

プラットフォーム事業者ヒアリングの結果（暫定版※）

※追加質問への回答が一部の国外事業者から未提出のため。

項目	結果（○：質問に対する一定の回答があったもの ●：質問に対して回答が不十分なもの）
1 ヒアリング対象サービスの規模	○ 半数の国内事業者及び一部の海外に本社を置く事業者（以下「国外事業者」）は、<u>日本国内における最近のMAU（月間アクティブユーザー数（1か月間に対象サービスに1度でもアクセスした利用者の数））</u>を回答。 ● 一部の国内事業者及び多くの国外事業者からは、<u>日本国内における最近のMAUの回答なし</u>。その理由として、<u>非公表であること、対象サービスの利用にログインが不要であることや、回答しない理由自体が不明</u>。 ● 全ての国内事業者及び国外事業者は、<u>日本国内における最近の月間合計投稿数の回答なし</u>。 <u>【日本国内の最近の月間アクティブユーザー数の回答状況】</u>（参考資料19-1-1 及び 19-1-2 Q1-1、Q2-1、Q2-4）
2 偽・誤情報の流通・拡散への対応方針	○ 一部の国内事業者及びほぼ全ての国外事業者は、<u>偽・誤情報と重複する場合がある偽アカウント・なりすまし・合成コンテンツ・欺瞞行為・詐欺、選挙・市民活動や医療等の分野とは別に、コンテンツモデレーション等</u>（削除、アカウント停止、表示順位の降格、収益不能化、アクセス不能化、警告表示・ラベリング、ファクトチェック結果の表示、投稿時の注記の義務付け等。以下同じ。）<u>による対応の対象となる偽・誤情報の範囲（定義）や類型（例示等）について、個別に策定し日本語で公表しているポリシー、ガイドライン等（以下「偽・誤情報ポリシー等」）</u>を回答。 【偽・誤情報の範囲（定義）や類型（例示等）[例]】 ■ 「Yahoo! ニュースコメントポリシー」 https://news.yahoo.co.jp/info/comment-policy 健康被害等をもたらす可能性のある偽情報であって、ファクトチェックにより反真実であることが明らかな投稿（明らかな偽情報）（「新型コロナウイルスのワクチンを接種すると、流産する。不妊になる。」、「ワクチン接種された実験用の動物が全て死亡した。」、「ワクチンを接種することでコロナウイルスに感染する。」など）

■Yahoo!知恵袋「利用のルール」 <https://chiebukuro.yahoo.co.jp/topic/guide/rule/>

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

■ファイナンス掲示板「ヘルプページ」 <https://support.yahoo-net.jp/ScFinance/s/article/H000011273>

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

■LINE オープンチャット「安心・安全ガイドライン」 <https://openchat-jp.line.me/other/guideline>

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

■「LINE VOOM コミュニティガイドライン」

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

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

■YouTube「誤った情報に関するポリシー」 <https://support.google.com/youtube/answer/10834785?hl=ja>

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

	により、国勢調査への参加を妨げる行為。) - 改ざんされたコンテンツ: ユーザーの誤解を招くように技術的に操作または改ざんされ(前後関係を見落としてクリップを切り抜く以上の操作が多い)、重大な危害を及ぼす可能性のあるコンテンツ(地政学的緊張を悪化させ、重大な危害を及ぼす可能性のある不正確に翻訳された動画の字幕。政府関係者の死を装うために技術的に操作(前後関係を見落としてクリップを切り抜く以上の操作が多い)された動画。重大な危害を及ぼす可能性のある事件を捏造するために技術的に操作(前後関係を見落としてクリップを切り抜く以上の操作が多い)された動画。) - 虚偽のコンテンツ: 過去の事象の古い映像を現在の事象のものであると虚偽の主張をすることで、重大な危害を及ぼす可能性のあるコンテンツ(実際は別の地域や事象に関するコンテンツを、特定の地域の人権侵害の記録として不正確に提示したコンテンツ。実際は数年前の映像であるにもかかわらず、そのコンテンツが現在の事象のものであるという虚偽の主張とともに、抗議活動への参加者に対する軍事的弾圧を示したコンテンツ。) ■Meta「偽情報」ポリシー https://transparency.meta.com/ja-jp/policies/community-standards/misinformation/ - 実際の危害や暴力: 弊社は、人々に対する差し迫った暴力または実際の危害のリスクに直接つながる可能性が高いと、専門家のパートナーが判断した偽情報および検証できない噂を削除します。偽情報とは、信頼できる第三者が虚偽であると判断する主張を含むコンテンツと定義されます。検証できない噂とは、専門家のパートナーによる情報元の特特定が極めて困難であるか不可能な主張、信頼できる提供元がない主張、その内容を証明するための具体性が不十分な主張、またはその内容があまりに信じがたい、もしくは不合理で信用できない主張と定義されます。弊社は、一見無害に思われる偽情報でも、特定の文脈では、死や深刻な怪我といった実際の危害の高いリスクにつながる可能性がある暴力的脅威など、オフラインでの危害のリスクにつながる場合があることを認識しています。弊社は、こうした各地域の動向について専門知識を有する非政府組織(NGO)、非営利団体、人道支援組織、国際機関のグローバルネットワークと連携しています。社会的暴力のリスクが高まっている国では、どの虚偽の主張が差し迫った身体的危害のリスクに直接つながるかを理解するために、現地のパートナーと積極的に協力しています。その上で、弊社のプラットフォームでそのような主張を
--	---

	するコンテンツを特定し削除します。例えば現地の専門家と協議して、文脈から切り離して虚偽の主張をし、暴力行為、暴力の被害者や加害者、武器や軍事用具を描写するメディアを削除する可能性があります。) 有害な健康関連の偽情報:弊社では、公衆衛生と安全に対する差し迫った危害に直接的な害をもたらす可能性の高い健康誤情報を特定するために、主要な保健機関と協議しています。弊社が削除する有害な健康関連の偽情報には次のようなものがあります。ワクチンに関する偽情報(弊社は、主にワクチンに関する偽情報について、保健当局がその情報は虚偽であり、差し迫ったワクチン接種の拒否を促進する可能性が高いと判断した場合は、その情報を削除します。)、公衆衛生上の緊急事態の間の偽情報(偽情報については、公衆衛生当局が、その情報が虚偽であり、差し迫った身体的危害のリスクに直接つながる可能性が高いと判断した場合、公衆衛生上の緊急事態の間、弊社はその情報を削除します。これには、個人が有害な疾病に感染したり、それを拡散させたりするリスクや、関連するワクチンを拒否するリスクにつながるものが含まれます。弊社では、世界と地域の保険機関と連携して公衆衛生上の緊急事態を特定しています。現在対象となるものとしては、ウイルスに関する緊急事態宣言が引き続き公に発表されている国における、新型コロナウイルス感染症関連の特定の虚偽の主張などがあります。新型コロナウイルス感染症およびワクチンについて、弊社が許可していない偽情報の種類をまとめたルール一覧についてはこちらをクリックしてください。)、健康上の問題に対する有害な「奇跡的な治療法」の宣伝または擁護(これには、医療の観点から推奨される使用法が、重傷または死亡のリスクに直接つながる可能性が高く、正当な医療上の使用法がない治療が含まれます(例: 漂白剤、消毒剤、黒色軟膏、苛性ソーダ)。) 投票者または国勢調査への干渉:選挙や国勢調査の健全性を促進する取り組みとして、弊社は、このようなプロセスに人々が参加する能力を妨害するリスクに直接つながる可能性が高い偽情報を削除します。これには次のようなものが含まれます。投票や有権者登録、国勢調査への参加に関する日付、場所、時間、方法に関する偽情報、投票できる人、投票資格、投票の有効性、投票するために提供しなければならない情報や書類に関する偽情報、候補者が立候補するか否かに関する偽情報、国勢調査への参加資格や、国勢調査へ参加するために提供しなければならない情報や書類に
--	--

	関する偽情報、個人の国勢調査の情報が別の(国勢調査の実施を担当していない)政府機関と共有されると述べること(ただし該当する場合)など、政府の国勢調査への関与についての偽情報、米国移民関税執行局(ICE)が投票所に待機しているとの偽りのまたは根拠のない主張、投票プロセスに参加すると新型コロナウイルス感染症(または別の伝染病)に感染するとする明確な偽の主張、選挙当局によって検証された、米国の投票所の現状に関して投票ができなくなるような偽りの主張。弊社では、暴力の呼びかけ、違法な参加の促進、選挙への協調的干渉の呼びかけを対象とする追加ポリシーを設けており、これらは弊社のコミュニティ規定の他のセクションに記載されています。 - 加工されたメディア:メディアはさまざまな方法で編集することができます。多くの場合、このような変更は、芸術的な理由による切り取りもしくは短縮、または音楽の追加など、無害なものです。ただ、特に動画コンテンツの場合、中には加工の有無がはっきりせず誤解を与えてしまう可能性が生じることもあります。このようなコンテンツはすぐに拡散されやすく、専門家の助言によれば、加工されたメディアに関する誤った思い込みはさらなる議論で修正できないことが多々あるため、弊社ではこのようなコンテンツを削除します。次の一定の条件を満たす動画については、本ポリシーに基づき削除します。 (1)明瞭さや画質・音質の調整にとどまらず、動画の被写体が実際には発言していない言葉を使ったと一般の利用者に誤解させるような編集または合成が、一般の利用者にはわからない形で行われている動画、および(2)ディープラーニング技術を含む人工知能または機械学習(例: 人為的なディープフェイク)によって制作され、特定の動画に他のコンテンツを統合、結合、置換、重ね合わせるなどして動画が本物に見えるように作成されている動画。 - 偽・誤情報ポリシー等に違反していないにもかかわらず、プラットフォームの信頼性と完全性(authenticity and integrity)を損なう誤情報については、独立した第三者ファクトチェック・パートナーのネットワークと協力して、誤情報の配信を減らし、強力な警告ラベルを表示し、誤情報に出くわした人、それを共有しようとした人、またはすでに共有した人に通知するというアプローチをとっている。1つのファクトチェックに基づき、誤りであることを暴いたストーリーと重複するものを特定する類似性検出方法を発動し、特定されたポストに対して、フィード上の流通を減らし、警告ラベルを表示し、利用者に通知するという同じペナルティを適用することができる。
--	---

■TikTok「コミュニティガイドライン」

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

- ・ 意図にかかわらず、個人や社会に重大な危害を及ぼし得る不正確な、誤解を招く、または虚偽のコンテンツ(重大な危害には、身体的、心理的、または社会的危害、および物的損害が含まれる。商業上の損害や風評被害はこれには含まれず、単なる不正確な情報や作り話も対象外)
- ・ 公共の安全に危険を及ぼしたり、危機または緊急事態についてパニックを引き起こし得たりする誤情報(以前に行われた攻撃の過去の映像を現在進行中であるかのように使用することや、特定の場所で基本的な生活必需品(食料や水など)が在庫切れを起こしていると誤った主張をすることなど)
- ・ 誤った医療情報(ワクチンに関する誤解を招く発言、生命を脅かす疾患に対して適切な治療を受けることを思いとどまらせる不正確な医療上の助言、公衆衛生に危険を及ぼすその他の誤情報など)
- ・ 確立した科学的コンセンサスを阻害する気候変動に関する誤情報(気候変動やその一因となっている要素の存在を否定することなど)
- ・ 暴力的またはヘイトに満ちた危険な陰謀論(暴力的な行動を呼びかける、過去の暴力行為と関連付ける、十分に立証されている暴力事件を否定する、保護属性を持つグループに対する偏見を引き起こすなど)
- ・ 個人を名指しして攻撃する、特定の陰謀論
- ・ 現実世界の出来事について人に誤解を与えるようなやり方で編集、接合、または合成された素材(動画や音声など)
- ・ 根拠がなく、特定の出来事や事態が「政府」や「秘密結社」などの秘密の集団や強力な集団によって引き起こされたとする一般的な陰謀論
- ・ 詳細がまだ明らかになっていない緊急事態や展開中の出来事に関連する未確認情報
- ・ ファクトチェックの審査中の、危険性が高い可能性のある誤情報

■LinkedIn「プロフェッショナルコミュニティポリシー」

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

- ・ 虚偽または誤解を招く内容。事実として提示された特定の主張が、明らかに虚偽であるか実質的

に誤解を招くものであり、害を及ぼす恐れがある場合は、これを削除。虚偽であるか実質的に誤解を招くものではあるものの、害を及ぼす恐れがないコンテンツの場合、投稿者のネットワーク外での配信は認められない。

■LinkedIn ヘルプセンター「虚偽または誤解を招くコンテンツ」

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

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

	イルスとその変異株が根絶され、パンデミックは終息した、COVID-19 の症状、死亡率、感染率は季節性インフルエンザよりも深刻なものではない、COVID-19 に感染して、死亡または重症化した人はいない) ○ 一部の国内事業者及び全ての国外事業者は、<u>偽・誤情報ポリシー等に定める禁止事項に違反した場合や同ポリシー等違反でない場合</u>における偽・誤情報への対応として、<u>モデレーション等の具体的な対応方法</u>を回答。 【モデレーション等 [例] ・ 投稿の削除 ・ アカウントの停止・削除 ・ 再取得したアカウントにおける投稿制限 ・ サービスの利用停止 ・ 他のサービスや機能へのアクセス制限 ・ プロフィールの編集要請 ・ 第三者が閲覧又はアクセスできない状態化 ・ 表示の降格 ・ 検索結果の上位表示の回避 ・ おすすめやトレンド等の対象外化 ・ 投稿時の注意喚起や警告ラベルの表示 ・ 閲覧を望まない第三者への非表示 ・ 第三者に対する変化の激しい出来事ガイド、信憑性未確認ラベルやファクトチェック結果の表示 ○ 一部の国内事業者及びほぼ全ての国外事業者は、<u>偽・誤情報に対するコンテンツモデレーション等についての考え方</u>や、<u>コンテンツモデレーション等のうち投稿の削除の対象となる偽・誤情報についての例示等</u>を回答。 【考え方 [例] ・ 偽情報である旨が明らかである投稿を禁止。
--	--

- 特定の種類の誤解を招くコンテンツまたは虚偽が含まれるコンテンツで、重大な危害を及ぼす可能性のあるものは許可されない。これには、現実の世界で危害を与える可能性がある特定の種類の誤った情報、技術的に操作された特定の種類のコンテンツ、民主的な手続きを妨害するコンテンツが含まれる。
- 偽情報は、包括的な禁止事項を明示する方法がなく、他の種類の発言とは異なる。例えば、過度な暴力描写やヘイトスピーチについて禁じている発言はポリシーに規定されているため、そのポリシーに賛成しない人でも従うことができるが、偽情報については、そのような方針を提供することができず、世界は絶え間なく変化し続けているため、ある時点では真実であっても、次の瞬間には真実でなくなることがあり、また、自身の周りの世界について異なるレベルの情報を有しているため、真実でない情報も真実だと信じてしまうことがある。ポリシーとして、偽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ダンスを設けるよう努めている。
- グローバルなコミュニティでは人々がさまざまな意見を持つのは自然なことであるが、事実や現実の共有に基づいて行動するべく努めているため、個人や社会に重大な危害を及ぼし得る不正確な、誤解を招く、または虚偽のコンテンツは、意図にかかわらず許可しない。重大な危害には、身体的、心理的、または社会的危害、および物的損害が含まれ、商業上の損害や風評被害はこれには含まれず、単なる不正確な情報や作り話も対象外となる。

【投稿の削除の対象に関する例示等【例】】

- 国勢調査の妨害: 国勢調査に関する時間、場所、方法、資格要件について参加者を誤解させることを目的としたコンテンツ、または国勢調査を著しく妨げる可能性のある虚偽の主張(国勢調査の参加方法に関して誤った手順を示す。回答者の在留資格が法執行機関に報告されるという誤った主張により、国勢調査への参加を妨げる行為。)
- 改ざんされたコンテンツ: ユーザーの誤解を招くように技術的に操作または改ざんされ(前後関係を見捨ててクリップを切り抜く以上の操作が多い)、重大な危害を及ぼす可能性のあるコンテンツ(地政学的緊張を悪化させ、重大な危害を及ぼす可能性のある不正確に翻訳された動画の字幕。政府関係者の死を装うために技術的に操作(前後関係を見捨ててクリップを切り抜く以上の操作が多い))

	された動画。重大な危害を及ぼす可能性のある事件を捏造するために技術的に操作(前後関係を見無視してクリップを切り抜く以上の操作が多い)された動画。) - 虚偽のコンテンツ: 過去の事象の古い映像を現在の事象のものであると虚偽の主張をすることで、重大な危害を及ぼす可能性のあるコンテンツ(実際は別の地域や事象に関するコンテンツを、特定の地域の人権侵害の記録として不正確に提示したコンテンツ。実際は数年前の映像であるにもかかわらず、そのコンテンツが現在の事象のものであるという虚偽の主張とともに、抗議活動への参加者に対する軍事的弾圧を示したコンテンツ。) - 実際の危害や暴力: 弊社は、人々に対する差し迫った暴力または実際の危害のリスクに直接つながる可能性が高いと、専門家のパートナーが判断した偽情報および検証できない噂を削除します。偽情報とは、信頼できる第三者が虚偽であると判断する主張を含むコンテンツと定義されます。検証できない噂とは、専門家のパートナーによる情報元の特定制が極めて困難であるか不可能な主張、信頼できる提供元がない主張、その内容を証明するための具体性が不十分な主張、またはその内容があまりに信じがたい、もしくは不合理で信用できない主張と定義されます。弊社は、一見無害に思われる偽情報でも、特定の文脈では、死や深刻な怪我といった実際の危害の高いリスクにつながる可能性がある暴力的脅威など、オフラインでの危害のリスクにつながる場合があることを認識しています。弊社は、こうした各地域の動向について専門知識を有する非政府組織(NGO)、非営利団体、人道支援組織、国際機関のグローバルネットワークと連携しています。社会的暴力のリスクが高まっている国では、どの虚偽の主張が差し迫った身体的危害のリスクに直接つながるかを理解するために、現地のパートナーと積極的に協力しています。その上で、弊社のプラットフォームでそのような主張をするコンテンツを特定し削除します。例えば現地の専門家と協議して、文脈から切り離して虚偽の主張をし、暴力行為、暴力の被害者や加害者、武器や軍事用具を描写するメディアを削除する可能性があります。) - 有害な健康関連の偽情報: 弊社では、公衆衛生と安全に対する差し迫った危害に直接的な害をもたらす可能性の高い健康誤情報を特定するために、主要な保健機関と協議しています。弊社が削除する有害な健康関連の偽情報には次のようなものがあります。ワクチンに関する偽情報(弊社は、主
--	--

	にワクチンに関する偽情報について、保健当局がその情報は虚偽であり、差し迫ったワクチン接種の拒否を促進する可能性が高いと判断した場合は、その情報を削除します。)、公衆衛生上の緊急事態の間の偽情報(偽情報については、公衆衛生当局が、その情報が虚偽であり、差し迫った身体的危害のリスクに直接つながる可能性が高いと判断した場合、公衆衛生上の緊急事態の間、弊社はその情報を削除します。これには、個人が有害な疾病に感染したり、それを拡散させたりするリスクや、関連するワクチンを拒否するリスクにつながるものが含まれます。弊社では、世界と地域の保険機関と連携して公衆衛生上の緊急事態を特定しています。現在対象となるものとしては、ウイルスに関する緊急事態宣言が引き続き公に発表されている国における、新型コロナウイルス感染症関連の特定の虚偽の主張などがあります。新型コロナウイルス感染症およびワクチンについて、弊社が許可していない偽情報の種類をまとめたルール一覧についてはこちらをクリックしてください。)、健康上の問題に対する有害な「奇跡的な治療法」の宣伝または擁護(これには、医療の観点から推奨される使用法が、重傷または死亡のリスクに直接つながる可能性が高く、正当な医療上の使用法がない治療が含まれます(例: 漂白剤、消毒剤、黒色軟膏、苛性ソーダ)。) 投票者または国勢調査への干渉: 選挙や国勢調査の健全性を促進する取り組みとして、弊社は、このようなプロセスに人々が参加する能力を妨害するリスクに直接つながる可能性が高い偽情報を削除します。これには次のようなものが含まれます。投票や有権者登録、国勢調査への参加に関する日付、場所、時間、方法に関する偽情報、投票できる人、投票資格、投票の有効性、投票するために提供しなければならない情報や書類に関する偽情報、候補者が立候補するか否かに関する偽情報、国勢調査への参加資格や、国勢調査へ参加するために提供しなければならない情報や書類に関する偽情報、個人の国勢調査の情報が別の(国勢調査の実施を担当していない)政府機関と共有されると述べる事(ただし該当する場合)など、政府の国勢調査への関与についての偽情報、米国移民関税執行局(ICE)が投票所に待機しているとの偽りのまたは根拠のない主張、投票プロセスに参加すると新型コロナウイルス感染症(または別の伝染病)に感染するとする明確な偽の主張、選挙当局によって検証された、米国の投票所の現状に関して投票ができなくなるような偽りの主張。弊社では、暴力の呼びかけ、違法な参加の促進、選挙への協調的干渉の呼びかけを対象とする追加ポリシーを設
--	--

	けており、これらは弊社のコミュニティ規定の他のセクションに記載されています。 - 加工されたメディア:メディアはさまざまな方法で編集することができます。多くの場合、このような変更は、芸術的な理由による切り取りもしくは短縮、または音楽の追加など、無害なものです。ただ、特に動画コンテンツの場合、中には加工の有無がはっきりせず誤解を与えてしまう可能性が生じることもあります。このようなコンテンツはすぐに拡散されやすく、専門家の助言によれば、加工されたメディアに関する誤った思い込みはさらなる議論で修正できないことが多々あるため、弊社ではこのようなコンテンツを削除します。次の一定の条件を満たす動画については、本ポリシーに基づき削除します。 (1)明瞭さや画質・音質の調整にとどまらず、動画の被写体が実際には発言していない言葉と言ったと一般の利用者に誤解させるような編集または合成が、一般の利用者にはわからない形で行われている動画、および(2)ディープラーニング技術を含む人工知能または機械学習(例: 人為的なディープフェイク)によって制作され、特定の動画に他のコンテンツを統合、結合、置換、重ね合わせるなどして動画が本物に見えるように作成されている動画。 - 公共の安全に危険を及ぼしたり、危機または緊急事態についてパニックを引き起こし得たりする誤情報(以前に行われた攻撃の過去の映像を現在進行中であるかのように使用することや、特定の場所で基本的な生活必需品(食料や水など)が在庫切れを起こしていると誤った主張をすることなど) - 誤った医療情報(ワクチンに関する誤解を招く発言、生命を脅かす疾患に対して適切な治療を受けることを思いとどまらせる不正確な医療上の助言、公衆衛生に危険を及ぼすその他の誤情報など) - 確立した科学的コンセンサスを阻害する気候変動に関する誤情報(気候変動やその一因となっている要素の存在を否定することなど) - 暴力的またはヘイトに満ちた危険な陰謀論(暴力的な行動を呼びかける、過去の暴力行為と関連付ける、十分に立証されている暴力事件を否定する、保護属性を持つグループに対する偏見を引き起こすなど) - 個人を名指しして攻撃する、特定の陰謀論 - 現実世界の出来事について人に誤解を与えるようなやり方で編集、接合、または合成された素材(動画や音声など)。
--	--

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

	の時々の大衆の興味に合わせたものになるという理由で<u>偽・誤情報について内容による類型定義を行なっていない旨、類型に応じて対応方法を変えずに投稿や行為の悪質性・公益性・投稿者のサービス利用状況に応じて偽・誤情報への対応を決定している旨</u>を回答。 ● ほぼ全ての国内事業者及び全ての国外事業者は、<u>偽・誤情報ポリシー等違反における類型ごとの具体的な対応方法について、コンテンツモデレーション等のうち投稿の削除の対象となる偽・誤情報の例示や同ポリシー等違反を繰り返す場合のアカウント停止等に関する基準等を除き、どのような場合にどのモデレーション等の対象となるかや公共の利益等の観点から例外としてコンテンツモデレーション等の対象とならないか等に関する具体的な考え方や基準に関する回答なし。</u> ● 一部の国外事業者は、<u>偽・誤情報が同ポリシー等に定める禁止事項ではない場合について、おすすめの対象外化等の対象となる偽・誤情報の例示や自らのプラットフォームの信頼性と完全性を損なう場合や信憑性の低い場合における第三者のファクトチェック機関との連携を除き、どのような場合にどのモデレーション等の対象となるか等に関する具体的な考え方や基準に関する回答なし。</u>また、<u>偽・誤情報ポリシー等がグローバルなものであり、言語や地域等を問わずに施行される旨</u>を回答。 【偽・誤情報ポリシー等の策定・公表状況等】(参考資料19-1-1 及び 19-1-2Q3-1、Q3-2) ○ 一部の国内事業者及び国外事業者は、<u>偽・誤情報ポリシー等について、定期的な見直しと第三者によるレビューの両方の実施、また、ほぼ全ての国内事業者及び国外事業者は、第三者機関によるレビューの実施</u>を回答。 ● ほぼ全ての国内事業者及び国外事業者は、<u>ポリシー等の定期的な見直しに関する回答なし。</u>また、一部の国外事業者は、<u>偽・誤情報ポリシー等がグローバルなものであり、言語や地域等を問わずに施行される旨</u>を回答。 【偽・誤情報ポリシー等の見直し・外部有識者等によるレビュー状況】(参考資料19-1-1 及び 19-1-2 Q3-3)
3 偽・誤情報の流通・拡散に	○ 全ての国内事業者及び国外事業者は、 <u>法令違反やポリシー等に違反するコンテンツ</u> について、

対するモデレーション等の 手続・体制	権利を侵害されている者及び発信者（投稿者）以外の<u>第三者からの日本語による通報を受け付ける窓口（以下「第三者通報受付窓口」）の設置</u>を回答。 ○ 一部の国内事業者及び国外事業者は、<u>公的機関によるファクトチェック済みの情報を添える等により、偽・誤情報ポリシー等に違反する偽・誤情報を選択して通報が可能</u>であることを回答。 ● 一部の国内事業者及び多くの国外事業者は、<u>サービスの ID やアカウントを取得しているユーザやサービスにログイン可能なユーザ等通報可能な主体に限定</u>がある旨を回答。 ● 一部の国内事業者及び多くの国外事業者は、<u>偽・誤情報ポリシー等に違反する偽・誤情報を選択して通報可能であるかどうかに関する回答なし</u>。 【日本語による第三者通報受付窓口の設置状況】(参考資料19-1-1及び19-1-2 Q5-1) ○ 全ての国内事業者は、第三者通報受付窓口について、<u>日本語通報への対応が可能な人数</u>を回答。 ● 全ての国外事業者からは、第三者通報受付窓口における日本語通報に対応可能な人数について、<u>随時変動しうることやセキュリティ及びビジネス上の理由から非公表とする等、明確な回答なし</u>。 【通報受付窓口において、日本語通報対応が可能な人数】(参考資料19-1-1 及び 19-1-2 Q5-2-(1)) ○ 全ての国内事業者及び一部の国外事業者は、日本語通報対応について、<u>AI その他の機械的手段は利用していない</u>と回答。 ● 一部の国外事業者は、<u>日本語通報対応における AI その他の機械的手段の利用について、当該手段の概要や利用手順（どのようなケースで用いるのか等）に関する明確な回答なし</u>。 【日本語通報対応における AI 等の利用状況】(参考資料19-1-1 及び 19-1-2 Q5-2-(3)) ○ 全ての国内事業者及び一部の国外事業者は、<u>日本語通報対応結果を通報者に通知等</u>していると回答。 ○ 半数の国内事業者及び一部の国外事業者は、<u>通報者からの不服申立等に対応</u>していると回答。 ● 多くの国外事業者は、<u>日本語通報対応結果を通報者に通知等しているかどうかに関する明確な</u>
-------------------------------	--

	<u>回答なし。</u> ● 半数の国内事業者及び多くの国外事業者は、<u>通報者からの不服申立等に対応しているかどうかに関する明確な回答なし。</u> 【日本語通報対応結果の第三者通報者への通知等状況】(出典:参考資料19-1-1 及び 19-1-2 Q5-2-(6)) 【日本語通報対応結果に関する第三者通報者からの不服申立等の専用窓口等の状況】(参考資料19-1-1 及び 19-1-2 Q5-2-(7)) ○ ほぼ全ての国内事業者は、<u>第三者通報受付窓口を通じた日本語による通報をどの程度の期間で処理しているか</u>について、<u>24時間以内や2営業日以内等の目標期間</u>を回答。 ● 全ての国外事業者は、処理の目標期間について、<u>報告される問題やトピックの複雑さによって異なることや非公表とする等、明確な回答なし。</u> 【日本語による通報への処理期間の目標の設定状況】(参考資料19-1-1 及び 19-1-2 Q5-2-(5)) ○ 一部の国内事業者及び多くの国外事業者は、<u>政府機関（法務省、警察、違法・有害情報相談センター（総務省）等）やNGO/NPO など特定の第三者通報主体からの通報の優先的取扱いの実施</u>を回答。 ● ほぼ全ての国内事業者及び一部の国外事業者は、<u>政府機関やNGO/NPO など特定の第三者通報主体からの通報の優先的取扱いは実施していないことを回答。</u> ● 全ての国内事業者及び国外事業者は、偽・誤情報ポリシー等に違反する偽・誤情報について、<u>特定の第三者通報主体からの通報の優先的取扱いを実施しているかどうかに関する回答なし。</u> 【特定の第三者通報主体からの日本語による通報の優先的取扱い】(出典:参考資料19-1-1 及び 19-1-2 Q5-2-(8)) ○ 一部の国内事業者及び国外事業者は、<u>24時間365日稼働で違反投稿のパトロール・検知・判定、投稿した瞬間と一定の再生回数に達した段階で検知・対応、随時の目視確認、技術による自動的な研修と人による組み合わせ等</u>により、自社で検知・対応していると回答。 ● 一部の国内事業者及び国外事業者は、<u>当事者や関係者からの問い合わせや通報等、ファクトチ</u>
--	---

	<u>エック団体との連携</u>により、検知・対応していると回答。 【第三者通報を待たずに自社による検知・対応状況】(参考資料19-1-1 及び 19-1-2 Q5-3(1)) ○ 一部の国内事業者及び多くの国外事業者は、<u>関連キーワードの抽出・自動表示、システムと人間の両方による検知・対応、自動モデレーション</u>において、<u>A I 等の機械的手段</u>を利用していると回答。 ● 多くの国内事業者及び一部の国外事業者は、<u>A I 等の機械的手段</u>を利用していないと回答。 【自社による検知・対応におけるAI等の機械的手段の利用状況】(参考資料19-1-1 及び 19-1-2 Q5-3(2)) ○ ほぼ全ての国内事業者及び一部の国外事業者は、自社で検知してからモデレーション等を実施するまでの目標期間について、<u>作業上の目安を設定し事案に応じた対応を行う旨、検知後速やかに対応する旨や検知・対応の結果 24 時間以内に削除された動画の割合を公表</u>（23 年 7～9 月期に日本で削除された動画の総数に占める割合は 76.8%）している旨等を回答。 ● 一部の国内事業者及びほぼ全ての国外事業者は、処理の目標期間について、<u>明確な回答なし</u>。 【自社による検知・対応における処理期間の目標の設定状況】(参考資料19-1-1 及び 19-1-2 Q5-2-(5)) ○ 一部の国内事業者及び国外事業者は、<u>透明性レポートとしての定期的な公表や自社のウェブページでの掲載</u>と回答。 ● ほぼ全ての国内事業者及び国外事業者は、<u>非公開等と回答</u>。 【第三者通報対応や自社による検知・対応の公開状況】(参考資料19-1-1 及び 19-1-2 Q5-4)
4 偽・誤情報の流通・拡散への対応状況	○ 半数の国内事業者及びほぼ全ての国外事業者は、偽・誤情報ポリシー等への違反を理由として<u>投稿の削除や非表示等</u>を行った日本国内における<u>全体の件数</u>を回答。 ● 全ての国内事業者及び国外事業者は、偽・誤情報ポリシー等への違反を理由としたコンテンツモデレーション等について、<u>投稿の削除以外のモデレーション等</u>を行った日本国内における<u>件数についての明確な回答なし</u>。 【偽・誤情報ポリシー等違反に関するモデレーション等の実施件数】(参考資料19-1-1 及び19-1-2 Q6-1(1))

○ 一部の国内事業者及び国外事業者は、偽・誤情報ポリシー等への違反を理由として投稿の削除等を行った日本国内における全体の件数のうち、A I 等の機械的手段のみによって検知・対応した件数又はA I 等の機械的手段と人間による組み合わせにより検知・対応した件数を回答。

● ほぼ全ての国内事業者及び国外事業者は、件数を取得していないことや公表していないこと等、明確な回答なし。

【AI 等の機械的手段を利用して検知した件数】(参考資料19-1-1 及び19-1-2 Q6-1-(2)~(3))

○ 一部の国外事業者からは、A I 等の機械的手段のみによって検知・対応した検知した結果に誤りがあることが事後的に判明した件数を回答。

● 全ての国内事業者及びほぼ全ての国外事業者は、件数を取得していないこと、削除済みコメントを全件チェックすることは困難であることや公表していないこと等、明確な回答なし。

【AI等の機械的手段による判定結果に誤りがあることが事後的に判明した件数】(参考資料19-1-1 及び19-1-2 Q6-1-(4))

○ 一部の国内事業者は、第三者通報の総数、第三者通報を契機としたモデレーション等の実施件数、第三者通報の受付からモデレーション等実施までの平均期間を回答。

● 全ての国外事業者は、第三者通報の総数、第三者通報を契機としたモデレーション等の実施件数、第三者通報の受付からモデレーション等実施までの平均期間について、また、全ての国内事業者及び国外事業者は、モデレーション等の有無・内容に関する通報者からの不服申立等の件数について、明確な回答なし。

【第三者通報の総数、それを契機としたモデレーション等の実施件数、その受付からモデレーション等実施までの期間、モデレーション等の有無・内容に関する通報者からの不服申立等の件数】(参考資料19-1-1 及び19-1-2 Q6-1-(9)~(12))

○ 一部の国内事業者及び国外事業者は、モデレーション等の対象となった投稿者（発信者）からの日本語による苦情や不服申立てを受け付ける窓口（以下「苦情等受付窓口」）を通じた苦情・不服申立て件数、そのうち実際にモデレーション等の撤回につながった件数を回答。

	● <u>ほぼ全ての国内事業者及び国外事業者は、苦情等受付窓口を通じた苦情・不服申立て件数、そのうち実際にモデレーション等の撤回につながった件数等について、明確な回答なし。</u> 【投稿者(発信者)からの日本語による苦情・不服申立て件数、苦情・不服申立てに基づくモデレーション等を撤回した件数、苦情・不服申立ての受付からモデレーション等の撤回までの期間、苦情・不服申立ての再審査要求の件数】(参考資料19-1-1 及び19-1-2 Q6-1(13)~(17)) ● <u>全ての国内事業者及び国外事業者は、投稿者(発信者)からの日本語による苦情・不服申立てのうち特定の発信者(投稿者)による苦情・不服申立ての件数等について、明確な回答なし。</u> 【投稿者(発信者)からの日本語による苦情・不服申立て件数のうち特定の発信者(投稿者)による苦情・不服申立ての件数、当該苦情・不服申立ての受付からモデレーション等の撤回までの期間、モデレーション等の有無・内容に関する発信者(投稿者)・通報者との間で訴訟等に発展した件数】(参考資料19-1-1 及び 19-1-2 Q6-1(18)~(20))
5 偽・誤情報の発信者(投稿者)の表現の自由等への配慮	○ <u>ほぼ全ての国内事業者及び全ての国外事業者は、発信者(投稿者)のマイページへの訪問、登録メールアドレスへの通知、アプリ内通知やカスタマーサービスへの問い合わせ等により、投稿したコンテンツに違法性があると法的請求を受けたこと、削除されたコンテンツ、投稿の削除等になったことや違反したポリシー等削除理由等の確認ができる</u>と回答。 ● <u>全ての国内事業者及びほぼ全ての国外事業者は、投稿の削除等以外のコンテンツモデレーション等について、発信者(投稿者)への実施の事実等の通知等に関する明確な回答なし。</u> 【発信者(投稿者)に対するモデレーション等の実施の事実や理由の通知等状況】(参考資料19-1-1 及び19-1-2 Q4-1) ○ <u>全ての国外事業者は、苦情等受付窓口について、再審査請求含め発信者(投稿者)からの日本語による専用の窓口を整備・公表</u>と回答。 ● <u>全ての国内事業者は、一般的な相談等と共通の対応窓口にて対応</u>しており、専用の窓口は整備していない。 ● <u>全ての国内事業者及びほぼ全ての国外事業者は、投稿の削除以外のコンテンツモデレーション等に関する苦情等受付窓口に関する明確な回答なし。</u> 【発信者(投稿者)からの苦情等受付窓口の設置状況】(参考資料19-1-1 及び19-1-2 Q4-2)

	○ ほぼ全ての国内事業者は、苦情等対応受付窓口について、<u>日本語対応が可能な人数を回答。</u> ● 全ての国外事業者からは、<u>苦情等対応受付窓口における日本語通報に対応可能な人数について、随時変動しうることや非公表とする等、明確な回答なし。</u> 【苦情等受付窓口において、日本語通報対応が可能な人数】(参考資料19-1-1 及び 19-1-2 Q4-3(1)) ○ 全ての国内事業者及び一部の国外事業者は、苦情等受付対応について、<u>A I その他の機械的手段は利用していないと回答。</u> ● 一部の国外事業者は、<u>苦情等受付対応におけるA I その他の機械的手段の利用について、当該手段の概要や利用手順（どのようなケースで用いるのか等）に関する明確な回答なし。</u> 【苦情等受付対応におけるAI等の利用状況】(参考資料19-1-1 及び 19-1-2 Q4-3(2)) ○ ほぼ全ての国内事業者は、苦情等受付窓口を通じた日本語による苦情等をどの程度の期間で処理しているかについて、<u>24時間以内や2営業日以内等の目標期間</u>を回答。 ● 全ての国外事業者は、<u>処理の目標期間について、非公表とする等、明確な回答なし。</u> 【苦情等受付における処理期間の目標の設定状況】(参考資料19-1-1 及び 19-1-2 Q4-3(5)) ○ ほぼ全ての国内事業者及び一部の国外事業者は、<u>苦情等受付対応結果を発信者（投稿者）に通知等</u>していると回答。 ○ 一部の国内事業者及び一部の国外事業者は、<u>発信者（投稿者）からの再審査に対応</u>していると回答。 ● 多くの国外事業者は、<u>苦情等受付対応結果を発信者（投稿者）に通知等しているかどうかに関する明確な回答なし。</u> ● ほぼ全ての国内事業者及び多くの国外事業者は、<u>発信者（投稿者）からの再審査に対応しているかどうかに関する明確な回答なし。</u> 【苦情等受付対応結果の発信者（投稿者）への通知等状況】(参考資料19-1-1 及び 19-1-2 Q4-3-(6)) 【苦情等受付対応結果に関する発信者（投稿者）からの再審査制度の状況】(参考資料19-1-1 及び 19-1-2 Q5-2-(7))
--	---

	○ 一部の国内事業者及び国外事業者は、<u>政府機関や NGO/NPO など特定の発信者（投稿者）からの苦情等の優先的取扱いの実施</u>を回答。 ● ほぼ全ての国内事業者及び一部の国外事業者は、<u>政府機関や NGO/NPO など特定の発信者（投稿者）からの苦情等の優先的取扱いは実施していないこと</u>を回答。 ● 全ての国内事業者及び国外事業者は、<u>偽・誤情報ポリシー等に違反する偽・誤情報について、特定の発信者（投稿者）からの苦情等の優先的取扱いを実施しているかどうかに関する明確な回答なし</u>。 【特定の発信者（投稿者）からの苦情等の優先的取扱い】（参考資料19-1-1 及び 19-1-2 Q4-3-(8)）
6 レコメンデーションやコンテンツモデレーション等に関する透明性・アカウントビリティ確保に向けた取組状況	○ 多くの国内事業者及び国外事業者は、<u>A I 等の機械的手段の利用によるコンテンツモデレーション等の自動的な対応について、不適切コメント対策、ポリシー等違反可能性コンテンツの検出や同違反コンテンツの特定、ポリシー違反ではないが潜在的に有害な誤情報等のボーダーラインコンテンツのおすすめ制限、わけづ・不快・不適切な画像やコメントの抽出・判定等において考慮する要素や用いられている主なパラメータやアルゴリズム</u>を回答。 ○ 一部の国内事業者及び国外事業者は、A I 等の機械的手段によってコンテンツモデレーション等を自動的に対応する際に考慮する要素や用いられている主なパラメータやアルゴリズムについて、<u>透明性レポートや自社ウェブページ等で公開</u>と回答。 ● 多くの国内事業者及び国外事業者は、<u>コンテンツモデレーション等のアルゴリズムやパラメータの重み付け等の詳細な内容について、非公開等、明確な回答なし</u>。 【コンテンツモデレーション等を自動的に決定している場合の主なパラメータ等の状況】（参考資料19-1-1 及び19-1-2 Q8-1～Q8-2） ○ 全ての国内事業者及び多くの国外事業者は、<u>A I 等の機械的手段の利用によるレコメンデーションの自動的な対応について、建設的・注目コメントや信頼できる情報源からの高品質な情報の優先表示、ユーザの属性情報等による投稿者・使用者双方が満足する掲出マッチング、多様なコ</u>

	<u>メント表示、ニュースコンテンツのランキング等、A I 等の機械的手段によってレコメンデーションを自動的に対応する際に考慮する要素や用いられている主なパラメータやアルゴリズム</u>を回答。 ○ 全ての国内事業者及び多くの国外事業者は、A I 等の機械的手段によってレコメンデーションを自動的に対応する際に考慮する要素や用いられている主なパラメータやアルゴリズムについて、<u>自社ウェブページや利用規約等で公開</u>と回答。また、一部の国外事業者は、と回答。 ● <u>多くの国内事業者及び国外事業者は、レコメンデーションのアルゴリズムやパラメータの重み付け等の詳細な内容について、非公開等、明確な回答なし。</u> 【レコメンデーションを自動的に決定している場合の主なパラメータ等の状況】(参考資料19-1-1 及び19-1-2 Q8-3～Q8-4) ○ 半数の国内事業者は、<u>レコメンデーションに使用している要素の公開可能性や必要に応じたアルゴリズムの開示可能性</u>を回答。 ○ 一部の国外事業者は、プラットフォームやコンテンツモデレーション等のアルゴリズムについて、<u>秘密保持契約を締結した上で限られた研究者にはコードの一部公開や、研究者を含む第三者におすすめのタイムラインのアルゴリズムを Github で開示</u>を実施していると回答。 ● <u>全ての国内事業者は、コンテンツモデレーション等やレコメンデーションのアルゴリズムについて、研究機関等ふくめ第三者に開示していないと回答。</u>また、ほぼ全ての国外事業者は、<u>米国・欧州で連携している一部の研究者のみへの研究者向けAPIの公開を除き、研究者等の特定の第三者への開示に関する明確な回答なし。</u> 【アルゴリズムの特定の第三者への開示状況】(参考資料19-1-1 及び19-1-2 Q8-5、Q8-6)
7 令和6年能登半島地震関連の偽・誤情報の流通・拡散への対応状況	○ 一部の国内事業者は、令和6年度能登半島地震に関連して、<u>投稿の削除（削除基準としては、例えば、明示的に虚偽情報を拡散させようとする投稿や、その説を強く信じ込み、他者に対してもそれを信じさせようとする意図が感じられる投稿）・非表示（例えば、人工地震と断定し流布するような投稿や募金を募る行為について利用規約等に照らして違反行為と認められた場合）やアカウント停止等を実施した件数</u>を回答。

- ほぼ全ての国内事業者及び全ての国外事業者は、令和6年能登半島地震に関連して、投稿の削除等のコンテンツモデレーション等を行った日本国内における全体の件数について、明確な回答なし。

【能登半島地震におけるモデレーション等の取組状況】(参考資料19-1-1 及び 19-1-2 Q7-1)

- 一部の国内事業者は、令和6年度能登半島地震に関連して、投稿の削除対象とする根拠として、ファクトチェック機関により明確に誤りとされていることを根拠として採用したものと削除を実施した件数を回答。

- ほぼ全ての国内事業者及び全ての国外事業者は、令和6年能登半島地震に関連して、投稿の削除等におけるファクトチェック機関との連携や削除等を実施した件数について、明確な回答なし。

【能登半島地震に関連するコンテンツにおけるファクトチェック状況】(参考資料19-1-1 及び 19-1-2 Q7-2)

- 一部の国内事業者及び国外事業者は、能登半島地震に際し、チーム編成や特別な対策等による震災関連投稿のパトロール・モニタリングの強化（ポリシー等違反コンテンツの一斉確認、危険なキーワード・ハッシュタグのモニタリング、偽・誤情報が含まれていないか注意を払うよう審査員チームに対する指導、審査の精度向上のためのガイドラインの共有等）、危機管理プロトコルによる対応、最新情報まとめページや特設ページの作成、警鐘を鳴らすトピックや図解の掲載、地震関連のデマの打ち消しのトピックスの作成、専門家やジャーナリストによる注意喚起や記事の紹介、キー局のライブ配信、ローカル局や地方新聞が運営するチャンネル等の信頼できる情報を見つけやすくする施策、偽情報に関する注意喚起の掲載・若年層向けの啓発動画キャンペーンの紹介、自治体や政府機関との連携など、偽・誤情報の流通への対応を強化したことを回答。

- 多くの国内事業者及び一部の国外事業者は、既存人数で対応可能等、対応を強化せず、また明確な回答なし。

【能登半島地震におけるコンテンツモデレーション等に関する体制強化の状況】(参考資料19-1-1 及び19-1-2 Q7-3)

	○ 一部の国内事業者及び国外事業者は、能登半島地震に際し、<u>業界団体（SMAJ）において他事業者と連携した注意喚起（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）、ファクトチェック機関との連携、民間の気象関連機関との連携による情報発信、内閣府・総務省・警察庁や地方自治体との連携等</u>を実施したことを回答。 ● 多くの国内事業者及び国外事業者は、<u>関係機関等との連携状況について明確な回答なし。</u> 【能登半島地震における関係機関等との情報共有等の状況】(参考資料19-1-1 及び19-1-2 Q7-4) ○ 一部の国内事業者及び多くの国外事業者は、今後の地震等災害において、<u>ユーザへの注意喚起・リテラシー向上、政府の公表情報やファクトチェック機関によるファクトチェック記事の活用、信頼できる情報空間の在り方に関する情報発信、技術・ユーザーの報告・専門家やファクトチェック機関からのトレンドレポートによる誤情報の検出、地域の誤情報トレンドの積極的な監視や新たな誤情報に類似するコンテンツの積極的なチェック等モニタリング強化やNPOや行政機関との連携による緊急コミュニケーションのサポート等</u>の取組を実施することを回答。 ● 多くの国内事業者及び一部の国外事業者は、<u>今後の対応について、強化する予定はないこと等明確な回答なし。</u> 【災害に関連する偽・誤情報への今後の対応】(参考資料19-1-1 及び19-1-2 Q7-5) ○ 一部の国内事業者及び国外事業者は、<u>緊急時においては一見して明らかな虚偽であるか信頼できる機関からの情報提供がないと真偽の判断が困難であること、ファクトチェック能力を有していないことや投稿の件数が膨大であることから偽・誤情報であることを理由に記事を削除することが難しいこと、政府としての公式情報をワンストップで確認できるページがあれば正しい情報への誘導がしやすくなること、情報は常に変化し専門家や当局の見解が一致しないこともあり、正確で権威ある情報を得ることが難しい場合があること等</u>を回答。 ● 多くの国内事業者及び国外事業者は、特に回答なし。 【災害に関連する偽・誤情報への対応における支障】(参考資料19-1-1 及び19-1-2 Q7-6)
8 選挙時の偽・誤情報の流	<別途整理予定>

通・拡散への対応状況	
9 なりすましへの対応状況	<別途整理予定>
10 広告の質の確保への対応状況	<別途整理予定>
11 広告配信先の質の確保への対応状況	<別途整理予定>
12 発信者への広告収入分配状況	<別途整理予定>
13 AI・ディープフェイク技術への対応状況	<別途整理予定>
14 ファクトチェックの推進に向けた取組状況	<別途整理予定>
15 マスメディア（新聞・放送）との連携状況	<別途整理予定>
16 利用者のICTリテラシー向上に向けた取組状況	○ 多くの国内事業者及び国外事業者は、<u>小中学校や大学等の教育機関、安心なインターネット環境づくり等に関する普及啓発機関等と連携・協力した取組を実施</u>と回答。また、一部の国外事業者より、インターネットを取り巻く環境は日々変化しており、その変化に対応するために必要なクリティカルシンキングやスキルを身につける機会を提供することが重要である一方、教育現場でそのようなスキルを教えられる教員や保護者の数は限られており、このような取り組みを行う一部のファクトチェック団体を含む非営利団体は、財政的に不安定な傾向があることから、このような取り組みには、指導者のための研修の機会や財政的な支援が必要との回答。 ● 全ての国内事業者及び国外事業者は、<u>消費者団体・利用者団体と連携・協力した取組に関する明確な回答なし。</u> 【取組【例】】 ■教育機関との連携

	▪ 「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情報モラル教材を開発するとともに、全国の学校や自治体に 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。 ▪ GIGA スクール構想の展開にあわせ、「情報モラル」と「情報活用」の育成や向上を図るため、2022 年 7 月、新たな活用型情報モラル教材「GIGA ワークブック」の汎用版を開発し、同年 9 月から、全国の小中学校で活用いただけるよう無償提供を開始。教材を導入いただいた自治体と連携し、学校現場での教材活用のサポートとして教員の方々へのオンライン研修(無償)も開始。 ▪ LINE みらい財団におけるユーザ啓発活動(GIGA ワークブック)として、啓発教材(GIGA ワークブック)を導入することとした自治体(教育委員会)を通じて、地域内の全ての公立学校(小学校・中学校・高校)へ教材導入の推進を案内。 ▪ 大学と連携し、リテラシー教育授業を実施。 ▪ 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。このカリキュラムでは、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っています。これまでに約 10 万人の中高生に教材を提供。 ▪ 中高生 15,557 人と、中高生の教員 119 名を対象にインターネット利用について調査し、この結果を「中高生インターネット利用白書 2021」として公開。調査では、利用時間や目的、インターネット利用で感じるメリットやデメリット、そして実際に経験したトラブル等について調査。 ▪ YouTube では、誤情報・フェイクニュースに関する取り組みとして、総務省並びに国際大学 GLOCOM のご協力の元、2023 年 4 月に「ほんとかな？があなたを守る」というテーマのキャンペーンを実施。ユーザーに向けて、フェイクニュースが自分の日常に潜む問題であると感じくきっかけをすること、そして、情報との向き合い方について考える機会を提供することを目指し、若者層に人気の高い 9 組の YouTube クリエイターの協力を得て、3 つのメッセージ(フェイクニュースは身近に存在すること)、「ファクトチェックが重要であること」、「安易な拡散が人に迷惑をかけてしまうリスクに繋がりがねないこと」)を伝えるショート動画を作成。 ▪ 学校・教育委員会や、自治体が行う青少年向けのカンファレンス、財団が行う研究事業などにおいて
--	--

	て、パネリストや講演者として参加し、ICTリテラシー向上に向けた啓発活動。 ・ マイクロソフトは Future Learning Lab プロジェクトのメンバーとして GLOCOM (国際大学グローバル・コミュニケーション・センター) と連携し、FuLLのイベント等にスピーカーを派遣するなど、関係者との意見交換・連携強化に向けたディスカッションに参加。また、FuLLの取り組みとして、児童(小学校中学年・高学年)がオンライン上で自己のデータを適切に管理するために注意すべきこと等を学べる学習コンテンツを作成。作成されたコンテンツは、経済産業省のSTEM Libraryに掲載。 ■普及啓発機関との連携 ・ 「安心ネットづくり促進協議会」への参加。 ・ 高校生 ICT カンファレンスに参加し、各地の高校において、ICTリテラシー向上に関する講演。 ■その他 ・ ウェブサイト上に LINE Safety Center -LINE の安心安全ガイド-を掲載。 ・ LINE みらい財団において、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。 ・ LINE みらい財団において、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・教材作成等を実施。 ・ 一般社団法人セーフアーインターネット協会主催の「Disinformation 対策フォーラム」へ 参加し、有効な対策について多様なプレイヤーと議論。 ・ 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り組みを実施。 ・ 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。 ・ 総務省がセーフアーインターネットデーに合わせて公開したリテラシーコンテンツ紹介サイト「ネット&SNS よりよくつかって 未来をつくろう～ICT 活用リテラシー向上プロジェクト～」に協力。 ・ Grow with Google という私たちの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。家族で実践安心・安全なインターネット利用のためにできることでは、家族を中心としたユーザーを対象に、情報の正確さを確認するためのヒントを含むオンライントレーニングを提供。あわせて個人で実践
--	--

	編も公開。 ・ こどもたちがテクノロジーを安全に、かつ最大限に活用できるようになるために、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。子どもが楽しみながら自発的に学べるオンラインゲーム「インターランド」のほか、保護者と教師向けのさまざまなリソースを含む。 ・ 有識者が主催する全国こどもネットフォーラム 2023 を後援し、子どもがインターネットの利用について自ら議論し学ぶ場を設けることを支援。 ・ Google が主幹事となり、あらゆる方のスキル開発を支援する取り組み「日本リスキリングコンソーシアム」を立ち上げ。総務省、経済産業省の協力のもと、200 を超えるパートナーとともに様々なリスキリングプログラムを提供。 ・ Google では、外部専門家の知見を取り入れる形で Grow with Google 「はじめてのメディアリテラシー 情報と向き合うとき、子どもも大人もすべきこと」として 10 本のトレーニング動画を公開している他、5 組の YouTube クリエイターの協力も得てより広い層に訴求する啓発動画を制作。 ・ 2020 年より、NPO 法人企業と教育協会 (ACE) と共同でデジタルリテラシー教育プログラム「みんなのデジタル教室#wethinkdigital」を実施。「フェイクニュースの見抜き方」「デジタルアイデンティティを考える」の 2 つのモジュールで、これまでに全国 24000 人以上の中高生を対象に授業を実施。 ・ 2023 年 7 月、弊社は既存の 2 モジュールのリニューアルを発表。誤情報に関する総務省のリテラシー資産を統合し、新モジュール「デジタルシチズンシップと情報発信」を導入。このプログラムは総務省情報通信局の支援。 ・ 総務省と共同で政策討議：2022 年以降の ICT 利活用のためのリテラシー向上に関する研究会。ICT 利活用のためのリテラシー向上プロジェクトに貢献。 ・ 専門家やクリエイターと連携して、多様な啓発活動を実施。例えば、サイバーセキュリティ月間、若年層の性暴力被害予防月間、自殺予防週間、感染症ワクチン啓発、研究者と連携した「TikTok クリエイター向けメンタルヘルス講習会」、TikTok 悩み相談「性の悩み」「不登校」「親との関係」、選挙に関するデジタルリテラシーキャンペーン、親子向け啓発イベント、ペアレンタルコントロール利用の手
--	--

	引きの公開、クリエイターと連携した、安心安全のための啓発動画の制作。 - Microsoft は、育て上げネットなどの国内外の非営利団体と協力して、女性やサービスの行き届いていないコミュニティの若者など、日本社会のメンバーに責任ある AI とサイバーセキュリティを促進するためのコンテンツとスキルを提供。 - LinkedIn は、一般社団法人 1mm Innovation と協力して、女性が自信を取り戻し、雇用の障害を乗り越えて、経済的な安全と自立を確保できるよう支援するプログラムを策定。 - マイクロソフトは、ウェブサイトでリテラシー学習教材を提供。たとえば、“デジタル セーフティ”に関する専用サイトを公開し、教師が授業で使える教材を提供。“ゲーム安全性ツールキット”も提供しており、安全にゲームをする方法や Xbox の安全性に関する機能の使い方について親子で学ぶことが可能。 - マイクロソフトは、自社のプラットフォームやパートナーシップを活用し、信頼できる情報を検索、利用、共有する方法について消費者に情報を提供。検索の練習と検索コーチは Teams for Education に組み込まれている無料のツールで、効果的な検索条件を指定し、信頼できる情報源を特定して確認する方法について、学生にリアルタイムのコーチングを実施。The Investigators は教育版マイクラフトの新しい世界で、ゲームを利用した学習を通じて学生が情報リテラシーとメディア リテラシーを身に付けられるよう支援。このゲームと補助教材は英語で提供が開始されており、2024 年には（日本語を含む）28 か国語で公開され、世界中で数百万人の学生と教師が利用できるようになる予定。 - 生成 AI を責任をもって安全に活用するためのリテラシー習得を目的として、中高生を対象とした「AI Classroom ツールキット」や、教師や教育者等を対象とした「AI Toolkit for Trainers」などを提供。 - LinkedIn サイバーセキュリティの基礎知識 by Microsoft x LinkedIn: サイバーセキュリティの基礎知識 by Microsoft x LinkedIn を通じたリテラシー教材も提供。 【利用者の ICT リテラシー向上に向けた取組状況】（参考資料19-1-1 及び19-1-2 Q13-1～13-3）
17 研究開発の推進に向けた取組状況	<別途整理予定>
18 サイバーセキュリティ関	○ 一部の国内事業者及び国外事業者は、 <u>偽サイトや偽アカウントに関する情報共有等</u> において、

係機関等との連携状況	<u>日本国内のサイバーセキュリティ関係機関等と連携・協力した取組を実施</u>と回答。 ● <u>全ての国内事業者及び国外事業者は、偽・誤情報の流通・拡散への対応という観点において、日本国内のサイバーセキュリティ関係機関等と連携・協力した取組に関する明確な回答なし。</u> 【取組【例】】 - 関係団体(日本サイバー犯罪対策センター(JC3)やフィッシング対策協議会)へ会員として参加。対象サービスに限らず、当社に関わるフィッシングサイト情報や、偽サイトを含む脅威情報の提供を受領。また、当社で発見したフィッシングサイト情報については協議会経由で JPCERT/CC へ共有を行い、サイト閉鎖に向けた対応。 - 社内での調査、社内外からの CSIRT 窓口への通報により、当社に関連する SNS 等の偽アカウントの発見することもあり、偽・誤情報の流通・拡散への対応のため、停止に向けた措置を実施。 - IPA から脆弱性情報の共有(偽情報対策という面において特に連携はしていない) - 偽・誤情報の流通・拡散への対応に特化した取り組みではないが、CyberAgent CSIRT の活動として日本シーサート協議会、FIRST (Forum of Incident Response and Security Teams)に加入して情報収集。 - 外部のインテリジェンス専門機関と、サイバーセキュリティに関するモニタリング及び情報提供において連携。この連携には、TikTok のコミュニティを欺いたりする可能性のあるアカウント(例:選挙の信頼性に悪影響を及ぼす可能性のあるアカウント)を探知するための協力などを含む。 - Tech Against Terrorism のメンバーとして、暴力的過激主義者がプラットフォームを利用して害を及ぼすことを防ぐ取り組みを実施。 - マイクロソフトはさまざまな省庁・団体と協力し、オンラインのセキュリティと安全性の確保に努力。日本サイバー犯罪対策センター(JC3)、JPCERT、ICT-ISAC、FISCACAS、Software ISAC、フィッシング対策協議会、日本サイバーセキュリティ・イノベーション委員会(JCIC)等。 - 専用フォームを通じて報告が可能な日本政府や法執行機関と協力しており、これには日本の警察も含む。違法・有害情報相談センターは X に報告することができ、関連する問題やトピックについて情報提供や議論を行うために連携。
-------------------	--

	【サイバーセキュリティ関係機関等との連携状況】(参考資料19-1-1 及び19-1-2 Q15-1～15-3)
19 行政機関や地方公共団体等との連携状況	○ 半数の国内事業者及び全ての国外事業者は、偽・誤情報の流通・拡散への対応という観点にかぎらないものも含め、<u>偽情報等の流通状況や拡散の仕組みに係る分析、事業者に求められる偽情報等への対応に係る検討、事業者の事例の発表・共有、インターネット悪用事例に関する情報交換、違法・有害情報相談センターに基づく対応、若年層向けの啓発動画キャンペーン、ICT 利用のためのリテラシー向上プロジェクト、総務省のデジタルリテラシーキャンペーン、観光・広報・文化・伝統芸能・中小事業者支援・防災啓発・平和教育・気候変動啓発など多岐にわたる政策テーマにおいて広報啓発・プロモーション、オンラインにおけるセキュリティと安全性の確保、警察庁や各都道府県警等との情報共有やサービス・製品の説明、防災目的のために公共機関のアカウントに無料の公共サービス API を提供等、日本の行政機関や地方公共団体等と連携・協力した取組を実施と回答。</u> ● 半数の国内事業者は、<u>偽・誤情報の流通・拡散への対応という観点において、日本の行政機関や地方公共団体等と連携・協力した取組に関する明確な回答なし。</u>なお、支障となっている点として、<u>連携に関する具体的な情報不足、担当者の確保が難しいこと</u>を回答。 【取組【例】】 - 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係る分析や、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係る検討を実施。 - 総務省「ICT 活用のためのリテラシー向上に関する検討会」にオブザーバー参加し、事業者の事例を発表・共有。 - テロ等に関連するインターネット悪用事例に関する情報交換(GIFCT ワークショップ参加)。 - 違法・有害情報相談センターの「青少年案件情報提供スキーム」に基づく対応。 - 常日頃から総務省を含む様々な省庁の関係者と対話をし、また数多くの研究会にも委員ないしはオブザーバーとして参画。今回のようにヒアリングの要請にも対応。 - 総務省「ICT 活用のためのリテラシー向上に関する検討会」(座長:山本龍彦 慶應義塾大学大学院法務研究科教授)の取り組みの一環として、セーフアーインターネットデー(2月6日)に合わせて

	「ネット&SNS よりよくつかって 未来をつくろう～ICT 活用リテラシー向上プロジェクト～」が公開された際、コンテンツの提供者として協力。 ▪ 昨年4月には、GLOCOM 国際大学が主催しましたイベント「フェイクニュースと日本一私たちにできること・社会としてできることー(G7 デジタル・技術大臣会合関連イベント)」を支援。そのイベントに合わせて、YouTube では、総務省並びに国際大学 GLOCOM ご協力の元、フェイクニュースに惑わされないための若年層向けの啓発動画キャンペーンをローンチ。本キャンペーンは、若者に人気のある 9 名の YouTube クリエイターの協力を得て作成された、「フェイクニュースは身近に存在すること」「ファクトチェックが重要であること」そして、「安易な拡散が人に迷惑をかけてしまうリスクに繋がりがねないこと」この三つのメッセージを伝えるショート動画。YouTube で配信された各クリエイターのこれらの動画は、合計で 1500 万回以上再生(2024 年 3 月現在)。YouTube は、「情報に対するリテラシーを高める」ことも大切と考えており、今後も真摯に取り組む。 ▪ 2020 年より、NPO 法人企業と教育協会(ACE)と共同でデジタルリテラシー教育プログラム「みんなのデジタル教室#wethinkdigital」を実施。「フェイクニュースの見抜き方」「デジタルアイデンティティを考える」の 2 つのモジュールで、これまでに全国 24000 人以上の中高生を対象に授業を実施。2023 年 7 月、弊社は既存の 2 モジュールのリニューアルを発表。誤情報に関する総務省のリテラシー資産を統合し、新モジュール「デジタルシチズンシップと情報発信」を導入。このプログラムは総務省情報通信局の支援。 ▪ 総務省と共同で政策討議: 2022 年以降の ICT 利活用のためのリテラシー向上に関する研究会。ICT 利活用のためのリテラシー向上プロジェクトに貢献。 ▪ 2024 年 2 月、「セーフアーインターネットデー」に向けて、透明性ツールや、広告を含むパーソナライゼーションがプラットフォーム上でどのように機能するかについての利用者教育キャンペーンを開始。このキャンペーンを IG のクリエイターと共に作成し、また同日に開始された総務省のデジタルリテラシーキャンペーンとも連携。キャンペーン・ランディングページは総務省のウェブサイトでも紹介。 ▪ 多様化する行政課題の解決を支援するため、TikTok では地方公共団体や行政機関と連携のもと、観光、広報、文化・伝統芸能、中小事業者支援、防災啓発、平和教育、気候変動啓発など多岐にわ
--	---

	たる政策テーマにおいて広報啓発・プロモーション等の取り組みを推進。直近 3 か年では、2022 年 32 件、2023 年 42 件、2024 年 9 件の計 83 件の連携事例を実施。 - 令和 5 年 7 月九州北部豪雨災害や能登半島地震を支援するチャリティーLIVE 開催および寄付贈呈等の復興支援プロジェクト。 - 広島県との「県政コミュニケーションに係る連携及び協力に関する連携協定」の締結および毎年 8 月 6 日に開催する「平和記念式典」の LIVE 配信。 - 札幌市と「魅力発信の取り組みに関する連携協定」を締結。 - 選挙ドットコムと連携した統一地方選挙に合わせて選挙教育プロジェクト。 - 埼玉県と共同で「家族三世代で楽しめる TikTok セミナー」を開催。 - 各省庁とも連携したプロジェクトも多く実施。 - マイクロソフトは各省庁と協力し、オンラインにおけるセキュリティと安全性の確保に努力。例えば、総務省リードによるリテラシーキャンペーン、テロ等に関連するインターネット サービスの悪用に関する勉強会、マイクロソフトが有する脅威情報を各省庁に提供（GSP: 政府セキュリティ プログラム）、各省庁との月 1 回のセキュリティ勉強会、専門知識の提供や各種セキュリティ ガイドラインの策定支援（内閣サイバーセキュリティセンター（NISC）クラウド設定ガイドライン、経済産業省クラウド セキュリティ ガイドライン、金融情報システムセンター（FISC）金融システム監査基準など）、地方自治体と提携し、工業高校でのリテラシー習得支援事業を実施。 - 警察庁や各都道府県警、またその他省庁と随時連携。特定の事項に関する情報共有や、私たちのポリシー、製品（コミュニティノートを含む）についての説明など。 - X は防災目的のために公共機関のアカウントに無料の公共サービス API を提供。 【行政機関や地方公共団体等との連携状況】（参考資料19-1-1 及び19-1-2 Q16-1～16-2）
20 国際機関等との連携状況	○ 一部の国内事業者及び全ての国外事業者は、偽・誤情報の流通・拡散への対応という観点にかぎらないものも含め、<u>G7 デジタル・技術大臣会合関連イベント</u>、<u>G7 内務・安全担当大臣会合への対面での参加やプレゼンテーション</u>、<u>International Fact-Checking Network (IFCN) の支援</u>、<u>IFCN 主催の世界ファクトチェック会議 Global Fact 10 のスポンサー及びセッション開催</u>、

	<u>Global Network Initiative (GNI)・Global Internet Forum to Counter Terrorism (GIFCT)・Digital Trust and Safety Partnership (DTSP)・Tech Coalition 等の創設、「2024 年選挙における AI の欺瞞的使用に対抗するための技術協定」を 2 月 16 日に業界各社と締結、テクノロジー企業のための自主的な Election Integrity Guidelines に署名、コンテンツが AI を使用して作成されたことを示す共通の技術標準への協力、UNDP と連携したハラスメント及びヘイトスピーチ防止キャンペーン、ダボス会議等におけるプラットフォームの安全性について各国の政府関係者と意見交換、GIFCT・Rome Call for AI Ethics (人工知能の倫理的ガイドライン)・Paris Call for Trust and Security in Cyberspace (サイバー空間の信頼性と安全性のためのパリ・コール)・2024 年選挙における人工知能の欺まんの使用に対抗するための技術協定等マルチステークホルダーアプローチを主導等、国際機関等と連携・協力した取組を実施と回答。</u> ● <u>ほぼ全ての国内事業者は、偽・誤情報の流通・拡散への対応という観点において、国際機関等と連携・協力した取組に関する明確な回答なし。</u>なお、支障となっている点として、<u>連携に関する具体的な情報不足、担当者の確保が難しいこと、また、一部の国外事業者は、関連する技術の堅牢性は業界全体の課題で、可能な限り取り組む必要があるが、何がどこまで可能なのかを見極める必要もあることを回答。</u> 【取組【例】】 - 旧 LINE および旧ヤフーは、G7 デジタル・技術大臣会合関連イベント「フェイクニュースと日本一私たちにできること・社会としてできること」(国際大学 GLOCOM 主催、総務省およびグーグル合同会社後援、2023 年 4 月開催)に協力参加。 - デジタルやメディア・リテラシープログラムやユーザーだけでなく、ジャーナリスト及び研究者向けトレーニング等の活動を通して、Google はプロダクトを超えて健全なジャーナリズムのエコシステムをサポートし、市民団体や研究者とパートナーシップを組み、将来リスクの一步先を行き、このエコシステムの一部として、世界中の事実確認の専門家やネットワークとの協力し、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。 - Global Network Initiative (GNI)、Global Internet Forum to Counter Terrorism (GIFCT)、Digital Trust
--	--

	and Safety Partnership (DTSP)、Tech Coalition など、多くの国際組織やイニシアティブの創設メンバーまたは参加メンバー。 ▪ ミュンヘン・セキュリティ・カンファレンスにおいて、「2024 年選挙における AI の欺瞞的使用に対抗するための技術協定」を 2 月 16 日に業界各社と締結し、欺瞞的 AI 選挙コンテンツに関連するリスクを軽減するための 8 つの具体的な約束。Adobe、Amazon、Google、IBM、Meta、Microsoft、OpenAI、TikTok、X を含む 20 のテクノロジー企業がこの協定に署名。 ▪ Google、Microsoft、The Rockefeller Foundation、Snap、TikTok とともに、International Foundation for Electoral Systems (IFES) が運営するテクノロジー企業のための自主的な Election Integrity Guidelines にも署名。このガイドラインは、企業と選挙管理当局が選挙の完全性を推進し、情報エコシステムに対する信頼を高めるために期待されること、および実践すべきことを共有。ガイドラインは、韓国ソウルで開催された民主主義サミットで発表。 ▪ コンテンツが AI を使用して作成されたことを示す共通の技術標準について、業界のパートナーと協力。これらのシグナルを検出できるようになれば、利用者が Facebook や Instagram に投稿する AI 生成画像にラベルを付けることが可能。現在この機能を構築中で、今後数ヶ月のうちに、各アプリがサポートするすべての言語でラベルの適用を開始する予定。弊社は、利用者がどのように AI コンテンツを作成し共有しているのか、利用者がどのような透明性に最も価値を見出すのか、そしてこれらの技術がどのように進化していくのかについて、さらに多くのことを学べることを期待。業界のベストプラクティス、そして弊社自身の今後のアプローチに反映。 ▪ 「みんなのデジタル教室」で一般の人々に対しても認識向上の活動。 ▪ 国際会議のスポンサーとして参加し、国際ファクトチェックネットワーク(IFCN)主催の世界ファクトチェック会議 Global Fact 10 のスポンサー及びセッション開催。 ▪ 国際機関とのメディアリテラシーキャンペーンのコラボとして、UNDP クリエイター・フォワード:UNDP と連携したハラスメント及びヘイトスピーチ防止キャンペーン。 ▪ 米国国立児童行方不明センター(NCMEC)との連携として、児童の性被害防止のために、TikTok のグローバルチャイルドセーフティチームは、24 時間体制で、疑わしい児童搾取コンテンツを積極的に
--	--

	米国国立児童行方不明センター (NCMEC) に報告。その後、NCMEC はこれらの情報を警察庁と共有。また、コンテンツに緊急性があるか、または深刻な場合は、該当チームが迅速に情報を提供するため警察庁に、サイバーチップを共有。 ▪ 世界経済フォーラム年次総会 (ダボス会議) 等に出席し、プラットフォームの安全性について各国の閣僚はじめ政府関係者とも意見交換。 ▪ デジタル関連の問題に対処する際、マイクロソフトはマルチステークホルダーアプローチを主導して奨励するよう努力。インターネット上の問題に効果的に対処するには、その性質上、世界規模の対策が必要になることが多い。例えば、マイクロソフトは GIFCT (テロ対策に関するグローバル インターネット フォーラム)、Rome Call for AI Ethics (人工知能の倫理的ガイドライン)、Paris Call for Trust and Security in Cyberspace (サイバー空間の信頼性と安全性のためのパリ・コール)、Tech Accord to Combat Deceptive Use of AI in 2024 Elections (2024 年選挙における人工知能の欺まんの使用に対抗するための技術協定) など、複数の利害関係者による国際的な枠組みを主導。マイクロソフトは、日本をはじめとする世界中の関連する政府、企業、学術機関、非営利組織 (NPO) を支持し、支援。 ▪ 2023 年 12 月の G7 内務・安全担当大臣会合への対面での参加やプレゼンテーション、また GIFCT のメンバーとしての活動など。 ▪ テロ関連の違反コンテンツを特定するため、業界が共有するハッシュデータベース (GIFCT が支援) を使用したり、報告がある前に PhotoDNA のような業界共有ハッシュを利用するなど、さまざまな方法を実施。オンラインでのテロリストや暴力的過激派のコンテンツをより効果的に検出し削除するための技術投資を続けることを約束。これには、デジタルフィンガープリント技術や AI 技術の発展や導入。クライストチャーチ・コール、GIFCT、EU インターネットフォーラム (EUIF) など、様々な団体との協力を通じて、テロリストや過激派がインターネットをどのように利用しているかの新たな動向を把握。 ▪ CSE (子供への性的搾取) やテロリストコンテンツを撲滅するための取り組みを実施。 【国際機関等との連携状況】(参考資料19-1-1 及び19-1-2 Q17-1～17-3)
21 その他のステークホルダーとの連携状況	○ ほぼ全ての国内事業者及び全ての国外事業者は、偽・誤情報の流通・拡散への対応という観点にかぎらないものも含め、<u>様々な分野の専門家・ジャーナリスト・クリエイター等と連携した情</u>

報発信、クリエイターへの情報・ノウハウの共有・エデュケーションプログラム・人材交流等支援、（一社）VRM コンソーシアム等への参加による他のプラットフォーム事業者やメタバース関連事業者との連携・協力、（一社）コンピュータエンターテインメント協会に参加、（一社）セーファーインターネット協会・（一社）安心協・モバイルコンテンツフォーラム等業界団体を通じての取り組み、アジアインターネット日本連盟（AICJ）や一般社団法人ソーシャルメディア利用環境整備機構（SMAJ）等業界団体における取り組み、ゼロレーティングの享受を通じた電気通信事業者との連携、10代や保護者向けの啓発活動、新型コロナウイルス感染症ワクチンに係る誤情報に関する啓発活動への協力、不登校生動画選手権の共催、プラットフォームにおける性被害の防止のための啓発等、その他のステークホルダーと連携・協力した取組を実施と回答。

- 一部の国内事業者及び国外事業者は、その他のステークホルダーとの連携において支障となっている点として、行われているイベントや勉強会に関する情報共有の仕組みがないことや、誤情報や偽情報のような問題には業界を超えた協力と社会全体からのアプローチが必要であることを回答。

【取組【例】】

- ・ 2023 年8月、さまざまな分野の専門家やジャーナリスト、クリエイターが自らの知見をもとにユーザーの新しい気づきや考えるヒント、行動につながる情報を発信するプラットフォーム（Yahoo!ニュース エキスパート）を新たに開始。記事や、コメントで各分野のエキスパートからの信頼性の高い情報を提供。クリエイター」のほか、Yahoo!ニュース個人の書き手である「オーサー」「コメンテーター」総勢約 2600 名が「Yahoo!ニュース エキスパート」に参加。
- ・ 多くのコンテンツ制作主体者に協賛・ご協力いただき、多種多様なイベントを開催。日本ネットクリエイター協会等、クリエイターの方々の支援に繋がる組織への参加。
- ・ 一般社団法人 VRM コンソーシアム等への参加により、他のプラットフォーム事業者様やメタバース関連事業者との連携・協力。
- ・ 一般社団法人コンピュータエンターテインメント協会に参加し、「CS品質向上委員会」に所属。各メーカーの顧客サポートに寄せられた課題・取組事例等の共有、消費者トラブルを未然に防ぐ取

	組。 ▪ セーフティーインターネット協会、安心協、モバイルコンテンツフォーラムなど、業界団体を通じての取り組みや、有志企業担当者との勉強会は随時開催。 ▪ アジアインターネット日本連盟 (AICJ) や一般社団法人ソーシャルメディア利用環境整備機構 (SMAJ) などの業界団体における取り組みを通じ、日本における革新的なビジネス及びインターネット産業の健全な成長向け連携・協力体制を構築。 • Instagram は 2020 年より、クリエイターと連携し、若年層ユーザーと一緒にインスタグラムの安全な使い方を考えるプロジェクト「#インスタアンゼン会議」を立ち上げ、10 代や保護者向けの啓発活動を実施。 ▪ 他の事業者との業界連携に積極的に貢献。SIA が主催した 2021 年の新型コロナウイルス感染症ワクチンに係る誤情報に関する啓発活動に積極的に協力。 ▪ 「不登校新聞」と連携し、『不登校生動画選手権』を 2023 年 5 月に共催。今年もこの取り組みは継続予定。 ▪ クリエイターとの連携・協力として、クリエイターへの情報・ノウハウの共有において、TikTok Creator Academy というエデュケーションプログラムを実施。クリエイターの人材交流という意味で、クリエイター同士のコミュニティづくりを重視した TikTok Creator Connect や、クリエイターとファンとの交流を目的とした TikTok Creative Festival も開催。 ▪ 日本におけるゼロレーティングの享受を通じた連携として、Softbank: ウルトラギガモンスタープラスプラン(2020 年3月 11 日をもち新規受付終了)、NuroMobile: Neo プラン、Povo2.0 PrePaid プラン: 期間限定のゼロレーティングを含むオプション。SNS データ使い放題(3/20-4/20, 2023)、SNS+動画データ使い放題(4/28-5.28, 2023)。 ▪ SNS等のプラットフォームを運営する事業者等から構成される一般社団法人ソーシャルメディア利用環境整備機構 (SMAJ) に理事会社として参加し、会員の事業者と、偽・誤情報の課題検討等、各種委員会・ワーキンググループにおいて意見交換やベストプラクティスの共有。 ▪ 他のプラットフォームと連携した、安全のための啓発活動。2021 年にはプラットフォームにおける性
--	---

	被害の防止を目的として、専門家や NPO に加えて旧 Twitter Japan 株式会社、旧 LINE 株式会社と連携したオンラインセミナーを開催し、TikTok LIVE で配信することで啓発メッセージを発信。 【その他のステークホルダーとの連携状況】(参考資料19-1-1 及び19-1-2 Q18-1～18-4)
--	--