改正法」という。）が 2024 年 5 月 17 日に成立した³。本改正法により、法の題名が「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」（平成 13 年法律第 137 号。以下「情報流通プラットフォーム対処法」という。）に改められるとともに、大規模プラットフォーム事業者に対し、権利侵害情報の削除対応の迅速化や運用状況の透明化に係る措置が義務付けられることとなった。

¹ https://www.soumu.go.jp/main_sosiki/kenkyu/digital_space/index.html

² https://www.soumu.go.jp/main_content/000966997.pdf

³ 同法は 2025 年 4 月 1 日に施行された。

- ・ また、諸外国においても、違法・有害情報対策を始めとするデジタル空間における情報流通の規律の在り方については様々な動きがある。欧州連合（EU）ではデジタルサービス法（Digital Services Act。以下「DSA」という。）が2024年に完全施行され、英国ではオンライン安全法（Online Safety Act。以下「OSA」という。）が2023年に成立している。各国で取組が先行しているようにも見える一方、まだ運用の詳細が固まっていない条項や施行されていない条項もある。また、豪州では、2024年に偽・誤情報の対策強化を目的とした通信法改正案が議会で審議されたが、政府による検閲の懸念に基づく反対があり、成立が見送られている。

このように、デジタル空間における情報流通の規律の在り方については、諸外国においても未だ試行錯誤の中、検討を重ねている段階といえる。

- ・ こうした現状を踏まえ、2024年10月、デジタル空間における情報流通に伴う様々な諸課題について、制度整備を含むその対処の在り方等を検討するため、「デジタル空間における情報流通の諸課題への対処に関する検討会」（座長：宍戸常寿 東京大学院法学政治学研究科教授。以下「諸課題検討会」という。）が開催され、議論を開始した⁴。2025年1月からは、制度整備の在り方を議論することを目的として、「デジタル空間における情報流通に係る制度ワーキンググループ」（以下「本WG」という。）が開催され、議論を開始した⁵。
- ・ 本WGでは、こうした諸外国の政策動向も踏まえつつ、情報流通プラットフォーム対処法において迅速化規律が整備された「権利侵害情報」のほか、
 - ① 「権利侵害情報」には該当しない「その他違法情報（法令違反情報）」
 - ② 「有害情報」の2種類の情報についても議論を深め、政府や事業者による対応の在り方について、12回にわたって検討を行った。
- ・ このうち3回は、プラットフォーム事業者から、闇バイトの募集投稿を含む違法情報等への対応の状況について、ヒアリングを行った。
- ・ 今後、我が国におけるインターネット上の違法・有害情報対策は、どのような方向に進めていくべきか。本取りまとめは、最新の諸外国の政策動向も踏まえつつ、健全性検討会で深掘りすべきとされた論点について本WGとしての一

⁴ https://www.soumu.go.jp/menu_news/s-news/01ryutsu02_02000418.html

⁵ https://www.soumu.go.jp/main_sosiki/kenkyu/digital_shokadai/index.html

定の考え方を示すとともに、制度的対応の在り方について幅広く提言を行うものである⁶。

⁶ なお、選挙における SNS 利用の在り方については、表現の自由や政治活動、選挙運動の自由にも関わる重要な問題であり、各党各会派において議論が行われるべき事柄である。2025 年 6 月現在、「選挙運動に関する各党協議会」において議論されている状況であり、本取りまとめはその制度的対応の在り方について方向性を示すものではない。

Ⅱ 検討の視点 ～レイヤーごとの整理～

- ・ インターネット上の違法・有害情報は、短時間で広範に流通・拡散し、国民生活や社会経済活動に重大な影響を及ぼし得る深刻な課題である。「違法情報」は、名誉権、著作権等の他人の権利を侵害する「権利侵害情報」と、権利侵害情報には該当しないものの刑法等の法令でその情報を流通させることが違法とされている「その他違法情報（法令違反情報）」に分類される。また、「有害情報」は、違法情報ではない情報のうち、例えば、青少年有害情報⁷（青少年の健全な成長を著しく阻害する情報（暴力、アダルト等））のほか、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、存在しない災害・事故・事件が存在するかのように見せかけた偽画像・偽動画、災害発生時に外国人が犯罪行為を行っているとする偽・誤情報などが含まれる。
- ・ インターネット上の違法・有害情報について、表現の自由との関係で議論が実施しやすいよう、本 WG では、「違法情報・有害情報といった情報の種類に着目した切り口」（図表 1 参照）と「プラットフォーム事業者のサービス設計に着目した切り口」の両面からの対策を検討するため、2つのレイヤー（①情報レイヤーと②サービス設計レイヤー）に分けて論点を整理し、議論・検討を進めた。
- ・ 「情報レイヤー」の検討では、プラットフォーム事業者が運営する SNS や動画共有サイト、その他、電子掲示板やブログなどにおいて流通する情報（権利侵害情報、その他違法情報（法令違反情報）・有害情報等を含む個々の情報の流通）自体に焦点を当てた。本 WG において検討対象となる情報の種類は、健全性検討会の際に整理された分類を踏襲し、権利侵害情報、その他違法情報（法令違反情報）及び有害情報とした。いわゆる偽・誤情報も、基本的には、この3つのカテゴリのいずれかに分類することが可能であり、それぞれの項目における検討に包含されるものである（図表 1 参照）。
- ・ また、「サービス設計レイヤー」の検討では、情報流通の場となるプラットフォームの設計、具体的には、提供されるサービスの内容（利用者同士の投稿

⁷ 「青少年有害情報」については、法令上、「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（平成 20 年法律第 79 号）第 2 条第 3 項で定義されている。

を通じた交流の方法など)、インターフェース(投稿の表示場所など)、利用者にどの情報を表示するかを決定するアルゴリズムなどに焦点を当てた。

図表 1 : 情報の種類

違法情報 有害情報	権利侵害情報	A	偽・誤情報 偽情報 ・人を混乱させ惑わすために意図的・意識的に作られたウソ、虚偽の情報 誤情報 ・勘違いや誤解により拡散された間違った情報
	その他違法情報（法令違反情報）	B	
	有害情報	C	

種類		具体例
A	権利侵害情報	<ul style="list-style-type: none"> 虚偽の事実の摘示による誹謗中傷（名誉毀損）等
B	その他違法情報（法令違反情報）	<ul style="list-style-type: none"> 虚偽・誇大広告（景表法・特商法等違反） いわゆる「闇バイト」等の犯罪の実行者を誘引（募集）する情報（職安法違反）等
C	有害情報	<ul style="list-style-type: none"> 青少年有害情報（青少年の健全な成長を著しく阻害する情報（暴力、アダルト等）） 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 存在しない災害・事故・事件が存在するかのように見せかけた偽画像・偽動画 災害発生時に外国人が犯罪行為を行っているとする偽・誤情報 等

Ⅲ 事業者ヒアリング

- ・ 本 WG では、本 WG の親会である諸課題検討会及び「ICT サービスの利用環境の整備に関する研究会 通信ログの保存の在り方に関するワーキンググループ⁸」（主査：鎮目征樹 学習院大学法学部教授。以下「ログ保存 WG」という。）との合同開催として、プラットフォーム事業者によるインターネット上の違法・有害情報への取組の現状を把握するため、2024 年後期に特に大きな社会問題となったインターネット上の闇バイトの募集投稿⁹を含む違法・有害情報対策について、プラットフォーム事業者からヒアリングを行った¹⁰。ヒアリング対象としたのは、アクティブユーザー数が多い事業者（当該事業者又はその関連事業者が日本国内における 2023 年度末時点の平均月間アクティブユーザー数が 1,000 万人以上）である Google、LINE ヤフー、Meta、TikTok、X の計 5 事業者（以下「ヒアリング対象事業者」という。）である。
- ・ インターネット上の闇バイトの募集投稿への事業者の対応については、2024 年 12 月に総務省が事業者に対して行った要請¹¹内容を中心にヒアリングを実施し、その結果について、2025 年 5 月 15 日、別紙のとおり「ヒアリング総括（案）」をまとめたところである。
- ・ 闇バイトの募集投稿以外の一般的な違法・有害情報へのヒアリング対象事業者の対応については、以下のとおりである。なお、以下の内容は、ヒアリング対象事業者からのヒアリングシート回答及びヒアリング結果に基づき整理したものである。

⁸ https://www.soumu.go.jp/main_sosiki/kenkyu/ICT_services/index.html

⁹ 闇バイトの募集投稿とは、SNS 等において、仕事の内容を明らかにせずに著しく高額な報酬の支払いを示唆するなどして犯罪の実行者を募集する投稿のことを指す。

¹⁰ このヒアリングの際には、いわゆる通信履歴（ログ）の保存の在り方についてもヒアリングを行ったが、その論点等については、「ICT サービスの利用環境の整備に関する研究会」及びその下に設置されたログ保存 WG で議論が行われており、本 WG では扱わないこととする。

¹¹ 総務省は、2024 年 12 月 18 日、SNS 等を提供する大規模事業者に対して、SNS 等におけるいわゆる「闇バイト」への対応について、文書により要請を実施。具体的には、①闇バイトの募集活動に係る投稿に対する利用規約等に基づく対応、②SNS 等のアカウント開設時における本人確認手法の厳格化、③SNS 等に対する照会への回答の円滑化、④SNS 等の利用者に対する注意喚起・周知活動、の 4 点について要請。

(https://www.soumu.go.jp/menu_news/s-news/01kiban18_01000242.html)

【違法・有害情報に関する事業者の対応状況（ヒアリング結果）】

（１）違法情報に対する利用規約等¹²に基づく対応

- 全てのヒアリング対象事業者から、利用規約等において、法令に違反する情報の投稿を禁止しているとの回答があった。過半のヒアリング対象事業者からは、利用規約等に照らして違法性の判断を行い、法令違反のおそれがあることを確認した際には、投稿の削除やアカウントの停止措置を実施しているとの回答があった。

（２）違法情報への対応の体制整備状況

（ｉ）通報窓口の整備状況

- 利用者からの情報削除等の申告への対応について、全てのヒアリング対象事業者から、オンライン上で利用可能な形式で受付を実施しており、24 時間 365 日対応可能な体制（審査体制も含む。）を整備しているとの回答があった。
- 行政機関からの情報削除等の申告への対応について、利用者からの申告と同様の対応をしていると回答した事業者、専用窓口を設けていると回答した事業者、窓口の形ではないが、一部の行政機関とのチャネルを持っていると回答した事業者があった。
- 適格性のある第三者機関からの情報削除等の申告への対応について、利用者からの申告と同様の対応をしているとの回答をした事業者、専門性等をあらかじめ認定した政府機関又は NGO からの申告について、優先対応を行っているとの回答した事業者があった。

（ii）検知方法

- 違法情報も含む問題のある情報の検知方法について、全てのヒアリング対象事業者において、AI を利用してスクリーニングを行う方法と人の目による検知を行う方法を併用しているとの回答があった。

（iii）標準処理期間

¹² 事業者によって、自社のサービスを利用する際のルールについては、利用規約やポリシーなど様々な呼び方をしているところ、本取りまとめでは「利用規約等」と表現することとする。

- 違法情報の削除に至るまでの標準処理期間について、具体的な期間について非公開で回答した事業者もあったが、公表していない旨の回答をした事業者があったほか、具体的な期間について回答がなく、できる限り早く削除している旨の回答があった事業者もあった。

(iv) 不服申立て

- 申告者に対する通知について、行っていると回答した事業者、行っていないと回答した事業者、無回答の事業者があったが、情報を削除した場合の発信者に対する通知については、全ての事業者が既に行っている又は行う予定であると回答した。なお、不服申立てについては、全ての事業者からオンラインで利用可能な形式で受け付けていると回答した。

(3) 違法・有害情報への共通する対応

- 全てのヒアリング対象事業者が、違法・有害情報の流通・拡散を抑止するための方策として、利用規約等に違反する内容の投稿を繰り返し発信する行為に対し、利用規約等に基づき、コンテンツの削除又はアカウント停止等の対応を行っていると回答した。また、過半のヒアリング対象事業者は、違反回数が一定を超えた場合には、アカウントの削除等、違反回数に応じた対応を行っていると回答した。
- 災害時等の特定の場面に着目した取組も行われている。例えば、ニュース、SNS、ユーザーレポート等を監視して、不適切なコンテンツに関連する新しいトレンドを検出し、社会問題化する前に対応できる体制を準備していると回答した事業者があった。また、災害時における被災地向けの正確な情報発信や、投稿ユーザーへの注意喚起を行っていると回答した事業者もあった。さらに、選挙時において信頼できる情報源への誘導を行っていると回答した事業者もあった。
- 諸外国における自らのサービスが及ぼす社会的影響の評価・軽減措置に関する制度に対する評価については、過半のヒアリング対象事業者が、EU(DSA)におけるリスク評価作業の負担への懸念があると回答したほか、我が国において同様の制度の導入を検討するに当たっては、規制の導入よりも定期的なレポートを通じた継続的な透明性の確保が望ましいと回答した事業者、日本の状況を踏まえて改めて回答したいと回答した事業者もあった。

IV 対応の方向性

1 情報レイヤーに関する検討①：権利侵害情報への対応の在り方

1-1 総論

情報流通プラットフォーム対処法

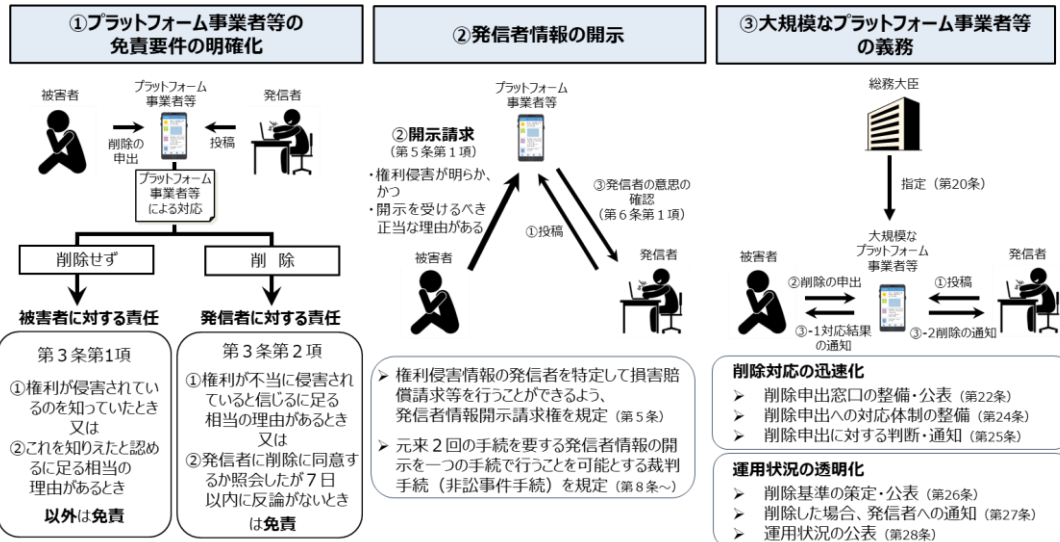
- ・ 情報流通プラットフォーム対処法においては、総務大臣が指定する一定規模以上のプラットフォームサービスを提供する事業者（以下「大規模プラットフォーム事業者」という。）に対し、①権利侵害情報の削除対応の迅速化や②運用状況の透明化に係る措置が義務付けられた。
- ・ 具体的には、①削除対応の迅速化に係る措置として、被侵害者から権利侵害情報の削除申出を受け付けるための窓口・手続の整備・公表、十分な知識経験を有する者の選任、削除申出に対する原則7日以内の判断・通知等が義務付けられている（以下「迅速化規律」という。）。
- ・ 併せて、②運用状況の透明化に係る措置として、削除基準の策定・公表、削除した場合の発信者への通知、運用状況の公表が義務付けられている（以下「透明化規律」という。）。透明化規律の対象情報について、同規律は、大規模プラットフォーム事業者が削除又はアカウント停止の対象としている情報に対して適用されることとなる。すなわち、当該事業者が、権利侵害情報のみならず、その他違法情報や有害情報についても、事業者が自ら定め公表する削除等の基準（以下「削除基準」という。）に含めている場合には、これらの情報に対しても透明化規律が適用されることとなる。

図表2：情報流通プラットフォーム対処法の概要

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）

（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（平成13年法律第137号））

インターネット上の違法・有害情報の流通が社会問題となっていることを踏まえ、「**被害者救済**」と発信者の「**表現の自由**」という重要な権利・利益のバランスに配慮しつつ、プラットフォーム事業者等がインターネット上の権利侵害等への対処を適切に行うことができるようするための法制度を整備するもの。



違法情報ガイドライン

- ・ 総務省では、令和6年改正法の施行と併せて、どのような情報を流通させることが権利侵害や法令違反に該当するのかを明確化する「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第26条に関するガイドライン」（以下「違法情報ガイドライン」という。）を策定した。同ガイドラインでは、①他人の権利を不当に侵害する情報の送信を防止する義務がある場合（権利侵害情報）と②その他送信防止措置を講ずる法令上の義務がある場合（法令違反情報）を例示列挙している。
- ・ 大規模プラットフォーム事業者においては、違法情報ガイドラインの内容を削除基準に盛り込むことで、利用者に対して運用状況の透明性を確保するとともに、権利侵害情報や法令違反情報に対して削除基準に基づき適切に対応することが期待されている。

図表 3：違法情報ガイドラインの概要

違法情報ガイドライン

情報流通プラットフォーム対処法第26条第1項第2号に定める「他人の権利を不当に侵害する情報の送信を防止する義務がある場合その他送信防止措置を講ずる法令上の義務（努力義務を除く。）がある場合」を例示することにより、

- どのような情報を流通させることが権利侵害や法令違反に該当するのかを明確化するとともに、
- 大規模特定電気通信役務提供者が送信防止措置の実施に関する基準を策定する際に盛り込むべき違法情報を例示する。
- あわせて、第28条に基づく運用状況の公表に当たり、可能な限り、申出の受付件数等をガイドラインの分類に基づいて区分することを求める。

1. 他人の権利を不当に侵害する情報の送信を防止する義務がある場合（権利侵害情報）

1-1. 対象となる権利・利益

- ・ 名誉権、名誉感情、プライバシー、私生活の平穩、肖像権、氏名権、パブリシティ権、著作権及び著作隣接権、商標権、営業上の利益について、どのような場合に各権利・利益の侵害が成立するかを明確化し、関連する裁判例もあわせて掲載する。

1-2. 情報の送信を防止する義務が生ずる場合

- ・ 大規模特定電気通信役務提供者に送信防止措置を講ずる義務が生ずる場合として、「人格権侵害其他法令の規定に基づく差止請求」及び「条理上の義務があると認められる場合」を規定する。

2. その他送信防止措置を講ずる法令上の義務（努力義務を除く。）がある場合（法令違反情報）

2-1. 対象となる情報

- ・ わいせつ関係、薬物関係、振り込め詐欺関係、犯罪実行者の募集関係、金融業関係、消費者取引における表示関係、銃刀法関係、その他の区分に基づき、関係法令を分類するとともに、どのような情報を流通させることが各法令に違反するのかを具体的に示す。

2-2. 情報の送信を防止する義務が生ずる場合

- ・ 大規模特定電気通信役務提供者に送信防止措置を講ずる義務が生ずる場合を規定する。

1-2 被侵害者以外の第三者からの申出への対応

- ・ 健全性検討会のとりまとめでは、被害者救済により資するという観点から、「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。」との提言がなされた¹³。

我が国における現行制度や対応状況

- ・ 現行の情報流通プラットフォーム対処法では、1-1で述べたとおり、大規模プラットフォーム事業者に迅速化規律が課せられるのは、被侵害者本人が侵害情報送信防止措置を講ずるよう申出を行った場合に限定されている（第22条等）。適切な代理権を有する者が申出を行うことは可能であるが、代理権を

¹³ 健全性検討会とりまとめ 92 ページ（注 156）を参照。

有さない第三者からの申出については、情報流通プラットフォーム対処法の迅速化規律は適用されない。

具体的検討

- ・ 本 WG では、被侵害者以外の第三者からの申出についても、大規模プラットフォーム事業者に対して削除の対応の迅速化を義務付けるべきか検討を行った。
- ・ まず、被侵害者以外の第三者からの申出に対して大規模プラットフォーム事業者による迅速な対応が求められる場面としては、被侵害者自身が被害を申し出ることが困難である場合、特に、被侵害者の精神的な損害が大きく、本人による申出が難しい場合等が想定される。そのような場合に、適切な第三者からの申出に対しても大規模プラットフォーム事業者による迅速な対応を求める必要があるかどうかも含め¹⁴、今後、総務省は、関係省庁と連携しつつ、被害者支援団体等の聴取などを通じた実態把握を行うことが適当である。
- ・ また、現時点においても、大規模プラットフォーム事業者は、権利侵害情報について、被侵害者以外の第三者からも削除申出を受け付けている。削除要請への対応に関して、被侵害者からの申出に限って情報流通プラットフォーム対処法の義務である迅速な対応（原則として7日間以内の判断・通知）等を行うべきか、あるいは被侵害者以外の第三者からの申出であっても情報流通プラットフォーム対処法の迅速化規律に則った対応を行うべきかについては、今後の大規模プラットフォーム事業者による対応も注視して検討することが適当である。
- ・ この点、情報流通プラットフォーム対処法の透明化規律により、被侵害者以外の第三者からの削除申出も含めて¹⁵、今後、大規模プラットフォーム事業者が対応状況を公表することとなっていることから、総務省は、情報流通プラッ

¹⁴ このほか、迅速な対応を義務付けるためには、第三者からの申出がどのような根拠に基づくのか等についても検討する必要がある。

¹⁵ 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施行規則（令和4年総務省令第39号）（以下「総務省令」という。）では、利用者からの通報を受けて送信防止措置を講じた件数・講じなかった件数や公的機関から送信防止措置を講ずるよう要請があった件数、当該要請を受けて送信防止措置を講じた件数及び講じなかった件数等を公表事項としている（総務省令第18条第5項第2号、第4号、第5号等）。

プラットフォーム対処法の透明化規律により公表された事項に関する適切なモニタリング等も通じて実態を把握しつつ、当該規律を設ける必要性等について、更に検討することが適当である¹⁶。

- ・ なお、検討の際には、あらゆる者から根拠不十分の大量の申出が濫用的になされ、大規模プラットフォーム事業者の対応が滞り、結果として本来迅速に対応すべき申出に適切に対応できなくなるおそれもあることから、「適切な」第三者としてどのような要件を満たす者・団体を想定すべきかについても併せて検討することが重要である。
- ・ この点、被侵害者以外の第三者が権利侵害情報に関する削除申出に携わっている事例として、例えば、著作権・商標権侵害への対応に関して、「情報流通プラットフォーム対処法ガイドライン等検討協議会」が策定する「著作権関係ガイドライン」及び「商標権関係ガイドライン」における「信頼性確認団体¹⁷」の取組が挙げられる。具体的には、プラットフォーム事業者等による、より迅速かつ適切な対応を目指し、特に権利侵害の有無等の判断につき「信頼性確認団体」を経由した場合には、プラットフォーム事業者等において権利侵害の有無等の判断が適切に確認されていると判断して送信防止措置を講ずることができるとされている。

また、法務省の人権擁護機関¹⁸においては、インターネット上の誹謗中傷等の情報が人権侵害に当たるとして相談を受けた場合には、相談者の意向に応じ、当該情報の違法性の有無を判断した上で、人権侵犯事件調査処理規程（平成16年法務省訓令第2号）に基づき、事業者への削除要請をするなどの対応を行っている。加えて、一部の地方公共団体においては、条例に基づき、専門家を含む審査会等の判断を踏まえた上で、プラットフォーム事業者に対して

¹⁶ なお、構成員からは、被侵害者以外の第三者が被侵害者の意思とは別に勝手に申出を行い、プラットフォーム事業者が当該情報を削除してしまった場合に、プラットフォーム事業者によっては、IPアドレス等のログが消え、その結果、被侵害者が発信者情報開示に基づく損害賠償請求をすることができなくなる可能性があるとの指摘があった。

¹⁷ 著作権等侵害であることの確認をプラットフォーム事業者に代わって適切に判断することを目的に、「著作権関係ガイドライン」及び「商標権関係ガイドライン」において規定されている団体。信頼性確認団体として活動するためには認定を受ける必要があり、認定要件として、法人であること、著作権等に関する専門的な知識及び相当期間にわたる十分な実績を有していること並びに著作権等の侵害等の確認等を適切に行うことができることが掲げられている。

¹⁸ 法務省の人権擁護機関とは、国の機関として人権擁護に取り組んでいる法務省人権擁護局、その地方支分部局である法務局・地方法務局及びその支局、法務大臣が委嘱する人権擁護委員を指す。

削除申出の要請を行うスキームも運用されている。

- ・ 被侵害者以外の第三者による削除申出に関する制度検討に際しては、これらのスキームを参考にしつつ、どのような要件を満たす者・団体であれば、濫用的な申出を防ぎつつ、適切なプロセスにより、大規模プラットフォーム事業者に迅速な対応を求めるに足ると判断できるかについて、引き続き検討することが適切である^{19, 20}。

2 情報レイヤーに関する検討②：その他違法情報への対応の在り方

- ・ 「その他違法情報」とは、権利侵害情報以外の法令違反情報を指すものであり、例えば、虚偽・誇大広告（景表法・特商法等違反）、犯罪実行者の募集情報（いわゆる闇バイトの募集情報）（職業安定法（昭和 22 年法律第 141 号。以下「職安法」という。）違反）等が該当する。本節では、文章中、権利侵害情報以外の「その他」情報であることを強調する必要がある場合以外は、分かりやすさの観点から、単に「法令違反情報」と表記する。
- ・ 1-1 で述べたとおり、情報流通プラットフォーム対処法の迅速化規律の適用は、被侵害者本人が権利侵害情報に関する送信防止措置を講ずるよう大規模プラットフォーム事業者に対し申出を行った場合に限られており、法令違反情報は対象とはならない。一方、透明化規律の適用については、削除基準に含まれている場合には、法令違反情報も対象となる。

2-1 行政機関からの通報への迅速な対応

- ・ 健全性検討会のとりまとめでは、行政法規に抵触する違法な偽・誤情報について、対応の迅速化を通じた実施の促進を図る観点から、『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』、『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

¹⁹ なお、構成員からは、被侵害者以外の第三者による削除要請については、被害者からの削除要請を当該第三者が代行する形となるため、弁護士法第 72 条の非弁行為との関係で整理が必要ではないかといった指摘があった。

²⁰ 「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における大規模特定電気通信役務提供者の義務に関するガイドライン」においては、被侵害者以外の者による削除申出についても、法の規定に準じて、速やかに対応を行うことが望ましいと記載されている。

送 PF 事業者を免責』等の対応を具体的に進めることが適当」であるとの提言がなされた。また、違法性の判断能力の観点から、「行政法規を所管する行政機関からの申出・要請を契機としたコンテンツモデレーションについて、具体化を進めることが適当である」との提言がなされた。

- ・ 健全性検討会のとりまとめでは、「違法な偽・誤情報」としているが、本 WG では、権利侵害情報以外の違法情報に一般化して検討を行った。
- ・ とりまとめの提言について、整理すると以下になるため、それぞれ検討する。
 - ① コンテンツモデレーションに関する外部からの申出・窓口を整備・公表すること、行政機関からの申出を契機としたコンテンツモデレーションの体制整備を検討すること
 - ② 一定の条件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プラットフォーム事業者を免責すること

2-1-1 窓口整備などの体制整備

我が国における現行制度や対応状況

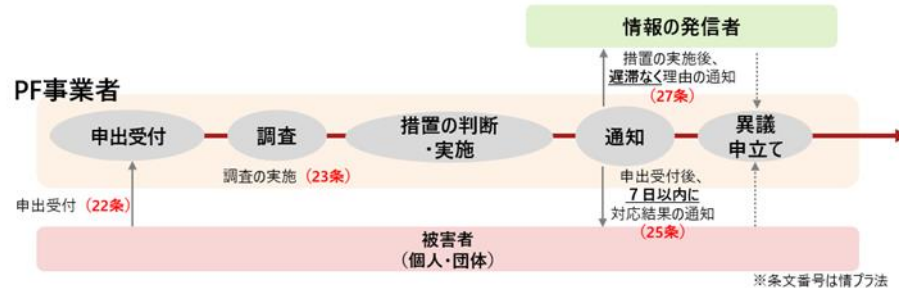
- ・ 情報流通プラットフォーム対処法では、立法時における検討に際しての課題として、プラットフォーム事業者の設置する削除申出受付窓口が被害者等から所在が分かりにくいなどの指摘があった²¹ことを踏まえ、大規模プラットフォーム事業者に対し、迅速化規律の一環として、権利侵害情報の削除申出を受け付けるための窓口・手続の整備・公表、十分な知識経験を有する者の選任、削除申出に対する原則 7 日以内の判断・通知等の体制整備を義務付けている（第 22 条、第 24 条、第 25 条等）。一方で、その他違法情報の通報に関しては、このような体制整備を義務付ける規定はない。両者の対応フローの差異を図にすると、以下のとおりとなる（図表 4）。

図表 4：権利侵害情報への対応とその他違法情報への対応 フロー比較

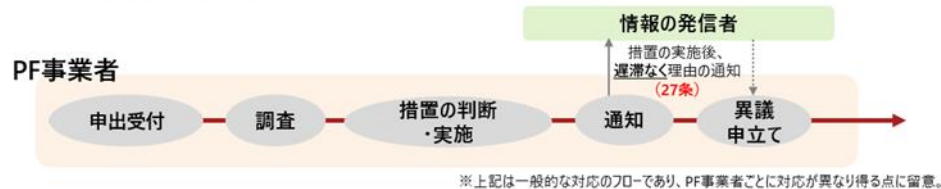
²¹ 総務省「プラットフォームサービスに関する研究会 第三次とりまとめ」（令和 6 年 1 月）9 頁

- 権利侵害情報の場合、窓口設置から対応結果の通知までについて、情プラ法に規定がある（削除対応の迅速化規律）。対応結果の通知までの期間は、原則 7 日。
- 一方、**その他違法情報については情プラ法の削除対応の迅速化規律の対象とはなっていない。**

＜情プラ法に基づき権利侵害情報に関して措置を行う場合の一般的な対応のフロー＞



＜利用規約等に基づきその他違法情報に関して措置を行う場合の一般的な対応のフロー＞



- また、現状、一部の行政機関や団体が、大規模プラットフォーム事業者に対して、法令違反情報の削除依頼や情報提供を行っている。例えば、都道府県警察では、違法・有害情報について削除依頼を行っている。また、各省庁が所管する行政規律の中には、特定の種類の情報流通を違法（法令違反情報）と位置付けるとともに、当該法令違反情報に関して警察組織以外の行政機関が大規模プラットフォーム事業者に削除要請を行うことができる規定が設けられている例がある。例えば、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（昭和 35 年法律第 145 号。以下「薬機法」という。）では、虚偽・誇大広告（第 66 条）や承認前医薬品等の広告（第 68 条）について、厚生労働大臣又は都道府県知事が大規模プラットフォーム事業者に対して削除要請を行うことができるとされている。
- 本 WG において実施した事業者ヒアリング（詳細は、Ⅲ 参照。）では、行政機関からの削除要請に対する大規模プラットフォーム事業者側の窓口整備の状況としては、過半の事業者から専門の窓口を設けて対応しているとの回答があった一方で、一般の利用者と同様に扱っている、すなわち行政機関用に特別な窓口を設置していないと回答した事業者も存在した。また、事業者が提供し

ている優先対応スキーム²²を利用する行政機関からの申告については、優先対応を行っているとは回答した社もあった。削除等の判断に関する申告者に対する通知等の状況は、事業者によって対応が区々であり、対応の迅速性に関しても、過半の事業者の標準処理期間は把握できない状況にある。

なお、全てのヒアリング対象事業者から、法令に違反する内容の投稿は禁止しているとの回答があり、また、過半の事業者から、法令に違反する投稿を確認した際は、削除対応を行っているとの回答があった。一方で、一部の事業者から、コンテンツが違法かどうかを評価することは困難であるといった回答もあった。

諸外国の状況

- ・ EU (DSA) では、インターネット上の違法情報への迅速な対応を図るため、ユーザー向けの通報窓口の設置を設置し公表することを事業者に義務付けている(第 12 条)。また、ユーザーが違法情報をオンラインで容易に通報(notice)できる仕組みを準備することを義務付けている²³ (第 16 条)。さらに、行政機関からの命令には、対応の有無と対応内容を遅滞なく当該行政機関に通知することを義務付けている (第 9 条)。

また、EU (DSA) では、違法コンテンツ対策の実施命令や情報提供命令といった司法又は行政当局が発する命令に限定されるものであるが、命令に関係するサービス利用者には、当局から命令とその効果について理由を含めて通知するという手続保障の規定が置かれている (第 9 条第 5 項、第 10 条第 5 項)。また、事業者には、違法コンテンツ対策の実施命令や情報提供命令の件数等について、少なくとも一年に一回、報告書を公開することを義務付けている (第 15 条)。

²² 例えば、YouTube の優先報告者プログラムにおいては、同プログラムに参加した政府機関と非政府組織 (NGO) から報告された動画について、YouTube が優先的に審査する取組等がなされている (<https://support.google.com/youtube/answer/7554338?hl=ja-JP>)。例えば、2021 年 8 月に新型コロナウイルス感染症や新型コロナウイルスワクチンに関する正確な情報を届ける医師のプロジェクト、「コロワくんサポーターズ」が参加している (<https://youtube-jp.googleblog.com/2021/08/jpcovidefforts.html>)。

²³ DSA 前文 (50) では、「・・・通報を促す、アクセスしやすく使いやすい通報と対応の仕組みを、その規模にかかわらずあらゆるホスティングサービスプロバイダが導入することが重要である。かかる仕組みは、明確に識別可能で、問題となる情報の近くに配置され、少なくともホスティングサービスプロバイダの利用規約に違反するコンテンツの通報の仕組みと同程度に簡単に見つけて使用できるものとしなければならない。」とされている。

- ・ 英国（OSA）では、OSAに基づく行動規範（Codes of practice）²⁴において、詐欺コンテンツについて、事業者がトラステッド・フラグガーから報告を受け付ける専用窓口を設けることが推奨されている。トラステッド・フラグガーとしては、警察、労働年金省、国家犯罪庁等の行政機関とされている。

具体的検討

- ・ 本 WG では、違法情報に関する行政機関からの削除要請に関して、①大規模プラットフォーム事業者の体制整備（削除申出受付窓口・手続の整備、優先対応、迅速対応等）、②優先対応スキームを利用する通報・要請の範囲、③発信者の手続保障について検討を行った。

① 大規模プラットフォーム事業者の体制整備

- ・ 事業者ヒアリングの結果、行政機関向けの申請窓口の設置状況が区々であること、必ずしも全てのヒアリング対象事業者が行政機関向けの申請窓口を設置し、又は特定の行政機関からの通報の優先対応スキームを設けている状況ではないことが明らかとなった。行政機関向けの申請窓口を設けていない大規模プラットフォーム事業者に対して行政機関が違法情報の削除要請を行う場合、一般利用者向けの窓口から申請を行うこととなり、多数の一般利用者からの削除要請と区別せずに判断が行われるため、大規模プラットフォーム事業者による判断や対応が必ずしも速やかに行われない可能性がある。行政機関向けの専用窓口を設けるなどの体制を整備することで、事業者の対応が円滑になることが期待される。例えば、警察庁の発表²⁵によると、闇バイト対策については、同庁が、ある事業者と協議し、同事業者において、インターネット・ホットラインセンター（IHC）からの犯罪実行者募集情報の通報に関する運用プロセスを変更した結果、削除依頼に対する 5 営業日以内の削除率が、90%を超えたとされている。
- ・ また、事業者ヒアリングでは、一部のヒアリング対象事業者からは、対応の困難性の理由として、違法性の判断が困難である旨との回答もあった。この点、

²⁴ 事業者が各種の義務を遵守するための内容を Ofcom（英国通信庁）が定めたもの。

²⁵ 「令和 6 年におけるサイバー空間をめぐる脅威の情勢等について」（2025 年 3 月 13 日）

（https://www.npa.go.jp/publications/statistics/cybersecurity/data/R6/R06_cyber_jousei.pdf）

一般論としては、個別法を所管している行政機関は法解釈権を持っており、厚生労働省における職安法の解釈の明確化²⁶などの例にもあるとおり、専門的かつ正確な判断を行い得る主体である。こうしたことを踏まえれば、法令違反情報への対応の在り方として、違法性の判断が可能な専門の行政機関からの通報を受け付け、優先的に対応する体制整備を大規模プラットフォーム事業者を求めることは、制度的対応の方向性として有効な手段であると考えられる。

- ・ しかし、事業者ヒアリングでは、行政機関からの削除要請に対して、違法情報の削除に至るまでの標準処理期間については、具体的な期間について非公開で回答した事業者もあったが、公表していない旨の回答をした事業者があったほか、具体的な期間について回答がなく、できる限り早く削除している旨の回答があった事業者もあった。他方で、詳細は後述（２－２）するが、行政機関以外に削除要請を行っている団体へのヒアリング結果によると、法令違反情報の削除要請への対応については、団体によって多少の違いはあるものの、大規模プラットフォーム事業者が一定程度迅速な対応を行っているといえる例も見られる。
- ・ 以上を踏まえると、法令違反情報について、当該法令違反情報の違法性の判断が可能な行政機関からの削除要請への大規模プラットフォーム事業者による迅速な対応に関する制度的な検討に当たっては、行政機関からの削除要請に対してプラットフォーム事業者がどの程度の処理期間で対応を行っているかを含め、今後、取組状況を把握・分析し、表現の自由に配慮しつつ、制度的な対応の必要性を精査していく必要がある。

② 優先対応スキームを利用する通報・要請の範囲

- ・ 仮に大規模プラットフォーム事業者に体制整備を求めるとしても、法令違反情報には様々な種類・性質のものが含まれており、また、その流通がもたらす社会的影響の大小や罰則の有無も様々であることや、事業者による迅速な対応を過度な負担なく実現する観点からも、今後、優先的対応が求められる法令違反情報や削除要請の主体となり得る行政機関の種類・要件を更に精査し、スキームを利用する通報・要請の範囲を絞り込むことが適当である。

²⁶ 職安法第５条の４第１項においては、労働者の募集に関する情報等については、虚偽の表示又は誤解を生じさせる表示をしてはならないこととされているところ、厚生労働省は、令和６年１２月に同項に反しないと認められるためには、雇用しようとする者の氏名・名称等の６情報を含める必要があるとの解釈の明確化を行った。

- ・ その際、例えば、薬機法など、既にプラットフォーム事業者への削除要請が特定の行政規律によって規定されている類型の法令違反情報については、当該行政規律を所管する行政機関に対して、対応の迅速化へのニーズについても把握していくことが適当である²⁷。

また、本 WG で実施した IHC 及び一般社団法人セーファーインターネット協会（SIA）からのヒアリング²⁸によれば、違法情報のうちプラットフォーム事業者への削除依頼を実施したものは、ポルノ関係や預貯金通帳等の譲渡等の勧誘・誘因の件数が多いなど、内容によって件数に差異がある状況であった。こうした実態も踏まえながら、特に優先的に対応すべき法令違反情報の絞込みを行った上で、要請を行う行政機関による透明性の確保も含め、引き続き制度的対応の要否や内容の深掘りを行うことが適当である。

- ・ なお、ヒアリング対象事業者の一部は、既に行政機関向けの専門窓口の整備や優先対応を実施している。今後の制度的検討と並行して、こうした取組を実施していない事業者も含め、例えば、事業者自らが特定の行政機関を認定し、当該行政機関からの通報を専用窓口で受け付け、通報を信頼性が高いものとして優先的に取り扱うプロセスを確立するなど、法令違反情報への対応を迅速に行うための枠組みを自ら整備して対応することも期待される。

③ 発信者の手続保障

- ・ 現行の情報流通プラットフォーム対処法では、図表 4 で記載したとおり、利用規約等に基づいてその他違法情報の送信防止措置を実施した場合、当該情報の発信者に対して、遅滞なく理由の通知を行うことが義務付けられている（第 27 条）。
- ・ この規定は、権利侵害情報について「被害者救済」と「発信者の表現の自由」

²⁷ この点、具体的な法制度の検討に当たっては、薬機法のように、個別の行政規律に行政機関からプラットフォーム事業者に対する削除要請の法的根拠が必要であるという考え方がある一方で、そのような個別の規律がなくとも、別表方式で対象となる法令違反情報を列挙するとともに、それらに掲げる法令違反情報共通の対応としてプラットフォーム事業者に対する窓口整備義務を規律する形式も考えられるとの指摘があった。なお、別表方式で規律する場合、各行政法規に新たに削除要請権限を与えているのと同義となる可能性もあり、そのような対応が可能かについても、更なる検討が必要であると考えられる。

²⁸ 制度 WG 第 3 回 資料 3－1 及び 3－2 を参照。

のバランスを図るための規定であると考えられるところ、仮にその他違法情報にも迅速化規律を適用することとした場合には、その他違法情報には被害者が存在しないため、「発信者の表現の自由」とバランスをとるべき法益は「社会的法益」となると考えられる。このため、発信者への手続保障として現行の第 27 条のみで十分であるか検討が必要となる。

- ・ この点、法令違反情報について、行政機関からの削除要請への迅速な対応の制度化を図る場合、発信者の表現の自由の保障の観点から、例えば異議申立手続を追加的に整備するなど、発信者への手続保障のための制度的対応について、特に検討を深めるべきである。
- ・ また、行政機関の削除要請については、透明性・アカウンタビリティを確保するとともに、過度な要請に対する歯止めについても検討が必要である。情報流通プラットフォーム対処法の透明化規律²⁹により、行政機関からの要請件数³⁰の公表状況など、同法の運用状況も注視する必要がある。

2-1-2 行政機関から法令違反情報に関する削除要請があった場合の免責

我が国における現行制度や対応状況

- ・ 現行の情報流通プラットフォーム対処法では、プラットフォーム事業者にならず責任が生じない場合を可能な範囲で明確にする観点から、権利侵害情報については、事業者が削除を行った際、権利が不当に侵害されていると信じるに足る相当の理由がある場合又は発信者に削除に同意するか、照会したが 7 日以内に反論がない場合には、事業者は賠償の責めに任じない規定がある（第 3 条第 2 項）。
- ・ 薬機法では、虚偽・誇大広告、承認前医薬品等の違法広告について、厚生労働大臣又は都道府県知事から要請を受けて削除した場合、発信者に生じた損害については、必要な限度において行われたものであるときは、賠償の責めに任じない規定がある（72 条の 6）。

²⁹ 情報流通プラットフォーム対処法では、法令違反情報も含め、自らの削除基準に基づいて行った削除等の運用状況の公表が事業者には義務付けられている（第 28 条）。

³⁰ 総務省令では、公的機関から送信防止措置を講ずるよう要請があった件数について、要請があった理由の別に応じて区分されたものを公表事項としている（総務省令第 18 条第 5 項第 4 号）。

- ・ 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（平成 26 年法律第 126 号）では、私事性的画像記録について、事業者が削除を行った際、撮影対象者から削除の申出があった場合、発信者に削除に同意するか、照会したが 2 日以内に反論がない場合には、事業者は賠償の責めに任じない規定がある。

諸外国の状況

- ・ EU（DSA）では、事業者（ホスティングサービス）は、違法なコンテンツを実際に知っていた又は認識した場合に、迅速に違法情報を削除するか、アクセスを無効化する措置をとることで、利用者の求めに応じて蓄積した情報について免責される（第 6 条）。ただし、自らのサービスが違法コンテンツの保存にも利用されているという事実を一般的な意味で認識しているという根拠のみでは「実際に知っていた又は認識した」ものとはみなされない³¹。
- ・ 英国では、OSA 上の免責規定はないものの、電子商取引規則 2002³²により、ホスティングについて、①違法な行為・情報に関する現実の知識を有せず、かつ、それらがサービスプロバイダに明白となる事実・状況の認識がない場合、②そのような知識・認識を得た際に、直ちに、当該情報の削除・アクセス停止をする場合には、蓄積の結果による損害やその他のいかなる金銭的救済、刑事的責任について、これを負わないこととされている³³。本規則は、EU の電子商取引指令を受けたものであるが、英国の EU 離脱後も英国法として維持されている³⁴。

具体的検討

- ・ 一般的に、プラットフォーム事業者が利用規約等に基づきコンテンツを削除した場合は同規約等に基づき免責される。本 WG では、発信者との関係で、行政機関からの要請を受けて迅速に違法情報を削除したプラットフォーム事業

³¹ DSA 前文（22）参照。

³² The Electronic Commerce (EC Directive) Regulations 2002
(<https://www.legislation.gov.uk/ukxi/2002/2013/signature>)

³³ プロバイダ責任制限法検証WG提言（平成 23 年 6 月）
(https://www.soumu.go.jp/main_content/000134914.pdf)

³⁴ 同法は、ビジネス貿易省が公表している、EU 離脱後も英国法として機能する法律として、Retained EU law and assimilated law dashboard のリストに掲載されている。
(<https://www.gov.uk/government/publications/retained-eu-law-dashboard>)

者を制度的に免責する必要はあるかについて、検討を行った。

- ・ 一般論として、情報流通プラットフォーム対処法における権利侵害情報に関するプラットフォーム事業者の免責規定（第3条）や、諸外国の法令における規定のように、自らは発信者ではなく、あくまで「場」を提供しているプラットフォーム事業者が、過度に躊躇することなく、自らの判断で適切な削除等の対応をとるよう促されるような法的枠組みとして、免責規定は有効な手段といえる。
- ・ しかし、例えば、行政機関からの法令違反情報に関する削除要請があった際に、プラットフォーム事業者が当該要請に応じて削除を行った場合の免責の法的枠組みを設けた場合、当該要請があった情報に関して自ら違法性の判断を行わずに機械的に削除するインセンティブを生み、実際には違法情報ではないが疑わしいとされる情報が削除され、投稿の過度な削除等が行われるおそれがある。また、事実上の行政による検閲のリスクが生じかねないことなど、利用者の表現の自由に対する制約をもたらす懸念もある。
- ・ 他方で、削除した場合の免責規定を一切設けなかった場合、事業者が誤って削除したときは、発信者に対して損害賠償責任を負うことになる可能性がある。そのため、2-1で検討したとおり、行政機関からの専用窓口を設けることや、迅速な対応に関する規律を設けることとしても、依然としてプラットフォーム事業者には違法性判断について慎重な対応が求められることから、迅速な対応につながらない可能性がある。また、事業者が依然として法的に不安定な位置に置かれたままとなる可能性もある。
- ・ ただし、こうした懸念に対しては、事業者が自ら定める利用規約において削除に関するポリシーや免責規定を設けていることが通常であり、当該利用規約に沿って削除等の対応を行う限り、発信者との関係で損害賠償請求を受けるリスクは想定し得ないという指摘もある。
- ・ これらを踏まえると、行政機関から法令違反情報に関する削除要請があった場合の免責については、削除対応の迅速化を実現する必要性と、過剰な削除による表現の自由に対する制約をもたらす懸念との間のバランスを慎重に検討しつつ更なる検討を行うことが適当である。なお、検討の際には、法令違反情報への対応に関して事業者が発信者から損害賠償請求を受けた裁判例等の調査や、薬機法等のプラットフォーム事業者に対する法令違反情報への対応に

関して免責規定を置いている法令に関する実態調査を行うべきである。

2-2 適格性のある第三者からの通報への対応

- ・ プラットフォーム事業者が法令違反情報に対してより迅速に対応することを促進するための方策としては、プラットフォーム事業者に対して、行政機関からの削除要請のみならず、その他の適格性のある団体からの削除要請に対してもプラットフォーム事業者が優先的に対応することを制度的に求めているという手段も考えられる。

我が国における現行制度や対応状況

- ・ 現状、我が国において、権利侵害情報やその他違法情報に関する第三者からの削除要請については、プラットフォーム事業者に優先的な対応を義務付ける制度は存在しない³⁵。
- ・ 事業者ヒアリング（詳細は、Ⅲ参照。）では、公益性の高いその他団体からの削除要請に対するプラットフォーム事業者側の窓口整備の状況としては、専用の窓口を設けて対応している事業者もあった一方で、専用の窓口は設けておらず、削除要請を行っている実績のある団体からの削除要請であっても一般の利用者と同様に扱っている事業者も存在した。事業者が提供している優先対応スキーム³⁶を利用する機関からの申告については、優先対応を行っているとは回答した事業者もあった。
- ・ また、本 WG では、プラットフォーム事業者に法令違反情報の削除要請を行っている団体からもヒアリングを行った。ヒアリング対象としては、インターネット上の違法・有害情報について通報を受け付け、国内外を問わず、サイト管理者等に削除を依頼している IHC³⁷と、同様に事業者に対して削除依頼を実施している「セーフライン」を運用する一般社団法人セーフアーインターネット協会（SIA）³⁸を対象とした。

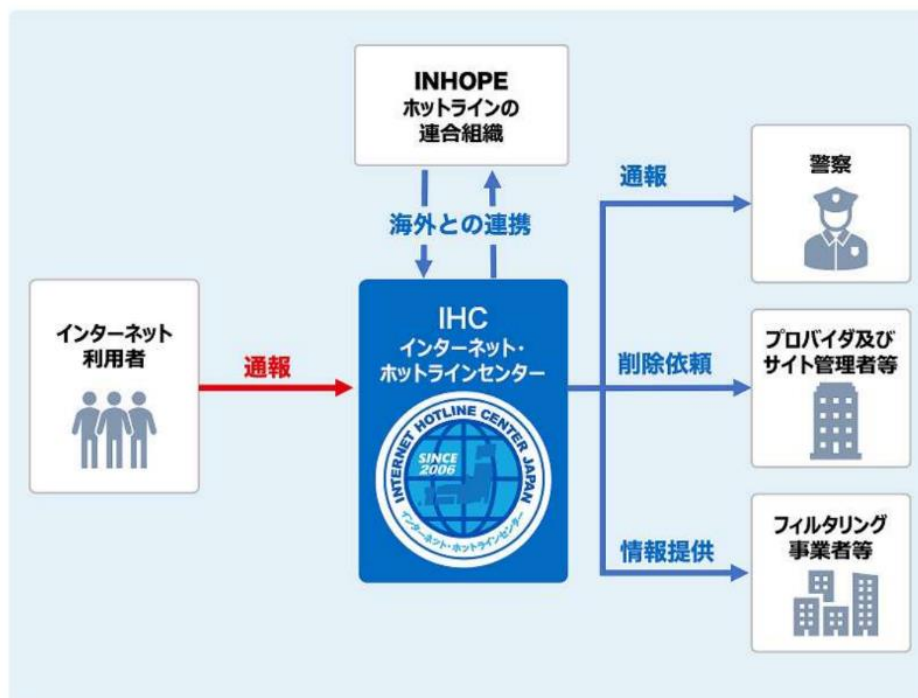
³⁵ なお、3で扱う有害情報に関してもこのような制度は存在しないため、2-2の議論は有害情報の場合にも援用できると考えられる。

³⁶ 前述の脚注 22 も参照。

³⁷ 制度 WG 第3回 資料3-2を参照のこと。

³⁸ 制度 WG 第3回 資料3-1を参照のこと。

図表 5 : IHC 説明資料（本 WG 第 3 回）



図表 6 : SIA 説明資料（本 WG 第 3 回）



- ・ IHC からのヒアリングによれば、同機関による削除依頼に対し、約 9 割は 5 営業日以内に削除されていることのことであった³⁹。一方で、SIA のセーフラ

³⁹ 本 WG 第 3 回会合における IHC 発表

インによる削除率は 14 日以内で 5 割、1 か月以内で 87%となっており⁴⁰、削除要請を行っている団体からの申出に対するプラットフォーム事業者の対応の処理に係る時間は区々であった。こうした差異が生ずる背景には、それぞれの団体が主に扱う情報の種類が異なるといった事情もあるものと考えられる。

諸外国の状況

- ・ EU (DSA) では、インターネット上の違法情報に迅速かつ信頼できる形で対応を行うため⁴¹、事業者に対して、専門的知見に基づき通報を行う者として各国のデジタルサービス・コーディネーター⁴² (DSC) が認定する「トラステッド・フラグガー」からの違法情報に関する通報を優先的に対応することを義務付けている (第 22 条)。この制度は、削除要請を受けて即削除というスキームではなく、あくまで他の要請と比較して優先的に取り扱うものである。また、「優先的に」対応というのは、必ずしも全ての情報に迅速に対応するというわけでもない。全ての通報を適時に慎重かつ非恣意的に処理し決定を下すという要件を損なわないという観点から⁴³、このトラステッド・フラグガーの条件として、違法コンテンツの検出・特定・通報に関する「専門性」・事業者からの「独立性」・判断と通報に「客観性」が求められている⁴⁴。

トラステッド・フラグガーの認定に関しては、ガイドラインが 2025 年の第二四半期に意見募集が開始され、同年末までに採択される予定⁴⁵であり、詳細

⁴⁰ 本 WG 第 3 回会合における SIA 発表

⁴¹ DSA 前文 (61) の冒頭で、「本規則で要求される仕組みに基づいて提出された**全ての**通報を適時に慎重かつ非恣意的に処理し決定を下すという要件を損なうことなく、指定の専門分野の範囲内で行為するトラステッド・フラグガーにより当該通報と対応の仕組みを用いて提出された通報を優先的に取り扱うことを保証するために必要な措置をオンラインプラットフォームプロバイダが講じる場合には、違法コンテンツへの対応はより迅速かつ信頼できる形で行うことができる。」としている。

⁴² 加盟国内における DSA の監督・執行の権限を有する組織 (ただし、VLOP/VLOSE に対する監督・執行は欧州委員会が担う)。EU 加盟国は DSA の適用および執行に責任を有する所管当局として DSC を指定する。DSC は DSA の執行の権限と責任を国内において負うとともに、欧州委員会ならびに他の加盟国との連携を行う。

⁴³ 脚注 29 参照。

⁴⁴ DSA 前文 (61) では、「・・・かかるトラステッド・フラグガーの地位を与える対象は、違法コンテンツへの対処における特別な専門知識と能力を有していること、及び慎重、正確かつ客観的な方法で作業していること等を実証した、個人ではない組織のみとすべきである。」とされている。

⁴⁵ <https://digital-strategy.ec.europa.eu/en/policies/trusted-flaggers-under-dsa>

な認定条件等は今後明らかになるものと考えられる⁴⁶。

- ・ 一方、英国（OSA）では、OSAに基づく行動規範（Codes of practice）において、事業者がトラステッド・フラグガーから詐欺コンテンツについて報告を受け付ける専用窓口を設けることが推奨されている。英国はEUと異なり、トラステッド・フラグガーとして、警察、労働年金省、国家犯罪庁等の行政機関が指定されている。

具体的検討

- ・ 本WGでは、プラットフォーム事業者が法令違反情報に対してより迅速に対応することを促進するため、諸外国の取組を参考として、「トラステッド・フラグガー制度」類似の制度を我が国においても導入し、プラットフォーム事業者に対して、適格性のある「第三者」からの通報を優先的に対応することを制度的に求めていく必要があるかについて検討を行った。
- ・ まず、我が国では、サイト管理者等に違法情報の削除依頼を行っているIHCによる削除依頼の削除率は5営業日以内で約9割、SIAのセーフラインによる通報の削除率は14日以内で5割、1か月以内で87%となっている。また、一部のプラットフォーム事業者は、これらの通報団体向けの専用窓口を設けて自主的に優先対応を行っている。
- ・ 上記の現状を踏まえると、我が国において、法令違反情報への対応として、トラステッド・フラグガー制度類似の制度、すなわち適格性のある第三者からの通報制度を直ちに導入する必要性は必ずしも高くないものと考えられる。
- ・ ただし、あくまで上記の状況は現状評価に過ぎないため、当該制度の必要性等については、今後も継続的にIHCやSIA等の通報団体の削除実績を注視し、実態把握を続けていくことが必要である。また、実態把握の際には、通報団体ごとの削除率やプラットフォーム事業者における対応の迅速性の差異の要因分析も併せて行っていくことが望ましい。

⁴⁶ なお、ガイドライン策定前に、先行してトラステッド・フラグガーを認定している国（フィンランド、オーストリア、スウェーデン、デンマーク、ルーマニア、ドイツ、フランス、リトアニア、ギリシャ、ハンガリー、マルタ及びイタリア）もあり、これらの国では、NPO、NGO、業界団体、労働組合、ファクトチェッカー等が認定されている。

- ・ その上で、仮に今後削除率の低下や迅速な対応が図られなくなるような状況が見られた場合には、それらの原因を深掘りするとともに、改めてトラステッド・フラグガー制度類似の制度の導入の必要性を検討することが適当である⁴⁷。
- ・ また、本 WG では、仮に制度を導入する場合、具体的な通報主体としてはどのような団体が適切か検討を行った。
- ・ 特に、プラットフォーム事業者においては、発信者の表現の自由を侵害しない限りで、違法情報に対応していくことが求められる立場であるため、1－2 や 2－1 での検討と同様、削除要請を行う団体としては、法令違反情報の違法性の判断に関して専門的知見を持っており、その「専門性」を持って行われる判断と通報に一定の信頼性があること・恣意性がないことが最低限求められる。
- ・ さらに、安定的・持続的な運営を図るため、適格性のある通報団体について、財政的基盤や人的リソースをどのようにして確保すべきかについても重要な課題である⁴⁸。
- ・ したがって、これらの「専門性」「独立性」「客観性」「財産基盤」といった要素を複合的に踏まえて、適格性のある通報団体の要件を引き続き深掘りしていくことが望ましい⁴⁹。
- ・ なお、ヒアリング対象事業者の一部は、既に通報団体向けの専門窓口の整備や優先対応を実施している。今後の制度的検討と並行して、こうした取組を実施していない事業者も含め、例えば、事業者自らが特定の通報団体を認定し、当該行政機関からの通報を専用窓口で受け付け、通報を信頼性が高いものとして優先的に取り扱うプロセスを確立するなど、法令違反情報への対応を迅速に行うための枠組みを自ら整備して対応することも期待される。

⁴⁷ なお、構成員からは、プラットフォーム事業者においては判断の難しさがあっても違法有害情報に対応していく社会的責任があるという立場から、表現の自由に関する官民関係を「制度化」していく意義はあるとする意見もあった。

⁴⁸ IHC は警察庁の委託事業となっている。一方、SIA は民間企業によって設立された団体であり、会員企業の入会金や年会費によって運営されている。

⁴⁹ なお、構成員からは、この点、既存の機関も含め、トラステッド・フラグガー制度の下に包含すれば、ガバナンスが効くのではないかとの意見や、トラステッド・フラグガーを適切に選定しないと誤った通報になりかねないとの意見等があった。

3 情報レイヤーに関する検討③：有害情報への対応の在り方

- ・ 「有害情報」とは、例えば、青少年有害情報（青少年の健全な成長を著しく阻害する情報（暴力、アダルト等））のほか、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、存在しない災害・事故・事件が存在するかのように見せかけた偽画像・偽動画、災害発生時に外国人が犯罪行為を行っているとする偽・誤情報といったものも含まれる、非常に広範な概念である。
- ・ 健全性検討会のとりまとめでは、有害性や社会的影響の重大性が大きい偽・誤情報に対する事業者の取組を促す観点から、これらの情報に対し、「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、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」との提言がなされた。

我が国における現行制度や対応状況

- ・ 有害情報は、一般に、人によって有害かどうかの判断が異なり得るだけでなく、有害か否かの線引き・判断に困難が伴うものが含まれる。この点、事業者は利用規約に基づいて、自らの判断で有害情報を削除することができるが、事業者による対応が十分ではない場合の制度的対応が行き過ぎれば、表現の自由との関係で懸念が生じることにもなる。

諸外国の状況

- ・ EU（DSA）では、有害情報について、削除対応を事業者を求める規律はない。一方、VLOP/VLOSE⁵⁰に対して、違法情報だけでなく有害情報についても、サービスの設計・機能・アルゴリズムシステムの影響、又はそのサービスの利用に起因するシステムリスクを特定・分析・評価することを義務付けており（第 34 条）、そのリスク評価を踏まえた軽減措置の実施等を義務付けている（第 35 条）。
- ・ 英国（OSA）でも、有害情報について、削除等の直接的な対応を事業者に求

⁵⁰ VLOP：大規模オンラインプラットフォーム事業者、VLOSE：大規模オンライン検索エンジン事業者

める規律はないが、こども（18 歳未満）にとって有害な情報に関しては、リスク評価・安全措置の実施等を義務付けている（第 12 条、第 13 条）。

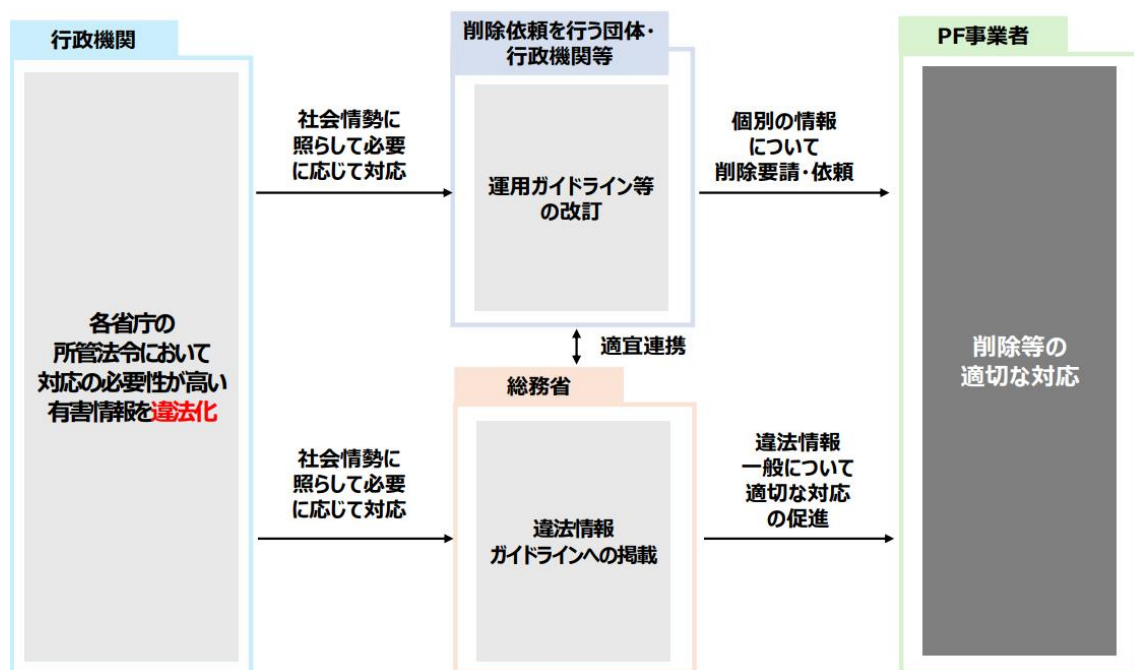
具体的検討

- ・ 有害情報に関しては、違法ではない情報であり、個別の情報の削除を事業者を求めることは、表現の自由の観点から極めて慎重な検討を要する。本 WG では、こうした基本的な認識に立って、有害情報への対応の在り方について検討を行った。
- ・ この点、本 WG では以下の 2 つの対応の方向性が示された⁵¹。
 - ① 「有害性や社会的影響の重大性が大きい」有害情報について、表現の自由への制約が大きい削除以外のコンテンツモデレーションを促進して対応する方向性
 - ② 法益侵害を発生させ、又は惹起が確実なものとして社会のコンセンサスが得られる情報について、当該法益を保護する個別法において違法化した上で、削除等対応を促進して対応する方向性
- ・ ①については、削除以外のコンテンツモデレーションとしては、警告表示、収益化停止、ラベル付与、表示順位低下などが想定されるが、これらの対応については 4 において後述する。
- ・ ②については、ある有害情報について、法益侵害を発生させ、または惹起が確実な情報としての社会的コンセンサスが得られた結果、違法情報と位置付けられることになれば、違法である範囲は明確化されることとなる。社会情勢等の必要性を踏まえた上で、各省庁が所管する個別法において表現の自由にも配慮しつつ、当該情報が違法であることを明確化したり、新たに違法化されるのであれば、事業者による削除等の適切な対応が図られると考えられる。
- ・ 一般に、法令によって情報の発信や流通を新たに違法化することは、表現の内容規制そのものであるため、十分な実態把握と国民の理解が欠かせないものと考えられる。

⁵¹ なお、本 WG では、有害情報の発信の繰り返し行為について何らかの対応ができないか検討されたが、有害情報について削除等を制度的に求めることは表現の自由の観点から困難であることから、結局、方向性①又は②で対応することとなるとの結論となった。

- ・ 仮に違法情報と位置付けられれば、削除依頼を行う団体や行政機関等の運用ガイドラインに、当該違法情報に関する判断基準等を掲載することにより、プラットフォーム事業者に対する削除依頼の実施が適切に行われると考えられる。また、違法情報ガイドラインに、当該違法情報を反映することにより、必要に応じて、プラットフォーム事業者が自らの削除基準に当該情報を盛り込むことで、プラットフォーム事業者による適切な対応の促進が図られるものと考えられる。

図表 7：違法化された場合における対応のイメージ



4 サービス設計レイヤーに関する検討：違法情報・有害情報共通の対応の在り方

- ・ 本 WG では、情報流通の場となるプラットフォームのサービス設計に関する制度的対応の在り方について、事業者による取組や諸外国における政策動向も参考としつつ検討を行った。
- ・ SNS 等のプラットフォームのサービス設計については、提供されるサービスの内容（利用者同士の投稿を通じた交流の方法、利用者に対する広告表示の有無等）、インターフェース（投稿の表示場所など）、利用者にどの情報を表示するかを決定するアルゴリズムなど様々なものがあるが、利用者との関係では、特に以下の設計・機能が重要である。
 - 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、また、手軽に情報を入手できる設計（**情報発信・受信に係る設計**）、
 - 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、経済的インセンティブ設計（**拡散促進機能等**）、
 - ③ 情報の受信者の興味・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合わせる機能（**レコメンデーション機能等**）
- ・ 上記のサービス設計は、利用者の利便性の向上に資するものであると考えられるが、同時に、違法・有害情報の流通・拡散を容易にし、また、利用者が触れる情報に偏りを生じさせるリスクも内包するものであり、サービスを提供するプラットフォーム事業者自身がサービス設計・提供の当事者として、こうした課題への適切な対応について責任を果たすことが求められる。
- ・ これらのリスクへの対応（リスク軽減措置）について、
 - 1) **情報発信・受信に係る設計**がもたらすリスクへの対応としては、①不適切な内容を投稿しようとする利用者に対してサービス上（画面上）で警告を表示することで当該投稿の発信を抑止する「警告表示」、②信頼できる情報を受信者側のおすすめ欄に優先的に表示することで信頼できる情報の流通の促進を図る「プロミネンス」、③投稿される画像・映像等が AI 生成物である場合にその旨を明示することで情報の信頼性等を見分けることを可能にする「ラベル付与」などの方法が考えられる。

- 2) **拡散促進機能等**がもたらすリスクへの対応としては、コンテンツやコンテンツに付随して表示される広告の閲覧数等に応じて、当該コンテンツの発信者に対し広告収入が分配される仕組みを変更する方法などが考えられる。
- 3) **レコメンデーション機能**がもたらすリスクへの対応としては、利用者に「おすすめ」するコンテンツを優先的に配信する「レコメンダシステム」のアルゴリズムを違法・有害情報の流通・拡散を助長しないように変更する方法などが考えられる。
- ・ 本 WG では、上記の例示を含むリスクへの対応について、プラットフォーム事業者に対し、どのように実施を促すべきか制度面からの議論・検討を行った。リスクへの制度的対応の方向性については、①法令等によりリスクへの対応の確実な実施や透明性の確保を図りつつも、その具体的な内容については、プラットフォーム事業者委ねる領域と、②法令等によりプラットフォーム事業者が実施すべき対応の具体的な内容も含む一定の規範を設ける領域があると考えられるが、その切分けのための考え方として、以下に示す3つの基本的な考え方を設定⁵²し、議論・検討を行った。
- (1) プラットフォーム事業者が提供するサービスの在り方や流通する情報の種類、ユーザー層などのサービス特性は様々であり、サービスから発生し得るリスクも様々である。そのため、サービスを最もよく知るプラットフォーム事業者が、サービス上で流通する違法・有害情報の状況やサービス設計に起因するリスクや社会的影響について自ら評価し、サービス設計上、リスクに応じた対応を自ら実施することが重要である。【違法・有害情報の流通状況とサービス設計がもたらすリスクの理解・リスクへの適切な対応】
- (2) インターネット上における情報の流通過程において、利用者は情報の発信者でもあり受信者でもあるが、プラットフォームのサービス上、今見ている情報がなぜ表示されているのか、プラットフォーム事業者による情報表示の考え方について、利用者の間で正確な理解が必ずしも十分に広がっているわけではない⁵³。したがって、プラットフォーム事業者は、情報表示

⁵² なお、今後の検討において別の視点から議論を行うことを排除するものではない。

⁵³ 本 WG 第 9 回における高口構成員発表資料参照。

や自らのサービスの仕組みについて更に透明性を高め、利用者に伝えることで、利用者に自由な情報摂取の機会を保障する環境を整備することが重要である。【適切な情報表示の在り方】

- (3) 偽・誤情報や誹謗中傷の投稿の発信を抑止するための方策として、また、犯罪捜査の観点からの発信者のトレーサビリティを確保するための方策として、アカウント開設時の本人確認等が考えられるところ、匿名表現の自由との関係も踏まえつつ、対応の在り方を整理することが重要である。【適切な利用者の確認の在り方】

4-1 違法・有害情報が流通するリスクとリスクへの対応の在り方

4-1-1 リスク評価・軽減措置に関する対応の在り方

- ・ 健全性検討会のとりまとめでは、「情報伝送プラットフォーム事業者は、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・実施することが適当」との提言がなされた⁵⁴。

我が国における現行制度や対応状況

- ・ 現行の情報流通プラットフォーム対処法では、大規模プラットフォーム事業者に対し、権利侵害情報に関する削除申出を行うための方法、権利侵害情報の削除の実施状況、送信防止措置を講ずるための人的・技術的体制の整備、削除基準の内容、削除基準により送信防止措置の対象となる情報の流通状況（違法・有害情報の流通状況）などの事項について、自ら行った評価⁵⁵（以下「自

⁵⁴ DSA 第 35 条でリスク軽減措置の具体例として掲げられている措置にはリスク評価が前置されているが、例えば、レコメンダシステムについては、透明性の確保の規律（第 27 条）がリスク評価を前置しない形で規定されている。本 WG では、このようにリスク評価を前提としない規律を含めて「リスク軽減措置」という用語を用いることとする。

⁵⁵ 情報流通プラットフォーム対処法における「自己評価」には、違法・有害情報の流通

己評価」という。)を毎年一回公表することを義務付けている⁵⁶(第28条第5号)。また、事業者は、自己評価を実施するに当たっては、自己評価の基準を定めて行うものとされている(総務省令第18条第6項)。

なお、法令上は事業者による自己評価の公表のみが求められているものであり、自己評価の結果を踏まえた対応の在り方については、上記の設計・機能の変更・見直しの適否やその内容も含め、事業者の自主的な判断に委ねられている。

- ・ また、自己による評価の他の例としては、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(令和2年法律第38号。以下「取引透明化法」という。)がある。同法においては、特定デジタルプラットフォームの更なる透明性及び公正性の向上に向けた自主的な改善のために、特定デジタルプラットフォーム提供者⁵⁷に対して、経済産業大臣に提出する報告書において、苦情の処理及び紛争の解決に関する事項、取引条件等の開示の状況に関する事項や相互理解の促進に向けた自主的な手続・体制の整備に関する事項について、自ら行った評価を含めることを義務付けている(第9条第1項第5号)。

諸外国の状況

- ・ EU(DSA)では、VLOP/VLOSEに対して、違法・有害コンテンツについて、サ

状況が評価項目として含まれており、EU(DSA)や英国(OSA)におけるリスク評価の評価項目と一部重なる部分もあると考えられる。一方、自己評価は、事業者が実施した送信防止措置を前提として、毎年1回、前年度の状況を事後的に評価するものであるため、重大な影響を及ぼす可能性のある機能を展開する前にも評価を実施するEU(DSA)や英国(OSA)とは、評価する段階や対象が異なる。

⁵⁶ 自己評価の公表は、議員修正により規定されたものであり、その提案理由は、「SNS等において深刻化する権利侵害等に対処するため、大規模特定電気通信役務提供者による削除等の運用状況について、より一層の透明化を図る必要。そこで、本修正では、大規模特定電気通信役務提供者が毎年一回公表しなければならない事項として、送信防止措置の実施状況及び当該実施状況について自ら行った評価を明記することとし(た)」とされた。

⁵⁷ デジタルプラットフォームのうち特に取引の透明性・公正性を高める必要性の高いプラットフォームを提供する事業者を「特定デジタルプラットフォーム提供者」として指定して規律の対象としている。現時点では、本WGのヒアリング対象事業者のうち、LINEヤフー、Google、Metaが指定されており、そのほか、アマゾンジャパン合同会社、楽天グループ株式会社等が指定されている。

(https://www.meti.go.jp/policy/mono_info_service/digitalplatform/provider.html)。

ービスの設計・機能・アルゴリズムシステムの影響又はそのサービスの利用に起因するシステムリスクを特定・分析・評価することが義務付けられている（第 34 条）。また、リスク評価の具体的な実施内容は、法律上で定められておらず、事業者委ねられている。

その上で、リスク評価を踏まえた軽減措置を行うことが義務付けられており（第 35 条）、法律上、「オンライン・インターフェースも含むサービス設計の適合」、「利用規約の変更」、「コンテンツモデレーションの改善」、「アルゴリズムの改善」、「生成されたコンテンツの表示」等が実施すべき措置事項として例示されている（図表 8）。その具体的な内容は、欧州委員会が策定する個別のガイドラインにおいて、一定程度の明確化が図られている。

- ・ 2018 年 10 月、欧州では、オンラインプラットフォーム、大手テック企業、広告業界の代表がオンライン上の偽情報の拡散に対処するための自主規制の枠組みとして行動規範（Code of Practice on Disinformation）を策定した。欧州の行動規範は、リスク軽減のための取組を記載したコミットメント集であり、複数事業者によって策定され、各事業者がコミット可能な取組を選択した上で署名をすることで拘束力を持つ枠組みとなっている。DSA に基づき認定された行動規範を遵守している場合は、DSA 上のリスク軽減措置の履行義務を遵守しているかどうかの判断材料の一つとなるとされており（第 35 条第 1 項 (h)）、その意味で法的な枠組みの下で策定・署名されるものである。

なお、この偽情報に関する行動規範は、2025 年 2 月に DSA に基づく行動規範として認定された。その中では、9 つの分野⁵⁸（例えば、①広告表示の精査（偽情報の収益化廃止等）、②サービスの完全性確保（なりすまし、ボットなどの許されない行動の共通理解等）、③ユーザーのメディア・リテラシーの向上、④研究者コミュニティ支援、⑤ファクトチェッカー支援、⑥モニタリング体制の強化（欧州委員会に対する報告等））で 44 のコミットメントが記載され、2025 年 3 月時点では、プラットフォーム事業者を含む 42 団体によって署名されている。

- ・ また、リスク評価及び軽減措置の遵守状況については、監査法人等の第三者機関による監査を受けることが事業者義務付けられており、監査での指摘への対応や、リスク評価結果・軽減措置について各国の DSC 及び欧州委員会に報告を行うことも義務付けられている（第 37 条、第 42 条）。

⁵⁸ 例で掲げた分野以外には、政治広告への対応、透明性センターへの協力、常設タスクフォースへの協力など、本 WG の対象外又は EU 独自のスキーム関係の分野がある。

- ・ 英国（OSA）では、全てのユーザー間サービス⁵⁹の提供事業者に対して、違法コンテンツとの遭遇、優先犯罪(priority offence)の実行又は助長への使用、個人に与える危害等についてのリスクレベルを評価することが義務付けられている（第9条）。その内容は、Ofcom に対して、ユーザー間サービスなどがもたらす危険を評価し、評価結果をリスクごとに、サービスのリスクプロファイル（危険度などの特徴）を作成・公表を義務付けており（第98条）、2024年12月に公表されている⁶⁰。

その上で、ユーザー間サービスや検索サービスを提供する事業者に対して、リスク評価を踏まえた措置を行うことが義務付けられている（第10条。具体的な安全義務については、図表9参照）。その内容は、Ofcom に対して、プラットフォーム事業者に課される義務に関する行動規範の公表が義務付けられており（第41条）、違法コンテンツに対する安全義務の詳細を記載した行動規範（Codes of Practice）⁶¹が2024年12月に公表された。

リスク評価及び安全義務の遵守状況について、ユーザー間サービスの提供事業者は、Ofcom に対して、透明性レポートを提出することが義務付けられている（第77条）。

- ・ 上記のとおり、DSA と OSA においては、対象事業者（大規模事業者に限定するか否か）、リスク評価の具体的内容（事業者に委ねるか、政府において具体化するか）、遵守の担保方法（監査を介するか否か）といった点で異なっている。

図表8：DSAにおけるリスク軽減措置の一覧

⁵⁹ ユーザー間サービスとは、ユーザーが作成・共有したコンテンツを、他のユーザーが閲覧できるサービスをいう（OSA 第3条第1項）。

⁶⁰ <https://www.ofcom.org.uk/siteassets/resources/documents/online-safety/information-for-industry/illegal-harms/risk-assessment-guidance-and-risk-profiles.pdf?v=390984>

⁶¹ 行動規範（Codes of practice）に明記された措置の内容としては、「優先的犯罪の実行又は助長に使用されるリスクの効果的な軽減・管理」や「個人への害のリスクの効果的に軽減・管理」など。

<p>【サービス設計の改善】 サービス設計、特徴、機能を（リスク評価結果に）適合させる（オンラインインターフェースも含む）（35条1項(a)）</p>
<p>【利用規約の改善】 利用規約と執行を（リスク評価結果に）適合させる（35条1項(b)）</p>
<p>【コンテンツモデレーションの改善】 コンテンツモデレーションプロセスを（リスク評価結果に）適合させる（特定の種類の違法コンテンツの通知の処理スピードや品質、適切な場合、違法ヘイトスピーチやサイバー暴力の観点で通知されたコンテンツの迅速な削除やアクセス禁止を含む。また、コンテンツモデレーションのための意思決定プロセスや専属リソースも（リスク評価結果に）適合させる。（35条1項(c)）</p>
<p>【アルゴリズムの改善】 アルゴリズムシステム（レコメンダシステムを含む）をテストし、（リスク評価結果に）適合させる。（35条1項(d)）</p>
<p>【広告システムの改善】 広告システムを（リスク評価結果に）適合させ、提供するサービスに関連した広告の表示を制限又は調整することを目的とした、的を絞った対策を採用する（35条1項(e)）</p>
<p>【内部プロセスの強化】 特にシステミックリスクの検知に関する、活動の内部プロセス、リソース、テスト、文書化又は監督を強化する（35条1項(f)）</p>
<p>【第三者機関との協力】 第22条に基づくトラステッド・フラグガーとの協力及び第21条に従った裁判外紛争処理機関の決定の実行を開始又は調整する（35条1項(g)）</p>
<p>【他の事業者との協力】 第45条及び第48条それぞれに基づき、行動規範や危機プロトコルを通じて他のオンラインプラットフォーム事業者やオンライン検索エンジンとの協力を開始又は調整する（35条1項(h)）</p>
<p>【意識啓発対策】 サービス利用者により情報を与えるため、意識啓発対策を措置し、オンラインインターフェースを（リスク評価結果に）適合させる（35条1項(i)）</p>
<p>【児童の保護】 年齢認証（age verification）やペアレンタルコントロールツール、適切な場合、未成年者が虐待を通報したり支援を得ることを支援することを含む、児童の権利を守るための的を絞った対策を措置すること（35条1項(j)）</p>
<p>【生成コンテンツの表示等】 情報がオンラインインターフェースに表示された際、既存の人物、物、場所、その他の実体または出来事に著しく類似し、真正または真実であるかのように見せかける、生成または操作された画像、音声または映像を含むかどうかを、目立つマーク（prominent markings）によって区別できるようにし、さらに、サービス利用者がそのような情報を示すことができる使いやすい機能を提供すること（35条1項(k)）</p>

図表 9：OSA における安全義務の一覧

■ ユーザー間サービスや検索サービスを提供する事業者に対して「違法コンテンツに関する安全義務」を義務付け（10条）。

措置内容	条文
サービスの設計または運営に関連する以下の措置を講じる義務	10条 2 項
<ul style="list-style-type: none"> 優先的違法コンテンツに個人が接触することを防止 優先的犯罪の実行または支援に使用されるリスクを、最新の違法コンテンツリスク評価に基づき効果的に軽減・管理 最新の違法コンテンツリスク評価で特定された個人への害のリスクを効果的に軽減・管理 	(a)～(c)
サービスを次のように運営する義務	10条 3 項
<ul style="list-style-type: none"> 優先的違法コンテンツが存在する期間を最小限に抑えるために適切なシステムおよびプロセスを利用してサービスを運用 違法コンテンツの存在を通知された場合、または存在を認識した場合、そのコンテンツを迅速に削除するようサービスを運用 	
上記の義務（10条2項・3項）について、以下の分野において措置を組み込む義務	10条 4 項
<ul style="list-style-type: none"> リスクの管理 機能・アルゴリズム設計 利用規約 特定コンテンツのアクセス管理 コンテンツモデレーション ユーザーサポート スタッフの方針と実務 	(a)～(h)
以下をサービスの利用規約に含め、適用する義務	10条5項、6項
<ul style="list-style-type: none"> 違法コンテンツから個人を保護する方法 	(a)
プロアクティブ技術に関する情報（技術の種類、使用タイミング、機能の仕組みを含む）をサービスの利用規約に含める義務	10条 7 項
利用規約の規定が明確で、アクセスしやすいことを確保する義務	10条 8 項
リスク評価の結果をサービス利用規約に要約し、利用規約に記載する義務(大規模なユーザー間サービス事業者（カテゴリー 1 サービス）のみ義務）	10条 9 項

図表 10：欧州 DSA、英国 OSA のリスク評価、情報流通プラットフォーム対処法の自己評価について

	対象情報	リスク評価項目		評価時期
		情報レイヤー	サービス設計レイヤー	
欧州DSA 第34条	<ul style="list-style-type: none"> 違法情報（加盟国において違法な情報） 有害情報（偽情報、未成年者への悪影響等） 	<ul style="list-style-type: none"> 違法コンテンツの流通・拡散 基本的権利(人間の尊厳、個人情報保護、メディアの自由と多元性等)の行使に対する悪影響 市民的言論や選挙プロセス、公共の安全に対する悪影響 ジェンダーに基づく暴力、公衆衛生・未成年者保護、個人の身体的・精神的健康に対する悪影響 	<ul style="list-style-type: none"> サービス設計（アルゴリズム、コンテンツモデレーションシステム、利用規約、広告表示、データ運用等）によって生じるリスク ユーザーの利用方法（botや偽アカウントを用いた不正利用等）によって生じるリスク 	<ul style="list-style-type: none"> VLOP/MLOSEの適用日まで（その後は少なくとも年1回） リスクに重大な影響を及ぼす可能性のある機能を展開する前
英国OSA 第9条、第11条	<ul style="list-style-type: none"> 違法情報（あらゆる種類の違法コンテンツ） 子どもに対する有害情報（ポルノコンテンツ、自殺誘引、摂食障害誘引 等） 	<ul style="list-style-type: none"> 優先的違法コンテンツへの接触可能性 優先的犯罪の実行等への使用可能性 等 子どもに対する有害なコンテンツへの接触可能性 子どもに対する有害なコンテンツによって生じるリスク 等 	<ul style="list-style-type: none"> 違法コンテンツの存在や拡散を容易にするサービスの機能性に係るリスク サービスの設計及び運用（ビジネスモデル、ガバナンス、プロアクティブ・テクノロジーの利用、利用者のメディアリテラシーを促進する措置及びサービスの安全な利用を促進するための措置、その他のシステム及びプロセスを含む）が、特定されたリスクをどのように軽減または増大させるか 等 	<ul style="list-style-type: none"> サービス開始（※）から3ヶ月以内 サービスの設計・運用に重大な変更をするとき 等

※ Ofcomが策定するリスク評価指針の公表時、すでに対象となるサービスがあるときは、リスク評価基準の公表日から3か月以内に最初のリスク評価を行う（指針の公表が2024年12月16日であったので、2025年3月16日まで）。

	対象情報	自己評価項目		評価時期
		情報レイヤー	サービス設計レイヤー	
情プラ法 第28条5号	違法・有害情報	違法・有害情報の流通状況	技術的体制の整備状況	年1回

※ このほかに、情プラ法においては、削除申出の方法、削除等の実施状況、削除しない場合等の理由の粒度、人的体制の整備状況、削除基準の内容についても自ら行った評価結果を公表することを義務付け。

具体的検討

(1) 現行制度に基づく自己評価への対応の在り方

- ・ 本 WG では、健全性検討会のとりまとめの提言を踏まえ、プラットフォーム事業者が「サービスアーキテクチャ」や「ビジネスモデル」がもたらす「社会的影響」をどのように理解・予測し、適切な対応をとることを促すべきかについて、検討を行った。
- ・ プラットフォームのサービス設計については、プラットフォーム事業者がサービスの設計・提供の当事者としての知見と対応能力を持っているところ、違法・有害情報の流通・拡散から利用者を保護するため、①プラットフォーム上の違法・有害情報の流通状況、②サービスがもたらすリスクを把握・分析・評価し、リスクを踏まえたサービス設計の見直し・改善によって適切に対応していくことが適当である。これは、健全性検討会で提言されたプラットフォーム事業者の役割・責務の一つでもある。
- ・ まず、①プラットフォーム上の違法・有害情報の流通状況の評価については、上述のとおり、現行の情報流通プラットフォーム対処法において、プラットフォーム事業者が自己評価を行った結果を公表することとされている。
- ・ 次に、②サービスがもたらすリスクについては、図表 10 にもあるとおり、現行の情報流通プラットフォーム対処法における自己評価の評価項目にはサービスがもたらすリスクそのものを評価する項目はないものの、例えば、送信防止措置を講ずるための人的・技術的体制の整備について自己評価を求めている、サービス設計レイヤーに関する評価項目が設けられている。
- ・ 総務省は、今後、プラットフォーム事業者が実施する自己評価について、実施プロセス・方法、評価結果の内容・粒度等を含め、適切なモニタリング等を通じて注視していくことが適当である。
- ・ なお、プラットフォーム事業者は、「技術的体制の整備」に関する自己評価の一環として、災害時など特定の場面において、ボット⁶²を使った違法情報や有害情報を繰り返し発信する行為もあることから、ボット対策についても評価を行うことが適当である。

⁶²「ボット」とは、特定の作業を自動で行うコンピュータープログラムのことをいう。

(2) リスク評価の実施を促す制度的な対応の在り方

- ・ 本 WG では、プラットフォーム事業者に対してリスク評価の実施を制度的に促す場合の考え方についても検討を行った。
- ・ プラットフォーム事業者によるリスク評価の実施を促す制度的対応の在り方については、リスク評価の結果を踏まえてプラットフォーム事業者が何らかの対策をとることを前提とする場合、①リスク評価が必要な状況、②リスク評価が必要な利用者層、③リスク評価を行うべき理由等について整理し、プラットフォーム事業者が具体的に何をどのような観点から評価すればよいのかが明確になっていることが必要である。これらの点や諸外国の政策動向を踏まえると、リスク評価制度の在り方としては、以下の2つの方向性があると考えられる。

<方向性A>

- ・ リスク評価の義務内容を法律で詳細かつ明確に規定する方向性

※ 我が国の一般的な法制のように、一定の主体に対して法律上の義務を課す場合には、その義務内容が具体的かつ明確であり、違反した場合には行政による執行が可能な規律であることが必要。

※ 例えば、英国（OSA）のように、英国通信庁（Ofcom）が法令に基づき策定する行動規範（Codes of Practice）においてリスク評価の具体的な内容を明記するような仕組みが、この考え方と一定の親和性があるものと考えられる。

<方向性B>

- ・ リスク評価について、具体的な内容は法律に規定せず、プラットフォーム事業者に評価の方法は委ねるが、例えばリスク評価が適切に行われない場合に、国が適切に行うよう促す仕組みを設ける方向性

※ 技術やサービスの移り変わりは激しく、サービスの在り方によって様々なリスクや社会的影響が発生し得るプラットフォームサービスに対しては、全ての事業者にとって一定のルールを一律に設定しては、新たに発生するリスクに制度的に迅速に対応することが困難である。リスク評価の内容は事業者

委ねつつ、それらの対応に一定の実効力を持たせるような共同規制⁶³的枠組みである。例えば、EU（DSA）のように、様々なリスク軽減措置について規定しているが、リスク評価自体の在り方はプラットフォーム事業者に委ねるような形式の制度⁶⁴が、この考え方と一定の親和性があるものと考えられる。

- ・ 総務省においては、（１）のモニタリングと並行して、DSA や OSA 等の諸外国における法律の執行状況も注視しつつ、リスク評価を導入するとした場合、どちらの方向性をとるべきかも含め、引き続き検討を深めていくことが適当である。

（３）リスク軽減措置の実施を促す制度的な対応の在り方

- ・ 「４」の冒頭で述べたように、プラットフォーム事業者によって提供される SNS 等のサービスがもたらす様々なリスクへの制度的対応については、①法令等によりリスクへの対応の確実な実施や透明性の確保を図りつつも、その具体的な内容については、プラットフォーム事業者に委ねる領域と、②法令等によりプラットフォーム事業者が実施すべき対応の具体的な内容も含む一定の規範を設ける領域があると考えられる。
- ・ 後述する「収益化停止措置」「レコメンダシステムに関する対応」「プロモーションに関する対応」「AI 生成物のラベル付与」等を含め、個別のリスク軽減措置がいずれに分類されるかは、プラットフォーム事業者が提供する個々のサービス設計に内在するリスクの種類・大小に加え、一定の規範の必要性（更なる利用者保護の必要性を含む。）・有効性・明確性・執行可能性・事業者の対応

⁶³ 生貝直人『情報社会と共同規制 インターネット上政策の国際比較制度研究』（勁草書房、22-23 頁）によると、2003 年の欧州議会・欧州理事会・欧州委員会の期間協定において、共同規制とは「立法機関によって定義された目的の達成を、その分野で活動する主体（経済的主体や社会的パートナー、NGO や共同体などを含む）に委ねる法的措置のメカニズム」と定義されている、と解説されている。

⁶⁴ 欧州委員会の DSA 執行担当である DG CONNECT によれば、プラットフォーム事業者のシステムの変化や技術の進歩に対応できるように、DSA 上はプラットフォーム事業者が取るべき措置を具体的に規定してないとのことである。また、欧州委員会がリスク評価・軽減措置義務の違反を評価する際には、それぞれの事業者の対応を比較し、不足している部分が無いかを確認しているとのことである（例えば、5つのプラットフォーム事業者を調査し、そのうち4つのプラットフォーム事業者において最先端の対策と考えられるものを実施しているが、残りの1つが実施していない場合、そのプラットフォーム事業者は違反している可能性がある」と判断する。）。

(https://www.soumu.go.jp/main_content/001002942.pdf)

可能性等を考慮して判断されるべきものであると考えられる。

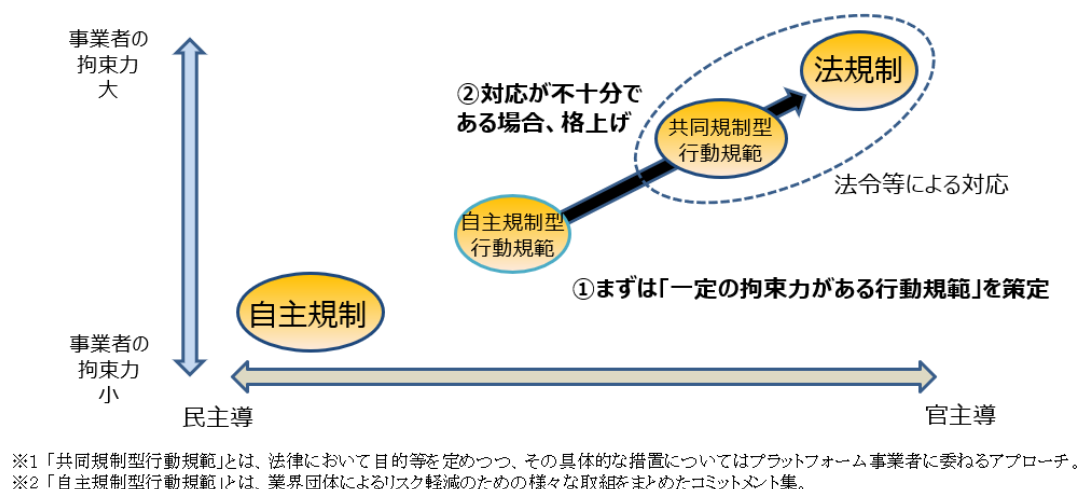
- ・ なお、上記①のような場合のリスクへの具体的な対応の在り方としては、
 - ・ プラットフォーム事業者に対し、リスク軽減のための様々な取組をまとめた「共同規制型行動規範」⁶⁵の策定を含め、リスク軽減措置の実施を法令等により求めつつ、その具体的な内容の検討・実施を事業者委ねること
 - ・ その上で、当該措置の実施状況について、法令等により、事業者自身による公表を求め透明性の確保を図るとともに、総務省がモニタリングを実施し、必要に応じ改善を求めることなどが考えられる。
- ・ この点、プラットフォーム事業者ごとにサービスの内容は様々であり、当該サービスに具備される機能がもたらす様々なリスクへの対応はサービスを設計する事業者自身が実施すべきものであること、一方、Ⅲで述べたように、諸外国で法令に基づき実施されるリスク評価・軽減措置について、過半のヒアリング対象事業者から作業負担への懸念等が示されたことを踏まえれば、まずは、業界団体が「一定の拘束力がある行動規範」(以下「自主規制型行動規範」という。)の策定を通じて、リスク軽減措置に積極的に取り組むことが期待される。
- ・ 「自主規制型行動規範」としては、例えば、SNS等のサービスがもたらすリスクへの対応として、業界団体がリスク軽減のための様々な取組をまとめたコミットメント集を策定し⁶⁶、各事業者がその中から実施する取組を選択し、実施を約束することが考えられる。総務省としては、この自主規制型行動規範実施状況について、定期的にモニタリングするべきである。
- ・ この「自主規制型行動規範」は速やかに策定されるべきであり、総務省においては、同枠組みの年内の策定に向けて、策定段階から積極的に支援等を行うべきである。
- ・ なお、行動規範策定後も、事業者による対応が不十分である場合には、政府

⁶⁵ 「共同規制型行動規範」とは、法律において目的等を定めつつ、その具体的な内容については、プラットフォーム事業者委ねるという共同規制的なアプローチを想定している。

⁶⁶ コミットメントを行う分野については、欧州の行動規範のように、ファクトチェックの支援・連携や利用者のリテラシー向上なども含まれ得る。

に事業者の取組状況をモニタリングするための新たな執行権限を設けること
や、リスク評価・軽減措置の実施を制度上明確に位置付けることなども含め、
更なる制度的対応を検討することが適当である。

図表 11：各種制度的対応のイメージ



（４）研究者に対するデータ提供の在り方

- ・ SNS 等における違法・有害情報の流通・拡散をはじめとして、プラットフォーム事業者が提供するサービスが社会にもたらすリスクの評価については、事業者自身によるリスク評価のみならず、専門的な知見を有する研究者によるリスク評価の実施とその結果の共有も有効な手段になると考えられる。
- ・ 具体的には、一定の目的・条件を具備した研究者⁶⁷が、当事者であるプラットフォーム事業者とは異なる独立した客観的な視点から、プラットフォームサービスにおける違法・有害情報の流通状況やサービス設計がもたらすリスクについて分析を行い、事業者や利用者と結果を共有することで、プラットフォーム事業者自身のサービスの改善や利用者による異なるプラットフォームサービス間のサービス選択に資することが期待される。

⁶⁷ 例えば、EU (DSA) では、VLOP/VLOSE に対して、第 34 条に関するシステミックリスクの検知や、第 35 条に関するリスク軽減措置の妥当性等に寄与する研究を行う目的に限って、特定の研究機関への所属や商業的な利害関係からの独立性、研究資金源の開示等の要件を満たす審査を経た研究者がデータにアクセスできるようにすることを義務付けている（第 40 条）。

- ・ こうしたリスクの評価に際しては、SNS 等における違法・有害情報の流通・拡散の状況等について、投稿内容、投稿数、時系列で見た場合の傾向等を含め、データを用いた分析を行うことが必要であるが、現状、主要なプラットフォーム事業者においては、グローバルでは一定の基準で適格性のある研究者へのデータ提供が行われているものの、我が国においては一部の事業者を除きこれらの対応が行われていない⁶⁸。
- ・ したがって、まずは、(3)で述べた「自主規制型行動規範」などに基づき、プラットフォーム事業者は、我が国においてもグローバルでの対応と同様に適格性のある研究者へのデータ提供を行うことが望ましい。
- ・ その上で、仮にプラットフォーム事業者による研究者への対応が不十分である場合には、独立した客観的な視点から、プラットフォームサービスにおける違法・有害情報の流通状況やサービス設計がもたらすリスクについて分析を行うことが困難となることから、行政が違法・有害情報の流通・拡散の状況等に関するデータを事業者から入手し分析を可能とするための更なる制度的対応も検討すべきである。
- ・ なお、その際には、個人情報や企業秘密等の機微な情報の取扱いに留意しつつ、制度の趣旨に照らして過剰な範囲を要求することのないよう、DSA 等も参考に、データ取得に関する執行状況の透明性確保や対象となるプラットフォーム事業者の異議申立手続の整備についても併せて検討することが適当である。また、仮に当該制度的対応を実施する場合には、行政側の執行体制に関して、専門人材によるデータ等の分析が可能な人員の強化の必要性も併せて検討が必要であると考えられる。

4-1-2 経済的インセンティブに起因するリスクへの対応の在り方

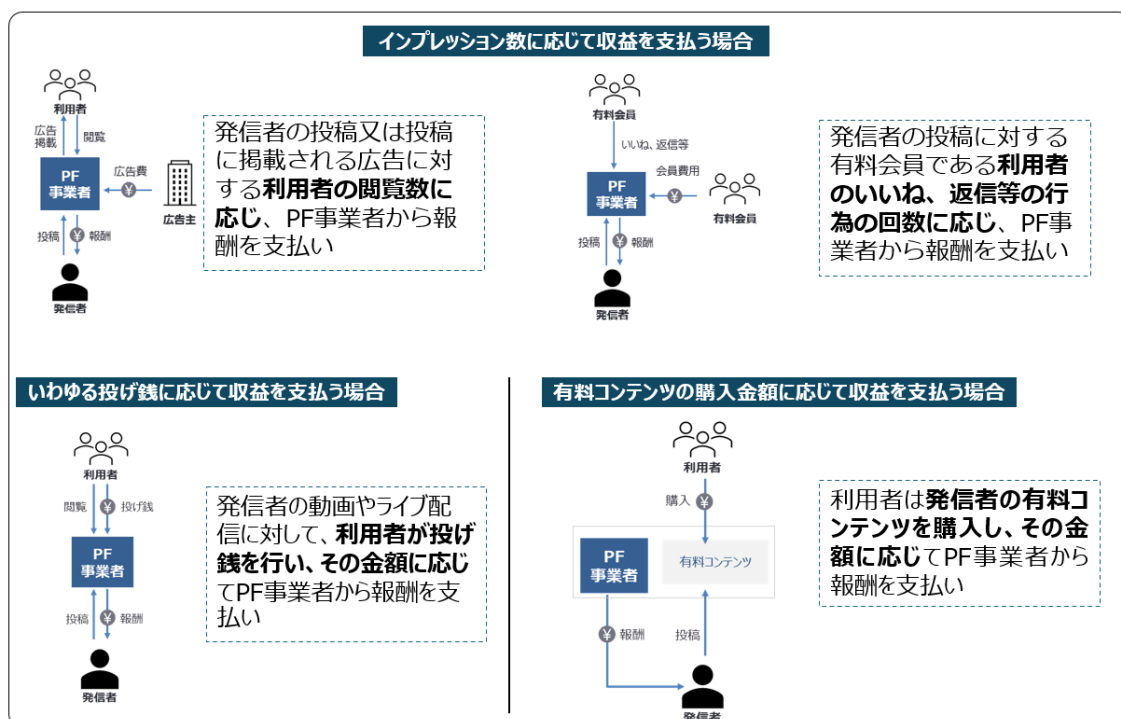
- ・ 健全性検討会のとりまとめでは、偽・誤情報の流通・拡散を直接的に抑止す

⁶⁸ 健全性検討会資料 WG25-1-3 (Code for Japan 発表資料) によると、プラットフォーム事業者がデータを研究者に無料で提供する「リサーチャープログラム」について、日本を含む世界中の研究者に提供するプログラムも存在するものの、X (Academic Research Access) は現在では利用不可、Google (Google Request Records) は欧州の研究者のみ、TikTok (Research API) は米国または欧州の研究者のみへの提供とされている。(https://www.soumu.go.jp/main_content/000950985.pdf)

るための措置として、「対象とする偽・誤情報の範囲に照らし比例的な対応を検討することが適当」とされ、情報の可視性に直接の影響がないものとしては、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したりすることにより、収益化の機会を失わせる措置）が挙げられた。

- プラットフォーム事業者が提供する収益化プログラムの内容は様々であり、また、内容も随時変更されるものである⁶⁹が、現段階における一つの分類方法としては、下図のように、
 - ① インプレッション（閲覧など）数に応じて収益を支払う場合、
 - ② いわゆる投げ銭に応じて収益を支払う場合、
 - ③ 有料コンテンツの購入金額に応じて収益を支払う場合
 の3類型に分類することができる（図表 12 参照）。

図表 12：SNS 等における収益化プログラムの種類



我が国における現行制度や対応状況

⁶⁹ 例えば、Xにおける投稿の収益化プログラムについては、2024年11月より、投稿に対して有料会員が「いいね」や返信等を行った数に応じて、有料会員から得られた会員収入の最大25%が支払われる仕組みとなっている。

- ・ 我が国の現行制度には、プラットフォーム事業者に対し、SNS 等における違法・有害情報の抑止を図る目的で収益化停止措置を求める制度は存在しない。
- ・ 事業者による自主的な取組としては、主要なプラットフォーム事業者は、収益化プログラムを提供する場合に、利用規約等において、利用者が事業者から収益を得るための要件に加え、当該プログラムにおいて収益化が停止される行為（例えば、違法なコンテンツ、紛争や自然災害に関するコンテンツ、暴力的なコンテンツ、性的コンテンツ、知的財産権を侵害するコンテンツの投稿等）を定めている。

諸外国の状況

- ・ EU（DSA）及び英国（OSA）では、プラットフォーム事業者に対し、SNS 等における違法・有害情報の抑止を図る目的で、個別の投稿について、ユーザーの収益化を停止する措置を直接的に義務付ける規定は存在しない。
- ・ ただし、EU では、行動規範の署名者であるプラットフォーム事業者は、署名者は、「偽情報の流布に資金を提供しない（defund）ことを約束し、①収益化の対象となるコンテンツの適格性、②収益化と広告掲載の管理、③広告掲載に関する管理とサービスの正確性と有効性を報告するためのデータを決定する、ポリシーとシステムを改善する」と約束している⁷⁰。

具体的検討

- ・ 本 WG では、収益化停止措置を含め、経済的インセンティブに起因するリスクへの制度的対応の在り方について、検討を行った。
- ・ 前述のとおり、プラットフォーム事業者の多くは、広告を主な収入源としているところ、投稿の閲覧数等に比例し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。インターネット上で膨大な情報が流通する中で、利用者からより多くのアテンションを集めてクリックされるために、プ

⁷⁰ 行動規範（2022 Strengthened Code of Practice on Disinformation）については、ヒアリング対象事業者のうち、Google、Meta、TikTok が署名している。2025 年 3 月 21 日現在、プラットフォーム事業者を含む 42 団体は署名している。

(<https://digital-strategy.ec.europa.eu/en/library/signatories-2022-strengthened-code-practice-disinformation>)

プラットフォーム上では過激なタイトルや内容、憶測だけで作成された事実に基づかない記事等が生み出されることがあり、より多くのアテンション（注目・関心）を集めて金銭的対価を得る「アテンション・エコノミー」と呼ばれる経済モデルの下では、違法・有害情報が広範に流通・拡散するリスクがある。この点、「アテンション・エコノミー」の下で流通・拡散する違法・有害情報から災害時の人命や民主主義の基盤をいかに守るかという視点も重要である。

- ・ また、少なくとも権利侵害情報が流通する場合については、プラットフォーム事業者は利用規約等に基づく削除対応によりリスクへの対応が可能であるため、収益化停止措置を選択する必要性がないと考えられる。このことは、違法情報が流通する場合にも当てはまると考えられる。
- ・ したがって、本節においては、有害情報に対する収益化停止措置について制度的な対応が必要か検討を行った。
- ・ この点、
 - ①収益化停止措置は、削除と同様、コンテンツモデレーションの一類型であり、（表現の自由の制約の程度は相対的に小さいと考えられるものの）SNS 等において行われる個々の表現内容に一定の制約を与える措置であること、
 - ②利用者が発言・投稿によって収益を得る機会を奪う点で経済活動の自由への制約ともなり得ること、
 - ③前述のとおり、利用規約等において収益化が停止される行為を定めるなど、事業者による一定の自主的な取組も行われていることから、プラットフォーム事業者に対し有害情報であることをもって収益化停止措置を一律に求めることについては、現時点においては、極めて慎重な検討を要する。
- ・ したがって、まずは、4-1-1（3）で述べた「自主規制型行動規範」を活用しつつ、プラットフォーム事業者が収益化停止措置に係る取組を自ら実施することが望ましい。総務省においては、その実施状況を把握・分析・評価した上で、対応が不十分である場合には、必要な改善を求めつつ、制度的対応も含め検討することが適当である。
- ・ ただし、①例えば、災害時のように、プラットフォーム事業者による速やかな対応が求められる状況において、その社会的な法益が、収益を得る権利を保護する法益を上回る場合、また、②そのような特定の場面を明確に特定し得る

場合には、法令等により、プラットフォーム事業者に対し、収益化停止措置を求めることも制度的対応として十分あり得ると考えられる。

- ・ この点については、災害時などの個々の場面ごとに、収益化停止措置によって得られる利益と、表現の自由やプラットフォーム事業者の経済活動の自由への制約との比較衡量をし、

- ① 対象となるサービス・収益の種類（ユーザーからのインプレッションや投げ銭に応じて収益を支払う場合など）
- ② 対象情報（どのような情報を対象とすべきか）
- ③ 海外からの発信者（法の対象とすることが可能か。現実的に、海外の発信者に対して支払われる収益を停止することが可能か。）
- ④ 対象者（目的・属性等に照らし一定の例外を設けるか）
- ⑤ 停止する期間（始期・終期等）

といった項目に関する考え方を丁寧に整理することが必要である。

- ・ なお、上記の論点のうち、特に、②の「対象情報」に関しては、具体的にどのようなキーワードや動画・画像が含まれていれば特定の場面に関わる情報であると判断できるか、といった点などについて精査が必要であるほか、定義の明確性、執行可能性、事業者による対応可能性の観点からも検討することが必要である。また、収益化停止をした場合の免責規定を設けること等により事業者には当該措置の実施を促すことも考えられる。

- ・ このように、プラットフォーム事業者に対して収益化停止措置を特定の場面に限定して導入することを検討する際には、上記の具体的な項目に関する考え方を整理した上で、検討していくことが適当である。

並行して、災害等の特定の場面における違法・有害情報の流通・拡散を抑止するために、収益化プログラムをはじめとする経済的インセンティブに関する利用規約等、その運用等をどのように見直すのか、その具体策についても、前述の「自主規制型行動規範」を活用しつつ、対応を促すことも考えられる。

4-1-3 レコメンダシステムに関する対応の在り方

- ・ 健全性検討会のとりまとめでは、「プラットフォーム事業者は影響を軽減するための措置を検討・実施することが適当」とされており、「影響を軽減するための措置」の例示として、レコメンデーション機能の変更等を通じた措置が挙げられている。

- ・ プラットフォーム事業者は、レコメンダシステム⁷¹（アルゴリズム）によって、利用者個人のクリック履歴など収集したデータを組み合わせて分析（プロファイリング⁷²）し、利用者に「おすすめ」するコンテンツを優先的に配信している。

我が国における現行制度や対応状況

- ・ 現状、プロファイリングについては、個人情報保護法上、「いわゆる「プロファイリング」といった、本人に関する行動・関心等の情報を分析する処理を行う場合には、分析結果をどのような目的で利用するかのみならず、前提として、かかる分析処理を行うことを含めて、利用目的を特定する必要がある」とされている⁷³。
- ・ また、アルゴリズムについては、取引透明化法において、一般利用者（消費者）が検索順位等の意味を理解し、合理的な購買行動を採ることを促すため、特定デジタルプラットフォーム提供者に対し、検索順位等の決定に用いられる主要な事項（金銭の支払が順位に影響を及ぼす可能性がある場合はその旨を含む）を一般利用者の開示することが義務付けられている（第5条第2項第2号イ）。

⁷¹ 「レコメンダシステム」とは、利用者によって開始された検索の結果として、又はその他の方法で表示される情報の相対的な順序又は優先順位を決定することを含め、ユーザーにオンライン・インターフェースでの特定の情報を提案し、又はその情報に優先順位をつけるためにオンラインプラットフォームによって使用される完全又は部分的に自動化されたシステムのことをいう。

⁷² 「プロファイリング」とは、個人データを利用して、自然人に関する特定の個人的側面を評価することである。特に、その自然人の仕事のパフォーマンス、経済状況、健康、個人的な好み、興味、信頼性、行動、場所または移動に関する側面を分析又は予測することからなる、個人データのあらゆる形式の自動処理を意味する（欧州一般データ保護規則（GDPR）第4条4項）。

⁷³ 「「個人情報の保護に関する法律についてのガイドライン」に関する Q&A」 Q&A2-1

⁷⁴ このほか、スマートフォン上の利用者情報の保護やセキュリティ確保の観点から関係事業者等が実施する望ましい事項を定めた「スマートフォン・プライバシー・セキュリティ・イニシアティブ（SPSI）」（ICT サービスの利用環境の整備に関する研究会 利用者情報に関するワーキンググループ、2024 年 11 月 29 日公表）では、スマホアプリ提供者がプロファイリングによりセンシティブ情報を予測・生成する行為は、センシティブ情報の取得につながるおそれもあると否定できないと考えられることから、原則として実施しないこととし、実施する場合には、利用者本人に対して個別の同意取得を行うことが望ましいとしている。

諸外国の状況

- ・ EU (DSA) では、プラットフォーム事業者がレコメンダシステムを使用している場合、使用する主なパラメータと利用者が当該パラメータを変更できる選択肢について、利用規約に平易かつ分かりやすい言葉で明記することが義務付けられている（第 27 条⁷⁵第 1 項）。

この主なパラメータには、レコメンドされる情報を決定する上で、最も重要な基準と当該パラメータがなぜ重要であるかの理由が含まれる必要がある（第 27 条第 2 項）。

さらに、複数のレコメンド方法を選べる場合、利用者がレコメンド方法をいつでも選択・変更できる機能を提供することが義務付けられている（第 27 条第 3 項）。

また、VLOP/VLOSE に対しては、レコメンダシステムをテストし、リスク評価結果に適合させることがリスク軽減措置の一環として例示（第 35 条第 1 項(d)）されるとともに、レコメンダシステムに少なくとも 1 つ、プロファイリングに基づかない選択肢を提供することが義務付けられている（第 38 条）。

一連の DSA の規定は、表示される情報がどのように優先順位付けされたものであるのかを利用者自身が理解できるようにすることの確保を目的とするものとされている⁷⁶。

また、アルゴリズムシステムの影響を研究するための専門機関が存在し（イタリアにある ECAT⁷⁷ (European Centre for Algorithmic Transparency)）、DSA の執行をサポートしている。

⁷⁵ DSA 第 27 条では、「主なパラメータ」には、なぜ特定の情報が利用者に推奨されるのかを説明する内容が含まれていなければならない。具体的には、少なくとも、(a) 利用者に推奨される情報を決定する上で最も重要な基準、及び、(b) それらの基準が重要とされる理由を含めるものとされている。

⁷⁶ DSA 前文 (70) では、「・・・サービス利用者が自身に対して情報がどのように優先順位を付けられているのかを確実に理解できるよう、レコメンダシステムのパラメータを簡単に理解できる方法で明確に提示する必要がある。当該パラメータには、情報がプロファイリングとそのオンライン上の行動に基づき優先順位付けされる場合を含め、少なくとも、サービス利用者に提案される情報を決定する際の最も重要な基準とそれぞれについての重要性の理由を含めなければならない。」とされている。

⁷⁷ https://algorithmic-transparency.ec.europa.eu/about_en。報道によれば、30～40 名の職員がいるとのこと。

(<https://fedscoop.com/european-union-ai-algorithm-transparency-united-states/>、<https://www.yomiuri.co.jp/world/20250327-OYT1T50018/>)

- ・ 英国（OSA）では、ユーザー間サービス提供事業者に対し、安全設計に関する措置を組み込むことが義務付けられており（安全義務）、その具体的な内容として、レコメンダシステムに関する措置は Ofcom が策定する行動規範 (Codes of practice) に明記されている⁷⁸。

具体的検討

- ・ 本 WG では、レコメンダシステムに関する制度的対応の在り方について検討を行った。
- ・ レコメンダシステムは、SNS 等の利用者の利便性の向上に資するものであると考えられる。一方で、アルゴリズムを用いて配信された情報を受け取り続けることにより、ユーザーが自身の興味のある情報だけにしか触れなくなる「フィルターバブル」を引き起こす可能性⁷⁹、自分と同じ考えの人とのみ繋がる「エコーチェンバー」に陥る可能性、また、本人の無意識のうちに依存的利用に陥ってしまう可能性がある⁸⁰など、レコメンダシステムに起因する一定のリスクの発生が明らかになりつつある。SNS 等においては通常レコメンダシステムが利用されている一方、利用者の中でレコメンダシステムを理解する者は半数以下にとどまるとの調査結果⁸¹があるなど、リスクへの理解が利用者の間に十分広がっているとはいえない状況にある。
- ・ このようリスクへの対応として、利用者が、レコメンダシステムを通じて表示される情報について、どのような基準やプロセスによって優先順位付けされたものであるか十分理解でき、また、同システムを利用するかどうかを利用者自身が自ら決定できるような環境とすることが、情報発信・受信のための

⁷⁸ 行動規範 (Codes of practice) に明記された措置の内容としては、「違法コンテンツへの接触可能性等の安全性指標の作成・分析」「レコメンダシステムを設計変更する際に、テストユーザーからフィードバックを受け付けるテスト環境の整備」など。

⁷⁹ 総務省「令和 6 年版情報通信白書」。

⁸⁰ 欧州委員会が TikTok Technology Limited に対して、TikTok のシステム（アルゴリズムシステムを含む。）について、依存症行動の助長やラビットホール効果 (Rabbit Hole Effects、延々と関連コンテンツを見続けてしまう現象を指す。) の発生に関するリスク評価・軽減措置義務違反の疑いがあるとして、法的手続 (DSA 第 66 条) を開始している。https://ec.europa.eu/commission/presscorner/detail/en/ip_24_926

⁸¹ 高口構成員発表資料（本 WG 第 9 回）では、回答者の 46.6% が、「履歴などをもとに利用者の好みに合ったコンテンツを提案・表示するレコメンダ」について知らなかったという調査結果であった。

公共的な基盤としてのプラットフォームの機能が重要性を増す中で、極めて重要である。

- ・ そのためには、サービスを提供するプラットフォーム事業者において、レコメンダシステムの透明性の向上を通じて利用者による理解の促進を図ること、利用者がレコメンダシステムを介さずに情報を受け取ることのできるオプションを提供することが肝要である。レコメンダシステムの透明性の確保については、外部から改善を求める社会的な声が上がることにより、改善が促されるという効果も期待される。
 - ・ 本 WG において紹介された調査結果⁸²をみても、レコメンダシステムの透明性を図るという方向性については、国民の受容性はあると考えられる。
 - ・ 以上を踏まえ、レコメンダシステムに関する制度的対応の在り方としては、プラットフォーム事業者に対し、①～③の取組を求める対応を検討することが適当である。なお、これらの取組は互いに排他的なものではなく、同時に求めることも可能である。
- ① プロファイリング又はレコメンダシステムの透明性を確保すること⁸³
(例：「いいね」等のユーザーのリアクション、視聴回数等のコンテンツ情報、言語設定などのユーザー情報等のレコメンダ機能に影響を与えている主要な要素)に加え、利用者にとって平易な言葉による説明等を行うこと
 - ② 複数のレコメンダ方法を選べる場合、任意のレコメンダ方法を選択したり変更したりすることが可能な機能を利用者に提供すること

⁸² 脚注 85 参照。レコメンダシステムについては、過半の回答者がレコメンダに関する機能を平易な言葉で説明する対策を「よい対策」と評価しており、説明の平易さにとどまらず、説明の表示の仕方、サービス利用者がとりうる対応についてサポートなどが工夫されればさらに「よい対策」とする割合が増える可能性があるということであった。また、レコメンダの選択肢の提示については、選択肢を提示されても選択が困難であると回答する層が一定数いるため、「よい対策」とする割合は4割程度であるが、選択肢の中に「プロファイリングによらないという選択肢」を含めると過半の回答者がよい対策と評価している。

⁸³ 注 74 の SPSI は、プロファイリングの透明性を確保する観点から、スマホアプリ提供者がプロファイリングを行う場合には、どのような取扱いが行われているかを利用者が予測・想定できる程度に利用目的を特定し、かかる分析処理を行うことを含めて利用目的を特定するとともに、プロファイリング結果に基づき、利用者にとって重要な決定が自動的に行われることがある場合には、その旨や当該決定に至る際に依拠する基準等を利用者に明示することが望ましいとしている。

③ プロファイリングに基づかない情報表示の選択肢を利用者に提供すること

- ・ これらの制度的な対応を実現するための具体的な方策としては、以下の2つの方向性があると考えられる。なお、方向性Bにより事業者による軽減措置の検討・実施を促したとしても、レコメンダシステムについて上記①～③のような取組を自主的に実施するインセンティブが事業者側には生じ得ないと考えられることに留意が必要である。

＜方向性A＞

- ・ プラットフォーム事業者に対し、上記①～③の取組の実施を法令等により義務付ける方法

※ アルゴリズムの透明化や選択肢を利用者に示す義務は、プラットフォーム事業者が提供するサービスの内容に関わらず、義務の内容も明解である。したがって、立法事実があることを前提として、法律上義務化することは我が国の法文化等に鑑みても問題はない。

＜方向性B＞

- ・ プラットフォーム事業者に対し、上記①～③の取組も含め、レコメンダシステムがもたらす影響を軽減するための措置を自ら実施することを法令等により促す方向性

※ レコメンダに用いるアルゴリズムの透明性確保や選択肢の提示の在り方は、プラットフォーム事業者のサービスの利用者層やサービス内容に応じて異なりうることから、方向性Aのように一律に義務化する場合、サービスを制約する程度が大きい可能性がある。

※ 他方で、我が国の一般的な法制のように、一定の主体に対して法律上の義務を課す場合には、その義務内容が具体的かつ明確であり、違反した場合には行政による執行が可能な規律であることが求められる。方向性Bのように「レコメンダシステムについて軽減措置を求める」といった抽象的な規律を設ける場合には、その点にも留意する必要。

4-2 適切な情報表示の在り方

- ・ 「適切な情報表示の在り方」には、情報の受信者に対する適切な情報表示の

在り方と、情報の発信者に対する適切な情報表示の在り方が考えられるが、本 WG では、考え方（２）のとおり、「利用者に自由な情報摂取の機会を保障する」という観点から、前者を中心に検討を行った⁸⁴。

4-2-1 プロミネンス（優先表示）に関する対応の在り方

- ・ 健全性検討会のとりまとめでは、「災害発生時等における社会的影響の軽減措置として、信頼できる情報源からの情報の伝送確保（プロミネンス⁸⁵）及びその基準の明確化など、更なる具体化を進めることが適当」との提言がなされた。

我が国における現行制度や対応状況

- ・ プラットフォーム事業者における取組の例としては、自社のサービス上において、国や地方公共団体等のアカウントや当該アカウントが発信した情報を目立つ形で表示する取組や、自社や他の機関が提供する信頼できる情報に誘導するため、リンクやバナー等を目立つ箇所に表示する取組がある。例えば、災害時に特定のキーワードで検索した際に、報道機関等のニュースが見つけやすいような表示をする取組が実施されている。
- ・ また、総務省では、2025 年 1 月から、インターネットや SNS における利用者の ICT リテラシー向上を目指し、官民連携での意識啓発プロジェクトとして「デジタルポジティブアクション」を行っている。このプロジェクトでは、総合的なリテラシーの向上に向けて、「世代に応じた多様な普及啓発」、「SNS、デジタルサービスにおけるサービス設計上の自主的な工夫」、「信頼性の高い情報にかかる表示上の自主的な工夫」を方向性として、これらの趣旨に賛同する企業・団体と総務省が連携して行っているものである。

具体的検討

- ・ 本 WG では、一部の事業者が災害時などに優先表示を既に自主的に実施して

⁸⁴ 情報の発信者に対する情報表示としては、例えば、利用規約等に違反すると判定されたコメントを複数回投稿している利用者に対して投稿前に注意のメッセージを掲出するなどの「警告表示」がある。

⁸⁵ この「プロミネンス」については、まだ訳語が定着していないと思われるが、ここでは「優先表示」という訳語をひとまず当てることとする。

いることも踏まえつつ、これらの対応に係る制度的対応の在り方について検討を行った。

- ・ 優先表示は、偽・誤情報の流通・拡散に対して、カウンター発信としての効果があると考えられる。しかし、どのような情報源・情報の内容を「信頼できる」情報と位置付けるべきか、また、当該情報の表示を「どのように優先」すべきか（文字の大きさ、表示場所等）といった点について、各プラットフォーム事業者に共通する一律の具体的な基準を示すことには、現時点で課題があると考えられる⁸⁶。
- ・ したがって、優先表示に関する制度的対応については、4-1-1（3）で述べた「自主規制型行動規範」を活用しつつ、各事業者が提供するサービスの性質等に応じた形で信頼できる情報の優先表示の実施を求めるかどうかも含め、更に検討を深めることが適当である（その場合の考え方は、上記4-1-1（リスク評価・軽減措置）を参照）。
- ・ 並行して、総務省のデジタルポジティブアクションなどの官民連携の意識啓発プロジェクトの中で、事業者の取組を把握し、効果のあった優先表示の事例など、優良な先進事例を事業者間で情報共有することを通じ、当該先進事例の取組が事業者間で広がるとともに、利用者に周知されていくことが望ましい。

4-2-2 AI生成物のラベル付与に関する対応の在り方

- ・ 健全性検討会のとりまとめでは、偽・誤情報の流通・拡散を直接的に抑止するための措置として、「対象とする偽・誤情報の範囲に照らし比例的な対応を検討することが適当」とされ、情報の可視性に直接の影響がないものとしては、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）等が挙げられた。
- ・ また、「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

⁸⁶ また、構成員から、現実の人間の情報摂取行動や認知過程において優先表示に関する取組が本当にどこまで正の効果をもたらしているのかに関する効果検証が必要ではないかといった指摘もあった。

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」との提言がなされた。

我が国における現行制度や対応状況

- ・ 事業者の取組としては、利用者がコンテンツ投稿時に自主的に AI 生成物である旨を明示できる機能を提供する取組や、AI 生成物であることを検知したりした場合に事業者側でラベル付けする取組が行われている。
- ・ 政府の取組としては、総務省及び経産省において、近年で更に発展した AI 技術の特徴及び国内外における AI の社会実装に係る議論⁸⁷を反映し、AI 活用に取り組む全ての事業者⁸⁸向けのガイドラインとして、2024 年 3 月、「AI 事業者ガイドライン」が取りまとめられた（2025 年 3 月に 1.1 版が公表⁸⁹。）。同ガイドラインでは、「全ての AI 関係者向けの広島プロセス国際指針」などを踏まえ、高度な AI システムに係る事業者⁹⁰に対して「技術的に可能な場合は、電子透かし⁹¹やその他の技術等、AI 利用者及び業務外利用者が、AI が生成したコンテンツを識別できるようにするための、信頼できるコンテンツ認証及び来歴のメカニズムを開発し、導入する」とされ、「透かしや識別子を利

⁸⁷ AI を巡る国際的な枠組みとしては、2023 年 5 月の G7 広島サミットの結果を受け、生成 AI に関する国際的なルールの検討を行うため、「広島 AI プロセス」が立ち上げられた。同年 10 月の「広島 AI プロセスに関する G7 首脳声明」を踏まえ、同年 12 月に「G7 デジタル・技術閣僚会合」において、「全ての AI 関係者向けの広島プロセス国際指針」がとりまとめられた。同指針では、技術的に可能な場合は、電子透かしやその他の技術等、ユーザーが AI が生成したコンテンツを識別できるようにするための、信頼できるコンテンツ認証及び来歴のメカニズムを開発し、導入することが挙げられている。

⁸⁸ AI 事業活動を担う主体について、AI 開発者（AI システムを開発する事業者（AI を研究開発する事業者を含む））、AI 提供者（AI システムをアプリケーション、製品、既存のシステム、ビジネスプロセス等に組み込んだサービスとして AI 利用者、場合によっては業務外利用者に提供する事業者）、AI 利用者は、（事業活動において、AI システム又は AI サービスを利用する事業者）に大別して整理されている。

⁸⁹ AI 事業者ガイドラインの第 1.0 版が 2024 年 4 月 19 日に公表され、その後、第 1.01 版が同年 11 月 22 日に公表、第 1.1 版が 2025 年 3 月 28 日に公表されている。

⁹⁰ 最先端の基盤モデル及び生成 AI システムを含む、最も高度な AI システムを開発・利用する組織のための指針を提供するものであり、組織には、学界、市民社会、民間部門、公共部門の主体等が含まれる。「透かしや識別子を利用する」取組を含め、AI 開発者のみに適用される内容もあるが、AI 利用者及び AI 提供者に対しても、適切な範囲で遵守することが求められている。

⁹¹ AI が生成したことを示す識別情報をコンテンツに埋め込む技術のこと。

用することに加え、各主体がこの分野の状況を前進させるために協力し、研究に投資すべき」と記載されている。

諸外国の状況

- ・ EU (DSA) では、リスク軽減措置の一環として、AI 生成物等へのラベリング機能⁹²（人物や事象に著しく類似し、誤認させるような画像、音声、動画がオンライン・インターフェースに表示される際に目立つマークによって区別できる機能）の提供が例示されている（第 35 条第 1 項(k)）。
- ・ また、米国（カリフォルニア州 AI 透明化法）では、対象事業者に対して、ユーザーへの AI 検出ツールの無償提供、自社システムで生成された AI コンテンツに対して AI 生成物であることを明記する選択肢の提供等が義務付けられている（セクション 22757）。

具体的検討

- ・ 本 WG では、AI 生成物のラベル付与に関する制度的対応の在り方について検討を行った。
- ・ AI 生成物のラベル付与については、生成 AI により動画、画像等を生成する際にラベルを付与する取組と、プラットフォーム事業者が生成 AI により生成された動画、画像等を識別した際にラベルを付与する取組がある。
- ・ 上述の「AI 事業者ガイドライン」では、高度な AI システムに係る事業者に対して、「透かしや識別子を利用する」よう奨励されており、AI 利用者及び AI 提供者に対しても、当該取組を適切な範囲で遵守することが求められている。
- ・ 同ガイドラインには「AI 生成物のラベル付与」については明示的な記載はないものの、「透かしや識別子を利用する」取組と類似するものと考えられる

⁹² EU の AI 規則第 50 条では、AI システムの提供者及び導入者の透明性義務として、ラベリング（AI によって人工的に生成されたものとして検出可能なことの保証）が義務付けられている。

(<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>)

ことから、プラットフォーム事業者⁹³による AI 生成物のラベル付与等の取組は、AI 生成物を識別できるようにするという同ガイドラインの趣旨に沿うものであり、取組が更に広がることが期待される。

- ・ 一方で、構成員からは、ラベル付与も含めた適切な情報表示の在り方については、効果検証をするなど実態把握をすることがまずは必要という指摘もあった。
- ・ この点、電子透かし⁹⁴や来歴管理⁹⁵といった関連技術も含めて、利用者に対して生成 AI の利用を明示する取組は未だ技術開発途上である。特に、生成 AI を用いて生成される偽・誤情報について、自社システムで生成された AI 生成物については、プラットフォーム事業者において AI 生成物かどうかの識別が可能である場合もあるが、他社システムで生成された AI 生成物については、プラットフォーム事業者において真偽の判断を行うことが技術的に困難な場合もあると考えられる。また、ラベル付与の具体的な手法についても、最も効果的な手法は事業者のサービス形態によって異なる可能性もある。
- ・ したがって、プラットフォーム事業者に対しては、電子透かしや来歴管理とラベル付与との技術的差異やそれぞれの効果や実現性の比較をしつつ、4-1-1（3）で述べた「自主規制型行動規範」を活用するか、引き続き、検討をしていくことが適当である（リスク軽減措置の一環としての対応の実施を求める場合の考え方は、上記 4-1-1（リスク評価・軽減措置）を参照）。なお、AI 生成物のラベル付与については、上述のとおり、EU（DSA）の軽減措置の例示の中にも含まれている。）。
- ・ 併せて、同ガイドラインにおいて、「AI 開発者」又は「AI 提供者」に該当する AI 事業者による AI 生成物に対するラベル付与等のコンテンツ識別の在り方や、当該 AI 事業者がプラットフォーム事業者に当たる場合の、同プラット

⁹³ 同ガイドラインにおいては、プラットフォーム事業者は、プラットフォーム事業者のサービス上において AI 生成物が投稿される場合、当該 AI 生成物が自社の開発する AI システムによって生成される場合は、「AI 開発者」として、自社の提供する AI サービスによって生成される場合は、「AI 提供者」として扱われる。（なお、他社が開発・提供する AI システム及び AI サービスによる AI 生成物については、プラットフォーム事業者は同ガイドラインの対象外となる。）

⁹⁴ AI が生成したことを示す識別情報をコンテンツに埋め込む技術

⁹⁵ 作成者等の来歴情報を検証可能な形でコンテンツに付与する技術

フォームへの投稿に対するラベル付与等の在り方について、同ガイドラインに明確に記載すべきか、引き続き、検討をしていくことが適当⁹⁶である。

4－3 利用者の確認に関する対応の在り方（アカウント開設時の本人確認等）

- ・ 健全性検討会のとりまとめでは、「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アカウント登録時等の本人確認の厳格化等が考えられるが、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、…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、…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」との提言がなされた。
- ・ 健全性検討会のとりまとめ後、SNS 等における闇バイトの募集投稿が問題となったことなどから、政府において、「被害に遭わせない」ための対策として、SNS 等を利用した犯罪の捜査上の課題に対応するため、SNS アカウントの開設時の本人確認の強化を含む措置について検討を行うこととされ⁹⁷、総務省からもプラットフォーム事業者に対して本人確認の厳格化の検討要請が行われたところである。

我が国における現行制度や対応状況

- ・ 現在、我が国において、プラットフォーム事業者に対してアカウント開設時に本人確認を義務付ける法律はない。
- ・ なお、一部のプラットフォーム事業者は、SNS 等のサービスにおけるアカウント開設時の本人確認の方法として、携帯電話による SMS 認証を採用している。同認証については、携帯電話契約において適切に本人確認がなされている

⁹⁶ 「AI 事業者ガイドライン」は、総務省「AI ネットワーク社会推進会議・AI ガバナンス検討会合同会議」及び経済産業省「AI 事業者ガイドライン検討会」にて議論の上、策定・更新等を行っている。なお、人工知能関連技術の研究開発及び活用の推進に関する法律（令和 7 年法律第 53 号）が第 217 回国会（常会）にて成立し、今後、同法に基づき AI 戦略本部が設置され、基本計画の策定、指針の整備等が進められるところ、これらとの関係についても関係府省が連携し引き続き検討していくことが適当である。

⁹⁷ 犯罪対策閣僚会議「いわゆる「闇バイト」による強盗事件等から国民の生命・財産を守るための緊急対策」

(https://www.kantei.go.jp/jp/singi/hanzai/kettei/241217/kinkyu_taisaku.pdf)

ことが前提となっている。携帯電話契約における本人確認については、匿名の携帯電話が振り込め詐欺等の犯罪に利用されていたことを受け、携帯音声通信事業者による契約者等の本人確認等及び携帯音声通信役務の不正な利用の防止に関する法律（平成 17 年法律第 31 号）において、携帯電話事業者が音声通信の契約者の身分証明書による本人確認等を行うことを義務付けている。

- ・ 近年の改正として、一見して判別できないほど精巧に偽変造された本人確認書類を利用して契約された事例の発生を受け、令和 7 年 3 月に同法施行規則⁹⁸が改正され、非対面の本人確認手法は、マイナンバーカードの公的個人認証に原則として一本化する本人確認の強化が図られたところであり、この意味で、SNS 等のサービスにおけるアカウント開設時に SMS 認証による本人確認を行っているプラットフォーム事業者における本人確認も厳格化されている。さらに、現在、総務省の有識者会合⁹⁹において、SMS 機能付きデータ通信専用 SIM 契約時の本人確認義務付けに関しても検討が進められているところである。

諸外国の状況

- ・ 英国（OSA）においては、大規模なユーザー間サービス事業者（カテゴリー 1）に対して、成人ユーザー向けの本人確認のオプションを提供することを義務付けている（第 64 条・第 65 条）。また、成人ユーザーが希望する場合、本人確認を行っていないユーザーをフィルタリングする機能を提供することを義務付けている（第 15 条第 9 項）。

韓国においては、オンライン掲示板の利用者向けの本人確認を義務付けていたこともあったが、一部規定については違憲判決が出され、削除されている¹⁰⁰。

具体的検討

⁹⁸ 携帯音声通信事業者による契約者等の本人確認等及び携帯音声通信役務の不正な利用の防止に関する法律施行規則（平成 17 年総務省令第 167 号）

⁹⁹ ICT サービスの利用環境の整備に関する研究会 不適正利用対策に関するワーキンググループ（主査：大谷和子 株式会社日本総合研究所執行役員法務部長）

¹⁰⁰ 情報通信網法において、オンライン掲示板の利用者向けの本人確認を義務付けていたが、憲法裁判所にて違憲判決が出され、民間事業者向けの本人確認義務規定は削除された。

また、公職選挙法において、インターネット報道機関の掲示板やチャットページにおける選挙期間中選挙に関する投稿を対象に実名認証と認証表示を義務付けていたが、憲法裁判所にて違憲判決が出され、本人確認義務規定については削除された。

- ・ 本 WG では、プラットフォーム事業者によるアカウント開設時の本人確認の在り方について、特に匿名表現の自由との関係という観点から、①アカウント開設後、プラットフォーム上において利用者の実名表示等を義務付けることが許容されるか、②プラットフォーム事業者に対してアカウント開設時の本人確認を義務化することが許容されるか検討を行った。
- ・ まず、匿名表現の自由と表現の自由との関係については、憲法第 21 条における表現の自由は、自己の素性を明らかにしないまま、匿名で表現活動を行うことにも及び（匿名性の保障）、匿名表現の自由も憲法第 21 条の保護領域に含まれる¹⁰¹と考えられる。また、下級審ではあるが、匿名による表現活動を行う自由は、憲法第 21 条第 1 項により保障されているものと解される旨の裁判例もある¹⁰²。
- ・ したがって、アカウント開設後、プラットフォーム上において利用者の実名表示等の義務付けることが許容されるかという点については、義務付けを行うことにより利用者は匿名での表現行為が禁止されることになるため、規制が必要不可欠と考えられる限定的な場面で合理性のある場合に限られると考えられ、慎重な検討が必要である。
- ・ 次に、プラットフォーム事業者に対してアカウント開設時の本人確認を義務化することが許容されるかという点については、本人確認を義務化しても匿名でのプラットフォーム上の表現行為が許容される場合には、プラットフォーム上における実名表示等の義務付けを行う場合に比べて、表現の自由への制約の程度が小さいといえる。しかし、匿名表現の自由は表現の自由として保障されることから、合憲性の評価については慎重な比較衡量を行うことが必要と考えられる¹⁰³。
- ・ 例えば、犯罪捜査における追跡可能性を確保するという目的に基づき、プラットフォーム事業者に対し、アカウント開設時の本人確認義務を課す場合、ど

¹⁰¹ 本 WG 第 4 回における曾我部真裕 京都大学大学院法学研究科教授発表資料参照。

¹⁰² 例えば、大阪市ヘイトスピーチ対処条例事件 1 審判決（大阪地判令和 2 年 1 月 17 日判例自治 468 号 11 頁）。曾我部真裕 京都大学大学院法学研究科教授の発表（本 WG 第 4 回資料）参照。

¹⁰³ 仮に本人確認を義務化したとしても、プラットフォーム事業者が強制捜査等により情報を開示するにあたっては、開示の必要性・許容性を個別に判断し、手続的保障も考慮することが必要である。

のような事業者に本人確認を義務付けるのか（例：プラットフォーム事業者）、対象となる利用者はどの範囲か、目的と手段の比例性は確保されているか、目的に照らして他に有効な手段はないか、規律に実効性はあるかといった点に関する検討を、立法事実を踏まえて行うことが適当である。

- ・ なお、上記①の利用者の実名表示等についても、上記②のアカウント開設時の本人確認についても、これらの措置を義務化するのではなく、利用者自身が実名表示・本人確認をするか否かを選択できるようにする措置をプラットフォーム事業者に義務付けることとする場合には、匿名表現の自由との関係では①②に比べて制約の程度が小さいものと考えられる。

4-4 小括

※ 4-3までの議論を踏まえて記載予定

5 その他

5-1 マルチステークホルダーによる協議会

- ・ 健全性検討会のとりまとめでは、インターネット上における情報流通の健全性確保に向け、「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」との提言がなされた。
- ・ また、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プラットフォーム事業者による恣意的な影響予測を防ぎ、かつ、情報伝送プラットフォーム事業者が実施する軽減措置の実効性を担保するため、「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」など具体化を進めることが適当とされた。
- ・ 本WGでは、マルチステークホルダーによる協議会を立ち上げる場合、協議会にはどのような役割を持たせるべきかについて検討を行った。

- ・ 構成員からは、情報の真偽・違法性の判断をプラットフォームだけに委ねるのではなく、信頼できる第三者の団体と様々なステークホルダーが関わりながら、判断の能力を社会全体で高めていくべきという指摘があった。
- ・ 一般論として、マルチステークホルダープロセスは多様な価値観をインプットするためには有効な手段である。一方で、このプロセスが有効に機能するためには協議会の目的と責任の所在を明確にしなければならない。このような点も踏まえつつ、引き続き、マルチステークホルダープロセスが有効に機能する場面について議論を継続していくことが適当である。

5-2 執行当局の執行権限・体制の強化

- ・ インターネット上における情報流通を巡る環境は変化しており、闇バイトやオンラインカジノなど、新たな社会問題なども発生している。このようなインターネット上の違法・有害情報に対応するためには、制度的整備に加え、対策技術の開発や利用者のリテラシー向上など、様々な取組が官民で実施されることが必要である。本 WG では特に制度的な対応の在り方について検討し、情報流通プラットフォーム対処法の適切な運用に加え、更なる制度整備の在り方についても検討を進めていく必要がある旨提言を行ったが、これらの提言に実効性を持たせるためにも、制度整備の議論と並行して、制度の運用を担う行政の執行権限や体制の強化についても検討が必要である。

我が国における現行制度や対応状況

- ・ 現行の情報流通プラットフォーム対処法では、総務大臣は、大規模プラットフォーム事業者に対し、同法に規定する義務規定の施行に必要な限度において、その業務に関し報告をさせることができるほか（第 29 条）、義務規定の違反が認められた場合に必要な措置を講ずることを勧告すること等ができる（第 30 条）。
- ・ 我が国では、総務省において、2023 年 7 月に情報流通行政局に情報流通適正化推進室が新設された¹⁰⁴。同室では、「情報の電磁的流通の適正な利用の促進」に関する事務をつかさどるとされており¹⁰⁵、情報流通プラットフォーム対

¹⁰⁴ https://www.soumu.go.jp/main_content/000890337.pdf

¹⁰⁵ 総務省組織規則（平成 13 年総務省令第 1 号）第 45 条第 4 項。

処法は同室において運用されているほか、政府におけるインターネット上の偽・誤情報を含む違法・有害情報対策の中心を担っている。

- ・ また、他省庁における、巨大 IT 企業を監督・規律している行政側の体制整備という観点では、例えば、公正取引委員会においては、スマホソフトウェア競争促進法の実効的な運用に向けて必要な体制を構築していくために、巨大 IT 企業の調査を担当する役職を新たに設置するとともに、常勤職員は 14 人から 51 人に増強したほか、民間採用の非常勤の IT 専門家「デジタルアナリスト」を 7 人から 10 人に増強している。

諸外国の状況

執行権限関係

- ・ EU (DSA) では、VLOP/VLOSE の義務違反の疑いがある場合は、欧州委員会は、事前調査（第 66 条）を開始し、情報提供要請（第 67 条）を行うことができる。事前調査後も違反の疑いがある場合は、法的手続を開始し、関係者へのインタビュー（第 68 条）、立入検査（第 69 条）等を実施し、違反を決定することとなっている。また、EU (DSA) では、欧州委員会は、DSA の遵守状況を監視するために、VLOP/VLOSE に対し、保有するデータベースとアルゴリズムへのアクセスを求め、説明を命じることができる（モニタリング措置。第 72 条）。この措置には、DSA の遵守状況を評価するために必要とみなされる全ての文書を保持する義務を課すことを含めることができる。
- ・ EU (DSA) では、リスク評価に基づく軽減措置として 4-1-1 で述べたような広範な内容をリスク評価と軽減措置の文脈で義務化しており、事実上、違法・有害情報への対応としてプラットフォーム事業者がどのようなプロセスや対応をとっているのか網羅的に聴取する権限がある。
- ・ 英国 (OSA) では、規制対象サービス提供事業者（プラットフォーム事業者等）の違反が継続している場合や Ofcom の措置要求に従っていない場合等の一定の場合、Ofcom は、裁判所に、プラットフォーム事業者等にサービスを提供する者（付随サービスと呼ばれる。課金システム提供者等）のサービス制限命令を出すよう申請することができる（第 144 条）。さらに、サービス制限命令の効果が不十分である場合又はサービス制限命令を出しても重大な被害を防げないと見込まれる場合に、裁判所に対して、アクセスプロバイダにアクセ

ス制限命令を出すよう申請することができる（第 146 条）。

体制関係

- ・ EU（DSA）では、アルゴリズム分析、法務、経済分析等の専門家を含む、DSA の執行担当者が 120 名超おり、本年末までに 200 名まで増強予定とされている。また、英国（OSA）の担当官の数についても、報道によれば、Ofcom の担当者曰く、2024 年 7 月は 466 人だったものを、2025 年 3 月末までに 557 人まで雇用を増やすとされている¹⁰⁶。

具体的検討

- ・ 本 WG では、インターネット上の違法・有害情報対策として、既存の執行権限に加えて、どのような権限が求められるか検討を行った。
- ・ 今後、総務省は、情報流通プラットフォーム対処法の透明化規律に基づき事業者から公表された事項のモニタリングも含め、大規模プラットフォーム事業者による同法の各規律の履行状況等について把握・分析を行っていくことが想定されるが、仮にデータが十分集まらないなど、モニタリングが適切に実施できない状況である場合は、透明化規律による公表事項の充実を図ること、行政に新たに調査権限を設けることなどを検討することが適切である。
- ・ また、EU（DSA）や英国（OSA）では、様々な執行権限が規定されているところであるが、更なる制度整備の内容に応じてどのような執行権限が効果的か、諸外国の動向の調査や欧州や英国の執行機関との意見交換等を通じて検討を深めることが適切である。
- ・ 次に、本 WG では、令和 6 年改正法の施行を受け、また、新しい制度を設ける場合に必要な執行体制の在り方について検討を行った。
- ・ 上述のとおり、今後、総務省は、情報流通プラットフォーム対処法の透明化規律に基づき事業者から公表された事項のモニタリングも含め、大規模プラットフォーム事業者による同法の各規律の履行状況等について把握・分析を

¹⁰⁶ <https://www.independent.co.uk/business/ofcom-hiring-more-online-safety-staff-amid-calls-to-tackle-misinformation-b2598960.html>

行うことが想定される。こうした分析や評価を適切に実施するためには、法務・データ分析の専門家を含む専門的な知見を有する人材を集めた体制の整備や、膨大な情報を分析・評価するのに十分な人員の確保が必要である。

- ・ 以上により、情報流通プラットフォーム対処法の適切な運用のための専門性の向上（法務やデータ分析の専門家の確保）と増員は喫緊の課題であり、今後も更なる制度的対応を行うこともあり得ることから、体制の強化について検討を進めていくことが適当である。

V 結論

※ IVまでの議論を踏まえて記載予定