

プラットフォーム事業者から追加質問への回答②

目次

1. 偽・誤情報に関するポリシーの運用について	・ ・ ・ P. 1
2. ニュースフィードについて	・ ・ ・ P. 25
3. 特定の通報者からの通報への優先対応について	・ ・ ・ P. 26
4. 信頼できる情報のプロミネンスについて	・ ・ ・ P. 31
5. 発信者への収益還元について	・ ・ ・ P. 35
6. 広告サービスについて	・ ・ ・ P. 40
7. なりすましアカウントやなりすまし広告への対応 について	・ ・ ・ P. 45
8. 生成 AI 対策について	・ ・ ・ P. 55
9. 児童その他脆弱な主体の保護について	・ ・ ・ P. 55
10. 研究機関等へのデータ提供について	・ ・ ・ P. 61
11. ガイドライン等の必要性・妥当性その他社会的傾向について	・ ・ ・ P. 63
12. ファクトチェック組織との連携について	・ ・ ・ P. 65
13. DSA に基づく調査への対応状況について	・ ・ ・ P. 67
14. ガイドライン等の必要性・妥当性について	・ ・ ・ P. 67

(設問)	Google	Meta	TikTok Japan	Microsoft	X
1. 偽・誤情報に関するポリシーの運用について					
Q 1－1 【Google への質問】 YouTube では、2023 年 7 月から 9 月にかけて、コミュニティガイドライン違反を理由として日本国内の IP アドレスからアップロードされた合計 106,620 本の動画を削除し、そのうち「スパム、欺瞞行為、詐欺に関するポリシー」違反により削除されたものが 1,262 本、「誤った情報に関するポリシー」違反により削除されたものが 1,654 本とのことですが、これらの数字に関し、以下の 6 点をご教示ください。なお、ご回答が難しい場合はその理由をご教示ください。 (1) 106,620 本、1,262 本、1,654 本というそれぞれの削除数について、日本国内における削除率(日本国内の IP アドレスからアップロードされたものとして同時期に存在した全ての動画の数に対する割合)をご教示ください。 (2) 透明性レポート¹を拝見すると、106,620 本という削除数は世界 10 位とのことですが、削除率は世界何位でしょうか。 (3) 106,620 本のうち、その他以下のポリシー違反により削除された動画の数及び削除¹率をそれぞれご教示ください。 ・ 選挙の誤った情報に関するポリ	[総合回答] 様々な理由から、通常、国レベルでの詳細なデータの開示は控えさせていただきます。そのような中、私たちはプラットフォームサービス研究会における貴省の取り組みに鑑み、その重要性を深く認識し、Google といたしましても可能な限りの貢献を自発的に行い、最も関連すると思われる一部の内容について日本独自のデータを作成し、ご提供させていただくことになりました。ご賢察の通り、このようなデータの作成には多大な人的リソースと時間を要しておりますが、Google といたしましては、貴省の取り組みに最大限の協力をさせていただきます。何卒ご理解賜りますよう、よろしくお願い申し上げます。 (4) (5) フラグ立てとは、YouTube が違反の可能性のある動画をどのように認識するようになったかを指しています。106,620 本という数字は、フラグ立てられの後に審査された結果、ポリシーに違反していると確認されたことにより削除された動画の数となります。 (6) ヒアリングシートにおいて提供したデータは、その時点では 2023 年第 4 四半期のデータの準備が整っていなかったため、2023 年第 3 四半期のも		(1) ①について TikTok透明性レポートによると、2023年7月～9月期の日本における公開済みの動画数に対する削除率は、1.4%となっています。 ②について TikTok透明性レポートによると、2023年7月～9月期の全世界における公開済みの動画数に対する削除率は、0.8%となっています。 (2) 日本において「誠実性と信頼性」のガイドライン違反により削除された動画の件数のうち、さらに自動的なシステムにより検知がされたものと外部からの通報により発見されたものの数については、公表しておりません。 (3) 上記(2)のとおり、さらなる詳細な件数については公表しておりません。		(1) 各国・地域の投稿総数は公開しておりません。また、ここで求められている形式で詳細に分解されたデータは用意がなく、国別の特定の削除に関する比較チャートも提供していません。 (2) ご質問の数値は、X のポリシーに基づきユーザーからの報告を通じたもののみではなく、弊社の積極的な検出を通じたものも含んでおります。 (3) あるコンテンツのポリシー違反の通報に対し、実際対応された際の根拠となるポリシーは異なる場合がございます。また一つのコンテンツに対し複数のユーザーからレポートが寄せられることもございます。 (4) SAMM ポリシーは、操作されたメディアが人々を欺くか混乱させ、害を及ぼす可能性がある場合に適用されます。私たちの対応は、投稿の削除やメディアにラベルを追加することが含まれる場合があります。日本においては、3 つのユニークな投稿にラベルを追加しました(最初の回答でこれを提供できなかったことをお詫び申し上げます)。 また、2023 年に日本で市民活動の阻

¹ <https://transparencyreport.google.com/youtube-policy/removals?hl=ja>

(設問)	Google	Meta	TikTok Japan	Microsoft	X
シー - 医学的に誤った情報に関するポリシー - なりすましに関するポリシー (4) 106,620 本のうち、自動検出によりフラグ立てされたものが103,531 本、ユーザーからの報告によりフラグ立てされたものが 2,549 本、YouTube 優先報告者からの報告等によりフラグ立てされたものが 540 本とのことですが、ここで言う「フラグ立て」の意味をご教示ください。「最終的に削除されたか否かにかかわらず、自動システムによる検知や外部からの報告が存在したこと」の意味でしょうか。 (5) 上記(4)の1 03,531 本、2,549 本、540 本という数字それぞれについて、同時期に「フラグ立て」された全ての動画の数に対する割合をご教示ください。 (6) 透明性レポートを拝見すると、2 023 年 1 0 月から 12 月にかけては、コミュニティガイドライン違反を理由として日本国内の IP アドレスからアップロードされた合計 9 4,418 本の動画を削除しており、この数は世界 14 位だったとのことですが、上記(1)～(5)でお尋ねした事項²²について、こ	のでした。以下が直近のデータとなります。 2023 年 10 月から 12 月にかけて、日本国内の IP アドレスからアップロードされた合計 94,418 本の動画を YouTube のコミュニティガイドラインに違反したことにより、YouTube から削除しました。そのうちの 96% (90,962 本)は自動検出によってフラグが立てられました。これは、機械学習技術への投資の結果です。残りの内 2,706 本はユーザーによって、そし 750 本は YouTube 優先報告者などによってフラグが立てられました。日本国内で削除された動画のうち、910 本は、「スパム、欺瞞行為、詐欺に関するポリシー」に違反したために削除されました。そして、2,868 本は「誤った情報に関するポリシー」に違反したために削除されました。また、日本国内で削除された動画の内 92% (86,418 本)は、削除された時点で再生回数が 100 回またはそれ以下でした。				害に関するポリシーを施行していなかったため、該当データは存在しません。 市民活動の阻害に関するポリシーは、私たちのポリシーや執行の一部に過ぎません。レビューに基づいて、私たちのエージェントは最も重大な違反に基づいて行動を起こすことがあります。これにより、最高レベルの是正措置を適用することを保証します。この点で、報告された投稿には異なるポリシー違反が含まれていた可能性があります。

² すなわち、①削除率、②削除率で見た場合の世界の中での順位、③5つのポリシー（「スパム、欺瞞行為、詐欺に関するポリシー」、「誤った情報に関するポリシー」、「選挙の誤った情報に関するポリシー」、「医学的に誤った情報に関するポリシー」、「なりすましに関するポリシー」）別の削除数及び削除率、④「フラグ立て」の契機別の内訳、⑤「フラグ立て」された全ての動画の数に対する割合。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
ちらの 9 4,418 本という最新の数字をベースにするとどのようになるか、ご教示ください。 【TikTok Japan への質問】 2023 年 7 月から 9 月にかけて、日本国内で合計 1,111,969 本の動画を削除し、そのうち「誠実性と信頼性」のガイドライン違反により削除されたものが約 1 6,680 本のことですが、これらの数字に関し、以下の 3 点をご教示ください。なお、ご回答が難しい場合はその理由をご教示ください。 (1) 1,111,969 本という削除数に関し、①日本における 削除率（日本国内で同時期に存在した全ての動画に対する割合）、② 上記①の率について、全世界における 削除率との関係（全世界の平均値との関係や各国別の順位等）をご教示ください。 (2) 約 16,680 本という削除数のうち、AI 等の自動的なシステムにより検知がされたものと、外部からの通報により発見されたものは、それぞれ何本になるか、ご教示ください。 (3) 上記（2）の外部からの通報により発見されたものに関し、外部からの通報自体は全部で何件ほどあったのか、ご教示ください。 【X への質問】 2023 年の 1 年間で、日本国内で「誤解を招くアイデンティティや虚偽のアイデンティティに関するポリシー」違反により 4,829					

(設問)	Google	Meta	TikTok Japan	Microsoft	X
件、「プラットフォームの操作とスパムに関するポリシー」違反により 4,451,797 件の投稿を削除等されたということですが、これらの数字に関し、以下の 4 点をご教示ください。なお、ご回答が難しい場合はその理由をご教示ください。 (1) 4,829 件、4,451,797 件というそれぞれの削除等の数について、①日本国内における 削除等の率（日本国内で同時期に存在した全ての投稿に対する割合）、②上記①の率について、全世界における削除等の率との関係（全世界の平均値との関係や 各国別の順位等）をご教示ください。 (2) AI その他の機械的手段のみによって検知・対応したものは、① 4,829 件のうち 900 件及び ② 4,451,797 件のうち 2,973,405 件とのことです。残りのそれぞれ ①3,929 件及び ②1,478,392 件はいずれも外部からの通報を契機として検知・対応したものと理解してよろしいでしょうか。 (3) 上記各ポリシー違反を理由として投稿の削除等を求める外部からの通報の総数と、そのうち実際に問題の投稿を削除等した件数をご教示ください。 (4) 上記各ポリシーに対し、日本					

(設問)	Google	Meta	TikTok Japan	Microsoft	X
国内で「市民活動の阻害に関するポリシー」又は「合成または操作されたメディアに関するポリシー」への違反により削除等された投稿の数は、2022 年・2023 年ともに 0 件とのことですが、ポリシー間でこのように該当する数に大きな開きがある理由として考えられるものは何でしょうか。また、削除等された投稿について、該当するポリシー違反は 1 つのポリシーだけで整理シラベリングされているのか、それとも複数のポリシーで整理シラベリングされているのか。 特に、後者の場合、日本国内で「市民活動の阻害に関するポリシー」又は「合成または操作されたメディアに関するポリシー」への違反により削除等された投稿の数が 2022 年・2023 年ともに 0 件である点について、その理由をどのように説明するのでしょうか。					
Q 1－2 【Microsoft への質問】 LinkedIn では、2024 年 1 月中に偽・誤情報であるとして通報があったり削除されたりしたコンテンツは 1 件もなかったということですが、データとして「2024 年 1 月」に限ったものを開示いただいているのは、2023 年以前のデータは集計していないからと理解してよろしいでしょうか。また、ヒアリングの中でも質問させていただきまし				LinkedIn は、能登地震に関連する偽・誤情報を回答に含めることができるように、2024 年 1 月を期間として選択しました。ご要望がございましたら、他の期間に関する情報も提供したいと思います。 詐欺を意味すると解釈される「なりすまし」に関しましては、LinkedIn には偽プロフィールを検出して削除する堅牢なプログラムがあります。 Automated Fake Account Detection at LinkedIn. LinkedIn では、無料	

(設問)	Google	Meta	TikTok Japan	Microsoft	X
たが、いわゆる「なりすまし」に関しても日本国内では認知件数 0 件という理解でよいのか、念のため確認させてください。				でプロフィールの確認するサポートを提供しています。 Verifications on your LinkedIn profile LinkedIn Help.	
Q 1－3 【Microsoft への質問】 Bing 検索では、「ユーザーを権威の低いコンテンツに誘導する可能性が高いことが判明した自動サジェストおよび関連キーワード」を削除しているとのことですが、日本国内における直近の削除件数をご教示ください（ご回答が難しい場合はその理由をご教示ください）。				ご指摘のデータの取得にはリソースを要し、また現時点で必要なリソース不足のため直ちに回答することが困難です。	
Q 1－4 【Microsoft への質問】 Microsoft Start では「誤解を招く虚偽のコンテンツ」、「公衆または個人の安全、身体、精神、財務上の健全性を損なうコンテンツ」などを禁止しているとのことですが、これらに該当するとして①日本国内で削除等の対応がされた直近の件数と、②日本国内における削除等の率（日本国内で同時期に存在した全てのコンテンツに対する割合）をご教示ください（ご回答が難しい場合はその理由をご教示ください）。				Start について、2023 年は 日本市場のユーザーから合計 2,641,528 件のコメントが寄せられました。うち 147,559 件 (5.6%) は、コミュニティ ガイドライン違反に対するプロアクティブなモデレーションシステムによって事前にブロックされ、残る 2,493,969 件のコメント (94.4%) が Start コミュニティに投稿されました。特性の重複や複数のモデレーションプロセスが存在するため、事前にブロックされたコメントを特性によって分類することは困難です。 また、2023 年は 日本市場ではユーザーから 43,947 件のコメントがコミュニティ ガイドラインの違反として報告され、13,507 件 (31%) のコメントがレビュー後に削除されました。誤った情報カテゴリで削除されたコメントの数は 503 件です。コミュニティ ガイドラインに違反したとして、8 人のユーザーが Start/MSN コミュニティを退出する	

(設問)	Google	Meta	TikTok Japan	Microsoft	X
				こととなりました。	
Q 1－5 【Google への質問】 日本において YouTube で削除された動画の削除理由について、他国と比べて、日本に特徴的な傾向があればご教示ください。 【Meta への質問】 Facebook と Instagram それぞれで日本において削除又は抑制されたコンテンツの削除・抑制理由について、他国と比べて、日本に特徴的な傾向があればご教示ください。 【TikTok Japan への質問】 日本において 削除やおすすめフィード対象外とされた動画について、他国と比べて、その理由に関して日本に特徴的な傾向があればご教示ください。 【Microsoft への質問】 日本において 各サービスで 削除したり 検索結果を上下 させたりしたコンテンツ について、他国と比べて、その理由に関して 日本に特徴的な傾向があればご教示ください 【X への質問】 日本において 投稿が 削除 等 された 理由について、他国と 比べて、日本に特徴的な傾向があればご教示ください。	2023 年 10 月から 12 月にかけて、日本国内の IP アドレスからアップロードされた合計 94,418 本の動画を YouTube のコミュニティガイドラインに違反したことにより、YouTube から削除しました。その内上位 3 分野は、有害または危険なコンテンツ（6 万 4 千本以上）、ヌードと性的なコンテンツ（8,200 本）、子どもの安全に関するコンテンツ（7,900 本）です。 これに対し、世界全体では、同期間で削除された動画の上位 3 分野は、有害または危険なコンテンツ、子どもの安全に関するコンテンツ、そして暴力的で生々しいコンテンツです。	特定の傾向は関知していません。日本のファクトチェック機関に照会することをとお勧めします。	削除やおすすめフィード対象外とされる理由については、その時々により傾向は変化します。 （※上記より後の回答については非公開）	LinkedIn では、プロフェッショナル コミュニティ ポリシーに違反するコンテンツはグローバルに削除します。このため、検索結果を下げることはありません。	私たちは、ポリシーに違反する投稿をユーザーに削除してもらうよう依頼するか、投稿のリーチを制限するか、または違反の重大性に応じてアカウントを停止することがあります。 プラットフォームの使用方法が進化し続ける性質上、それに伴う文脈、および関連する害を考慮して、最新のポリシーガイダンスを装備したフロントラインのモデレーターを確保することに注意を払っています。追加のポリシーガイダンスの提供は、モデレーターが 1 つ以上のポリシーエリアで特別なトレーニングやリフレッシャーを必要とすると考えられる事象がきっかけでしばしば行われます。 そのようなリフレッシャーは、現在の出来事に基づくもの、または文化的小および社会的文脈に関連するものである可能性があります。 スパムを通じたプラットフォームの操作は進化する脅威であり、私たちが対処している傾向です。これは日本で発生していますが、それに限定されるものではありません。これは世界的な傾向であり、増加する数とそれに対する私たちの増加する行動に表れています。
Q 1－6 【Google への質問】 日本と外国との間で、問題のあるコンテンツの傾向は異なると思われますが（例えばヘイトが問題になり	Google では、偽情報・誤情報が世界的に問題になる中、日本における実態調査を進めることが、社会全体における適切な対策を検討する上で必要だと考えました。とりわけ日本で	特定の傾向は関知していません。日本のファクトチェック機関に照会することをとお勧めします。	（※回答非公開）	LinkedIn では、日本のコンテンツに関して特に懸念はありません。	当社のポリシーと執行ガイドラインは全世界で適用され、コンテンツのレビュー時にモデレーターが使用する資料を含んでいます。X のスケールドオペレーションチームは、多様

(設問)	Google	Meta	TikTok Japan	Microsoft	X
やすい国とそうでない国等)、日本について特に懸念すべき傾向のコンテンツがあればご教示ください(やはり災害関連が特徴的でしょうか)。 【Meta への質問】 日本と外国との間で、問題のあるコンテンツの傾向は異なると思われますが(例えばヘイトが問題になりやすい国とそうでない国等)、日本について特に懸念すべき傾向のコンテンツがあればご教示ください(やはり災害関連が特徴的でしょうか)。 【TikTok Japan への質問】 日本と外国との間で、問題のあるコンテンツの傾向は異なると思われますが(例えばヘイトが問題になりやすい国とそうでない国等)、日本について特に懸念すべき傾向のコンテンツがあればご教示ください(やはり災害関連が特徴的でしょうか)。 【Microsoft への質問】 日本と外国との間で、問題のあるコンテンツの傾向は異なると思われますが(例えばヘイトが問題になりやすい国とそうでない国等)、日本について特に懸念すべき傾向のコンテンツがあればご教示ください(やはり災害関連が特徴的でしょうか)。 【X への質問】 日本と外国との間で、問題のあるコンテンツの傾向は異なると思われますが(例えばヘイトが問題になりやすい国とそうでない国等)、日本について特に懸念すべき傾向のコンテンツがあればご教示ください	は、新型コロナウイルス感染症が拡大する中、新型コロナウイルスやコロナワクチンに関する誤情報が拡散され、調査の必要性が強まりました。そこで、Google が支援する国際大学 GLOCOM の共同研究プロジェクトでは、2019 年以降、4 年連続で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行っていました。これまでの研究結果については、こちらをご確認ください。				なスキル、経験、およびツールを持っており、すべてのルールとポリシーにわたる報告の効果的なレビューと対応を可能にしています。 コンテンツモデレーターが世界中で一貫してコンテンツをレビューし、ポリシーを適用できるように、関連する文脈とガイダンスを持つことが重要であると考えています。このため、投稿のレビューに必要な適切な労働技能と職務知識をコンテンツモデレーターに提供するための堅牢なトレーニングプログラムとシステムを設けています。これらの焦点領域は、コンテンツモデレーションのライフサイクル前後にコンテンツモデレーターが適切に対応するための準備を整えることを保証します。これには以下が含まれます： ・エージェントと学習ニーズに焦点を当てたトレーニングの分析/設計 ・専門トレーナーによる教室研修 ・新しいスキルを適用する準備期間 ・クロススキリングの機会 ・アップスキリングの機会 ・リフレッシュプログラム ・新規立ち上げ/更新のロールアウトプロセス ・改善計画 この質問で提供された例は必ずしも偽・誤情報に関連しているものではないと思いますが、その流れで言及しますと日本において特に対処しているカテゴリーは以下となります： 違法な商品や違法なサービス(薬物、売春など)、個人融資、闇金、闇バイ

(設問)	Google	Meta	TikTok Japan	Microsoft	X
(やはり災害関連が特徴的でしょうか)。					ト、口座買収などを促進する内容等。
Q 1ー7 【Meta への質問】 Facebook と Instagram それぞれで、次の各カテゴリーに該当すると して昨年1 年間(又はその他の任意 の期間)のうちに削除された日本語 のコンテンツの数及び削除率(日本 語のコンテンツの総数に対する割 合)をご教示ください。ご回答が難 しい場合は、その理由を具体的に (※) ご教示ください。 (1) 差し迫った身体的危害を及 ぼす危険性のある誤情報 (2) 投票行動抑制・投票に係る不正行為 (3) 選挙に係る協調的干渉 (4) 政治過程の機能の妨害を直接助長する可能性のあるコンテンツ (5) 非常に紛らわしい特定の加工されたメディア (6) 故意に人を欺く、故意に虚偽の説明をする、又は金銭や財産を搾取することを意図したコンテンツ ※ ヒアリングにて、「日本語のコンテンツであるというだけでは日本に関するコンテンツであるとは判断できない」といった趣旨のご意見を伺いましたが、一旦、「日本に関するコンテンツ」と言えるかどうかという論点は措いていただいて、とにかく単純に「日本語のコンテンツ」の数をご教示いただくことが技術的に可能なのか、できないとすれば何が理由なのかをご教示いただけますと幸いです。		デジタルコミュニケーションが国境にとられないグローバルプラットフォームです。それゆえ、弊社のコミュニティ規定施行の努力の有効性を測定するアプローチは、私たちがコミュニティ規定等をどれだけグローバルに施行しているかという疑問に答えることを目的としたグローバルな測定基準から始まります。これらのポリシーの施行に関する評価は、コミュニティ規定施行レポート(CSER)でご覧いただけます。 さらに、四半期ごとの Adersarial Threat report (敵対的脅威に関するレポート)では、組織的偽装行為ポリシーに違反したアカウント、ページ、グループのネットワークに関する国別の詳細と数値を共有しています。			

(設問)	Google	Meta	TikTok Japan	Microsoft	X
Q 1－8 【Google への質問】 偽・誤情報に関するポリシー違反の有無が問題になりやすいコンテンツについて、日本のトレンドを踏まえて重点的にチェックを行っている分野はございますでしょうか。例えば能登半島地震については、YouTube でチームを編成して対応されたとのことですが、同様の取組をされている分野があれば、ご教示ください。 【Meta への質問】 偽・誤情報に関するポリシー違反の有無が問題になりやすいコンテンツについて、日本のトレンドを踏まえて重点的にチェックを行っている分野はございますでしょうか。 【TikTok Japan への質問】 偽・誤情報に関するポリシー違反の有無が問題になりやすいコンテンツについて、日本のトレンドを踏まえて重点的にチェックを行っている分野はございますでしょうか。 【Microsoft への質問】 偽・誤情報に関するポリシー違反の有無が問題になりやすいコンテンツについて、日本のトレンドを踏まえて重点的にチェックを行っている分野はございますでしょうか。 【X への質問】 偽・誤情報に関するポリシー違反の有無が問題になりやすいコンテンツについて、日本のトレンドを踏まえて重点的にチェックを行っている分野はございますでしょうか。	上記をご参照ください。	弊社は、IFCN に認定された独立したファクトチェック機関と共に、日本でサードパーティファクトチェックプログラムを開始する予定です。ファクトチェック機関は、一次情報源へのインタビュー、公的データの参照、写真やビデオを含むメディアの分析など、独自の報道を通じて記事の正確性を評価し、評価を行います。ファクトチェックパートナーから評価を受けた後は、その誤情報の人々に見せないようにするために、私たちのプラットフォーム全体で当該記事にラベルを表示するといった対策を講じます。	コミュニティガイドライン違反になる可能性のあるコンテンツのトレンドや、ユーザーの行動は、常に変わり続けています。そのため、何か一つ決まったものがあるわけではなく、その時々においてリスクの高いコンテンツの傾向やユーザーの行動をできるだけ早く把握して、モデレーションに反映させるように取り組んでいます。	LinkedIn では、世界的なイベントについては、予定（選挙など）、予定外（地震など）を問わず、重点的にコンテンツ チェック活動を行っております。	新しいポリシーを作成するか、既存のポリシーを変更するには、オンライン行動のトレンドに関する徹底的な調査が必要です。これには、許可される内容に関する期待を明確に示す外部言語の開発と、何百万もの投稿に対応できるレビュアーのための施行ガイドラインの作成が含まれます。このような取り組みは、オンラインスピーチの変化する性質に対するグローバルな視点を考慮するために不可欠であり、異なる文化のおよび社会的文脈での私たちのルールの適用と解釈の仕方を理解することを含みます。
Q 1－9 【Google への質問】 日本のユーザーのオンライン上の	上記参照をご参照ください。	特定の見解は持ち合わせていません。	コミュニティガイドライン違反になる可能性のあるコンテンツのトレンドや、ユーザーの行動は、常に変わ	Start では、コメントに関するすべてのユーザーの報告（43,955 件）のうち、上位 3 つの特性は、ヘイト	先ほどの回答の通り、日本で特に対処しているカテゴリーは以下となります：違法な商品や違法なサービス

(設問)	Google	Meta	TikTok Japan	Microsoft	X
行動について、何かしら留意すべき特徴があればご教示下さい。 【Meta への質問】 日本のユーザーのオンライン上の行動について、何かしら留意すべき特徴があればご教示下さい。 【TikTok Japan への質問】 日本のユーザーのオンライン上の行動について、何かしら留意すべき特徴があればご教示下さい。 【Microsoft への質問】 日本のユーザーのオンライン上の行動について、何かしら留意すべき特徴があればご教示下さい。 【X への質問】 日本のユーザーのオンライン上の行動について、何かしら留意すべき特徴があればご教示下さい。			り続けています。そのため、何か一つ決まったものがあるわけではなく、その時々においてリスクの高いコンテンツの傾向やユーザーの行動をできるだけ早く把握して、モデレーションに反映させるように取り組んでいます。	スピーチ（13.8%）、冒涇（12.8%）、いじめと嫌がらせ（9.7%）です。Bing では、Web コンテンツのリンク付けに影響するシグナルと、検索結果でのコンテンツの表示方法に影響するアクションに関する公開情報を提供しております（「Bing での検索結果の提供方法」をご参照ください）。	（薬物、売春など）、個人融資、闇金、闇バイト、口座買収などを促進する内容等。
Q 1－10 【Google への質問】 Google 検索・YouTube それぞれに関し、個別のコンテンツについて、審査の結果、ポリシー違反の有無について疑義が生じた場合、どのようなプロセスで、削除／検索順位を下げる／今回は対応しない、という判断を行っているか、ご教示ください。ご回答が難しい場合はその理由をご教示ください。 【Meta への質問】 個別のコンテンツについて、審査の結果、ポリシー違反の有無について疑義が生じた場合、どのようなプロセスで、削除や抑制等の対応を行う／今回は対応しない、という判断を行っているか、ご教示ください。ご回答が難しい場合はその理由をご教示ください。	●情報アクセスを提供することは Google の中核的なミッションであり、私たちは誰もがインターネット上の情報を利用できるようにするために日々取り組んでいます。可能な限り最も関連性と信頼性の高い情報を返せるようにシステムを設計していますが、検索結果にはオープンウェブのコンテンツが含まれているため、ユーザーが検索する内容によっては、好ましくないコンテンツや不快なコンテンツが検索結果に含まれることがあります。 ● Google は情報へのオープンアクセスを約束していますが、それと同時に法令遵守とユーザー保護にも強い使命と責任感を持っています。現地法に違反しているコンテンツは、Google 検索でアクセスできないようにしています。	弊社は「誤情報」を単純に禁止するポリシーを持っているわけではありません。その代わりに、ポリシーは異なるカテゴリの誤情報を明確にし、それをどのように取り扱うかについての明確な指針を提供しようとしています。各カテゴリにおいて、弊社のアプローチは表現の自由、安全性の確保、人間としての尊厳、真実性（authenticity）、そしてプライバシーといった各価値の間でバランスを取る試みを反映しています。 差し迫った身体的危害のリスクに直接寄与する可能性がある誤情報は削除します。また、政治的プロセスの機能に直接干渉する可能性のあるコンテンツや、非常に欺瞞的な操作されたメディアも削除します。これらのカテゴリにおける誤情報を判断す	コミュニティガイドライン違反について疑義があるコンテンツについては、自動モデレーション技術と人間の両方で対応を行っています。違反の可能性が検出された場合には、自動モデレーションシステムは追加審査のために、人間の審査担当者、つまりモデレーターに情報を送信してさらに審査をする場合もありますし、自動モデレーション技術による審査の段階でコンテンツがTikTokのコミュニティガイドラインに違反しているという確証の度合いが高い場合には、自動的にコンテンツを削除する場合もあります。	LinkedIn では、プロフェッショナルコミュニティ ポリシーに違反するコンテンツはグローバルに削除し、検索結果を下げることはありません。 Microsoft Start では、コミュニティ ガイドラインに違反する広告やコメントは削除し、検索結果を下げることはありません。 Start では、コミュニティ ユーザーは Start のユーザー プロフィール ページに移動すると、コミュニティ ガイドライン違反で削除されたコメントを表示できます。Start では、現在のところメールによる通知はサポートしておりません。コミュニティ 投稿のユーザー レポートからコメントが削除された場合、コミュニ	私たちが執行行動を取る際、特定のコンテンツ（例えば個々の投稿やダイレクトメッセージ）またはアカウントに対してこれを行うことがあります。これらのオプションを組み合わせ使用することもあります。ほとんどの場合、これは行動がXのルールに違反しているためです。ポリシー違反がある場合、違反の罰を決定する際には、違反の重大性や個人の過去のルール違反の記録など、多くの要因を考慮しますが、これに限定されません。各ポリシーのページには、可能な執行行動に関する情報が含まれています。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
【TikTok Japan への質問】 個別のコンテンツについて、審査の結果、コミュニティガイドライン違反の有無について疑義が生じた場合、どのようなプロセスで、削除を行う／おすすりフィード対象外とする／今回は対応しない、という判断を行っているか、ご教示ください。ご回答が難しい場合はその理由をご教示ください。 【Microsoft への質問】 個別のコンテンツについて、審査の結果、ポリシー違反の有無について疑義が生じた場合、どのようなプロセスで、削除を行う／検索結果を上下させる／今回は対応しない、という判断を行っているか、ご教示ください。ご回答が難しい場合はその理由をご教示ください。 【X への質問】 個別のコンテンツについて、審査の結果、ポリシー違反の有無について疑義が生じた場合、どのようなプロセスで、削除を行う／アカウントを凍結する／投稿を一部非表示にする／今回は対応しない、という判断を行っているか、ご教示ください。ご回答が難しい場合はその理由をご教示ください。	● 全体として、情報の質とウェブページ削除に関しては、人々が必要な情報にアクセスできるようにする一方で、オンライン上の有害な情報から人々を守るために最善を尽くす、この2つのバランスを取ることを目指しています。その方法を簡単にまとめます。 ● 法律上の義務に基づき検索結果から特定のページを削除することに関して、Google は高い基準を設けています。検索結果からどのページを削除するかについては、現地法を主な指針としています。現地法には、オンライン上で発見できるべき情報とそうでない情報に関する当該地域社会の集合的な判断が反映されているからです。プライバシーの侵害や名誉毀損といった多くの問題は複雑で、そのような問題の取り扱い方は国や地域によって判断が異なるため、Google が負う法的義務も国によって異なる可能性があります。特に法的な判断が難しい場合は、自分たちで法律や事実を評価することはせず、裁判所の判断に従うことにしています。 ● 法律で義務付けられる通りにコンテンツを削除するだけでなく、特にオープンウェブに表示される個人のコンテンツに関しては、法律要件よりも厳しいポリシーを適用しています。たとえば、金融情報、医療情報、政府発行の ID、同意なしに公開された親密な画像などが、そのようなコンテンツに該当します。 ● このようなタイプのコンテンツは、一般的に非公開が望まれ、個人情報の盗難といった深刻な被害をも	る際、独立した専門家と提携して、コンテンツの真偽とそれが差し迫った危害のリスクに直接寄与する可能性があるかどうかを評価します。 他のすべての誤情報については、デマやバイラルな誤情報の拡散を減らし、利用者を権威ある情報に導くことに焦点を当てています。この取り組みの一環として、サードパーティファクトチェック機関と提携して、プラットフォーム上で最もバイラルなコンテンツの正確性を評価しています。弊社のファクトチェックプログラムの詳細は下記サイトをご覧ください。 https://transparency.fb.com/en-gb/features/how-fact-checking-works/ さらに、誤情報はしばしばヘイトスピーチ、いじめや嫌がらせなど、他の有害なコンテンツカテゴリにもまたがり、そのコンテンツがこれら他のポリシーに違反する場合には削除されることがあります。		ティ ューザーは異議を申し立てることができます。異議申し立ては、Start のコメント モデレーションチームによって審査され、そのコメントがコミュニティ ガイドラインに違反していないと判断された場合には、コメントが復活します。Start ューザーが削除されたコメントに対する異議申し立てを行えるページは、日本語で表示されます。	

(設問)	Google	Meta	TikTok Japan	Microsoft	X
	たらず可能性があるものです。そのため Google では、ユーザーが Google 検索結果からの削除を請求できるようにしています。 ● また Google では、オンライン上のユーザー保護をさらに強化するため、これらのポリシーを慎重に拡大するための新たな方法を模索しています。例えば、不当な削除方針が示されているサイトにある個人に関するコンテンツを Google から削除するポリシーや、個人情報をオンラインでより多くのユーザーに意図的に公開する晒し行為のコンテンツを Google 検索から削除するポリシーに基づき、ユーザーが自分の個人情報が掲載されているページの削除を要請できるようにしています。このようなケースでは、ユーザーが有益な情報を探したり、そのサイトのポリシーや運営方法を確認したりするために、そのようなサイトへのアクセスを望む可能性はありますが、そのページ自体に価値や公共の利益はほとんどなく、Google が防止を目指す風評被害や物理的な被害にもつながる可能性があります。 [YouTube] ● 私たちは、YouTube の検索結果、レコメンデーション、情報パネルにおいて、信頼できる情報源からの高品質な情報の表示順位を上げ、ユーザーが正確で有益な情報を見つけられるようにしています。これは特に、信頼性の高い情報源に人々を繋ぐことに関連します。現在、数百万件の YouTube の検索クエリにこの処理が施されており、対象となるトピックや国は継続的に拡大しています。過				

(設問)	Google	Meta	TikTok Japan	Microsoft	X
	去数年間、YouTube はランキング付けの際に信頼性を考慮するシステムや、質の高いコンテンツが YouTube で常に表示されるためのパートナーシップに多大な投資を行ってきました。 ● ポリシーに違反しないものの、それに近いコンテンツ（いわゆる「ボーダーラインコンテンツ」）は YouTube で視聴されているコンテンツのごく一部に過ぎません。YouTube では、機械学習を用いて、潜在的に有害な誤情報を含むこの種のコンテンツが視聴者の目に触れる機会を減らしています。 ● どのコンテンツが、ボーダーラインコンテンツに該当するかを判断することは困難です。そのため、動画の品質について重要な情報提供を世界各地にいる外部評価者に依頼しています。外部評価者は、一般に公開されている詳細な評価ガイドラインに従ってトレーニングを受けています。ボーダーラインコンテンツであると決定するには、評価者はコンテンツが不正確、誤解を招く、欺瞞的、無神経、不寛容、有害または害をもたらす可能性がある、などの要素について評価します。その結果を総合して、動画に有害な誤情報が含まれている可能性、またはボーダーラインである可能性をスコア化します。そして、そのコンテンツをチャンネル登録していないユーザーにすすめることを制限します。 ● 一方で、信頼性について判断するために、評価者はいくつかの重要な点を確認します。コンテンツの裏付けは取れているか、目的を達成				

(設問)	Google	Meta	TikTok Japan	Microsoft	X
	しているか、動画が目的を達成するにはどのような専門知識が必要か、動画に登場するスピーカーやそのチャンネルの評判はどうか、動画の主なトピックは何か（ニュース、スポーツ、歴史、科学など）、コンテンツは主に風刺を意図しているのか、これらの点を確認した上で動画の信頼性を判断します。ニュースや情報コンテンツの場合、スコアが高い動画ほどより宣伝されることとなります。				
Q 1－1 1 【Google への質問】 ポリシー違反のコンテンツの検知・対応に従事する人員のうち、日本語のコンテンツに対応可能な方の人数をご教示ください (EU において公用言語ごとに対応可能な人数を開示されているのと同じ基準で結構です)。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 【Meta への質問】 ポリシー違反のコンテンツの検知・対応に従事する人員のうち、日本語のコンテンツに対応可能な方の人数をご教示ください (EU において公用言語ごとに対応可能な人数を開示されているのと同じ基準で結構です)。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 【TikTok Japan への質問】 ポリシー違反のコンテンツの検知・対応に従事する人員のうち、日本語のコンテンツに対応可能な方の人数をご教示ください (EU において公用言語ごとに対応可能な人数を開	Google は、質の高いコンテンツを推奨し、報いると同時に、偽情報を含む潜在的に有害なコンテンツの削減・削除に全力を注いでおります。プラットフォームの安全性を優先することは、全従業員の責務です。また、YouTube を含む Google のプラットフォームの保護に専念する約 2 万人の従業員を擁しております。 Google は常にこの問題を真摯に受け止め、今後もユーザー保護のために投資を続けてまいります。 同時に、日本語で寄せられた問題の審査・承認は、日本語話者だけで行われるのではなく、時差のある複数の地域に拠点を設け、それぞれの専門知識を持つさまざまな部署のメンバーが集まったグローバルチームとなって、24 時間 365 日体制で対応しています。コンテンツの審査・措置にあたっては、さまざまな分野の専門知識が必要であり、国を超えて、担当者が知識や経験を持ち寄ること、適切かつ必要な対応を行うことができます。また、こうした体制をとることで、違反報告に対して迅速	弊社では、AI 検出技術、人の目によるレビュー、利用者からの報告の組み合わせを使用して、ポリシーに違反するコンテンツを特定しています。AI 検出技術が成熟し、賢くなるにつれて、人の目によるレビューや利用者からの報告に依存する割合は減少します。 プラットフォーム上のコンテンツの量と規模を考えると、主に AI を使用して違反コンテンツを特定し削除しているところ、その多くの場合、人の目・利用者がそれを見つける前に削除されます。人間のレビュアーは、主に、特に投稿が違反しているかどうかを技術では確信できない場合に、レビューすることが不可欠なケースで作業します。これらの人間のレビュアーは、多言語能力を含むさまざまなスキルと専門知識を持っていますが、すべてのコンテンツが言語専門知識を必要とするわけではありません（たとえば、ヌードや性的コンテンツは言語に依存しません）。したがって、特定の言語に対する人間のレビュアーの数は、弊社の能力	TikTok では透明性レポート（コミュニティガイドライン実施レポート）において、日本語を主要言語とするモデレーターの割合を公表しており、2023 年 10 月～12 月期においては、全世界のモデレーターのうち、2 %が日本語を主要言語としております。日本語を主要言語とする人数については、現時点では公表しておりません。	LinkedIn には、日本語を含む言語の専門知識を持つコンテンツ レビューアーのグローバル チームがあります。特定の言語の専門知識を持つレビュアーの数は、その言語の LinkedIn のコンテンツの量によって異なります。 マイクロソフトのサービス全体にわたって、特定の言語のコンテンツの量に応じて、その言語の専門知識を持つコンテンツ レビューアーのグローバルチームがあります。 （※上記より後の回答については非公開）	セキュリティおよびビジネス上の理由から、日本語での苦情を取り扱うために雇用されている人数は公開しておりません。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
示されているのと同じ基準で結構です)。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 【Microsoft への質問】 LinkedIn、Bing 検索、Microsoft Start それぞれにおいて、ポリシー違反のコンテンツの検知・対応に従事する人員のうち、日本語のコンテンツに対応可能な方の人数をご教示ください (EU において公用言語ごとに対応可能な人数を開示されているのと同じ基準で結構です)。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 【X への質問】 ポリシー違反のコンテンツの検知・対応に従事する人員のうち、日本語のコンテンツに対応可能な方の人数をご教示ください (EU において公用言語ごとに対応可能な人数を開示されているのと同じ基準で結構です)。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。	に対応することができます。そのため、日本語のコンテンツに関する問題に従事する担当者的人数をお伝えすることはできません。何卒ご理解賜りますよう、よろしくお願い申し上げます。	やポリシー施行の質を示す指標ではありません。			
Q 1－1 2 【Google への質問】 ポリシー違反のコンテンツの検知・対応に従事する人員について、地理的・言語的に日本固有と言える問題状況（例えば地政学的なリスクや日本固有の差別的表現の存在など）を理解している方を確保・育成することが重要と考えられるが、日本においてこのような人員を確保・育成することの課題、そして、この課題解決のために御社として行っている取組があれば、詳細ご教示くださ	上記をご参照ください。	弊社では、異なる視点、文化、信念を包含する、特に、見過ごされがちまたは疎外される可能性のある人々やコミュニティのためのポリシーを注意深く作成しています。 弊社のコミュニティ規定等、その他のポリシーの開発は、日本語を含む異なる専門知識、文化的小および言語的背景を持つ幅広い内部の人員によって支えられています。また、NGO、学者、その他の専門家、特に日本の組織などの外部の利害関係者とも協	TikTok では世界各国にコンテンツモデレーションチームが存在しており、<u>すでに日本にもコンテンツモデレーションチームがございます</u>。そのチームにおいて、日本固有と言える問題状況を理解し、日本語を主要言語とする審査担当者が審査を行っています。	Start では、信頼できるさまざまなファクトチェック組織の検証済み情報を利用し、分析することで、プラットフォームに投稿されたコメントに含まれる誤情報を評価し、対処しています。	X のスケールドオペレーションチームは、多様なスキル、経験、およびツールを持ち合わせており、全てのルールやポリシーに基づく報告に対して効果的なレビューと対応が可能です。X は、コンテンツモデレーターがレビューする報告の中でどの言語からの報告が多いかを精査し、それらの言語に熟達したコンテンツモデレーション専門家を採用しています。これには日本語も含まれます。チームメンバーは、ツール、プロセス、ルール、ポリシーに関する定期

(設問)	Google	Meta	TikTok Japan	Microsoft	X
い。また、このような人員を確保・育成するにあたって日本政府に期待することがあればご教示ください。そして、御社のコンテンツモデレーションの拠点について、日本においてどのような環境や条件があれば、日本国内に設置することが考えられるのかご教示ください。 【Meta への質問】 ポリシー違反のコンテンツの検知・対応に従事する人員について、地理的・言語的に日本固有と言える問題状況（例えば地政学的なリスクや日本固有の差別的表現の存在など）を理解している方を確保・育成することが重要と考えられるが、日本においてこのような人員を確保・育成することの課題、そして、この課題解決のために御社として行っている		力して、ポリシー開発に現地の文脈やニュアンスを反映させています。これらの内外の協力者には、ヘイトスピーチ、子供の安全、テロなどの問題に関する専門家、刑事訴追、レイプ危機カウンセリング、学者、人権、法律および教育の経験を持つ人々が含まれます。ポリシーの開発方法の詳細はこちらをご覧ください： https://transparency.meta.com/en-gb/policies/improving/deciding-to-change-standards/ 数十億のコンテンツが毎日投稿されるグローバルなプラットフォームとして、コンテンツモデレーションの実践は 24 時間 365 日スケールしてグローバルに運用できる必要があります。これが、多くのコンテンツが AI 検出技術によってモデレートされ、人間のレビュアーが特定のスキルと専門知識（日本の文化および言語の専門知識を含む）を持って AI 検出技術が投稿がポリシーに違反しているかどうかを判断できない場合に作業するための補完とされている理由です。また、ヌードや性的行為など、言語の専門知識が不要なコンテンツをレビューする言語に依存しないレビュアーもいます。これらの専任のレビュアーチームは、24 時間 365 日のカバレッジを提供するために、世界中のさまざまな場所に分散しており、80 以上の言語で対応しています。			的なトレーニングと再トレーニングを受けており、文化のおよび歴史的背景に関する特別セッションも含まれます。X のチームに参加する際には、各個人がオンボーディングプログラムを受け、この期間中およびその後も、外部従業員は品質保証プログラムを通じ、内部従業員は社内外の法律顧問を通じて個別のメンタリングを受けます。 チームメンバーは雇用期間全体を通じて、充実したトレーニングやワークフローのドキュメントに直接アクセスでき、いつでもトレーナーやリーダー、専門の法律・ポリシーチームおよびマネジメントから指導を受けることが可能です。 私たちは投資の評価と検討を継続しており、これに関するさらなる議論を歓迎します。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
取組があれば、詳細ご教示ください。また、このような人員を確保・育成するにあたって日本政府に期待することがあればご教示ください。そして、御社のコンテンツモデレーションの拠点について、日本においてどのような環境や条件があれば、日本国内に設置することが考えられるのかご教示ください。 【Microsoft への質問】 ポリシー違反のコンテンツの検知・対応に従事する人員について、地理的・言語的に日本固有と言える問題状況（例えば地政学的なリスクや日本固有の差別的表現の存在など）を理解している方を確保・育成することが重要と考えられるが、日本においてこのような人員を確保・育成することの課題、そして、この課題解決のために御社として行っている取組があれば、詳細ご教示ください。また、このような人員を確保・育成するにあたって日本政府に期待することがあればご教示ください。そして、御社のコンテンツモデレーションの拠点について、日本においてどのような環境や条件があれば、日本国内に設置することが考えられるのかご教示ください。 【X への質問】 ポリシー違反のコンテンツの検知・対応に従事する人員について、地理的・言語的に日本固有と言える問題状況（例えば地政学的なリスクや日本固有の差別的表現の存在など）を理解している方を確保・育成することが重要と考えられるが、日本においてこのような人員を確保・育成することの課題、そして、この課題解					

(設問)	Google	Meta	TikTok Japan	Microsoft	X
決のために御社として行っている取組があれば、詳細ご教示ください。また、このような人員を確保・育成するにあたって日本政府に期待することがあればご教示ください。そして、御社のコンテンツモデレーションの拠点について、日本においてどのような環境や条件があれば、日本国内に設置することが考えられるのかご教示ください。					
Q 1－1 3 【Google への質問】 YouTube の「選挙の誤った情報に関するポリシー」について、以下の 4 点をご教示ください。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 (1) 直近の日本の国政選挙に際しての違反の認知件数 (2) 上記 (1) のうち、外部からの通報による認知件数 (3) 上記 (1)・(2) のうち、実際に削除等の対応を実施した件数 (4) 上記 (3) に含まれる具体的な事例」 【TikTok Japan への質問】 選挙関連のコミュニティガイドラインも整備されていると伺いましたが、これに関し以下の 4 点をご教示ください。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 (1) 直近の日本の国政選挙に際しての違反の認知件数 (2) 上記 (1) のうち、外部からの通報による認知件数 (3) 上記 (1)・(2) のうち、実	[総合回答] 様々な理由から、通常、国レベルでの詳細なデータの開示は控えさせていただいております。そのような中、私たちはプラットフォームサービス研究会における貴省の取り組みに鑑み、その重要性を深く認識し、Google といたしましても可能な限りの貢献を自発的に行い、最も関連すると思われる一部の内容について日本独自のデータを作成し、ご提供させていただくことになりました。何卒ご理解賜りますよう、よろしくお願い申し上げます。		(1)～(4)につきましては、特に直近の国政選挙だけに限定した件数の把握などは行なっていないため、ご回答は難しい状況です。	前回の選挙から時間も経過しており、LinkedIn としてはこの詳細データをご提供することは困難です。 前回の選挙から時間も経過しており、マイクロソフトとしてもこの詳細データをご提供することは困難です。選挙に関するマイクロソフトのプロセスはこの数年で大きく進化しました。詳細は次のご質問をご覧ください。 Microsoft Start では、2021 年の日本の選挙に関する選挙固有のガイドラインとポリシーに基づく 日本市場におけるコンテンツのモデレーションは行っていませんでした。現在、Start では信頼できるさまざまなファクトチェック組織の検証済み情報を利用して、Start エコシステムに投稿されたコメントに含まれる情報を評価し、対処しています。ファクトチェック組織のコンテンツはグローバルであり、1 週間から 1 四半期の定期的な周期で更新されます。これらのファクトチェックされた誤情報のナラティブは、Start コミュニティのモデレーションで自動的に考慮されます。Start は、日本	2023 年に日本で市民活動の阻害に関するポリシーを施行していなかったため、該当データは存在しません。日本の主要な選挙イベント、たとえば 2024 年または 2025 年の国政選挙についてはポリシー運用を検討します。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
際に削除等の対応を実施した件数 (4) 上記(3)に含まれる具体的な事例 【Microsoft への質問】 各サービスにおける選挙関連のポリシーに関し、以下の4点をご教示ください。もしご回答が難しいということでしたら、その理由を具体的に教示ください。 (1) 直近の日本の国政選挙に際しての違反の認知件数 (2) 上記(1)のうち、外部からの通報による認知件数 (3) 上記(1)・(2)のうち、実際に削除等の対応を実施した件数 (4) 上記(3)に含まれる具体的な事例 【X への質問】 「市民活動の阻害に関するポリシー」について、以下の4点をご教示ください。もしご回答が難しいということでしたら、その理由を具体的に教示ください。 (1) 直近の日本の国政選挙に際しての違反の認知件数 (2) 上記(1)のうち、外部からの通報による認知件数 (3) 上記(1)・(2)のうち、実際に削除等の対応を実施した件数 (4) 上記(3)に含まれる具体的な事例				市場におけるユーザーの誤情報報告を引き続きサポートしていきます。 2021 年では、ユーザーからコミュニティ ガイドライン違反が報告された 33,328 件のコメントのうち、上位3つのカテゴリは、冒涇 (35.5%)、ヘイトスピーチ (15.48%)、陰謀論 (11.6%) でした。 Bing 検索では、選挙の完全性の保護に対して多面的なアプローチを取っており、リスク、トレンド、技術革新の進化に適応するために、そのプロセス、ポリシー、プラクティスを定期的に更新しています。このアプローチとして、防御的な検索介入、常にユーザーを権威ある高品質の情報源に誘導すること、ユーザーを権威の低いコンテンツに誘導する可能性が高いと考えられる自動サジェストおよび関連検索キーワードの削除、研究機関や非営利団体と提携して、脅威インテリジェンスを維持し、潜在的なアルゴリズム介入に関する情報を提供し、幅広い検索コミュニティに貢献すること、特別な情報パネルと回答による選挙や投票に関する権威ある情報源へのユーザーの誘導、全社的な選挙イニシアチブに専任で対応する社内ワーキンググループ、コンテンツの出所と選挙に関する社内調査の実施、選挙および政治コンテンツに関する生成 AI 機能の評価とレッドチームテストの実施、すべての AI 機能に対する責任ある AI レビューの実施 (Bing の Web 検索と Copilot の両方に対する対象を絞った監視の開発と継続的改善)、追加の対策、指標、緩和策の実	

(設問)	Google	Meta	TikTok Japan	Microsoft	X
				施の必要性についての定期的な評価などが挙げられます。	
Q 1－1 4 【Google への質問】 選挙期間中に選挙関連のポリシー違反のコンテンツや候補者等の権利を侵害するコンテンツの存在を外部から指摘された場合、削除等するか否かを判断するための検討期間として、フランスでは法律で2日以内が義務づけられていると理解しているところ、日本では、技術面や運用面等を踏まえると最短でどのくらいの日数が必要か、ご教示ください（ご回答が難しい場合はその理由をご教示ください）。 【Meta への質問】 選挙期間中に選挙関連のポリシー違反のコンテンツや候補者等の権利を侵害するコンテンツの存在を外部から指摘された場合、削除等するか否かを判断するための検討期間として、フランスでは法律で2日以内が義務づけられていると理解しているところ、日本では、技術面や運用面等を踏まえると最短でどのくらいの日数が必要か、ご教示ください（ご回答が難しい場合はその理由をご教示ください）。 【TikTok Japan への質問】 選挙期間中に選挙関連のポリシー違反のコンテンツや候補者等の権利を侵害するコンテンツの存在を外部から指摘された場合、削除等するか否かを判断するための検討期間として、フランスでは法律で2日以内が義務づけられていると理解しております。この点、日本では技術面や運用面等を踏まえると最短	個別のコンテンツの削除に要する手間は千差万別であり、その相当の部分はプラットフォームがコントロールしえない要素によっているため回答は困難です。例えば、コンテンツの媒体（文字媒体か動画か）、削除申請がされているコンテンツの時間的な長さ、削除申請における申請理由の記載の明確さ、判断に必要なコンテンツ外の情報の有無及び入手可能性、対象となる問題の性質が AI による処理に適しているか否か、などにより処理期間が変わり得ます。	デジタルプラットフォームは、サービス上の有害な行為やコンテンツに対処する際に、安全性と整合性の努力を最大化するためのリソースと投資を配分する柔軟性を持つと最も効果的です。特定の行動を取るための固定時間枠を課す要求は、安全でより安全なエコシステムを促進することよりも企業が規制遵守を優先してしまうことを促すことになり、意図しない悪影響をもたらす可能性があります。 たとえば、プラットフォームが違法コンテンツを削除するように利用者からの報告に応答することを要求するのは、非常に非現実的であり、運用上不可能であるだけでなく、プラットフォーム上の有害または問題のあるコンテンツを削減する問題を解決することもできません。利用者からの申立てには、プラットフォームがグローバルプロセスの下で処理およびレビューするために必要な情報や証拠が十分に含まれていないことが多くあります。デジタルプラットフォームは、裁判所や政府機関のような調査権限を持っていないため、そのような利用者からの報告を調査する能力は限られています。これらの状況では、ある一定の時間枠を設定することは無意味です。 さらに、すべての利用者からの報告が同等のリスクと害を引き起こすわけではありません。いくつかの報告は、即時の害がなく、表示回数が少	・2023 年 7 月～9 月期の TikTok 透明性レポートによると、選挙関連に限らない全てのコンテンツを対象とした数字ですが、通報から 2 時間以内に対応を完了したコンテンツの割合は 96.9%となっており、通報に対して非常に迅速な対応を行っております。 ・しかし、プラットフォームに投稿されるコンテンツは非常に多様であり、コミュニティガイドラインに違反するか否かの判断が非常に困難なコンテンツも多くあります。このような場合、TikTok では、ただ形式的に判断するのではなく、モデレーションチーム内でどのように審査すべきかを議論することもあります。 ・TikTok では、このように、それぞれの地域の文化や背景に配慮して、必要であれば社内で議論も行うといった、丁寧なモデレーション体制を構築しています。そのため、モデレーションの質を考慮せずただスピードだけ早ければ良いという観点での議論は、真摯に モデレーションに取り組んでいるプラットフォームにおける、丁寧なモデレーションの機会を奪うことになります。 ・また、ファクトチェック機関に照会など行う場合には、通常よりも時間がかかる可能性があります。 ・ <u>プラットフォームにより、どれだけ各地域の文化や背景を理解・反映したモデレーションを重視しているか、またモデレーションのための人</u> <u>員体制も、方針も、基準も、全く異</u>	LinkedIn の対応に要する日数は、世界の地域によっても問題の重大度によっても異なります。日本の選挙関連コンテンツのレビューは、通知を受け取ってから 2 営業日以内の実施にコミットしています。 2023 年 11 月、マイクロソフトは今後の日本の選挙にも適用される一連の選挙主要原則を発表いたしました。 ・ 有権者には、選挙に関する透明で権威ある情報を得る権利があります。 ・ 立候補者は、コンテンツが自らの活動に由来することを主張したり、選挙期間中に自らの画像やコンテンツが国民を欺く目的で AI によって歪められた場合に対処可能であるべきです。 ・ 政治活動はサイバー脅威から自らを守り、手頃な価格で簡単に導入できるツール、トレーニング、サポートを利用して AI をナビゲートできるようにする必要があります。 ・ 選挙当局は安全で強靱な選挙プロセスを確保し、このプロセスを可能にするツールやサービスにアクセスできるようにする必要があります。 また、立候補者が自分自身のディープフェイクに関する懸念をマイクロソフトに報告できる新しいウェブサイト「Microsoft-2024 Elections」	ポリシー違反の可能性に関する報告を受けた場合、できるだけ迅速に対応するよう努めています。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
でどのくらいの日数が必要かについて、ヒアリング当日にお尋ねしたところ回答できないとのことでしたが、検討期間をできるだけ短くするという観点から障害となっている要因や課題をご教示いただけますでしょうか（ご回答が難しい場合はその理由をご教示ください）。 【Microsoft への質問】 選挙期間中に選挙関連のポリシー違反のコンテンツや候補者等の権利を侵害するコンテンツの存在を外部から指摘された場合、削除等するか否かを判断するための検討期間として、フランスでは法律で2日以内が義務づけられていると理解しているところ、日本では、技術面や運用面等を踏まえると最短でどのくらいの日数が必要か、ご教示ください（ご回答が難しい場合はその理由をご教示ください）。 【X への質問】 選挙期間中に選挙関連のポリシー違反のコンテンツや候補者等の権利を侵害するコンテンツの存在を外部から指摘された場合、削除等するか否かを判断するための検討期間として、フランスでは法律で2日以内が義務づけられていると理解しているところ、日本では、技術面や運用面等を踏まえると最短でどのくらいの日数が必要か、ご教示ください（ご回答が難しい場合はその理由をご教示ください）。		ないかまったくない場合がありますが、他の投稿は急速に拡散して大きな害を引き起こす可能性があります。すべての報告を平等に扱い、固定期間内にブロックすることを要求することは、意図しない悪影響を引き起こすリスクがあります。たとえば： - 固定時間枠を守るためにレビューの精度を低下させることを余儀なくされ、過剰な執行や表現の自由を抑制する可能性がある。 - 利用者からの報告をすべて平等に扱い、個人に大きな害を引き起こす可能性のある報告を優先するのではなく、すべての報告を同等に扱うことを余儀なくされる。 迅速さと害のリスクおよび利用者の権利のバランスを取ることが重要です。	なっているという事実を踏まえて、スピードだけでなく質の高いモデレーションを実施するという観点からも、検討期間について考えていくことが、重要であると考えます。	を立ち上げる予定です。いわば、これは世界中の立候補者によるグローバルでのディープフェイク検出を支援するものです。 この取り組みに加えて、Digital Safety Unit を拡大して立ち上げ直します。これは、既存のデジタル セーフティ チームの活動を拡張するものです。当チームは、児童に影響を与えたり、過激派の暴力を助長したりする、虐待的なオンライン コンテンツや行為に長年取り組んできました。このチームは、大規模な銃乱射事件などの反社会的コンテンツを弊社サービスから速やかに削除するために、24 時間 365 日対応しています。 Microsoft Start では、日本語を専門とする人間のレビュアーに加え、AI / ML ソリューションも活用してモデレーションを行います。コメントのプロアクティブなモデレーションの平均待ち時間は 22 秒未満です。人間によるレビューの 95% は、概ね 75 分以内に完了しております。	
Q 1－1 5 【Google への質問】 いずれのサービスかを問わず、貴社のポリシーやガイドライン等の基準に則って削除やアカウント停	YouTube では、コミュニティに著しく有害な影響を及ぼす悪質な行為は許されません。クリエイターには、ポリシーに遵守し、YouTube の内外において視聴者、他のクリエイター、	ポリシーのコンテンツおよびアカウントに対する施行に加えて、弊社は広範な外部ステークホルダーと協力し、プラットフォームの整合性を保護するための他の行動も取っていま	例えば自殺予告の動画などを検知した場合に、警察などに弊社よりご連絡をすることがあります。件数等については公表しておりません。	LinkedIn は、問題の重大度に応じて世界の法執行機関に協議しています。 マイクロソフトは、日本を含む世界	第三者のサイトへのリンクを含む、児童性的搾取を描写または促進するコンテンツが報告された場合、それらは事前の通告なしに削除され、失踪・被搾取児童センター（NCMEC）に

(設問)	Google	Meta	TikTok Japan	Microsoft	X
止の対象となった投稿・アカウントについて、削除・アカウント停止以外の対応（例えば警察との協力、摘発、被害を受ける可能性のある人・組織等への連絡など）がなされたことはあるでしょうか。ある場合、件数及び主な事例をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 【Meta への質問】 いずれのサービスかを問わず、貴社のポリシーやガイドライン等の基準に則って削除やアカウント停止の対象となった投稿・アカウントについて、削除・アカウント停止以外の対応（例えば警察との協力、摘発、被害を受ける可能性のある人・組織等への連絡など）がなされたことはあるでしょうか。ある場合、件数及び主な事例をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 【TikTok Japan への質問】 貴社のポリシーやガイドライン等の基準に則って削除やアカウント停止の対象となった投稿・アカウントについて、削除・アカウント停止以外の対応（例えば警察との協力、摘発、被害を受ける可能性のある人・組織等への連絡など）がなされたことはあるでしょうか。ある場合、件数及び主な事例をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 【Microsoft への質問】 いずれのサービスかを問わず、貴社のポリシーやガイドライン等の基準に則って削除やアカウント停	広告主に敬意を払うことが求められます。 クリエイターがポリシーに違反した場合、収益化の一時停止やアカウントの停止といった措置が講じられる可能性があります。この措置は、クリエイターの既存の全チャンネル、新たに作成するチャンネル、およびクリエイターが定期的に登場するチャンネルにも適用される場合があります。 また、いずれかのチャンネルが停止されたり収益化が無効にされたりした場合、こうした制限を回避するために新しいチャンネルを作成する（あるいは既存のチャンネルを使用すること、停止期間中に関連するチャンネルで YouTube パートナープログラムに申し込むことは認められません。そのような行為は、全チャンネルの停止につながる可能性があります。クリエイターの責任について詳しくは、こちらをご確認ください。	す。これには、以下のことが含まれます。 - サービス利用規約に違反する者に対する法的対応。たとえば ： https://about.fb.com/news/2023/01/leading-the-fight-against-scraping-for-hire/ https://about.fb.com/news/2022/02/taking-legal-action-against-financial-services-scams/ https://about.fb.com/news/2021/12/taking-legal-action-against-phishing-attacks/ - オンラインで子供を標的とする犯罪者などの情報を業界の仲間や法執行機関との共有。 https://about.fb.com/news/2023/12/combating-online-predators/ https://about.meta.com/actions/safety/audiences/law - 利用者をサポートするために地元の安全およびその他の NGO と提携してのリソース開発。 https://about.fb.com/news/2024/02/helping-teens-avoidsextortionscams/		の法執行機関と協力しています。また、関係する省庁からコンテンツ削除検討依頼を受け取ることができるプロセスの確立も目指しています。	報告されます。 自傷行為または自殺の意図を表明したという報告を受けた場合、当該個人に直接連絡を取り、サポートを求めることを奨励し、専用のオンラインリソースやホットライン情報を提供します。状況に応じて、法執行機関と協力することもあります。たとえば、私たちの法執行機関ガイドラインで定義された有効な緊急開示要求を受けた場合などです。 法執行機関向けのガイドライン

(設問)	Google	Meta	TikTok Japan	Microsoft	X
止の対象となった投稿・アカウントについて、削除・アカウント停止以外の対応（例えば警察との協力、摘発、被害を受ける可能性のある人・組織等への連絡など）がなされたことはあるでしょうか。ある場合、件数及び主な事例をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 【X への質問】 貴社のポリシーの基準に則って削除やアカウント停止の対象となった投稿・アカウントについて、削除・アカウント停止以外の対応（例えば警察との協力、摘発、被害を受ける可能性のある人・組織等への連絡など）がなされたことはあるでしょうか。ある場合、件数及び主な事例をご教示ください（ご回答が難しい場合はその理由をご教示ください）。					
Q 1－1 6 【Google への質問】 外国における偽・誤情報のチェック（例えば、日本の災害について、海外で外国語の誤情報が拡散しているような場合は、自国内の偽・誤情報と比較して判断が難しい部分があるのではと推測します。何か工夫がありましたらご教示ください。 【Meta への質問】 外国における偽・誤情報のチェック（例えば、日本の災害について、海外で外国語の誤情報が拡散しているような場合は、自国内の偽・誤情報と比較して判断が難しい部分があるのではと推測します。何か工夫がありましたらご教	各国で見られた新たな問題やトレンドをグローバルチームに随時共有し、グローバルで行っているポリシー開発や対策に活かしていくことも多々あります。日本は世界でも早くから YouTube を楽しんでいた国であり、多くのユーザー、クリエイターの方がいらっしゃいます。そのため、日本で見られたトレンドがグローバルのオペレーション改善の参考になることは実際に多々あります。 例えば、YouTube チームは Google の Threat Analysis Group の担当者と密接に連携して、私たちのプラットフォームで組織的な影響力の行使が行われている可能性がある事例を	弊社はテクノロジー企業が真実の仲裁者であるべきではないと考えているため、独立したファクトチェック機関に誤情報のレビューと評価を依頼しています。グローバルに、90 以上の機関と 60 以上の言語で提携しています。日本では、IFCN に認定された現地の機関と共に、サードパーティファクトチェックプログラムを近日中に開始する予定です。この機関は、日本語のコンテンツをレビューし評価することとなります。	TikTok では、日本を含む世界各国にモデレーションチームが存在しており、常にモデレーションを担当するチーム内において情報共有を行ない、他国のモデレーターとも意見交換ができる環境を構築する工夫をしています。	マイクロソフト（Bing および Start）、LinkedIn 共に、偽・誤情報が広がっている地域に関係なく、グローバルベースで活動しております。 Microsoft Start では、信頼できるさまざまなファクトチェック組織の検証済み情報を利用して、同プラットフォームに投稿されたコメントに含まれる誤情報を評価し、対処しております。さまざまなファクトチェック組織の誤情報コンテンツ（誤情報のカテゴリ/特性など）はグローバルであり、定期的に更新され、Start コミュニティのモデレーションで自動的に考慮されます。JA-JP 市場におけるユーザー報告も引き続き	私たちのポリシーに違反しないコンテンツについては、日本を含むユーザーにコミュニティノートを提供しています- https://x.com/CommunityNotes/status/1676767470961532928 コミュニティノートはメディアにも利用可能です- https://x.com/CommunityNotes/status/1663609484051111936 2023 年 4 月 4 日現在、69 カ国でコミュニティノートが利用可能です。この投稿をご覧ください。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
示ください。 【TikTok Japan への質問】 外国における偽・誤情報のチェック（例えば、日本の災害について、海外で外国語の誤情報が拡散しているような場合）は、自国内の偽・誤情報と比較して判断が難しい部分があるのではと推測します。何か工夫がありましたらご教示ください。 【Microsoft への質問】 外国における偽・誤情報のチェック（例えば、日本の災害について、海外で外国語の誤情報が拡散しているような場合）は、自国内の偽・誤情報と比較して判断が難しい部分があるのではと推測します。何か工夫がありましたらご教示ください。 【X への質問】 外国における偽・誤情報のチェック（例えば、日本の災害について、海外で外国語の誤情報が拡散しているような場合）は、自国内の偽・誤情報と比較して判断が難しい部分があるのではと推測します。何か工夫がありましたらご教示ください。	特定し調査しています。悪意のある人物を特定してそのチャンネルとアカウントを停止し、他のテクノロジー企業と協力してインテリジェンスとベストプラクティスを共有し、脅威に関する情報を法執行機関に報告します。また、こうした調査の結果として、YouTube で行った対応に関する情報を四半期ごとの TAG Bulletin で提供しています。TAG Bulletin は 2020 年から公開されています。			きサポートしていきます。	
2. ニュースフィード について					
Q 2 - 1 【Microsoft への質問】 Microsoft Start のニュースフィードのパーソナライゼーションについて、ウォッチリストを用いたパーソナライゼーションのためのアルゴリズムはユーザー側で選択できないとヒアリング当日に伺いました。それでは、どのようなアルゴリズムでパーソナライズされて				Microsoft Start では、フィードのパーソナライズやトピック選択のほか、ユーザーが 日本市場のパブリッシャーをフォロー/フォロー解除することができます（以下をご参照ください）。	

(設問)	Google	Meta	TikTok Japan	Microsoft	X
いるのかについて、ユーザーが知ることはできますでしょうか。					
Q 2－2 【Microsoft への質問】 2022 年・2023 年に日本において、ポリシー違反を理由としてアカウントの停止や投稿の削除等をした件数のうち、次の 5 点をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 （１）日本の政府や自治体、放送事業者や新聞社のアカウントからの投稿に対する削除等の件数 （２）日本の政府や自治体、放送事業者や新聞社のアカウントに対して行った凍結・停止措置の件数 （３）（２）の措置に至った理由の例 （４）（２）のうち、事後的に凍結・停止措置を解除した件数 （５）（４）の解除措置に至った理由の例				Microsoft Start では、一般ユーザーの Start コミュニティ アカウント（コメント アカウント）を、政府やパブリッシャーのアカウントと区別しておりません。そのため、この情報を上記の方法で分類することは困難です。	
３． 特定の通報者からの通報への優先対応について					
Q 3－1 【Google への質問】 ヒアリングシート Q5-2 で言及いただいた優先報告者プログラムについて、DSA の Trusted Flagger の仕組みとの関係（DSA の策定・施行を見据えて導入されたものなのか、それとも完全に日本において独自に導入されたものなのか）をご教示ください。 【Meta への質問】 ヒアリングシート Q5-2 で第三者ファクトチェックパートナーを紹介	YouTube の優先報告者プログラムは、YouTube のコミュニティガイドラインに違反する可能性のあるコンテンツを YouTube に報告するために、コンテンツポリシー分野で特定の専門知識を持つ政府機関や非政府組織（NGO）に強力なツールを提供する YouTube 独自のグローバルで展開されているプログラムです。 YouTube のコミュニティガイドラインに違反するコンテンツについて YouTube に伝えるのに特に効果的な NGO や政府機関は、お住まいの地域	弊社の技術は、さまざまなシグナルに基づいて誤情報である可能性のある投稿を検出できます。これには、人々がどのように反応し、コンテンツがどのくらいの速さで拡散しているかが含まれます。また、Facebook、Instagram、および Threads の利用者がコンテンツを「誤情報」としてフラグを立てたり、投稿に対するコメントで不信感を表明したりするかどうかも考慮します。ファクトチェック機関は独自にレビューするコンテンツも特定し	ヒアリングシート Q5-2 でご紹介した取り組みは、DSA 以前から実施している取り組みであり、導入経緯としては DSA というよりも日本の文化を尊重した削除対応をすることを目的とした取り組みです。		このフォームは、法執行機関／政府関係者を対象としており、以下をカバーすることを目的としています： -アカウント情報の要求 -データ保持の要求 -緊急開示の要求 -コンテンツ削除の要求

(設問)	Google	Meta	TikTok Japan	Microsoft	X
いただいておりますが、第三者ファクトチェックパートナーからの通報は他の通報に比べて優先的に取り扱われるのでしょうか。その場合、日本国内で具体的にどの第三者についてどのような優先的取扱いをしているのか（報告後〇日以内に審査・フィードバックを行う、など）をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 【TikTok Japan への質問】 ヒアリングシート Q5-2 で言及いただいた、「信頼できる第三者による通報を優先的に取り扱う仕組み」について、DSA の Trusted Flagger の仕組みとの関係（DSA の策定・施行を見据えて導入されたものなのか、それとも完全に日本において独自に導入されたものなのか）をご教示ください。 【X への質問】 ヒアリングシート Q5-2 で、通報窓口に関し「日本の政府や法執行機関専用のフォームを用意しています」と回答いただいておりますが、これは偽・誤情報であることを理由とする通報を受け付けるフォームということでしょうか。その場合、具体的にどのような機関からの通報が多いか、概ねの件数を含めてご教示ください（ご回答が難しい場合はその理由をご教示ください）。	の YouTube 連絡窓口を通じて直接やりとりを行い、特定のポリシー分野のトレーニングを受けます。 これは、DSA の Trusted Flagger プログラムとは異なる設定です。DSA の Trusted Flagger プログラムは、EU で違法とされるコンテンツを報告するためのもので、Trusted Flaggers は、EU 加盟国の Digital Services Coordinators (DSC) によって決定されています。	ます。 誤情報であると予測されるコンテンツは、一時的にフィード内で低く表示されることがあります。詳細はコンテンツ配信ガイドラインをご覧ください： https://transparency.meta.com/en-gb/features/approach-to-ranking/types-of-content-we-demote/			
Q 3 - 2 【X への質問】 ファクトチェック機関など、「日本の政府や法執行機関」以外の機関は通報できないのでしょうか、上					政府機関以外からのリクエストについては、X ルールおよび利用規約に違反している可能性のある事例の報告方法に関するウェブページを参照してください。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
記「フォーム」を通じて通報を行う主体となるための資格や登録基準・条件をご教示ください（ご回答が難しい場合はその理由をご教示ください）。					https://help.twitter.com/en/rules-and-policies/x-report-violation コミュニティノートについては、資格基準を満たす X のアカウントを持つ誰でも登録して協力することができます。 https://help.twitter.com/en/using-x/community-notes
Q 3－3 【Google への質問】 優先報告者プログラムを通じた通報について、日本国内で具体的にどの第三者についてどのような優先的取扱いをしているのか（報告後〇日以内に審査・フィードバックを行う、など）をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 【Meta への質問】 仮に第三者ファクトチェックパートナーからの通報を優先的に取り扱う場合、その仕組みについて、DSA の Trusted Flagger の仕組みとの関係（DSA の策定・施行を見据えて導入されたものなのか、それとも完全に日本において独自に導入されたものなのか）をご教示ください。 【TikTok Japan への質問】 上記仕組みを通じた通報について、日本国内で具体的にどの第三者についてどのような優先的取扱いをしているのか（報告後〇日以内に審査・フィードバックを行う、など）をご教示ください（ご回答が難しい場合はその理由をご教示ください）。	Google の製品でポリシー違反のコンテンツを見つけたユーザーから報告をいただけることはとてもありがたいことです。例えば YouTube では、YouTube 優先報告者プログラム を設け、コミュニティガイドラインに違反しているコンテンツの報告に関してその分野の知識が多い個人、政府機関、非政府組織（NGO）にツールを提供しています。日本では自殺予防に取り組む日本の非営利団体 OVA や、 法務省人権擁護局 と連携し、優先報告者として、プラットフォーム上の問題のあるコンテンツをフラグしていただいております。優先報告者から報告された動画は、YouTube のコミュニティガイドラインに基づいて審査されます。優先報告者から報告されたからといって、自動的に削除されたり、異なるポリシーによって対処されたりするわけではなく、他のユーザーから受け取った報告と同じ基準が適用されます。ただし、その高い正確性により、YouTube 優先報告者からの報告は優先的に審査されます。	弊社は、DSA 導入前からファクトチェックおよびトラステッド・パートナーと協力しています。トラステッド・パートナープログラムは、113 カ国以上からの 400 以上の非政府組織、人道機関、人権擁護者および研究者のネットワークであり、ソーシャルメディアの監視能力を強化し、ポリシー、施行プロセスおよび製品を改善するために積極的に協力しています。コンテンツを報告するだけでなく、トラステッド・パートナーは、弊社のコンテンツポリシーおよび施行に関する現地の専門的な文脈とフィードバックを提供し、利用者の安全を確保する努力を支援します。	上記仕組みを通じて通報いただいた場合には、直接それに対応するチームが社内であり、迅速かつきめ細やかに対応できる（担当チームと通報者が直接に意見交換ができる）という特徴があります。		上記フォームを通じたリクエストは法執行機関／政府関係者のみが提出できます。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
【X への質問】 上記「フォーム」を通じた通報について、日本国内で具体的にどの第三者に対してどのような優先的取扱いをしているのか（報告後〇日以内に審査・フィードバックを行う、など）をご教示ください（ご回答が難しい場合はその理由をご教示ください）。					
Q 3－4 【Google への質問】 日本国内において優先報告者プログラムを通じて通報を行う主体となるための資格や登録基準・条件をご教示ください（ご回答が難しい場合はその理由をご教示ください）。また、当該基準・条件として、特に偽・誤情報対策の観点から設けているもの（政府機関であること／政府機関から独立していること、ファクトチェックに関する知見・経験など）がもしあれば、詳細ご教示ください。 【Meta への質問】 日本国内において第三者ファクトチェックパートナーとなるための資格や登録基準・条件をご教示ください（ご回答が難しい場合はその理由をご教示ください）。また、当該基準・条件として、特に偽・誤情報対策の観点から設けているもの（政府機関であること／政府機関から独立していること、ファクトチェックに関する知見・経験など）がもしあれば、詳細ご教示ください。 【TikTok Japan への質問】 日本国内において上記仕組みを通じて通報を行う主体となるための	上記をご参照ください。	ファクトチェック機関には、IFCN によって認証されることを求めています。これにより、IFCN の原則に従う必要があります。詳細はこちらをご覧ください。 https://ifcncodeofprinciples.poynter.org/	（※回答非公開）		このフォームは法執行機関および政府関係者のみが利用可能であり、受け取ったリクエストは随時審査しています。 また、児童性的搾取のケースは優先的に扱います。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
資格や登録基準・条件をご教示ください（ご回答が難しい場合はその理由をご教示ください）。また、当該基準・条件として、特に偽・誤情報対策の観点から設けているもの（政府機関であること／政府機関から独立していること、ファクトチェックに関する知見・経験など）がもしあれば、詳細ご教示ください。 【X への質問】 日本国内で上記「フォーム」を通じて通報を行う主体となるための資格や登録基準・条件として、特に偽・誤情報対策の観点から設けているもの（政府機関の中でも特にファクトチェックに関する知見・経験があることなど）がもしあれば、詳細ご教示ください。					
Q 3－5 【X への質問】 上記「フォーム」について、DSA の Trusted Flagger の仕組みとの関係（DSA の策定・施行を見据えて導入されたものなのか、それとも完全に日本において独自に導入されたものなのか）をご教示ください。					X の専用法的提出フォームは、政府および法執行機関向けのものであります。これらは日本国内および国際的に使用されています。
Q 3－6 【Google への質問】 特に偽・誤情報の流通・拡散の観点で、日本における優先報告者プログラムの導入による日本国内における効果として把握されているものがあればご教示ください。 【Meta への質問】 特に偽・誤情報の流通・拡散の観点で、日本における上記仕組みの導入による日本国内における効果として把握されているものがあればご教示ください。	YouTube の優先報告者は、深い知識を各分野にわたって持つパートナーであり、YouTube のシステムを補完する重要な存在であると考えています。優先報告者プログラムに参加した政府機関や NGO は、YouTube のチームと直接連絡を取り、報告されたコンテンツの審査結果や、報告された動画の優先的な審査を表示します。また、YouTube コンテンツ領域に関する継続的なディスカッションとフィードバックにも参加します。	現在共有できるデータはありません。研究者が研究 API を使用して興味のある分析を行うことを歓迎します。前述のように、IFCN に認証された独立したファクトチェック機関と共に、日本でサードパーティファクトチェックプログラムを開始する予定です。	特に偽・誤情報に限ったものではありませんが、これまで信頼できる第三者として通報いただいた事案に関しては、全件について適切な審査を行い、削除等の対応をしています。また、効果としては日本固有の状況について特に審査における判断が困難なケースについても、信頼できる第三者の知見を反映した審査を実施できるという効果があります。		このフォームは、法執行機関／政府関係者を対象としており、以下をカバーすることを目的としています： -アカウント情報の要求 -データ保持の要求 -緊急開示の要求 -コンテンツ削除の要求

(設問)	Google	Meta	TikTok Japan	Microsoft	X
ばご教示ください。 【TikTok Japan への質問】 特に偽・誤情報の流通・拡散の観点で、日本における上記仕組みの導入による日本国内における効果として把握されているものがあればご教示ください。 【X への質問】 特に偽・誤情報の流通・拡散の観点で、日本における上記「フォーム」の導入による効果として把握されているものがあればご教示ください。					
4. 信頼できる情報のプロミネンスについて					
Q 4－1 【Google への質問】 英国でメディアのコンテンツの表示の方法に関するプロミネンスと言われる法整備がされており、日本でも関心が上がってきております。特に災害発生時等の非常事態において、地方自治体等の公的機関、新聞や放送等の伝統メディアのコンテンツやファクトチェック記事、その他信頼性の高い発信源が発信する情報を優先的に表示するといった取組をもしされていれば、詳細ご教示ください。また、今後そのような取組を行うことの可能性について、お考えをお聞かせください。 【Meta への質問】 英国でメディアのコンテンツの表示の方法に関するプロミネンスと言われる法整備がされており、日本でも関心が上がってきております。特に災害発生時等の非常事態において、地方自治体等の公的機関、新聞や放送等の伝統メディア	YouTube では、YouTube の検索結果、おすすめ動画、情報パネルにおいて、信頼できる情報源からの高品質な情報の表示順位を上げ、ユーザーが正確で有益な情報を見つけられるようにします。その一環として、例えば、偽・誤情報が生まれやすい話題やニュースに関連するコンテンツについて説明する情報パネルや、ニュースメディアが発信する情報を提供しています。さらに、YouTube のニュース速報セクションや最新ニュースセクションなど、信頼できるニュースの情報源となる関連動画を紹介する専用のサービス機能も用意しています。 また、YouTube のポリシーに違反してはいないものの、有害な方法でユーザーを誤解させる可能性のあるコンテンツの拡散を制限しています。これには、陰謀論の動画や、許容範囲の限界事例となるようなコンテンツが含まれることもあります。このようなコンテンツは YouTube に残	まず、弊社はパブリッシャーではありません。 さはさりながら、利用者が正確で権威ある情報にアクセスできるようにするために、選挙、新型コロナウイルス感染症、気候変動科学などのさまざまなトピックに関する正確で信頼できる情報と利用者に届けるためのプロダクト開発をしています。			

(設問)	Google	Meta	TikTok Japan	Microsoft	X
のコンテンツやファクトチェック記事、その他信頼性の高い発信源が発信する情報を優先的に表示するといった取組をもしされていれば、詳細ご教示ください。また、今後そのような取組を行うことの可能性について、お考えをお聞かせください。	るとしても、YouTube がユーザーに推奨したり、検索結果の上位に表示することはありません。 そして、YouTube は収益化のより高い基準を満たすクリエイターに還元します。YouTube は質の低いコンテンツの収益化を防ぎ、視聴者に常に信頼性の高い高品質のコンテンツを提供するクリエイターに収益化の機会を提供します。				
Q 4－2 【Google への質問】 ヒアリングシート Q7-1 への回答に関連して、災害時や選挙期間においては、偽・誤情報に対するコンテンツモデレーション（例：削除）だけでなく、信頼できる情報を ³ 目立つ位置に表示する措置（いわゆる「プロミネンス」）が重要になると思います。YouTube では、ニュース速報セクションが表示される場合がありますが、ここで表示されるニュースを供給している「信頼できるニュース提供元」 ³ は、どのような基準で選ばれているのでしょうか。	信頼できる質の高いニュース提供元を識別して ニュース機能 で取り上げるために、YouTube のシステムはチャンネルの品質や、最近のニュースや関連事件の報道範囲など、さまざまなシグナルを使用しています。また、チャンネルは Google 検索結果の機能に関するポリシー と Google ニュースのコンテンツポリシー を遵守している必要があります。 Google は機械学習技術を使用して、ニュース機能のニュース提供元を特定するためのシグナルを生成しています。 サードパーティの評価担当者 からのフィードバックにより、これらのシグナルを継続的に強化、評価、改善しています。				
Q 4－3 【TikTok Japan への質問】 偽の救助要請等は容易に見抜けるものではないため対応が難しいという点はよく理解しています。ただし、災害時において、地方自治体等の公的機関や新聞や放送等の信頼できる提供元からの情報を目			令和6年能登半島地震に際しては、以下のような取り組みを行いました。 ・モデレーションの強化 ・令和6年能登半島地震のインシデントは高リスクとして扱われ、通常 のモデレーションに加えて特別な対策が講じられました。		信頼できる情報源からの情報を強調するような協力体制は敷かれておりません。しかしプレゼンテーションでご説明しましたが、能登半島地震後に Yahoo/日本財団等に提供したような緊急告知をサポートするために無償広告枠の活用等は今後も検討可能と考えております。

³ <https://support.google.com/youtube/answer/9057101?sjid=5553189098828327477-AP>。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
立たせることが重要であり、偽情報に対応するためには、災害現場の地方自治体等の信頼できる提供元との情報共有・連携が必要不可欠だと考えています。令和6年能登半島地震に際して、そのような取組は行われたのでしょうか。また今後、制度が整備されれば、そのような取組を行いやすくなるのでしょうか。 【Xへの質問】 偽の救助要請等は容易に見抜けるものではないため対応が難しいという点はよく理解しています。ただし、災害時において、地方自治体等の公的機関や新聞や放送等の信頼できる提供元からの情報を目立たせることが重要であり、偽情報に対応するためには、災害現場の地方自治体等の信頼できる提供元との情報共有・連携が必要不可欠だと考えています。令和6年能登半島地震に際して、そのような取組は行われたのでしょうか。また今後、制度が整備されれば、そのような取組を行いやすくなるのでしょうか。			・令和6年能登半島地震のインシデントでは、高リスクなインシデントを担当するチームが、偽・誤情報を含むコミュニティガイドライン違反のコンテンツの一斉確認を実施し、また危険なキーワード/ハッシュタグのモニタリングを実施しました。 ・(※このポツは回答非公開) ・理事を務める業界団体の SMAJ において他の事業者と連携して、SMAJ の X アカウントおよび HP において、実際の被害と異なる救助要請といった虚偽情報の拡散や、送金を促す詐欺行為について、注意喚起を発信しました。 また、制度の整備につきましては、制度の内容が分かりかねますため現時点では回答を控えさせていただきますが、プラットフォームの仕組みはサービスにより大きく異なるため、ある取り組みが一つのプラットフォームには有効でも他のプラットフォームでは機能上実装が不可能であるといったことも考えられます。そのため、今後日本政府にて制度の検討をされる場合には、事業者の意見を広く反映して議論を進めることが必要であると考えます。		
Q 4 - 4 【Google への質問】 2022 年・2023 年に日本において、ポリシー違反を理由としてアカウントの停止や投稿の削除等をした件数のうち、次の 5 点をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 （１）日本の政府や自治体、放送事業者や新聞社のアカウントからの投稿に対する削除等	YouTube は、コミュニティガイドラインに基づく削除件数について、ご質問でいただいたような属性に基づく管理は行なっておらずデータを提出することができません。何卒ご理解賜りますよう、よろしくお願い申し上げます。	リクエストされた統計は持っていません。しかし、企業がプラットフォームの特定のポリシーやリスクに最も関連し適切なメトリクスを決定する柔軟性を提供することが重要です。この質問でリクエストされたメトリクスは、プラットフォームにとってあまりにも詳細であり、かつ特定の項目に寄り過ぎています。	特に日本の政府や自治体、放送事業者や新聞社のアカウントだけに限定した削除等の件数の把握などは行っていないため、ご回答は難しい状況です。		X による大規模なセーフティの取り組みにおいては、時に偶発的な誤検知が発生することがあります。誤って対応された投稿やアカウントがあると把握した場合、できるだけ迅速に修正するよう努めています。 なお、X のルールと利用規約はすべてのアカウントに適用されます。またプライバシーの理由から、個々のアカウントに関する点についてはコメント致しません。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
の件数 (2) 日本の政府や自治体、放送事業者や新聞社のアカウントに対して行った凍結・停止措置の件数 (3)(2)の措置に至った理由の例 (4)(2)のうち、事後的に凍結・停止措置を解除した件数 (5)(4)の解除措置に至った理由の例 【Meta への質問】 2022 年・2023 年に日本において、ポリシー違反を理由としてアカウントの停止や投稿の削除等をした件数のうち、次の 5 点をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 (1) 日本の政府や自治体、放送事業者や新聞社のアカウントからの投稿に対する削除等の件数 (2) 日本の政府や自治体、放送事業者や新聞社のアカウントに対して行った凍結・停止措置の件数 (3)(2)の措置に至った理由の例 (4)(2)のうち、事後的に凍結・停止措置を解除した件数 (5)(4)の解除措置に至った理由の例 【TikTok Japan への質問】 2022 年・2023 年に日本において、ポリシー違反を理由としてアカウントの停止や投稿の削除等をした件数のうち、次の 5 点をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 (1) 日本の政府や自治体、放送事業者や新聞社のアカウントか					

(設問)	Google	Meta	TikTok Japan	Microsoft	X
らの投稿に対する削除等の件数 (2) 日本の政府や自治体、放送事業者や新聞社のアカウントに対して行った凍結・停止措置の件数 (3) (2) の措置に至った理由の例 (4) (2) のうち、事後的に凍結・停止措置を解除した件数 (5) (4) の解除措置に至った理由の例 【X への質問】 2022 年・2023 年に日本において、各ポリシー違反を理由としてアカウントの停止や投稿の削除等をした件数のうち、次の 5 点をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 (1) 日本の政府や自治体、放送事業者や新聞社のアカウントからの投稿に対する削除等の件数 (2) 日本の政府や自治体、放送事業者や新聞社のアカウントに対して行った凍結・停止措置の件数 (3) (2) の措置に至った理由の例 (4) (2) のうち、事後的に凍結・停止措置を解除した件数 (5) (4) の解除措置に至った理由の例					
5. 発信者への収益還元について					
Q 5 - 1 【Google への質問】 YouTube パートナープログラム YPP を通じた収益化には コミュニティガイドラインや広告掲載に適	[総合回答] ご質問で頂いたようなデータは公開しておりません。何卒ご理解賜りますようお願い申し上げます。なお YouTube は収益化のより高い基準を	対象となるアカウントは、Facebook および Instagram で収益化が可能です。これらのアカウントは、まず招待を受け、一連の基準を満たす必要があります :	(1) 前段について ・ Creator Rewards Program の参加対象となるクリエイターは 18 歳以上で、フォロワー数 1 万人以上と動画再生回数 10 万以上を満たし、コ		

(設問)	Google	Meta	TikTok Japan	Microsoft	X
したコンテンツのガイドラインを遵守することが必要であり、これらのガイドラインへの違反に対しては広告掲載数の減少や YPP の資格喪失、チャンネル削除といった対応を行っていると同いました。それを前提に、以下の 4 点をご教示ください。もしご回答が難しいということでしたら、その理由を具体的に教示ください。 (1) 日本において、YPP に参加し収益が得られるようになるためには、具体的にどのような基準により、どのような手法で、ボットかどうかの本人確認含め審査を行っているのでしょうか。また、この審査において、不適当として参加できなかったものの件数と、世界全体で不適当とされたものの総数に対する日本における当該件数の割合をご教示ください。 (2) 日本国内で実際に広告掲載数の減少／ YPP 資格喪失／チャンネル削除それぞれの対応を行った件数 と、世界全体における当該件数における日本国内の件数の割合 (3) 上記 2) のうち、特に偽・誤情報に関するポリシーへの違反を理由として対応を行った件数 (4) 上記 1)) ～ (に含まれる 具体的な事例 【Meta への質問】	満たすクリエイターに還元します。YouTube は質の低いコンテンツの収益化を防ぎ、視聴者に常に信頼性の高い高品質のコンテンツを提供するクリエイターに収益化の機会を提供します。詳しくは、こちらをご覧ください。YouTube パートナープログラムの概要と利用資格については、こちらをご覧ください。YouTube コミュニティガイドラインの適用についてはこちらをご参照ください。	https://www.facebook.com/business/help/686805875660832?id=376980407544978 しかし、これらのアカウントは、誤情報ポリシーに違反するコンテンツを繰り返し投稿するなど、弊社のポリシーに違反した場合、収益化の資格を失う可能性があります。Facebook または Instagram で独立したサードパーティのファクトチェッカーによって「誤り」または「改変」に当たると評価されたコンテンツを繰り返し投稿するアカウントは、90 日間のペナルティを受けます。たとえば、Facebook のフィードや Instagram のフィードおよびストーリーズでこれらの投稿を下位に表示し、利用者がそれを見る可能性を減らします。これらのコンテンツを引き続き投稿する場合、さらに下位に移動します。ページ、グループ、ウェブサイトも収益化の能力や広告・ブーストコンテンツを使用して閲覧数を増やすことができなくなります。詳細はこちらをご覧ください：https://www.facebook.com/business/help/297022994952764?id=673052479947730	コミュニティガイドラインを遵守し、クリエイティブな動画投稿を継続されている方です。 また、プログラムへの参加申し込みの際に、身分証明書の提出を義務付けているため、ボットのアカウントはございません。 (1) ～ (4) の件数について ・(1) ～ (4) に挙げられた件数だけを分類して把握することはしていないため、ご回答は難しい状況です。		

(設問)	Google	Meta	TikTok Japan	Microsoft	X
Facebook や Instagram では、リアル広告が掲載されたコンテンツの発信者（投稿者）において収益化を可能とするプログラムが存在する理解ですが⁴、正しいでしょうか。正しい場合、以下の 5 点をご教示ください。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 (1) コンテンツの内容によって収益化を停止したり無効化したりするポリシーが存在するか否か、及び存在する場合の具体的な内容 日本において どのような基準により、どのような手法で、ポストかどうかの本人確認含め審査を行っているのかを含む) (2) 上記 (1) の審査において不適当として収益化プログラムに参加できなかったものの件数と、世界全体で不適当とされたものの総数に対する日本における当該件数の割合 (3) 日本国内で実際に 上記 (1) のポリシーを適用して収益化を停止・無効化した件数 と、世界全体における当該件数における日本国内の件数の割合 (4) 上記 3) のうち、特に コンテンツの内容が虚偽であることを 理由として対応を行った件数 (5) 上記 (4) に含まれる					

⁴ <https://ja-jp.facebook.com/business/help/646155026631347?id=376980407544978>

(設問)	Google	Meta	TikTok Japan	Microsoft	X
具体的な事例 【TikTok Japan への質問】 Creator Rewards Program では、コミュニティガイドライン違反のコンテンツをそもそも対象外にしていると伺いました。それを前提に、以下の 4 点をご教示ください。もしご回答が難しいということでしたら、その理由を具体的に教示ください。 (1) 日本において、当該 Program に参加し収益が得られるようになるためには、具体的にどのような基準により、どのような手法で、ポットかどうかの本人確認含め審査を行っているのでしょうか。また、この審査において、不適当として参加できなかったものの件数と、世界全体で不適当とされたものの総数に対する日本における当該件数の割合をご教示ください。 (2) 日本国内で実際に コミュニティガイドライン違反を理由に収益化を停止・無効化した 件数 と、世界全体における当該件数における日本 国内の件数の割合 (3) 上記 (2) のうち、特に 偽・誤情報に関する ガイドライン への違反を理由として対応を行った件数 (4) 上記 (3) に含まれる具体的な事例					
Q 5 - 2 【X への質問】					申し訳ありませんが、現時点でご提供できる情報はございません。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
広告レベニューシェアについて、収益化を停止・無効化した件数などの定量的な統計は現時点では存在しないと伺いましたが、将来的に、少なくとも日本における当該件数のデータを蓄積していく予定がありますでしょうか。					
Q 5 - 3 【X への質問】 広告レベニューシェアを受けられるアカウントは青色のバッジを付されている理解ですが、日本において、そうしたアカウントに対しては具体的にどのような審査基準により、どのような手法で、ポットかどうかの確認含め本人確認を行っているのでしょうか。また、この本人確認において、不適当としてバッジを付与できなかったものの件数と、世界全体で不適当とされたものの総数に対する日本における当該件数の割合をご教示ください。そして、バッジ付与後においても、信頼性に関する定期的な確認は行われているのでしょうか。					X の広告レベニューシェアについては、Ads Revenue Sharing、また Ads Revenue Sharing Terms をご参照ください。 広告レベニューシェアでは、X 上で投稿したコンテンツへの返信に表示される広告の認証済みユーザーのオーガニックインプレッションから収益を共有できます。これは、人々が X を通じて直接生計を立てる手助けをするための取り組みの一環です。 資格を得る方法 クリエイター広告収益共有の対象となるには、以下の条件を満たす必要があります： ・ X プレミアムまたは認証組織に加入していること。 ・ 過去 3 ヶ月間の投稿の累計で少なくとも 500 万のオーガニックインプレッションを有していること。 ・ 少なくとも 500 人のフォロワーを有していること。 また、コミュニティノートがついた投稿は、レベニューシェアの対象外となります。 Blog comment on Ads Revenue Sharing
6. 広告サービスについて					
【Google への質問】 貴社は広告主向け・パブリッシャー向けそれぞれの広告サービス（Google 広告、Google	(1) Google 広告コードを使用してコンテンツを収益化する場合、以下のポリシーに準拠する必要があります。	広告が Facebook または Instagram で表示される前に、弊社の広告基準に基づいてレビューします。これは、広告が表示される前に自動的に	TikTok においては、外部の広告媒体メディア向けの広告を配信する、といったサービスはございません。		現段階で、私たちの広告インベントリは 100% 当プラットフォーム内にあります。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
AdSense)を展開し、いわゆる広告仲介プラットフォームの運営を行っている理解ですが、こうした広告サービスに関し、以下の4点をご教示ください(ご回答が難しい場合はその理由をご教示ください)。 (1) 広告配信先のパブリッシャーが運営するメディア(広告媒体)について、日本においては、事前に、どのような基準により、どのような手法で、貴社の仲介により広告配信が可能で問題ないものとして審査しているのでしょうか。 (2) 上記(1)の事前審査において、偽・誤情報等を発信・拡散するメディア(広告媒体)という理由で、広告配信先として問題ありとされたものの件数と、問題ありとされたものの全体における当該件数の割合をご教示ください。 (3) 広告配信先のパブリッシャーが運営するメディア(広告媒体)について、偽・誤情報等を発信・拡散するような悪質なサイトであった場合に、広告主から当該サイトへの広告掲載を停止するよう求める通報窓口のようなものは設けていますでしょうか。 (4) 上記(3)で設けている場合、通報後の対応フローと、実際の通報件数や対応した件	本ポリシーに違反すると、コンテンツに基づいた広告の表示がブロックされたり、アカウントが停止または閉鎖されたりする場合があります。本ポリシーは、Google パブリッシャー向けサービスのご利用に関するポリシーの附則として適用されます。 Google は、自由で開かれたインターネットを実現するために、パブリッシャー様のご自身のコンテンツを収益化できるよう、また広告主様が有益で関連性の高い商品やサービスの広告を通じて潜在顧客にリーチできるよう、サポートしています。広告エコシステムの信頼性を保つためには、収益化の対象となるコンテンツに制限を課すことが必要になります。Google パブリッシャー向けポリシーは、次のカテゴリに分類されます。 コンテンツポリシー 行動ポリシー 動画ポリシー プライバシーに関するポリシー 要件とその他の基準 詳しくは、こちらをご覧ください。 (2) 2024年3月27日に、広告主とパブリッシャー向けポリシーの施行における進捗を共有し、広告収入によって支えられる健全なインターネットを責任を持って維持する取り組みを示す Ads Safety Report2023 年版を公開しました。そのハイライトについて、こちらをご覧ください。	行われます。 弊社のプラットフォーム上で表示される数百万の広告に対して広告基準を適用するために、主に自動化技術に拠っています。また、個別のレビューアーが自動化技術を改善およびトレーニングするために働いており、場合によってはマニュアルで広告をレビューします。 以前に承認された広告が再レビューの対象となることがあります。その理由は、利用者がその広告を非表示にしたり、ブロックしたり、報告したり、その他の否定的なフィードバックを行った場合などがあります。これらのフィードバックは、レビュー過程で見逃したものを示す手助けとなる可能性があります。たとえば： - 精度を確保するためにランダムに行われることがあります。 - 特にブロックや非表示にされる場合に、予想外に高いエンゲージメントを受けた場合。 - レビュー過程で見逃した場合、報告などのシグナルの組み合わせを使用して原因を特定します。 既に表示されている広告が弊社のポリシーに違反している場合、当該広告は取り下げられ、広告の配信は停止されます。 現在アクティブなすべての広告は広告ライブラリで閲覧できます： https://www.facebook.com/ads/library/			(1) (現段階で、私たちの広告インベントリは100%当プラットフォーム内にありますので、ご提供できる情報はございません) (2) (現段階で、私たちの広告インベントリは100%当プラットフォーム内にありますので、ご提供できる情報はございません) (3) (現段階で、私たちの広告インベントリは100%当プラットフォーム内にありますので、ご提供できる情報はございません) (4) (現段階で、私たちの広告インベントリは100%当プラットフォーム内にありますので、ご提供できる情報はございません)

(設問)	Google	Meta	TikTok Japan	Microsoft	X
数など、日本国内における運用状況をご教示ください。 【Meta への質問】 貴社は広告主向け・パブリッシャー向けそれぞれの広告サービスを展開し、いわゆる広告仲介プラットフォームの運営を行っている理解ですが、こうした広告サービスに関し、以下の4点をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 (1) 広告配信先のパブリッシャーが運営するメディア（広告媒体）について、日本においては、事前に、どのような基準により、どのような手法で、貴社の仲介により広告配信が可能で問題ないものとして審査しているのでしょうか。 (2) 上記(1)の事前審査において、偽・誤情報等を発信・拡散するメディア（広告媒体）という理由で、広告配信先として問題ありとされたものの件数と、問題ありとされたもの全体における当該件数の割合をご教示ください。 (3) 広告配信先のパブリッシャーが運営するメディア（広告媒体）について、偽・誤情報等を発信・拡散するような悪質なサイトであった場合に、広告主から当該サイトへの広告掲載を停止するよう求める通報窓口のよ	(3) Google は、常に安全な広告が表示されるように尽力していますが、不適切な広告が審査をすり抜けで表示されてしまうこともあります。Google の広告ポリシーに違反していると思われる広告を見つけた場合の、報告できる仕組みがございます。詳しくは、こちらをご覧ください。 (4) 2024年3月27日に、広告主とパブリッシャー向けポリシーの施行における進捗を共有し、広告収入によって支えられる健全なインターネットを責任を持って維持する取り組みを示す Ads Safety Report2023 年版を公開しました。そのハイライトについて、こちらをご覧ください。	社会問題、選挙、政治に関する広告は、アクティブでなくなっても7年間広告ライブラリで閲覧可能です。利用者は、ポリシーに違反していると思われる広告を広告ライブラリから直接調査し報告することができます。			

(設問)	Google	Meta	TikTok Japan	Microsoft	X
うなものは設けていますでしょうか。 (4) 上記(3)で設けている場合、通報後の対応フローと、実際の通報件数や対応した件数など、日本国内における運用状況をご教示ください。 【TikTok Japan への質問】 貴社は広告主向けに広告サービスを展開している理解ですが、いわゆるパブリッシャー（外部の広告媒体メディア）向けの広告配信サービスは展開（いわゆる広告仲介プラットフォームを運営）しておりますでしょうか。もし展開している場合、以下の4点をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 (1) 広告配信先のパブリッシャーが運営するメディア（広告媒体）について、日本においては、事前に、どのような基準により、どのような手法で、貴社の仲介により広告配信が可能で問題ないものとして審査しているのでしょうか。 (2) 上記(1)の事前審査において、偽・誤情報等を発信・拡散するメディア（広告媒体）という理由で、広告配信先として問題ありとされたものの件数と、問題ありとされたもの全体における当該件数の割合をご教					

(設問)	Google	Meta	TikTok Japan	Microsoft	X
示ください。 (3) 広告配信先のパブリッシャーが運営するメディア（広告媒体）について、偽・誤情報等を発信・拡散するような悪質なサイトであった場合に、広告主から当該サイトへの広告掲載を停止するよう求める通報窓口のようなものは設けていますでしょうか。 (4) 上記(3)で設けている場合、通報後の対応フローと、実際の通報件数や対応した件数など、日本における運用状況をご教示ください。 【X への質問】 貴社は広告主向けに広告サービスを展開している理解ですが、いわゆるパブリッシャー（外部の広告媒体メディア）向けの広告配信サービスは展開（いわゆる広告仲介プラットフォームを運営）しておりますでしょうか。もし展開している場合、以下の4点をご教示ください（ご回答が難しい場合はその理由をご教示ください）。 (1) 広告配信先のパブリッシャーが運営するメディア（広告媒体）について、日本においては、事前に、どのような基準により、どのような手法で、貴社の仲介により広告配信が可能で問題ないものとして審査している					

(設問)	Google	Meta	TikTok Japan	Microsoft	X
のでしょうか。 (2) 上記(1)の事前審査において、偽・誤情報等を発信・拡散するメディア(広告媒体)という理由で、広告配信先として問題ありとされたものの件数と、問題ありとされたもの全体における当該件数の割合をご教示ください。 (3) 広告配信先のパブリッシャーが運営するメディア(広告媒体)について、偽・誤情報等を発信・拡散するような悪質なサイトであった場合に、広告主から当該サイトへの広告掲載を停止するよう求める通報窓口のようなものは設けていますでしょうか。 (4) 上記(3)で設けている場合、通報後の対応フローと、実際の通報件数や対応した件数など、日本における運用状況をご教示ください。					
Q 6 - 2 【Meta への質問】 広告主から掲載の申し込みを受けた広告クリエイティブを審査する際、政治広告など、広告の形態で何らかの思想・言説を発信・拡散しようとするクリエイティブの投稿を制限するような基準は設けておりますでしょうか。設けている場合、その具体的な内容(例えば真実でない場合は制限する、ターゲティング広告としては制限するがマス広告としては許可する、な		社会問題、選挙、政治に関する広告については、追加のステップを踏む必要があります。これらの広告を配信しようとする方は、認証プロセスを完了し、広告に「Paid for by (支払者)」のディスクレーマーを付ける必要があります。これには、政治的な人物、政党、選挙(「投票へ行こう」といったキャンペーンを含む)に言及する広告を作成、修正、公開、または一時停止するすべての人が含まれます。これらの広告は7年間広告ライブラリに保存さ			

(設問)	Google	Meta	TikTok Japan	Microsoft	X
ど)をご教示ください(ご回答が難しい場合はその理由をご教示ください)。		れ、内容を精査できます。社会問題、選挙、政治に関する広告および透明性の取り組みについての詳細はこちらをご覧ください： https://www.facebook.com/business/help/1838453822893854?helpref=faq_content また、出会い系サービス、オンライン薬局、オンラインギャンブルなど、特定の種類のビジネスや製品に関連する広告も制限しています。詳細はこちらをご覧ください： https://transparency.meta.com/en-gb/policies/ad-standards/			
7. なりすましアカウントやなりすまし広告への対応について					
Q7-1 【Google への質問】 一般のアカウントやコンテンツにしても、広告にしても、有名人や有名企業になりすましたり、そうした人・企業の肖像やロゴなどを無断で使用したりしているケースに対しては、ポリシー違反として削除やアカウント停止といった対応を行っているとのことでした。それを前提に、以下をご教示ください。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 (1) アカウント開設について、なりすまし防止のため、日本においては、事前に、どのような基準により、どのような手法で、審査しているのでしょうか。そして、この事前審査において、なりすましとして開設できな	Google には、Google の広告製品を使用するユーザーやパートナーの安全を守るためのポリシーを策定し、それを施行する責任があります。 ● Google は、そのエコシステムを不正から守り、危険なヘイト行為や誤情報を減らし、子供たちを不適切な広告体験から保護するために、積極的にポリシーを施行しています。 ● 2023 年には、55 億以上の広告を削除し、69 億以上の広告を制限し、1270 万以上の広告主アカウントを停止しました。アカウント対応に関するこれらの数字は、前年度から約 2 倍に増加しています。 2023 年、広告の安全を確保するための取り組みを大幅にアップグレードし、ポリシー施行において最先端の大規模言語モデル(LLM)の活用を拡大させました。 ● Google は、AI によって広告の安全性に新たな課題が生まれたこと	まず第一に、誤情報と詐欺は別個の問題です。対策のアプローチも異なります。 弊社は、以下のような場合には、サービスの使用を許可せず、アカウントや他のエンティティ(ページ、グループ、イベントなど)を制限または無効にします。 – スクリプトやその他の不正な手段でアカウントを作成または使用する場合、 – 以前の削除を回避するためにアカウントやエンティティを作成または再利用する場合、 – 他の人を誤解させたり欺いたりするために意図的に身元を偽るアカウントを作成または使用する場合 誤解を招く身元の偽装を評価する際には、次のような要素を考慮します。	(1) ・ 広告発信のためのアカウント開設においては、すべてのアカウントについて審査をしています。その審査において、過去に広告ポリシーに違反した経歴のあるアカウントの情報と照らし合わせることで、なりすましを含む広告ポリシー違反となるアカウントが開設されることを防止しています。 ・ 広告アカウントのなりすましについては、非常に厳格な対応(長期間にわたる出稿禁止など)が適用されています。 (2) ・ 広告発信のためのアカウント開設においては、すべてのアカウントについて審査をしています。その審査において、過去に広告ポリシーに違反した経歴のあるアカウントの情報と照らし合わせることで、なりすまし		(1) 申し訳ありませんが、現時点でご提供できる情報はございません。 (2) 広告主が X 広告を使ってコンテンツを宣伝する場合、そのアカウントとコンテンツは承認プロセスの対象となります。この承認プロセスは、X 広告プラットフォームの品質と安全性を保持するために設計されています。このプロセスにより、X は広告主が広告ポリシーを遵守しているかを確認します。承認プロセスについての詳細はこちらをご覧ください。 https://business.x.com/ja/help/ads-policies.html (3) 申し訳ありませんが、現時点でご提供できる情報はございません。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
かったものの件数と、問題ありとされたもの全体における当該件数の割合をご教示ください。 (2) 配信される広告について、偽広告の防止のため、日本においては、事前に、どのような基準により、どのような手法で、審査しているのでしょうか。そして、この事前審査において、偽広告として配信できなかったものの件数と、問題ありとされたもの全体における当該件数の割合をご教示ください。 (3) 偽のなりすましアカウントが開設されたり、偽広告が配信された場合における事後の対応として、日本国内で実際に偽広告の削除や偽のなりすましアカウント停止といった対応を行った件数と、削除やアカウント停止の全体における当該件数の割合 (4) 上記(3)の対応の契機(自動システムによる検知か、なりすまされた本人からの報告か、その他消費者団体等外部からの報告か)別の内訳 (5) 上記(3)について、特に、なりすまされた本人と、その他消費者団体等外部からの報告を契機として対応を行った件数について、それらの報告総数に対する実際に対応を行った件	を認識していますが、昨年はAI領域の最新のイノベーションを取り入れ、LLMの力を活用してポリシー施行の精度と効率を向上させました。 ● LLMは大量のコンテンツを素早く確認し解釈できるだけでなく、そのコンテンツに含まれる重要なニュアンスを捉えることもできます。また、それと同様に重要なこととして、LLMの活用によって、今までよりも少ないデータでより素早く、違反コンテンツにポリシーを適用することができます。そのためGoogleは、これまでより迅速にポリシーを施行し、新たな問題や今後発生してくる問題にも素早く正面から立ち向かえるようになりました。 ● 広告の安全性に与えるAIの影響を示すデータを紹介します。昨年実施されたページへのポリシー適用の90%以上が、最新のLLMを含む機械学習モデルによって開始されたものです。 2023年、Googleは特に詐欺広告対策に集中的な投資を行いました。 ● Googleの不実表示に関するポリシーでは、広告主による「広告主のビジネス、商品、サービスに関する情報について隠蔽または虚偽記載を行ってユーザーを欺く」広告や「広告のリンク先で『フィッシング』の手口を使ってユーザー情報を収集する」広告の運用は許可されません。 ● Googleは、新たな詐欺や不正の脅威からユーザーをより確実に保護するため、新しいポリシーをいくつか策定したほか、ポリシーの施行方法を更新しました。 ○ 2023年、不正表示に関する	– 名前や年齢などの身元情報の繰り返しまたは重大な変更 – 誤解を招くような自己紹介内容や位置情報などプロフィール情報 – ストック画像や盗まれた写真の使用 – その他の関連するアカウント活動 以下のような方法でなりすましを行う場合、 – 他人の写真を使用して他人を装う、 – 他人を欺くためにアカウントを作成、 – 他の人やエンティティのために発言することを装ってページの作成 (2) 独立したサードパーティのファクトチェッカーによってファクトチェックされ、評価された広告を拒否します。現時点では共有できるデータはありません。 (3)–(6) データを公開していません。悪意のあるアクターに悪用される可能性があるためです。	しを含む広告ポリシー違反となるアカウントが開設されることを防止しています。 ・ 広告アカウントのなりすましについては、非常に厳格な対応((※括弧内は回答非公開))が適用されています。 (3) 事後の対応に関する詳細な件数は公表しておりません。ただし、昨今問題になっている著名人のなりすまし広告に関しては、確認したところ、(※「、」と「著しく」の間の箇所は回答非公開) 著しく少なくなっています。(2024年4月時点) (4) (5) (6) (4)～(6)については、件数や事例について公表していないため、ご回答は難しい状況です。		(4) 申し訳ありませんが、現時点でご提供できる情報はございません。 (5) 申し訳ありませんが、現時点でご提供できる情報はございません。 (6) 申し訳ありませんが、現時点でご提供できる情報はございません。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
数の割合 (6) 上記(1)～(3)それぞれについて、日本国内で実際に対応を行った具体的な事例 【Meta への質問】 一般のアカウントやコンテンツにしても、広告にしても、有名人や有名企業になりすましたり、そうした人・企業の肖像やロゴなどを無断で使用したりしているケースに対しては、ポリシー違反として削除やアカウント停止といった対応を行っているとのことでした。それを前提に、以下をご教示ください。もしご回答が難しいということでしたら、その理由を具体的に教示ください。 (1) アカウント開設について、なりすまし防止のため、日本においては、事前に、どのような基準により、どのような手法で、審査しているのでしょうか。そして、この事前審査において、なりすましとして開設できなかったものの件数と、問題ありとされたもの全体における当該件数の割合をご教示ください。 (2) 配信される広告について、偽広告の防止のため、日本においては、事前に、どのような基準により、どのような手法で、審査しているのでしょうか。そして、この事前審査において、偽広告として配信できなかったものの件数と、問題ありと	Google のポリシーに違反しているものとして、2 億 600 万以上の広告を削除しました。 ○ また今年に入ってから、著名な人物、ブランド、組織の支持を受けている、またはそれらと提携関係にあると虚偽の宣伝をしている悪質な広告主のアカウントを素早く永久停止できるようにするポリシーの更新も発表しました。 ○ さらに、新たに広告配信の制限ポリシーを施行し、新たな広告主には「会社確認」を行う期間を設けました。この期間中は、広告インプレッションが制限される可能性があります。(当面このポリシーは、広告キャンペーンでブランド名を対象にしている広告主にのみ適用されません。つまり、広告主とその広告で言及されているブランドの関係性が不明確な場合や、その広告が一般的な内容である場合に、このポリシーが適用されます。) ● Google では、新たな問題が見つかった場合に、即座に専属チームを編成して対応することがよくあります。 手始めに、悪質な行為者の行動パターンを正確に特定し、それと一致する広告やそれと類似した広告があれば自動的に検出のうえ削除するように、ポリシー施行モデルを訓練しました。 ● 2023 年初頭にマルウェア詐欺が発生した際は、Google のチームが迅速に対応し、何万もの悪質な広告を削除することができました。 Google は、機敏な対応とあらゆる手段を通じて、人々を利用したり欺				

(設問)	Google	Meta	TikTok Japan	Microsoft	X
されたものの全体における当該件数の割合をご教示ください。 (3) 偽のなりすましアカウントが開設されたり、偽広告が配信された場合における事後の対応として、日本国内で実際に偽広告の削除や偽のなりすましアカウント停止といった対応を行った件数と、削除やアカウント停止の全体における当該件数の割合 (4) 上記(3)の対応の契機(自動システムによる検知か、なりすまされた本人からの報告か、その他消費者団体等外部からの報告か)別の内訳 (5) 上記(3)について、特に、なりすまされた本人と、その他消費者団体等外部からの報告を契機として対応を行った件数について、それらの報告総数に対する実際に対応を行った件数の割合 (6) 上記(1)～(3)それぞれについて、日本国内で実際に対応を行った具体的な事例 【TikTok Japan への質問】 一般のアカウントやコンテンツに しても、広告にしても、有名人や有名企業になりすましたり、そうした人・企業の肖像やロゴなどを無断で使用したりしているケース に対しては、ガイドライン	こうとしたりする悪質な広告を防ぐために懸命に取り組んでいます。 ● 広告に支えられた健全なインターネットを維持するため、何千人もの Google 社員が 24 時間体制で対応しています。 ● 悪質な行為者同士が連携し、高度な手口を使って検出とポリシーの適用を逃れていることも少なくありません。問題はそれぞれ異なりますが、Google チームは、新たなポリシー、迅速なポリシーの施行、広告認定資格の拡大、使いやすい透明性ツールといった手段を組み合わせ、そうした新たな脅威や不正行為に対抗しています。 ● 2023 年には、新たな問題や今後発生してくる問題にも正面から取り組めるように、広告配信の制限に関するポリシーや AI で生成した選挙広告の情報開示要件などを含む 31 の新しいポリシーを策定しました。				

(設問)	Google	Meta	TikTok Japan	Microsoft	X
違反として削除やアカウント停止といった対応を行っていると言いました。それを前提に、以下をご教示ください。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 (1) アカウント開設について、なりすまし防止のため、日本においては、事前に、どのような基準により、どのような手法で、審査しているのでしょうか。そして、この事前審査において、なりすましとして開設できなかったものの件数と、問題ありとされたもの全体における当該件数の割合をご教示ください。 (2) 配信される広告について、偽広告の防止のため、日本においては、事前に、どのような基準により、どのような手法で、審査しているのでしょうか。そして、この事前審査において、偽広告として配信できなかったものの件数と、問題ありとされたもの全体における当該件数の割合をご教示ください。 (3) 偽のなりすましアカウントが開設されたり、偽広告が配信された場合における事後の対応として、日本国内で実際に偽広告の削除や偽のなりすましアカウント停止といった対応を行った件					

(設問)	Google	Meta	TikTok Japan	Microsoft	X
数と、削除やアカウント停止の全体における当該件数の割合 (4) 上記(3)の対応の契機 (自動システムによる検知か、なりすまされた本人からの報告か、その他消費者団体等外部からの報告か)別の内訳 (5) 上記(3)について、特に、なりすまされた本人と、その他消費者団体等外部からの報告を契機として対応を行った件数について、それらの報告総数に対する実際に対応を行った件数の割合 (6) 上記(1)～(3)それぞれについて、日本国内で実際に対応を行った具体的な事例 【Xへの質問】 一般のアカウントやコンテンツにしても、広告にしても、有名人や有名企業になりすましたり、そうした人・企業の肖像やロゴなどを無断で使用したりしているケースに対しては、ポリシー違反として削除やアカウント停止といった対応を行っていると言いました。それを前提に、以下をご教示ください。もしご回答が難しいということでしたら、その理由を具体的にご教示ください。 (1) アカウント開設について、なりすまし防止のため、日本においては、事前に、どのような基準により、どのような手法で、審査してい					

(設問)	Google	Meta	TikTok Japan	Microsoft	X
るのでしょうか。そして、この事前審査において、なりすましとして開設できなかったものの件数と、問題ありとされたもの全体における当該件数の割合をご教示ください。 (2) 配信される広告について、偽広告の防止のため、日本においては、事前に、どのような基準により、どのような手法で、審査しているのでしょうか。そして、この事前審査において、偽広告として配信できなかったものの件数と、問題ありとされたもの全体における当該件数の割合をご教示ください。 (3) 偽のなりすましアカウントが開設されたり、偽広告が配信された場合における事後の対応として、日本国内で実際に偽広告の削除や偽のなりすましアカウント停止といった対応を行った件数と、削除やアカウント停止の全体における当該件数の割合 (4) 上記(3)の対応の契機 (自動システムによる検知か、なりすまされた本人からの報告か、その他消費者団体等外部からの報告か)別の内訳 (5) 上記(3)について、特に、なりすまされた本人と、その他消費者団体等外部からの報告を契機として					

(設問)	Google	Meta	TikTok Japan	Microsoft	X
対応を行った件数について、それらの報告総数に対する実際に対応を行った件数の割合 (6) 上記(1)～(3)それぞれについて、日本国内で実際に対応を行った具体的な事例					
Q7-2 【Google への質問】 ポリシーに違反する広告に対して、貴社が精力的に対応していることは理解しています。にもかかわらず、日本において、著名人になりすました詐欺的な広告をはじめ、明らかにポリシー違反と見られる広告が、見逃されて、フィード上に多数残っているように見えます。日本において、明らかにポリシーに違反する広告が多数残っている主な原因は何でしょうか(例えば、人員などのリソース不足、検知アルゴリズムの限界等)。 【Meta への質問】 ポリシーに違反する広告に対して、貴社が精力的に対応していることは理解しています。にもかかわらず、日本において、著名人になりすました詐欺的な広告をはじめ、明らかにポリシー違反と見られる広告が、見逃されて、フィード上に多数残っているように見えます。日本において、明らかにポリシーに違反する広告が多数残っている主な原因は何でしょうか(例えば、人員などのリソース不足、検知アルゴリズムの限界等)。	上記をご参照ください。	弊社は、日本の利用者を対象とした詐欺へ厳正に対応し、これを防止するという日本政府と同じ思いを持っています。このため、詐欺と戦うために多大な投資を行い、プラットフォーム上の詐欺から利用者を保護するための多面的なアプローチを開発しています。これには、この種の行動を禁止または抑止するポリシーとシステム、プラットフォーム内外での対応、詐欺を報告し自己防衛するためのツールや機能、各国政府および非政府組織との教育キャンペーン若しくはパートナーシップ構築が含まれます。 詐欺は非常に敵対的な領域であり、オンラインにおける悪意ある行為のパターンの変化に対応するために弊社は技術を進化させ続けています。最大かつ最も持続的な効果をもたらすために、弊社は独自の探知活動およびトラステッド・パートナーからの報告といった幅広いシグナルを利用して、持続的で組織的に脅威を企図する者への捜査及び対象者の壊滅をターゲットにしています。たとえば、2023 年 1 月から 2024 年 1 月の間に、数十万ものアカウントに対して削除等の対策をとりました。	・(※前半部分は回答非公開) TikTok では、常に新しいタイプの違反広告に関する情報を自動モデレーション技術に反映させ、また審査員のトレーニングにおいても最新情報の共有を行なうことで、意図的にコンテンツを調整しながら違反広告を配信しようとする傾向をキャッチアップして審査に反映するよう取り組んでいます。		広告主は、Xの広告ポリシーに従う必要があります。Xの「虚偽のコンテンツや詐欺的コンテンツ」ポリシーは、詐欺的または詐欺的な行動に関連するコンテンツを全世界的に禁止しています。私たちは、広告主がポリシーを遵守しているかを確認するために、機械学習アルゴリズムと人間によるレビューの組み合わせを使用しています。ポリシーに違反するコンテンツがプラットフォーム上で宣伝されるのを防ぐために、今後も私たちは常にモデルとレビューメカニズムを強化していきます。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
【TikTok Japan への質問】 ポリシーに違反する広告に対して、貴社が精力的に対応していることは理解しています。にもかかわらず、日本において、著名人になりすました詐欺的な広告をはじめ、明らかにポリシー違反と見られる広告が、見逃されて、フィード上に多数残っているように見えます。日本において、明らかにポリシーに違反する広告が多数残っている主な原因は何でしょうか（例えば、人員などのリソース不足、検知アルゴリズムの限界等）。 【X への質問】 ポリシーに違反する広告に対して、貴社が精力的に対応していることは理解しています。にもかかわらず、日本において、著名人になりすました詐欺的な広告をはじめ、明らかにポリシー違反と見られる広告が、見逃されて、フィード上に多数残っているように見えます。日本において、明らかにポリシーに違反する広告が多数残っている主な原因は何でしょうか（例えば、人員などのリソース不足、検知アルゴリズムの限界等）。		また、ポリシーに違反するコンテンツやアカウントを識別するために、新しい機械学習技術などさまざまな方法を引き続き展開しています。			
Q 7 - 3 【Google への質問】 特に、生成 AI で作られたと思われる有名人の偽動画・音声を活用した偽広告への、日本国内における対策方法及びその効果をご教示ください。 【TikTok Japan への質問】 特に、生成 AI で作られたと思われ	上記をご参照ください。		・ 生成 AI で作られたものも含めて、著名人のなりすましや肖像・ロゴを無断で使用した広告は、TikTok 広告ポリシーにおいて禁止されています。そのため、広告が配信される前に、これらのコンテンツは事前審査において不承認となり、配信が阻止されます。 ・		AI によって生成されたメディアやコンテンツの共有は、X 上でそれ自体が違反ではありません。AI が生成したコンテンツは X の利用規約および／または X のルールの対象となり、私たちはそのコンテンツの生成元に関わらず、私たちのポリシーを適用します。有害で違法なコンテンツや行動に関しては、AI を使用し

(設問)	Google	Meta	TikTok Japan	Microsoft	X
る有名人の偽動画・音声を活用した偽広告への、日本国内における対策方法及びその効果をご教示ください。 【X への質問】 特に、生成 AI で作られたと思われる有名人の偽動画・音声を活用した偽広告への、日本国内における対策方法及びその効果をご教示ください。			https://ads.tiktok.com/help/article/tiktok-ads-policy-misleading-andfalse-content?lang=ja ・ https://ads.tiktok.com/help/article/tiktok-ads-policy-intellectual-property-infringement?lang=ja ・ また、生成 AI など新しい技術を用いた偽広告を活用した悪意ある活動については、日々新たな手法が生まれています。これらの活動は、認識されてはじめて審査に反映することが可能になりますが、常に新たな傾向を把握することが重要です。そのため TikTok では、スピーディに対応するために生成 AI に特化したポリシーを作成しており、新たな傾向に関する審査員へのトレーニングも継続的に行なっています。 <参考> ・ 広告への適用については現在検討中ですが、通常コンテンツにつきまして、TikTokでは、コンテンツの出所と信頼性に関する標準化団体 Coalition for Content Provenance and Authenticity (C2PA) と連携し、他のプラットフォームで作成されたAI生成コンテンツにも自動的にラベル付けを開始いたしました。この取り組みは、C2PAの技術である「コンテンツクレデンシャル」機能を動画プラットフォームとして初めて実装するものです。今後も、広告を含むコンテンツの真正性の確保に積極的に取り組んでまいります。 ・ 弊社リリース		て生成されたかどうかにかかわらず、それを検出し対処できるかどうかは私たちの優先事項です。 また、私たちは AI 選挙協定の共同署名者です。これらの問題が一つの企業によって解決されることはなく、業界間の協力が重要であると考えています。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
			https://newsroom.tiktok.com/ja-jp/partnering-with-our-industry-to-advanceai-transparency-and-literacy-jp ・ C2PAを代表したAdobeからのリリース https://blog.adobe.com/en/publish/2024/05/09/advancing-trust-transparencyonline-with-tiktok		
8. 生成 AI 対策について					
Q 8－1 【Meta への質問】 貴社は投稿された画像が生成 AI で生成された場合、ラベルを表示する取り組みを開始すると発表しています。これは投稿者の自己申告ではなく、サービス上で自動でラベル付けされるという理解でよいでしょうか（もし自動でない場合は、自動とする予定はあるのか、あるいはない場合は何が壁になっているか教えてください）。		2 月に、弊社は映像や音声を含む AI コンテンツを識別するための共通の技術標準について、業界のパートナーと協力していることを発表しました。AI 生成の映像、音声、画像に対する「Made with AI」ラベルは、AI 画像の業界共有シグナルの検出や、AI 生成コンテンツをアップロードしている旨の利用者の自己申告に基づいています。 また、「Imagined with AI」ラベルを、Meta AI 機能を使用して作成されたフォト・リアリスティックな画像に追加していきます。			
9. 児童その他脆弱な主体の保護について					
Q 9－1 【Google への質問】 児童の保護について特別な配慮を行っているようでしたらご教示ください（プロファイリングに基づく広告規制等）。なお、児童のリテラシー教育については既にヒアリングシートへご回答いただいておりますので、リテラシー教育以外で何かあるかをご教示いただけますと幸いです。 【Meta への質問】 児童の保護について特別な配慮を	YouTube は、あらゆる年齢の若者に安全なオンライン体験を提供できるような環境作りが必須だと考えています。YouTube は長年かけて確固たるポリシーを確立し YouTube コミュニティ全体を守るとともに、ユーザーが信頼性の高い情報源につながれるよう努めてきました。たとえば、摂食障害や自殺を助長または美化するようなコンテンツや、危険行為に挑戦するコンテンツを禁止する取り組みを行っています。また、YouTube Kids（13 歳以下）や、も	弊社は、Meta のサービスを利用する利用者、特に若者の安全を守る責任を認識しています。 弊社は、プラットフォームを若者にとって安全で年齢に適したものにするために、青年の発達、心理学、および精神衛生の専門家と定期的に相談し、どの種類のコンテンツが若年層の利用者にとって不適切である可能性があるかを理解するために努力しています。ポリシーに違反するコンテンツを削除することに加えて、若年層の利用者が年齢に不適切なコ			X は子供やティーンエイジャー向けの主要なプラットフォームとして設計されていないため、子供向けのビジネスユニットもありません。X のユーザーは 13 歳以上でなければなりません。13 歳未満の人がアカウントを作成しようとする、と、登録することができません。 未成年者は X のユーザーベースの一部を占めるだけですが、オンラインでより脆弱な対象であるこのグループの保護に全力を尽くしています。私たちは、プラットフォーム上の若

(設問)	Google	Meta	TikTok Japan	Microsoft	X
行っているようでしたらご教示ください（プロファイリングに基づく広告規制等）。 【X への質問】 児童の保護について特別な配慮を行っているようでしたらご教示ください（プロファイリングに基づく広告規制等）。	う少し年齢の高い子どもの保護者向け管理機能を通じて、年齢に応じたコンテンツライブラリを利用できるようにしているほか、子どもの YouTube 利用時間の管理に役立つ各種機能を提供しています。子どもの発達の専門家と協力して開発したコンテンツの利用者（ログイン・ログアウト両方を含む）は、ひと月あたり 1 億人を超えています。さらに、18 歳未満のユーザー向けに、休憩およびおやすみ時間のリマインダーや、若年層にはふさわしいとは言えないコンテンツに年齢制限を設けるなど、さまざまな安全機能を提供しています。こうした安全に向けた取り組みは、今後も引き続き徹底していく所存です。 暮らす場所にかかわらず、安全かつ質の高いオンライン体験を得る権利が誰にでもあり、子どもや青少年は皆、インターネットを通じた機会へのアクセスを平等に与えられるべきです。若年層のオンライン利用方法は常に進化し続けており、YouTube もそれに従ってサービスとポリシーを進化させ続けています。以下では、YouTube が運営上の指針としている 5 つの基本的な考え方を紹介します。しかし、この取り組みは YouTube だけで実現できることではありません。だからこそ私たちは、政策立案者、研究者、専門家、企業、家族が一丸となって、オンラインで若年層にサービスを提供する企業のための、一貫した基準を定義することを支援しています。 1: 子どもと青少年のプライバシー	ンテツを見にくくすることも目指しています。年齢に不適切なコンテンツとは、生物学的、社会感情的、認知的な発達の違いから若年層の利用者に特有のリスクをもたらす可能性のあるコンテンツです。 若年層の利用者にとって、潜在的にセンシティブなコンテンツを見つけにくくしています。本人がフォローしている人が共有した場合でも、です。 若年層の利用者に対して年齢に適した保護を提供する一方で、彼らを含む利用者が自分に合った方法でソーシャルメディアの体験を管理し、ウェルビーイングをサポートできるようにすることも重要だと認識しています。これが、利用時間の管理、望まないやり取りの防止、見たいコンテンツやアカウントの制御ができる機能を導入した理由です。また、いじめを防ぐためのツールも開発し、利用者がアカウントを管理しはじめに遭遇しないようにしています。			年層のユーザーを保護するためにさまざまなツールとポリシーを導入しています。 当社の年齢保証プロセスは、自己申告による年齢確認と追加の技術的措置を組み合わせ、アカウント保持者の年齢が本物であり、子供を保護するための適切なコントロールが設置されていることを確認します。 13 歳から 17 歳のユーザーは、デフォルトでプライバシー、安全、セキュリティの設定が強化されています。例えば、彼らはグラフィックコンテンツ、成人の裸体、性的行動を含む敏感なメディアを見ることができず、ダイレクトメッセージが制限され、位置情報もオフになっています。広告主はこの年齢層をターゲットにすることが禁止されています。 親や他の人からアカウント保持者が 13 歳未満であるとの確認が得られた場合、私たちのポリシーはそのアカウントを永久に停止することです。さらに、ユーザーが 18 歳未満であることを示す生年月日を入力すると、新しい生年月日を再入力することができなくなります。 ここにいくつかの具体的な措置を挙げます： 年齢制限コンテンツ- 敏感なメディア設定： 私たちは、未成年者が知っているアカウントやプロフィールに生年月日を含めていないユーザー

(設問)	Google	Meta	TikTok Japan	Microsoft	X
	一、身体的安全、メンタルヘルス、ウェルビーイングを守るために、オンラインにおいて特別な対策が必要。 YouTube は、若年層のために素晴らしい経験を築くことを目指しているため、彼らの安全性、ウェルビーイングや、メンタルヘルスは、当プラットフォームと継続的な開発の中核となっています。また、若年層による利用方法の実情が、これらのサービスやポリシーに正確に反映されるようにしたいと考えています。具体的には、ネットいじめや未成年者の性的対象とするコンテンツを禁止しているほか、YouTube 上の不正行為を早い段階で特定できるよう改善を続け、こうした行為の阻止に努めています。 2: 特に 12 歳以下の子どもがいる家庭内でのオンライン利用ルールを定めるにあたり、保護者は重要な役割を担う。 保護者は、子どものインターネットの利用時間を明確に制限したり、オンラインで責任のある選択をするように教育したりするなど、子どもの行動に大きく影響を与えることができます。こうした家庭の役割を強化できるよう、YouTube でもできる限りの取り組みを行っています。特に、青少年による YouTube の利用を管理する新しい方法を検討しており、既存の保護者による使用制限を拡充していく予定です。				に対して、成人向けコンテンツなど特定の敏感なメディアの視聴をインタースティシアルを通じて制限しています。 年齢制限コンテンツ- コンテンツ通知: X は通知やインタースティシアルの背後に敏感なメディアを隠します。これには未成年者が成人向けコンテンツを閲覧するのを制限する年齢制限コンテンツポリシーが含まれます。 年齢制限コンテンツ- セーフサーチ: X は自動的に、未成年者や生年月日が記載されていないアカウントの検索結果から、潜在的に敏感なメディアコンテンツ（ユーザーがミュートまたはブロックしたアカウントも含む）を除外します。 年齢制限コンテンツ- 広告: 未成年者向けに禁止されているコンテンツを含む広告は「ファミリーセーフではない」とタグ付けされ、21 歳以上のユーザーにのみ表示されます。この基準は現時点で全世界で一貫しています。 アカウントのデフォルト保護設定: 未成年者のアカウントはデフォルトで「保護された投稿」に設定されます。これは、新しい人がフォローを希望した際に未成年者がリクエストを受け（承認または拒否が可能）、彼らの投稿がフォロワーにのみ表示され、彼らとそのフォロワーにのみ検索可能であることを意味します（つまり、公開検索には表示されません）。 ダイレクトメッセージのデフォルト保護設定: 未成年者のアカウントはデフォルトでフォローしているア

(設問)	Google	Meta	TikTok Japan	Microsoft	X
	3: すべての子どもは、個人的な興味やニーズに応じて、年齢にふさわしい質の高いコンテンツを無料で視聴する機会を与えられるべきである。 子どもの年齢に適した安全な動画を見せるだけで十分とはいえません。YouTube は、若年層に人生を豊かにし、想像力を刺激するようなコンテンツを視聴してもらいたいと考えています。そのために、子どもの発達、デジタル学習、デジタルシチズンシップの分野の専門家と協力しながら、今後もコンテンツを拡充させていく予定です。子ども向けの質の高いコンテンツのガイドラインを基盤に、青少年のニーズにより的確に応えていけるよう尽力していきます。 もちろん、どんなに質の高いコンテンツであっても、誰もが手頃にアクセスできなければあまり意味がありません。YouTube は、家庭の経済事情に関わらずあらゆる子どもが豊富なコンテンツを平等に視聴できるよう努めており、これは広告を基盤として成り立っています。ただし、YouTube Kids と、YouTube を利用する 18 歳未満のユーザーに対してパーソナライズド広告は配信しておらず、今後も配信する予定はありません。 4: 子どもの発達上のニーズは青少年とは大きく異なり、オンライン体験に反映させる必要がある。				カウントからのダイレクトメッセージのみを受け取るよう制限されます。 年齢ロック：新しいユーザーが 18 歳未満であるとする生年月日を入力した場合、そのアカウントで新たに生年月日を再入力することはできなくなります。 自殺および自傷行為の援助：X は、ユーザーが自殺や自傷行為を促進または奨励することを禁じるポリシーを開発しました。自殺または自傷に関連する用語を検索すると、最上位の検索結果は助けを求めることを奨励する通知が表示されます。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
	青少年は自立心がつくに従って、アイデンティティやコミュニティを求めたり、縛られることなく幅広い重要なコンテンツを探究したり、自分を理解して受け入れてくれるような関係を構築したりします。また、この世代は視聴にとどまらず、投稿やコメントをして積極的に関わっていきます。しかし、青少年の脳はまだ発達途中であり、特に意思決定や自制に関しては、大人の脳とは異なる働きをします。 青少年のオンライン上での行動が絶えず変化していくなかで、YouTube はこうした行動の変化についていくよう努めながら、サービスやポリシー全体を通じて幅広い対策を講じていきます。たとえば、サービスの利用中に状況に適した学びを盛り込んで、健全な利用を促すといった方法があります。具体的には、動画を視聴中のユーザーに「休憩」を促す通知を表示する、動画を作成中のユーザーに個人情報の公開を控えるよう注意喚起する、などの方法が含まれます。 5: 適切な安全保護対策があれば、革新的なテクノロジーは子どもや青少年に恩恵をもたらすことができる。 最新のデバイスにしろ新しい生成AI ツールにしろ、今日の若者は、革新的なテクノロジーを試したり、使いこなしたりすることに貪欲です。YouTube は、青少年に関する原則に沿って、ウェルビーイングに役				

(設問)	Google	Meta	TikTok Japan	Microsoft	X
	立つ機能や自動学習ツールを盛り込むなど、テクノロジーの開発および利用を進めてまいります。インターネットが若年層に及ぼす影響について研究が進むなかで、その分野の専門家と協力しながら、共同でできる取り組みを引き続き模索していきます。 また、2023 年 10 月に「子どもと青少年をオンライン上で保護するための法的枠組みへの提言」を発表しました。詳しくは、こちらをご覧ください。				
Q 9 - 2 【Google への質問】 児童に限らず、貴社のサービスとの関係で、特に脆弱な主体の保護について留意すべきことがあればご教示下さい。 【Meta への質問】 児童に限らず、貴社のサービスとの関係で、特に脆弱な主体の保護について留意すべきことがあればご教示下さい。 【TikTok Japan への質問】 ヒアリング当日、児童の保護についての取組を種々ご紹介いただきましたが、児童に限らず、貴社のサービスとの関係で、特に脆弱な主体の保護について留意すべきことがあればご教示下さい。 【Microsoft への質問】 児童に限らず、(高齢者など)脆弱な主体が有害な情報に触れないようにするための工夫について何かしらの取組があればお知らせ下さい。 【X への質問】	例えば YouTube では、個人を対象とした身体的特徴または保護対象グループへの所属（年齢、障がい、民族、性別、性的指向、人種など）に基づく長期に及ぶ侮辱または中傷を含むコンテンツは許可されていません。その他の有害な行為（脅迫や晒し行為など）も許可されません。未成年を対象とするコンテンツに対しては、より厳しく対処されます。 Google では、様々なプロダクトやサービスにおいてユーザー補助機能を提供しています。詳しくは、こちらをご覧ください。	弊社は、Meta のサービスを利用する利用者、特に若者の安全を守る責任を認識しています。 弊社は、プラットフォームを若者にとって安全で年齢に適したものにするために、青年の発達、心理学、および精神衛生の専門家と定期的に相談し、どの種類のコンテンツが若年層の利用者にとって不適切である可能性があるかを理解するために努力しています。ポリシーに違反するコンテンツを削除することに加えて、若年層の利用者が年齢に不適切なコンテンツを見にくくすることも目指しています。年齢に不適切なコンテンツとは、生物学的、社会感情的、認知的な発達の違いから若年層の利用者に特有のリスクをもたらす可能性があるコンテンツです。 若年層の利用者にとって、潜在的にセンシティブなコンテンツを見つけにくくしています。本人がフォローしている人が共有した場合でも、です。	TikTok に限らず多くのプラットフォームに広く共通する留意事項として、オンラインを通じた性被害や誹謗中傷、自殺に関するコミュニケーション、デジタルリテラシーの格差などが挙げられます。これらの課題解決のために、TikTok ではサポートを必要とする人とサポートを提供できる人を繋げる架け橋として、NPO 等への相談を支援するキャンペーンや、専門家やクリエイターと連携した啓発動画の作成、青少年保護のための機能の開発などに取り組んでいます。	LinkedIn では、プロフェッショナル コンテンツ ポリシーに違反するコンテンツを削除することで、高齢者を含むすべてのメンバーを保護しております。 マイクロソフトでは、提出済みの回答でご説明しましたさまざまなポリシーに違反するコンテンツを削除しております。児童が有害なコンテンツに触れないように保護していただけるように、家族向けに Bing セーフサーチと Microsoft ファミリーセーフティ設定を用意しています。 Bing のセーフサーチ機能は、露骨な画像やテキストを検索結果から除外するものです。セーフサーチは、ほとんどの市場では既定で「標準」に設定されており、成人向けの画像が検索結果から除外されます。「高レベル」モードでは、露骨な表現を含むテキストと画像の両方が検索結果から除外されます。ファミリーセーフティ設定も、保護者が	X の使命は、誰もがアイデアや情報を創造し共有し、意見や信念を障害なく表現できる力を与えることです。表現の自由は人権です-私たちは、誰もが声を持ち、それを使う権利があると信じています。私たちの役割は、多様な視点を代表する公共の会話を支えることです。人々が X 上で虐待を経験すると、自己表現の能力が危険にさらされることがあります。研究によると、特定のグループの人々がオンラインで過剰に虐待の対象になっています。複数の代表性の低いグループに属する人々にとって、虐待はより一般的で、性質がより重く、より有害である可能性があります。私たちは、特に過去に疎外されてきた人々の声を封じ込めようとする虐待に対抗することに尽力しています。この理由から、保護されたカテゴリーに属すると見なされる個人またはグループを標的にした虐待行為を禁止しています。 また、自殺と自傷行為に関するポリ

(設問)	Google	Meta	TikTok Japan	Microsoft	X
児童に限らず、貴社のサービスとの関係で、特に脆弱な主体の保護について留意すべきことがあればご教示下さい。		若年層の利用者に対して年齢に適した保護を提供する一方で、彼らを含む利用者が自分に合った方法でソーシャルメディアの体験を管理し、ウェルビーイングをサポートできるようにすることも重要だと認識しています。これが、利用時間の管理、望まないやり取りの防止、見たいコンテンツやアカウントの制御ができる機能を導入した理由です。また、いじめを防ぐためのツールも開発し、利用者がアカウントを管理していじめに遭遇しないようにしています。		Microsoft アカウント、デバイス、またはローカルネットワークで、子どもの検索エクスペリエンスを特定のセーフサーチ設定に設定できるようにするものです。さらに、これらの機能を使用してデバイスまたはネットワーク上で Bing の生成 AI 機能を無効にすることも可能です。 Bing は、ユーザーが検索結果のコンテンツによって誤情報を得たり、損害を受けたりしないようにするための機能を多数備えています。たとえば、ユーザーが潜在的に問題のあるコンテンツを明示的に探している場合、Bing は信頼できる回答、公共サービス広告、または検索コンテンツの潜在的な問題に関する警告を表示することがあります。同様に、潜在的に有害または違法なコンテンツを検索する意図を示すクエリを入力したユーザーに対しては、多くの場合、追加の信頼できる情報を提供する公共サービス広告を表示します。また、ユーザーはビジュアル検索機能を使用してアダルト コンテンツを検索することもできません。	シーもあります。報告を繊細に扱うために、私たちのアプリ内報告は、自傷行為を行う意図を表現している可能性のある人々と、自傷行為や自殺を励ますまたは促進するコンテンツのために、別々のオプションを提供しています。
10. 研究機関等へのデータ提供について					
Q10-1 【Google への質問】 ヒアリングシート Q14 に関連して、日本国内の研究者への API 公開について、2022 年と 2023 年における最近の実績（提供した研究者数や提供目的等）をご教示ください。また、日本における API 公開によるデータ提供について、その可否や範囲等を検討するに当たっての悩み・課題などをご教示ください。	研究者とのエンゲージメント Google では、大学や研究機関との協力関係を育むため、学术界や外部の研究コミュニティにリソースやサポートを提供するさまざまなプログラムを運営しています。データ、ツール、出版物を研究者に提供したり、選ばれた研究者が応募できる研究者プログラムを用意する等、多くのイニシアチブを実施しています。詳しくはこちらをご覧ください。	研究者との契約に関する詳細情報は、秘密保持契約により共有できません。			

(設問)	Google	Meta	TikTok Japan	Microsoft	X
さい。 【Meta への質問】 ヒアリングシート Q14 に関連して、日本国内の研究者への API 公開について、2022 年と 2023 年における最近の実績（提供した研究者数や提供目的等）をご教示ください。また、日本における API 公開によるデータ提供について、その可否や範囲等を検討するに当たっての悩み・課題などをご教示ください。					
Q 1 0 - 2 【TikTok Japan への質問】 日本国内の研究者への API 公開によるデータ提供について、欧米の一部では既に行われているところ、日本における提供も基本的には可能と考えられるが、なぜ日本では行われていないのか、そして、具体的に、どのような環境や条件であれば日本においても可能なのか、現在 検討中と伺っておりますが、今後のスケジュールも含め、ご教示いただけますと幸いです。			日本を含む各国における API 公開については、現在検討中であり、今後のスケジュールなどについて現時点で公表できる事項はございません。		
Q 1 0 - 3 【Microsoft への質問】 ヒアリングシートで、日本の研究機関や研究者との連携は「該当なし」とお答えいただいておりますが、今後、例えば API 公開を通じたデータ提供 の連携・協力の可能性があるのか、その可否 や範囲等を検討するに当たっての条件や 悩み・課題などをご教示ください。				マイクロソフトは、GLOCOM および日本ファクトチェックセンター等と議論、連携しています。こうしたパートナーシップを今後どう発展させることができるのかを模索しており、マイクロソフトとの提携を目指す他の機関からのアプローチを歓迎します。総務省のリードにより、潜在的なニーズや組織を特定することで、こうした連携がより進めやすくなると考えています。	
Q 1 0 - 4 【Microsoft への質問】				マイクロソフトは、安全かつ責任ある AI の利活用を確保するために	

(設問)	Google	Meta	TikTok Japan	Microsoft	X
ヒアリングシート Q.14-1 への回答に関連する質問です。偽情報の蔓延はもちろん、将来的に AI の発展が社会にもたらす様々なリスクを検討し、対応を行っていくためには、企業、大学、政府等の連携が不可欠だと考えています。マイクロソフト社は、すでに東京大学と AI の研究推進に関する連携を行っています。ELSI の観点から、法学や倫理学といった分野と連携を行っていくことも検討されていますでしょうか。またこうした連携を、日本国内でさらに広げていくことは検討されていますでしょうか。				は、ご指摘の法学や倫理学などの ELSI 視点からの検討は重要であり、企業、大学、政府等のマルチステークホルダーアプローチが必要であると考えております。具体的には、Partnership on AI や Roma Call for AI Ethics の創設メンバーの 1 つであり、OECD や UNESCO 等の国際機関における AI の議論にも積極的に関わっています。特に Roma Call for AI Ethics については、日本からも中央大学が署名しており、引き続き、日本を含めてこうした活動が広がるように積極的に参加していきたいと考えております。	
Q 1 0 - 5 【X への質問】 日本国内の研究者への API 公開に係る具体的な条件を書面で回答いただける旨伺いましたので、ご回答いただけますと幸いです。					現時点でご提供できる情報は過日ご案内させていただきました以下に限られます。 ----- 私たちの API は研究者を含む誰でも申し込むことができます。過去 1 年間で利用可能なサービスの範囲を拡大し、無料、ベーシック、プロ、エンタープライズを含む 4 つのアクセスレベルを提供するようになりました。 https://developer.twitter.com/en/docs/twitter-api ----- またアップデートがありましたらご報告させていただきます。
1 1. ガイドライン等の必要性・妥当性その他社会的傾向 について					
Q 1 1 - 1 【Microsoft への質問】 偽・誤情報について、日本の政府がコンテンツモデレーションにどのように関与することが効果的な対処になるか、例えば行動規範を				マイクロソフトは、情報エコシステム全体の主な改善目標に重点を置いた、成果に基づくリスクに比例したアプローチを推奨いたします。こうしたアプローチでは、すべてのオンライン サービスは固有の特徴を有	

(設問)	Google	Meta	TikTok Japan	Microsoft	X
策定することなどについて、日本の文化的背景への評価を含め、貴社のお考えをお聞かせ下さい。加えて、偽・誤情報に対する規制のあり方について、貴社が日本政府に期待されることをお聞かせ下さい。				し、ユーザーを偽・誤情報にさらすリスクもさまざまであることを認識する必要があります。また、サービスプロバイダーにはそれぞれのサービスに関する偽・誤情報に対応する独自の方法があり、そのような成果に基づくアプローチを通じて適切に対応できる柔軟性を維持すべきであることを認識する必要もあります。 EUのような国際的な枠組みに沿った原則の重視は、歓迎すべきものと考えます。ただし、リスク評価と報告の要件は、プロバイダーが毎年適切かつ有意義に対応できるよう、管理しやすいように設計することが推奨されます。 また、既存の規制制度と整合し、さまざまな要件を調和させ、さまざまな規制当局間の連携を促進する選択肢を検討することも推奨します。数多く存在する規制制度間の調和が取れていない状態は、グローバルに事業を展開する企業に、法律の「つぎはぎ対応」状態をもたらします。法律間の整合性の欠如により、企業のコンプライアンス負担が増大し、プロバイダーが市場参入の規模縮小を余儀なくされ、特定市場で消費者に提供されるサービスオファリングが制限される可能性があります。	
Q 1 1－2 【Microsoft への質問】 社会的な傾向の評価についての質問です。プレゼンテーションの最初に表現の自由が重要とのお話がありましたが、昨今の社会的な傾向として、表現の自由をいかに保				また、既存の規制制度と整合し、さまざまな要件を調和させ、さまざまな規制当局間の連携を促進する選択肢を検討することも推奨します。数多く存在する規制制度間の調和が取れていない状態は、グローバルに事業を展開する企業に、法律の「つぎ	

(設問)	Google	Meta	TikTok Japan	Microsoft	X
護すべきか、というよりは、偽・誤情報のような問題のあるコンテンツの蔓延やなりすまし等に対処する必要性の方が重視されるようになっている、というご認識でしょうか。				はぎ対応」状態をもたらします。法律間の整合性の欠如により、企業のコンプライアンス負担が増大し、プロバイダーが市場参入の規模縮小を余儀なくされ、特定市場で消費者に提供されるサービス オファリング が制限される可能性があります。 Bing は、法的義務に応じてコンテンツ モデレーションを行います が、主に防御的なアルゴリズムを利用して潜在的な偽・誤情報の表示を制限します。防御的な検索では、分類器を利用して、検索クエリの意図と、ユーザー クエリに関連する Bing インデックス内の Web コンテンツの「品質」または「権威」を理解します。これらの分類器モデルが危害目的の意図を識別した場合、それらの防御的分類器に一致する内容を含むクエリに関連する Web サイトは「降格」され、トップページに表示されなくなります。検索結果のランク付けの決定にあたっては、クエリの意図の強さとクエリに関連する Web サイトの「品質」または「権威」とのバランスを図ります。これは、ブラウザのモード（セーフサーチ設定など）によっても異なります。一般に、ユーザーが防御的なトピックを特に検索していない場合、Bing は「防御的な」文書を検索結果の上位に表示しないようにします。	
12. ファクトチェック組織との連携について					
Q12-1 【Google への質問】 日本ファクトチェックセンターへの投資や YouTube におけるファク	日本ファクトチェックセンターを含め、第三者機関の専門家と引き続き連携して参ります。				

(設問)	Google	Meta	TikTok Japan	Microsoft	X
トチェックビデオの作り方等の情報共有の取組をご紹介いただきましたが、さらに進んで、ファクトチェック組織と連携して YouTube の動画や Google 検索の検索結果にラベル付けをしたり、ファクトチェック結果をまとめたページを公開するといった取組を日本国内で進める予定はございますでしょうか。また、そのような連携に当たって障害となっていることがもしあれば、それは何か、具体的にご教示ください。					
Q 1 2 - 2 【Meta への質問】 第三者ファクトチェック・プログラムに日本のファクトチェック機関も参加できるようになるということですが、こちらのプログラムでは参加するファクトチェック機関に対して何らかの経済的インセンティブを付与することがあるのでしょうか。また、参加するための要件を具体的にご教示ください。また、参加するための要件を具体的にご教示ください。		すべてのサードパーティファクトチェッカーは、IFCN によって認定される必要があります。 弊社は、ファクトチェック団体が新しいスキルを開発し、イノベーションを追求し、オンライン上の誤情報に対処するための努力を拡大するのを支援することで、ファクトチェック業界の長期的な強化をサポートすることにコミットしています。 弊社は、どのプラットフォームよりも大きなグローバルファクトチェックネットワークを構築しており、2016 年以降、ファクトチェック活動を支援するプログラムに 1 億ドル以上を拠出しています。 https://www.facebook.com/formedia/blog/third-party-fact-checking-industry-investments ファクトチェック機関との具体的な契約の詳細は事業上の秘密としています。			
Q 1 2 - 3 【Meta への質問】 第三者ファクトチェック・プログ		弊社のサードパーティファクトチェックプログラムは、単一のグローバル基準に基づいて運営されていま			

(設問)	Google	Meta	TikTok Japan	Microsoft	X
ラムに参加するファクトチェック機関が貴社からの独立性を保つために工夫されていることがあれば、ご教示ください。		す。すべてのサードパーティファクトチェック団体は、IFCN によって認証され、非党派性等を含む原則に従う必要があります。原則の詳細については、こちらをご参照ください。 https://ifcncodeofprinciples.poynter.org/application/public/japan-fact-check-center/30E22DC8-0D31-EF97-8173-CB48728E8E8D			
Q 1 2 - 4 【TikTok Japan への質問】 コミュニティガイドラインの中の「ファクトチェックを行う独立したパートナー」を挙げておられますが、 https://www.tiktok.com/safety/ja/jp/safety-partners/ 相談対応機関のようなところがリストアップされているように思います。これらの機関がファクトチェックに果たす役割についてご教示下さい。			・(※前半部分は回答非公開) ご教示いただいたリンクは、ファクトチェック機関ではなく、セーフティパートナーをご紹介するウェブページです。連携しているファクトチェック機関は、こちら (https://www.tiktok.com/transparency/ja-jp/combating-misinformation/) に掲載されております。		
Q 1 2 - 5 【X への質問】 貴社の偽・誤情報等の対策については、外部有識者やファクトチェック団体、その他の関連団体との連携を行っておられないようですが、それはどのような理由に基づくのでしょうか(ご回答が難しい場合はその理由をご教示ください)。					こちらの ブログ と 投稿 をご確認ください。
1 3. DSA に基づく調査への対応状況について					
Q 1 3 - 1 【TikTok Japan への質問】 貴社の欧州拠点では、昨年からの欧州委員会からの DSA に基づく報			欧州委員会からの DSA に基づく報告要求並びに正式調査につきましては現在欧州チームにおいて対応中であるところ、調査中であることか		

(設問)	Google	Meta	TikTok Japan	Microsoft	X
告要求に引き続き、2月には正式調査が開始されたと理解していますが、現在の対応状況や今後の予定について、ご教示いただけますと幸いです（ご回答が難しい場合はその理由をご教示ください）。			ら、現時点では詳細のご回答は控えさせていただきます。		
Q13-2 【Xへの質問】 貴社の欧州拠点では、昨年からの欧州委員会からのDSAに基づく報告要求に引き続き正式調査が開始されたと理解していますが、現在の対応状況や今後の予定について、ご教示いただけますと幸いです（ご回答が難しい場合はその理由をご教示ください）。					他の管轄区域で進行中の案件についてのコメントは差し控えさせていただきますが、引き続き連携させていただければと思います。
14. ガイドライン等の必要性・妥当性について					
Q14-1 【Metaへの質問】 EUでは、DSAに基づくシステミックリスク評価・軽減義務に関連して、選挙のインテグリティに関するリスク評価・軽減措置実施のためのガイドラインの策定が進められているところですが、日本においても、選挙に限らず、例えば自然災害発生時における偽・誤情報リスクに関し、これを評価・軽減するためのガイドライン等を政府やその他のステークホルダーが策定・公表することの必要性・妥当性について、お考えをお聞かせください。 【TikTok Japanへの質問】 EUでは、DSAに基づくシステミックリスク評価・軽減義務に関連して、選挙のインテグリティに関するリスク評価・軽減措置実施のためのガイドラインの策定が進めら		弊社は、複数の管轄地域で業界やセクターの関与を経て形成する自主規制または業界自主規制としての実践規範に積極的に参加してきましたし、またこうしたことを支持しています。これらの実践規範には、EUおよびオーストラリアの誤情報実践規範で見られるように、以下のような利点があります。 － 社としてのポリシーの執行を監視、観察、テストし、将来のポリシー開発に活かすための理由を収集する。 － 政府、業界、その他の利害関係者間の継続的な対話と協力を促進する。 － 地元の懸念に対処しながら、グローバルな業界標準と整合させる。 － 既存の法律を補完し、より広範な政策および立法フレームワークをサポートする。 － 絶えず変化し、または新たに発生	・プラットフォームは、それぞれのサービスごとに、基本的な仕組み（例えば、ソーシャルグラフに基づいて組み立てられているか否か）も、モデレーションやレコメンデーション、広告審査などの手法・社内体制・哲学も、大きく異なります。 ・そのため、あるガイドラインが一つのプラットフォームには有効でも他のプラットフォームの仕組みには当てはまらないといったことも考えられます。そのため、ガイドライン等の検討をされる場合には、事業者の意見を広く反映して議論を進めることが必要であると考えます。		私たちのコンテンツモデレーションシステムは、サービスの使用や基本的な権利、特に表現の自由を不必要に制限することなく、体系的なリスクを軽減するように設計され、調整されています。コンテンツモデレーション活動は、原則に立脚したポリシーに基づいて実施され、多様な介入セットを活用して、私たちの行動が合理的かつ効果的であることを目指しています。私たちのコンテンツモデレーションシステムは、自動化と人間のレビューを組み合わせたもので強固な上訴システムを備えており、ユーザーが潜在的なモデレーションの異常やミスを迅速に指摘できるようにしています。

(設問)	Google	Meta	TikTok Japan	Microsoft	X
れているところですが、日本においても、選挙に限らず、例えば自然災害発生時における偽・誤情報リスクに関し、これを評価・軽減するためのガイドライン等を政府やその他のステークホルダーが策定・公表することの必要性・妥当性について、お考えをお聞かせください。 【X への質問】 EU では、DSA に基づくシステミックリスク評価・軽減義務に関連して、選挙のインテグリティに関するリスク評価・軽減措置実施のためのガイドラインの策定が進められているところですが、日本においても、選挙に限らず、例えば自然災害発生時における偽・誤情報リスクに関し、これを評価・軽減するためのガイドライン等を政府やその他のステークホルダーが策定・公表することの必要性・妥当性について、お考えをお聞かせください。		する脅威に迅速に対応し、改善、適応するための柔軟性を持つ。 しかしながら、技術的要件や義務が適切で実現可能であり、各市場におけるプラットフォームリスクに適していることを確保するためには、業界の代表者や企業自身がこれらの実践規範の開発に関与することが重要です。			