

# プラットフォームサービスに係る違法・有害情報 (誹謗中傷、偽情報)への対策に関する主な論点 (案)

---

2022年6月9日  
事 務 局

## 1. 誹謗中傷等への対応に関する現状と課題

### （1）現状と課題

- ①これまでの対策と議論の方向性
- ②流通状況
- ③各ステークホルダーの取組状況

### （2）プラットフォーム事業者等による対応のモニタリング結果（2022年）

### （3）海外動向

- 誹謗中傷及びフェイクニュース・偽情報への対応について、2021年9月の中間とりまとめ、「緊急提言」及び「政策パッケージ」に基づき、プラットフォーム事業者をはじめとする産学官民の多様なステークホルダーによる取組は適切に進捗しているか。
- 諸外国ではどのような制度的検討や対策が進められているか。

## 3. 今後の取組の方向性

### （1）違法・有害情報全般への対応

### （2）偽情報への対応

- 当該実態を踏まえ、誹謗中傷や偽情報といった違法・有害情報への対策として、プラットフォーム事業者をはじめとする多様なステークホルダーにより、今後さらに具体的にどのような取組が進められることが適当か。
- 特に、プラットフォーム事業者について、主体的な情報の削除等の対応及び取組の透明性やアカウントビリティの確保が適切に行われるよう促しつつ、それらの取組状況をモニタリングし検証・評価を行うという共同規制的なアプローチを効果的に実現するために、どのような枠組み・方策が必要か。

# 1. 誹謗中傷等への対応に関する現状と課題

---

## 論点 (前提)

- インターネット及び携帯電話・スマートフォン等の急速な普及は、個人間のコミュニケーションを容易にするとともに、個人がソーシャルメディアにおいて様々な発信を行うことを可能とした。特に、SNS（ソーシャルネットワーキングサービス）等のソーシャルメディアの利用時間や利用率が近年大きく伸びており、SNS等のソーシャルメディアは我が国の日常生活や社会経済活動において大きな役割を果たしている。今後、新たなサービスの登場も予期される場所である。
- しかし、インターネット上では、依然として、違法な情報や有害な情報の流通も認められ、昨今、特定の個人に対して多くの誹謗中傷の書き込みが行われるいわゆる「炎上」事案が発生したり、震災や新型コロナウイルス感染症などの社会不安に起因する誹謗中傷が行われるなど、特にSNS上での誹謗中傷等の深刻化が問題となっている。
- これらの状況を踏まえ、当研究会において、2020年8月に「インターネット上の誹謗中傷への対応の在り方に関する緊急提言」（以下「緊急提言」という。）を公表し、これを受けて、総務省において2020年9月に「インターネット上の誹謗中傷への対応に関する政策パッケージ」（以下「政策パッケージ」という。）を策定・公表した。
- また、インターネット上の誹謗中傷への対策について、プラットフォーム事業者等の取組が十分か、また、その透明性・アカウントビリティが十分果たされているかを検証するために、2021年、当研究会において、日本国内でサービスを展開するプラットフォーム事業者等に対してモニタリングを実施し、プラットフォーム事業者等における自主的な誹謗中傷に関する対応状況と、それらの取組に関する透明性・アカウントビリティの確保状況の把握を図ったところ、プラットフォーム事業者の誹謗中傷への対応に関する透明性・アカウントビリティ確保状況には差異が見られた。

## 論点（前提）（続き）

- モニタリングの結果を踏まえ、2021年9月に公表した「プラットフォームサービスに関する研究会 中間とりまとめ」（以下「中間とりまとめ」という。）では、個別の投稿の削除を義務づけることには極めて慎重であるべきとする一方、我が国における透明性・アカウントビリティ確保が実質的に図られない場合には、透明性・アカウントビリティの確保方策に関する行動規範の策定及び遵守の求めや法的枠組みの導入等の行政からの一定の関与について、具体的に検討を行うことが必要ととりまとめられた。
- また、行政からの一定の関与の検討に際しては、①リスクベースアプローチ、②リスク分析・評価と結果公表、③適切な対応の実施と効果の公表、④継続的モニタリング、⑤データ提供といった大枠としての共同規制的枠組みの構築を前提に検討を進めることが適当とされた。
- 上記の経緯を前提に2022年に実施したモニタリングの結果を踏まえ、プラットフォーム事業者の誹謗中傷対策に関する取組は十分か、また、その透明性・アカウントビリティは十分確保されているか。 中間とりまとめ、緊急提言や政策パッケージを踏まえ、産学官民の取組は適切に進められているか。

- 社会問題となっているインターネット上の誹謗中傷に対応するため、総務省において、2020年9月に「政策パッケージ」を公表。

## 1. ユーザに対する情報モラル及びICTリテラシーの向上のための啓発活動

- ①「インターネットトラブル事例集(2020年版)追補版」を作成・公表し、全国の総合通信局等や教育委員会等を通じて子育てや教育の現場へ周知【2020年9月公表・周知済】
- ②「e-ネットキャラバン」の講座内容にインターネット上の誹謗中傷に関するものを追加【2020年9月実施済】
- ③「#NoHeartNoSNS 特設サイト」(主催:総務省・法務省人権擁護局・(一社)SMAJ・(一社)SIA)の拡充等により、社会全体における情報モラルやICTリテラシーが高まるようにするための取組を強化【継続的に実施中】

## 2. プラットフォーム事業者の自主的取組の支援と透明性・アカウントビリティの向上

- ①実務者検討会を開催し、法務省人権擁護機関からの削除依頼に対する事業者の円滑な対応を促進【定期的で開催中】
- ②事業者及び事業者団体との意見交換を通じ、誹謗中傷対策の実施や有効性の検討を働きかけ【継続的に実施中】
- ③自主的な取組の報告等により、事業者による透明性・アカウントビリティ確保方を促進し、取組の状況把握や評価方法の検討を実施【PF研等の場を通じ継続的に実施】
- ④国際的な制度枠組みや対応状況を注視し、国際的な対話を深化【継続的に実施中】

## 3. 発信者情報開示に関する取組

- ①電話番号を開示対象に追加する省令改正の実施を踏まえ、弁護士会照会に応じて電話番号に紐付く氏名・住所を回答可能である旨をガイドラインで明確化【2020年11月実施済】
- ②新たな裁判手続の創設や特定の通信ログの早期保全のための方策について、法改正を実施【2021年4月成立、2022年10月施行予定】
- ③開示対象となるログイン時情報を明確化するため、法改正を実施【2021年4月成立、2022年10月施行予定】
- ④要件該当性の判断に資する民間相談機関の設置やガイドラインの充実に関する民間の取組を支援【2021年4月ガイドライン公表】

## 4. 相談対応の充実に向けた連携と体制整備

- ①違法・有害情報相談センターについて、相談員の増員等による体制強化を図るとともに、相談件数・内容の分析を実施【2021年度から実施中】
- ②相談内容に応じて相談機関間で紹介を行うなど、他の相談機関との連携対応を充実【継続的に実施中】
- ③複数の相談窓口の特徴やメリットを記載した案内図の作成など、ユーザにとって分かりやすい相談窓口の案内を実施【2020年12月公表済】

## 1. ユーザに対する情報モラル及びICTリテラシーの向上のための啓発活動

- ・実態把握や分析結果に基づき、産学官民が連携し、引き続きICTリテラシー向上施策が効果的となるよう取り組む必要
- ・e-ネットキャラバン等の青少年向けの取組に加え、大人も含め幅広い対象に対してICTリテラシー向上のための取組を実施することを検討していくことが必要
- ・総務省や各ステークホルダーによるICTリテラシー向上の取組状況を把握し、ベストプラクティスの共有や更なる効果的な啓発を行うことが必要

## 2. プラットフォーム事業者の自主的取組の支援と透明性・アカウントビリティの向上

### <プラットフォーム事業者の自主的取組の支援>

- ・プラットフォーム事業者が自らのサービス上での違法・有害情報の流通状況について実態把握とリスク分析・評価を行うことが必要
- ・トラステッドフラグガーの仕組みの導入・推進にむけて検討を行うことが望ましい。人権擁護機関からの削除要請に関し、削除に関する違法性の判断基準・判断方法や個別の事業者における削除実績等について関係者間で共有し、行政側・事業者側双方の削除に関する対応についての透明性を向上させ、円滑な削除対応を促進することが必要
- ・プラットフォーム事業者は、一定の短期間の間に大量の誹謗中傷が集まった場合へのアーキテクチャ上の工夫について、既存の機能や取組の検証や新たな対応の検討を行うことが望ましい

### <プラットフォーム事業者による取組の透明性・アカウントビリティの向上>

- ・我が国における透明性・アカウントビリティ確保が図られていない事業者に関しては、特に透明性・アカウントビリティ確保の取組を進めることが強く求められる。総務省はモニタリングと検証評価を継続的に行っていくことが必要

## 3. 発信者情報開示に関する取組

- ・法施行に向けて関係政省令の策定を進め、関係事業者及び総務省の間で新制度の具体的な運用に関する協議を進めることが必要
- ・プラットフォーム事業者・行政側の双方で、発信者情報開示に関する申請や開示件数等について集計・公開することが求められる

## 4. 相談対応の充実に向けた連携と体制整備

- ・違法・有害情報相談センターにおいて引き続き相談対応を行い、システム更新、相談機関間の連携強化、相談窓口の周知が必要



- ・プラットフォーム事業者による取り組みの透明性・アカウントビリティ確保について、次回以降のモニタリングにおいて、依然として事業者が自主的な報告を行わない場合や、我が国における透明性・アカウントビリティ確保が実質的に図られない場合には、透明性・アカウントビリティの確保方針に関する行動規範の策定及び遵守の求めや法的枠組みの導入等の行政からの一定の関与について、具体的に検討を行うことが必要
- ・①リスクベースアプローチ、②リスク分析・評価と結果公表、③適切な対応の実施と効果の公表、④継続的モニタリング、⑤データ提供、といった大枠としての共同規制的枠組みの構築を前提に検討を進めることが適当

## 論点

- 我が国におけるインターネット上の誹謗中傷の実態が適切に把握されているか。具体的にどのような情報を把握すれば誹謗中傷の流通状況を明らかにすることが可能か。
- 総務省が運営を委託している違法・有害情報相談センターで受け付けている相談件数は高止まり傾向にあり、令和2年度の相談件数は、受付を開始した平成22年度の相談件数の約4倍に増加している。令和3年度の相談件数は例年より増加し、6千件を上回った。令和3年度において相談件数が多い事業者/サービス上位5者は、Twitter、Google、Meta、5ちゃんねる、LINE。
- 法務省が相談等を通じて調査救済手続を開始したインターネット上の人権侵害情報に関する人権侵害事件は、引き続き高水準で推移している。法務省は、インターネット上の人権侵害情報について、法務省の人権擁護機関による削除要請件数と削除対応率のサイト別の数値を令和4年に初めて公表した。平成31年1月～令和3年10月の期間内に、人権侵害事件として処理されたのは5,136件であり、そのうち、法務局において、当該情報の違法性を判断した上で、実際に削除要請を実施した件数の合計は1,173件、削除対応率は69.74%。さらに、投稿の類型別（私事性的画像情報、プライバシー侵害、名誉毀損、識別情報の摘示）の削除要請件数及び削除対応率についても公表を行った。
- 一般社団法人セーファーインターネット協会（以下「SIA」という。）が運営する誹謗中傷ホットラインへの連絡件数について、2021年1月1日から12月31日までの受領件数は、2,859件（1,516名）であった。また、サイトの属性別には、SNSが最多の28%であり、次いで、匿名掲示板が19%、地域掲示板が7%であった。

論点

<MRI発表（第36回会合）>

- SNSユーザーを対象としたアンケート調査結果では、「他人を傷つけるような投稿（誹謗中傷）」について、約半数（50.1%）が目撃しており、投稿を目撃したサービスとしては、Twitter（52.6%）が最も多く、匿名掲示板（39.7%）、Yahoo!コメント（32.0%）、YouTube（28.2%）がこれに続いている。また、過去1年間にSNS等を利用した人の1割弱（8.9%）が「他人を傷つけるような投稿（誹謗中傷）」の被害に遭っており、年代別に見ると20代で最も多く（16.4%）、10代及び30代がともに1割強でこれに次ぐ。40代以上は相対的に少なかった。

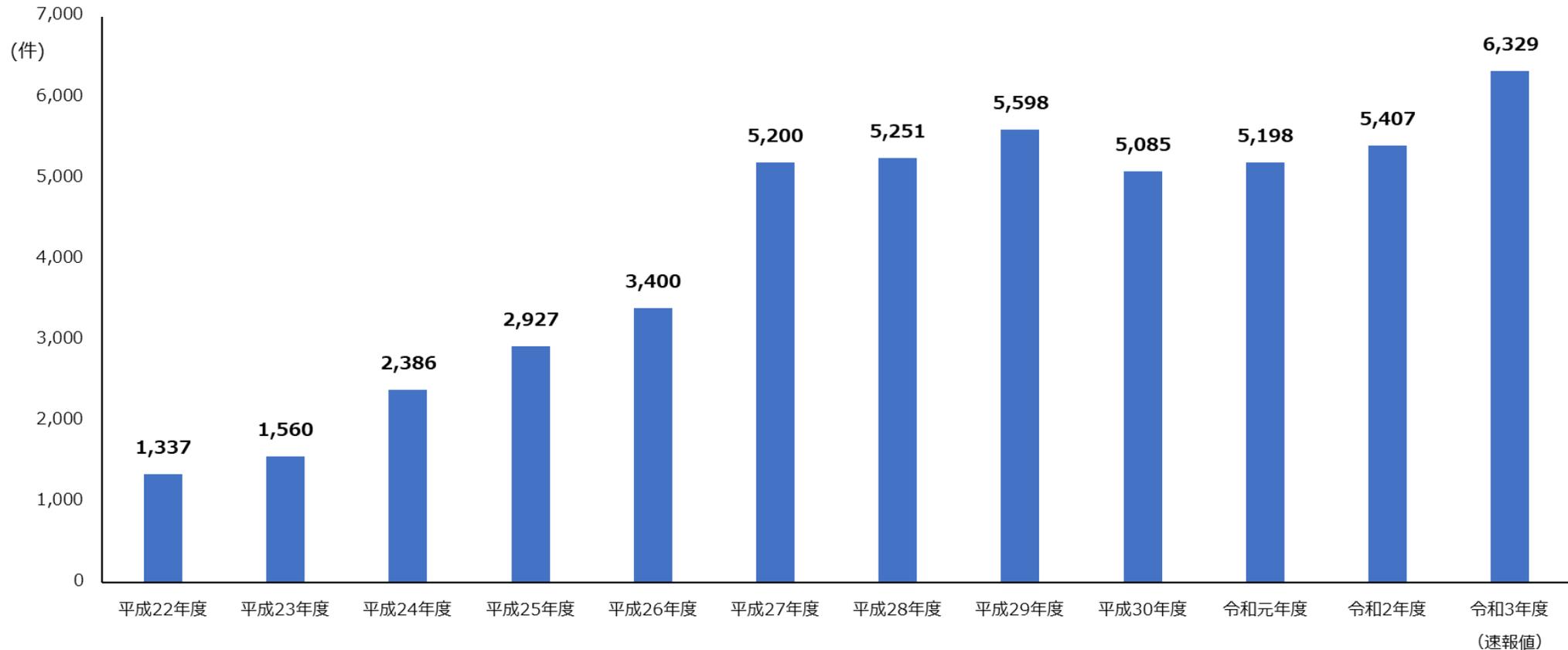
<UUUM発表（第35回会合）>

- クリエイターが受ける誹謗中傷被害について、具体性のある身体生命の加害投稿等が見られる他、侮辱、名誉毀損、プライバシー侵害、著作権侵害が6割から7割程度と大半を占める。中には、法的責任を超えた不当要求も見られる。
- 被害を受ける場としては、クリエイターが活動するSNSプラットフォームがその場所の一つとなっている。

## 違法・有害情報相談センターにおける相談件数の増加

- ・ 違法・有害情報相談センターで受け付けている相談件数は増加傾向にあり、令和2年度の相談件数は、平成22年度の相談件数の約4倍に増加している。
- ・ 速報値であるものの、令和3年度の相談件数は、6,000件を上回った。

違法・有害情報相談センターにおける相談件数の推移 <平成22年度～令和3年度（速報値）>

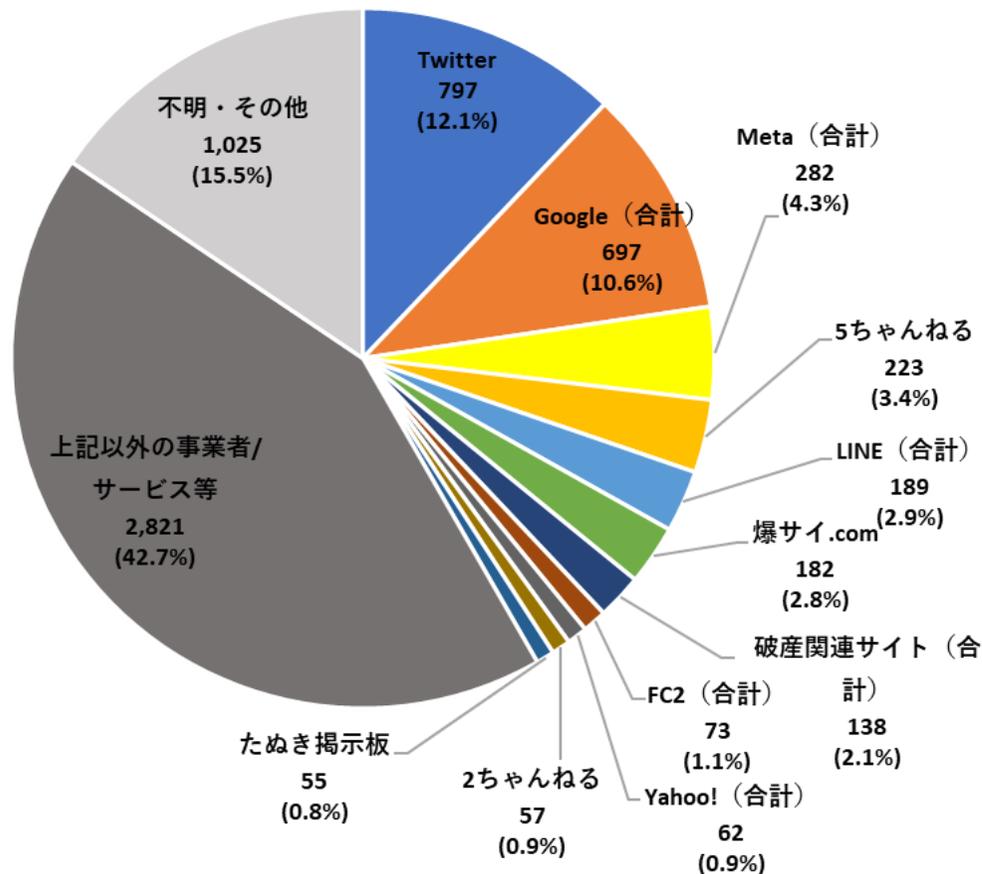


## (3) 相談(作業)件数の内訳について

相談(作業)件数の内訳: 事業者/サービス別 (n=6,601)

&lt;令和3年度(速報値)&gt;

※相談(作業)件数 6,329件を対象



事業者/サービス名等	件数	割合	
Twitter	797	12.1%	
Google (合計)	検索	450	6.8%
	YouTube	112	1.7%
	map	108	1.6%
	その他	27	0.4%
Meta (合計)	instagram	216	3.3%
	facebook	66	1.0%
5ちゃんねる	223	3.4%	
LINE(合計)	livedoorサービス	128	1.9%
	LINEアプリ内サービス	61	0.9%
	爆サイ.com	182	2.8%
破産情報掲載関連サイト(合計)	138	2.1%	
FC2(合計)	73	1.1%	
Yahoo!(合計)	オークション	14	0.2%
	ニュース	13	0.2%
	知恵袋	12	0.2%
	検索	9	0.1%
	その他	14	0.2%
2ちゃんねる	57	0.9%	
ためき掲示板	55	0.8%	
上記以外の事業者/サービス等	2,821	42.7%	
その他・不明	1,025	15.5%	

注1 相談(作業)件数を集計したものであり、個別の相談が権利侵害にあたるか相談センターでは判断していない。

注2 作業件数につき、複数のサービスを回答する場合もあるため、作業件数と上記グラフの総計が一致しない。

注3 独自ドメインを利用しているものがあり、実際のドメインが判明しない場合がある。

## 【インターネット上の人権侵害情報】法務省の人権擁護機関による削除要請と削除対応率 (サイト別)

番号	サイト名	(種別)	要請件数 ○+△+× =□(件)	削除合計 ○+△(件)		削除せず ×(件)	全部削除率 ○/□(率)	削除対応率 ○+△/□(率)	
				全部削除 ○(件)	一部削除 △(件)				
1	FC2	ブログ	31	31	29	2	0	93.55%	100.00%
2	2ちゃんねるブックマーク	掲示板のコピーサイト	19	19	18	1	0	94.74%	100.00%
3	爆サイ.com	掲示板	134	132	124	8	2	92.54%	98.51%
4	2ch勢いランキング	掲示板のコピーサイト	27	26	26	0	1	96.30%	96.30%
5	ログ速	掲示板のコピーサイト	21	20	20	0	1	95.24%	95.24%
6	Imgur	画像・動画の共有サイト	15	12	12	0	3	80.00%	80.00%
7	ライブドアブログ	ブログ	24	19	17	2	5	70.83%	79.17%
8	2ちゃんねる(2ch.sc)	掲示板	78	53	41	12	25	52.56%	67.95%
9	Yahoo!知恵袋	Q & A サイト	21	13	9	4	8	42.86%	61.90%
10	ホストラブ	掲示板	13	8	6	2	5	46.15%	61.54%
11	Twitter	SNS	86	29	24	5	57	27.91%	33.72%
12	5ちゃんねる	掲示板	80	21	17	4	59	21.25%	26.25%
13	YouTube	画像・動画の共有サイト	33	8	6	2	25	18.18%	24.24%
14	2ch2.net	掲示板	14	2	1	1	12	7.14%	14.29%
15	その他		577	425	403	22	152	69.84%	73.66%
	全体		1173	818	753	65	355	64.19%	69.74%

※ 件数は、個別のプロバイダ等に対する削除要請の件数であり、個別の投稿の件数ではない。通例は、同一の被害者について、特定のサイト等に複数の人権侵害性のある投稿がなされ、そのような複数の投稿について、まとめてプロバイダ等に削除要請を行うところ、このようなプロバイダ等1社に対する要請1回を1件としてカウントしたもの。このうち、全部が削除された場合を「全部削除」、一部が削除されたものを「一部削除」とし、その合計を要請件数で除した数値を削除対応率として示した。

※ 削除には、被害者や地方公共団体による削除依頼に基づく削除のほか、投稿者による自主的な削除もある。

※ 法務省の人権擁護機関による削除要請と削除との条件関係は、厳密に特定できるものではない。

※ 対象期間は、平成31年1月～令和3年10月。対象期間中に処理を終えた要請件数が10件以上のサイト名を掲げた（閉鎖が確認された破産者情報サイトを除く。）。

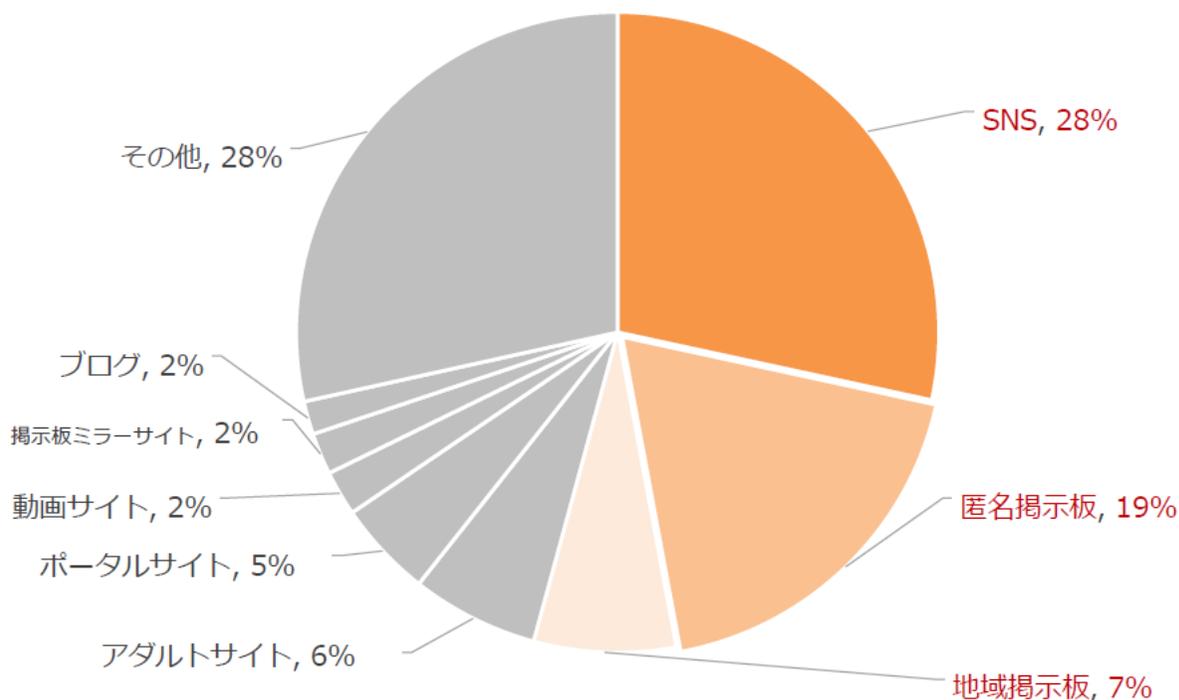


## 受領件数

期間：2021/1/1-12/31

受領件数は**2,859件 (1,516名)**。

サイトの属性では、**SNSが最多 (28%)**。次いで匿名掲示板 (19%)、地域掲示板 (7%)。サイト別では、大手短文投稿サイトが17%を占め、上位5サイトで全体の44%を占める。



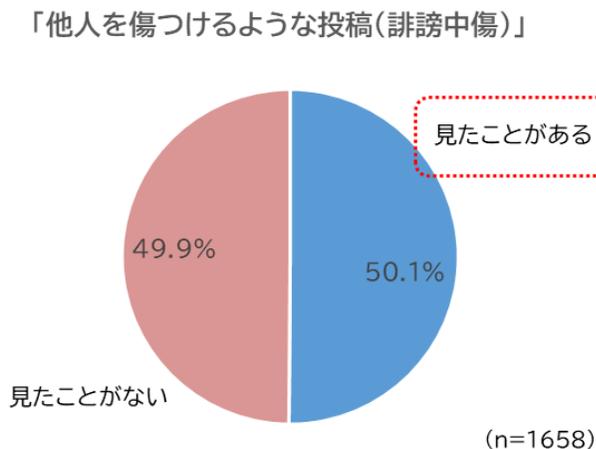
### サイト別 (上位5サイト)

サイト	件数	割合
大手短文投稿サイト	484件	<b>17%</b>
匿名掲示板1	228件	8%
地域掲示板	204件	7%
匿名掲示板2	194件	7%
アダルトサイト	154件	5%
		<b>44%</b>

- 調査実施期間: 2022年3月16日(水)~3月18日(金)
- 調査方法: インターネットアンケート調査
- 調査対象: 日本国内の15歳から99歳までの男女
- 回答者数: 2,000名
- 回答割付: 回答者の性別・年代による回答に特定の傾向がある場合に、その影響を少なくするために、(均等割付でなく)インターネットの利用者の性別・年代別の分布に即して回答者数を割り付けた: 具体的には、1) 住民基本台帳人口(2021年1月1日付)の性別・年齢階級別人口に、通信利用動向調査(令和2年度)のインターネット利用率を掛け合わせてインターネット利用人口を算出。2) 回収総数を2,000人とし、算出したインターネット利用人口の構成比に基づいて、性別・年代別の回収数を割付

## 誹謗中傷等に関する投稿の目撃経験

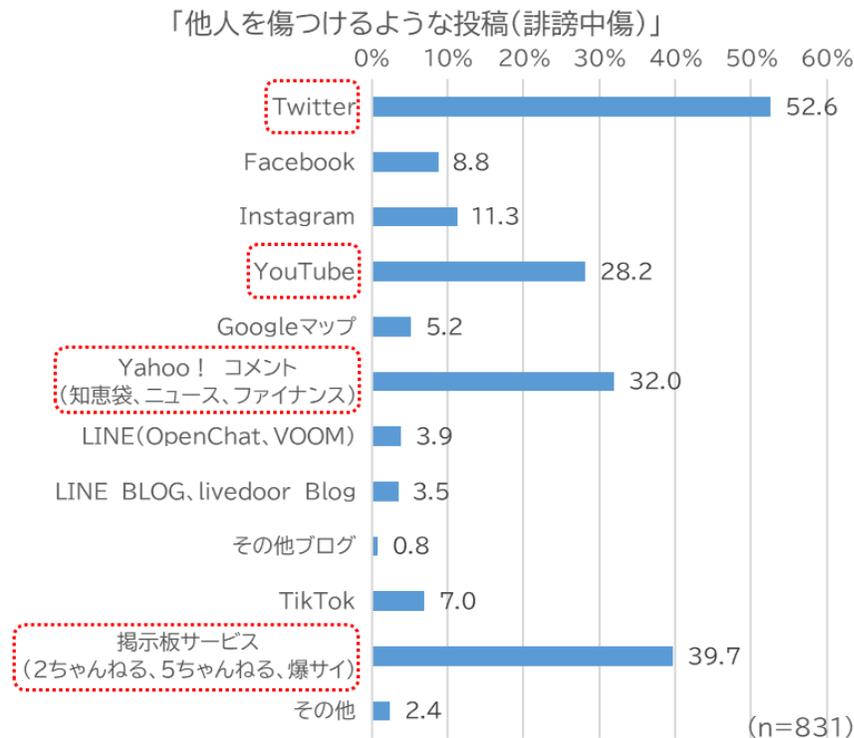
Q あなたはインターネット上で次のような投稿を**見たこと**がありますか。以下の投稿について、それぞれ「見たことがある」又は「見たことがない」のいずれかで教えてください。(SA)



※ 過去1年間にいずれかのSNSなどのサービスを利用したと答えた回答者を抽出して集計

## 目撃した際のサービス

Q 前問で「見たことがある」と回答した投稿について、どのサービスで見ましたか。(MA)



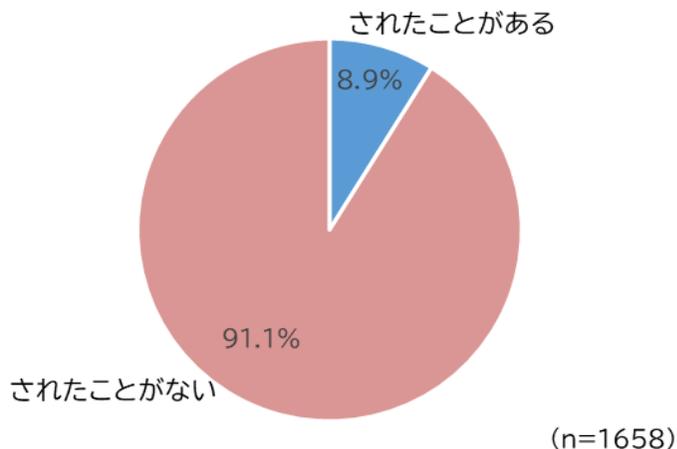
※「掲示板サービス」は、( )内の各サービスでの目撃経験の合計。

- 調査実施期間: 2022年3月16日(水)~3月18日(金)
- 調査方法: インターネットアンケート調査
- 調査対象: 日本国内の15歳から99歳までの男女
- 回答者数: 2,000名
- 回答割付: 回答者の性別・年代による回答に特定の傾向がある場合に、その影響を少なくするために、(均等割付でなく)インターネットの利用者の性別・年代別の分布に即して回答者数を割り付けた: 具体的には、1) 住民基本台帳人口(2021年1月1日付)の性別・年齢階級別人口に、通信利用動向調査(令和2年度)のインターネット利用率を掛け合わせてインターネット利用人口を算出。2) 回収総数を2,000人とし、算出したインターネット利用人口の構成比に基づいて、性別・年代別の回収数を割付

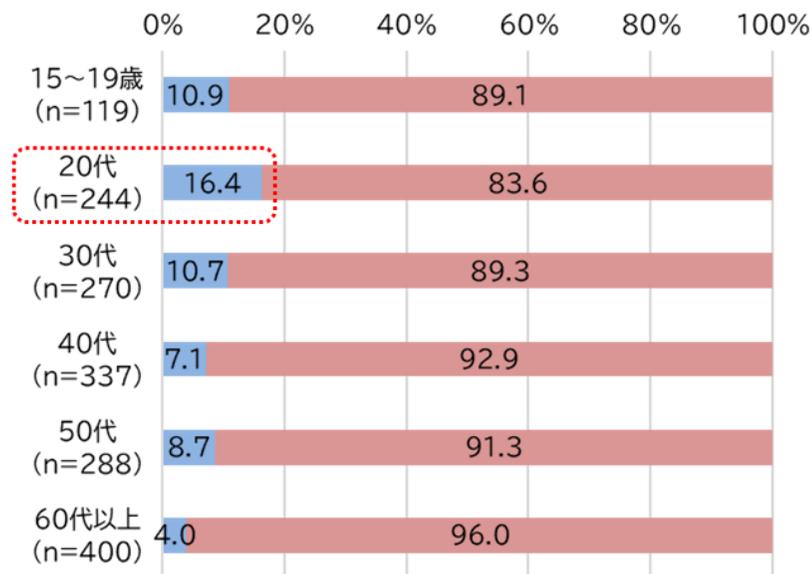
## 誹謗中傷等に関する投稿の被害経験

Q あなたはインターネット上で次のような投稿をされたことがありますか。以下の投稿について、「されたことがある」又は「されたことがない」のいずれかでお答えください。(SA)

### 「他人を傷つけるような投稿(誹謗中傷)」



### 他人を傷つけるような投稿(誹謗中傷)・年代別



※ 過去1年間にいずれかのSNSなどのサービスを利用したと答えた回答者を抽出して集計

■ されたことがある ■ されたことがない (n=1658)  
(第36回会合 資料5より抜粋)

## 論点

## ＜政策パッケージの進捗状況について＞

## 【1. ユーザに対する情報モラル及びICTリテラシーの向上のための啓発活動】

## ○ 政府及び各ステークホルダーにおける啓発活動は十分か。

- 総務省では、ユーザに対する情報モラル及びICTリテラシーの向上のための様々な啓発活動を実施。
- 2022年3月にインターネット上の誹謗中傷に関する注意事項等をまとめた「インターネットトラブル事例集(2022年版)」を作成・公表し、学校等への周知を実施。2022年版では、SNS等での誹謗中傷に関するトラブル事例に加え、誹謗中傷を受けた場合に相談できる窓口のフローチャートを掲載。関係省庁・関係団体との連携のもと、様々なチャネルを活用し、全国の学校、SNSユーザ等へ周知を実施。
- 学校等での出前講座「e-ネットキャラバン」の講座内容にインターネット上の誹謗中傷に関するものを追加。2021年度は、2,559件の講座を実施し、約40万人が受講。
- SNSで悩んでしまった際に役立ててもらうため、「#NoHeartNoSNS 特設サイト」を開設し、周知を実施。（主催：総務省、法務省人権擁護局、一般社団法人ソーシャルメディア利用環境整備機構、一般社団法人セーフインターネット協会、協力団体：電気通信事業者団体等7団体）
- 2021年12月に#NoHeartNoSNSの取組の一環として、『秘密結社 鷹の爪』とコラボし、「鷹の爪団の#NoHeartNoSNS大作戦」特設サイトを開設。総務省の広報ツール、広告配信、都内の駅前大型ビジョンでの動画放映、関係省庁・関係団体と連携したSNS投稿など、様々なチャネルを活用して周知を実施。

## ① インターネットトラブル事例集

子育てや教育の現場での保護者や教職員の活用に資する、インターネットに係る誹謗中傷等のトラブル事例とその予防法等をまとめた事例集。2009年度より毎年更新・作成し公表。総合通信局等や教育委員会等を通じて、全国の学校等への周知を実施。

インターネットトラブル事例集(2022年版)(抜粋)

◆情報発信

### 16 SNS等での誹謗中傷による慰謝料請求

有名人の顔口を再投稿・拡散したら

発信者として特定され慰謝料請求

大嫌いなアイドルの顔口をSNSで再投稿したら、本人から慰謝料請求された。

虚偽の投稿内容による苦情をきっかけに、本人の投稿者だけでなく再投稿・拡散した本人も慰謝料請求された。

考えてみよう！

いらい立ちを覚えたり、自分の中の正義感が高くなり、過激な投稿で個人攻撃をする人がいます。こうした加害行為(再投稿)も含まれるのを念頭に注意したいことは？

A. 誹謗中傷≠批判意見  
ほとんどのSNSには「誹謗中傷禁止」の利用規約があります。相手の人格を否定する言葉や言い回しは、批判ではなく誹謗中傷。その違いを正しく見極め、必要に応じて再投稿しないです。

B. 匿名性による気のゆるみ  
対面や実名では言えないのに、匿名だと言えたため攻撃性が強まりました。たとえ匿名でも、投稿の発信者を特定できる仕組みがあり、民事上・刑事上の責任を問われる可能性があります。

C. カカとわっても立ち止まらなくていい  
怒りは人の自然な感情ですが、はけに吐き出すのはSNS。炎上したのが悔しかったから「みんな投稿しなればよかった」と悔みながら投稿は控えてください。

「目立つ存在だから仕方がない」という主張は通用しない

SNS上で、悪意を感じる投稿を見かけるといいますが、中には「正義感からやめたことと主張する人もいますが、「立憲」や「事実かどうか」を問わず、人格を否定するような投稿は正解ではありません。近年、YouTubeやTwitterなどで、事件・事故の現場写真、被害者の顔や名前を無断で公開するなどの主体的投稿をする以外の「炎上」に再投稿・拡散する人も増えています。本人の同意が得られれば、集団攻撃となり人を傷つける可能性があります。たとえ相手がどんな人であっても、民事上・刑事上の責任(損害賠償請求、名誉毀損による慰謝料、侮辱罪)による罰則(罰金、懲役)の可能性があり、この点に十分に注意して、誹謗や利用規約等のルールやマナーを尊重し、正しい利用をお願いします。

ワンポイントアドバイス

誹謗中傷は、再投稿でも「広めること」に加担したとみなされます。投稿・再投稿する前に「自分が言われたらどう思うか」を考えて！

29

● 法務省「インターネット人権相談受付窓口」 <https://www.jinken.go.jp/>  
● 厚生労働省「まもろうとここ」 <https://www.mhlw.go.jp/mamoroukokoro/>

発信者は、#NoHeartNoSNS「ハートがなければSNSじゃない！」をスローガンとするインターネット上の誹謗中傷対策の一環として、人気キャラクター「魔法少女まどかまろ」の「魔法少女まどかまろ」の#NoHeartNoSNS大発表「特設サイト」を開設しています。(2022年12月20日開催予定)

● インターネット上の誹謗中傷への対策 ◎ (総務省サイト内/相談フローも掲載) 報道するさまざまな資料や情報へのリンクがまとめられています。併せて活用ください！>>> [https://www.soumu.go.jp/main\\_sosaku/joho\\_tsusinai\\_syohi/hiboutyusyou.html](https://www.soumu.go.jp/main_sosaku/joho_tsusinai_syohi/hiboutyusyou.html)

## ② e-ネットキャラバン

誹謗中傷を含む、子どもたちのインターネットの安全な利用に係る普及啓発を目的として、児童・生徒、保護者・教職員等に対する、学校等の現場での「出前講座」。情報通信分野等の企業・団体と総務省・文部科学省が協力して全国で開催。

e-ネットキャラバン教材(抜粋)

危険4個人情報漏えい

### e-ネットキャラバン

## 写真や動画での個人情報漏えい

写真や動画の共有や生中継で簡単に漏えい。さらに、過激な投稿の要求や誹謗中傷に繋がることも。

そのほか、情報通信自由があるといわれてきた身近な人に個人情報を提供して利用されている。

そのほか、投稿者によって個人情報が漏えいする可能性があります。事前に「正義感からやめたことと主張する人もいますが、「立憲」や「事実かどうか」を問わず、人格を否定するような投稿は正解ではありません。近年、YouTubeやTwitterなどで、事件・事故の現場写真、被害者の顔や名前を無断で公開するなどの主体的投稿をする以外の「炎上」に再投稿・拡散する人も増えています。本人の同意が得られれば、集団攻撃となり人を傷つける可能性があります。たとえ相手がどんな人であっても、民事上・刑事上の責任(損害賠償請求、名誉毀損による慰謝料、侮辱罪)による罰則(罰金、懲役)の可能性があり、この点に十分に注意して、誹謗や利用規約等のルールやマナーを尊重し、正しい利用をお願いします。

e-ネットキャラバン講座実施の様子



## ③ #NoHeartNoSNS

法務省・SNS事業者団体(SMAJ)と共同で2020年7月に特設サイトを開設。2020年12月には新たに電気通信事業者団体等(7団体)を協力団体として追加し、2021年2月にはインターネット関連事業者団体(SIA)を主催団体として追加。

#NoHeartNoSNS特設サイト(抜粋)

ハートがなければSNSじゃない!

# #NoHeartNoSNS

SNSはハートをつなげるもの  
誰かを傷つけるためにあるんじゃない!

#NoHeartNoSNS #傷つけてしまいそうなら

もし、あなたが誰かを傷つけてしまいそうなら  
覚えておいてほしい言葉は刃(やいば)にもなり  
集まれば銃弾の雨にもなって誰かの全てを奪ってしまうこともある。  
SNSであなたがしたかったことは何ですか？

NoHeartNoSNS #傷ついたなら

もし、あなたが傷ついたなら  
あなたを守る方法があることを  
覚えておいて欲しい。

○ 「鷹の爪団の#NoHeartNoSNS大作戦」のコンテンツについては、総務省の広報ツール、広告配信、大型ビジョンでの動画放映、関係省庁・関係団体と連携したSNS投稿など、様々なチャネルを活用して周知を実施。

## ○ 総務省の周知広報

- ・総務省の公式Twitter(総務省・ICTツイート)、公式Facebookにて継続的に周知
- ・総務省広報誌(令和4年2月号)に掲載 ※全国の都道府県・市区町村等に対して発行
- ・総務省内に設置された電子パネルにメインビジュアルを表示

<総務省公式Twitterでの投稿>



<広告配信>



## ○ 広告配信

- ・令和4年1月1日～3月10日の間、Twitter・YouTubeにて広告を配信

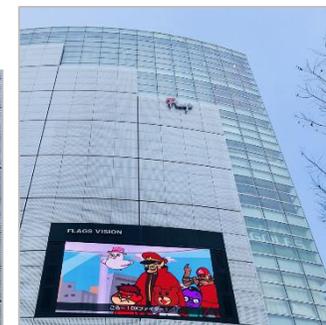
## ○ 都内の大型ビジョンでの動画放映

- ・令和4年2月11日～17日の間、新宿南口エリアの大型ビジョンにて動画を放映(今後も実施予定)

<大型ビジョンでの動画放映>

## ○ 関係省庁・関係団体と連携した周知広報

- ・全国の総合通信局・総合通信事務所:
  - 管内のe-ネットキャラバン主催者、講師及び受講者に周知、
  - 管内会議・イベントにて資料やパンフレット配布、SNSでの投稿、
  - スポーツ試合での動画放映
- ・関係団体・事業者:
  - ホームページにリンク掲載、SNSでの投稿、会員向けに情報発信



これまでの主な意見

【普及啓発活動について】

- 特に啓発がどの程度効果があるのかがよく分からない。実際にどういった対策を行うと、誰のどのような意識や行動を変容させることができるのか。例えば、投稿に慎重になったや、被害を受けても毅然と対応できるようになったと考えると、いった行動変容、意識変容を喚起させていくためには、どういった対策を取る場合が一番効果があるのかが見えてくるのではないか。今後はこれまでやってきたことに対して効果のKPIをしっかり立てる等、どのような効果があるのかについて分析をしっかりとやっていく必要があるのではないか。【寺田構成員（第32回会合）】
- リテラシーについての教育は、小さい頃からやっておくべき内容だと強く感じる。個人情報を取り扱う意味からも必要であるし、倫理観や道徳観に根差しているものにもつながることだと思う。全国的な展開をするという流れは大事だが、やはり授業の中でこういう項目を全国的に小中学生にも情報系の教育をしっかりやっていく時代に完全に入っている。ネットワークを使うと、一般に会話している内容とは全く違って、拡散の問題といった全体的なことを踏まえて、書き込む際の慎重さや判断の仕方を小さい頃から知っていただくということが今後非常に重要になる。世代別の普及啓発をどういうふうに今後組み立てていくのか、この分野においてどう広げていったらいいのかというところが重要だと感じた。【手塚構成員（第32回会合）】

論点

【2. プラットフォーム事業者の自主的取組の支援と透明性・アカウントビリティの向上】

○ 法務省による削除要請に対してプラットフォーム事業者は適切に対応しているか。

- 法務省の人権擁護機関は、インターネット上の人権侵害事案において、名誉毀損、プライバシー侵害等に該当する場合には、被害者からの人権相談を端緒として削除要請をサイト運営者に対して行っている。
- 総務省は法務省と連携し、サイト運営者による迅速な削除等の対応が行われる環境を整備するため、大手海外事業者等のPF事業者や業界団体との意見交換の場となる実務者検討会を継続的に開催。法務省の人権擁護機関への相談事案に関する認識について関係者間で共有を図るとともに、法務省の人権擁護機関からの申告について透明化の向上を行い、サイト運営者による円滑な対応を促進。
- 法務省は、インターネット上の人権侵害情報について、法務省の人権擁護機関による削除要請件数と削除対応率のサイト別の数値を令和4年に初めて公表した。平成31年1月～令和3年10月の期間内に、人権侵犯事件として処理されたのが5,136件。そのうち、法務局において、当該情報の違法性を判断した上で、実際に削除要請を実施した件数の合計は1,173件、削除対応率は69.74%。さらに、投稿の類型別（私事性的画像情報、プライバシー侵害、名誉毀損、識別情報の摘示）の削除要請件数及び削除対応率についても公表を行った。  
（再掲）
- 法務省人権擁護局は、公益社団法人商事法務研究会が主催し、2021年4月から開催されている「インターネット上の誹謗中傷をめぐる法的問題に関する有識者検討会」に参加し、削除要請に関する違法性の判断基準や判断方法等の議論に積極的に関与。議論の結果は、5月に公表された。同有識者検討会には、総務省も参加。

これまでの主な意見

【法務省人権擁護機関による削除要請について】

- 人権擁護機関の削除要請の中で、被差別部落等の識別情報の摘示の問題があった。その中では権利侵害情報になるものと、そうでないものが出てきてしまう。これを海外のプラットフォームに説明できるのかどうか。もちろんプラットフォームとしては、表現の自由と権利侵害を天秤にかけて判断しなければいけないが、我が国の特殊な被差別部落の人権侵害とその対策の歴史をなかなか説明しにくい。プラットフォームとしても「はい、分かりました、消します」というわけには当然いけないと思う。権利侵害を理由に削除を要請するという構成とは別に、部落差別の解消の推進に関する法律のような公法によって削除権限を持っていただくという考え方もある。【森構成員（第32回会合）】

論点

【2. プラットフォーム事業者の自主的取組の支援と透明性・アカウントビリティの向上（続き）】

○ 民間の相談機関からの削除要請に対してプラットフォーム事業者は適切に対応しているか。

- SIAは、2020年6月29日より、「誹謗中傷ホットライン」を運用開始。インターネット上で誹謗中傷を受けている被害者からの連絡を受け、コンテンツ提供事業者に、各社の利用規約に基づき削除等の対応を促す通知を行っている。
- 立場の弱い私人や個人商店等が対象。コロナ罹患者やアスリートに対する情報も対象。
- SIAが運営する誹謗中傷ホットラインへの連絡件数について、2021年1月1日から12月31日までの受領件数は、2,859件。サイトの属性別には、SNSが最多の28%、次いで、匿名掲示板が19%、地域掲示板が7%であった。（再掲）
- そのうち、ガイドラインに基づき削除通知対象となる「特定誹謗中傷情報」に該当するものが796件（27.8%）、非該当が2,063件（72.2%）
- 1,414URLに対して削除等の対応を促す通知を行った結果、一週間後に削除が確認されたものは1,046URL（削除率74%）

論点

【2. プラットフォーム事業者の自主的取組の支援と透明性・アカウントビリティの向上（続き）】

○ 誹謗中傷等への対応に関して、国際的な対話が深められているか。

○ 2022年2月3日に、第27回日EU・ICT政策対話を実施。また、2021年11月17日に第12回、2022年4月8日に第13回の日EU・ICT戦略ワークショップを実施。日本側からは、「インターネット上の誹謗中傷への対応に関する政策パッケージ」及び中間とりまとめについて紹介。EU側からは、Digital Services Act(DSA: デジタルサービス法)及びDigital Market Act(DMA: デジタル市場法)の概要や欧州内での議論状況等について紹介があった。

○ 2021年11月11日に、第12回インターネットエコノミーに関する日米政策協力対話（政府間会合）を実施。日本側からは、「インターネット上の誹謗中傷への対応に関する政策パッケージ」及び中間とりまとめについて紹介。米国側とインターネット上の違法・有害情報対策についての問題意識を共有した。

○ 2022年3月23日には、第6回日独ICT政策対話を実施。日本側からは、インターネット上の誹謗中傷及び偽情報に関する政策動向を紹介。

○ 2022年4月28日には、「未来のインターネットに関する宣言」立ち上げイベントが開催され、開かれた、自由で、グローバルで、相互運用可能で、信頼性のある、安全な、一つのインターネットへの支持を呼びかけられた。その一環として、違法・有害情報対策における人権及び基本的自由の保護※について確認された。

※ 政府、規制当局、およびオンラインプラットフォームを含むデジタルサービスにより、オンライン上の違法・有害なコンテンツや活動を削減するために取られる行動は、検閲、嫌がらせ、または脅迫の恐れなしに、意見の多様性と多元性を奨励しながら、表現の自由の権利を含む国際人権法に合致するべきであるという我々のコミットメントを再確認する。

○ 2022年5月には、金子総務大臣が欧州委員会ヴェステアー上級副委員長及びブルトン委員、英国DCMSナディーン・ドリーズ大臣と会談し、デジタル空間における利用者保護について意見交換を行い、協力関係の深化を図る方針を確認した。

○ 2022年5月10日、11日には、G7デジタル大臣会合がドイツにおいて開催され、事業者の違法・有害情報への対応措置に関する透明性・アカウントビリティを世界・国・地域のレベルにおいて、該当するポリシーごとに果たすことを含む、eSafety等について議論。

誹謗中傷を始めとしたインターネット上の違法・有害情報対策に関する国際的な制度枠組みや対応状況を注視し、対応方針について国際的な調和(ハーモナイゼーション)を図るため、国際的な対話を深化させる。

## 日EU・ICT政策対話、同戦略ワークショップ

- 2022年2月3日(木)、第27回日EU・ICT政策対話を実施。日本側からは、「インターネット上の誹謗中傷(Online Harassment)」及び「フェイクニュース(“Fake news” & disinformation)」に関する政策動向について説明。EU側からは、Digital Services Act(DSA: デジタルサービス法)及びDigital Market Act(DMA: デジタル市場法)について紹介があった。

## インターネットエコノミーに関する日米政策協力対話

- 2021年11月11日(木)、第12回インターネットエコノミーに関する日米政策協力対話(政府間会合)を実施。日本側からは、「インターネット上の誹謗中傷(Online Harassment)」に関する政策動向を説明。米国側とインターネット上の違法・有害情報対策についての問題意識を共有した。

## 金子大臣の会談

- 2022年5月4日(水)、英国デジタル・文化・メディア・スポーツ省ナディーン・ドリーズ大臣と金子総務大臣との間で会談を実施。両大臣は、昨今の国際情勢においてデジタル空間の安全性・信頼性の確保が一層重要になっている背景を踏まえ、日英のICT分野での協力深化の必要性を確認。具体的には、デジタル空間の安全性及び信頼性の確保を日英が協力して国際的に推進していくことを確認。
- 2022年5月2日(月)、欧州委員会のヴェステアー上級副委員長及びブルトン委員と金子総務大臣との間で会談を実施。これまでの日EU・ICT政策対話や日EU・ICT戦略ワークショップ等による日EUの緊密な連携を確認。デジタル空間における利用者保護等に関連して、表現の自由の重要性と透明性確保の重要性を確認。

### 未来のインターネットに関する宣言

- 2022年4月28日、「未来のインターネットに関する宣言」立ち上げイベントが開催された（米国主催、日・豪・カナダ・EU・英国を含む国と地域が賛同）。日本側からは金子総務大臣が出席し、自由で開かれた分断のないグローバルなインターネットの強化、マルチステークホルダー・アプローチによるインターネットガバナンスの重要性及び2023年に我が国で開催予定のインターネット・ガバナンス・フォーラムとの連携可能性等について発言した。
- 同宣言は、開かれ、自由で、グローバルで、相互運用性があり、信頼でき、安全な、未来のインターネットへの支持を呼びかけるという目的を持つ。同宣言のパートナーは、既存の各フォーラムで原則を促進したり、原則の内容を具体的な政策や行動を通じて実現していくこと等が期待される。

### 日EUデジタルパートナーシップ

- 2022年5月12日、第28回日EU定期首脳協議において「日EUデジタルパートナーシップ」が立ち上げられた。日本側はデジタル庁、総務省、経済産業省、EU側は欧州委員会通信ネットワーク・コンテンツ・技術総局が中心となり、日EUのデジタル分野における共同の優先事項を扱う。
- オンライン上の違法・有害コンテンツに関するプラットフォームにおける透明性と説明責任の確保に向けて、意見交換等を通じて協力を進めることとなった。

論点

【3. 発信者情報開示に関する取組】

○ 法制度整備にむけて順調に準備が進んでいるか。

- 総務省は、インターネット上の誹謗中傷などによる権利侵害についてより円滑に被害者救済を図るため、①発信者情報開示について新たな裁判手続（非訟手続）を創設し特定の通信ログを早期保全すること、②開示対象となるログイン時情報を明確化すること等を内容とする、プロバイダ責任制限法の改正を実施（2021年4月に改正法成立）。2022年（令和4年）10月1日施行予定。
- SIAにおいて設置された有識者会議において、2020年7月から2021年3月までの間に、任意開示の促進に向けた施策の検討が行われ、2021年4月5日に「権利侵害明白性ガイドライン」が策定・公表された。また、同ガイドラインに関する理解を深めるため、プロバイダからの同ガイドラインに関する相談を受け付ける窓口が設置された。同ガイドラインは、適切な任意開示の促進につながるよう、容易に名誉毀損が明白であると判断可能な類型を示すことで、プロバイダにおける任意開示の判断に際して参考となる判例を集積したもの。
- コンテンツプロバイダ・アクセスプロバイダ・有識者・専門性や実務的知見を有する者が協力して発信者の特定手法について支援協力を行える体制やノウハウ共有を行う場として、2021年2月以降、「発信者情報開示に関する実務者勉強会」が継続的に開催されている。

これまでの主な意見

【発信者情報開示制度について】

- 新たな裁判手続の創設により、以前よりはハードルが低くなったと思うが、やはり一般の利用者にとってはまだまだハードルが高く、これをどうやって下げていくのか。もう少し使いやすいものになっていかないかと思う。【木村構成員（第32回会合）】

## 論点

## 【4. 相談対応の充実にに向けた連携と体制整備】

- 相談窓口に関するユーザへのわかりやすい案内は十分か。相談機関間の連携は十分か。
- 2021年度より、違法・有害情報相談センターにおける相談員の増員等により体制強化を図るとともに、相談件数・内容の更なる分析を実施。ユーザにとって分かりやすく相談窓口を案内し、利便性を高めるために、相談システムからの通知や各種相談窓口をチャットボットで案内できるよう、公式LINEアカウントを開設。
- 相談内容に応じて相談機関間（法務局、警察、インターネットホットラインセンター、セーフライン、誹謗中傷ホットライン、消費者生活相談窓口等）で紹介を行うなど、他の相談機関との連携対応を充実。2021年度からは、違法・有害情報相談センターにおいて違法・有害情報相談機関連絡会を開催し、各機関の取組に関する相互の理解の促進と、機能の相互補完を目指した連携を図っている。
- ユーザにとって分かりやすい相談窓口の案内を実施するため、インターネット上の誹謗中傷に関する複数の相談窓口（違法・有害情報相談センター（総務省）、人権相談（法務省）、誹謗中傷ホットライン（SIA））に関する案内図を作成し、2020年12月に公表。様々な方法で全国に周知広報を実施。
- SNSユーザを対象としたアンケート調査結果では、不適切な投稿をされたときの相談窓口について、「利用したかったが、分からなかった」・「知らなかった」という人が7割弱（69.2%）を占め、「知っており、利用したことがある」人は非常に少なかった（2.4%）。また、具体的な相談窓口についての認知度はいずれもあまり高くないが、「みんなの人権110番」、「地方自治体の各種相談窓口」が相対的に高かった。（相談窓口を知っている層の半数以上が、前述の2つを知っている。）

**これまでの主な意見****【相談体制の充実について】**

- インターネット上の誹謗中傷に関する相談窓口に関する取組は大変良いこと。利用者の心が折れているときに、自分がどこに相談していいのかを悩まなければならないというのは大変なこと。どこに相談しても、こうした案件であればここに相談すればいいという風に、適切なところに案内してくれるような連携が必要ではないか。【木村構成員（第32回会合）】
- 違法・有害情報センターに苦情が増えているというのは、苦情を言えるようになって届けやすくなったのか、それとも単に苦情そのものが増えているのか、微妙なところだと思うが、普及啓発も大変大事なこと。特に子供はすぐに成長するため、同じ小学校であっても生徒はどんどん替わっていくため、普及啓発は本当に大事なこと。いまだに、やはり写真を送ってしまうとか、大人にとってみればしてはいけないと当然分かっていることでも知らなかったとか、居場所がない子供さんたちが居場所を求めてSNSなどでいろんな被害に遭ってしまうということを聞くと、本当に切ない思いがする。適切に周知する必要性があり、学習会とかイベントだけではなく、色々な機会で普段からあると良い。18歳未満の児童の写真を送ったりするのは児童ポルノ法違反ですとか、18歳未満の連れ出しは略取誘拐罪に当たる可能性がありますと、テレビ番組の画面に出たことは周知効果がある。いろいろな機会を捉えて、そうした啓発をすることは非常に大事。今後、様々な方がネットでのコミュニケーションをますます取っていくことになっていくと思う。ぜひ様々な取組を引き続きお願いしたい。【木村構成員（第32回会合）】

- インターネット上に流通した情報による被害に係る一般利用者などからの相談を受け付け、具体的な削除要請の方法等についての的確なアドバイス等を行う「違法・有害情報相談センター」を設置・運営。
- 2021年度より、「違法・有害情報相談機関連絡会」を新設し、相談機関間の円滑な連携による相談者への最適な解決策の提供等を目的に、相談機関間での連携強化を図っている。



## 違法・有害情報相談機関連絡会の新設

- 違法・有害情報相談センターでは、インターネット上の人権侵害等に関する被害者救済を図るため、他の相談機関との連携強化を実施。
- 2021年度(令和3年度)から、違法・有害情報相談機関連絡会を新設。2021年11月に第1回会合を開催、2022年3月に第2回会合を開催。今後、年2回程度の開催を予定している。

### <参加機関>



(事務局)



インターネット  
ホットラインセンター



セーフアール  
インターネット協会



東京都



全国消費生活  
相談員協会

### <オブザーバ>

- ・警察庁
- ・個人情報保護委員会事務局
- ・総務省
- ・法務省
- ・国民生活センター

## ○ インターネット上の誹謗中傷に関する複数の相談窓口(※)に関する相談窓口の案内図を作成

(※)「違法・有害情報センター」(総務省)、「人権相談」(法務省)、「誹謗中傷ホットライン(セーフターインターネット協会)」

## ○ 相談窓口の周知広報

- ・総務省(新型コロナウイルス関連情報ページ含む)、法務省のホームページにおける周知広報
- ・「インターネットトラブル事例集(2022年版)」に掲載
- ・e-ネットキャラバンの相談窓口に関する参考資料に追加
- ・#NoHeartNoSNS特設サイトにおける掲載
- ・内閣府政府広報室と連携した周知広報(※)
- ・「鷹の爪団の#NoHeartNoSNS大作戦」コンテンツに掲載
- ・全国の総合通信局等と連携した周知広報
- ・関係団体(※)と連携した周知広報

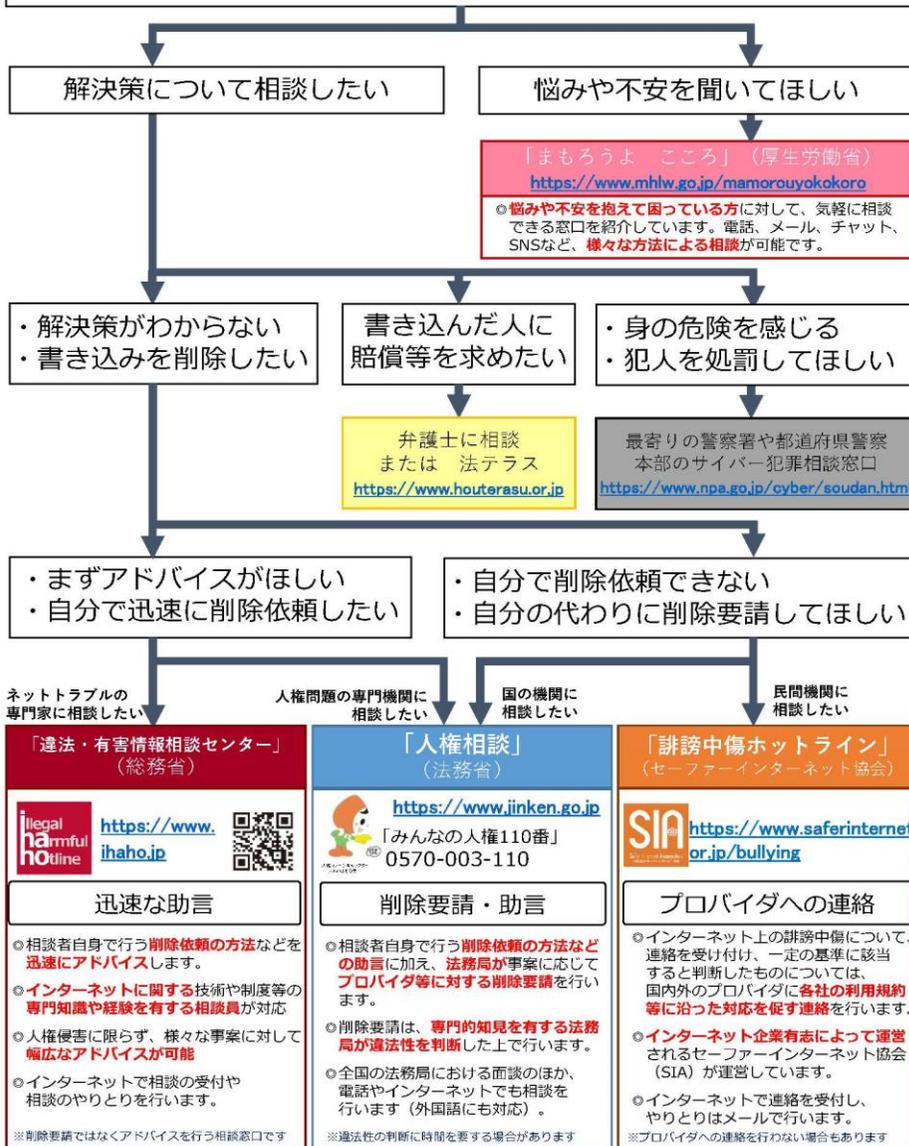
(※)通信事業者関連では、(一社)安心ネットづくり促進協議会、(一社)全国携帯電話販売代理店協会、全国高等学校PTA連合会、(一社)ソーシャルメディア利用環境整備機構、(一社)テレコムサービス協会、(一社)電気通信事業者協会、(一財)マルチメディア振興センター等

その他、アスリートへの写真・動画による性的ハラスメントへの対応として、(公財)日本オリンピック協会、(公財)日本スポーツ協会、(公財)日本障がい者スポーツ協会、(一社)一般社団法人大学スポーツ協会、(公財)全国高等学校体育連盟、(公財)日本中学校体育連盟、(独法)日本スポーツ振興センター

等

## インターネット上の誹謗中傷に関する相談窓口のご案内

インターネットの書き込みにより、誹謗中傷などの被害にあわれた場合



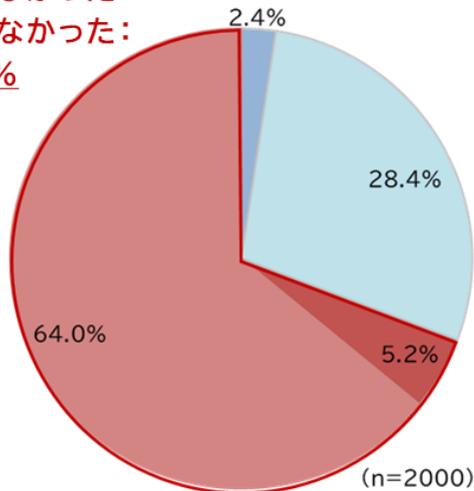
※上記のほか、学校や地方公共団体にある相談窓口も活用してください。

- 調査実施期間: 2022年3月16日(水)~3月18日(金)
- 調査方法: インターネットアンケート調査
- 調査対象: 日本国内の15歳から99歳までの男女
- 回答者数: 2,000名
- 回答割付: 回答者の性別・年代による回答に特定の傾向がある場合に、その影響を少なくするために、(均等割付でなく)インターネットの利用者の性別・年代別の分布に即して回答者数を割り付けた: 具体的には、1) 住民基本台帳人口(2021年1月1日付)の性別・年齢階級別人口に、通信利用動向調査(令和2年度)のインターネット利用率を掛け合わせてインターネット利用人口を算出。2) 回収総数を2,000人とし、算出したインターネット利用人口の構成比に基づいて、性別・年代別の回収数を割付

## 相談窓口の認知度

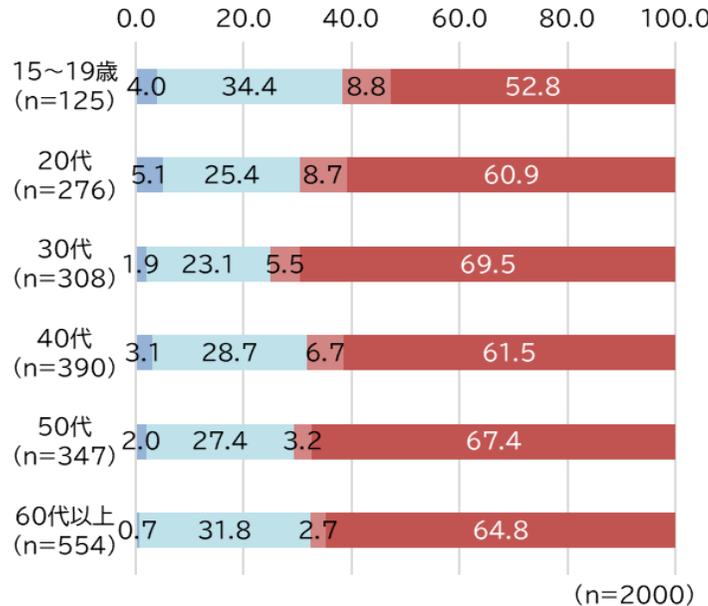
Q あなたは、SNSなどのインターネットのサービス上で不適切な投稿をされたときに、相談できる窓口があることを知っていますか。(SA)

分らなかった  
+ 知らなかった:  
69.2%



- 知っており、利用したことがある
- 知っているが、利用したことはない
- 利用したかったが、分らなかった
- 知らなかった

## 相談窓口の認知度(年代別)

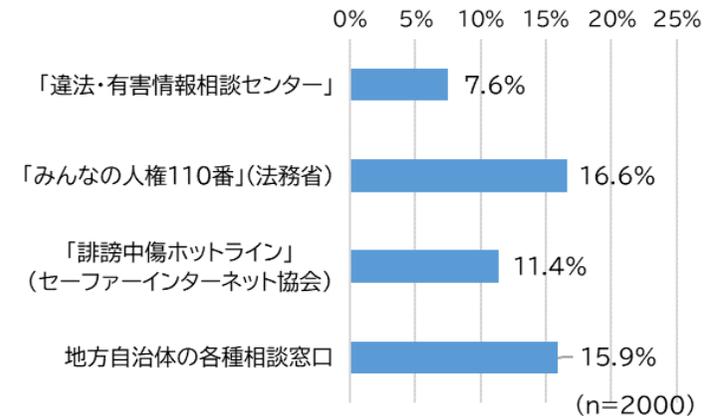


- 知っており、利用したことがある
- 知っているが、利用したことはない
- 利用したかったが、分らなかった
- 知らなかった

## 相談窓口別の認知度

Q あなたは、以下の相談窓口について知っていますか。以下のそれぞれの相談窓口について、「知っている」又は「知らない」のいずれかで教えてください。(※)

※ 原設問は、相談窓口の存在を知っていると答えた回答者を対象としていたが、「分らなかった」「知らなかった」回答者は具体的な窓口も知らないとして、全回答者における認知度を算出。



\* 上記以外に、YouTube運営、SNSの相談窓口、警察署、ベンダーの苦情窓口、が各1件ずつ挙げられた。

※ 過去1年間にいずれかのSNSなどのサービスを利用したと答えた回答者を抽出して集計

○ PF事業者による誹謗中傷等の情報への対応に係る透明性・アカウントビリティの確保が図られているかについて、対応に係る各プロセスにおいて数値等の公開の有無を検証する。

### <主なモニタリング項目>

#### 1. サービス概要とポリシー策定状況

- ・ 他人を傷つける情報や行為に関するポリシーでの禁止・制限の設定状況

#### 2. 削除申告等の受付・審査の体制及びプロセス

- ・ 端緒となる投稿の申告・探知の件数等
- ・ 対応のプロセス、審査基準
- ・ 対応の体制
- ・ 削除等の理由の説明の有無等

#### 3. 削除等の対応の実施状況及び機能上の工夫

- ・ 削除された投稿の件数等

#### 4. プロバイダ責任制限法に基づく請求への対応

- ・ PFによる発信者情報の開示件数

#### 5. 取組の情報分析・公開

- ・ 一連の対応に係る数値等の公表の有無
- ・ PF内の対応や体制に係る公表の有無

### <PF事業者による対応のフロー(イメージ)>

同意・  
利用開始

- ・ PFは自らのサービス利用に関するポリシー・利用規約をあらかじめ策定。
- ・ 全てのユーザはポリシー等に同意して利用を開始。

投稿

- ・ 発信者が誹謗中傷などの情報をPFのサービス上に投稿

申告・検知

- ・ 被害者が発信者の投稿をポリシー等に反する情報として申告・削除要請
- ・ 又は、PFが主体的にポリシーに反する情報を検知

審査・決定

- ・ 申告・要請の対象となった情報が、ポリシー等に反しているか審査
- ・ コンテンツモデレーションの実施の有無を決定

通知・  
異議申立て

- ・ 実施に関する通知(申立人・発信者)や異議申立ての受け付け

削除等

- ・ 当該情報や発信者に対するモデレーション(投稿の削除、ラベルの付与、アカウント停止等)

一部の投稿は  
裁判手続きでの  
解決に移行

### <裁判手続きによる課題解決のフロー>

発信者  
情報開示

- ・ 被害者がPFによる任意での発信者情報の開示を請求
- ・ 裁判手続きによる開示

訴訟等

- ・ 開示された情報に基づき、被害者が発信者に対して損害賠償請求等

結果の公表

- ・ 一連の対応結果や体制に関する透明性レポートの公表等

## プラットフォームサービス事業者における誹謗中傷等への対策状況ヒアリング (2022年) 質問項目

### 1. サービス概要とポリシー

- ① 対象サービス名
- ② 加入者数、月間アクティブユーザ数又は書き込み数
- ③ (i)誹謗中傷等に関して禁止とする情報や行為【★】  
(ii) (i)で挙げた行為に対して取り得る対応【★】
- ④ ポリシー等の見直し状況及び外部レビューの有無とそのタイミング

### 2. 削除申告等の受付・審査のプロセス

- ① (i)一般ユーザからの申告・削除要請への受付窓口・受付態勢、対応プロセス【★】  
(ii)主体的な誹謗中傷等の情報の検知態勢
- ② 申立時における申立者への通知の有無
- ③ 申立てに対するポリシー違反等の審査プロセス、案件ごとの優先度の決定方法【★】
- ④ 一般ユーザからの申告や削除要請に対応する部署・チームの規模・人数及び部署・チームに関する日本国内の拠点の有無、日本における責任者の有無【★】
- ⑤ (i)対応決定時の申立者への通知の有無とその理由説明内容  
(ii)対応決定時の対象となる情報の発信者への通知の有無とその理由説明内容及び異議申立て方法の案内の有無
- ⑥ 削除等への異議申立や問い合わせに対する受付態勢及び処理プロセス【★】
- ⑦ ⑥の異議申立・問い合わせ件数【★】
- ⑧ 不正な申告や削除要請への対策の方法・仕組み、対応件数

### 3. 削除等の対応の実施状況及び機能上の工夫

- ① 誹謗中傷等に関する申告や削除要請の件数【★】
- ② 誹謗中傷等に関する申告や削除要請に対する  
(i)削除件数【★】、(ii)削除以外の対応件数【★】、  
(iii)対応を行わなかった件数、  
(iv)削除までの平均時間及び/又は削除対象の閲覧数の抑制程度
- ③ (i)主体的な削除件数(AIを用いた自動検知機能の活用【★】)  
(ii)主体的な削除以外の対応件数、  
(iii)主体的な対応までの平均時間及び/又は対象情報の閲覧数の抑制程度
- ④ 誹謗中傷等個人の権利侵害を防止・低減するためのサービス上の機能の工夫(特に炎上対策)

### 4. 発信者情報開示への対応

- ① 発信者情報開示請求を受けた件数(裁判上/訴外)【★】
- ② 開示を行った件数(裁判上/訴外)【★】

### 5. 情報分析・公開

- ① 透明性レポートの公開、日本語で閲覧可能か【★】
- ② 取組の効果分析【★】
- ③ 誹謗中傷等の流通状況に関する分析・調査及びユーザに与える影響の分析・調査
- ④ 研究者への情報提供、利用条件

### 6. その他の取組

- ① 普及啓発
- ② 誹謗中傷以外にポリシーで禁止又は制限される情報・行為
- ③ 意見・補足

### 論点

#### 【プラットフォーム事業者による取組 透明性・アカウントビリティの向上】

##### ○ 主な項目（案）に関する事業者による自主的な誹謗中傷対策のための取組

- 1.関係 ・ 誹謗中傷等に関する利用規約・ポリシーについて、適切に設定されているか
- 2.関係 ・ 誹謗中傷等に適切に対応するための態勢は適切か。特に、日本国内における態勢が適切か。
  - ・ ポリシーに違反する投稿等について適切にユーザ申告の受付又は事業者による検知が行われているか。
  - ・ 申告受付又は検知された投稿について、適切に審査されているか。
  - ・ 審査結果等に関する発信者や申立者に対する異議申し立ての受付態勢が適切か。
- 3.関係 ・ 利用規約・ポリシー等について、適切にポリシー等に基づいて削除等の対応がなされているか。
- 4.関係 ・ 発信者情報開示について、適切に対応されているか。
- 5.関係 ・ 透明性レポートや取組の効果分析は適切に公開されているか。特に、日本国内において公開されているか。

##### ○ 主な項目（案）以外の項目に関する事業者による自主的な誹謗中傷対策のための取組

- ・ AIの活用やアーキテクチャ上の工夫が進められているか
- ・ 過剰な削除や不当なアカウント停止等の行き過ぎた対応が行われていないか（個別の説明は十分か）
- ・ 各事業者や事業者団体における啓発活動は十分か
- ・ 誹謗中傷等の流通状況に関する分析・調査や、研究者への情報提供が行われているか

##### ○ 上記の各項目に関する、事業者による誹謗中傷対策に関する透明性・アカウントビリティ

- ・ 利用者が安心・信頼してサービスを利用できるよう、透明性・アカウントビリティ確保が十分に図られているか
- ・ 定量的な数値等を用いて、客観的に透明性・アカウントビリティの確保が十分に図られているか
- ・ グローバルな取組だけでなく、我が国における透明性・アカウントビリティの確保が十分に図られているか
- ・ 前回ヒアリングと比較して、事業者による透明性・アカウントビリティ確保のための取組に進展は見られるか

- プラットフォーム事業者の誹謗中傷への対応に関する透明性・アカウントビリティ確保状況は、前回ヒアリング状況から一部進展が見られるものの、一部項目において、依然、透明性・アカウントビリティの確保が十分とは言えない状況。
- ヤフー及びLINEは、我が国における誹謗中傷への対応について、具体的な取組や定量的な数値を公表しており、透明性・アカウントビリティ確保に向けた施策が進められている。ヤフーでは透明性レポートが新たに公表された。
- Googleは、我が国における定量的な削除件数等が一部新たに示されおり、透明性・アカウントビリティ確保に向けた施策に進展が見られる。一方、対応のための部署に関する情報や発信者情報開示に関する情報は明らかではない。
- Meta及びTwitterは、我が国における定量的な削除件数等が事後的に一部示され透明性・アカウントビリティ確保に向けた施策に一定の進展が見られる。一方、対応のための部署に関する情報や発信者情報開示に関する情報は明らかではない。

### プラットフォーム事業者における誹謗中傷等への対応に関する透明性・アカウントビリティ確保状況の主な項目の評価(案)

	ヤフー	Google	LINE	Meta	Twitter
誹謗中傷等に関するポリシー(1-③(i)(ii)関係)					
一般ユーザからの申告や削除要請に対応する部署・チームや日本国内の拠点・責任者(2-①③④関係)					
削除等への苦情等対する苦情受付態勢及び苦情処理のプロセス、件数(2-⑥⑦関係)					
日本における削除要請件数や削除件数(3-①②③関係)					
発信者情報開示の件数(4-①②関係)					
取組の効果分析(5-②関係)					
透明性レポートの公開(5-①関係)					

→今回のモニタリングの質問項目について、事業者が「回答を控えた理由」は次ページの通り

(参考)プラットフォーム事業者による対応のモニタリング結果(2022年)  
「回答を控えた理由」及び「今後の対応方針」(1/2)

事業者名	回答を控えた理由及び今後の対応方針
ヤフー	<p>&lt;Yahoo!知恵袋&gt; 4. (3) 削除等の対応(①偽情報等に関する申告や削除要請の件数、②(ii)アカウントの停止数) 理由: 申告理由について、「偽情報」という切り口で件数をカウントしていないため。 対応方針: 対処すべき投稿がどのようなものかを含め、偽情報への対応につき有識者の意見も聞きながら検討を進める。</p> <p>&lt;Yahoo!ニュース&gt; 4. (3) 削除等の対応(①偽情報等に関する申告や削除要請の件数～⑥不正な申告や削除要請への対策の方法・仕組み、対応件数) 理由: 2022年3月にYahoo!ニュースコメントポリシーを改定したばかりであり、件数等定量データ収集中であるため。 コメントポリシー&lt;<a href="https://news.yahoo.co.jp/info/comment-policy">https://news.yahoo.co.jp/info/comment-policy</a>&gt; コメントポリシーの改定&lt;<a href="https://news.yahoo.co.jp/newshack/information/comment_policy_20220322.html?msckid=443a99f8a9bf11ec9a2f08d6d1a109ec">https://news.yahoo.co.jp/newshack/information/comment_policy_20220322.html?msckid=443a99f8a9bf11ec9a2f08d6d1a109ec</a>&gt; 対応方針: 来年度以降、回答準備を進めていく。</p>
Google	<p>&lt;誹謗中傷&gt; *2-4 一般ユーザからの申告や削除要請に対する部署・チームの規模・人数 及び部署・チームに関する日本国内の拠点の有無、日本における責任者の有無 →日本の案件の審査と承認は、日本語話者だけではなく、さまざまな部署のグローバルなチームがそれぞれの専門性を持って対応しておりますため、日本の案件だけに従事する人間の数を正確にお示しすることができませんが、日本をカバーする人数は必要に応じて拡充していきます。 *3-3-2 (ii)主体的な削除以外の対応(非表示化やラベルの付与等)を行った件数 →このような取り組みの個別件数については、グローバルに開示しておりません。引き続きユーザーの安全を守るために、どのような仕組みが機能しているかについては、さまざまな形で開示しています。Google 検索について詳細は、How Search Works(Google 検索の仕組みを確認する) &lt;<a href="https://www.google.com/intl/ja/search/howsearchworks/">https://www.google.com/intl/ja/search/howsearchworks/</a>&gt; をご覧ください。YouTube についての詳細は、How YouTube Works (YouTube の仕組みとは) &lt;<a href="https://www.youtube.com/howyoutubeworks/">https://www.youtube.com/howyoutubeworks/</a>&gt; をご覧ください。Google マップについての詳細は、Google のクチコミの仕組みについてのブログ &lt;<a href="https://japan.googleblog.com/2022/02/google.html">https://japan.googleblog.com/2022/02/google.html</a>&gt; をご覧ください。 *4-1 発信者情報開示請求を受けた件数 (i)裁判上の開示請求件数、(ii)裁判外の開示請求件数 および *4-2 開示を行った件数 (i)裁判上の開示件数、(ii)裁判外の開示件数 →政府からの情報開示リクエストに対して、Google がどのような対応をとったかについて透明性を持たせることを主目的としている弊社の透明性レポートでは、その目的に沿って裁判所を含む政府機関(民事訴訟を除く)におけるユーザー情報の開示についてのデータのみ提供しております。</p> <p>&lt;フェイクニュース&gt; *4-(2)-3 一般ユーザからの申告や削除要請に対応する部署・チームの規模・人数 →日本の案件の審査と承認は、日本語話者だけではなく、さまざまな部署のグローバルなチームがそれぞれの専門性を持って対応しておりますため、日本の案件だけに従事する人間の数を正確にお示しすることができませんが、日本をカバーする人数は必要に応じて拡充していきます。 *4-(3)-2-ii →何をもち「日本」のアカウントとするかは、技術的な課題がございます。弊社は、継続して利用可能な情報を拡大しますが、正確さとデータの妥当性には最大限の注意を払って行います。</p>

(参考)プラットフォーム事業者による対応のモニタリング結果(2022年)  
「回答を控えた理由」及び「今後の対応方針」(2/2)

事業者名	回答を控えた理由及び今後の対応方針
LINE	<p>偽情報・フェイクニュース対策のヒアリングにおける質問6-④「日本におけるファクトチェック推進に向けた取組」について、 「回答を控えた理由」 当初設問には「6-④」が無かったところ、他方「6-⑤」において特定非営利活動法人ファクトチェック・イニシアティブ(FIJ)に正会員として参加し、日本におけるファクトチェック推進に寄与している旨既往回答済みであったことから、凡そ同様の趣旨の回答となることによる重複を避けた方が良いものと判断させて頂いたため。 「今後の対応方針」 政府や関係団体等と連携して、日本におけるファクトチェック推進に向けた検討を進める。</p>
Meta	<ul style="list-style-type: none"> <li>・誤情報は、弊社のコミュニティ規定上の他のタイプの言論とは異なり、該当する内容の包括的なリストがありえないことに留意する必要があります。</li> <li>・ポリシーで単に「誤情報」を禁止したとしても、弊社のサービスの利用者にとって有益な通知にはなりませんし、弊社は全ての情報を知り得るわけではないので、そのようなポリシーを行使することもできません。</li> <li>・このため、弊社のポリシーでは、誤情報をさまざまなカテゴリーに分類し、そのような言動を確認したときにどのように扱うかについて明確な指針を示すよう努めています。</li> <li>・また弊社は、新型コロナウイルス感染症に関連する誤情報については、日本における措置件数を具体的に示しています。</li> <li>・弊社は、今後も外部のステークホルダーと協力し、ポリシーの構築と実施、そして誤情報に対応するための投資を続けて参ります。</li> </ul>
Twitter	<ul style="list-style-type: none"> <li>・日本における誹謗中傷や偽情報に関するデータについては、2022年5月11日に提供いたしました。</li> <li>・削除要請等に対応する部者の規模や拠点について、社員の安全面等を考慮し、会社の方針として個別具体の部署の人数は(今回の当該部署に限らず)公表しておりません。また拠点についても、セキュリティの観点から公表しておりません。</li> </ul>

### 論点

【プラットフォーム事業者による取組 透明性・アカウントビリティの向上】

(総論 (案) )

○ 主な項目 (案) に関する事業者による自主的な誹謗中傷対策のための取組に関する評価

我が国における誹謗中傷対策のための取組は適切に行われているか、また、その透明性、アカウントビリティの確保が図られているかという点について、次のとおり。

- すべての事業者において、誹謗中傷対策のために、あらかじめポリシーに定めた行為や情報について、ユーザや第三者からの申請やAIなどを活用した自主的な検知に基づき、投稿の削除やアカウント停止などの対応を行っている。
- 前回ヒアリングと比較して、プラットフォーム事業者によるヒアリングシートへの回答では、例えば、我が国における削除件数について、Googleが構成員限りであった定量的な情報を新たに公表したり、Meta、Twitterもヒアリング後に追加で数値を提供するなど、一部で進展が見られる。
- 一方で、我が国における誹謗中傷のための投稿の削除やアカウント停止などの対応について、具体的には、我が国における誹謗中傷対策に十分なリソースが投入されているかという点や削除要請等に対する審査のプロセス、発信者情報開示請求に関する対応状況、サービス上の取組に関する効果分析、透明性レポートの公表状況等について、Google、Meta、Twitterによる透明性・アカウントビリティの確保状況には差異がある状況であり、誹謗中傷への対応のために、我が国において十分なリソースが投入され、文化的、社会的、政治的背景を踏まえた措置が行われているか、十分に明らかではない点がある。
- 各社の回答において、回答対象とするサービスや回答に用いる指標、期間等には、差異が見られる。

### 論点

#### 【プラットフォーム事業者による取組 透明性・アカウントビリティの向上 (続き)】

- 各プラットフォーム事業者による取組状況とその透明性・アカウントビリティ確保状況について、ヒアリングシート及び事後に追加提出のあった情報によれば、次のとおり。
  - 誹謗中傷等に関するポリシー (1-③(i)(ii)関係)  
すべての事業者において、前回ヒアリングに引き続き、誹謗中傷を含む一定の種類の行為・情報について禁止規定を定めており、削除・警告表示・アカウント停止等の対応が規定されていることが明らかになっている。
  - 一般ユーザからの申告や削除要請に対応する部署・チームや日本国内の拠点・責任者 (2-①③④関係)  
一般ユーザからの申告・削除要請について、すべての事業者において、前回ヒアリングに引き続き、受付窓口・受付態勢、対応プロセスが整備されていることが明らかになっている。また、主体的なポリシー違反等の情報の検知について、すべての事業者において、これを行うための仕組みが設けられていることが明らかになっている。  
対応する部署・チームの規模・人数や日本国内の拠点・責任者について、前回ヒアリングから引き続き、情報開示には差が見られる。ヤフー及びLINEは、日本における体制について、前回ヒアリングに引き続き、定量的に示している。他方で、Google、Meta、Twitterは、前回ヒアリングに引き続き、グローバルな体制については公表しているものの、また、我が国における具体的な体制が示されておらず、誹謗中傷への対応のために、我が国に対して、十分なリソースが投入されているか、また、文化的、社会的、政治的背景を踏まえた措置が行われているか、明らかではない。
  - 削除等の対応に対する異議申し立ての機会とその受付態勢 (2-⑥⑦関係)  
すべての事業者において、削除等への苦情、問い合わせの受付窓口や手続きを設けていることが明らかになっているが、削除等の措置を受けた発信者向けのものか、削除等の措置を申請した申立者へのものが明らかではない。ヤフーでは、今後、削除理由の開示フローを検討することとされている。Metaでは、削除されたコンテンツについて異議申し立ての手続きを経ている場合には、監督委員会に申し立てることができる。ヤフー、LINEでは、日本国内における異議申し立ての件数が、Googleではグローバルでの一部サービスに関する件数が、それぞれ明らかにされている。Meta、Twitterは、決定に対する異議申し立ての件数がグローバル、日本国内いずれも明らかではない。

### 【プラットフォーム事業者による取組 透明性・アカウントビリティの向上 (続き)】

- 日本における削除要請件数や削除件数等 (3-①②③関係)

ヤフー及びLINEは、日本における削除要請件数及び要請に基づく削除件数と主体的な削除件数について、前回ヒアリングに引き続き定量的に示しており、透明性・アカウントビリティ確保が進められている。

Googleは、国内における削除要請件数について、前回ヒアリングと比較して新たにサービスごとに定量的に示している。また、国内における削除件数について、一部サービスについての回答が明確ではないものの、前回ヒアリングでは構成員限りの非公開としていた情報を公開するなど、透明性・アカウントビリティ確保について進展が見られる。

Metaは、国内における削除要請件数について、前回ヒアリングに引き続き、定量的に示していない。国内における削除件数について、ヒアリング後に、我が国におけるいじめや嫌がらせに関するポリシー違反として削除された投稿件数について、ユーザからの申請に基づくものと主体的なものが明確ではないものの、新たに定量的な数値を示すなど、透明性・アカウントビリティ確保について一部進展が見られる。

Twitterは、国内における削除要請件数について、前回ヒアリングに引き続き、定量的に示している。国内における削除件数について、ユーザからの申請に基づくものと主体的なものの区別が明確ではないものの、ヒアリング後に、我が国におけるTwitterが定めるポリシー違反として削除されたアカウント件数や投稿件数について、新たに定量的な数値をポリシー別に示すなど、透明性・アカウントビリティについて一部進展が見られる。

削除以外の対応を行った件数については、Metaを除いた各社において定量的に示されているものの、各社によって行った対応の内容や国内における対応状況に関する数値の開示状況には差異が見られる。

### 【プラットフォーム事業者による取組 透明性・アカウントビリティの向上 (続き)】

#### ・ 発信者情報開示の件数 (4-①②関係)

ヤフー及びLINEは、日本におけるプロバイダ責任制限法に基づく発信者情報開示請求を受けた件数について、前回ヒアリングに引き続いて、裁判上/裁判外の件数をいずれも定量的に示している。また、開示を行った件数について、裁判上/裁判外の件数がいずれも示されており、透明性・アカウントビリティ確保が進められている。

Google、Metaは、発信者情報開示請求を受けた件数及び開示した件数について、前回ヒアリングに引き続いて、いずれも示されておらず、透明性・アカウントビリティが確保されていない。

Twitterは、前回ヒアリングと比較して、政府、執行機関・捜査機関、第三者からのアカウント情報の開示の請求を受けた件数について定量的な数値を新たに公開したが、裁判上/裁判外の請求であるかの区別や、プロバイダ責任制限法に基づく発信者情報開示請求や捜査関係事項照会などの区別が明確ではない。

#### ・ 取組の効果分析 (5-②関係)

ヤフーは、AIを活用した取組に関して、前回ヒアリングに引き続いて、日本におけるサービス上の取組の効果分析を定量的に示している。LINEは、前回ヒアリングに引き続いて、サービス上の取組の効果分析を示していないが、教育に関する取組の効果を示している。Google・Meta・Twitterは、前回ヒアリングに引き続いて、サービス上の取組の効果分析を公開しているが、グローバルの数値のみであり、我が国における効果分析を示しておらず、透明性・アカウントビリティ確保について進展は見られない。

#### ・ 透明性レポートの公開 (5-①関係)

ヤフーは、前回ヒアリングと比較して、透明性レポートを新たに作成・公表している。

LINEは、前回ヒアリングに引き続いて、日本語による透明性レポートを公開している。

Google及びTwitterは、前回ヒアリングに引き続いて、日本語による透明性レポートを公開しているが、一部を除き、グローバルの数値のみが公開されている。

Metaは、前回ヒアリングと引き続いて、日本語による透明性レポート (コミュニティ規定施行レポート) を公開しているが、すべてグローバルの数値であり日本における数値は公開されていない。

### 【プラットフォーム事業者による取組 透明性・アカウントビリティの向上 (続き)】

○ 主な評価項目 (案) 以外の項目に関して、アーキテクチャ上の工夫やAIを用いた削除対応が行われているかという観点について、次のとおり。

#### ➤ AI等を用いた削除対応について (3. ④関係)

各事業者において、積極的にAIを活用した削除等の取組が進められている。

- ・ヤフー : 「Yahoo!ニュース コメント」において2014年から機械学習による不適切投稿への対応を開始。AIによる検知を通じて、1日平均約29万件の投稿のうち、約2万件の不適切な投稿 (記事との関連性の低いコメントや誹謗中傷等の書き込みなど) の削除を実施。  
一定以上のコメント数がある記事のコメント欄を対象に、AIが判定した違反コメント数などの基準に従い、コメント欄を自動的に非表示にする機能を実装。
- ・Google : 機械学習を活用して不適切なコンテンツを検出。有害なコンテンツのほとんどがシステムによって一度も視聴されずに削除されている。ボーダーライン上のコンテンツのリコメンデーションを通じた視聴数の低下に取り組んでいる。
- ・LINE : 機械的なチェックにより、禁止用語やルールと照合し、規約や法令に反した投稿かどうか確認し、自動で非表示化。全サービスにおいて、わいせつ、出会い系、不快画像等について、AIを活用した「違反画像」を検知。
- ・Meta : AIを活用して不適切なコンテンツを検出。AI は、コンテンツレビュアーがレビューするケースに優先順位をつけて、最も有害で時間的な問題のあるコンテンツを最初に処理できるようにしている。
- ・Twitter : テクノロジー (PhotoDNA、社内の独自ツールなど) を活用し、違反コンテンツを特定。

### 【プラットフォーム事業者による取組 透明性・アカウントビリティの向上 (続き)】

○ 主な評価項目 (案) 以外の項目に関して、アーキテクチャ上の工夫やAIを用いた削除対応が行われているかという観点について、次のとおり。

➤ 削除以外の対策 (事業者の創意工夫による対応) について (3. ④関係)

- ・ ユーザ自らが他のユーザのアカウントやコメント等を非表示・ブロックする機能については、多くのサービスですでに導入されている (ヤフー、Google、LINE、Meta、Twitter)。
- ・ 投稿内容について投稿時等に再考・再検討を行う機会を設ける機能、ユーザの選択に応じたコンテンツフィルタリング機能、返信やDMのやりとり相手を選択できる機能などについては、一部のサービスで導入が進められるなど、サービスの性質に応じた様々な仕組みが設けられている。上記の仕組みにより誹謗中傷全般について対応できる可能性はあるものの、一定の短期間に大量の誹謗中傷が集まった場合に対応する特別な仕組みには引き続き多くのサービスにおける導入が期待される。

ヤフー : 自身の選択により書き込みや利用者の非表示・ブロック、低品質投稿の機械的検出と折りたたみ表示 (知恵袋)、AIを活用した投稿時における注意メッセージの掲出 (ニュースコメント、2020年度開始)、一度投稿停止措置を受けたユーザが再度アカウントを作成した場合の投稿制限等。一定以上のコメント数がある記事のコメント欄を対象に、AIが判定した違反コメント数などの基準に従い、コメント欄を自動的に非表示にする機能を実装。

Google : 利用規約上ボーダーライン上のコンテンツ等をおすすめ機能に表示しない機能。動画における低評価の数の表示を撤廃。

LINE : 自身の選択により書き込みや利用者の非表示・ブロック、18歳未満のユーザ検索機能制限、オープンチャットにおける誹謗中傷などについてAIを活用してNGワードを検知し投稿前に警告する機能を導入

Meta : 自身の選択により書き込みや利用者の非表示・ブロック、タグ付けや返信等を許可する相手を選択する機能、ブロックした人の別アカウントによる望まないやりとりの自動検知・防止、ポジティブなコメントを固定、不適切なコメントを自動的に非表示するフィルタ機能。コメントやダイレクトメッセージ (DM) を送るアカウントを一時的に制限し、急激に注目を集めた利用者が望まないやりとりを防ぐことができる「抑制」機能。

Twitter : 自身の選択により書き込みや利用者の非表示・ブロック、返信できるユーザの範囲を選択する仕組み、センシティブな内容を非表示にするフィルタ機能 (セーフサーチ)

### 【プラットフォーム事業者による取組 透明性・アカウントビリティの向上 (続き)】

○ 主な評価項目 (案) 以外の項目に関して、過剰な削除や不当なアカウント停止等のいきすぎた対応が行われていないかという点について、次のとおり。

➤ 対応決定時における通知の内容、理由の記載の程度について (2. ⑤関係)

すべての事業者において、システム上可能な場合には対応決定時において報告者や利用者に通知を行っている、理由の記載の程度については、濫用防止を理由に削除可否について詳細な理由を伝えることはできない、どのガイドラインに抵触したかについての情報を記載しているといった回答があった。また、児童ポルノなどの例外を除いて、対応決定の通知の際に、異議申立のプロセスの案内を行っているという回答があった。

➤ 不正な申告や削除要請への対策の方法・仕組み、対応件数について (2. ⑧関係)

一部事業者からは、特に違法性がないと思われる不正な申告や削除要請への対策の方法・仕組みについて回答があった (例: Metaでは、アカウントやコンテンツの過剰な報告は規約違反であり、システムの悪用 (大量の報告) を防ぐため重複報告を認識する技術を導入)。濫用的な報告に関する定量的な件数については、LINEのみが数値を公表している。

### 【プラットフォーム事業者による取組 透明性・アカウントビリティの向上 (続き)】

○ 主な評価項目 (案) 以外の項目に関して、誹謗中傷等の流通状況に関する分析・調査や、研究者への情報提供が行われているかという観点について、次のとおり。

➤ 誹謗中傷等の流通状況に関する分析・調査について (5. ②③関係)

一部事業者から、有識者との意見交換や、内部での分析・検討が行われているという回答があった。分析結果については、透明性レポート等により公開している事業者が見られた。

➤ 研究者への情報提供について (5. ④関係)

ヤフー : 大学、公的研究機関の研究者向けに一部ソフトウェアとデータを公開。

Google : 学術関係者はAPIを研究目的で使用可能。

LINE : 個別に対応。

Meta : 学術関係者などに対して選挙、民主主義等に焦点を当てた、Metaが世の中に及ぼす影響を研究するために必要なツールとデータの提供等を実施。

Twitter : 条件に合致する研究者に対する無料APIの提供を実施。

### 【プラットフォーム事業者による取組 透明性・アカウントビリティの向上 (続き)】

○ 主な評価項目 (案) 以外の項目に関して、各事業者や事業者団体における啓発活動は十分かという観点について、次のとおり。

#### ➤ 普及啓発について (6. ①関係)

・ 各事業者や団体において、自らのサービスのガイドの公開や、モラル教育等、様々な啓発活動が行われている。

ヤフー : サービスの理解を深めるための各種ガイドをウェブサイトで公開。各ガイドでは削除される投稿例の追記などを行い、ユーザーの理解を推進。コメントポリシーの遵守を呼びかける文言を投稿ページの目立つ位置に表示。

Google : 無料のデジタルスキルトレーニングプログラム「Grow with Google」を実施。その他、各種オンライン教材や、サービスの理解を深めるためのウェブサイト「YouTubeのしくみ」の公開。ネット利用に関する調査結果を「中高生インターネット利用白書 2021」として公開し、オンライン教材を拡充。

LINE : サービスの理解を深めるための各種ガイドをウェブサイトで公開。全国の学校や自治体に講師を派遣するワークショップ授業・講演活動等を2012年から累計で約10,000回以上実施。LINEみらい財団では、研究者共同で、独自の情報モラル教育教材の開発を行い、ウェブサイトで公開。

Twitter : サービスの理解を深めるためのガイドをウェブサイトや公式アカウントで公開。UNESCOと合同でメディア情報リテラシーのハンドブック「Twitterを活用した教育と学習」を作成。悩み相談を行うNPOへの無償広告枠提供等を通じた支援。

Meta : サービスの理解を深めるための各種ガイドをウェブサイトで公開。いじめ防止やオンラインでの嫌がらせなどのリソースを提供するプログラム「みんなのデジタル教室」を公開。特定非営利活動法人企業教育研究会の協力のもと、デジタルリテラシーに関する授業を提供。若年層の利用者やクリエイターと一緒に安全な利用を考える取り組み「#インスタNZENカイギ」等を実施。

SMAJ : 法務省人権擁護局・総務省と共同で、SNSのより良い利用環境実現に向けたスローガンである #NoHeartNoSNS を発表。特設サイトを開設し、被害を防ぐ仕組みや相談先を掲示。会員社による周知を実施。また、SMAJのウェブサイトにおいて、会員社が運営するサービスについて、統一されたフォーマットにより、サービスの概要、利用上の注意点、ユーザーを保護する機能を紹介。

	ヤフー	Google	LINE	Meta	Twitter
対象サービス名	Yahoo!知恵袋 Yahoo!ニュース(コメント欄) Yahoo!ファイナンス掲示板	Google 検索 YouTube Google マップ	LINE VOOM LINEオープンチャット LINE LIVE LINE BLOG、livedoor Blog	Facebook Instagram	Twitter
2-① (i) 一般ユーザからの申告・ 削除要請への受付窓口・ 受付態勢	<ul style="list-style-type: none"> <li>ご意見・ご要望フォームを通じてカスタマーサポート窓口にて受付。</li> <li>法的な判断が必要であれば専門部署にエスカレーション。</li> <li>プロバイダ責任制限法に基づく請求を受け付ける専用の書面受付窓口も用意。</li> </ul>	<ul style="list-style-type: none"> <li>■Google検索 コンテンツ削除をリクエストするためのウェブフォームを提供。</li> <li>■YouTube 不適切と思われるコンテンツを匿名で報告できる。</li> <li>■Googleマップ ポリシー違反のクチコミをビジネスプロフィールから報告できる。</li> </ul>	<ul style="list-style-type: none"> <li>■LINE VOOM、LINEオープンチャット、LINE LIVE CSフォームからの問い合わせという形で連絡を受け、対応マニュアルに則り対応。</li> <li>■LINE BLOG、livedoor Blog 入信フォームや郵送で依頼を受け、対応。</li> </ul>	<ul style="list-style-type: none"> <li>コミュニティからの報告、コンテンツ・モデレーション・チームによる審査、技術を組み合わせてコミュニティ規定を施行。</li> <li>違反コンテンツの特定には主に人工知能(AI)を利用。AIは、案件の優先順位付けにも役立ち、最も有害で一刻を争うコンテンツを優先的に審査。優先順位付けは、重大性、バイラリティ、違反の可能性などいくつかの要素に基づいて行われる。</li> </ul>	<ul style="list-style-type: none"> <li>Twitterルールまたは利用規約に違反しているツイート、リスト、ダイレクトメッセージを報告できる。</li> <li>違反の内容によって、個別のツイート、リスト、またはプロフィールから直接報告(スパム、攻撃的または有害なコンテンツ、不適切な広告、自傷行為、なりすましなど)。</li> </ul>
2-① (ii) 主体的な誹謗中傷等の 情報の検知態勢	<ul style="list-style-type: none"> <li>■サービス全体 カスタマーサポートチームが24時間365日、目視によるパトロール</li> <li>■知恵袋 機械学習による低品質な投稿の判定・自動削除を実施。</li> <li>■ヤフーニュース 機械学習による不適切なコメントの解析・検知・自動削除を実施。</li> </ul>	<ul style="list-style-type: none"> <li>■Google検索 <ul style="list-style-type: none"> <li>自分に関する情報を含むページを検索結果から削除することをリクエストできる。</li> <li>降格シグナルを使用し、不当な方針が示されているサイトの順位が検索結果の中で下がるようにしている。</li> </ul> </li> <li>■YouTube <ul style="list-style-type: none"> <li>人間と機械学習を組み合わせ問題のあるコンテンツを大規模に検出。</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>機械的なチェックでは、事前にシステムに登録している禁止用語やルールと照合し、規約や法令に反した投稿の場合、投稿後ただちに自動で投稿が非表示にする仕組み。</li> <li>人の目によるチェックでは、機械的なチェックでは判断のつかなかったケースについて、事前に用意した判断基準や事例集と照合して規約や法令に反する投稿がないかを担当者が確認。コミュニティ規約や法令に反すると判断した場合には、その投稿が非表示。</li> <li>AIを活用した違反画像の検知(現状、違反画像検知AIをLINE全サービスで適用中。LINE全サービスで自動処理を実装予定)</li> <li>AIを活用した違反テキストの事前検知(ユーザへのポップアップ警告)(各サービスで逐次開発予定)</li> </ul>	<p>独自のテクノロジーを利用して違反コンテンツを積極的に検出し、利用者によって報告される前にその大半を削除している。</p>	<ul style="list-style-type: none"> <li>Twitterでは人による調査とテクノロジーを組み合わせてTwitterルールの徹底に取り組んでいる。</li> <li>日本語を含む世界の言語で24時間365日、グローバルに対応する専任のスペシャリストチームが、ポリシーに則り、削除も含めたさまざまな強制的対応を用いて公正かつ公平にルールを執行。</li> <li>Twitterチームが会話の流れを理解し、執行措置を実施する前に必要な情報を確実に得るため、構造的な人から直接話を聞かなければならない場合がある。</li> <li>攻撃的な行為をより効率的に検出するための技術的に優れたツールも活用しており、現在では、利用者からの報告がなくても、プラットフォームに投稿される攻撃的なコンテンツの65%以上が、スペシャルチームによる目視確認が必要として、システムに検知されるまでになっている。なお、2021年1月1日～6月30日に、Twitterルールに違反した470万件のツイートを削除。</li> </ul>
2-③ 申立てに対するポリシー 違反等の審査プロセス、 案件ごとの優先度の決定 方法	<ul style="list-style-type: none"> <li>原則、申し立て順に対応。</li> <li>申し立て内容の権利侵害の重大性に鑑み、緊急的に対応することもある。</li> </ul>	<ul style="list-style-type: none"> <li>■YouTube <ul style="list-style-type: none"> <li>現実社会で生じうる実害が大きいと想定されるケースを優先。</li> <li>公認報告者による報告は、YouTubeのポリシーとその施行に関するスペシャリストが優先的に審査し、必要に応じて対処。</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>入信フォームや郵送で依頼があったものから順次対応。</li> <li>捜査機関から削除要請(殺害予告等)や情報開示請求を受けた場合、投稿記事削除及び発信者情報開示仮処分命令申立書を受けた場合は優先して対応。</li> <li>モニタリングは、ユーザのアクセスが多いチャットやコメント欄における投稿から優先的に対応。投稿内容に応じてポイントを付与し、高いものから順に対応。</li> </ul>	<p>どのコンテンツを公開し、どのコンテンツを削除するかを決定するための一貫した枠組みである、コミュニティ規定に則り対応。</p>	<ul style="list-style-type: none"> <li>Twitterルールに違反する行為があった場合、対応。</li> <li>特定の国において、権限のある機関から適切な適用範囲を定めた有効な要請を受けた場合に、措置を実施することができる。</li> </ul>



	ヤフー	Google	LINE	Meta	Twitter
対象サービス名	Yahoo!知恵袋 Yahoo!ニュース(コメント欄) Yahoo!ファイナンス掲示板	Google 検索 YouTube Google マップ	LINE VOOM LINEオープンチャット LINE LIVE LINE BLOG、livedoor Blog	Facebook Instagram	Twitter
2-④ 一般ユーザからの申告や削除要請に対応する部署・チームの規模・人数及び部署・チームに関する日本国内の拠点の有無、日本における責任者の有無	<p>【日本】</p> <ul style="list-style-type: none"> <li>投稿の削除等のパトロール業務を専門とするパトロール部隊に約70名が在籍名し、24時間365日対応。</li> <li>プロバイダ責任制限法に基づく法的な判断をする部署については、7～8名程度で対応。</li> <li>日本国内の拠点：有</li> <li>日本における責任者：有</li> </ul>	<p>【日本】</p> <p>日本の案件だけに従事する人間の数を正確に示せないが、日本をカバーする人数は必要に応じて拡充していく。</p> <p>【グローバル】</p> <p>Google 全体で20,000人を超えるスタッフが24時間365日体制でGoogleのプラットフォーム全体でコンテンツの確認と削除に取り組んでいる。</p>	<p>【日本】</p> <ul style="list-style-type: none"> <li>削除要請については、各サービスにつき10名程度で対応。</li> <li>モニタリングについては、各サービスにつき15名～90名程度で対応。</li> <li>東京(LINE株式会社)及び福岡(LINE Fukuoka株式会社)に拠点。</li> <li>日本における責任者を置いている。</li> </ul>	<p>【日本】</p> <p>回答なし</p> <p>【グローバル】</p> <ul style="list-style-type: none"> <li>審査チームは、コンテンツ審査を担当する社員と、パートナーに従事するコンテンツ審査担当者によって構成。</li> <li>審査チームには退役軍人や法律専門家に加え、児童の安全、ヘイトスピーチ、対テロ対策などの各種ポリシー分野に精通する規定施行の専門家などがある。</li> <li>さまざまな企業から参画している15,000人以上の審査担当者と協力。</li> <li>Metaの審査チームはグローバル体制の下、24時間年中無休で審査。</li> <li>20以上の拠点でコンテンツの審査を行うこれらのチームは全体で日本語を含む数十種類に及ぶ言語に対応。</li> </ul>	<p>【日本】</p> <p>回答なし</p> <p>【グローバル】</p> <ul style="list-style-type: none"> <li>社内のグローバルチームは、日本語を含むTwitterの全サポート言語において、年中無休でTwitterルールの強制的対応を管理。</li> <li>適切なプロセスを提供し、公正で偏りがなく、バランスが取れ、人権を尊重する方法でTwitterルールが適用されるよう注力。その基盤が、コンテンツモデレーションでの透明性と説明責任に関するサンタクララ原則および複数関係者が関わるその他のプロセスの精神。今後もTwitterルールの適用方法に関して将来の報告書で提供できる情報範囲の拡大にリソースを投じる。</li> </ul>
2-⑥ 削除等への苦情や問い合わせに対する苦情受付態勢及び苦情処理プロセス	<p>■サービス全体</p> <ul style="list-style-type: none"> <li>ご意見・ご要望フォームを通じてカスタマーサポート窓口にて受付。</li> <li>法的な判断が必要であれば専門部署にエスカレーション。</li> </ul> <p>■ヤフーニュース</p> <ul style="list-style-type: none"> <li>削除に疑義が生じた場合に、運営者に削除理由などを問い合わせるための専用手続きを用意。</li> <li>理由の通知のみならず、措置理由が合理的であったか否かの判断も含めて確認。</li> </ul> <p>■他のサービス</p> <p>削除されたユーザーからの異議申し立てについて、専用窓口を設置するとともに、削除理由の開示フローなどについて検討を進める。</p>	<p>■サービス全体</p> <p>再審査請求をすることで、その措置に不服を申し立てることができる。</p> <p>■Google検索</p> <p>ウェブサイトの管理者向けのGoogle Search Consoleを通じて通知。</p> <p>■YouTube</p> <ul style="list-style-type: none"> <li>動画が削除されたクリエイターに通知が発行され、再審査を請求するための手順が記載されたリンクを提供。</li> <li>クリエイターが再審査請求を提出すると、その請求は審査担当者に送られたのち、判定が維持または破棄される。</li> <li>再審査は動画の削除について最初の判断を行った審査担当者とは異なる担当者によって行われる。</li> </ul>	<p>■LINE VOOM、LINEオープンチャット、LINE LIVE</p> <p>異議申し立てがあった場合は、その内容に応じて適宜対応。</p> <p>■LINE BLOG、livedoor Blog</p> <p>モニタリングの削除基準に該当しないものについては、プロバイダ責任制限法に則って対応。</p>	<ul style="list-style-type: none"> <li>異議申し立てを行うには、自らのコンテンツが削除または警告付きでほかが入れられたことについて通知を受けた後、「審査のリクエスト」を選択。</li> <li>異議申し立てがなされた場合、コミュニティ・オペレーション・チームが再度審査し、判断が正しかったか確認。判断が誤っていた場合、利用者にその旨を知らせ、投稿、写真、ビデオを復元。</li> <li>報告されたにもかかわらず対処されなかったコンテンツについても、異議申し立ての機会を提供し始めている。</li> <li>児童の搾取に関連する画像など重大な安全上の懸念がある違反は、異議を申し立てることはできない。</li> </ul>	<ul style="list-style-type: none"> <li>スパム行為や攻撃的な行動がみられるという報告を受けてアカウントを一時的に利用できない状態になった場合、一定期間ツイートできなくなったり、自分自身に関する特定の情報を認証するよう求められたりすることがある。</li> <li>手違いでアカウントが凍結またはロックされていると考えられる場合には、異議申し立てをすることができる。</li> </ul>
2-⑦ 異議申立・問い合わせ件数	<p>【日本】</p> <p>サービスにより異なるが、1日あたり数件～1か月あたり数十件程度。また、投稿復帰する機能の有無はサービスによって異なる。</p>	<p>【日本】</p> <p>■YouTube</p> <p>2021年10月から12月の期間に1945件の動画が日本で異議申し立てされています。その結果、812件が元に戻されました。</p>	<p>【日本】</p> <p>2021年1月～12月 [LINE VOOM] 0件 [LINEオープンチャット] 7件 [LINE LIVE] 52件 [LINE BLOG] 18件 [livedoor Blog] 3,952件</p>	<p>【日本】</p> <p>回答なし</p> <p>【グローバル】</p> <p>(URLを記載)</p>	<p>【日本】</p> <p>回答なし</p> <p>【グローバル】</p> <p>回答なし</p>

# プラットフォーム事業者における回答状況（3-①②③関係）（2022年）

	ヤフー	Google	LINE	Meta	Twitter
対象サービス名	Yahoo!知恵袋 Yahoo!ニュース(コメント欄) Yahoo!ファイナンス掲示板	Google 検索 YouTube Google マップ	LINE VOOM LINEオープンチャット LINE LIVE LINE BLOG、livedoor Blog	Facebook Instagram	Twitter
3-① 誹謗中傷等に関する一般ユーザからの申告や削除要請の件数	【日本】 特定の1ヶ月間の件数 ■知恵袋 約5万7千件 ■ヤフーニュース 約21万8千件 ■ファイナンス掲示板 約2万7千件	【日本】 ■Google検索 約2,300件 (2021年1月～6月) ■YouTube 世界で7番目に報告が多い。 ■Google マップ 218,870件 (2021年1月～10月)	【日本】 2021年1月～12月 ■LINE VOOM: 693件 ■LINEオープンチャット: 4,701件 ■LINE LIVE: 52件 ■LINE BLOG: 9件 ■livedoor Blog: 3,644件	【日本】 回答なし 【グローバル】 回答なし	【日本】 18,503件 【グローバル】 42,023件 (2021年1月～6月、 誹謗中傷以外の削除件数 も含む)
3-② (i) 誹謗中傷等に関する一般ユーザからの申告や削除要請に対する削除件数	【日本】 特定の1ヶ月間の件数 ■知恵袋 約5千件 ■ヤフーニュース 約1万2千件 ■ファイナンス掲示板 約1万3千件	【日本】 ■YouTube 24,153本 (2021年10月～12月) ※ 上記のうち、ヘイトスピーチ620 件、嫌がらせやネットいじめ626件 ■Google マップ 32,190件 (2021年1月～10月) ※ 上記のうちヘイト、不適切、ハラ スメントなど3,648件	【日本】 2021年1月～12月 ■LINE VOOM: 20件 ■LINEオープンチャット: 56件 ■LINE LIVE: 9件 ■LINE BLOG: 0件 ■livedoor Blog: 0件	【日本】 ■Facebook 約5万件 ■Instagram 約10万件 (2021年第1四半期～第4四半 期) ※ 日本からのものと予測さ れるユーザーが作成(また は所有)したコンテンツ 【グローバル】 ■Facebook ・ いじめや嫌がらせ: 920万件 (表示頻度0.14%～0.15%) ・ ヘイトスピーチ: 2,230万件 (表示頻度0.03%) ■Instagram ・ いじめや嫌がらせ: 780万件 (表示頻度0.05%～0.06%) ・ ヘイトスピーチ: 600万件 (表示頻度0.02%) (2021年第3四半期)	【日本】 回答なし ※ <u>ヒアリング後提出</u> (p51～54参照) 【グローバル】 5,913,337件 (2021年1月～6月、 誹謗中傷以外の削除件数 も含む)
3-③ 主体的な削除件数 (AIを用いた自動検知機能の活用等)	【日本】 特定の1ヶ月間の件数 ■知恵袋 約2万6千件 ■ヤフーニュース 約3万件 ※上記に加え、AIによる投稿削 除数: 約14万9千件 ■ファイナンス掲示板 約1万7千件	【日本】 3-② (ii)と同回答のため割愛 【グローバル】 ■YouTube 370万本以上の動画と 12億件以上のコメントを削除 (2021年10月～12月)	【日本】 ■LINE VOOM: 0件 ■LINEオープンチャット: 0件 ■LINE LIVE: 0件 ■LINE BLOG: 0件 ■livedoor Blog: 0件 ※主体的な削除以外の対応(非表示化やラベルの付 与等)を行った件数(ヒアリングシート3-③(ii)) ■LINE VOOM: 約78万件 ■LINEオープンチャット: 約180万件 ■LINE LIVE: 約3万件 ■LINE BLOG: 約1,700件 ■livedoor Blog: 約16万件 (2021年1月～12月)	【グローバル】 ■Facebook ・ いじめや嫌がらせ: 920万件 (表示頻度0.14%～0.15%) ・ ヘイトスピーチ: 2,230万件 (表示頻度0.03%) ■Instagram ・ いじめや嫌がらせ: 780万件 (表示頻度0.05%～0.06%) ・ ヘイトスピーチ: 600万件 (表示頻度0.02%) (2021年第3四半期)	【日本】 回答なし 【グローバル】 ・ 攻撃的な行為をより効率的に検出 するための技術的に優れたツール も活用しており、現在では、利用者 からの報告がなくても、プラッ トフォームに投稿される攻撃的なコン テンツの65%以上が、スペシャル チームによる目視確認が必要として、 システムに検知。 ・ なお、2021年1月1日～6月30日に、 Twitterルールに違反した470万件 のツイートを削除。
3-② (ii) 誹謗中傷等に関する申告 や削除要請に対して 削除以外の対応を行った 件数(非表示化や ラベルの付与等)	【日本】 2021年10月19日～12月18日 ■ヤフーニュース コメント欄が非表示となった記 事数: 計216件、 1日あたり平均3.5件	【日本】 ■Google マップ 約120 万件のレビューが公開前 に機械学習によってブロック (2021年1月～10月)	【日本】 2021年1月～12月 ■LINE VOOM: 2件 ■LINEオープンチャット: 0件 ■LINE LIVE: 0件 ■LINE BLOG: 9件※ ■livedoor Blog: 1,364件※ ※記事内容のうち問題のある記 載の削除など	【日本】 回答なし 【グローバル】 回答なし	【日本】 回答なし 【グローバル】 【グローバル】 何らかの強制対応が行わ れたアカウント数: 4,826,539 件 凍結されたアカウント数: 1,240,148件 (2021年1月～6月)

	ヤフー	Google	LINE	Meta	Twitter
対象サービス名	Yahoo!知恵袋 Yahoo!ニュース(コメント欄) Yahoo!ファイナンス掲示板	Google 検索 YouTube Google マップ	LINE VOOM LINEオープンチャット LINE LIVE LINE BLOG、livedoor Blog	Facebook Instagram	Twitter
4－① 発信者情報開示請求を受けた件数 (i)裁判上の開示請求件数	【日本】 23件 ※ 2021年4月1日～12月31日 までに名誉毀損を理由とした 請求があったもののうち、 2022年2月21日時点で結論が 出ているもの	【日本】 回答なし  【グローバル】 回答なし  【参考となる数値】 日本におけるユーザー情報の 開示リクエストは合計74件、アカ カウント数は135件、何らかの情報の 開示に至ったリクエストの割合 は54%（民事訴訟を除く） （2020年7月～12月）	【日本】 14件 （LINEオープンチャット5件、 livedoor Blog9件） （2021年1月～12月）	【日本】 回答なし  【グローバル】 回答なし  【参考となる数値】 日本における政府からの データ開示要求への対応に ついては、透明性レポートに て開示。	【日本】 2,298件 （2021年1月～6月）  【グローバル】 12,396件 （2021年1月～6月）
4－① 発信者情報開示請求を受けた件数 (ii)裁判外の開示請求件数	【日本】 28件 ※ 2021年4月1日～12月31日 までに名誉毀損を理由とした 請求があったもののうち、 2022年2月21日時点で結論が 出ているもの		【日本】 24件 （LINE LIVE2件、LINE BLOG1件、 livedoor Blog21件） （2021年1月～12月）		
4－② 開示を行った件数 (i)裁判上の開示件数	【日本】 18件（うち一部開示6件） ※ 2021年4月1日～12月31日 までに名誉毀損を理由とした 請求があったもののうち、 2022年2月21日時点で結論が 出ているもの	【日本】 回答なし  【グローバル】 回答なし	【日本】 8件 （LINEオープンチャット1件、 livedoor Blog7件） （2021年1月～12月）	【日本】 回答なし  【グローバル】 回答なし  【参考となる数値】 日本における政府からの データ開示要求への対応に ついては、透明性レポートに て開示。	【日本】 2,298件 （2021年1月～6月）  【グローバル】 12,396件 （2021年1月～6月）
4－② 開示を行った件数 (ii)裁判外の開示件数	【日本】 2件 ※ 2021年4月1日～12月31日 までに名誉毀損を理由とした 請求があったもののうち、 2022年2月21日時点で結論が 出ているもの	【日本】 回答なし  【グローバル】 回答なし	【日本】 9件 （LINEオープンチャット0件、 livedoor Blog9件） （2021年1月～12月）		

# プラットフォーム事業者における回答状況（5-①②関係）（2022年）

	ヤフー	Google	LINE	Meta	Twitter
対象サービス名	Yahoo!知恵袋 Yahoo!ニュース(コメント欄) Yahoo!ファイナンス掲示板	Google 検索 YouTube Google マップ	LINE VOOM LINEオープンチャット LINE LIVE LINE BLOG、livedoor Blog	Facebook Instagram	Twitter
5-① 透明性レポートの公開 日本語で閲覧可能か	<ul style="list-style-type: none"> <li>2020年度の透明性レポートを2021年12月公表。</li> <li>日本語で閲覧可能。</li> </ul>	日本語で公開。	LINE Transparency Reportを公表。	コミュニティ規定施行レポートを日本語に翻訳し公表。	「Twitterの透明性に関するレポート」を年2回発行。
5-② 誹謗中傷等の取組の効果分析	<p>■ニュースコメント</p> <p>投稿時注意メッセージの取り組みについて効果を検証しその結果を公開。</p>	<p>■YouTube</p> <ul style="list-style-type: none"> <li>2019年、ガイドラインのボーダーライン上のコンテンツの拡散を制限するため、リコメンデーションシステムの変更を発表。その結果、同年米国では、ガイドラインに違反しないがボーダーライン上のコンテンツに関して、おすすめ動画等からの視聴時間が(チャンネル登録者へのおすすめは除く)70%減少。他の市場でも、おすすめ動画をきっかけとして見つけたボーダーライン上のコンテンツの視聴時間が低下。</li> <li>「YouTubeコミュニティガイドラインの適用について」というレポートを発表。ポリシーに違反した動画やチャンネル、コメントのグローバル全体の削除、再審査請求とそれによって元に戻された動画、人間と機械による報告に関するデータを提供。また、ハラスメントやネットいじめ、暴力的過激主義、ヘイトスピーチといった主要なコンテンツポリシー領域について詳細な分析も提供。</li> </ul>	LINEみらい財団は、青少年のネット利用実態把握を目的とした調査を実施。小中高生に対するネット上のコミュニケーションに関する教育の効果についての分析をまとめた報告書を2021年4月に公開。	FacebookとInstagramにおけるポリシーの施行状況を公的かつ透明性をもって追跡し、FacebookとInstagramを安全で包括的なものにするための継続的な取り組みを示すため、コミュニティ規定施行レポートを四半期ごとに発行。	<ul style="list-style-type: none"> <li>特定のツイートの削除が必要な強制的対応に関する新しい測定基準としてインプレッション(削除前にツイートが獲得した閲覧数)を導入。</li> <li>2021年1月1日～6月30日に、Twitterルールに違反した470万件のツイートを削除。削除されたツイートのうち、削除前のインプレッションが100未満のものは全体の68%、100～1,000のツイートは24%、1,000を超えるツイートは8%。今回の報告期間におけるルールに違反するツイートのインプレッションが全ツイートのインプレッションに占める割合は0.1%未満。</li> </ul>

※本資料は、各社提出のヒアリングシート等から事務局において主な項目を抜粋したものであり、詳細は各社提出資料を参照。

## Twitter : japan-specific numbers for January-June 2021

Total Unique Accounts Reported for JP (Jan - Jun 2021):	~2.16M	
By Rule:		
<b>Twitter Rule</b>	<b>Country</b>	<b>Unique Accounts Reported</b>
Abuse/harassment	Japan	879029
Child sexual exploitation	Japan	23093
Hateful conduct	Japan	839167
Impersonation	Japan	96023
Non-consensual nudity	Japan	202359
Private information	Japan	136253
Promoting suicide or self-harm	Japan	243066
Sensitive media	Japan	692244
Violence and extremism	Japan	551191

"**Accounts reported**" reflects the total number of unique accounts that users reported for potentially violating the Twitter Rules.

**Note:**

1. Countries are assigned by inferring an account's country based on the following data from a user; sign-up location, user selected location and IP addresses. While every effort is made to present accurate data, we cannot guarantee that the assigned country reflects the true location of a user at a specific instance when an enforcement action occurred.
2. For definitions of Twitter Rules, see <https://transparency.twitter.com/en/reports/rules-enforcement.html#2021-jan-jun>
3. There might be scenarios where the same account can be reported or can be actioned for more than one policy. In that case, the account will be counted against all policies that it was reported or actioned for.
4. Data is as of January 2022.

Total Unique Accounts Actioned for JP (Jan - Jun 2021):	~280K	
By Rule:		
<b>Twitter Rule</b>	<b>Country</b>	<b>Accounts Actioned</b>
Abuse/harassment	Japan	39589
COVID-19 misleading information	Japan	133
Child sexual exploitation	Japan	77325
Civic integrity	Japan	11
Hateful conduct	Japan	29563
Illegal or certain regulated goods or services	Japan	39981
Impersonation	Japan	5652
Non-consensual nudity	Japan	2785
Platform manipulation - multi account abuse	Japan	34
Private information	Japan	3436
Promoting suicide or self-harm	Japan	11042
Sensitive media	Japan	79051
Terrorism/violent extremism	Japan	868
Violence	Japan	1484

“**Accounts actioned**” reflects the number of unique accounts that were suspended or had some content removed for violating the Twitter Rules.

Total Accounts Suspended for JP (Jan - Jun 2021):	~130K	
By Rule:		
<b>Twitter Rule</b>	<b>Country</b>	<b>Accounts Suspended</b>
Abuse/harassment	Japan	3989
COVID-19 misleading information	Japan	4
Child sexual exploitation	Japan	76839
Hateful conduct	Japan	2061
Illegal or certain regulated goods or services	Japan	33309
Impersonation	Japan	5103
Non-consensual nudity	Japan	621
Platform manipulation - multi account abuse	Japan	34
Private information	Japan	520
Promoting suicide or self-harm	Japan	479
Sensitive media	Japan	7077
Terrorism/violent extremism	Japan	868
Violence	Japan	963

“**Accounts suspended**” reflects the number of unique accounts that were suspended for violating the Twitter Rules.



Twitter : japan-specific numbers for January-June 2021

Unique Content Removed for JP (Jan - Jun 2021):	~246K	
By Rule:		
Twitter Rule	Country	Unique Content Bounced
Abuse/harassment	Japan	72302
COVID-19 misleading information	Japan	180
Child sexual exploitation	Japan	1127
Civic integrity	Japan	11
Hateful conduct	Japan	43732
Illegal or certain regulated goods or services	Japan	17868
Impersonation	Japan	603
Non-consensual nudity	Japan	6274
Private information	Japan	6772
Promoting suicide or self-harm	Japan	15310
Sensitive media	Japan	80771
Violence	Japan	1966

“**Content removed**” reflects the number of unique pieces of content (such as Tweets or an account’s profile image, banner, or bio) that Twitter required account owners to remove for violating the Twitter Rules.

### これまでの主な意見

#### 【誹謗中傷への対応に関する事業者ヒアリング・ヒアリングシートについて】

- ツイッターからの、なぜ日本の拠点、責任者に対する情報を求めるのかという質問については、基本的には日本において十分なリソースをこの問題に対して割いているか、そのための指標として知りたいということ。そのため、拠点は日本になく外国にあるが、日本語対応の責任者や、日本語対応の不適切投稿の責任者が何人いるか、ということでもよい。【森構成員 (第33回会合)】
- 日本固有の情報が公開できない理由として、日本語をやればいいのか、日本居住者をやればいいのか、日本アカウントをやればいいのかという話があったが、それが決まらないから公開できないというのは納得し難い。AIを利用する関係上、言語別のデータは取れるのではないかと推測するため、日本語の数を示していただければそれで結構。我々の場合は、日本語、日本人、日本の国土というのがわりとそろっているので、それで十分意味のあるデータになる。もし日本語での数が出せないということであれば、日本アカウントなどそういうことでも結構。【森構成員 (第33回会合)】
- Metaの説明で責任のある数字を公表したいということだが、それは当然だと思う。しかし、責任のある数字を公表したいから、何も公表できないということは納得できない。何らかのレポートを公表されている、数字を示されているという話があったが、そうであったとしても、我々の検討会のヒアリングシートでは回答がなかったということ。責任のある数字を公表したいのであれば、それはどこかで公表しているということではなく、まさしく政府の検討会において公表すべき。【森構成員 (第33回会合)】
- 表示頻度についてもMetaから説明いただき、どれだけ見たか、どれだけ見られたかということで理解した。それははっきり言って、その時点におけるそのSNSの危険性を示す指標にすぎないと思う。例えば、政府、学校による周知啓発が効果を発揮すれば、頻度は自動的に下がる、安全になるんだと思う。もちろんその安全性についても教えていただくのは結構なことだが、我々が本当に知りたいのは、不適切情報に対する対策、取組をSNS事業者がどのようにやっているか、日本語についてどの程度やっているかということが知りたいので、表示頻度では不十分。【森構成員 (第33回会合)】

### これまでの主な意見

#### 【誹謗中傷への対応に関する事業者ヒアリング・ヒアリングシートについて (続き)】

- 昨年の内部告発問題についての言及がなかったのは残念。日本でも広く報道されていて、ポジティブという言葉が一貫して使われていました。ただここでいうネガティブ、ポジティブが、一般的な日本人が考えるそれと同じかということは疑問だと思っている。怒りや偏見を招くコンテンツが、ネガティブなものとしてこれまで捉えられてきたのかということ。昨年の内部告発事件においては、そういうものが放置されていたのではないかということが明らかになったわけなので、そのカテゴリーの再分類、ネガティブを捉え直すというようなことができたのかどうかということに、我々は興味を持っている。【森構成員 (第33回会合)】
- なぜ誹謗中傷あるいは違法有害情報対策について、プラットフォームの事業者の方々に透明性をお願いしてきたか。これは繰り返しになるが、令和2年の緊急提言、それから昨年の中間取りまとめでも事業者の方々と話をし、また事業者の方々が加入している事業者団体からのパブリックコメントなどでも丁寧なやり取りをさせていただいたと私としては思っていました。要するに、事業者の方々が日本で非常にしっかり取り組んでいるということは確かだろうと思うが、それがどのように具体的に成果が上がっているのかどうか、あるいは他の要因や他のプレーヤーと、どこに責任があって、どのような連携をしているのか、誹謗中傷の問題、あるいは違法有害情報対策がうまくいくのか。そういうこととの関係で、情報の提供や透明性の確保を消費者に対して、あるいはこの検討会について、お願いをしているという状況。【穴戸座長 (第33回会合)】
- 数字がない、あるいは数字が取れないということになると、例えば日本のプラットフォーム上で、現実にはこういう被害があって、例えば自殺などを含む被害が発生しているという数字は具体的に出てきてしまう。それによって、プラットフォーム事業者の方々の対策が取れていないのではないのか。また、日本においては、オフラインで何かがあったというものもあると思いますが、オンラインの誹謗中傷が完全にオンラインで、そうでないものの数字がかなり具体的に報道されるなど見えている部分があるため、こういう状況を日本法人によく理解いただき、また日本法人から本社にしっかり説明をいただいて、意味のある対話をさせていただきたい。具体的にこういう指標がより実効的なのではないかということがあればその御提案をいただいて、この場で議論をさせていただきたい。【穴戸座長 (第33回会合)】
- 同じ方向を向いているはずなので、それに向けてどのような連携ができるかということについて、引き続き協力をお願いをしたいと思う。【穴戸座長 (第33回会合)】

### これまでの主な意見

#### 【日本における透明性確保について】

- 去年の10月頃のメタの内部告発の問題は海外や日本で報道されており、本検討会としてもそれを無視してはならない。特にこの調査との関係では、内部告発の中に、言語や国によって取組のリソースのかけ方が全然違うということもあったので、日本においてどうなのかということを明らかにする必要性は以前にも増して高くなっていると思う。【森構成員（第35回会合）】
- 方向性として自由放任は難しく、共同規制で考えていくという方向性は異論ないが、その中で、方向性そのものについてどう考えていくのかがはっきりしていない。ゴールをきちんと定める必要がある中で、ゴールを普遍的・一般的なものにするというのが相当難しく、その議論を延々と続けるのもかなり難しいのではないと思う。逆に社会的なコンセンサスをどうプラットフォームやSNSに反映させるかという仕組みを作るなど、様々な方向性で、SNSが自ら考えていくための仕組みのようなものを作っていくことが先ではないかと思う。特に透明性でこういった評価をしていくところ、重要なポイントが1つあり、例えばこの評価を見て、ここはできている・できていないや、優劣という判断ではなく、この方法だとこのようリスクが残っていると、いったリスク評価をしていくようなモニタリングとその評価が、実は一番重要なのではないかと思う。その中で、各社がリスクに対応するための仕組みを義務化していくといった方向性で考えるほうが、現実的なのではないか。【寺田成員（第35回会合）】

### これまでの主な意見

#### 【森構成員 (第35回会合Meta社に対する事後質問)】

- 2021年10月の内部告発によって、貴社サービスの投稿の表示順位を決めるアルゴリズムについて以下のことが判明したとの報道がありますが、事実ですか。①より多くのクリックやコメント、シェアを得られるものがアルゴリズム上重視されており、②その結果、怒りや分断を招くものがアルゴリズム上重視されている
  - 前記内部告発を受けて、表示順位を決めるアルゴリズムを変更しましたか。変更した場合にはどのように変更したか教えてください。
  - 前記の内部告発によって、言語別のヘイトスピーチ対応のための費用の割合は、英語 (米国) が38%と最も多く、英語以外の言語への対応は極端に少ないこと、および偽情報対策の予算は、米国が87%を占めていること、が報道されています。言語別の①ヘイトスピーチ対応、②誤情報対応、③誹謗中傷いじめ等対応において、日本語が占める予算の割合を①～③のそれぞれについて教えてください。
  - インスタグラムの若者に対する悪影響について、「日本を含む6カ国の約2万人を対象にした調査」が貴社によって2019年ごろに行われたと報道されています。①この報道は事実ですか。②事実である場合、日本における調査結果と調査結果を受けた貴社の対応について教えてください。

#### 【Meta社回答】

- Metaは、利用者と利用者にとって最も重要な人とコト、例えば、友人、興味、そして世界をひとつにまとめて、つながることができます。弊社のゴールは、利用者にとって最も価値のあるコンテンツを確実に見ることができるようになることであり、それはフィードを眺めているときや、検索で特定のものを探しているとき、グループ機能でコミュニティとつながっているとき、そしてウォッチ機能でビデオを見ているときなどを含みます。そのために、それぞれの画面上で表示されるすべてのコンテンツを整理し、利用者が個人的に最も興味を持つと思われるものを、それぞれの画面上の一番上に表示するようにしています。

### これまでの主な意見

#### 【Meta社回答 (続き)】

- つまり、Facebookでコンテンツに出会うとき、そのコンテンツの表示のされ方は利用者に合わせてカスタマイズされています。利用者につながっている友達、参加したグループ、「いいね！」を押したページ、コメント、「この投稿は時間をかける価値がありますか」などのフィード調査に対する回答はすべて、利用者にとって何が最も重要かを示す情報です。弊社が使用するアルゴリズムは、こうしたすべての活動に基づいて利用者が最も関心を持ちそうなものを予測し、何を表示するかを決定しています。このようにパーソナライズすることにより、利用者は体験に積極的に参加することができるようになります。
- 弊社は常に、有害なコンテンツをプラットフォームから排除するビジネス上のインセンティブを持っています。そのため、弊社はプラットフォーム上で許可されないものについて明確なルールを設け、安全とセキュリティのために本年だけでも50億ドル以上を費やす予定です（この規模は、事業規模を考慮して調整したとしても、他のどのテクノロジー企業よりも多いものです。）。さらに弊社では、4万人以上もの者が、弊社のアプリ上で利用者の安全を守るという業務のみに従事しています。
- 弊社のアプリで利用者がより適切で価値のあるコンテンツを見ることによって、利用者が弊社のアプリでポジティブな体験をするお手伝いをするには、長期的にみて弊社ビジネスの最善の利益となります。弊社のシステムは、挑発的なコンテンツを薦めるようには設計されていません。むしろ、その逆のを行うように設計されています。センセーショナルであったり、誤解を招くコンテンツ、不当にクリックさせようとするコンテンツ、独立したファクトチェック団体によって虚偽であると判断されたりしたコンテンツは、コンテンツの配信が減らされ、その結果、フィードで下位に表示されることとなります。
- 弊社のランキングへのアプローチの詳細については、下記URLをご参照ください。
- <https://transparency.fb.com/features/ranking-and-content/>
- 弊社は、オープンさと透明性を保ち、コンテンツの配信方法を利用者がよりコントロールできるようにすることにより、人々がアルゴリズムの仕組みについて理解を深め、より信頼と自信を持つことにつながると考えています。このため、弊社は、フィードに表示するコンテンツを決定する際の考え方・価値観、ガイドライン、システムに関する情報を定期的に公開しています。
- <https://about.fb.com/news/2021/12/changes-to-news-feed-in-2021/>

### これまでの主な意見

#### 【Meta社回答 (続き)】

- 例えば、昨年9月には、コンテンツ配信ガイドラインに関する情報を公開しました。このガイドラインでは、フィードでの配信が制限されるコンテンツの種類（コミュニティ基準に違反する可能性が高いコンテンツ、ポリシー違反を繰り返す人が投稿したコンテンツ、性的または衝撃的な内容を含むページへのリンクを含むコンテンツなど）について説明しています。
- <https://about.fb.com/news/2021/09/content-distribution-guidelines/>
- また、利用者がフィードに表示されるコンテンツをより自由にコントロールできるようにするための新しいツールの導入も続けています。例えば、次のようなものがあります。
- 「Why Am I Seeing This?」ツールは、特定の投稿がフィードに表示される理由についての情報を利用者に提供し、そのページまたはアカウントからの投稿を見続けるかどうかをカスタマイズできるようにしています。
- Facebookでは、利用者はフィードを時系列で表示し、最新の投稿を一番上に表示することを選択することができます。これは、新しいもの順にコンテンツをランク付けするものです。また、利用者は、自分にとって最も重要な友人やページを選択し、「お気に入り」別にフィードを表示することもできます。
- インスタグラムでは、利用者は「お気に入り」や「フォローフィード」の機能により見る投稿を選択することができます。「お気に入り」は、利用者が選択した特定のアカウントの一覧から最新の投稿を表示し、「フォロー」は、利用者がフォローしている人の投稿を表示します。
- 弊社は、利用者とアルゴリズムの関係をよりよく理解することが、すべての人の利益につながると考えています。それは、オープンで透明性の高い、よりコントロールしやすい環境を利用者に提供することから始まると考えています。

### これまでの主な意見

#### 【ヒアリングの依頼方法等について】

- 参考資料 2 に明記されたとおり、本ヒアリングは、法的に強制された義務ではなく、「任意」でのお願いになる。事業者の皆様からの回答をいただいた時に、改めて感じたが、特にアメリカに拠点を置く事業者の皆様は、合衆国憲法第 1 修正上の表現の自由の原理との関係を踏まえた上でもなお、本研究会の取組の趣旨を理解いただいて、今後一層の理解と協力をいただくためには、本研究会からの依頼方法について少し検討が必要。【山口構成員 (第33回会合)】
- 例えば、参考資料 2 の2頁の※印の英語説明について、「誹謗中傷」の訳をOnline Harassmentとすると、本研究会での取組の趣旨が、アメリカに拠点を置く事業者の皆様—特に本社の責任ある立場の皆様—に直接的に伝わるかは悩ましいと、本日のヒアリングでも改めて認識した。また、同頁のdignityに関する記載について、日本の憲法・民法等における「尊厳」という文言の解釈の関係でも、再考の余地がある。そこで、Online Harassmentよりはむしろ、日本市場のサービス利用者の名誉毀損やプライバシー侵害等の "Dignitary Harms" とされてはいかがか。【山口構成員 (第33回会合)】
- こうした「尊厳」に係る深刻な被害に対する、「任意」での取組については、政府だけで何かできることではなく、また、日本だけで何かできるわけでもない。グローバルに事業を展開される関係の皆様と手を取り合って、車の両輪で諸対策を進める必要がある。特に日本法の下では違法 (Illegal) となる情報について、被害者救済は、ユーザー保護という意味でも、関係事業者の皆様にとって最優先の経営課題と思う。その辺りについて、本研究会からの説明も、もう少し工夫できる余地があるかと思う。【山口構成員 (第33回会合)】

### 論点

#### 【業界団体による取組】

- 業界団体等の取組により、ノウハウやベストプラクティスを事業者間で共有されているか。

ヤフーや業界団体（SIA・SMAJ）による取組が進められている。

- ・ヤフー：不適切コメントへの対策として導入している AIのAPIを他社に対して無償提供し、業界全体の不適切投稿の削減につなげようという取組を行っている。
- ・SIA：前述のとおり、「誹謗中傷ホットライン」を運営し、被害者からの誹謗中傷に関する連絡を受け付け、事業者に対し、各社の利用規約に沿った削除等の対応を促す通知を行っている。
- ・SMAJ：SMAJ会員各社のSNS等のサービスにおいて、名誉毀損や侮辱等を意図したコンテンツを投稿するなどの行為を禁止し、利用規約に記載するとともに、禁止事項について啓発広報を実施することを推進。禁止行為への対応の実効性向上のため、これまで各事業者が実施してきたソーシャルメディアにおける課題に対する取り組みを体系化した「ユーザー保護ナレッジデータベース」構築等を通じて、継続的かつ実効性のある利用者保護施策を検討・実施。会員各社の代表的な取組事例は以下のとおり。
  - ・ByteDance：Rethink（コメントする前に再考を促す）機能、コメントをユーザー自身が削除できる機能、コメントフィルター機能（不適切なコメントが自動でフィルターされて非表示になる機能や、すべてのコメントが承認しない限り表示されなくなる機能、自分で設定したキーワードを含むコメントを自動で非表示にできる機能）、ペアレンタルコントロール機能
  - ・株式会社ニュースピックス：フォロワー外のユーザーに対して自分のコメントを表示させる場合は、実名認証を必須とする。また、フォロワーユーザーに対してもアカウント登録後のみコメント表示が可能となる
  - ・モイ：モデレーターを設定することで配信者以外でもコメントの管理を行える

等

## 【諸外国における誹謗中傷等の違法・有害情報に関する政策動向・総論】

- 欧州や米国における政策動向として、日本と同様に、プラットフォーム事業者の違法・有害情報対策に関する透明性やアカウントビリティ確保を求める方向性で検討が進められている。特に、欧州においては、法的規制による透明性確保の方向性での議論が進められている。また、欧州のうちドイツやフランスなど一部では立法によりプラットフォーム事業者に対して削除の義務付け等の法的規制が導入・検討されているが、表現の自由への萎縮効果の懸念の声があることや、立法後に違憲判決により当該規定が削除されるといった状況。

## 【EUにおけるDigital Services Act (DSA)】

- 欧州委員会が提示した法案に大幅に条文を追加した上で、2022年4月に、欧州委員会、欧州議会、欧州理事会は、Digital Services Act (DSA) に関する暫定的政治合意に至った（条文は未公開）。DSAは全ての仲介サービス提供者（プラットフォーム事業者等）に対して、違法コンテンツの流通に関する責任を規定するとともに、事業者の規模に応じたユーザ保護のための義務を規定。
- 対象事業者は、仲介サービス（ISP等）、ホスティングサービス、オンライン・プラットフォーム（オンラインマーケットプレイス、アプリストア、SNS等）、超大規模オンラインプラットフォーム（検索エンジンを含む）を提供する事業者。
- 仲介サービス提供者の違法コンテンツに対する責任について、「単なる導管」、「キャッシング」、「ホスティング」の3類型に分けて違法コンテンツに対する免責条件を規定。一般的モニタリング義務は無いものの、司法及び行政当局からの削除等の措置命令・情報提供命令への報告義務を規定。
- そのほか、自動化ツールの使用義務がないこと、E2E暗号サービスの提供を妨げないこと、サービスの匿名利用を制限する義務を課してはならないこと、個人データを収集しないサービスの利用や支払いを可能にするための合理的な努力することなどを規定。また、権利を侵害された者の申し出によって関連する仲介サービス提供者に対する当該コンテンツの削除または無効にする差止命令を可能にすることや、異議申し立ての機会の確保（当該コンテンツの復元、措置の合法性（必要性及び比例性を含む）への異議を含む）を規定。

## 【EUにおけるDigital Services Act (続き)】

○ 透明で安全なオンライン環境のため、以下の義務を規定。事業者に対して、利用者保護や透明性確保の観点から規制を拡充する方向で、欧州委員会案から追加されている。

① 全ての仲介サービス提供者に対する義務：

- ・ 利用者向け連絡窓口の設置、ダークパターンの禁止、未成年者に配慮した機能・デザイン 等
- ・ わかりやすい利用規約（要約の作成を含む）、サービス提供国の言語による作成 等
- ・ 違法コンテンツ仲介に関する免責条件、異議申し立ての機会の確保、一般モニタリング義務の解除 等
- ・ 透明性報告義務（※1、2）、連絡窓口・国内法定代理人の設置 等

※1 違法コンテンツの通知件数・削除件数（国・地域別の内訳）、信頼できる旗手によって提出された通知の数、措置を取るために必要な時間の中央値、自発的なコンテンツモデレーションの内容、対応に関する苦情の件数等。

※2 VLOPIに該当しない零細・小規模事業者は、透明性報告義務を負わない。

② ホスティングサービス提供者の義務：違法コンテンツの通知受付体制整備、違法コンテンツへの対応に関する理由の通知、刑事犯罪の疑いの当局への通知、コンテンツの無効化又は削除 等③ オンライン・プラットフォームの義務：異議申し立ての対象拡大、障害者への配慮、取引事業者の追跡、違法コンテンツへの対応に関する苦情受付体制整備、裁判外紛争解決機関の利用、信頼された旗手、不正な通知・反論に対する対策、オンライン広告の透明性確保、特別カテゴリー情報（性的指向、宗教、民族性など）を用いたターゲティング、未成年に対するターゲティングの禁止 等④ 超大規模オンライン・プラットフォームの義務：サービスのリスク評価実施・リスク軽減措置の実施、外部監査の実施・公表、レコメンダー・システム、オンライン広告の透明性の追加、規制当局及び研究者のデータアクセス、コンプライアンス・オフィサー設置、行動規範、危機対応、ディープフェイクの通知義務、手数料 等

○ モニタリング及びエンフォースメントに関して、各加盟国はDSAの執行責任者であり調査権限等を持つデジタルサービス調整官を設置。欧州委員会は超大規模オンライン・プラットフォームをモニタリング。義務違反の場合、前年度の総売上高の最大6%の罰金等を科すことが可能。

**これまでの主な意見****【インターネット上の違法・有害情報を巡るEUの最新動向について】**

- (DSAの議会修正案について) 利用規約の内容及び形式というところについて正面から踏み込んできたのが今回極めて重要なところ。まさにこの約款規制に近いようなものを、これからルールとしてどのようにプラットフォームの方々にしっかり基本権保護を含めてつくっていただくかといったようなこと、非常に大きな論点になってくる。それからDSAは、オンライン・プラットフォーム、超巨大プラットフォームに対して、実際追加的な義務を課すということが非常に大きなポイントだが、より上位レイヤーの概念である仲介サービス全体に対する義務もかなり様々な形で変わってくる。特に議会では、いわゆるBKYC、ビジネスカスタマーの身元確認義務が当初案ではオンライン・プラットフォーム等に限定されていたのを、CDNやDNS等のサービスに対してもしっかり課していくべきと言われている。海賊版対策等において情報発信者をどのようにつかまえるのかといった意味でも非常に大事なところであるから、今後の動きをフォローしていく価値がある。  
【生貝構成員 (第32回会合)】
- 透明性報告義務がしっかり法的義務としてあるところは、このプラットフォーム研究会でも再三出てきた論点であるため注視する必要がある。オンライン・プラットフォームのレコメンダー・システムの透明性の後半にある「レコメンダー・システム」については、ターゲティング広告と並んで非常に重要なところだが、どういう仕組みで透明性を確保しているのか、若干難しい。【森構成員 (第32回会合)】

## 【英国におけるOnline Safety Bill】

○ 2022年3月17日、英国政府（DCMS）は、プラットフォームなどのオンライン企業による自主規制（※1、2）に依存せず、政府が規制（注意義務の設定など）を行い、当該規制が守られているかをOfcomが監視するといった内容の法案を議会に提出。

※1 規制対象コンテンツ：違法コンテンツ（テロリズム、CSEA、DCMS国務大臣が定めるコンテンツ）、子どもにとって有害なコンテンツ、大人にとって有害なコンテンツ（DCMS国務大臣が定める優先有害コンテンツ、プロバイダが定めるコンテンツ）

※2 規制対象サービスは、英国と関連があり、除外されるサービスではない次のサービス。

- ・ユーザ間サービス（サービスのユーザーによって生成され、またはサービスの利用者によってサービスにアップロードされ、もしくはサービス上で共有されるコンテンツに、サービスの他のユーザーが遭遇する可能性のあるインターネットサービス）
- ・検索サービス（「検索エンジン」であるか、それを含み、ユーザー間サービスでないインターネットサービス）

## ①プラットフォーム事業者の注意義務

ユーザーを安全に保ち、サービスに対する違法で有害な行為などに対処するための合理的な措置を講じるために、プラットフォーム事業者を対象とする以下の義務を定める。

- ・ リスク評価（自社サービスに関連する違法コンテンツのリスク評価）
- ・ 違法コンテンツへの対応（違法なコンテンツを認識・ブロック又は削除し、安全性向上のシステムやプロセスを導入）
- ・ 表現の自由とプライバシーへの配慮（システムやプロセスの導入時にオンライン上の利用者の権利を考慮）
- ・ 報告、救済の仕組み（ユーザが利用しやすく効果的な報告・救済の仕組みを設定）
- ・ 記録の保存（上記の義務について対応結果等を記録）
- ・ 透明性確保（年次の透明性レポートを作成・公表）、Ofcomへの手数料支払い

## ②監視・評価、罰則・罰金の執行（Ofcom）

プラットフォームに課した注意義務がプラットフォーム事業者に遵守されているかを監視・評価する権限、及び罰則や罰金等を課す執行権限をOfcomに付与。

## ③行動規範の作成・提示（Ofcom）

Ofcomは、設定された注意義務の履行・遵守方法を概説したものを行動規範として作成・提示。プラットフォーム事業者は、行動規範に規定されているガイドラインに従うか、あるいは自社の代替アプローチがどの程度効果的に同等以上の影響をもたらすかを説明し、正当化する必要がある。

### 【オーストラリアにおけるオンライン安全法】

- 2021年7月、オンライン上の安全促進及び被害防止のための従来法を継承し、新たな被害に対処するための新法「オンライン安全法」(Online Safety Act)がオーストラリアで成立した。
  - 「オンライン安全法」により、eセーフティ・コミッショナーの権限が従来法よりも拡充された。
    - ・ 成人を対象としたネット暴力や、暴力的行為を助長する等の書き込みなど、従前よりも幅広い範囲（※1、2）の有害なコンテンツについて、ユーザはeコミッショナーに対して苦情申し立てが可能となった。
    - ・ eセーフティ・コミッショナーによる削除通告への対応義務が24時間以内に短縮され、一部の内容のコンテンツに関する削除通告や証拠提示通告に基づく要求を従わない場合、民事罰が科される。
- ※1 規制対象コンテンツ：オーストラリア人の児童を対象としたネットいじめ、オーストラリア人の成人を対象としたネット暴力、性的画像の同意なき共有、忌まわしい暴力的行為を助長する等の書き込み、1995年分類(出版、映画、ゲーム)法で指定される有害コンテンツ
- ※2 規制対象サービスは、オンライン上のあらゆるサービス（ソーシャル・メディア・サービス（SNS事業者）、関連電子サービス（電気通信事業者）、指定インターネットサービス（検索エンジン事業者を含む））。

### 【オーストラリアにおけるソーシャルメディア(荒らし対策)法案】

- 2022年2月10日、連邦通信大臣は「ソーシャルメディア(荒らし対策)法案」を提出したが、議会在解散されたことにより2022年4月11日付けで失効している。
- 2021年9月8日の連邦最高裁判決を受け、メディアが第三者の投稿による名誉毀損の賠償責任を負わなければならない状況を是正するため、SNS事業者の責任を明確にし、SNS上のページ所有者の保護と被害者の救済を図るため、同法案は提出された。
  - ① 発行者(publisher)の地位：第三者による中傷的な書き込みについて、SNS事業者がその発行者として最終的な責任を負う。
  - ② 苦情申立制度：SNS事業者は、中傷被害の申立てに対し72時間以内に発信者に申立てがあったことを通知しなければならない。
  - ③ 発信者情報開示命令：発信者が情報開示を認めないなどの場合、被害者は裁判所を通じてSNS事業者に情報開示を請求できる。
  - ④ 指定法人設置義務：国外のSNS事業者は、豪国内でサービスを提供する場合、指定法人を国内に設置しなければならない。

## 2. フェイクニュースや偽情報への対応に関する 現状と課題

---

## 論点

- 近年では、誹謗中傷の問題に加えて、インターネット上でのフェイクニュースや偽情報（以下「偽情報」という。）の流通の問題が顕在化しており、特に、新型コロナウイルス感染症や米国大統領選挙に関する偽情報に接触する機会が増加している。近日は、主に海外において国家が主体となる偽情報の流布の事例が見られるとの指摘もある。
- この点、特にSNSを始めとするプラットフォームサービスは、経済活動や国民生活などの社会基盤になりつつあり、情報流通の基盤にもなっていること、また、プラットフォームサービスの特性が偽情報の生成・拡散を容易にし、これらの問題を顕在化させる一因となっていると考えられることから、特にこうしたプラットフォーム上の偽情報への適切な対応が求められる。
- これらの状況を踏まえ、我が国における偽情報への対応の在り方について記載した当研究会の報告書を2020年2月に策定・公表した。
- また、インターネット上の偽情報への対策について、プラットフォーム事業者等の取組が十分か、また、その透明性・アカウントビリティが十分果たされているかを検証するために、2021年、当研究会において、日本国内におけるプラットフォーム事業者等に対してモニタリングを実施し、プラットフォーム事業者等における自主的な偽情報に関する対応状況と、それらの取組に関する透明性・アカウントビリティの確保状況の把握を図ったところ、プラットフォーム事業者の偽情報への対応に関する透明性・アカウントビリティ確保状況の取組の進捗は限定的であった。
- モニタリングの結果を踏まえ、2021年9月に中間とりまとめを公表し、個別の投稿の削除を義務づけることには、より一層極めて慎重であるべきとする一方で、我が国における透明性・アカウントビリティ確保が実質的に図られない場合には、透明性・アカウントビリティの確保方策に関する行動規範の策定及び遵守の求めや法的枠組みの導入等の行政からの一定の関与について、具体的に検討を行うことが必要とされた。
- 偽情報対策について、プラットフォーム事業者の取組は十分か。2021年の中間とりまとめなどを踏まえ、産学官民の取組は適切に進められているか。

## 1 自主的スキームの尊重

・民間による自主的な取組を基本とした対策を進めていくとともに、総務省はモニタリングと検証評価を継続的に行っていくことが必要

## 2 我が国における実態の把握

・PF事業者の認識や実態把握と調査結果とのギャップが生じていることから、プラットフォーム事業者は、自らのサービス上で生じている我が国における偽情報の問題について適切に実態把握を行い、研究者が分析を行うために必要な情報の無償で情報提供が行われることが望ましい

## 3 多様なステークホルダーによる協力関係の構築

・「Disinformation対策フォーラム」「Innovation Nippon」等において継続的に議論・研究が行われることが望ましい

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

・プラットフォーム事業者は、リスク分析・評価に基づき偽情報への対応を適切に行い、それらの取組に関する透明性・アカウントビリティ確保を進めていくことが求められる

・総務省は、これらの取組に関するモニタリングと検証評価を継続的に行っていくことが必要。どのような方法や情報により偽情報への適切な対応が図られているかどうかを評価することが可能かについて引き続き検討が必要

## 5 利用者情報を活用した情報配信への対応

・広告の種類・対応に応じてリスクや問題の差異を分析したうえで、特に、偽情報を内容とする広告の配信やターゲティング技術の適用については、そのリスクを踏まえ、より注意深い対応と、それに伴う透明性・アカウントビリティ確保が求められる

## 6 ファクトチェックの推進

・プラットフォーム事業者・ファクトチェッカー・ファクトチェック推進団体・既存メディア等が連携し、取組がさらに進められることが期待される

・我が国におけるファクトチェック結果を積み重ねて分析を行うことにより、偽情報の傾向分析やそれを踏まえた対策の検討が行われることが望ましい

## 7 情報発信者側における信頼性確保方策の検討

・現代のメディア環境に対応した情報の信頼性の確保の在り方について、既存メディア・ネットメディア・プラットフォーム事業者など関係者の間で検討を深めていくことが望ましい

・ミドルメディアを中心とした偽情報の生成・拡散・流通メカニズムに関する実態把握と分析も踏まえ、検討を深めていくことが望ましい

## 8 ICTリテラシー向上の推進

・偽情報の特徴を踏まえながら引き続きICTリテラシー向上施策が効果的となるよう取り組むことが必要

## 9 研究開発の推進

・ディープフェイク等に対抗するための研究開発や事業者の対応が進められることが望ましい

## 10 国際的な対話の深化

・偽情報に関する政策について国際的な対話の深化を深めていくことが望ましい



・次回以降のモニタリングにおいて、偽情報への対応に関して、事業者が自主的な報告を行わない場合や、我が国における透明性・アカウントビリティ確保が実質的に図られない場合には、透明性・アカウントビリティの確保方策に関する行動規範の策定及び遵守の求めや法的枠組みの導入等の行政からの一定の関与について、具体的に検討を行うことが必要

論点

【1. 多様なステークホルダーによる協力関係の構築】

- 多様なステークホルダーによる多面的な議論が行われているか。プラットフォーム事業者、ファクトチェック機関、メディアなど関係者間の協力が進んでいるか。
- 偽情報流通の実態を正確に把握し、その対応について多面的に検討すべく、産学官民の多様なステークホルダーによる協力関係の構築を図り、対話の枠組みを設けることを目的として、2020年6月に「Disinformation対策フォーラム」が設立された（事務局：SIA）。
- 議論のスコープについてはインターネット上のSNS等で個人のユーザが発信する「デマ」の類を対象とし、①フェイクニュースを巡る実態・最新の研究成果、②事業者における取組、③関連団体における取組、④今後の取組の方向性を内容とする中間とりまとめが2021年3月に公表された。
- さらに、ファクトチェックとリテラシー向上を中心として、民間部門の自主的かつ実効的な取組の方向性や具体化に向けた取組や方向性について議論を重ね、結果をとりまとめた報告書が2022年3月に公表された。

これまでの主な意見

【ステークホルダー間の連携について】

- **改めて団体同士の連携が大事で、一丸となってDisinformation対策をしていくことがよく分かった。【木村構成員（第34回会合）】**

## 論点

- 我が国におけるインターネット上の偽情報等の実態はどうか。

<みずほリサーチ&テクノロジーズ発表（第36回会合）>

- 2022年2月の調査によると、偽情報関連用語の認知度は、「フェイクニュース」（94.6%）、「ディープフェイク」（58.1%）、「ファクトチェック」（46.4%）であるが、過去調査（3期分）から時系列に比較するといずれの用語も認知度は上昇。
- 「インターネット上のメディア（SNSやブログなど）」において偽情報を見かける頻度（週1回以上）は31.1%、インターネット以外のメディア（テレビや新聞など）においては26.9%、「まとめサイト（例：2ちゃんねる・5ちゃんねるまとめブログ、●●速報など）」においては39%。
- 2022年2月調査と2021年3月調査を比較すると「インターネット上のメディア（SNSやブログなど）」は10.1ポイント増、「インターネット以外のメディア（テレビや新聞など）」は6.9ポイント増。
- 情報の真偽を見分ける自信は、調査対象国（日、米、英、仏、独、韓）中最も低い（28.8%）。年代別に見ると、「自信がない」が「自信がある」よりも高くなった年代の方が多く、「自信がある」が高くなったのは「10代」のみ。
- 偽情報対策に取り組むべき主体として期待されるのは、「報道機関、放送局、ジャーナリスト」（48.2%）、「個人（信用性の低い情報を拡散しないなど、自分自身のリテラシー向上）」（42.8%）、「政府機関」（42.2%）、「ソーシャルメディアサービスを提供している事業者」（34.1%）、「インターネット検索サービス事業者」（25.6%）、「わからない・誰でもよい」（12.5%）の順に多い。
- 新型コロナウイルスに関する情報について、特に信用できる情報源やメディア・サービスは、「自国の政府機関のウェブサイトや情報配信」（38%）、「公共放送局（テレビ・ラジオ・ウェブサイトなど）」（35.7%）、「民間放送局（テレビ・ラジオ・ウェブサイトなど）」（31.5%）、世界保健機関（WHO）や専門機関のウェブサイトや情報発信（22.2%）、新聞（新聞・ウェブサイトなど）（18.5%）の順に多い。

論点

【2. ファクトチェックの推進】

○ ファクトチェックの活性化のための環境整備が進められているか。海外においてファクトチェックはどのように推進されているか。

- Disinformation対策フォーラムにおける報告書において、ファクトチェックに関してさらなる取組を進めるにあたっての留意点や限られたリソースを効果的に用いるために踏まえるべき点について整理され、公表されている。
- 同報告書では、取組を進めるにあたっての留意点として、ファクトチェックを行う目的の明確化と表現の自由の尊重や、ファクトチェック結果をいかに拡散するか、また、国際ファクトチェックネットワーク（IFCN）との整合性が挙げられている。また限られたリソースを効果的に用いるために踏まえるべき点として、ファクトチェックの総合的な実施、対象分野・案件の選定、ファクトチェック団体に求められる運用とガバナンス（中立性と公平性の担保や定期的な第三者レビュー等）等が挙げられている。

## 論点

## 【3. 情報発信者側における信頼性確保方策の検討】

- インターネット上におけるメディア全体の情報の信頼性の確保方策について、メディアやプラットフォーム事業者等の関係者間で検討が進められているか。
- 「Disinformation対策フォーラム」において、一般社団法人日本新聞協会・日本放送協会・一般社団法人日本民間放送連盟がオブザーバ参加し、プラットフォーム事業者・メディア関係団体・有識者との対話や情報共有が進められている。
- Disinformation対策フォーラムにおいて共有されたメディア関連団体の取組は以下のとおり。
- 新聞社の取組：全国に取材網や取材拠点を有し、発信前に社内でも何重にもチェックするほか、不確かな情報に対する取材や検証、デマを打ち消す記事の発信とともに、誤った情報によって引き起こされた差別や偏見に対しても、対応する記事を発信。教育界と協力し、NIE（Newspaper in Education）という、教育現場で新聞を活用して情報リテラシーの向上を推進する取組を実施。
- 放送事業者の取組：放送法の規律に加え自律的な取組を行っており、日本民間放送連盟は放送倫理基本綱領（日本放送協会と共同作成）、放送基準、報道指針等を定めるほか、SNS上の情報については、投稿者のプロフィールの確認や投稿者へのコンタクト等を通じて、事実か否かを確認の上での報道を行う等の対応を実施。日本放送協会はSoLT（Social Listening Team）というチームを立ち上げ、SNSの情報をリアルタイムで観察し、事件事故の最新の状況や変化をいち早く捉え、キャッチした一次情報を報道に繋げている。また、「フェイク・バスターズ」という番組の放送により、積極的にファクトチェック情報を発信し、これら取組の方針は、「NHK 放送ガイドライン2020（インターネットガイドライン統合版）」において公表。  
また、東京大学鳥海不二夫研究室、TDAI Lab、日本放送協会では、ネット上の盛り上がりには留意が必要なニュースに注意喚起のマークの掲出や、意見の分断状況や少数者による拡散状況を可視化などを行う「公共放送コンテンツのオンライン配信を通じた健全な情報空間の形成に関する共同研究」を社会実証として実施。

## 論点

## 【4. ICTリテラシー向上の推進】

- 政府や各ステークホルダーにおけるリテラシー向上の推進に向けた活動は十分か。既存のICTリテラシー向上の取組に係る整理や様々な主体の連携促進が行われているか。
- 総務省では、偽情報に対抗するリテラシーの向上のための様々な啓発活動を実施。
  - 「インターネットトラブル事例集（2022年版）」において、エコチェンバー・フィルターバブル等のSNSの仕組みや、情報を鵜呑みにしないための確認方法を記載。関係省庁・関係団体との連携のもと、様々なチャネルを活用し、全国の学校、SNSユーザ等へ周知を実施。
  - 「上手にネットと付き合いおう！～安心・安全なインターネット利用ガイド～」を公開。ネットの時代における偽情報に関する特集ページに、情報を鵜呑みにしないための確認方法や日本における偽情報の実態などを周知。
  - 「e-ネットキャラバン」の講座内容に、偽情報への対応を追加。

## 論点

## 【5. 研究開発の推進】

- プラットフォーム事業者は、コンテンツモデレーションに関して、AIを活用した技術について研究開発を推進しているか。
- ディープフェイクなどの新たな技術による偽情報に対抗する技術に関する研究が進められているか。

- 2020年に科学技術振興機構（JST）の戦略目標「信頼されるAI」のもとで、「インフォデミックを克服するソーシャル情報基盤技術」が採択された（研究代表者：国立情報学研究所（NII）越前功教授。2020年12月から5.5年間）。同研究は、「AIにより生成されたフェイク映像、フェイク音声、フェイク文書などの多様なモダリティによるフェイクメディア（FM）を用いた高度な攻撃を検出・防御する一方で、信頼性の高い多様なメディアを積極的に取り込むことで人間の意思決定や合意形成を促し、サイバー空間における人間の免疫力を高めるソーシャル情報基盤技術を確立する。」ことを目的とする。

## &lt;NII越前教授発表（第31回会合）&gt;

- 顔、音声、身体、自然言語などの人間由来の情報をAIが学習することで、本物と見まがう顔を対象としたフェイクメディアの生成が可能になっており、フェイクメディアを用いたなりすまし事例が発生している。
- 顔を対象としたフェイクメディアの生成には、顔全体の合成、顔の属性操作、顔映像の表情操作、顔映像の話し方操作、顔の入れ替えの5つのタイプが存在する。この生成方法を学習することで、顔を対象としたフェイクメディアの判別と改ざん領域の推定を行う手法を開発している。
- ディープフェイク自動検出モデルを利用できるAPIを開発するほか、CRESTのサイトでは、プレプリント、プログラム、データセットが公開され、誰でも簡便・迅速にファクトチェック可能な環境を目指している。
- 一方で、解析に当たって、SNS共有時の圧縮による品質劣化や、多種多様なフェイクメディア作成手法の出現による学習モデルのアップデートの必要性などの課題がある。画像を対象とした自動ファクトチェックを行うことや、将来的に機械学習による真贋判定とファクトチェックを相補的に活用することが重要である。

【5. 研究開発の推進（続き）】

<Spectee発表（第31回会合）>

- SNS等のデータを解析し、災害発生に係る情報の可視化や予測を行うサービスを官公庁や企業に提供。
- サービスの提供に当たり、SNS上のデータについて、過去のデマ情報を基に学習したAIを用いて、自然言語解析や画像解析を実施するとともに、誇張表現や勘違い等デマ情報のパターン分けを実施。
- SNSのデータにアクセスするに際し、例えばTwitterの情報については、基本的にTwitter社から購入。購入費用は高額なため、災害分野に絞ってキーワードベースで取得しているが、画像映像情報もあり、テキストに書かれているキーワードだけの情報抽出は困難。ほとんどのプラットフォーム事業者はAPIを公開しており、キーワードでの絞り込みが多い。別の情報の取得の仕方をプラットフォーム側もできることが望ましい。

これまでの主な意見

【「顔を対象としたフェイクメディアの生成と検出について」について】

- 一般には、画像を人間は視認性でチェックするが、システムの仕掛けとして有効なデータ、または検証ができるようなデータ構造にしていかないと、フェイクかどうかについて非常に深遠なテクノロジーだけでカバーしていくのには限界が出てのではないかと。ほとんど同じにできてしまえば、全く見分けがつかなくなってしまう。データに対しても見分けられるようにすると、究極的には単なるローデータではなくて、そこに有効な情報を付加するという、そのようなデータの形で流通させていく世界を考えることが重要【新保座長代理（第31回会合）】

## 論点

## 【6. 国際的な対話の深化】

## ○ 偽情報に関する政策について、国際的な対話が深められているか。

- 2022年2月3日に、第27回日EU・ICT政策対話を実施。また、2021年11月17日に第12回、2022年4月8日に第13回の日EU・ICT戦略ワークショップを実施。日本側からは、「インターネット上の誹謗中傷への対応に関する政策パッケージ」及び中間とりまとめについて紹介。EU側からは、Digital Services Act(DSA: デジタルサービス法)及びDigital Market Act(DMA: デジタル市場法)の概要や欧州内での議論状況等について紹介があった。
- 2021年11月11日に、第12回インターネットエコミーに関する日米政策協力対話（政府間会合）を実施。日本側からは、「インターネット上の誹謗中傷への対応に関する政策パッケージ」及び中間とりまとめについて紹介。米国側とインターネット上の違法・有害情報対策についての問題意識を共有した。
- 2022年3月23日には、第6回日独ICT政策対話を実施。日本側からは、インターネット上の誹謗中傷及び偽情報に関する政策動向を紹介。
- 2022年4月28日には、「未来のインターネットに関する宣言」立ち上げイベントが開催され、開かれた、自由で、グローバルで、相互運用可能で、信頼性のある、安全な、一つのインターネットへの支持を呼びかけられた。その一環として、違法・有害情報対策における人権及び基本的自由の保護※について確認された。  
※ 政府、規制当局、およびオンラインプラットフォームを含むデジタルサービスにより、オンライン上の違法・有害なコンテンツや活動を削減するために取られる行動は、検閲、嫌がらせ、または脅迫の恐れなしに、意見の多様性と多元性を奨励しながら、表現の自由の権利を含む国際人権法に合致するべきであるという我々のコミットメントを再確認する。
- 2022年5月には、金子総務大臣が欧州委員会ヴェステアー上級副委員長及びブルトン委員、英国DCMSナディーン・ドリーズ大臣と会談し、デジタル空間における利用者保護について意見交換を行い、協力関係の深化を図る方針を確認した。
- 2022年5月10日、11日には、G7デジタル大臣会合がドイツにおいて開催され、事業者の違法・有害情報への対応措置に関する透明性・アカウントビリティを世界・国・地域のレベルにおいて、該当するポリシーごとに果たすことを含む、eSafety等について議論。

○ プラットフォーム事業者の偽情報への対応については、一部で取り組みの進展が見られるものの、及び透明性・アカウントビリティ確保の取組の進捗は限定的。多様なステークホルダーによる協力関係の構築、ICTリテラシー向上に関しては、まだ十分とは言えないものの、我が国においても取組が進められつつある。

**偽情報対策に関するプラットフォーム事業者における透明性・アカウントビリティ確保状況の主な評価項目の比較**

項目	ヤフー	Google	LINE	Meta	Twitter
我が国における実態の把握 (2. 関係)					
多様なステークホルダーによる協力 関係の構築(3. 関係)					
プラットフォーム事業者による適切な 対応及び透明性・アカウントビリティ の確保 (4. (2)(3)(4)関係)					
利用者情報を活用した情報配信への 対応(5. 関係)					
ファクトチェックの推進(6. 関係)					
情報発信者側における信頼性確保 方策の検討(9. 関係)					
ICTリテラシー向上の推進 (7. 関係)					
研究開発の推進(8. 関係)					

→今回のモニタリングの質問項目について、事業者が「回答を控えた理由」は次ページの通り

# (再掲) プラットフォーム事業者による対応のモニタリング結果 (2022年) 「回答を控えた理由」及び「今後の対応方針」 (1 / 2)

事業者名	回答を控えた理由及び今後の対応方針
ヤフー	<p>&lt;Yahoo!知恵袋&gt; 4. (3) 削除等の対応 (①偽情報等に関する申告や削除要請の件数、②(ii)アカウントの停止数) 理由: 申告理由について、「偽情報」という切り口で件数をカウントしていないため。 対応方針: 対処すべき投稿がどのようなものかを含め、偽情報への対応につき有識者の意見も聞きながら検討を進める。</p> <p>&lt;Yahoo!ニュース&gt; 4. (3) 削除等の対応 (①偽情報等に関する申告や削除要請の件数～⑥不正な申告や削除要請への対策の方法・仕組み、対応件数) 理由: 2022年3月にYahoo!ニュースコメントポリシーを改定したばかりであり、件数等定量データ収集中であるため。 コメントポリシー&lt;<a href="https://news.yahoo.co.jp/info/comment-policy">https://news.yahoo.co.jp/info/comment-policy</a>&gt; コメントポリシーの改定&lt;<a href="https://news.yahoo.co.jp/newshack/information/comment_policy_20220322.html?msckid=443a99f8a9bf11ec9a2f08d6d1a109ec">https://news.yahoo.co.jp/newshack/information/comment_policy_20220322.html?msckid=443a99f8a9bf11ec9a2f08d6d1a109ec</a>&gt; 対応方針: 来年度以降、回答準備を進めていく。</p>
Google	<p>&lt;誹謗中傷&gt; *2-4 一般ユーザからの申告や削除要請に対する部署・チームの規模・人数 及び部署・チームに関する日本国内の拠点の有無、日本における責任者の有無 →日本の案件の審査と承認は、日本語話者だけではなく、さまざまな部署のグローバルなチームがそれぞれの専門性を持って対応しておりますため、日本の案件だけに従事する人間の数を正確にお示しすることができませんが、日本をカバーする人数は必要に応じて拡充していきます。 *3-3-2 (ii)主体的な削除以外の対応 (非表示化やラベルの付与等) を行った件数 →このような取り組みの個別件数については、グローバルに開示しておりません。引き続きユーザーの安全を守るために、どのような仕組みが機能しているかについては、さまざまな形で開示しています。Google 検索について詳細は、How Search Works(Google 検索の仕組みを確認する) &lt;<a href="https://www.google.com/intl/ja/search/howsearchworks/">https://www.google.com/intl/ja/search/howsearchworks/</a>&gt; をご覧ください。YouTube についての詳細は、How YouTube Works (YouTube の仕組みとは) &lt;<a href="https://www.youtube.com/howyoutubeworks/">https://www.youtube.com/howyoutubeworks/</a>&gt; をご覧ください。Google マップについての詳細は、Google のクチコミの仕組みについてのブログ &lt;<a href="https://japan.googleblog.com/2022/02/google.html">https://japan.googleblog.com/2022/02/google.html</a>&gt; をご覧ください。 *4-1 発信者情報開示請求を受けた件数 (i)裁判上の開示請求件数、(ii)裁判外の開示請求件数 および *4-2 開示を行った件数 (i)裁判上の開示件数、(ii)裁判外の開示件数 →政府からの情報開示リクエストに対して、Google がどのような対応をとったかについて透明性を持たせることを主目的としている弊社の透明性レポートでは、その目的に沿って裁判所を含む政府機関(民事訴訟を除く)におけるユーザー情報の開示についてのデータのみ提供しております。</p> <p>&lt;フェイクニュース&gt; *4-(2)-3 一般ユーザからの申告や削除要請に対応する部署・チームの規模・人数 →日本の案件の審査と承認は、日本語話者だけではなく、さまざまな部署のグローバルなチームがそれぞれの専門性を持って対応しておりますため、日本の案件だけに従事する人間の数を正確にお示しすることができませんが、日本をカバーする人数は必要に応じて拡充していきます。 *4-(3)-2-ii →何をもち「日本」のアカウントとするかは、技術的な課題がございます。弊社は、継続して利用可能な情報を拡大しますが、正確さとデータの妥当性には最大限の注意を払って行います。</p>

(再掲) プラットフォーム事業者による対応のモニタリング結果 (2022年)  
「回答を控えた理由」及び「今後の対応方針」 (2/2)

事業者名	回答を控えた理由及び今後の対応方針
LINE	<p>偽情報・フェイクニュース対策のヒアリングにおける質問6-④「日本におけるファクトチェック推進に向けた取組」について、 「回答を控えた理由」 当初設問には「6-④」が無かったところ、他方「6-⑤」において特定非営利活動法人ファクトチェック・イニシアティブ (FIJ) に正会員として参加し、日本におけるファクトチェック推進に寄与している旨既往回答済みであったことから、凡そ同様の趣旨の回答となることによる重複を避けた方が良いものと判断させて頂いたため。 「今後の対応方針」 政府や関係団体等と連携して、日本におけるファクトチェック推進に向けた検討を進める。</p>
Meta	<ul style="list-style-type: none"> <li>・誤情報は、弊社のコミュニティ規定上の他のタイプの言論とは異なり、該当する内容の包括的なリストがありえないことに留意する必要があります。</li> <li>・ポリシーで単に「誤情報」を禁止したとしても、弊社のサービスの利用者にとって有益な通知にはなりませんし、弊社は全ての情報を知り得るわけではないので、そのようなポリシーを行使することもできません。</li> <li>・このため、弊社のポリシーでは、誤情報をさまざまなカテゴリーに分類し、そのような言動を確認したときにどのように扱うかについて明確な指針を示すよう努めています。</li> <li>・また弊社は、新型コロナウイルス感染症に関連する誤情報については、日本における措置件数を具体的に示しています。</li> <li>・弊社は、今後も外部のステークホルダーと協力し、ポリシーの構築と実施、そして誤情報に対応するための投資を続けて参ります。</li> </ul>
Twitter	<ul style="list-style-type: none"> <li>・日本における誹謗中傷や偽情報に関するデータについては、2022年5月11日に提供いたしました。</li> <li>・削除要請等に対応する部者の規模や拠点について、社員の安全面等を考慮し、会社の方針として個別具体の部署の人数は(今回の当該部署に限らず)公表しておりません。また拠点についても、セキュリティの観点から公表しておりません。</li> </ul>

## プラットフォームサービス事業者における偽情報等への対策状況ヒアリング（2022年） 質問項目

## 1. 前提（サービス概要）

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

## 2. 「我が国における実態の把握」関係

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

## 3. 「多様なステークホルダーによる協力関係の構築」関係

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

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

## (1) 偽情報等に関するポリシー

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

## (2) 削除要請や苦情に関する受付態勢・プロセス

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

## (3) 削除等の対応

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

## (4) 透明性・アカウントビリティの確保

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

## 5. 「利用者情報を活用した情報配信への対応」関係

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

## 6. 「ファクトチェックの推進」関係

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

## 7. 「ICTリテラシー向上の推進」関係

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

## 8. 「研究開発の推進」関係

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

## 9. 「情報発信者側における信頼性確保方策の検討」関係

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

## 10. その他

- ① 意見・補足

### 【偽情報等に対するプラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保】

#### ○ ヒアリング結果に関する全体的な傾向

- ・ すべての事業者において、我が国における偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は、前回ヒアリングと比較して、一部で進展が見られるものの、未だ限定的。
- ・ 多様なステークホルダーによる協力関係の構築や、特定のトピックに関する偽情報や誤解を招く情報の流布に関するポリシーの設定、ファクトチェック推進、ICTリテラシー向上に関しては、まだ十分とはいえないものの我が国においても取組が進められつつある。

#### ○ 「我が国における実態の把握」関係（2. 関係）

- ・ すべての事業者において、前回ヒアリングに引き続いて、我が国における偽情報の流通状況に関する実態把握とその結果の分析・公開は行われていない。
- ・ 実態把握に資する取組として、Twitterは研究者に向けて有償・無償の学術研究用データ提供を実施している。

#### ○ 「多様なステークホルダーによる協力関係の構築」関係（3. 関係）

- ・ SIAは「Disinformation対策フォーラム」を主催し、ヤフー、Google、Meta、Twitterがフォーラムに参加。Zホールディングスは「デジタル時代における民主主義を考える有識者会議」を開催。Googleは、Googleニュースイニシアチブを通じたジャーナリズム支援や、国際大学GLOCOMの研究プロジェクト「Innovation Nippon」の支援などを実施。

### 論点

【偽情報等に対するプラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保（続き）】

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

<モニタリング結果>

- ・ ヒアリングによると、いずれの事業者においても、自らのサービス上で生じている我が国における偽情報の問題について適切に実態把握・公表が行われていない。一方で、偽情報についても、すでにプラットフォーム事業者では、偽情報等の不適切な情報への措置を講じる必要性が認識され、あらかじめ対応方針や基準となるポリシーを自主的に設定し、投稿の削除やアカウントの停止等のコンテンツモデレーションと呼ばれる措置を行っている。
- ・ プラットフォーム事業者が自主的に定めている偽情報に関連するポリシーとしては、次のような情報や行為がポリシーで禁止するケースがみられた。
  - 明らかな偽情報（特に健康情報に関するファクトチェックの結果、反事実であることが明らかな情報）
  - COVID-19感染症やワクチンに関する健康衛生に関する誤情報
  - 選挙や市民活動の操作や妨害
  - 武力行使等に関する虚偽の情報
  - 気候変動の存在に関する不確かな情報
  - なりすまし
  - 合成または操作されたコンテンツ（ミスリーディングな改変等）
- ・ 偽情報は、様々な文脈で触れられ、何が誤情報か判断するのが難しく、事業者からは、誤情報について包括的な禁止事項を定めるのは困難であるという回答がみられた。また、誤情報全般を画一的にポリシーで禁止するのではなく、特定のトピック（COVID-19感染症関連やウクライナ侵攻等）に関するポリシーを設定する例が見られた。
- ・ また、プラットフォーム事業者が定めるポリシーについて、社内外の専門家による見直しが行われている例が見られたが、どのような基準でポリシーが設定されるかは十分明らかではなかった。

### 【偽情報等に対するプラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保（続き）】

- 「プラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保」関係（4. 関係）（つづき）
  - ・ すべての事業者において、我が国におけるプラットフォーム事業者による偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は、前回ヒアリングに引き続いて、限定的である。一方で、Google、Meta、Twitterにおいて、新型コロナウイルスに関する誤情報に関するポリシーに基づく投稿等の我が国における削除件数を公表するなど、一部で進展が見られる。
  - ・ヤフー：一般ユーザーが投稿するコメント機能のポリシーにおいて、新型コロナウイルスに関連する情報を念頭に、ファクトチェックにより偽情報であることが明らかな偽情報の投稿を禁じるポリシーが定められた。前回ヒアリングと比較して、明らかな偽情報に関する投稿に関する削除件数について、新たに公開。
  - ・Google：新型コロナウイルス関係の偽情報に関するグローバルなポリシーを具体的に設けており、削除等の対応を行っている（新型コロナウイルス関係以外は関連する他のポリシーでの対応）。前回ヒアリングと比較して、日本からのスパム、誤解を招くコンテンツ、詐欺のポリシーに違反する可能性のあるコンテンツの件数と削除件数について、新たに公開。
  - ・LINE：新型コロナウイルスに関連して、ガイドラインにおいて、健康に深刻な被害をもたらす誤情報や社会的混乱をもたらす情報や、政府が公式に否定する情報等について削除を実施しており、前回ヒアリングと比較して、コロナワクチン接種のデマ情報拡散に関する削除件数について、新たに公開。
  - ・Meta：新型コロナウイルス関係や選挙・政治関係等に関して、グローバルのポリシーを具体的に設けており、削除・警告表示・表示順位抑制等の対応を行っている。前回ヒアリングと比較して、新型コロナウイルス感染症及びワクチンの誤情報に関するポリシーに違反として削除された件数について、グローバルの件数及び日本における件数を、本年のヒアリング後に、新たに公開。
  - ・Twitter：新型コロナウイルス関係・選挙関連・ディープフェイク等に関して、グローバルのポリシーを具体的に設けており、削除やラベルの付与を実施。前回ヒアリングと比較して、グローバルの削除件数も具体的に公開しているが、新型コロナウイルスに関する誤解を招く情報に関するポリシーについて、日本における件数についても、本年のヒアリング後に新たに公開。

### 【偽情報等に対するプラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保（続き）】

#### ○ 「利用者情報を活用した情報配信への対応」関係（5. 関係）

- ・ 広告表示先（配信先）の制限：いずれの事業者においても、一定の禁止規定や特定の基準に満たない媒体・サイトへの広告配信を制限する規定が設けられている。
- ・ 広告の出稿内容に関する制限：いずれの事業者においても、偽情報を内容とする広告について、一定の禁止規定が設けられている。
- ・ 政治広告に関する制限：LINE及びTwitterでは政治広告が禁止されている（Twitterはグローバルに一律禁止）。Googleでは、日本において一般的な政治広告は許可されているが、選挙広告は禁止。
- ・ ターゲティング技術の適用に関する規定：上記の「広告の出稿内容に関する制限」に該当する偽情報の広告や政治広告はそもそも配信が禁止されているためターゲティング技術の対象とならない旨を説明している事業者が多い。なお、ヒアリング結果からは、（出稿が許されている）政治広告について、どのようなターゲティング技術に関する対応が行われているかについては明確になっていない（ヤフーは政治的意図をもってターゲティングを行うことを禁止する規定を2022年1月から適用、Googleは広告主に対して政治所属に基づくターゲティングを禁止）。
- ・ 出稿者の情報や資金源の公開、透明性レポートの公開等：すべての事業者において、広告に関する何らかの透明性確保方策が行われている。

ヤフー：広告出稿者の情報の明示を広告サイト内に明示することを広告掲載基準に規定。政治広告の資金源公開については、今後必要に応じて検討。広告審査等に関する透明性レポートを公開。

Google：広告主や所在を公開する広告主認証プログラムを導入し、2021年5月に日本でも導入済み。米国等では選挙に関する透明性レポートを公開（日本では選挙広告は禁止のため未提供）。

LINE：ターゲティング広告に使用される「みなし属性」について、推定のためのアルゴリズムに影響する要素（友だち登録した公式アカウント、購入したスタンプ等）に関する説明をプライバシーポリシー等において明記。

Meta：「広告ライブラリ」において、出資者・金額・リーチした利用者層などの情報を7年間保存し検索可能。

Twitter：政治に関するコンテンツを広告に掲載することを全面禁止されている。

### 【偽情報等に対するプラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保（続き）】

#### ○ 「ファクトチェックの推進」関係（6. 関係）

ヤフー及びGoogleでは、我が国におけるファクトチェック推進団体やファクトチェッカーとの連携等が進められている。

MetaやLINEでは、我が国における具体的な取組は行われていないが、諸外国では取組が行われている。

- ・ヤフー : ヤフーニュースでは、FIJと情報共有連携を行い、メディアや専門家の記事を掲載するほか、20年度に引き続き、21年度も情報連携及び資金面での支援を実施。FIJと情報共有や相互のリンク設置などの連携を行ったほか、メディアや専門家のフェイクニュース関連記事を掲載
- ・Google : Google検索及びGoogleニュースにおいて、一定の基準を満たす第三者ファクトチェッカーにより作成されたファクトチェックのラベルが表示される。APAC Trusted Media Summit 2021 の一環として、日本で誤報会議「DigitJournalism Dojo Summit」を開催し、メディア関係者が参加。そのほか、学生を対象にしたファクトチェック・チャレンジや記者も対象に含めたファクトチェック・ワークショップを実施。
- ・LINE : LINE Taiwanでは、偽情報等を抑制するための官民パートナーシップを2019年に締結し、台湾行政院および4つのファクトチェック機関と協力し、ニュース配信プラットフォームであるLINE TODAY上に偽情報等のファクトチェック検証の結果を表示する取組を実施。ユーザー自らが疑わしい投稿等をLINEから送信し、検証結果を得ることができるチャットボットを提供。我が国ではファクトチェッカーとの連携は行われていない。また、特定のニュースについて外部のファクトチェック結果を掲載。
- ・Meta : グローバルな取組として、独立したファクトチェッカーがコンテンツを審査し、ラベル付けを行う。ラベル付けされたコンテンツをシェアしようとする利用者や過去にシェアした利用者に通知。ファクトチェッカーが偽情報と評価したコンテンツは、表示順位低下などの表示抑制を実施。偽情報を繰り返し配信する違反者に対して、配信数抑制や収益化や広告機能停止など措置を実施。国際ファクトチェックネットワーク(IFCN)に認定された独立した第三者ファクトチェッカーと提携を行っているが、我が国では、IFCN認定ファクトチェッカーが存在しないため未実施。
- ・Twitter : 誤解を招くツイートに対してユーザが意見や文脈を追加できるBirdwatch機能を米国において実装。今後、日本においても導入予定。

### 【偽情報等に対するプラットフォーム事業者による適切な対応及び透明性・アカウントビリティの確保（続き）】

#### ○ 「ICTリテラシー向上の推進」関係（7. 関係）

すべての事業者において、我が国において偽情報の問題に対応したリテラシー教育に関する取組が行われている。

- ・ヤフー : 2021年度において教育現場（大学）において情報リテラシー教育を実施する予定。こうしたリテラシー教育の授業・研究を通じて、啓蒙コンテンツを作成しYahoo!を中心にグループで活用。“フェイクニュース”などの不確かな情報に惑わされないなどのための「Yahoo!ニュース健診」を提供開始。
- ・Google : 「Google News Initiative」を筆頭に、数多くのメディア・リテラシープログラムを立ち上げ。我が国において、ファクトチェックを含む主要なオンラインリテラシーのトピックを扱うオンライン・リテラシー・カリキュラムをこれまでに10万人以上の中学生・高校生に提供。その他、「Grow with Google」「ウェブレンジャー」等のプロジェクトを実施。その他研究活動やジャーナリズム人材育成に係る活動を実施。
- ・LINE : ワークショップ授業・講演活動等を2012年より累計で約10,000回以上実施。LINEみらい財団では、教育工学や授業デザインを専門とする研究者と共同で、独自の情報モラル教育教材の開発を行い、ウェブサイトで公開。
- ・Meta : アジア太平洋地域の専門家と協力して、「みんなのデジタル教室」を立ち上げ。日本の中学生・高校生を対象に、偽情報が発信される動機や、情報を受け取る側の視点を考え、偽情報を見分けるための様々なポイントについて、アクティビティを通じて学ぶ授業を展開。新型コロナウイルス感染症に関する誤情報から身を守るために必要な知識を共有するために、「新型コロナウイルス感染症に関する誤情報に対処するための6つのヒント」を展開。
- ・Twitter : 公式アカウントやヘルプセンターページでの情報提供を実施。@UNESCOの「Global Media and Information Literacy Week 2019」に合わせたハンドブックの公開や、はAds For Goodプログラムを通じてメディアリテラシーに取り組むNGOを支援。

#### ○ 「情報発信者側における信頼性確保方策の検討」関係（9. 関係）

- ・「Disinformation対策フォーラム」において、既存メディアや有識者との情報共有や協議が進められている。
- ・いずれの事業者も、新型コロナウイルス感染症関係の政府（内閣官房や厚労省等）など、公共性の高い情報をサービス内で優先表示させる仕組み等を積極的に実施。
- ・いずれの事業者も、ロシアによるウクライナ侵攻に関して、信頼性のあるニュースメディアや国連機関、第三者ファクトチェック機関などによる情報の優先順位を上げたり特設ページを開設するなど取り組みを実施。
- ・ニュースサービスでは、提携媒体の選定基準策定やテクノロジーの活用等により信頼性の高い情報を掲載。

# プラットフォーム事業者における回答状況（2 関係）

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
2-① 偽情報等の発生・拡散状況を把握できる体制分析・調査の有無	<p>■ サービス全体</p> <p>偽情報等の発生・拡散状況の把握については、いずれの情報が偽情報といえるのか判断が難しいものも多くあるが、特定のサービスに限定せず、今後調査、分析のための体制づくりを検討していく予定。</p> <p>■ Yahoo!知恵袋</p> <p>偽情報に限るものではないが、パトロール部隊による違反投稿の監視などを通じて、偽情報の検知、調査、対応の検討を行っている。</p> <p>■ Yahoo!ニュース</p> <ul style="list-style-type: none"> <li>・ ニュース部門、政策企画部門、リスクマネジメント部門など部署間連携によるフェイクニュース対策PJを組成し、対策。</li> <li>・ 海外でのフェイクニュース事例・諸問題など分析調査を実施。</li> <li>・ ニュースサービスにおいて信頼性の高い媒体社とのみ契約し、配信記事について情報の間違いなどがある場合、媒体社に速やかに連絡をするなどの対処。</li> </ul>	<p>■ YouTube</p> <ul style="list-style-type: none"> <li>・ 偽情報を含む、有害となる可能性のあるコンテンツを特定するため人間と機械によるアプローチを組み合わせて活用。有害なコンテンツを積極的に検出してフラグ付けをし、過去に削除されたコンテンツと酷似しているコンテンツを自動的に削除する、2つの柱の方法で機械学習を導入。いずれも収集した情報を使って、コンテンツのパターンを識別できるようにシステムを訓練し、そのパターンを使って、新たにマッチする事例を予測。</li> <li>・ ハッシュ(または「デジタルフィンガープリント」)を使い、既知の違反コンテンツのコピーを、ユーザーに視聴可能となる前に捕らえることもある。ポリシー違反である確度が高いコンテンツ(スパムなど)のみをシステムが自動的に削除し、残りは人間によるレビューを行うためフラグを立てる。</li> <li>・ 特定のポリシー分野に貢献する各国のエキスパートがいる場合には、政府機関やNPOなどに公認報告者プログラムに参加してもらい、審査チームがより迅速にポリシー違反の可能性のあるコンテンツに関する報告を審査できる体制をとっている。日本では、新型コロナウイルス感染症に関する医学的に間違った情報に関するポリシー分野において、コワくんサポーターズ等の専門家が参加。</li> </ul>	<p>■ コンテンツ共有サービス関係</p> <ul style="list-style-type: none"> <li>・ 偽情報等の事例を社内でも共有するなど、発生・拡散状況の把握、調査を実施。</li> <li>・ 各サービスにおいて、機械的なチェックや人の目によるチェックを通じて偽情報も含めて規約や法令に反するコンテンツが拡散されないように努めている。</li> </ul> <p>■ ニュース配信サービス関係</p> <ul style="list-style-type: none"> <li>・ 偽情報等のチェック等を目的として、配信記事の弊社ガイドラインに対するモニタリング、ファクトチェックも含む校正校閲、SNSを通じた情報収集を実施。</li> <li>・ 外部・読者からの問い合わせについて迅速に関係部署に連携するようにしており、実際に発生したトラブルにも素早く対処できる体制を構築。</li> </ul>	<ul style="list-style-type: none"> <li>・ 誤情報は、包括的な禁止事項を明示する方法がないため、コミュニティ規定で取り扱う他の種類の発言とは異なる。「誤情報」を単純に禁止するポリシーでは、情報への完全なアクセスができないため、サービスを利用する人びとに役立つ通知を提供することができず、ポリシーに強制力を持たせることもできない。代わりに、ポリシーでは、誤情報についてさまざまなカテゴリーを明確にし、対象となる発言を見つけたときの対処法を示した明確なガイドラインを設けるよう努めている。それぞれのカテゴリーにおいてのアプローチには、表現、安全、尊厳、真実性、プライバシーに対する弊社の価値観のバランスをとる試みを反映。</li> <li>・ 差し迫った実際の危害のリスクを直接助長する可能性が高い誤情報は削除。また、政治プロセスの機能の妨害を直接助長する可能性のあるコンテンツや、非常に紛らわしい特定の加工されたメディアも削除。このようなカテゴリーにおける誤情報の要素を判断するにあたり、知識と専門性を有する独立した専門家と連携し、コンテンツの真実性や、差し迫った危害のリスクを直接助長する可能性が高いかどうかを評価。</li> <li>・ その他のすべての誤情報について、その情報の表示頻度の抑制に力を入れている。第三者のファクトチェック団体と連携し、プラットフォームで最も拡散しやすいコンテンツの正確性についてレビューおよび評価を実施。また、利用者が閲覧、信頼、共有するコンテンツを自分自身で決められるよう、メディアとデジタルのリテラシー向上のためのリソースも提供。</li> <li>・ 誤情報の拡散と重複することが多い他の分野でのコンテンツや言動を禁止。例えば、コミュニティ規定により、偽アカウント、不正行為、組織的な偽装行為は禁止。</li> </ul>	有

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
2-② 日本における偽情報等の発生・拡散状況、結果公表	<p>【①社会的混乱関係】、 【②健康・医療関係】、 【③選挙・政治関係】 回答なし</p> <p>【④全般・その他】 ■Yahoo!知恵袋 調査結果は内部限りのものとし、外部への公表はない。</p> <p>■Yahoo!ニュース ・信頼性の高い既存メディアと契約の上で掲載しているため、偽情報等の発生・拡散は行われていない。 ・ニューストピックスなどにおいて間違った情報を掲載した場合は、適宜お詫び・訂正を公表。</p>	<p>【①社会的混乱関係】、 【②健康・医療関係】、 【③選挙・政治関係】、 【④全般・その他】 「透明性レポート」で国別の政府によるコンテンツの削除リクエストを含む、コンテンツ削除に関するデータなどを公開。</p> <p>■YouTube 四半期ごとに「透明性レポート」の一部として「YouTube コミュニティガイドラインの適用についてのレポート」を公表。動画やチャンネルの削除、ユーザーによる報告、再審査請求と動画の再公開、そしてポリシーに違反するコンテンツ視聴の割合 (Violative View Rate (VVR)) の公開データを提供。</p>	<p>【①社会的混乱関係】、 【②健康・医療関係】、 【③選挙・政治関係】 回答なし</p> <p>【④全般・その他】 ■コンテンツ共有サービス関係 ・コロナウイルスのワクチン接種における健康被害や社会的混乱を招く可能性のある偽情報の拡散が発生。 ・発生・拡散状況は公表していない。</p> <p>■ニュース配信サービス関係 ・これまで特筆すべき偽情報等が発生・拡散した事実はない。 ・結果の分析などは現在していないが、今後そのようなケースが発生した場合には、情報の公開を含め検討する。</p>	<p>【①社会的混乱関係】、 【②健康・医療関係】、 【③選挙・政治関係】、 【④全般・その他】 回答なし</p>	<p>【①社会的混乱関係】 災害時に偽情報の拡散を防ぐために様々な取り組みを実施。例えば、9月1日の「防災の日」および8月30日～9月5日の「防災週間」に際し、様々な取り組みを実施。</p> <p>【②健康・医療関係】 ・日本でもCOVID-19に関するオープンな会話を保護する取り組みについて最新情報を公開。COVID-19に関して誤解を招く非常に有害な情報をTwitter上から削除する継続的な取り組みに加え、COVID-19ワクチンに関する誤解を招く情報を含むツイートへのラベルを適用。 ・さらなる措置の必要性を判断するためのストライク制の導入を進めている。ポリシーに関するユーザーの理解を促し、有害または誤解を招く恐れのある情報の拡散（特に度重なる中程度および重度のルール違反）を抑制する効果が期待される。</p> <p>【③選挙・政治関係】 ・選挙やその他の市民活動の操作や妨害を目的としてサービスを利用することを禁止。これには、市民活動への参加を抑圧するようなコンテンツや、市民活動に関する日時、場所、方法について誤解を招くコンテンツの投稿や共有が含まれる。市民活動について誤った情報や誤解を招く情報が含まれるツイートには、背景情報を提供するために、ラベルを付けたり、拡散を抑制する場合がある。</p>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
2-② 日本における偽情報等の発生・拡散状況、結果公表					<p>(続き)</p> <ul style="list-style-type: none"> <li>市民活動を操作したり、妨害したりする行為を禁止。これには、市民活動の手続きや環境に関する誤った情報や、誤解を招く情報による妨害行為が含まれる。誤解を招く情報によって、市民活動が直接的に操作されたり、妨害することを目的としていないものの、サービスに混乱を引き起こしたりする場合は、背景情報を提供するためにツイートにラベルを付ける場合がある。選挙に関する重要な情報についての重大な混乱のリスクを考慮して、ツイートに風刺的またはユーモア溢れる要素が含まれている(または含まれようとしている)場合でも、これらの措置を取ることがある。</li> <li>ウクライナ情勢をめぐるツイートに関しては、日本を含む全世界で様々な対策を実施。 <ul style="list-style-type: none"> <li>○情報の信頼性の向上</li> <li>○国家当局関係メディアに関する取り組み</li> <li>○政府アカウントラベルに関する取り組み</li> <li>○ポリシーに基づく強制的対応の拡大</li> <li>○広告に関する取り組み</li> <li>○現地コミュニティへの支援</li> </ul> </li> </ul> <p>【④全般・その他】 誤報への対処方法をまとめた新しいヘルプページを用意。</p>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト(YouTube)</li> <li>・検索エンジン(Google 検索)</li> <li>・デジタル広告プラットフォーム(Google 広告)</li> <li>・オンラインマップ(Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス(LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス(LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
3-① 産学官民の協力関係の構築	<p>■サービス全体</p> <ul style="list-style-type: none"> <li>・「デジタル時代における民主主義を考える有識者会議」(主催:Zホールディングス株式会社)において、デジタルプラットフォームが民主主義に及ぼす影響やデジタルプラットフォーム事業者に期待される役割等について、フェイクニュースなどの例をもとに検討。</li> </ul> <p>■Yahoo!ニュース</p> <ul style="list-style-type: none"> <li>・「Disinformation対策フォーラム」(主催:一般社団法人セーフティネット協会)へ参加し、有効な対策について多様なプレイヤーと議論。</li> <li>・ファクトチェック・イニシアティブ(FII)の活動に賛同し、寄付の実施・法人会員として参加。偽情報・誤情報の検証結果を特設サイトからリンクするなど連携。</li> <li>・媒体社と連携し、ファクトチェック記事の配信を拡充する取組を実施。</li> <li>・有識者とフェイクニュース対策について議論、対策コンテンツ制作の助言を受ける活動を実施。</li> <li>・大学と連携し、フェイクニュースに対するリテラシー教育授業を実施。</li> </ul>	<p>■サービス全体</p> <ul style="list-style-type: none"> <li>・質の高いジャーナリズムを支援すること、また、この分野における多くの関係者間のパートナーシップを支援することに多くのリソースを投入。</li> <li>・さまざまなステークホルダーと協力し、科学的に厳密で学際的なアプローチを用いて、AI開発における思慮深いリーダーシップを促進。より多くの人々が有用なAIアプリケーションを開発できるよう、教材、ベストプラクティス、研究を公開し、AIの知識を共有。</li> <li>・新しいコンテンツポリシーを策定したり、特定の製品機能を検討する際には、誤報やヘイトスピーチなどの分野で外部の専門家と協議。</li> <li>・プラットフォームの完全性を確保し、危機への対応や公共の安全を脅かす可能性のある不正行為を検出した場合に迅速に行動するため、法執行機関からの情報を頼りにしている。プラットフォーム上での組織的な影響力の行使の防止など、国土安全保障上の重要な問題に関して、各国政府とのパートナーシップから恩恵を受けている。</li> </ul> <p>【日本】</p> <p>■Google ニュースイニシアチブ</p> <p>ジャーナリスト向けのデジタルスキル習得にフォーカスした世界的なトレーニングプログラムを提供し、偽情報と戦うための業界団体と連携し、メディアリテラシープログラムに貢献し、より多様で包摂的なニュースエコシステムを支援するためのプログラムを実施。</p> <p>(次ページへ続く)</p>	<p>研究者・官公庁・業界団体等とともに、偽情報等の流通状況や拡散の仕組みにかかる分析や、プラットフォーム事業者に求められる偽情報等への対応にかかる検討を実施。</p>	<ul style="list-style-type: none"> <li>・数十億人にも及ぶプラットフォームの利用者全員と協力関係を構築することは不可能であるため、市民社会組織や活動家グループなどの他者の利害を代表する組織、ならびにデジタル、公民権、差別禁止法、言論の自由、その他の基本的人権の分野における専門家と協力し合う機会を模索。</li> <li>・新型コロナウイルス感染症やワクチンについて、主要な保健機関が否定している誤った主張を削除し、事実確認機関が虚偽と判断したコンテンツの配信を減らし、ワクチン受容の向上に役立つ権威ある情報を提供。</li> </ul>	<ul style="list-style-type: none"> <li>・様々なステークホルダーと協力。</li> <li>・一般社団法人セーフティネット協会が設立したDisinformation対策フォーラムに加盟。</li> </ul>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
3-① 産学官民の協力関係の構築		<p>(続き)</p> <ul style="list-style-type: none"> <li>■ Disinformation対策フォーラム Disinformation対策フォーラムのメンバーとして活動。同フォーラムの第3回セッションにおいて、偽情報対策向けのツールを発表しており、今後も積極的に参加していく。</li> <li>■ 研究 国際大学GLOCOMの研究プロジェクトInnovation Nipponを支援。2019年以降、日本におけるフェイクニュースの実態と対処策について研究を実施。</li> <li>■ 対話 有識者との公式、非公式を問わず意見交換に積極的に参加し、製品やコンテンツモデレーションに関する情報提供を実施。直近、マスコミ倫理想談会全国協議会とこの問題に関する意見交換を実施。</li> </ul>			

# プラットフォーム事業者における回答状況（4（2）関係）

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
4- (2) ①一般ユーザからの申告・削除要請への受付窓口・受付態勢、対応プロセス	<p>○受付窓口・受付態勢</p> <ul style="list-style-type: none"> <li>・申告窓口を用意し、オンラインの申告フォームの他、書面でも受付。</li> <li>・受付窓口はカスタマーサポートが対応。違反申告がなされた投稿についてはパトロール部隊が24時間365日体制で確認を実施。</li> </ul> <p>○プロセス</p> <ul style="list-style-type: none"> <li>・受け付けた申告については、1件1件パトロール部隊が目視により確認し、ガイドライン違反の有無について確認。</li> <li>・偽情報等に関するガイドライン違反に当たるか否かの判断は、公的機関等が公表している情報を参考にし、人間が判断。判断に悩む事案は社内のチームにおいてレビューを実施。</li> <li>・ニュース記事に関する申告については、媒体社に伝達するとともに、ニュースサービスの担当者にも転送され、媒体社で対応がされているかを確認。媒体社および弊社での確認は人間が実施。</li> </ul> <p>○ファクトチェック機関との連携</p> <ul style="list-style-type: none"> <li>・ファクトチェック機関からの専用の申告窓口を設けているわけではないが、もし申告を受けた場合には、外部からの申告一般と同様に、パトロール部隊が1件1件、確認。</li> </ul>	<p>■Google 検索</p> <ul style="list-style-type: none"> <li>・コンテンツを削除する場合、以下の2つの理由がある。:</li> <li>① 法律に違反している場合</li> <li>② その製品やサービスの「ルール」に違反している場合</li> <li>・ 事業を展開している各国の法律を遵守し、その国のプラットフォーム上の違法なコンテンツを削除。さらに、製品やサービスごとに、どのような種類のコンテンツや行動が許容されないかを概説するポリシーを策定し維持。</li> <li>・ このコンテンツポリシーは公開されており、CSAM や法律に基づくリクエストを受けたコンテンツの削除なども扱っている。コンテンツの安全性と責任に関する現実的な懸念事項と、検索エンジンによって情報アクセスの機会を提供する必要性とのバランスを取るためこのコンテンツポリシーは慎重に策定されている。</li> <li>・ 検索機能(ナレッジパネル、オートコンプリート、ウェブ回答など)については、不正行為、操作されたメディア、科学的または医学的な統一見解に基づくベストプラクティスと矛盾する医療のコンテンツなどを除外するため、より高いハードルを設定。ユーザーは、コンテンツポリシーに違反する可能性があると思われる検索コンテンツについてフィードバックを送ることができる。</li> </ul> <p>(次ページへ続く)</p>	<p>○通報対応</p> <p>サービス内の通報ボタンから通報が可能。ユーザからの通報に基づき、ガイドライン違反の可能性があるコンテンツを確認・判断し、必要な措置を実施。</p> <p>○問い合わせ対応</p> <p>お問い合わせフォーム等を通じて連絡いただき、スタッフが対象となるコンテンツの確認や本人確認などを行い、判断。</p> <p>■LINE VOOM、LINEオープンチャット、LINE LIVE</p> <p>CSフォームからの問い合わせという形で連絡を受け、対応マニュアルに則り対応。</p> <p>■LINE BLOG、livedoor Blog</p> <p>入信フォームや郵送で依頼を受け、対応。</p>	<ul style="list-style-type: none"> <li>・ コミュニティ規定を設けている。これは、どのコンテンツを公開し、どのコンテンツを削除するかを決定するための一貫した枠組みを提供するもの。</li> <li>・ コミュニティ規定の施行には、コミュニティからの報告、コンテンツモデレーションチームによる審査、およびテクノロジーを組み合わせて用いる。</li> </ul> <p>○コミュニティからの報告</p> <ul style="list-style-type: none"> <li>・ Facebookでは、ページ、グループ、プロフィール、投稿、写真、ビデオ、コメント、広告など、すべてのコンテンツを報告可能。</li> <li>・ Instagramでも、アプリ内のフィード投稿、ビデオ、ストーリーズ、リール、コメントなど、あらゆるコンテンツを報告可能。</li> <li>・ 世界各地の市民団体とパートナーとしてのネットワークを構築し、専用のチャンネルを通じて連絡を取り、新たな問題を知らせたり、チームが知らないような重要な情報を受領。</li> <li>・ 毎週、数百万件のレポートを処理し、大半のレポートは24時間以内にレビュー。</li> </ul> <p>○人間によるレビュー</p> <ul style="list-style-type: none"> <li>・ 現在、4万人が安全・セキュリティに従事。このチームはグローバルに活動しており、日本語を含む数十の言語で24時間365日コンテンツをレビュー。</li> </ul> <p>(次ページへ続く)</p>	<p>違反の内容によって、個別のツイート、リスト、またはプロフィールから直接報告可能(スパム、攻撃的または有害なコンテンツ、不適切な広告、自傷行為、なりすましなど)</p>

	ヤフー	Google	LINE	Meta	Twitter
<p>サービス分類 (対象サービス名)</p>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
<p>4-(2) ①一般ユーザからの申告・削除要請への受付窓口・受付態勢、対応プロセス</p>		<p>(続き)</p> <ul style="list-style-type: none"> <li>■ YouTube             <ul style="list-style-type: none"> <li>・ ユーザーは、製品内の報告ツールを使って、不適切と思われるコンテンツを匿名で報告することができる。</li> <li>・ 問題を報告しても、自動的にコンテンツが削除されるわけではなく、ガイドラインに沿って審査される。</li> </ul> </li> <li>○その他法的問い合わせ             <ul style="list-style-type: none"> <li>現地の法律に準拠するためのプロセスも用意。</li> </ul> </li> <li>■ Google 広告             <ul style="list-style-type: none"> <li>悪質な広告もしくはパブリッシャーを報告することができる。</li> </ul> </li> </ul>		<p>(続き)</p> <ul style="list-style-type: none"> <li>・ コンテンツモデレーターはさまざまなバックグラウンドを持っているが、人材の採用及び最適化において最も重要と考えているのは、言語と文化的背景。</li> <li>・ コンテンツモデレーターは、公正かつ正確にコミュニティ規定を実施するために不可欠であり、特にコンテンツを取り巻く文脈が重要な場合には重要。</li> <li>○テクノロジーを使って有害なコンテンツを主体的に識別             <ul style="list-style-type: none"> <li>・ 主に人工知能 (AI) を利用して、違反コンテンツを特定。コンテンツの一部がコミュニティ規定に違反していることを確信して、(多くの場合、誰かの目に触れる前に) 自動的に削除することがある。</li> <li>・ AIは審査するコンテンツの量が多いため、コンテンツモデレーターが審査する案件の優先順位付けにも役立っており、最も有害で一刻を争うコンテンツを優先的に審査。優先順位付けは、バイラリティや重大性、違反の可能性などの要素に基づいて実施。</li> <li>・ 四半期ごとに発行されるコミュニティ規定施行レポート (Community Standards Enforcement Report) で、AIを使った進捗状況を公開。</li> </ul> </li> </ul>	

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
4- (2) ③一般ユーザからの申告や削除要請に対応する部署・チームの規模・人数	<p>【日本】</p> <ul style="list-style-type: none"> <li>■UGC機能を有するサービス</li> <li>・パトロール業務を専門とするパトロール部隊には、約70名が在籍し、24時間365日対応。</li> <li>・プロバイダ責任制限法に基づく法的な判断をする部署については、7～8名程度で対応。</li> <li>・そのほか、社外弁護士複数名に諮問して対応。</li> </ul>	<p>【日本】</p> <ul style="list-style-type: none"> <li>・ 回答なし (理由:日本の案件のレビューや承認は、日本語を話す人のみに限定せず、さまざまな部署からそれぞれの専門性を持った人が集まって、グローバルなチームワークで行っているため、日本語の案件だけを担当する人数を正確に伝えることはできない。)</li> <li>・ 今後も必要に応じて、日本への投資を続けていく。</li> </ul> <p>【グローバル】</p> <ul style="list-style-type: none"> <li>・ 2020年にコンテンツの削除等の取り組みに約10億ドルを費した。</li> <li>・ YouTubeを含むプラットフォームの保護を確実にするために、約20,000人の従業員が従事。</li> </ul>	<p>【日本】</p> <ul style="list-style-type: none"> <li>通報対応:120名程度</li> <li>問い合わせ対応:50名程度</li> </ul>	<p>【日本】</p> <ul style="list-style-type: none"> <li>回答なし</li> </ul> <p>【グローバル】</p> <ul style="list-style-type: none"> <li>回答なし</li> </ul> <p>【参考となる数値】</p> <ul style="list-style-type: none"> <li>・ 審査チームは、数多くの職務の一環としてコンテンツ審査を担当する社員と、パートナーに従事するコンテンツ審査担当者によって構成。プラットフォームのコミュニティの多様性の豊かさを反映して、審査チームには退役軍人や法律専門家に加え、児童の安全、ヘイトスピーチ、対テロ対策などの各種ポリシー分野に精通する規定施行の専門家など、さまざまな背景と職歴を持つ人たちが集結。</li> <li>・ 現在、4万人が安全・セキュリティに従事。</li> </ul>	<p>【日本】</p> <ul style="list-style-type: none"> <li>回答なし</li> </ul> <p>【グローバル】</p> <ul style="list-style-type: none"> <li>回答なし</li> </ul> <p>【参考となる数値】</p> <ul style="list-style-type: none"> <li>・ グローバルチームは日本語を含むサポート言語において、年中無休でTwitterルールの強制的対応を管理。</li> <li>・ Twitterルールを客観的に、そして一貫性を持って適用することを目指している。</li> <li>・ Twitterルールに違反すると判断されるコンテンツには強制的な措置が取られる。</li> <li>・ 適切なプロセスを提供し、公正で偏りがなく、バランスが取れ、人権を尊重する方法でTwitterルールが適用されるよう注力。その基盤となるのが、コンテンツモデレーションでの透明性と説明責任に関するサンタクララ原則および複数関係者がかかわるその他のプロセスの精神である。</li> <li>・ 今後も、Twitterルールの適用方法に関して将来の報告書で提供できる情報範囲の拡大にリソースを投じていく。</li> </ul>
4- (2) ④一般ユーザからの申告や削除要請に対応する部署・チームに関する日本国内の拠点の有無、日本における責任者の有無	有	<ul style="list-style-type: none"> <li>・ 具体的なオフィスの場所などは非公開。</li> <li>・ 日本を含めた各国の文化背景を理解できるスタッフも働いている。そうした理解の必要性をYouTubeでも強く認識しており、適切な体制(日本語のネイティブの審査担当者等)を整えている。</li> </ul>	日本国内に拠点:あり 日本における責任者:置いている	審査チームはグローバル体制の下、24時間年中無休でコンテンツを審査。ドイツ、アイルランド、ラトビア、スペイン、ポルトガル、フィリピン、米国など、弊社が全世界に展開している20以上の拠点でコンテンツの審査を行うこれらのチームは全体で日本語を含む数十種類に及ぶ言語に対応。	
4- (2) ⑤削除等への苦情や問い合わせに対する苦情受付態勢及び苦情処理プロセス	<ul style="list-style-type: none"> <li>■UGC機能を有するサービス</li> <li>・ ご意見・ご要望フォームを通じてCS窓口にて受付。</li> <li>・ 回答に当たり法的な判断が必要なものであれば、専門部署へエスカレーションが行われ、回答指示を行う。</li> <li>■Yahoo!ニュース</li> <li>・ コメントの削除に疑義が生じた場合に、運営者に削除理由などを問い合わせるための専用手続きを用意。このプロセスにおいて、理由の通知のみならず、措置理由が合理的であったか否かの判断も含めて確認。</li> </ul>	<ul style="list-style-type: none"> <li>■サービス全体</li> <li>・ 再審査請求をすることで、その決定に不服を申し立てることができる。</li> <li>■YouTube</li> <li>・ 再審査請求を認めている。</li> <li>・ 動画が削除されたクリエイターには通知が発行され、再審査を請求するための手順が記載されたリンクが提供される。</li> <li>・ 再審査請求を提出すると、その請求は審査担当者へ送られたのち、判定が維持または破棄される。</li> <li>・ この再審査は、動画の削除について最初の判断を行った審査担当者とは異なる担当者によって行われる。</li> <li>・ クリエイターには、結果を知らせるフォローアップメールが送られる。</li> </ul>	<ul style="list-style-type: none"> <li>■LINE VOOM、LINEオープンチャット、LINE LIVE</li> <li>・ 異議申し立てがあった場合は、その内容に応じて適宜対応。</li> <li>■LINE BLOG、livedoor Blog</li> <li>・ モニタリングの削除基準に該当しないものについては、プロバイダ責任制限法に則って対応。</li> </ul>	<ul style="list-style-type: none"> <li>・ 利用者がFacebookやInstagramで報告した場合、その報告の状況をサポート受信箱 (Facebook) やサポートリクエスト (Instagram) から確認可能。</li> <li>・ 報告を受け取った日時、報告の理由の確認、内容についてのレビューの結果を通知。</li> <li>・ コンテンツがコミュニティ規定に反すると判断した場合、コンテンツを削除。</li> <li>・ そのコンテンツがポリシーに違反していない場合には、その旨を通知。その際、コンテンツの掲載を継続するという判断に同意できない場合、通常、再審査をリクエストする機会を提供。利用者が再審査をリクエストした場合、コミュニティ運営の専門チームがコンテンツを再度審査。間違いが見つかった場合、利用者に通知し、報告されたコンテンツは非表示または削除。</li> <li>・ 最も有害な種類のコンテンツ(児童搾取画像など)の一部については再審査の申請不可。</li> </ul>	<ul style="list-style-type: none"> <li>・ 異議申し立てをすることによって、アカウントを凍結解除できる場合がある。</li> <li>・ 手違いでアカウントが凍結またはロックされていると考えられる場合には、異議申し立て可能。</li> </ul>

# プラットフォーム事業者における回答状況（４（３）関係）

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
4- (3) ①偽情報等に関する申告や削除要請の件数	<p>【日本】</p> <ul style="list-style-type: none"> <li>■Yahoo!知恵袋 回答なし (理由:「偽情報」という切り口で申告数を計上することが困難)</li> <li>■Yahoo!ニュース 回答なし (理由:2022年3月にYahoo!ニュースコメントポリシーを改定したばかりで件数等定量データ取集中のため。)</li> </ul> <p>※ 来年度以降、回答準備を進めていく予定。</p>	<p>【日本】</p> <ul style="list-style-type: none"> <li>■YouTube 463,135本の動画</li> </ul> <p>※ 日本のユーザーによって、スパム、誤解を招くコンテンツ、詐欺のポリシーに違反する可能性があるとしてフラグが立てられた件数 (2021年10月～12月)</p>	<p>【日本】</p> <p>該当なし (2021年)</p>	<p>【日本】</p> <p>回答なし</p> <p>【グローバル】</p> <p>回答なし</p>	<p>【日本】</p> <p>法的要請による削除要請件数: 18,518件</p> <p>※誤情報に限らない (2021年1月～6月)</p> <p>【グローバル】</p> <p>法的要請による削除要請件数: 43,387件</p> <p>※誤情報に限らない (2021年1月～6月)</p>
4- (3) ②(i)偽情報等に関する申告や削除要請に対する削除件数	<p>【日本】</p> <ul style="list-style-type: none"> <li>■Yahoo!知恵袋 約60件 (2021年12月2日～2022年3月1日)</li> </ul>	<p>【日本】</p> <ul style="list-style-type: none"> <li>■YouTube 24,153本の動画 (2021年10月～12月)</li> </ul> <p>※ 日本国内のIPアドレスからアップロードされたもの</p> <p>※ うち4,302本はユーザーによる最初の報告、さらに1,225本は公認報告者による最初の報告。</p> <p>※ 5,285本はスパム、誤解を招くコンテンツ、詐欺に関するポリシー違反</p> <p>【グローバル】</p> <p>透明性レポート、Ads Safety Report (広告の安全性レポート)</p>	<p>【日本】</p> <p>該当なし (2021年)</p>	<p>【日本】</p> <p>23,000件以上 (パンデミック開始以来)</p> <p>※ 新型コロナウイルス感染症及びワクチンの誤情報に関するポリシーに違反</p> <p>【グローバル】</p> <p>四半期ごとにコミュニティ規定施行レポート(CSER)を発行。CSER第10版(2021年8月)では、FacebookとInstagramにおいて全世界で2400万件以上のコンテンツを新型コロナウイルス感染症に関する誤情報に関連するポリシーに違反していると判断し、削除した旨公表。</p>	<p>【日本】</p> <p>回答なし</p> <p>※ヒアリング後提出 (p51～54参照)</p> <p>【グローバル】</p> <p>COVID-19を巡る誤解を招く情報: 27,935件(対応したアカウント) 33,761件(削除されたコンテンツ)</p> <p>市民の清廉性: 581件(対応したアカウント) 593件(削除されたコンテンツ)</p>
4- (3) ③偽情報等に関する主体的な削除件数 (AIを用いた自動検知機能の活用等)	<p>【日本】</p> <ul style="list-style-type: none"> <li>■Yahoo!知恵袋 約50件 (2021年12月2日～2022年3月1日)</li> </ul>	<p>【日本】</p> <ul style="list-style-type: none"> <li>■Google広告 日本を拠点とする広告主やバイヤーから350万件以上の新型コロナウイルス感染症関連の広告(ショッピング広告を含む)をブロックまたは削除(2020年)</li> <li>■YouTube 18,626本の動画 (2021年10月～12月)</li> </ul> <p>【グローバル】</p> <ul style="list-style-type: none"> <li>■Google広告 2億7,500万件以上の新型コロナウイルス感染症関連の広告を削除(2020年)</li> </ul>	<p>【日本】</p> <ul style="list-style-type: none"> <li>■LINE VOOM: 0件</li> <li>■LINEオープンチャット: 34件(コロナワクチン接種のデマ情報拡散)</li> <li>■LINE LIVE: 0件</li> <li>■LINE BLOG: 0件</li> <li>■livedoor Blog: 0件 (2021年)</li> </ul>	<p>【日本】</p> <p>回答なし</p> <p>【グローバル】</p> <p>新型コロナウイルス感染症とワクチンの誤情報に関するポリシーに違反するコンテンツを全世界で2400万件以上削除。また、同ポリシーに繰り返し違反したアカウント、ページ、グループを3,000以上削除。</p> <p>【参考となる数値】</p> <p>措置を講じた違反コンテンツのうち、利用者の報告を受ける前に検出したコンテンツの数は、上記CSERの中において「事前対応率」として報告。</p>	<p>【日本】</p> <p>回答なし</p> <p>【グローバル】</p> <p>回答なし</p>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
4- (3) ④削除の方法・仕組み (AIを用いた自動検知機能の活用等)	<p>■Yahoo!知恵袋</p> <p>パトロール部隊が人の目によって確認を行い、専門部署と連携しながら削除を行う。</p>	<p>■YouTube</p> <ul style="list-style-type: none"> <li>・人間と機械学習を組み合わせ問題のあるコンテンツを大規模に検出。</li> <li>・2021年10月から12月にかけて、YouTube はコミュニティガイドラインに違反するとして、全世界で37億5421万本以上の動画を削除。これらの動画の91%は、人間ではなく機械によって最初に通知。</li> <li>・機械によって検出されたもののうち、31.7%は1回も視聴されず、38%は1~10回視聴。</li> </ul>	<ul style="list-style-type: none"> <li>・偽情報等に関する削除については、AIの活用を行っていない。</li> <li>・児童ポルノやスパム行為といった違法・有害コンテンツに対しては、AIを用いた自動検知機能を活用。</li> </ul>	<ul style="list-style-type: none"> <li>・独自のテクノロジーを利用して違反コンテンツを積極的に検出し、利用者によって報告される前にその大半を削除。</li> <li>・エンジニア、データサイエンティスト、審査チームはこのテクノロジーを更新してさらに洗練されたものにするために協力して取組。</li> <li>・審査チームはこのテクノロジーを活用して、コンテンツの審査優先順位の設定にも役立っている。</li> <li>・削除している違反投稿・違反アカウントの数は1日あたり数百万件に達する。これらの大半は、背後にあるテクノロジーにより自動的に実行され、しかもそのほとんどが利用者に表示される前に処理。</li> <li>・ときには、違反の可能性のあるコンテンツが検出された後、確認と措置の実施のために審査チームに転送されることもある。</li> </ul>	<ul style="list-style-type: none"> <li>・ポリシーに違反する行為を行っているすべてのアカウントに対して、日本語を含む世界の言語で24時間365日、グローバルに対応する専任のスペシャリストチームが、ポリシーに則り、削除も含めたさまざまな強制的対応を用いて公正かつ公平にルールを執行。</li> <li>・チームは利用者からの報告を受けだけでなく、攻撃的な行為をより効率的に検出するための技術的に優れたツールも活用。その結果、現在、利用者からの報告がなくとも、プラットフォームに投稿される攻撃的なコンテンツの65%以上が、スペシャリストチームによる目視確認が必要として、システムに検知。</li> <li>・ポリシー違反行為を検知するためのシステムへの投資は、継続的に実施。公正さや一貫性をもってルールを適用するため、担当チームが利用者からの報告を受けた際、十分な準備を整えた上で対応できるよう、長期間にわたる徹底したトレーニングプログラムを確立。</li> </ul>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
4- (3) ⑤削除以外の取組 (i)警告表示	<p>■Yahoo!知恵袋</p> <p>新型コロナウイルス関連の投稿がなされた際に、当該投稿のページ上部に「新型コロナウイルスに関する内容の可能性がある記事です。」等の文言を掲出し、厚生労働省等の公的機関のHPを案内する取組を実施。</p>	<p>■YouTube</p> <ul style="list-style-type: none"> <li>・誤った情報が発生しやすいピックアップやニュースに関連するコンテンツへのコンテキストを提供する様々な情報パネルと、パブリッシャー自身を紹介。</li> <li>・情報パネルは、重要なコンテキストを提供するだけでなく、信頼できる情報源をユーザーに紹介。</li> <li>・信頼できる情報源を提供できるよう、「医療/健康情報セクション」、「医療/健康情報パネル」の提供を開始。</li> </ul>	<p>■LINE VOOM</p> <p>偽情報等に限らずモニタリング基準に沿って非表示対応を行なった場合、投稿者のポスターに「LINEの利用規約に違反する可能性のあるコンテンツが含まれていたために削除した」旨のメッセージが表示。</p> <p>■LINEオープンチャット</p> <p>情報の拡散につながるおそれがあるオープンチャットをキーワード単位でピックアップし、オープンチャット内にポップアップで警告を出す取組を実施。</p> <p>■LINE BLOG</p> <p>■livedoor Blog</p> <p>偽情報等に限らずモニタリング基準に沿って警告を送信。</p>	<p>虚偽の情報の拡散抑制に全力で取組。一部の国では、虚偽の情報を識別して審査するため、IFCNが認証した独立した第三者ファクトチェック団体と連携。ファクトチェック団体は事実確認を行い、情報の正確性を評価。</p> <p>○虚偽の情報の配信の抑制</p> <ul style="list-style-type: none"> <li>・フィードでの誤情報の表示順位を下げる</li> <li>・繰り返し虚偽の情報を配信する違反者に対するアクション</li> <li>・虚偽の情報のコピーを見つけるテクノロジーを使用する</li> </ul> <p>○虚偽の情報に関する追加情報の提供</p> <ul style="list-style-type: none"> <li>・虚偽の情報に関する補足情報の提供</li> <li>・フェイクニュースをシェアした利用者への通知</li> <li>・ファクトチェック団体による評価の種類</li> </ul> <p>○虚偽の情報を識別し、フィードバックする手段の提供</p> <ul style="list-style-type: none"> <li>・虚偽の情報の特定方法を確認する</li> <li>・フェイクニュースだと思ふ投稿についてフィードバック</li> <li>・ファクトチェック団体によるコンテンツの評価に異議を申し立てる</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツがTwitterのポリシーに違反しているかどうか、Twitterのチームが確認し、違反していると判断された場合にラベル提示が適用。この評価は自動検知ツールに共有され、Twitter内で同様のコンテンツを特定しラベルを掲示する一連のプロセスをより能動的に行うために利用。</li> <li>・機械学習と自動言語処理の効果があらわれるまでには時間がかかるため、まずは英語のコンテンツを対象とし、徐々に他の言語や文化的文脈に対して同じプロセスを適用。</li> <li>・ラベルは、設定されている表示言語で表示され、Twitterモーメントのガイドラインと方針や正式な公衆衛生またはTwitterルールにリンク。</li> <li>・ストライク制は市民活動の阻害に関するポリシーに定められているストライク制と類似。ラベル適用またはツイート削除要求がアカウントレベルでの追加措置に発展した場合、利用者に直接通知。COVID-19に関するポリシーへの違反が繰り返された場合、その違反によって課された累積ストライク数に応じた強制的対応を実施。</li> </ul> <p>・ ツイートの拡散を抑制</p> <p>・ おすすめから除外</p>
4- (3) ⑤削除以外の取組 (ii)表示順位の低下	<p>■Yahoo!知恵袋</p> <p>表示順位の低下の施策は講じていない。</p>	<p>■Google 検索</p> <ul style="list-style-type: none"> <li>・ ライブテスト、および世界各地の何千人もの訓練された外部の検索品質評価者の両方を含む、厳格なプロセスを採用。Search Quality Rater Guideline (検索品質評価者ガイドライン)に基づいて、評価者は、ユーザーにとって有害、グループにとって有害、誤解を招く、信頼できない、スパムである可能性があるページに対して最低評価を与えるよう指示されている。</li> </ul>	<p>■LINEオープンチャット</p> <p>偽情報等の拡散につながるおそれがあるトークルームについては、メイン画面(オープンチャットサービスの公式画面)やその検索結果に表示されないよう、またホームタブの「おすすめ」エリアなどサービス上で推薦されないよう表示制御を実施。</p> <p>■LINE BLOG</p> <p>■livedoor Blog</p> <p>偽情報等に限らずモニタリング基準に沿って表示順位を低下。</p>		
4- (3) ⑤削除以外の取組 (iii)その他の取組内容	<p>回答なし</p>	<p>■YouTube</p> <p>コミュニティガイドラインに違反しそうに見えるものの明確には違反しないように見えるコンテンツ、いわゆる「ボーダーラインコンテンツ」は、米国でYouTube が視聴しているコンテンツの1%にも満たない数だが、たとえそれ以下でも多すぎると考え、ボーダーラインコンテンツや、有害な方法でユーザーに誤解を与える可能性のある動画がおすすめ動画として視聴者の目に触れる機会を減らすよう取組。</p>	<p>特になし</p>		<ul style="list-style-type: none"> <li>・ いいね、返信、リツイートを無効にする</li> <li>・ まとめられたランディングページ (Twitterモーメント) や関連するTwitterのポリシーなどに、補足説明のためのリンクを表示</li> </ul>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
4- (4) ①コンテンツモデレーションのアルゴリズムに関する透明性・アカウントビリティ確保の方策	<ul style="list-style-type: none"> <li>■Yahoo!ニュース                             <ul style="list-style-type: none"> <li>パートナー記事に関する確認行為はAIは利用しておらず、人的リソースで品質等の確認。</li> <li>誤った情報を発見した場合はパートナーに連絡するフローを導入。</li> <li>コメント欄におけるモデレーションは、記述内容の品質等を判断するAIを活用しているが、虚偽情報を抽出するAI活用は行っていない。</li> <li>ガイドラインに定める不正行為や禁止事項等は、AI活用と人的リソースによる確認を実施。</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>■Google 検索                             <ul style="list-style-type: none"> <li>検索インデックスに登録されている膨大な数のウェブページを分類し、最も関連性の高い有用なものを瞬時に検索して、探している情報を見つけやすい形でユーザーに表示するランキングシステムは1つのアルゴリズムではなく、一連のアルゴリズムで構成。</li> <li>検索アルゴリズムが関連性や品質の高い基準を満たせるように、ライブテストや、世界各国にいる熟練した外部の検索品質評価者による評価など、厳正なプロセスを整備。品質評価者は、Google が規定する検索アルゴリズムに関する厳しいガイドラインに準拠。このガイドラインは一般公開されており、誰でも見ることができる。</li> </ul> </li> </ul>	<p>現状、AIを活用したコンテンツモデレーションは導入していない。</p>	<p>回答なし</p>	<ul style="list-style-type: none"> <li>グローバルチームは日本語を含む全サポート言語において、年中無休でTwitterルールの強制的対応を管理。</li> <li>Twitterルールを客観的に、そして一貫性を持って適用することを目指している。</li> <li>Twitterルールに違反すると判断されるコンテンツには強制的な措置が取られる。</li> <li>適切なプロセスを提供し、公正で偏りがなく、バランスが取れ、人権を尊重する方法でTwitterルールが適用されるよう注力。その基盤となるのが、コンテンツモデレーションでの透明性と説明責任に関するサンタクララ原則および複数関係者がかかわるその他のプロセスの精神。</li> <li>今後も、Twitterルールの適用方法に関して将来の報告書で提供できる情報範囲の拡大にリソースを投じていく。</li> </ul>
4- (4) ②透明性レポート 日本語で閲覧可能か	<ul style="list-style-type: none"> <li>2020年度の透明性レポートにおいては、偽情報等に関する取組状況については記載していない。</li> <li>透明性レポートにおける記載事項については、ユーザーや外部有識者からご意見を頂戴し、継続的な改善を検討。</li> </ul>	<ul style="list-style-type: none"> <li>2010年に初めて透明性レポートを公開。</li> <li>透明性レポートは日本語でご覧いただける。</li> </ul>	<p>LINE Transparency Reportの「違反投稿への対応」にて、偽情報等への対応を含めた、弊社における全世界でのコンテンツモニタリングの仕組みと対応実績を日本語で公開。</p>	<p>コミュニティ規定施行レポートを日本語に翻訳し公表。</p>	<p>日本語の透明性レポートを公表。</p>
4- (4) ③取組の効果分析	<ul style="list-style-type: none"> <li>■Yahoo!ニュース                             <ul style="list-style-type: none"> <li>誤った情報の流布に対して、ニューストピックスを中心に正しい情報を迅速かつ丁寧にわかりやすく伝えることを実践。</li> <li>より多くの人に打消しの情報を提供することで誤情報流布の抑止につなげることを企図。</li> <li>偽情報など単一 이슈に絞っての定量的な効果測定は行っていないが、長時間掲載によるimpsやCT数などに着目しながら、一定の抑止になるよう心掛け運営を行っている。</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>「YouTube コミュニティガイドラインの適用について」でポリシーに違反した動画やチャンネル、コメントのグローバル全体の削除、再審査請求とそれによって元に戻された動画、人間と機械による報告に関するデータを提供。</li> <li>「YouTube コミュニティガイドラインの適用について」の新しいデータポイント、Violative View Rate (VVR、違反コンテンツ視聴の割合)の公開を開。この指標は、四半期ごとに報告され、YouTubeでの視聴のうち、ポリシーに違反するコンテンツの視聴の割合を示す。</li> <li>2021年第4四半期の違反視聴率は0.12~0.14%の範囲にあり、YouTubeの1万再生回数のうち、違反コンテンツからの視聴は12~14%。</li> </ul>	<p>現状、偽情報等に該当して非表示(削除)した案件そのものがほぼなく、効果について分析を行っていない。</p>	<p>回答なし</p>	<ul style="list-style-type: none"> <li>世界的な大規模模倣の際と同様に、専門チームは潜在的なリスクについて積極的に監視。特に被害に遭いやすいのはジャーナリスト(特に女性ジャーナリスト)、活動家、政府機関、政府関係者などの知名度の高いアカウント。標的を定めたアカウントの乗っ取りや操作による被害を軽減する取り組みを進めており、英語、ウクライナ語、ロシア語でアカウントのセキュリティに関するリソースを公開。また、Trust and Safety Councilのメンバーを含む世界中のパートナーと協力して、支援を必要としているジャーナリストや活動家などからの緊急の支援要請を特定し、段階的に拡大して対応。</li> <li>昨年、プラットフォーム上の誤情報対策として、誤解を招く可能性のあるツイートにほかの利用者が意見や情報ソースを追加できる新機能「Birdwatch」を米国で試験的にリリース。日本では現時点で未導入。</li> </ul>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト(YouTube)</li> <li>検索エンジン(Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
5-① 広告表示先の制限	<p>「①社会的混乱関係(災害等)」 「②健康・医療関係(コロナ関係等)」 「③選挙・政治関係」</p> <ul style="list-style-type: none"> <li>Yahoo! JAPAN広告配信ガイドラインにおいて、広告の配信先は、メディアとしての知名度・実績があることと、広告配信先サイト内の情報に信頼性があることを条件としており、偽情報等のあるようなサイトへの配信はお断りしている。</li> </ul>	<p>Google の広告プラットフォームにおける偽情報への対応に関するポリシーでは、政治や時事問題に関する内容の正確性の判断は避け、広告主やパブリッシャーによる誤った表現や有害な行動に焦点を当てるアプローチを採用しています。ここでは、コンテンツを「偽情報」として特別に分類していませんが、Google のプラットフォーム上での詐欺的または低品質のコンテンツを防止することを目的に、長年にわたり多数のコンテンツポリシーを設けています。これらのポリシーは、冒頭で記載した対策を補完し、さらに発展させたものです。これらのポリシーは、欺瞞的なコンテンツに対応するための行動的アプローチを反映しています。具体的な主張の真意を判断するのではなく、誤った表現や有害なコンテンツに関連する広告主やパブリッシャーの行為に対してポリシーを適用しています。</p>	<p>LINEメディア外の3rd party mediaへの広告配信においては、以下の利用規約にてメディアの審査基準を定めており、基準に満たない場合は登録拒否及び削除を行う。</p> <ul style="list-style-type: none"> <li>LINE 広告ネットワーク利用規約</li> <li>LINE広告ネットワーク利用ガイド</li> </ul>	<p>コンテンツ収益化ポリシーは、第三者ファクトチェック団体により虚偽と判断されたコンテンツは収益化の対象として認めていない。</p>	<ul style="list-style-type: none"> <li>COVID-19(またはその派生語)に関連する不快な表現は禁止。</li> <li>センセーショナルまたはパニックを引き起こしかねないコンテンツは禁止。</li> <li>COVID-19に関連した商品の値上げを誇張して伝えることは禁止。</li> <li>COVID-19関連のニュースは、政治に関する広告コンテンツのポリシーが適用除外されているメディアパブリッシャーからのみ許可。</li> </ul>
5-② 広告出稿制限	<p>Yahoo! JAPAN 広告掲載基準において規定。</p> <p>「①社会的混乱関係(災害等)」 事実と異なる内容の広告は掲載をお断りしている。</p> <p>「②健康・医療関係(コロナ関係等)」 薬機法等の広告規制に基づき、虚偽広告は掲載をお断りしている。</p> <p>「③選挙・政治関係」 選挙および政治に関する広告は、選挙管理委員会や政党、政治団体による広告のみ掲載を認めている。その他の団体や個人による政治に関する広告は、虚偽か否かに関わらず掲載をお断りしている。</p>	<ul style="list-style-type: none"> <li>選挙運動の一環であるサイトを宣伝する広告を日本では表示しないようにしている。</li> <li>公衆衛生上の緊急事態など、社会的、文化的、政治的に重大な影響を及ぼすセンシティブな事象から利益を得る可能性のある広告、それを営利目的に利用しようとする広告は許可していない。</li> </ul>	<ul style="list-style-type: none"> <li>虚偽やフェイク広告については禁止事項としてガイドラインに掲載。</li> <li>違反が見られた場合は利用規約にて、広告の一部、あるいはサービスの利用停止 (LINE広告利用規約第13条10号「虚偽又は捏造した情報を含む情報を配信したとき」) を掲げている。</li> </ul>	<p>広告ポリシーにおいて、第三者ファクトチェック団体によって虚偽であると証明された主張、また場合によっては特定の専門性を持つ組織によって虚偽であると証明された主張を含む広告を禁止。虚偽と判断された情報を繰り返し投稿する広告主はプラットフォームでの広告掲載に制限がかけられる可能性がある。</p>	<ul style="list-style-type: none"> <li>政治に関するコンテンツを広告に掲載することを全世界で禁止。</li> <li>不適切なコンテンツを含む広告を全世界で禁止。</li> </ul>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索 (ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
5-③ ターゲティング技術の適用に関する規定	<ul style="list-style-type: none"> <li>偽情報等に関する広告は掲載を認めていないので、ターゲティング技術の適用に関する規定は設けていない。</li> <li>ただし、広告掲載基準において、政党による政治広告の場合にはターゲティングの設定を制限し、いわゆる行動ターゲティング広告を禁止している。これは特定の政治的意図をもって行うターゲティングを禁止する趣旨で2022年1月より適用している。</li> </ul>	<ul style="list-style-type: none"> <li>選挙運動の一環であるサイトを宣伝する広告を、日本では表示することを許可していない。また、広告主が政治的所属に基づくターゲティング広告を行うことも許可していない。</li> <li>ポリシーに違反する広告は許可されず、ひいてはターゲットにすることもできない。</li> </ul>	特に規定はない。	<p>広告ポリシーにおいて以下の定めを置いている。</p> <p>7. ターゲット設定</p> <ol style="list-style-type: none"> <li>利用者を差別、侮辱、挑発、攻撃する目的で、または略奪的な広告活動を行う目的でターゲット設定オプションを使用してはいけません。</li> <li>カスタムオーディエンスを広告のターゲットにする場合は、オーディエンス作成時に利用規約に準拠する必要があります。</li> </ol>	<ul style="list-style-type: none"> <li>広告主が自社コンテンツに興味関心を持つお客様にリーチしやすいよう幅広いターゲティングオプションが用意。</li> <li>ユーザーのプライバシーを尊重しているため、ユーザーはサポートされている端末で、Twitterの興味関心ターゲティングによる広告サービスをオプトアウト可能。</li> </ul>
5-⑤ 出稿者の情報や資金源の公開	<ul style="list-style-type: none"> <li>広告出稿者の社名・団体名・個人名等および所在地、連絡先を広告サイト内に明示することをYahoo! JAPAN広告掲載基準に規定している。</li> <li>特に政治広告の資金源の公開については、現状求めているが、世情に合わせ今後必要があれば検討していきたい。</li> </ul>	<ul style="list-style-type: none"> <li>日本では一般的な政治広告を許可していますが、日本では選挙広告を許可していない。</li> <li>2020年、すべての広告主から身元や事業内容に関するより多くの情報を取得するため、汎用的な広告主向けの適格性確認プログラムを導入し、特定の広告の背後にある広告主や広告主の所在地を「広告の表示について」(場所によっては「この広告について」)メニューで開示。本機能は、2021年5月より日本国内でも利用可能。</li> </ul>	特に規定を設けていないが、広告主の運営者情報をLP内に記載、もしくは運営者情報のサイトに遷移できなければならないという審査基準がある。	<ul style="list-style-type: none"> <li>広告ライブラリでは掲載されている広告を検索することができ、表示された広告に関する情報を確認可能。</li> <li>社会問題、選挙または政治に関連する広告の場合、アクティブではない広告も検索可能。</li> <li>選挙への介入を防ぐためには透明性が重要であるため、広告ライブラリには広告費用の出資者、広告費の金額範囲、広告がリーチした利用者層など広告についての追加情報も表示。これらの広告は、ライブラリに7年間収録。</li> </ul>	連邦選挙や総選挙に関する広告は、現在は禁止。
5-⑦ その他の透明性・アカウントビリティ確保方策	<p>広告審査および広告配信先の審査に関する透明性レポートを定期的に公開。実際にお断りした件数や内容の内訳をまとめている。</p>	<p>誰もがオンラインで見る広告を理解できるよう、選挙関連の広告主の認証済みの支出に関する情報を含む透明性レポートを公開。</p>	<p>上記に記載した審査でのチェックやユーザからの通報対応以外は、特にツールの提供などは行っていない。</p>	<p>Transparency Centerにおいて、誤情報に対するアプローチを詳しく説明。</p>	<ul style="list-style-type: none"> <li>広告パートナーと連携することで、興味関心に基づいた広告配信を行い、ユーザーの役に立ち興味を引くと思われる広告を表示。広告パートナーから提供される情報を利用して広告をカスタマイズすることがある。</li> <li>Twitter内外での興味関心に基づいた広告表示を希望しない場合は、この機能をオフにできる。</li> <li>Twitter広告に関連する報告のためのフォームを用意。</li> </ul>

## プラットフォーム事業者における回答状況（6関係）

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
6-① ファクトチェック結果の表示 具体的な仕組み・基準	<p>■UGC機能を有するサービス</p> <ul style="list-style-type: none"> <li>投稿内容もさまざまであって、真偽の確認も困難であることが多いため、ファクトチェック機能は導入していない。</li> <li>ユーザーや企業による申告をきっかけに事実と反する投稿が確認できた場合、あるいは、パトロールにより検知された偽情報(公式に示されている情報に明らかに反するもの)については、既存のガイドラインに抵触する限りにおいて対応を行う。</li> </ul> <p>■Yahoo!ニュース</p> <ul style="list-style-type: none"> <li>信頼性の高いメディアと契約を締結し、コンテンツを調達・掲載。また、Yahoo!ニューストピックにおいては、誤情報・デマに対して正しい情報を積極的に掲載。</li> <li>「新型コロナ特設サイト」では、コロナ関連のデマへの注意喚起を行うコーナーを設置。</li> <li>FIJと情報共有や相互のリンク設置などの連携を行ったほか、メディアや専門家のフェイクニュース関連記事を掲載。あわせて消費者庁や国民生活センターへのリンクも設置。各所への掲載にあたっては編集部が複眼的に確認をした上で掲載を行っている。</li> <li>ロシアのウクライナ侵攻に関しては、最新のニュースや専門家の解説等を集約した特設サイトを作成。同サイトにおいては、「ウソやデマに注意を」と呼びかけるとともに、FIJのウクライナ関連のファクトチェック特設サイトへのリンクを設置。</li> </ul>	<ul style="list-style-type: none"> <li>Google のファクトチェックにおいては、ユーザーに選択権を与えるアプローチを採用。ファクトチェックのラベルは Google 検索及び Google ニュース結果に表示され、ユーザーがコンテンツのどの部分がポジティブにファクトとしてチェックされているかが確認でき、より詳細を調べたい場合も簡単により情報を探ることができる。</li> <li>Googleニュースにはファクトチェックボックスがあり、Google ニュースの新型コロナウイルス感染症用ニュースエキスプレックスには専用のファクトチェック表示がある。</li> <li>これらの機能は、第三者パブリッシャーのオープンなネットワークに依存 (IFCN を含む特定の1つのファクトチェックネットワークのみに依存するものでもない)。</li> <li>世界中の事実確認の専門家やネットワークとの協力し、Poynter やFirst Draft など、各国のファクトチェック団体の連合組織である International Fact-Checking Network (IFCN) のような組織を支援。</li> </ul> <p>■YouTube</p> <p>ファクトチェックを学ぶ動画を公開するなど、誤った情報や意図的に操作された情報が拡散する「ミスインフォメーション(誤った情報)」に対して、誤った情報を扱っている動画の特定と削除や、ガイドラインに違反するかどうかの際どいボーダーライン上のコンテンツや有害な誤情報がユーザーにおすすめされることを減らす取組を実施。</p>	<p>■LINE NEWS</p> <ul style="list-style-type: none"> <li>万が一誤報などがあった場合の記事の訂正、削除などについては原則提供元の対応に沿い、記事内容の反映や削除などを実施。なお、特集枠やダイジェスト配信などへの掲出があった場合、LINE NEWS側でも別途、訂正情報の掲出等を行うケースも想定。</li> <li>記事の編集権を媒体側に開放している「オープンキャスト」枠については、編集部側でガイドライン違反がないか確認。なお、編集権限が媒体側にあるため、謝罪・訂正等の掲出を含めて媒体側が判断し実施することとなるが、内容によってはLINE NEWS側においても対応。</li> <li>特定のニュースに対し、誤報であると報道や注意喚起がされている場合には、その旨を可能な範囲で目立つ場所に掲載し、誤報による影響を最小にできるよう編成。</li> <li>ロシア・ウクライナ関連のフェイクニュースや偽情報への注意喚起やファクトチェック結果(ファクトチェック・ナビ)を掲載。</li> </ul> <p>○LINE Taiwanの取組</p> <ul style="list-style-type: none"> <li>LINE TaiwanのLINEファクトチェックサービスは、4つのファクトチェック機関と連携。このうち2社(The Taiwan Fact Check Center、MyGoPen)は国際ファクトチェックネットワーク(IFCN)認証を取得。全てのファクトチェック記事は、包括的な指針と原則(例:The Taiwan Fact Check Centerのファクトチェック作業ガイドンス)を備えたファクトチェック専門家を通じて発行。また、台湾行政院報道室と連携し、最新の政策を反映した政府公式見解のニュースを発表。</li> <li>LINEファクトチェッカーにおいて、個人のプライバシー保護とセキュリティ基準のモニタリングに関する内部レビューガイドラインがある。ファクトチェックパートナーがバックエンドシステムにアップロードした後、ガイドラインを遵守するためAIフィルタリングとLMPシステムを使用。</li> <li>LINE TODAYの編集者は、LINEファクトチェッカーにおいてファクトチェックされた記事を同時にストリーミングしながら、1,800万人のMAUユーザーに対してLINE TODAYで公開することが正しいかどうかを再度レビュー。</li> </ul>	<p>○誤情報を特定</p> <p>利用者からのフィードバックなどに基づいて誤情報の可能性のあるコンテンツを特定し、ファクトチェック団体に表示。また、ファクトチェック団体がコンテンツを特定したうえで審査する場合もある。</p> <p>○コンテンツを審査</p> <p>ファクトチェック団体がコンテンツを審査し、事実を検証したうえでその正確性を評価。これは独立して行われ、発信元や公開データを参照したり、動画と画像を認証したりなどの作業が含まれる。</p> <p>○誤情報を明確にラベル付けして利用者へ周知</p> <p>ファクトチェックパートナーに審査されたコンテンツにラベル付けを行い、利用者が追加の背景情報を閲覧できるようにする。また、ラベル付けされたコンテンツをシェアしようとする利用者や過去にシェアした利用者に通知。</p> <p>○誤情報を目にする利用者の数を抑制</p> <p>ファクトチェック団体が「虚偽」、「改変」、「一部虚偽」と評価したコンテンツは、ニュースフィードでの表示順位が下がり、Instagramでの発見タブの表示対象から除外され、フィードやストーリーズで目立たないようにする。これにより、誤情報を見る利用者の数を大幅に抑制可能。また、ファクトチェック団体によって評価されたコンテンツの広告は却下される。</p> <p>○誤情報を繰り返し配信する違反者への措置</p> <p>「虚偽」または「改変」と評価された誤情報を繰り返し配信するページやウェブサイトに対して、配信数の抑制などの制限措置を実施。また、それらのページやウェブサイトは収益化や広告に関する機能へのアクセスやニュースページとしての登録を一定期間取消。</p>	<ul style="list-style-type: none"> <li>信頼できるパートナー(APおよびロイターとのパートナーシップ、その他の報道機関、公共保健機関、政府など)と密接に連携してポリシーを適用。さまざまなパートナーとオープンな対話を行い、助言を求めたり、さまざまなメディアや主張を検証。</li> <li>オーストラリア、韓国、米国において、利用者が誤解を招くと思われるツイートを報告できる新しい報告機能のテストを開始。この実験/検証の一環として、[問題を報告する]を選択すると、「It's misleading(誤解を招いています)」という文言がオプションに表示されるようになる。</li> </ul>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
6-② ファクトチェックを容易にする ツールの開発及び提供	<p>■Yahoo!ニュース ツールの開発および提供はしていない。</p>	<ul style="list-style-type: none"> <li>ファクトチェックエクスプローラー及びファクトチェックマークアップの2つのツールで構成されるファクトチェックツールを提供。</li> <li>イメージの由来、偽情報対策のクリティカルなユースケースを確認するオープンソースでモバイル向けに最適化されたStoryfulによるSourceの発展をサポート。</li> <li>画像検索は改変された画像のオリジナルを見つけるためにファクトチェッカーに活用されている他、Google マップやGoogle Earthなどの地形情報も情報検証に欠かせないツールとして利用されている。それらのツールのファクトチェックへの活用方法を解説するセミナーや動画なども実施。</li> </ul>	<p>■LINE NEWS ファクトチェックツールの開発及び提供は行っていない。</p> <p>○LINE Taiwanの取組 LINEにおける偽情報の拡散を最小限に抑え、メディアリテラシーとフェイクニュースを区別するためLINEファクトチェッカーを開発。LINEファクトチェッカーでは、ユーザが疑わしい偽情報を転送することで、公式アカウントで直ちに真偽を判断することが可能。また、ユーザがすべての偽情報と正しい情報を閲覧できるようにウェブサイト(ダッシュボード)を構築。</p>	<ul style="list-style-type: none"> <li>利用者の反応やコンテンツの拡散速度など、さまざまなシグナルに基づいて、誤解を招く可能性のある投稿を検出。</li> <li>利用者は、ファクトチェック団体が投稿を詳しく見るために、コンテンツの一部にフラグを立てることが可能。</li> <li>その他、虚偽の情報を特定するためのシグナルとし投稿に対する不信感を示すコメントなどがある。</li> <li>IFCNの認証を受けているファクトチェック団体はCrowdTangle(自社ツール)にもアクセス可能。出版社、ジャーナリスト、研究者、学術関係者もCrowdTangleを使用して、ソーシャルメディア上の公開コンテンツを追跡、分析、報告可能。</li> </ul>	<ul style="list-style-type: none"> <li>昨年、プラットフォーム上の誤情報対策として、誤解を招く可能性のあるツイートにほかの利用者が意見や情報ソースを追加できる新機能「Birdwatch」を米国で試験的にリリース。日本では現時点で未導入。</li> </ul>
6-③ ファクトチェックを実施する人材の育成	<p>■Yahoo!ニュース ファクトチェックのための人員配置および育成は行っていないが、メディア事業における編集部が運用のなかでその役割を担っている。</p>	<p>回答なし</p>	<p>○LINE Taiwanの取組</p> <ul style="list-style-type: none"> <li>ファクトチェックは連携している外部機関により実施されるため、ファクトチェック自体を行う人材の育成は行っていないが、ソフトウェアを開発。また、Junyi Academy Foundation (非営利組織)等と協力して小・中・高校生がファクトチェックを自ら行えるように、ビデオ、ラーニングシート等のメディアリテラシー教材を開発。</li> <li>2021年に行われたLINE Taiwanの偽情報に係る調査によると、約41%が事実確認に積極的に乗り出すことが判明(2019年は約37%)。また、LINEファクトチェッカーなどのファクトチェックツールの2021年の利用率は、2019年に比べて2倍。</li> <li>台湾最大のゲーム教育プラットフォームであるPaGamOと協力し、学生に最新の偽情報を教え、各課題を通じて批判的思考能力を訓練するクイズを提供。2020年から2年間でLINE公式クイズミッション「バガモ」に参加した学生は計38万人を超える。</li> </ul>	<p>回答なし</p>	<p>3月3日(米国時間)より、米国のTwitter利用者の小規模(かつ無作為に抽出された)グループが、一部のツイートについてのBirdwatchノートを直接閲覧可能。これらのユーザはノートを評価することもでき、Birdwatchがさまざまな立場の人に役立つ文脈を追加する能力を向上させるための情報を提供。</p>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
6-④ 日本におけるファクトチェック推進に向けた取組	<p>■Yahoo!ニュース</p> <ul style="list-style-type: none"> <li>・ Disinformation対策フォーラムの議論に参画し、産官学民の連携を実践。</li> <li>・ FIJの活動に賛同し、寄付の実施や法人会員として支援連携。</li> </ul>	<ul style="list-style-type: none"> <li>・ FIJ、InFact及びBuzzFeed Japan に対し、ファクトチェックに関するワークショップを行い、Google のファクトチェックツールを使用するサポートを実施。BuzzFeed Japanは現在、ファクトチェックツールに登録され、ファクトチェックコンテンツがツール内で検索可能。</li> <li>・ 2021年に主に学生を対象にしたファクトチェック・チャレンジや記者も対象に含めたファクトチェック・ワークショップなど年間を通じたセミナーを実施。この取組は2022年も対象を広げて実施する予定。</li> <li>・ 毎年秋にTrusted Media Summit(信頼されるメディアサミット)を開催し、ミスインフォメーションの問題にどう取り組むかをアジア太平洋地域のファクトチェッカーやメディアリテラシーの専門家らを招いて実施。2020年、2021年はオンライン開催だが、日本からも多数の記者や有識者が参加。</li> </ul> <p>■You Tube</p> <p>ファクトチェックに関するクリエイターワークショップを2019年11月21日に行い、20のクリエイターが参加。Poynter Institutionとのパートナーシップにより開催。</p> <p>○Google News Lab</p> <p>日本のNews Labフェローの協力のもと、ジャーナリストや大学生を対象に、偽情報や誤った情報の識別、検索の有効活用、地図や衛星画像の活用など、ジャーナリズムにおけるデジタル技術の活用方法に関する研修を実施。2015年以降、Google News Lab が提供するトレーニングには、バーチャルおよび直接参加で22,000人が参加。</p>	回答なし	回答なし	<ul style="list-style-type: none"> <li>・ 一般社団法人セーフアインターネット協会のDisinformation対策フォーラムに加盟。</li> <li>・ 今後Birdwatch機能も日本で導入する予定。</li> </ul>
6-⑤ ファクトチェック機関との連携	上記のほか、連携相談や活動内容の共有などの定期的なMTGを実施している。	<p>無党派組織で米国、ドイツ、ブラジル、アルゼンチン、南アフリカ、インドその他のファクトチェック組織を束ねる Poynter の国際的ファクトチェックネットワーク (IFCN)と提携。</p> <p>【日本】</p> <p>2021年12月、APAC Trusted Media Summit 2021の一環として、日本で誤報会議「Digit Journalism Dojo Summit」を開催し、日本全国、地域、地元の主要メディアから123名が参加。</p>	<p>■LINE NEWS</p> <ul style="list-style-type: none"> <li>・ FIJに正会員として参加しており、活動を支援。</li> <li>・ ディスインフォメーション対策として、Zホールディングス株式会社、ヤフー株式会社と定期的に情報交換を行っており、プラットフォームとして、どのような対策ができるかを議論。</li> </ul> <p>○LINE Taiwanの取組</p> <p>偽情報等の疑わしい情報源を確認するため、以下4つの独立した第三者のファクトチェック機関と協力。</p> <ul style="list-style-type: none"> <li>・ The Taiwan Fact Check Center</li> <li>・ Rumor &amp; Truth</li> <li>・ MyGoPen</li> <li>・ Cofacts</li> </ul>	多くの国と地域で、この種のコンテンツの識別、審査および措置の遂行のために、IFCNに認証された独立したサードパーティーファクトチェック団体と提携。	ルールの執行だけでなく、専門家と協力し、利用者が信頼できる医療情報を取得できるようにすることを重視。そのため、地方、国、世界の公共医療機関と世界中で緊密に協議しながら、取組を実施。

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト(YouTube)</li> <li>・検索エンジン(Google 検索)</li> <li>・デジタル広告プラットフォーム(Google 広告)</li> <li>・オンラインマップ(Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス(LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス(LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
7ー① 普及啓発の取組・投資	<p>■Yahoo!ニュース</p> <ul style="list-style-type: none"> <li>・ 21年6月より大学と連携し、フェイクニュース対策としてのリテラシー向上授業を実践。これまでに中学高校大学やNPO団体イベントなどで開催し、今後も継続。</li> <li>・ 啓蒙啓発コンテンツを制作し、ニュースや検索をはじめメディア面を中心に誘導を強化。</li> </ul>	<ul style="list-style-type: none"> <li>・ 「Google News Initiative」を筆頭に、数多くのメディアリテラシープログラムを立ち上げ、支援。</li> <li>・ Google.org (Google の慈善事業部門) は、イギリス、ブラジル、カナダ、インドネシアなどですでに支援してきたプログラムを拡充する形で、世界中のメディアリテラシーを支援するための1,000万ドルのグローバルイニシアチブを2018年に立ち上げ。</li> </ul> <p>【日本】</p> <ul style="list-style-type: none"> <li>・ 2021年、学生にファクトチェックと検証のスキルを学んでもらうプログラム「ファクトチェック・チャレンジ」を実施。</li> <li>・ 中学生・高校生にオンライン・リテラシー・カリキュラムを提供。</li> <li>・ Grow with Google という無料のデジタルスキルトレーニングを提供するプログラムを通じて、ユーザーのインターネットリテラシー教育を強化するために、無料のオンライン講座を提供。</li> <li>・ ウェブレンジャーと呼ばれる、10代の若者にオンラインセーフティのためのプロジェクトを奨励するメディアリテラシープログラムを2015年、2016年、2018年に実施。</li> <li>・ お子様向けインターネットリテラシープログラム Be Internet Awesome (日本語版) を公開。</li> </ul>	<ul style="list-style-type: none"> <li>・ 「インターネット上でのコミュニケーション」と「対面のコミュニケーション」の違いを子どもたちに学んでもらうためのオリジナル教材を開発するとともに、全国の学校や地方公共団体からの申し込みを受け付けてLINEが講師を派遣するワークショップ授業・講演活動等を2012年より全国で開始し、これまでに累計で約10,000回以上実施。2020年1月からはコロナ禍の影響によりオンライン講演への対応へ変更し、2021年は788講演(情報モラル教育と情報防災教育の合計)を実施。</li> <li>・ オンラインではLINE Safety Center-LINEの安心安全ガイド(<a href="https://linecorp.com/ja/safety/index">https://linecorp.com/ja/safety/index</a>)を用意し、禁止行為などのガイドを確認できるようにしている。</li> <li>・ 2019年12月に設立したLINEみらい財団では、教育工学や授業デザインを専門とする研究者と共同で、独自の情報モラル教育教材の開発を行い、ウェブサイトで公開。</li> </ul>	<ul style="list-style-type: none"> <li>・ アジア太平洋地域の専門家と協力して、FacebookやInstagramといったオンライン上での嫌がらせやいじめなどに対処し、責任あるデジタル市民によるグローバルコミュニティを構築するためのリソースを提供するオンライン出張プログラム「みんなのデジタル教室」を立ち上げ。日本では特定非営利活動法人企業教育研究会の協力のもと、デジタルリテラシーに関する出張授業を国内の中学校・高等学校などで実施。</li> <li>・ 2021年12月末までに13,000名以上の学生が授業を受講しており、90%以上が「授業を受けて、インターネットやアプリ、SNSへの関心が高まった」と回答。</li> <li>・ 新型コロナウイルス感染症に関する誤情報から身を守るために、「新型コロナウイルス感染症に関する誤情報に対処するための6つのヒント」を展開。</li> </ul>	<ul style="list-style-type: none"> <li>・ 10月24日から@UNESCOの「Global Media and Information Literacy Week 2019(グローバルメディア情報リテラシーウィーク2019)」が始まるにあたって、教育者向けの新しいハンドブック「Twitterを活用した教育と学習」の提供を開始。</li> <li>・ UNESCO関連のNGOネットワークや各国の政府教育機関、欧州ネットワークセンター、日本では国連広報センター(UNIC)やアジア太平洋メディア情報リテラシー教育センター(AMILEC)など、世界中にある数多くのパートナーと連携して、学校教材としても配布。</li> <li>・ #世界メディア情報リテラシーウィーク、#シェアする前に考えよう、#クリックする前に考えようの3つのハッシュタグで期間限定の絵文字を作成。</li> <li>・ 全米メディアリテラシー教育協会(National Association for Media Literacy Education、NAMLE)とパートナーシップを組み、ハッシュタグ#MediaLitWkを使って米国におけるメディアリテラシーウィークを支援。</li> <li>・ @UNESCOと関連する他のNGOに無償広告支援プログラム「Ads For Good」を提供して、メディアリテラシーウィークの認知度を向上。</li> </ul>
7ー② 他のステークホルダーとの連携・協力・投資	<p>■Yahoo!ニュース</p> <ul style="list-style-type: none"> <li>・ 「Disinformation対策フォーラム」へ参加し、リテラシー向上について多様なプレイヤーと議論。</li> <li>・ 総務省ICTリテラシー向上施策研究会において、ヤフーの実践内容の共有と、今後のリテラシー教育についての提言を行った。</li> </ul>	<p>○研究活動</p> <p>国際大学GLOCOMとの共同研究であるInnovation Nipponで、プラットフォームや情報リテラシーに関する様々な研究を実施。</p> <p>○トレーニング</p> <ul style="list-style-type: none"> <li>・ FIJ、InFact、BuzzFeed Japanに対して、ClaimReviewの使い方のワークショップを行い、Google ファクトチェックツールの利用を支援。</li> <li>・ Google News Labは、ジャーナリストや大学生を対象に、デジタル技術をジャーナリズムに活用するためのトレーニングを実施。</li> </ul>	<ul style="list-style-type: none"> <li>・ 情報モラル・情報リテラシーの啓発活動の強化やネットトラブル防止を目的に、地方公共団体や専門家と協力しながら、調査研究・教材作成等実施。</li> <li>・ GIGA スクール展開にあわせた児童・生徒の情報モラル・リテラシーの強化と向上を目的とし、地方公共団体・研究者と共同で開発し、同教材を公開。</li> <li>・ 2021年6月に、神奈川県警察本部および神奈川県教育委員会と、情報モラル教育教材「SNSの上手な使い方を考えよう！」をリリース。</li> <li>・ 2022年2月に、災害時の適切な情報発信の方法について学ぶ情報リテラシー×防災の教材「情報防災訓練」の第2弾「情報発信編」を開発、公表。</li> </ul>	<p>「みんなのデジタル教室」では、日本では特定非営利活動法人企業教育研究会の協力のもと、デジタルリテラシーに関するオンライン出張授業を提供。</p>	<p>様々なステークホルダーと連携。</p>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
8-① AI技術に関する研究開発	<p>■Yahoo!ニュース 偽情報に関する機会学習の開発研究は行っていないが、コメント欄などUGC機能における品質観点でのAI開発は引き続き実施。</p>	<ul style="list-style-type: none"> <li>・ 有害なコンテンツを積極的に検出してフラグ付けをする、そして以前に削除されたコンテンツと酷似しているコンテンツは自動的に削除するという、2つの柱となる方法で機械学習を導入。</li> <li>・ 「合成メディア」(「ディープフェイク」と呼ばれることも多い)の検出にAIがどのように役立つかを理解するための研究に投資し、世界中のこの分野における主要な専門家と協力。</li> </ul>	<p>NGセンテンス検出技術の研究開発を実施。</p>	<p>AI研究チームと協力し、学者から学び、サードパーティーのファクトチェック団体とのパートナーシップを拡大し、他の組織(他のプラットフォームを含む)と協力する方法について話し合っている。</p>	<p>回答なし</p>
8-② 「ディープフェイク」対策の研究開発	<p>■Yahoo!ニュース 行っていないが、ヤフー上に掲載される映像コンテンツは契約しているメディア、個人からのものに限っている。</p>	<ul style="list-style-type: none"> <li>・ 2019年、高性能なフェイクオーディオ検出器を開発するための国際的な取組を支援するため、合成された音声に関するデータセットの公開を発表。このデータセットは、取組の一環として150以上の研究機関や業界団体によってダウンロードされ、現在は自由に公開。</li> <li>・ Jigsawと共同で、Googleが作成したビジュアルディープフェイクの大規模なデータセットを発表。このデータセットは、ミュンヘン工科大学とナポリのフェデリコ2世大学が新たに開発したFace Forensics benchmarkに組み込まれた。</li> </ul>	<p>現状は推進していない。</p>	<ul style="list-style-type: none"> <li>・ ディープフェイクは検出が最も難しい。</li> <li>・ 2019年には「Deep Fake Detection Challenge」を立ち上げ、世界中の人々がディープフェイクを検出するためのより多くの研究やオープンソースのツールを生み出すことに拍車をかけている。1000万米ドルの助成金で支援されたこのプロジェクトには、Partnership on AI、コーネル大学、カリフォルニア大学バークレー校、MIT、WITNESS、マイクロソフト、BBC、AWSなど、市民社会やテクノロジー、メディア、学術のコミュニティに所属する数名の組織の横断的な連合が参加。</li> <li>・ 2021年、ミシガン州立大学(MSU)と共同で、ディープフェイクの検出・帰属の研究手法を新たに発表。これは、AIが生成した1枚の画像から、その画像を生成するために使用した生成モデルをリバースエンジニアリングするもの。この方法により、ディープフェイク画像そのものが検出器の唯一の情報であることが多い実世界でのディープフェイク検出とトレースが容易になった。</li> </ul>	<p>Adobe, Twitter, NYTがディープフェイクの対策の取り組みを開始。誰が画像や動画を作成し、どのような変更が加えられたかを明確にするための取組を実施。</p>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
9-① 信頼性の高い情報の表示	<p>■Yahoo!知恵袋</p> <p>新型コロナウイルス関連の投稿がなされた際に、当該投稿のページ上部に「新型コロナウイルスに関する内容の可能性がある記事です。」等の文言を掲出し、厚生労働省等の公的機関のHPを案内する取組を実施。</p> <p>■Yahoo!ニュース</p> <ul style="list-style-type: none"> <li>トピックスについて</li> <li>ユーザーの課題解決に資するため、公共性の高い情報をサービスの最も目立つ場所であるYahoo!ニューストピックスやトピックス詳細ページに掲載。</li> <li>オーサーコメントについて</li> <li>特定分野の専門性を有するオーサーと契約を締結し、専門分野にかかる記事について、補足的な見解を示して、記事の内容についてのユーザーの理解を向上させることを狙いとす。</li> <li>なお、専門家の選定にあたっては、特定分野における専門性、評判、知名度などを基準にした審査を行っており、また専門家による投稿は全件、担当者によるチェックを行うことで、一定の品質を保っている。</li> </ul> <p>■Yahoo!検索</p> <p>①社会的混乱関係(災害等)</p> <p>大規模な災害が発生した際には、関連するワードで検索した際に、各自治体や省庁などが発信している情報等をはじめ、信頼性が高い情報をまとめ、検索結果画面の上部に掲出する取組を実施。</p> <p>(次ページへ続く)</p>	<p>■Google 検索</p> <ul style="list-style-type: none"> <li>「ナレッジパネル」や、YouTubeの「情報パネル」は、人物や物事についての重要な事実を提供。</li> <li>ラベルやスニペットを使うことにより、あるコンテンツがファクトチェックを受けたものであることがユーザーにすぐに分かるようにする。</li> </ul> <p>○Google News</p> <p>「Full Coverage」機能は、ユーザーがパーソナライズされていない、掘り下げたニュース記事へのアクセスを、ワンタップで可能にしている。</p> <p>■YouTube</p> <p>情報パネルは、YouTube でニュースを視聴しているユーザーに対し、ニュースコンテンツの情報源をよりよく理解するための追加情報を提供。</p> <p>■Google広告</p> <p>「広告の表示について」ラベルは、ユーザーになぜその広告が表示されたのかを説明し、また、パーソナライズド広告を変えるために好みをどう管理すべきか、あるいは、パーソナライズド広告を停止する方法を案内。</p>	<p>■ニュース配信サービス関係</p> <ul style="list-style-type: none"> <li>公共性・公益性の高い情報をユーザの目に入りやすい位置に優先的に配置。</li> <li>LINE NEWSに新型コロナウイルス関係の最新状況まとめサイトを設置し、速報や国内の感染状況一覧、国内の感染者数推移、世界の状況等の情報を記載。</li> </ul> <p>○その他</p> <ul style="list-style-type: none"> <li>新型コロナウイルスに関する便乗商法が広告として表示されないよう、広告を監視。</li> </ul>	<ul style="list-style-type: none"> <li>メタ・ジャーナリズム・プロジェクトは、世界中の出版社と協力し、ジャーナリストとコミュニティとのつながりを強化するために活動。</li> <li>また、ニュース業界が抱えるビジネス上の中核的な課題への対処も支援。</li> <li>ニュースを通じてコミュニティを構築すること、世界中のニュースルームを訓練すること、ニュース出版社や非営利団体と提携して誤情報と戦い、ニュースリテラシーの促進、新しい取り組みへの資金提供、プラットフォームでのジャーナリズムを改善することの3点で活動。</li> <li>新型コロナウイルス感染症情報センターのほか、FacebookおよびInstagram上での通知画面の表示を通じて、20億人の利用者を専門家からの信頼できる情報につなぐ。</li> <li>米国においてはワクチン検索ツールを通じてワクチン接種を申し込めるようにし、実際に400万人が利用するなど、ワクチン接種を促すために信頼できる情報を提供。</li> </ul>	<ul style="list-style-type: none"> <li>利用者が適切な情報に出会えるよう、2020年1月以降、COVID-19専用の検索表示機能を導入。検索結果の最上位に各国の公的保健機関へのリンクを掲載するという機能「#ThereIsHelp」を実装(日本は厚生労働省のサイトへのリンク)。この機能は世界80か国以上に展開され、現在では29の言語で利用可能、</li> <li>一部の国では、これにCOVID-19のワクチンに関する情報にリンクするボタンが追加されている。</li> </ul>

	ヤフー	Google	LINE	Meta	Twitter
<p>サービス分類 (対象サービス名)</p>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>・ニュース配信サービス (Yahoo!ニュース)</li> <li>・検索サービス (Yahoo!検索(ウェブ検索))</li> <li>・広告</li> </ul>	<ul style="list-style-type: none"> <li>・動画共有サイト (YouTube)</li> <li>・検索エンジン (Google 検索)</li> <li>・デジタル広告プラットフォーム (Google 広告)</li> <li>・オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>・コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>・ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>・ソーシャルメディア (Twitter)</li> </ul>
<p>9-① 信頼性の高い情報の表示</p>	<p>(続き)</p> <p>②健康・医療関係(コロナ関係等) 医療に関する信頼性の高い情報の提供を行うため、特定のワード(例えば「新型コロナウイルス」関連の主要なワードや、各種がんの病名等)で検索した際、公的機関が発信している情報等、信頼性が高い情報をまとめて、検索結果画面の上部に掲出する取組を実施。 同様にワクチンの情報を求めるようなワード(「ワクチン」「ワクチン反対」「ワクチン危険」等)においても、厚生労働省等と連携して、検索結果画面の上部で適切な案内や誤情報への注意喚起をするような施策を実施。</p> <p>④全般・その他 「フェイクニュース」と検索した際に、概要や注意喚起を促す情報を掲出することで、不確かな情報に惑わされないような啓蒙につなげる取り組みを実施。</p> <p>■その他 Yahoo!トップページ(ニューズピックスおよびTOPタイムライン)の取り組みとして、新型コロナ等、国民の生命財産に関わる重大事項が発生している場合、Yahoo!トップページにおいて、メディアから提供を受けたコンテンツを誘導。首相会見や都知事会見等の動画のエンベッド(埋め込み)による提供や、特設サイトへの誘導など、ユーザーが認知しやすい場所に掲載。災害においては、地震速報やユーザーのお住まいの地域ごとにアラート情報を掲出。</p>				

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
9-② ニュースの選別・編集に関する透明性・アカウントバリティ確保方策	<p>■Yahoo!ニュース</p> <ul style="list-style-type: none"> <li>トピ掲載基準の公開</li> <li>トピックスに掲載する記事の基準として、「公共性」と「社会的関心」を2つの大きな柱としており、その他の掲載基準についても公開。</li> </ul>	<p>Google ニュースやYouTube で提示されるニュースコンテンツのランキングシステムには、Google 検索と同じウェブクロールとインデックス登録の技術を採用。</p>	<p>■ニュース配信サービス関係</p> <ul style="list-style-type: none"> <li>記事の選定にあたっては報道・メディア経験者を中心としたメンバーで策定した方針・ガイドラインに沿いながら、「公共性・公益性」を重視して編集者が選定を行っている。また、編集部内のチェックフローとは別に、独立した校閲の専門チームを組織しており、必要に応じて一次ソースの確認も実施。</li> <li>LINE NEWS編集部によるニュースの選定基準や、レコメンドエンジンによる記事掲出の仕組みについては、公式ブログ、公式SNS等での発信を中心にを行っている。</li> <li>提携媒体に対しては、定期的なコミュニケーションのほか、セミナーの実施などを行って方針等の共有・開示に努めている。</li> </ul>	<p>回答なし</p>	<p>モーメントでは、さまざまな内容の魅力なオリジナルコンテンツを取り上げる。モーメントに掲載するツイートや動画の選定では、内容や画像/動画に加え、投稿者のプロフィール画像やユーザー名も検討の対象となる。</p> <p>○公平性 議論を呼びそうな話題に関するツイートを選定する際には、データに基づいて決定を行い、すでにTwitterで大きな反応を呼んでいるツイートを見つける。公の議論に関するトピックの場合、可能な限り多様な視点を提示するように務める。キュレーターは、Twitterでの、ありのままの会話を客観的に映し出すことを目指し、公平性を期すためのトレーニングを継続的に受けている。</p> <p>○正確性 キュレーションチームは、ソーシャルメディアの検証方法に関して定期的にトレーニングを受けている。また、ほとんどのモーメントは、レビューのプロセスを経てから取り上げられる。</p> <p>○誤情報の訂正 万が一、取り上げた情報が誤りだったことが判明した場合には、モーメントに目に見える形で訂正を加え、その旨をツイートで発表。ただし、稀に当該モーメントを削除し、関連する国のTwitterアカウントを使用して、誤情報を撤回する旨をツイートする場合がある。</p> <p>(次ページへ続く)</p>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
9-② ニュースの選別・編集に関する透明性・アカウントビリティ確保方策					<p>(続き)</p> <p>○基準</p> <p>ニュース性の高い会話を紹介するために必要な場合を除いて、冒険、暴力、裸体など、センシティブであるとみなされるコンテンツは選定の対象から外される。また違法行為を描写したり、助長したりするようなコンテンツは採用しない。センシティブであると受け取られる可能性のあるコンテンツがモーメントに含まれる場合、警告を表示。</p> <p>○利害関係の対立の回避</p> <ul style="list-style-type: none"> <li>キュレーションチームは、収益を伸ばしたり、アカウント数を増やしたり、Twitterのパートナー関係を管理したりする責任は負わない。ツイートや動画の紹介や、モーメントのトピックの選択は、Twitterのユーザーにとっての利益を最も重視して行われる。こうした決定が、広告主、パートナー、またはTwitterの事業関係者の利益の影響を受けることはない。</li> <li>キュレーションチームは基本的に、Twitterが属する業界や自社、Twitterの競合企業を扱ったモーメントは作成しない。TwitterやTwitterの競合企業がTwitter上の会話の中で重要な要素となった場合(たとえば、当該企業の何らかの問題がトレンドトピックになるなど)、Twitterのモーメント編集チームは、当該企業からのリアクションがあればそれも含め、事実に基づく概要をまとめることがある。</li> </ul>

	ヤフー	Google	LINE	Meta	Twitter
サービス分類 (対象サービス名)	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (Yahoo!知恵袋、Yahoo!ニュース(コメント機能))</li> <li>ニュース配信サービス (Yahoo!ニュース)</li> <li>検索サービス (Yahoo!検索(ウェブ検索))</li> <li>広告</li> </ul>	<ul style="list-style-type: none"> <li>動画共有サイト (YouTube)</li> <li>検索エンジン (Google 検索)</li> <li>デジタル広告プラットフォーム (Google 広告)</li> <li>オンラインマップ (Google マップ)</li> </ul>	<ul style="list-style-type: none"> <li>コンテンツ共有サービス (LINE VOOM、LINEオープンチャット、LINE LIVE、LINE BLOG、livedoor Blog)</li> <li>ニュース配信サービス (LINE NEWS)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Facebook、Instagram)</li> </ul>	<ul style="list-style-type: none"> <li>ソーシャルメディア (Twitter)</li> </ul>
9-③ メディアとの連携体制構築 具体的検討・取組	<p>信頼できる質の高い情報を増やしていくため、伝統的メディアやネットメディア等と連携体制を構築。</p>	<ul style="list-style-type: none"> <li>Google News Initiativeをはじめ、さまざまな方法でニュースエコシステムの他のステークホルダーと協力。</li> <li>日本新聞協会向け窓口の設置。</li> </ul>	<p>■ニュース配信サービス関係</p> <ul style="list-style-type: none"> <li>アライアンスを専門に行う組織を設置しており、日頃から各提携媒体様と密に連携。</li> <li>特に影響の大きなメディアとは定期的に情報交換をしており、情報の信頼性に関する話題を取り上げることもある。</li> </ul>	<p>回答なし</p>	<ul style="list-style-type: none"> <li>東日本大震災の発生から10年となった2021年より、毎年3月に国内の報道機関をはじめとするパートナーと連携し、大震災の記憶を語り継ぎ、その教訓を未来へつなぐキャンペーンを展開。2022年は、「#防災いまできること」「#PrayForTohoku」の共通ハッシュタグを使い、3月10日・11日にはパートナーが同じタイミングでツイート。</li> <li>ウクライナ情勢につきましては、キュレーションチームがメディアからの正確な情報をまとめている。</li> </ul>

※本資料は、各社提出のヒアリングシート等から事務局において主な項目を抜粋したものであり、詳細は各社提出資料を参照。

### これまでの主な意見

#### 【モニタリングについて】

- (LINEの拡散のおそれのある情報をユーザーの目につきにくくする取組に関し、) せっかく取り組んでいるので、効果測定などを工夫してやっていただきたいと思う。【木村構成員】
- (Metaの6つのヒントなど情報に関し、) せっかく取り組んでいるので、情報が利用者の方によりよく届くよう、お願いしたい。【木村構成員】
- ITの技術をこの世界でどのように活用していくのかという点についても、できれば話していただけるとありがたい。例えばコンテンツであれば、内容的真実性と手続的真正性という分類があるのではないか。内容的真実性については、まさに今皆さんがファクトチェックなどでやっているということはよく分かる。手続的真正性となると、つまり、どういう人がコンテンツ、またはデータのトラストをどのように考えた上で、それらを利用していくという環境を提供していくのか。非常に匿名性の高いサービスというのを売りにするというのもあれば、非常に厳格なサービスにするということで、そこでトラストは大分違ってくると思う。そのようなことも今後の考え方にと考えた。【手塚構成員】

#### 【EUにおける欧州民主主義行動計画（EDAP）（中間とりまとめまでの動き）】

- 欧州委員会は、EU全域でより強靱な民主主義を構築することを目的に、2020年12月に欧州民主主義行動計画を公表。①自由で公正な選挙の促進、②メディアの自由の強化、③偽情報への対抗措置の3つの柱により構成される。
- 偽情報への対抗措置に関しては、
  - ①偽情報の発信者に対するコストを科すための取組
  - ②プラットフォーム事業者が署名した「the Code of Practice on Disinformation（偽情報に関する行動規範）」の見直し、co-regulatory framework（共同規制）化
  - ③上記に伴い、オンラインプラットフォームマーに向けて、「行動規範を強化するためのガイダンス（guidance to enhance the Code of Practice）」を発行（2021年春）。その後、EUは新たな行動規範の実施状況をモニタリング
  - ④EU内外の偽情報へのメディア・リテラシー向上プロジェクトへ支援と資金提供等の取組について記載。
- プラットフォーム上の偽情報への取組は、自主的な取組である「偽情報に関する行動規範」に加え、Digital Services Act（DSA）が加わり、2段構造。EDAPはDSAを補完するものとされており、DSAに示された措置をEDAPにおいて具体化。
- EDAPの3つの柱のうち、「自由で公正な選挙の促進」、「偽情報への対抗措置」がDSAと関連。広告に関して、DSAではすべての広告についての透明性規律等が規定されているが、EDAPでは、政治広告について、政治的文脈におけるスポンサー付コンテンツの分野における透明性の向上を確保するための法律について提案。

### 【EUにおける欧州民主主義行動計画（EDAP）（中間とりまとめ以降の動き）】

#### ① 自由で公正な選挙の促進

- 2021年11月、政治広告の透明性とターゲティングに関するEU規則案が提示された（2023年4月発効予定）。同規則案は、政治的広告及び関連サービス提供者等を対象として、透明性義務及びターゲティングと増幅（アンプリフィケーション）の厳格な条件などを課す。
- DSAと比較して、政治広告の文脈で開示されるべき情報のカテゴリーと、関連するサービス・プロバイダーの範囲が拡大。DSAがオンラインプラットフォームに透明性要件を課しているのに対し、本イニシアチブは政治広告の出版社や、政治広告の準備、配置、宣伝、出版、普及に関与するその他の関連サービスプロバイダー全体を対象とする。

#### ② メディアの自由の強化

- 2021年12月、欧州委員会が欧州メディア自由法案について協議。同法案は、欧州域内の各国の規則の違い、規制当局間の協力体制が不十分、メディアへの官民の干渉、メディア多元主義のセーフガード不足などの課題に取り組むもの。
- 2022年1月、欧州メディア自由法案のパブリックコンサルテーションを開始（2022年第3四半期に委員会で採択予定）。

#### ③ 偽情報への対抗措置

- 同規則案は、現在ガイダンスに沿った改訂が進められている「偽情報に関する行動規範」の枠組みの中で実施されることを予定。
- 当初は2021年秋に署名者が強化された行動規範の最初の草案を委員会に提出する予定であったが、2021年12月までに合計66社（当初16社、ガイダンス公表後50社追加）が改訂に加わり、行動規範の更新プロセスは2022年3月末まで延長された。

#### 【フランスにおける情報操作との戦いに関する法律】

- 2018年12月、情報操作との戦いに関する法律が制定。2021年7月、外国の偽情報等によるデジタル干渉に対抗する機関としてViginumが創設。同年10月、運用開始。
- Viginumは選挙に限らず、幅広い分野を監視の対象とする組織であり、500万人以上の月間ユーザーを超えるオンラインプラットフォームを監視する。フランスの基本的利益を害する可能性のある外国のデジタル干渉活動を特定し、その影響を分析、特に選挙期間中に市民が触れる情報を変更する可能性が高い場合には、省庁間レベルでそれに対する国家の保護を指導し調整する責任を負う。

#### 【オーストラリア政府の偽情報等に対する取組】

- 2018年末、選挙プロセスへの海外からの干渉に対抗措置として「国家安全保障法改正(スパイ活動と外国干渉法)」が施行。
- 2021年2月、デジタル産業業界団体のDIGIは「The Australian Code of Practice on Disinformation and Misinformation」(オーストラリアの偽・誤情報に関する行動規範)を公表、これを受け2021年5月には署名企業8社が透明性レポートを公表。署名企業は、オンライン上の偽情報の安全措置と拡散を減らすための措置を講じる。
- 2022年3月、政府はオンラインでの有害な偽情報や誤情報に取り組む法律「New disinformation laws」を制定することを公表。
- 2022年3月、政府の「New disinformation laws」発表に合わせて、通信メディア庁(ACMA)は「misinformation report Fact sheet」を公表。Fact Sheetでは、1. 「デジタルプラットフォームの偽情報とニュース品質対策の適切性」報告書 主な調査結果、2. 行動規範の評価、3. 次なる取り組みが記載される。

#### 【海外におけるファクトチェック団体の動向（中間とりまとめまでの動き）】

- 全世界でファクトチェック団体は2020年10月時点で304団体が活動しており、1年前の調査結果と比較して100団体増加。2016年と比較するとインドネシア、韓国、インド等のアジア地域での増加。
- 大きく分けると、自らファクトチェックを行う「ファクトチェック団体」（Full Fact、FactCheckNI、Poynter Institute、SnopesMedia Group Inc.、台湾ファクトチェックセンター等）と、ファクトチェック活動の推進等を行う「ファクトチェック関連団体」（ソウル大学ファクトチェックセンター等）に分類が可能。
- 代表的なファクトチェック団体では、ファクトチェック体制（人数）は十数名～最大30名弱程度。ファクトチェック実績は、年間数十～最大2000件程度。Facebookのサードパーティファクトチェックプログラムに加盟するなど、メディアやプラットフォーム事業者との連携が行われている。
- ファクトチェック以外の活動としては、ファクトチェッカー等の人材育成や、ファクトチェック教育コンテンツの作成・公開、ファクトチェックに関するワークショップの開催等が行われている。
- 収入源としては、①寄付や助成金、②IFCN署名済団体におけるFacebookからのファクトチェック作業委託費（年間数百万～数千万円）、③ファクトチェックコンテンツの外部販売や広告収入、④研修事業の外部提供による収入確保、⑤他のファクトチェック団体からの活動支援が挙げられる。
- ファクトチェック関連団体であるソウル大学ファクトチェックセンター（SNU）は、メディアと大学が共同で実施する非政治的・非営利型の公共情報サービス。SNUでは、韓国内の新聞・放送・通信・オンラインメディア(30社が登録)のファクトチェック結果を集め、大手SNS事業者ネイバーを通じて公表している。メディア各社が同じフォーマットでファクトチェック結果を登録し、公表されることが特徴。ネイバーからの資金援助を受けている（年間1億円）。また、COVID-19関連のファクトチェック結果については、ソウル特別市のウェブサイト「コロナ19報道機関ファクトチェック」に連動して表示されている。

### 【海外におけるファクトチェック団体の状況（中間とりまとめ以降の動き）】

- IFCNが署名機関のファクトチェック団体を対象として調査した結果によると、2021年度の調査結果のポイントとして、「営利目的のファクトチェック団体が減少したことで、非営利団体が再びファクトチェックコミュニティの原動力となり、回答者の50%を占めた」と述べている。また、Meta（旧Facebook）の「サードパーティファクトチェックプログラム」（44.2%）が最も主要な収益源となっており、次いで、「助成金、寄付、会費」収入（33.7%）であった。
- 回答団体のファクトチェック活動のための年間予算額（組織全体の予算額ではない）は「10～50万ドル」（約1,200-6,200万円）規模の団体が最も多くなった。次に多かったのは、「0～2万ドル」（約0-250万円）規模の団体であった。なお、「10～50万ドル」の団体は2019年以降増加傾向を示しており、2021年は2019年よりも16ポイント増加した。逆に、2018年に最も多くを占めた「5～10万ドル」（約600-1,200万円）規模の団体は減少傾向が続いている。

### 【海外におけるファクトチェック支援団体の状況】

- IFCNの2021年の収益合計は約2.8億円であった。2021年にIFCNが受けた助成金・拠出金プログラムの上位5つをみると、Google、Facebook、WhatsAppのプラットフォーム事業者が上位を占めていた。他には、非営利団体である「Institute for War and Peace Reporting」（IWPR）からも支援を受けている。
- IFCNの2021年の費用をみると合計で2.7億円であった。内訳をみると、総人件費（教員や外部プロフェッショナル分を含む）約7,600万円（28%）、「プログラム運営費及びその他費用合計」約1.9億円（72%）であった。

※2021年1月1日～11月30日までの収益

※1ドル=123円で換算（みずほ銀行TTSレート。2022年3月31日時点）

#### 【海外における偽情報対策教育プログラムについて（中間とりまとめまでの動き）】

- 欧州ではメディアリテラシー教育に関するプロジェクトが複数立ち上がっており、EUによる資金援助も活発に行われている。
- 欧州委員会によるメディアリテラシー向上の取組を支援する目的のプログラムである「Preparatory Action on Media Literacy for All」では、偽情報に対する教材やサービスの開発、および攻撃者への認識を高める取組を実施。2016年から2020年までの期間に毎年50万ユーロまでの予算を確保し、各プロジェクトの総費用の60%を上限に資金援助を実施。2021-27年には、文化・クリエイティブ産業振興策である「Creative Europe」の枠組みの中で予算を拡大し、新たなメディアリテラシープロジェクト支援が行われている。
- 例えば、ベルギーの非営利団体AISBLでは、偽情報とフェイクニュースに関するワークショップの提供を通じ、若者から祖父母まで幅広い年齢層のメディアリテラシーの向上を図ることを目的としたプロジェクト「GET YOUR FACTS STRAIGHT! (GETFACTS)」が実施されている。GETFACTSの主たるターゲットグループは、若者（14～16歳）とその親と祖父母、また、偽情報への取り組みとメディアリテラシースキルの習得に関心のある人。2つのセッションから構成される、合計10時間のトレーニングであり、オンラインで実施（一部オフライン）。

### 3. 今後の取組の方向性

---

## 論点

## 【0. 議論のスコープについて】

- インターネット上の誹謗中傷や偽情報といった違法・有害情報の流通に関しては、依然としてSNS等のプラットフォームサービスの影響が大きく、プラットフォーム事業者を中心とした対応が求められる。
- しかしながら、現在のインターネット環境においては、CDN事業者によるコンテンツのキャッシュや、ホスティング事業者によるコンテンツのホストなど、ネット環境の担い手が多岐にわたっており、情報の削除や発信者特定など、違法・有害情報対策の実務では、これらのネット環境の複雑化に伴う実効性の低下等が問題となっている。また、プラットフォームサービス以外の、中小の掲示板や、まとめサイト等のミドルメディアにおける違法・有害情報も問題となっている。
- したがって、プラットフォームサービス以外の、CDN・ホスティング（クラウドサービス）・アプリマーケット・ミドルメディア等も射程に含め、コンテンツ流通メカニズム全体を踏まえながら、引き続き違法・有害情報対策に関する検討を行っていくことが必要ではないか。
- さらに、ヘイトスピーチ・部落差別・性被害・自殺誘引等、様々な種類の違法・有害情報が問題となっていることから、これらの誹謗中傷や偽情報以外にも含む違法・有害情報全般について対策を行っていくことが必要ではないか。

## これまでの主な意見

## 【ネット環境の全体構造を見据えた対応について】

- CDN事業者によるコンテンツのキャッシュや、ホスティング事業者によるコンテンツのホストなど、ネット環境の担い手が多岐にわたっている。そのため、情報の削除や発信者特定など、違法有害情報対策の実務では、手続きに時間がかかり煩雑になっているなど、実効性が伴わない場合もある。現在、アクセスプロバイダやコンテンツプロバイダによる単純なネット構造を前提とした制度が整えられてきたが、現実はずっと複雑。CDN事業者の対応も含めて、コンテンツ流通メカニズム全体を踏まえて、PF研の議論の方向性を見極めることが必要。現状を単純な切り口で捉えて制度設計をすればかえって遠回りの場合もある。PF研の射程を広げていく努力も必要ではないか。【大谷構成員(第27回会合)】

#### 論点

【誹謗中傷等の違法・有害情報の流通状況の把握について】

- インターネット上の誹謗中傷や偽情報といった違法・有害情報の流通に関しては、プラットフォーム事業者において、自社サービス上でどのような違法・有害情報が流通しているのか、自ら実態把握とリスク評価を行うことが必要ではないか。
- 総務省は、誹謗中傷といった違法・有害情報の流通について、相談機関等における相談件数や相談内容に関する傾向の把握や、目撃経験や被害経験に関するユーザ調査等を通じ、実態把握を継続して行うことが必要ではないか。

## 論点

### 【ユーザによる主体的な情報環境の選択】

- ユーザ自らが望ましいと判断する情報環境を目指すための環境整備が重要ではないか。
- 個人がデジタルを通じて自身が触れる情報の自律的なコントロールを可能にするための環境整備が重要ではないか。

### <鳥海教授発表（第35回会合）>

- アテンションエコノミーの経済圏においては、プラットフォーム側からみるとより多くのユーザのアクセスを集め、ユーザ側からみると欲しい情報の収集の効率化につながるため、短期的には双方にメリットが生じるが、長期的・社会的には、エコーチェンバーやフィルターバブルの発生によって、インフォデミックやフェイクニュースの拡散、社会的分断といった問題を生じる。
- この状況について、食をアナロジーに捉え、自らを律して健康的な食事を実現することで長期的な健康維持のために短期的な欲望に勝つように、情報についても、自らが摂取する情報について長期的な利得のための短期的なアテンションに勝つという、情報的健康を考えることができるのではないか。
- どのような状態が情報的健康であるかについて定義は容易ではなく、どのモデルケースを目指すかは個人の自己判断であるが、複数のモデルケースが提示可能ではないか。また、個人がそのモデルケースを目指すためにも、例えば、情報カロリー表示<sup>\*1</sup>や情報ドック<sup>\*2</sup>といった個人が今どのような状態にあるかを可視化することなどを通じた支援が必要であり、個人が摂取する情報に関する自分自身の状況を理解することが必要となる。
  - \*1 情報のメタ情報の提示。例えば、新聞社やメディアの透明性などの評価を提示する機能やニュースがSNS上でどのように理解されているかを提示する機能。
  - \*2 人間ドックのような情報的健康度の可視化。例えば、エコーチェンバー可視化システムやYahoo!ニュース検診。
- この実現のためには、プラットフォームが有するデータの活用が不可欠である。多様な情報を見るユーザの方がサービス継続率が高い可能性があることなども踏まえ、プラットフォーマーによる協力が期待される。

## 論点

## 【ユーザによる主体的な情報環境の選択（続き）】

## ＜山本構成員発表（第35回会合）＞

- 憲法の「知る権利（知る自由）」は、①「さまざまな」意見等に接し、これを摂取する自由であると解されており、また、②「さまざまな」意見を摂取する自由の目的は人格発展と民主主義の原理の実現とされている。情報的健康は「さまざまな」情報をバランス良く摂取することでフェイクニュース等に対する免疫を獲得している状態と暫定的に定義づけられるが、「さまざまな意見、知識、情報」を摂取することを、国家に強制されてはならない。**国家が、何が健康的な情報で何が不健康な情報かを定義してはならず、国家に求められるのは、多様な情報へのアクセスの機会を保障することであり、コンテンツに立ち入らないことが重要。**
- フィルターバブルやエコーチェンバーは、ユーザが、フィルターバブルに他律的に閉じ込められると、自己決定の機会や人格発展の機会がそがれ、民主主義や国家安全保障にも関わる問題を生じうる。大規模なプラットフォーム事業者には、フィルターバブルによる弊害を防止する措置をとることが期待される。また、ファクトチェック団体との緊密な連携とファクトチェック記事の効果的な表示が期待される。さらに、コンテンツの表示順位を決定するアルゴリズムやパラメータの透明化が必要である。この透明化を図ることで、悪質なアルゴリズムを用いていることが判明した事業者はユーザ市場やメディア等の記事配信・提供先から淘汰されることが期待される。
- その他、経営とコンテンツの編成の間にファイアウォールを設けることで、コンテンツの提供・配信等について、経営側や社外のステークホルダーによる不当な干渉を阻止することが期待できるのではないかと。また、個人データの収集やUIなどにユーザの基本的な人権を侵害する不適切なものがないかをチェックする多様な分野の専門家から構成される倫理審査委員会を設置するべきではないかと。配信元の妥当性などを監督するコンテンツ審査委員会の設置も期待される。
- また、自社サービス上で発生している問題の実態を適切に把握する必要があり、調査の結果は営業秘密に抵触しない限り定期的に公表し、生データを含む詳細は報道機関や研究者と共有されるべきではないかと。
- これらの取組をどこまで法律で強制するかは慎重な議論が必要だが、**少なくともプラットフォーム間の競争に委ねるための前提となる、当該プラットフォームの取組に関する透明性（情報開示）については、法律で要請する必要がある。**その他の取組は基本的にプラットフォームの自主性を尊重しつつその効果等を外部から検証・監査する共同規制的アプローチが考えられる。

#### 論点

##### 【1. ユーザに対する情報モラル及びICTリテラシーの向上のための啓発活動】

- それぞれのユーザが他人を個人として尊重し、SNSを始めとするインターネット上での自らの書き込みに対して他人が傷つく可能性を想像し、誹謗中傷を行わないよう心がけるなど、ユーザ自身の情報モラルが最も重要である。誰もが誹謗中傷の加害者になり得るし、誰もが偽情報を拡散する可能性があることを認識することが重要である。
- 実態把握や分析結果に基づき、産学官民が連携し、引き続きICTリテラシー向上施策が効果的となるよう取り組み、体系的で多角的なリテラシー啓発を実施することが必要ではないか（分析結果の例：ごく少数の者がネット炎上によるネット世論を作る、書き込む動機は正義感、多くの方は自分が誹謗中傷を書いていると気づいていない、等）。
- 総務省は、これまでのe-ネットキャラバン等の青少年向けの取組に加え、大人も含め幅広い対象に対してICTリテラシー向上のための取組を実施することを検討していくことが必要ではないか。様々な主体により行われている既存リテラシー施策について整理し、様々な主体の連携を促進することが必要ではないか。こうした総務省及び各ステークホルダーによる取組状況を把握し評価を行うことが必要ではないか。
- 普及啓発を実施するにあたっては、ユーザのどのような行動変容を促すのかといった目標の設定と効果分析の仕組みを設けることが重要ではないか。
- 個人が自分自身の触れる情報について自覚的になることへの支援が重要ではないか。その際、国が個人に対して、どのような情報が健康的であるかといった内容を押しつけることがあってはならない。

## 論点

### 【2. (1) プラットフォーム事業者の自主的取組の支援】

#### <違法・有害情報全般に共通する対応>

- 誹謗中傷や偽情報以外も含む違法・有害情報全般（ヘイトスピーチ・部落差別・自殺誘引等）に共通する対応として、まず、違法な情報に対して、プラットフォーム事業者をはじめとするサイト運営者は、プロバイダ責任制限法による免責規定や民間団体が策定するガイドラインを踏まえ、迅速に削除等の対応を行うことが求められるのではないか。
- 法務省人権擁護機関等の関係機関からの削除要請を受けた場合には、それらの手続の正当性や専門性も踏まえ、迅速に削除等の対応を行うことが求められることから、プラットフォーム事業者は、要請された内容に関するポリシーに基づく措置や我が国におけるトラステッドフラグガーの仕組みの導入などについて検討することが望ましいのではないか。
- プラットフォーム事業者・総務省・法務省人権擁護機関による実務者検討会の継続的な開催等により、削除に関する違法性の判断基準・判断方法や個別の事業者における削除実績等について関係者間で共有し、行政側・事業者側双方の削除に関する対応についての透明性を向上させ、円滑な削除対応を促進することが必要ではないか。
- 個別の書き込みが違法な情報か有害な情報かの判断が難しい場合も多いこと等を前提に、違法ではないが有害な書き込みについては、自らのポリシーや約款に基づき、適切に削除等の対応を行うことが求められるのではないか。
- 削除以外にも、それぞれのサービスの特性に応じた、アーキテクチャ上の工夫による違法・有害情報対策を進めることが期待されること、ヒアリング結果を踏まえ、特に一定の短期間に大量の誹謗中傷が集まった場合に、既存の機能・取組において効果的な対応が可能なのかという点について自ら検証を行い、仮に効果が見られない場合には、更なるアーキテクチャ上の工夫の導入について検討を行うことが望ましいのではないか。

## 論点

### 【2. (2) プラットフォーム事業者による取組の透明性・アカウントビリティの向上】

#### <モニタリング結果>

- ヒアリングによると、プラットフォーム事業者の誹謗中傷への対応に関する透明性・アカウントビリティ確保状況は、前回ヒアリング状況から一部進展が見られるものの、一部項目において、依然、透明性・アカウントビリティの確保が十分とは言えない状況ではないか。
- 我が国において、コンテンツモデレーションが過不足なく実施されているかに関する透明性・アカウントビリティ確保が図られていない事業者に関しては、特に透明性・アカウントビリティ確保の取組を進めることが強く求められるのではないか。

#### <透明性確保の必要性>

- 国は、引き続きプラットフォーム事業者等による自主的な削除等の対応を促進すべきではないか。ただし、プラットフォーム事業者等に対して削除義務を課すことや、個別のコンテンツを削除しなかったことに対して罰則等を設ける法的規制を導入することは極めて慎重な検討を要するのではないか。
- 総務省は、プラットフォーム事業者によるコンテンツモデレーションが過不足なく行われているかに関する検証可能性を確保する観点から、行動規範の策定及び遵守の求め又は法的枠組みの導入等の行政からの一定の関与を検討することが必要ではないか。

## 論点

【2. (2) プラットフォーム事業者による取組の透明性・アカウントビリティの向上 (続き)】

＜共同規制的枠組みの構築＞

- 自由放任モデル、伝統的規制モデル、共同規制モデルについて、実効性や表現の自由の侵害のおそれなどの観点から比較したとき、共同規制モデルを基本に透明性確保のための枠組みを検討するのが最も適切ではないか。
- ①リスクベースアプローチに基づく検討、②特に、リスクの大きい巨大プラットフォームサービスについて、自らのサービスのリスク評価の実施及び結果の公表、③リスクを低減するための合理的・比例的・効果的な対応の実施とその結果及び効果の公表、④政府及び外部研究者等による継続的なモニタリング、⑤モニタリングを可能とするデータ提供、といった大枠としての共同規制的枠組みの構築を前提に検討を進めることが適切ではないか。
- 総務省は、プラットフォーム事業者による自らのサービスに関するリスク評価やその低減のための措置と効果、結果等に関して、プラットフォーム事業者の取組状況に対する政府及び外部研究者による継続的なモニタリングを可能にするための枠組みを設ける必要があるのではないか。
- 総務省は、継続的に国際的な法的規制枠組みの検討状況を把握し、国際的対話を深めていくことが適切ではないか。グローバルにサービスを提供するプラットフォーム事業者における適切な対応について、諸外国の情報通信担当部局等と連携しながら、実効的な対応を検討していくことが必要ではないか。G7における成果文書等も踏まえ、グローバルにサービスを提供するプラットフォーム事業者の透明性・アカウントビリティ確保に関して、海外政府や国際機関における議論と協調して実施することが重要ではないか。

## 論点

【2. (2) プラットフォーム事業者による取組の透明性・アカウントビリティの向上 (続き)】

<その他>

- プラットフォームサービス以外のサービス (CDN・ホスティング (クラウド) ・アプリストア等) における違法・有害情報対策に係る取組についても、必要に応じて今後ヒアリングを行い、透明性・アカウントビリティ確保を求めていくことが望ましいのではないか。また、その必要性を判断するための実態の把握が必要ではないか。
- ヒアリング結果によると、AIの活用に関して、各社において深層学習を用いた自然言語処理モデルを活用した違法・有害情報への対応がすでに進められていることから、引き続き、これらの取組を進めることが有用ではないか。他方で、AIの活用によるオーバーブロッキング等の懸念もあることから、AIの活用に関して具体的な透明性・アカウントビリティ確保を図っていくことが望ましいのではないか。
- 具体的なモニタリング事項や法的枠組みの検討に関しては、既存のヒアリングシートを基本として、プラットフォーム事業者と対話を行いながら検討することが適当ではないか。より適切な指標や項目があると考えられる場合には、プラットフォーム事業者は、自らのサービスの特性を踏まえ、代替案となる指標や取組を積極的に示すことが望ましいのではないか。

## 論点

### 【3. 発信者情報開示関係】

- 改正プロバイダ責任制限法の施行に向けて、関係者及び総務省の間で、具体的な運用に関する協議を継続することが必要ではないか。
- その際、現在のインターネット環境においては、CDN事業者によるコンテンツのキャッシュや、ホスティング事業者によるコンテンツのホストなど、ネット環境の担い手が多岐にわたっており、情報の削除や発信者特定など、違法・有害情報対策の実務では、これらのネット環境の複雑化に伴う実効性の低下等が問題となっていることを踏まえ、プラットフォームサービス以外の、CDN・ホスティング（クラウドサービス）事業者等も含めた協議を継続するべきではないか。
- 円滑な発信者情報開示制度の運用にむけて、プラットフォーム事業者は、2.（2）に記載の透明性・アカウントビリティ確保の取組の中で、我が国における発信者情報開示に関する請求や開示件数等について集計・公開することが望ましいのではないか。また、総務省は、法務省や裁判所等と連携し、行政側でも現行制度及び新制度に関する発信者情報開示の件数等を把握することが求められるのではないか。

### 【4. 相談対応の充実】

- 違法・有害情報相談センターにおいて、引き続き被害者救済のための運用を着実に行うとともに、ユーザビリティに資するシステム更新等を随時検討していくことが望ましいのではないか。
- 総務省は、複数の相談機関間における連携強化を一層深めていくことが必要ではないか。
- また、相談を必要としている被害者に対して違法・有害情報相談センター等の必要とされる相談機関の相談窓口に関する情報が届くよう、複数の相談窓口の案内図について広く周知を行うなど、引き続き、被害者にとって相談窓口を分かりやすく示すための取組を行うことが必要ではないか。

## 論点

## 【5-1. 透明性・アカウントビリティ確保の重要性について】

- 多くのユーザが自由な情報の発信・受信を可能にするプラットフォーム事業者は、自社サービス上で誹謗中傷や偽情報といった違法・有害情報も多く流通する中、情報が流通している状況を覚知した場合には、削除やアカウントの停止、ラベルの付与等のコンテンツモデレーションを実施するなど、情報流通の適正について一定の責任を果たすことが期待されるのではないか。一方で、プラットフォーム事業者は、ユーザによる表現を預かる立場でもあり、ユーザの表現の自由の尊重について一定の責任を果たすことが期待されるのではないか。
- また、プラットフォーム事業者は、サービスの特性や誹謗中傷等の情報がユーザに与えるリスクを分析した上で、文化的、社会的、政治的背景を踏まえた、コンテンツモデレーションの実施に係るポリシーの設定とその実施に必要な体制の構築をはじめとするリソースの確保や、自社サービス上で生じた紛争解決のための発信者情報開示などの裁判手続きへの適切な協力などが期待されるのではないか。こうしたコンテンツモデレーションについては、削除以外の手法による対応も含め、事業者による自律的な創意工夫による対応が行われることが望ましいのではないか。
- すでに大規模なプラットフォーム事業者では、誹謗中傷等の不適切な情報への措置を講じる必要性が認識されており、あらかじめ対応方針や基準となるポリシーを自主的に設定し、投稿の削除やアカウントの停止等のコンテンツモデレーションを行っている例が見られる。こうした措置については、措置の対象とされるべき情報に対して確実に措置が行われることが望ましい一方で、行過ぎた措置や恣意的な措置といった不適切な運用によってユーザの表現の自由が損なわれることがないよう、過不足なく実施される必要があるのではないか。

## 論点

### 【5-1. 透明性・アカウントビリティ確保の重要性について（続き）】

- 違法・有害情報への対応が、対応されるべき情報が適切に対応されるとともに、ユーザの表現の自由に対する過度な制約とならないように、過不足なく行われるためには、ポリシーの設定状況やその運用状況、対応結果といった項目に関する透明性を確保し、「言論空間のガバナンスに対するガバナンス」、プロセスの透明性を確保することが必要ではないか。
- その透明性の確保にあたっては、大規模なプラットフォームサービスが情報流通について公共的役割を果たしていることから、当該サービスのユーザだけではなく、すべての者からの検証可能な環境が確保される必要があるのではないか。
- また、透明性の確保にあたっては、サービス上における誹謗中傷の発生件数等の流通実態やその抑制のための対策とその効果に関する総量的な数値等の把握というマクロの観点と、個別具体の誹謗中傷等の情報に対する権利回復のための裁判手続きへの対応や、十分に措置が行われないケースや行過ぎた措置が行われたケースが発生した場合の反論や異議申し立ての機会の確保とその対応状況の把握というミクロの観点の両面から、ユーザ及びユーザ以外の者からの検証可能性の確保が必要ではないか。
- プラットフォーム事業者は、G7における成果文書やサンタクララ原則のような国際的な議論を参考に、国別の数値やモデレーション実施に関する体制の確保などについて、透明性を確保することが望ましいのではないか。なお、サンタクララ原則は行政による規律のひな形ではないと言及されていることに留意する。

## 論点

## 【5-2. 透明性・アカウントビリティ確保のための枠組みの方向性】

- これまで当研究会が実施してきた、大規模なプラットフォーム事業者に対する透明性・アカウントビリティ確保に関する任意での回答の求めには、その回答状況には一部で進展が見られるものの、プラットフォーム事業者によるコンテンツモデレーションが過不足なく行われているかを判断するという観点からは、透明性・アカウントビリティの確保について不十分な点があるのではないか。
- 研究会において実施したモニタリングにおいて、プラットフォーム事業者による透明性・アカウントビリティ確保の取組には進展が見られるところ、引き続きその取組に期待するとともに、透明性が確保されることが望ましい事項を明確化することが望ましいのではないか。また、プラットフォーム事業者による透明性・アカウントビリティの確保について、一過性のものではなく継続的に行われることが担保されることが必要ではないか。
- また、プラットフォーム事業者に対して対応状況等に係る透明性確保や報告を求めることについては、プラットフォーム事業者側からの予見可能性確保の観点からも、行動規範の策定及び遵守の求めや法的枠組みなどの根拠に基づき行われる必要があるのではないか。
- プラットフォーム事業者による透明性に関する報告等について、コンテンツモデレーションが過不足なく実施されているかを把握するために透明性が十分確保されているか、ユーザやユーザ以外の者にとって判断可能な内容となっているか、ユーザがリスクを十分に理解・受容した上でサービスを利用できる環境が整っているか、継続的に把握し、評価する仕組みが必要ではないか。
- 以上を踏まえ、コンテンツモデレーションが過不足なく行われていることがユーザやユーザ以外の者に対しても明らかになるように、総務省は、透明性・アカウントビリティの確保方策に関する行動規範の策定及び遵守の求めや法的枠組みの導入等の行政からの一定の関与について、具体的に検討を行うことが必要ではないか。
- この際、表現の自由や検閲の禁止といった規定に十分に留意する必要があるのではないか。また、国は、ユーザやプラットフォーム事業者に対して投稿の削除を義務づけることについて、過剰削除の懸念から表現の自由の萎縮の観点から、引き続き、極めて慎重であるべきではないか。

## 論点

### 【5-2. 透明性・アカウントビリティ確保のための枠組みの方向性（つづき）】

- プラットフォーム事業者によるコンテンツモデレーションが過不足なく行われているかの検証可能性を確保するためには、
  - 1 誹謗中傷等に関するポリシー
  - 2 一般ユーザからの申告や削除要請に対応する部署・チームや日本国内の拠点・責任者
  - 3 削除等への苦情等に対する苦情受付態勢及び苦情処理プロセス
  - 4 日本における削除要請件数や削除件数
  - 5 発信者情報開示の件数
  - 6 取組の効果分析
  - 7 透明性レポートの公開

といった要素について、当研究会において実施したヒアリングシートの項目を中心に、海外における透明性確保に係る議論において対象とされる情報などを参考にしつつ、透明性を確保すべき対象情報について、プラットフォーム事業者の自主的な取組を尊重しながら、透明性・アカウントビリティが確保されるべき最低限の項目が示される必要があるのではないか。

- 透明性が確保すべき情報については、機械可読な形で、標準化され相互運用性のある形で透明性が報告されることが重要ではないか。
- G7における成果文書等も踏まえ、グローバルにサービスを提供するプラットフォーム事業者においては、グローバルのみならず我が国における透明性・アカウントビリティ確保が行われることが重要ではないか。

## 論点

### 【透明性・アカウントビリティ確保のための枠組みのあり方】

<水谷准教授発表（第31回会合）>

- デジタル革命以降、万人が公平に利用可能な情報発信、受信のインフラが登場したものの、思想内容から刺激の競争を背景に情報資本主義、アテンションエコノミーが到来し、それまで憲法学と親和的であったはずの思想の自由市場は機能不全に至り、民主主義を支える熟議が機能しにくくなっている。
- デジタルプラットフォーム事業者（DPF）は、自身のユーザに対して、心地よく場を利用してもらうため、ひいては自らの経済的な利益のために場の管理等を行い、ユーザが「知るべき」よりもユーザが「知りたい」情報をDPFが「設計した場」に流通させている。ユーザの自由は、DPFの「手のひらの上の自由」となっている。近年、DPFは、自身が定めるルールとその実行組織などを備えた、オンライン言論の新たな統治者（the New Governors）と評されている。
- DPFは、コンテンツモデレーションについて、その規模の大きさとツールの問題から、AIを用いた自動化と、個人の権利よりもリスクと利益を重視する確率論とシステム的な発想に基づいて実施する。こうしたシステム的な手法においては、必然的に生じるエラーを許容した上でガバナンスの設計に組み込むことが重要。
- これらを踏まえ、今後、憲法学においても、「**オンライン言論ガバナンスに対するガバナンス**」が重要であり、特に、**プロセスの透明性を高め、その設計がユーザにとって正統性を有するものかと問う、DPFに対するデュープロセス的観点**が重要となる。その観点からの国際的な議論としては、市民団体や学識者によるマニラ原則とサンタクララ原則が策定されており、DPFが果たすべき透明性と説明責任に関する基本原則が議論されている。
- **現代においても政府による発信者に対するコンテンツ規制の対象拡大や厳罰化には慎重を期すべきであり、デジタルプラットフォーム事業者に対する特定のコンテンツ削除義務を課す手法についても、デジタルプラットフォームの「検閲代理人化」とオーバーブロッキングの懸念がある。**
- **DPFによるオンライン言論ガバナンスは、削除に限らない様々な手法を駆使できる利点があり、事業者の機能的な「自律性」を認めつつ、政府はDPFのプロセスの透明性や適正性の実現を促進し、ひいてはDPFによるユーザの統治の「正統性」を確保するためのルール形成に注力するべきである。**また、政府自身もDPF規制に関しての透明性を確保しなければならない。

## 論点

### 【透明性・アカウントビリティ確保のための枠組みのあり方】

<曾我部教授発表（第35回会合）>

- **SNS事業者は、情報流通の媒介者であり社会的責任として、利用者の表現の自由の尊重が求められる一方で、違法・有害情報の流通も当然想定される中での情報流通の適正について一定の責任が求められる立場である。**また、基本権や憲法上の権利の主体でもあり、SNS事業者に規制を課す場合には法律の根拠を有する。
- 規律の構想にあたっては、国家・事業者・個人の三面関係を踏まえることを要する。（国民は国家とPFに対する表現の自由を主張する立場、PFは国民に対する利用規約やアーキテクチャによる規制等と国家に対する表現の自由と営業の自由を主張する立場、国家は国民に対する一般法に基づくPF上の行動の規制とPFに対する削除義務などの直接的な義務づけ等の規律を主張する立場）
- 誹謗中傷対策には、総量を減らすというマクロの視点と、個別の被害救済というミクロの視点がある。このマクロの視点には3つのモデルが考えられる。
  - ①自由放任モデル：特段の規律を設けず、一般法（民法、刑法など）の規律に委ねる。
  - ②伝統的法規制モデル：（監視義務や）削除義務を課し、その違反には責任を問う。
  - ③共同規制（規律された自主規制）モデル：情報流通の適正確保の義務を課し、透明性の確保等を通じ担保。
- 自由放任モデルには、違法・有害情報も含む情報の拡散力の飛躍的向上やアテンションエコノミーの発生、法執行の困難性、思想の自由市場のフィクション性の顕現による国家の介入による自由な言論空間の確保の必要性などの観点から課題がある。伝統的法規制モデルには、行政による表現内容への介入のおそれや執行リソースの大きさや、判断の困難性と過剰削除のおそれや削除以外の対応という事業者の創意工夫の余地を阻害する点から課題がある。
- 共同規制（規律された自主規制）モデルでは、透明性・説明責任の義務を通じて、情報流通の適正確保の義務の担保を図る。通報窓口や異議申し立てなどの体制の整備など、具体的な取組に関して義務づけるか事業者の創意工夫に委ねるかについては、事業者に対する予測可能性や過剰・過少削除のおそれの観点からバランスを図る必要がある。近年の欧州の立法などをみても透明性確保による情報流通の適正確保はトレンドといえる。**透明性の確保は、国民による監視機能の強化や義務違反の制裁の容易化の効果があり、政府による評価はこれを補助する。**放送法における番組準則を各事業者が定める規律に、共同規制モデルは一定の類似性がある。

## これまでの主な意見（続き）

## 【透明性確保のための枠組みのあり方について】

- 方向性として自由放任は難しく、共同規制で考えていくという方向性は異論ないが、その中で、方向性そのものについてどう考えていくのかがはっきりしていない。ゴールをきちんと定める必要がある中で、ゴールを普遍的・一般的なものにするというのが相当難しく、その議論を延々と続けるのもかなり難しいのではないかと思う。逆に社会的なコンセンサスをどうプラットフォームやSNSに反映させるかという仕組みを作るなど、様々な方向性で、SNSが自ら考えていくための仕組みのようなものを作っていくことが先ではないかと思う。特に透明性でこういった評価をしていくところ、重要なポイントが1つあり、例えばこの評価を見て、ここはできている・できていないや、優劣という判断ではなく、この方法だとこのようリスクが残っているとといったリスク評価をしていくようなモニタリングとその評価が、実は一番重要なのではないかと思う。その中で、各社がリスクに対応するための仕組みを義務化していくといった方向性で考えるほうが、現実的なのではないか。【寺田構成員（第35回会合）】

## 1 自主的スキームの尊重

・民間による自主的な取組を基本とした対策を進めていくとともに、総務省はモニタリングと検証評価を継続的に行っていくことが必要ではないか

## 2 我が国における実態の把握

・PF事業者の認識や実態把握と調査結果とのギャップが生じていることから、プラットフォーム事業者は、自らのサービス上で生じている我が国における偽情報の問題について適切に実態把握を行い、研究者が分析を行うために必要な情報の無償で情報提供が行われることが望ましいのではないかと

## 3. 多様なステークホルダーによる協力関係の構築

・「Disinformation対策フォーラム」「Innovation Nippon」等において継続的に議論・研究が行われることが望ましいのではないかと

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

・プラットフォーム事業者は、リスク評価に基づき偽情報への対応を適切に行い、それらの取組に関する透明性・アカウントビリティ確保を進めていくことが求められるのではないかと

・総務省は、これらの取組に関するモニタリングと検証評価を継続的に行っていくことが必要ではないか。どのような方法や情報により偽情報への適切な対応が図られているかどうかを評価することが可能かについて引き続き検討が必要ではないかと

・国は、引き続きプラットフォーム事業者等による自主的な削除等の対応を促進することとし、プラットフォーム事業者等に対して削除義務を課すことや、個別のコンテンツを削除しなかったことに対して罰則等を設ける法的規制を導入することは極めて慎重な検討を要するのではないかと。

## 5 利用者情報を活用した情報配信への対応

・広告の種類・対応に応じてリスクや問題の差異を分析したうえで、特に、偽情報を内容とする広告の配信やターゲティング技術の適用については、そのリスクを踏まえ、より注意深い対応と、それに伴う透明性・アカウントビリティ確保が求められるのではないかと

## 6 ファクトチェックの推進

・プラットフォーム事業者・ファクトチェッカー・ファクトチェック推進団体・既存メディア等が連携し、取組がさらに進められることが期待されるのではないかと

・我が国におけるファクトチェック結果を積み重ねて分析を行うことにより、偽情報の傾向分析やそれを踏まえた対策の検討が行われることが望ましいのではないかと

## 7 情報発信者側における信頼性確保方策の検討

・現代のメディア環境に対応した情報の信頼性の確保の在り方について、既存メディア・ネットメディア・プラットフォーム事業者など関係者の間で検討を深めていくことが望ましいのではないかと

・ミドルメディアを中心とした偽情報の生成・拡散・流通メカニズムに関する実態把握と分析も踏まえ、検討を深めていくことが望ましいのではないかと

## 8 ICTリテラシー向上の推進

・偽情報の特徴を踏まえながら引き続きICTリテラシー向上施策が効果的となるよう取り組むことが必要ではないかと

## 9 研究開発の推進

・ディープフェイク等に対抗するための研究開発や事業者の対応が進められることが望ましいのではないかと

## 10 国際的な対話の深化

・偽情報に関する政策について国際的な対話の深化を深めていくことが望ましいのではないかと

#### 論点

#### 【0. 議論のスコープ・偽情報への対応の方向性について】

- 引き続き、2021年9月の中間とりまとめにおいて記載の偽情報への対策に関する10の方向性について、各ステークホルダーの取組状況やプラットフォーム事業者のモニタリング状況を踏まえ、継続的に議論を深めていくべきではないか。

#### 【1. 自主的スキームの尊重】

- 我が国における偽情報への対応の在り方の基本的な方向性としては、引き続き、まずはプラットフォーム事業者を始めとする民間部門における関係者による自主的な取組を基本とした対策を進めていくとともに、プラットフォーム事業者等による自律的な対応及び自主的な報告を求め、総務省はモニタリングと検証評価を継続的に行っていくことが必要ではないか。
- 特に、プラットフォーム事業者等に対して削除義務を課すことや、個別のコンテンツを削除しなかったことに対して罰則等を設ける法的規制を導入することは、誹謗中傷と比べても、極めて慎重な検討を要するのではないか。
- 誹謗中傷等と同様に、偽情報についても、すでにプラットフォーム事業者では、偽情報等の不適切な情報への措置を講じる必要性が認識されており、あらかじめ対応方針や基準となるポリシーを自主的に設定し、投稿の削除やアカウントの停止等のコンテンツモデレーションを行っている。こうした措置については、措置の対象とされるべき情報に対して措置が確実に行われることが望ましい一方で、行過ぎた措置や恣意的な運用によってユーザの表現の自由が損なわれることがないよう、過不足なく実施される必要があるのではないか。
- そして、コンテンツモデレーションが過不足なく実施される状況を確認するためには、ユーザやユーザ以外の者からの透明性を確保することが必要ではないか。また、その検証のためには、ポリシーの設定状況やコンテンツモデレーションの実施のために投入されるリソース、実施の結果等の情報といった情報について、定量的な数値でもって説明されることが重要ではないか。

#### 論点

#### 【2. 我が国における実態の把握】

- 研究者による偽情報に関する実態把握や分析が進められているものの、偽情報の流通状況が明らかになっているとは言えないため、偽情報特有の問題の検討に資するよう、以下の点に着目しつつ、引き続き偽情報の実態把握を行うことが適当ではないか。
- 特に、研究者によるサービス上の情報流通についての調査によると、我が国において偽情報の問題が顕在化しているにもかかわらず、ヒアリング結果によると、プラットフォーム事業者は自らのサービス上の偽情報の流通状況について、我が国において適切な実態把握がなされていない。したがって、プラットフォーム事業者は、自らのサービス上で生じている我が国における偽情報の問題について適切に実態把握を行うとともに、研究者が分析を行うために必要な情報についてプラットフォーム事業者から無償で情報提供が行われることが望ましいのではないか。
- プラットフォーム事業者の中には、実態把握が進まない理由として、日本国内の情報として切り分けることの難しさや、どの情報を虚偽の情報として取り扱うべきかについての基準の判断が難しい旨の理由の説明があった。我が国に特有の情報の真偽を検証するファクトチェック団体の存在が期待されるのではないか。
- コンテンツ側の偽情報流通の実態把握・分析に当たっては、プラットフォームサービスのみならず、ミドルメディア等も含めた情報流通環境全体を捉えた視点により実施されることが望ましいのではないか。特に、ミドルメディアを中心とした偽情報の生成・拡散・流通メカニズムに関して、実態把握と分析を進めていくことが必要ではないか。
- SNS上の偽情報が偏った複数の集団・ネットワーク（クラスター）に分かれて拡散しているケースが見られることから、これらの既存の分析をもとに、引き続き偽情報の拡散状況や特徴について実態把握及び分析を行い、偽情報の拡散に有効なプラットフォームサービス上のアーキテクチャ上の工夫や、偽情報に関するリテラシー向上に関する取組など、必要な対応について検討していくことが必要ではないか。

#### 論点

#### 【3. 多様なステークホルダーによる協力関係の構築】

- 偽情報への対応に当たっては、多面的な解決策を検討していくことが必要であり、その際、プラットフォーム事業者のみならず、産学官民の多様なステークホルダーによる協力関係の構築を図り、対話の枠組みを設けることが重要と考えられる。
- ヒアリング結果によると、「Disinformation対策フォーラム」や「Innovation Nippon」などにより、我が国において、産学官民の多様なステークホルダーによる協力関係の構築や偽情報への対策の検討が進められている。
- 引き続き、「Disinformation対策フォーラム」や「Innovation Nippon」等の産学官民の連携の場において、継続的に偽情報への対策に関する議論や研究が行われることが望ましいのではないか。
- 特に、「Disinformation対策フォーラム」においては、当該フォーラムによる報告書の記載のとおり、今後、①伝統的なメディア・学術研究機関等と協力を行いながら、ファクトチェックを持続的かつ総合的に行う主体についての具体的検討とその実施を進めること、②情報リテラシーの向上のため、多様な主体によるコンテンツの作成や、総務省や関連団体とも協力して総合的な普及啓発の取組を促進することが期待されるのではないか。

#### 論点

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

<モニタリング結果>

- ヒアリングによると、いずれの事業者においても、自らのサービス上で生じている我が国における偽情報の問題について適切に実態把握・公表が行われていない。一方で、偽情報についても、すでにプラットフォーム事業者では、偽情報等の不適切な情報への措置を講じる必要性が認識されており、あらかじめ対応方針や基準となるポリシーを自主的に設定し、投稿の削除やアカウントの停止等のコンテンツモデレーションと呼ばれる措置を行っている。
- 事業者によるポリシー設定においては、明らかな偽情報やCOVID-19感染症に関する健康・衛生関連に関するポリシー、ウクライナ侵攻などの武力を伴う混乱に関するポリシーなど、身体・生命に重大な危機を及ぼしうる事項に関するポリシーが定められていた。また、これらのポリシーについては、適時に社内外の専門家によって見直されている例があった。

#### 論点

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

##### <コンテンツモデレーションに関する透明性確保の必要性>

- プラットフォーム事業者による我が国における偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は限定的であった。これらの状況を踏まえ、プラットフォーム事業者は、我が国において生じている自らのサービス上の偽情報の問題について適切に実態把握とリスク評価を行った上で、そのリスクに応じて偽情報への対応を行うこととし、また、こうした措置については、措置の対象とされるべき情報に対して措置が確実に行われることが望ましい一方で、行過ぎた措置や恣意的な運用によってユーザの表現の自由を損なうことがないよう、その自らの取組に関する透明性・アカウントビリティ確保を進めていくことが求められるのではないか。
- 偽情報に関するポリシー設定については、プラットフォーム事業者が、関係者と連携の上、情勢等に応じて適時適切に、自主的に設定することが求められるのではないか。
- 国は、引き続きプラットフォーム事業者等による自主的な削除等の対応を促進することとし、プラットフォーム事業者等に対して削除義務を課すことや、個別のコンテンツを削除しなかったことに対して罰則等を設ける法的規制を導入することは極めて慎重な検討を要するのではないか。
- 総務省は、これらの取組に関するモニタリングと検証評価を継続的に行っていくことが必要ではないか。この際、プラットフォーム事業者に対して具体的にどのような対応や情報公開を求めることにより、偽情報への適切な対応が図られているかどうかを評価することが可能かについて、引き続き十分な検討が必要ではないか。
- G7における成果文書等も踏まえ、グローバルにサービスを提供するプラットフォーム事業者においては、グローバルのみならず我が国における透明性・アカウントビリティ確保が行われることが重要ではないか。

#### 論点

#### 【5. 利用者情報を活用した情報配信への対応】

##### <ターゲティング技術の悪用による問題について>

- 行動ターゲティングの技術による政治広告における悪用事例のように、利用者の脆弱性につけ込み偽情報を拡散する使い方が可能であるといった指摘がある。
- 前回ヒアリングに引き続き、いずれの事業者においても、偽情報を内容とする広告や、政治広告の出稿について、一定の制限を設けている。特に、LINE及びTwitterでは政治広告が禁止されており、Googleでは日本において選挙広告は禁止されている。この点、複数の事業者が、偽情報を内容とする広告や政治広告はそもそも配信が禁止されているためターゲティング技術の対象とならない旨を説明している。なお、(出稿が許されている)政治広告について、どのようなターゲティング技術に関する対応が行われているかについては明確になっていない。
- 行動ターゲティング広告における利用者情報の取扱いについては、利用者情報WGにおいても議論が行われているところ、広告の種類・対応に応じてリスクや問題の差異を分析したうえで、特に、偽情報を助長しうるターゲティング技術の適用については、そのリスクを踏まえ、より注意深い対応と、それに伴う透明性・アカウントビリティ確保が求められるのではないか。

##### <偽情報を掲載しているサイトへの広告配信の問題について>

- また、コンテンツの内容が正しくなくてもページビューを稼げば広告収入で儲かる仕組みが偽情報の生成を支えているといった指摘がある。
- ヒアリングによると、複数の事業者において、偽情報に関連する内容の媒体・サイトや、特定の信頼性基準に満たない媒体・サイトへの広告配信を制限する規定が設けられている。
- したがって、広告収益を絶つことにより偽情報拡散防止の観点も踏まえ、これらのポリシーに基づき、実効性のある対応が行われることが求められるのではないか。
- 次回以降のモニタリングにおいて、偽情報に関する行動ターゲティング広告についての対応及び偽情報を掲載しているサイトへの広告配信の問題に関して、引き続き、我が国における透明性・アカウントビリティ確保が実質的に図られているか、実態把握を継続する必要があるのではないか。

#### 論点

#### 【6. ファクトチェックの推進】

- 我が国においても、ファクトチェックの取組が徐々に広がってきているものの、ファクトチェッカーやファクトチェック記事件数は十分とは言えず、ファクトチェックの担い手や社会的認知度・理解度が不足しているという課題は依然として残されていることから、引き続き、プラットフォーム事業者、ファクトチェッカー、ファクトチェック推進団体等が連携し、ファクトチェックの活性化のための環境整備を推進していくことが適当ではないか。
- プラットフォーム事業者においては、①ファクトチェッカー・ファクトチェック推進団体との連携強化（サービス上の情報へのファクトチェック結果の紐付け等）、②資金提供等の取組がさらに進められることが期待されるのではないか。
- ファクトチェッカーにおいては、③国際ファクトチェックネットワーク（IFCN）への加盟、④ファクトチェック体制等に関する透明性・アカウントビリティ確保が進められることが期待されるのではないか。
- ファクトチェック推進団体においては、⑤人材の育成、⑥市民のリテラシー向上、⑦ファクトチェック支援システム等によるファクトチェックの容易化、⑧透明性・アカウントビリティ確保、⑨複数のファクトチェッカーによる複合的なファクトチェック環境の構築に向けた取組等を推進していくことが期待されるのではないか。
- そのほか、今後、我が国におけるファクトチェック結果を積み重ねて分析を行うことにより、偽情報の傾向分析やそれを踏まえた対策の検討（例えば、更なるファクトチェック支援のための技術導入、プラットフォーム事業者におけるポリシー策定や削除等の対応など）が行われることが望ましいのではないか。
- 「Disinformation対策フォーラム」において、引き続き、プラットフォーム事業者・新聞や放送などの既存メディア・学術研究機関等との間で協力を行いながら、同フォーラムの報告書の内容も踏まえながらファクトチェックを持続的かつ総合的に行う主体についての具体的検討を進めることが望ましいのではないか。その際、新聞や放送などの既存メディアにおいても、これまでの信頼性のある情報発信の取組の一環として自律的なファクトチェックの担い手となるとともに、ファクトチェッカーやファクトチェック推進団体等に対して、これまで培ってきた知見やノウハウの共有等の支援を行うことなどが期待されるのではないか。

#### 論点

#### 【7. 情報発信者側における信頼性確保方策の検討】

- プラットフォーム上における偽情報の問題に対抗するためには、前述のファクトチェックの推進とともに、信頼性の高い情報の流通を増やし、利用者が容易にそれらを参照できるような環境整備が必要である。
- この点、我が国においては、全国的に大きな影響力を持つ新聞社や放送事業者等の既存メディアにおいて、自律的なチェック体制に基づく信頼性のある情報発信がこれまで行われてきたほか、例えば日本放送協会における「SoLT」のようなSNS時代に対応した報道態勢が構築されつつある。その他、「Disinformation対策フォーラム」におけるプラットフォーム事業者等との対話も進められている。
- これらの状況も踏まえつつ、新聞社や放送事業者等の既存メディアにおける情報の信頼性の確保のための取組やノウハウも参考とし、これをネットメディアにも広げていくという考え方も取り入れつつ、引き続き、現代のメディア環境に対応した情報の信頼性の確保の在り方について、メディア関係者の間で検討を深めていくことが望ましいのではないかと。
- プラットフォーム事業者においては、新型コロナウイルス感染症等に関して信頼できる政府機関・専門機関・メディア等の情報やそれらへのリンクをサービスの目立つ場所に掲示するなどの取組が積極的に行われているが、認知度が低い（1割程度）ことから、これらの取組を引き続き積極的に行うことが望ましいのではないかと。また、前述のとおり、我が国におけるファクトチェックメディア等との連携により、偽情報に対してファクトチェック結果を紐付けて表示する等の取組を進めていくことが望ましいのではないかと。
- ニュース配信プラットフォームサービスにおいては、ニュースや情報に関する選別・編集責任等に関するサービスの性質を踏まえながら、利用者のニーズに応じて信頼性の高い情報配信が行われるよう引き続き努めるとともに、情報配信に関する透明性やアカウントビリティの確保方策を適切に実施することが望ましいのではないかと。
- 偽情報の拡散要因について、インターネットにおけるニュースの生態系の問題として、ミドルメディアが大きな影響を与えていることがこれまでの分析により判明しつつある（いわゆる「こたつ記事」や「フェイクニュース・パイプライン」問題）。
- したがって、インターネット上におけるメディア全体の情報の信頼性をどのように確保していくかについて、ミドルメディアを中心とした偽情報の生成・拡散・流通メカニズムに関する実態把握と分析も踏まえ、「Disinformation対策フォーラム」等の場も活用しつつ、伝統的なメディア・ネットメディア・プラットフォーム事業者等の関係者間で、ミドルメディア運営事業者との連携可能性等も含め、今後検討をさらに深めていくことが望ましいのではないかと。

#### 論点

#### 【8. ICTリテラシー向上の推進】

- 総務省や一部事業者においては我が国においても偽情報に資するリテラシー向上の取組が行われているが、その他の事業者の既存のリテラシー向上に関する取組は、偽情報に効果的な内容となっているか明らかではない。
- 違法・有害情報全般に関するICTリテラシー向上の推進に向けた取組に加えて、偽情報特有の問題への対応のため、以下の点に着目して、偽情報に対抗するリテラシー向上の推進に向けた取組が行われることが望ましいのではないかと。
- 偽情報の実態把握や分析結果に基づき、スーパースプレッダーへの効果的な働きかけ、偽情報に対して効果のある「情報リテラシー」の分析、人間の認知の仕組み、偽情報のジャンル別に異なる有効な情報検証行動、インターネット上の情報の偏りやメディア環境の分析など、偽情報自体の特徴や偽情報が拡散する要因等を踏まえながら引き続きICTリテラシー向上施策が効果的となるよう取り組み、産学官民が連携し、体系的で多元的なリテラシー啓発を実施することが必要ではないかと。
- 偽情報に対抗するために必要なリテラシーについては、例えば、EU等で先行する取組も参考にしつつ、コンテンツを作成し実施していくことが望ましいのではないかと。
- 偽情報は青少年だけでなく大人であっても拡散しているケースが見られるという分析結果を踏まえ、青少年だけでなく、大人も含め幅広い対象に対して実施することが必要ではないかと。その際、総務省は、e-ネットキャラバンやトラブル事例集等の青少年向けの取組に加え、例えば、デジタル活用支援員の仕組みも活用した取組を検討していくことが必要ではないかと。
- プラットフォーム事業者は、日本向けのリテラシー向上のための取組を実施するとともに、多様な利用者に対して効果的にアプローチするため、行政機関・関連団体・研究者等と協力し、Disinformation対策フォーラム等も活用して、総合的な普及啓発の取組を促進することが望ましいのではないかと。

#### 論点

#### 【9. 研究開発の推進】

- ディープフェイクなどの新たな技術を悪用した偽情報が我が国においても出現している。ヒアリングによると、Google・Meta・Twitterでは、ディープフェイク対策の研究開発が行われている。我が国においては、プラットフォーム事業者によるディープフェイク対策の研究開発は現時点で行われていないものの、研究者による偽情報対策の研究開発が行われている。
- ディープフェイクに対抗するための研究開発や技術コンテスト等の取組が我が国においても進められることが望ましいのではないか。
- プラットフォーム事業者は、ディープフェイク対策のための研究開発を引き続き行うとともに、ディープフェイク等の偽情報に対応したポリシーを設けて、悪質度合いに応じて削除やラベルの付与等の適切な対応を行うことが望ましいのではないか。この際、日本語への対応や、日本人ユーザに対しても適切な形で対応可能な技術について研究開発を推進していくことが望ましいのではないか。
- 総務省は、諸外国及び国内での偽情報対策に資する研究開発の状況を継続的に把握し、関係者に対して情報共有を行うとともに、我が国における研究者が偽情報への対応に関する研究開発を行う際に必要なデータがプラットフォーム事業者から適切に提供されるよう、プラットフォーム事業者及び研究者、プラットフォーム事業者以外の民間企業と継続的な協議を行うことが望ましいのではないか。

#### 【10. 国際的な対話の深化】

- 違法・有害情報全般に関する政策に加えて、偽情報に関する政策についても、国際的な対話の深化を深めていくことが望ましいのではないか。
- 引き続き、諸外国における検討状況を把握しつつ、我が国においても、偽情報に特有の問題や政策的対応について検討していくことが必要ではないか。同時に、我が国における偽情報の流通状況や、産学官民の検討状況について諸外国に共有を行い、それらの対話を通じて、特にグローバルにサービスを提供する事業者における適切な対応について、諸外国の情報通信担当部局等と連携しながら、実効的な対応を検討していくことが必要ではないか。

#### 論点

#### 【11. 今後の取組方針】

- プラットフォーム事業者において、違法・有害情報となりうる偽情報への対応については、前述3. (1) 違法・有害情報対策の方針も踏まえて、透明性・アカウントビリティが確保されることが必要ではないか。
- 偽情報対策としては、引き続き、中間とりまとめにおいて記載の偽情報への対策に関する10の方向性について、各ステークホルダーの取組状況やプラットフォーム事業者のモニタリング状況を踏まえ、継続的に議論を深めていくべきではないか。
- 偽情報は、その外延や個別の情報偽情報であるか、また、その流通による社会に対する影響について、我が国における実態が未だ明らかではなく、一層、我が国における実態の把握を要するのではないかと。また、偽情報の流通状況について、モニタリングにおいて十分明らかではなかったことを踏まえ、プラットフォーム事業者は、引き続き、実態を把握しその結果を公表することが求められるのではないかと。
- 偽情報についても、すでにプラットフォーム事業者では、偽情報等の不適切な情報への措置を講じる必要性が認識されており、あらかじめ対応方針や基準となるポリシーを自主的に設定し、投稿の削除やアカウントの停止等のコンテンツモデレーションと呼ばれる措置を行っている。こうした措置については、措置の対象とされるべき情報に対して措置が確実に行われることが望ましい一方で、行過ぎた措置や恣意的な運用によってユーザの表現の自由を損なうことがないよう、過不足なく実施される必要があるのではないかと。
- プラットフォーム事業者による我が国における偽情報への対応及び透明性・アカウントビリティ確保の取組の進捗は限定的であった。これらの状況を踏まえ、プラットフォーム事業者は、我が国において生じている自らのサービス上の偽情報の問題について適切に実態把握とリスク評価を行った上で、そのリスクに応じて偽情報への対応を行うとともに、その透明性を確保を進めていくことが求められるのではないかと。