

1. 前提: 対象サービスの範囲について

SNS	X
Search	
Video Sharing	

News portal	
Blog	

Q1.1 上記【想定対象サービス】欄に挙げられたもの以外に、貴社又は貴社の関係会社が事業上運営しているサービスであって、以下の①～③のいずれにも当てはまるもの(上記【想定対象サービス】欄に記載の類型以外を含む)があれば、御教示下さい。

- ① インターネットを通じ、日本国内の不特定の者によって受信されることを目的としてコンテンツ(文字、画像、映像、音声など)を媒介・送信するものであること。
 - ② 上記①のコンテンツに自社以外の第三者が作成・投稿したものが含まれること。
 - ③ 広告収入を主たる運営基盤としていること(※)。
- ※ 広告収入以外の収入(購読料、寄付金など)を主たる運営基盤とするサービスは、いわゆるアテンション・エコノミーの影響を受けにくいと考えられるため、対象外とします。

- ◆ 以下の質問に対しては、上記1の【想定対象サービス】欄に挙げられたサービス(日本国内向けのものに限ります)及びQ1.1で御回答頂いたサービス(以下、併せて「対象サービス」)について御回答下さい。
- ◆ 各質問に対しては、原則としてそれぞれの対象サービスごとに分けて御回答下さい。ただし、対象サービスの全部又は一部に共通の御回答内容となる場合は、その旨を付記して御回答下さい。
- ◆ 各質問にある取組を現時点で実施していない場合でも、実施する予定があるときは、可能な限り当該予定の内容についても御回答内容に含めて頂けますと幸いです。

A 該当するものはございません。

2. 対象サービスの規模について

Q 2.1 対象サービスのうち、コンテンツの投稿に当たって利用者登録が必要なものについては、日本国内における(※)最近の月間アクティブユーザー数として把握している数字を、いつの時点の数字かと併せて御教示下さい。
※1 後記Q2.4参照。

A 日本の収益化できる1ヶ月あたりのアクティブ利用者数(mMAU):6,658万(2023年1月1日～2023年3月31日の平均値)

*mMAU = monetizable MAU(収益につながるMAU) 広告配信可能なx.com(ウェブブラウザ版)またはXアプリに1ヶ月にログインする利用者の数。

Q 2.2 対象サービスのうち、コンテンツの投稿に当たって利用者登録が不要なものについては、日本国内における(※)最近の月間合計投稿数として把握している数字を、いつの時点の数字かと併せて御教示下さい。
※ 後記Q2.4参照。

A	Xでは、ユーザーは投稿や返信をするためにまずアカウントを作り、そのアカウントにログインする必要があります。
Q 2.3	Q2.1及び2.2いずれの数字も把握していない対象サービスについては、日本国内における(※)事業規模を示す数字として把握しているその他の数字(登録ユーザー数など。推計値でもかまいません。)を御教示下さい。 ※ 後記Q2.4参照。
A	—
Q 2.4	Q2.1～2.3で御回答頂いた数字について、「日本国内における」数字であると整理された理由・基準を御教示下さい。 例:利用者登録時に入力された住所、投稿言語、投稿時の位置情報など
A	サービス利用時のIP アドレスに基づいております。

	3. 偽・誤情報の流通・拡散への対応方針について
3	対象サービス上での日本国内における(※1)偽・誤情報(※2)の流通・拡散への対応について、次の3点を御教示下さい(※3)。 ※1 文字や音声については少なくとも日本語を含むもの、画像や映像については日本に関連する人・物・場所等の画像・映像を含むものを想定しております。それ以外の基準により「日本国内における」実態を把握している場合は、当該基準を明示いただいた上でお答えください。以下、「日本国内における」「日本国内において」「日本に関する」とある箇所につき同様。 ※2 ここではひとまず大まかに、「内容や発信主体等に関して虚偽や誤解を招く事項を含む情報」とお考え下さい。 ※3 既に公開している規約、ポリシー、ガイドライン等(以下「規約等」)の中で説明している場合は、当該規約等のタイトル及びURLを付記して、該当箇所を抜粋下さい。
Q3.1 (a)	何らかの対応が必要と考える「偽・誤情報」の範囲(定義)・類型 例:災害関係、健康・医療関係、選挙・政治関係、国際紛争関係、青少年関係、自殺誘発関係、犯罪(詐欺など)誘発関係、ディープフェイクなど
A	Xは、例えば以下のポリシーを含む、プラットフォームの信頼性に関わる様々なポリシーを持っています。 <u>プラットフォームの操作とスパムに関するポリシー:</u> 情報を人為的に拡散または隠蔽したり、Xのユーザー体験や、Xによるプラットフォーム操作の防止策を操作または侵害する行為に関与したりする意図で、Xのサービスを利用することは禁じられています。 <u>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:</u> 他の利用者を欺く目的で、個人、集団、組織の身元情報を流用したり、自身の身元を詐称したりすることはできません。 <u>合成または操作されたメディアに関するポリシー:</u> 利用者を欺いたり、混乱させたりして、損害をもたらす可能性のある、合成または操作されたメディアや、文脈から切り離されたメディア(以下、「誤解を招くメディア」といいます)を共有することは禁止されています。さらに、Xでは誤解を招くメディアが含まれるツイートにラベル付けを行い、利用者がツイートの信ぴょう性や文脈を把握するのに支援する場合があります。

	<p>市民活動の阻害に関するポリシー: 選挙期間中に、市民活動への参加を抑圧したり、市民活動に関する日時、場所、参加方法について誤解を招いたり、現実世界の暴力を引き起こしたりするおそれのあるコンテンツをポストまたは共有するなど、選挙またはその他の市民活動の操作や妨害を目的にXのサービスを利用することを禁じます。</p> <p>質問に記載された例は、Xのサービス利用規約、ルールとポリシーの下での他の報告およびレビューの領域に該当する可能性もあります。犯罪や詐欺、特定の青少年問題などに対処するための専用のポリシー、フォーム、リソースを用意しています。</p> <p>なお、Q.19で提供されたコミュニティノートについての情報も考慮いただけますと幸いです。コミュニティノートは、Xにおける誤情報や誤解を招く情報に対処する上で重要な役割を果たします。</p>
Q3.1 (b)	<p>上記(a)の類型ごとに想定される典型事例</p>
A	<p>プラットフォームの操作とスパムに関するポリシー: 他者の快適性を損ねる目的でXの商品機能を悪用することは禁じられています。例えば、大量、過剰、膨大な返信、@ツイート、またはダイレクトメッセージを一方的に送信することなど。</p> <p>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: 例えばあなたが誰か、または誰を代表しているかについて、他の利用者を欺くため、実在する人物、集団、または組織を詐称すること。または、他の利用者を欺くため、実在しない人物を詐称すること。</p> <p>合成または操作されたメディアに関するポリシー: ・利用者を欺くことを意図して、大幅に改ざん、操作、またはねつ造されたメディアを含んでいるか、あるいは ・利用者を欺くことを意図した方法を使ったり、虚偽の文脈を伴ったりして共有されているメディアを含んでいて、なおかつ ・公共の問題について広範囲にわたって混乱をもたらしたり、公共の安全に影響を及ぼしたり、深刻な損害をもたらしたりする可能性が高いメディアを含んでいる</p> <p>市民活動の阻害に関するポリシー: 市民活動への参加の手続きについて誤解を招く情報(ポスト、テキストメッセージ、メール、または電話での投票ができない地域で、それが可能であると誤解させる情報など)</p>
Q3.1 (c)	<p>上記(a)の類型ごとの具体的な対応方法 例: 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など</p>
A	<p>プラットフォームの操作とスパムに関するポリシー:</p> <p>このポリシーへの違反に対する処分は、その違反の重大さや違反者の過去の違反歴によって決まります。確認されたスパム活動の種類も、措置に影響します。Xが取る措置としては、次のものなどがあります。</p> <p>アンチスパム対策 不審な行為を検知した場合、アカウントをロックし、追加情報(電話番号など)の提供またはreCAPTCHAの解答を求めることがあります。</p> <p>リンクの拒否リストへの登録</p>

Xが安全でないと判断したリンクについては、拒否リストへの登録か、または警告が行われます。安全でないリンクについての詳細もご覧ください。Xがリンクを安全でないと誤って判断した場合に、異議申し立てをする方法を記載しています。

ツイートの表示の制限

こうしたポリシーに違反していると思われるツイートについては、X商品の特定の場所に表示されなくなることや、Xのおすすめから除外されたり、拡散されなくなったりすることがあります。ツイートの表示が制限されることのあるその他のケースの詳細については、こちらをご覧ください。

ツイートの削除と一時的なアカウントのロック

プラットフォームの操作やスパムに関する違反が1回きり、または1回目の違反である場合、1つまたは複数のツイートの削除要請からアカウントの一時的なロックまで、Xはいくつかの措置を取る可能性があります。プラットフォームの操作に関して、続けて違反があった場合、アカウントは永久に凍結されます。複数のアカウントの使用が主な違反であれば、存続させるアカウントを1つ選択できる場合があります。残りのアカウントは永久に凍結されます。

Xの商品または機能へのアクセス権の一時的喪失

Xでは、ツイート、ツイートの編集、ダイレクトメッセージ、スペース、コミュニティ、ライブなどのXの機能または商品(ただしこれらに限定されません)へのアクセスを一時的に制限する場合があります。

永久凍結

重大な違反については、1回目の発見でアカウントを永久に凍結します。重大な違反の例を以下に挙げます。

- ・運用するアカウントで、上記のポリシーへの違反となる行為を繰り返すこと
- ・選挙の清廉性を損ねる目的で、このページに記載された行為のいずれかを実行すること
- ・アカウントを購入または販売すること
- ・凍結されたアカウントを代替または模倣するアカウントを作成すること
- ・Xルールに違反する組織であるとXが断定するアカウントを運用すること

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:

このポリシーの違反に対する処分は、その違反の重大さや種類、違反者の過去の違反歴によって決まります。Xが取る措置としては、次のものがあります。

プロフィールの修正

このポリシーに準拠していないパロディ、コメントリー、ファンアカウントなど、プロフィールに掲げられている対象との関係について、混乱を招くおそれのあるアカウントについては、プロフィールのコンテンツを編集するよう要請することがあります。最初の警告を受けた後に再度このポリシーに違反した場合、アカウントは凍結されることになります。

凍結

なりすまし行為に加担しているアカウントや、虚偽の身元情報を使用しているアカウントは凍結される場合があります。このポリシーに違反している可能性があるかと判断した場合、アカウントを復元するために、政府発行の身分証明書(運転免許証、パスポートなど)の提示を要請する場合があります。

合成または操作されたメディアに関するポリシー:

Xの合成または操作されたメディアに関するポリシーの違反に対する処分は、その違反の重大さによって異なります。

ツイートの削除

誤解を招くメディアに、個人やコミュニティに損害をもたらす深刻な危険性がある場合など、当ポリシーへの違反の程度が深刻な場合、当該コンテンツの削除を要請します。

ラベル付け

当ポリシーに違反するコンテンツを削除しない場合は、X上に表示されている誤解を招くメディアを共有するツイートに背景情報を追加する場合があります。具体的には、以下の措置が取られる場合があります。

- ・当該ツイートにラベルや警告メッセージを適用する
- ・当該ツイートの共有やいいねを行おうとする利用者に警告を表示する
- ツイートの拡散を抑制したり、おすすりから除外したりする
- ・いいね、返信、リツイートが無効にする
- ・たとえば関連するXポリシーなど、補足説明のためのリンクを表示する

多くの場合、ラベル適用の対象となるツイートに対し、上記の措置のいくつかが組み合わされて執行されます。

アカウントのロック

特定のアカウントが、誤解を招く有害な説明を流布しているか継続的に共有しており、それがXの合成または操作されたメディアに関するポリシーに違反していると判断された場合、そのアカウントは一時的に表示を抑制されるか、ロックまたは凍結される場合があります。

市民活動の阻害に関するポリシー：

このポリシーに基づき、強制措置の対象となったポストは、以下の対応によりXにおけるリーチが制限されます。

- ・検索結果、おすすり、トレンド、通知、ホームタイムラインから当該ポストを除外します
- ・[おすすり] および [フォロー中] タイムラインから当該ポストを削除します
- ・当該ポストが作者のプロフィールにだけ表示されるようにします
- ・当該ポストのいいね、返信、リポスト、引用、ブックマーク、共有、プロフィールへの固定表示、編集を行えなくします
- ・返信における当該ポストの表示順位を下げます

このポリシーに基づき、強制措置の対象となったポストには、当該ポストの拡散が制限されている旨を知らせるラベルが適用されます。このラベルは、ポストの作者と閲覧者の双方に表示されます。ポストの作者は、自分のポストに誤って表示制限が課されていると思われる場合には、ラベルに関する異議申し立てを提出することができます。

Q3.2 Q3.1で御回答頂いた対応方針の背景となる基本的な考え方(重視している価値や哲学)があれば、御教示下さい。

A Xの目的は、公共の場における会話に寄与することです。暴力、嫌がらせ、およびその他の類似行為は、利用者の自己表現を抑制し、その結果、世界中で行われる公共の会話の価値を下げることとなります。Xでは、すべての利用者が自由に、安心して公共の会話に参加できるよう、ルールを設けています。

その上でXでは、世界各地でリアルタイムの会話が行われており、攻撃的であったり、物議をかもしたり、偏見に満ちていたりすると思われる意見が述べられていることがあります。Xサービスでは、すべての利用者に自由な自己表現を認めています。嫌がらせや脅し、または恐怖を与えて他の利用者が発言できないようにする行為は禁止しています。

私たちの理念とプロダクトとエンフォースメントを更に説明するものとして以下の情報も合わせて提示します。

[ポリシーの策定に対するXのアプローチと対応についての理念](#)

	コミュニティノート Freedom of Speech. Not Reach
Q3.3	Q3.1で御回答頂いた対応方針について、定期的な見直しや外部有識者・第三者機関等によるレビューを行っている場合は、(a)当該見直し・レビューの頻度・タイミング、及び(b)レビューの実施主体を御教示下さい。
A	<p>Xには、ユーザーの安全性と同時にプラットフォームとアカウントの信頼性を保護するポリシーがあります。Xのルールとポリシーは、ヘルプセンターで公開されており、誰でも簡単に理解できるように書かれています。また、ルールを変更するたびに、ヘルプセンターを定期的に更新しています。</p> <p>さらに、ポリシー開発プロセスとルールの施行哲学についての説明もヘルプセンターで見つけることができます。新しいポリシーの作成やポリシーの変更には、オンライン行動のトレンドに関する徹底的な調査、許可される内容に関する期待を設定する明確な外部言語の開発、そして何百万ものコンテンツやアカウントにスケールできるレビュアー向けの施行ガイダンスの作成が必要です。私たちのポリシーはダイナミックであり、それらが最新の状態であり、必要であり、比例していることを常にレビューしています。</p>

	4. 偽・誤情報の発信者(投稿者)の表現の自由等への配慮について
Q4.1	<p>Q3.1で御回答頂いた対応方針の対象となる偽・誤情報(日本国内において流通するものに限り、以下「対象偽・誤情報」)の流通・拡散に対し、Q3.1(c)で御回答頂いた対応(以下「モデレーション等」)を実施した場合、実施の事実や理由を当該対象偽・誤情報の発信者(投稿者)に通知し、又は発信者(投稿者)の知り得る状態に置いている場合、その方法(※1)、言語及び内容(※2)を御教示下さい。</p> <p>※1 例えば、書面、電子メール、対象サービス上での表示など。</p> <p>※2 例えば、モデレーション等を実施した事実及び理由のほか、後記5の第三者通報によるものか、第三者通報を待たず自社で検知・対応したものかといった実施の端緒や、Q4.3の苦情・不服申立て受付窓口に関する情報など。</p>
A	<p>報告されたコンテンツがXの利用規約やXルールに違反しているかどうか、またXの利用規約やXルールに違反していなくても特定の国では違法であるという申し立てがあるかどうかによって、結論は変わります。</p> <p>報告されたコンテンツがXの利用規約やXルールに違反している場合</p> <ul style="list-style-type: none"> このポリシー違反に対するペナルティを決定する際、違反の重大性、誰かが標的にされているか(例:メンションされたり、フルネームで言及されたり、写真で示されたりするなど)、個人の過去のルール違反の記録など、複数の要因を考慮します。以下は、このポリシーに違反するコンテンツに対する潜在的な施行オプションのリストになります。 <ul style="list-style-type: none"> X上でのコンテンツの可視性を下げる。 投稿の削除を要求する。 アカウントの停止。 ログインするときに、アカウントを凍結またはロックしたというメッセージが表示されます。また、特定の操作を行わなければアカウントを利用できないというメッセージが表示されることもあります。 重大な違反を行ったり、違反行為を繰り返したりするとアカウントが永久凍結される可能性があります。アカウントの凍結についてはこちら、アカウントのロックについてはこちらのヘルプセンター記事をご覧ください。 <p>コンテンツにセンシティブなメディアが含まれていると報告された場合</p> <ul style="list-style-type: none"> そのコンテンツはセンシティブなコンテンツと見なされ、他のユーザーに対してそれを見るかどうかの選択肢が提示される場合があります。 <p>Xの利用規約やXルールには違反していないが、一部の国では違法であると申し立てられた場合</p>

	<ul style="list-style-type: none"> ・お使いのアカウントから投稿されたコンテンツに違法性があると申し立てる法的請求を受けたことをユーザーに通知します(通知が可能な場合)。 ・禁止されていなければ、その法的請求の内容をお知らせします。 ・これは、X側で問題のコンテンツを処理する前に、ユーザー自身が問題のコンテンツを削除したり、可能であれば請求者に直接対応したり、弁護士などに依頼して請求に異議を申し立てたりするなどの選択肢があることをユーザーに通知するためのものです。 ・このような通知は、特定のコンテンツを削除すべきという法的助言ではありません。通知はあくまでもXアカウントの特定のコンテンツに対する法的請求が提出されたことをユーザーに伝えるためのものであり、ユーザーに対応方法を検討していただくためにお知らせしています。また、報告されたコンテンツに対して、X側で法的請求に基づく処置を行う必要が生じる場合があることをお知らせするのも、通知の目的の1つです。
4.2	対象偽・誤情報の流通・拡散に対して実施したモデレーション等について、当該対象偽・誤情報の発信者(投稿者)からの苦情や不服申立てを受け付ける窓口を設置している場合、当該窓口に関し、次の3点を御教示下さい。
Q 4.2(a)	(a)苦情・不服申立ての受付方法(※) ※ 例えば、オンラインフォーム、書面、電子メール、電話など。なお、予め定められている申請フォーマットがある場合は、当該フォーマットを添付(又はURLを付記)下さい。
A	いくつかの方法があります。 ロックまたは制限されたアカウントに関するヘルプ ユーザーは、自分のアカウントが誤ってロックまたは凍結されたと考える場合、 異議申し立てを提出 することにより、その決定に対して異議を唱えることができます。
Q 4.2(b)	日本語による苦情・不服申立てを受け付けているか否か
A	Xは日本語で報告・苦情を受け付けています。
Q 4.2(c)	窓口情報を日本語で公開している場合はその方法(※) ※ オンライン上で公開している場合はURLを付記下さい。
A	該当情報は日本語で公開されています。 ロックまたは制限されたアカウントに関するヘルプ
4.3	Q4.2で窓口を設置しているか否かにかかわらず、対象偽・誤情報の流通・拡散に対して実施したモデレーション等について、当該対象偽・誤情報の発信者(投稿者)から日本語で苦情や不服申立てを受けた場合における対応(以下「日本語苦情等対応」)(※1)に関し、次の8点を御教示下さい(※2)。 ※1 苦情・不服申立て内容の合理性や問題となったモデレーション等の適否を調査・判断し、場合によってはモデレーション等を撤回するなどの対応を指します。 ※2 問題となったモデレーション等が後記5の第三者通報によるものか、第三者通報を待たず自社で検知・対応したものかといった実施の端緒によって御回答内容が異なる場合は、当該実施の端緒別に御教示下さい。
Q 4.3(a)	日本語苦情等対応が可能な人員として常用している人数
A	当社のルールを施行するために、機械学習と人間によるレビューの組み合わせを使用しています。当社のシス

	<p>テムは、重要な文脈を使用して潜在的なルール違反についての決定を下す人間のモデレーターにコンテンツを提示することができます。この作業は、24時間体制で多言語に対応できる国際的なクロスファンクショナルチームによって主導されています。また、発生する可能性のある任意のエラーに対する苦情処理プロセスも持っています。日本語での苦情対応に従事している人数は公開しておりません。</p>
Q 4.3(b)	<p>日本語苦情等対応にAIその他の機械的手段を利用している場合は、当該手段の概要及び利用手順(どのようなケースで用いるのかなど)</p>
A	<p>基本的には人による目視にて確認しています。</p>
Q 4.3(c)	<p>日本語苦情等対応にファクトチェック機関やマスメディア(新聞・放送)を関与させている場合は、当該機関等の名称及び関与手順(どのようなケースで関与させており、その意見等を日本語苦情等対応へどのように反映させているのかなど)</p>
A	<p>苦情を取り扱う際に、ファクトチェック機関やマスメディアを特別に関与させてはおりません。</p>
Q 4.3(d)	<p>上記(c)以外の外部有識者・団体を日本語苦情等対応に関与させている場合は、当該外部有識者・団体の名称及び関与手順(どのようなケースで関与させており、その意見等を日本語苦情等対応へどのように反映させているのかなど)</p>
A	<p>特別な対応はしておりません。</p>
Q 4.3(e)	<p>日本語苦情等対応の開始から完了までの目標期間を定めている場合は、当該目標期間(「開始」及び「完了」の定義を含む)</p>
A	<p>可能な限り迅速に対応できるよう努めています。</p>
Q 4.3(f)	<p>日本語苦情等対応の結果を発信者(投稿者)に通知し、又は発信者(投稿者)の知り得る状態に置いている場合は、その方法、言語及び内容</p>
A	<p>ポリシーおよび関連する報告に対応するオプションがあり、その中で日本語で送信者に通知をします。方法には、電子メール経由、または報告フローを持つXアプリ内のユーザーに対しても含まれます。</p>
Q 4.3(g)	<p>日本語苦情等対応の結果に対する再審査制度を設けている場合は、その概要(日本語による再審査要求を受け付けているか否か、及び外部有識者や第三者機関を関与させている場合はその詳細を含む)</p>
A	<p>報告を受け取り、異議申し立てに対応するプロセスがございます。アカウントのロックまたは凍結に関する異議申し立て</p>
Q 4.3(h)	<p>特定の発信者(投稿者)(※)からの日本語による苦情・不服申し立てを優先的に取り扱うこととしている場合は、当該発信者(投稿者)の類型及び具体的な優先的取扱いの内容</p> <p>※ 例えば、国会、裁判所、行政機関、地方公共団体、その他公的機関(独立行政法人等)、国際機関、専門機関、ファクトチェック機関など</p>

A	送信者やアカウントの所有者が誰であるかに基づいて優先順位をつけることはありません。 しかし、例えば政府や法執行機関によって送信された緊急の報告に優先的に対応するための 専用のチャネル とチームも持っています。
Q4.4	Q4.3で御回答頂いた内容を公開している場合は、その方法(※)を御教示下さい。 ※ オンライン上で公開している場合はURLを付記下さい。
	執行機関／捜査機関向けガイドライン

	5. 偽・誤情報の流通・拡散に対するモデレーション等の手続・体制について
5.1	対象偽・誤情報の流通・拡散について、それにより権利を侵害されている者及び発信者(投稿者)以外の第三者からの通報(以下「第三者通報」)を受け付ける窓口を設置している場合、当該窓口に関し、次の4点を御教示下さい。
Q 5.1(a)	受け付けている通報手段(※) ※ 例えば、オンラインフォーム、書面、電子メール、電話など。なお、予め定められている申請フォーマットがある場合は、当該フォーマットを添付(又はURLを付記)下さい。
A	私たちは、第三者による報告のための標準的なチャネルと手段を持っており、以下に例を挙げます。 ヘルプセンター https://help.twitter.com/ja/forms 法的要請 https://legalrequests.twitter.com/forms/landing_disclaimer
Q 5.1(b)	通報できる主体に限定を設けている場合は、当該主体 例: 利用者登録済みユーザー、ログイン済みユーザー、国会、裁判所、行政機関、地方公共団体、国際機関、専門機関、ファクトチェック機関など
A	政府や法執行機関専用のフォームを用意しています。 https://legalrequests.twitter.com/forms/landing_disclaimer なお、ユーザーは、コミュニティノートに参加することで、誤情報やデマ情報に情報を書き足すこともできます。日本にいる誰もがノートを閲覧し、評価し、貢献するために登録することができます。(詳細はQ.19にて)
Q 5.1(c)	日本語による通報を受け付けているか否か
A	日本語で受け付けています。
Q 5.1(d)	窓口情報を日本語で公開している場合はその方法(※) ※ オンライン上で公開している場合はURLを付記下さい。

A	以下の情報をご参照ください。 https://help.twitter.com/ja/forms
Q5.2	Q5.1の窓口を設置している場合、日本語による第三者通報を受け付けた後の対応(以下「日本語通報対応」)(※1)に関し、次の8点を御教示下さい(※2)。 ※1 通報内容の合理性や問題となったコンテンツの対象偽・誤情報該当性を調査・判断し、場合によってはモデレーション等を実施するなどの対応を指します。 ※2 通報主体や対象偽・誤情報の類型(Q3.1(a)参照)毎に異なる場合は、当該主体別又は類型別に御教示下さい。
Q 5.2(a)	日本語通報対応が可能な人員として常用している人数
A	私たちは、日本語を含む複数の言語に対応できる、24時間体制の国際的なクロスファンクショナルチームを持っています。また、発生する可能性のある任意のエラーに対する苦情処理プロセスも持っています。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。
Q 5.2(b)	日本語通報対応にAIその他の機械的手段を利用している場合は、当該手段の概要及び利用手順(どのようなケースで用いるのかなど)
A	私たちはルールの施行のため、機械学習と人のレビューを組み合わせています。私たちのシステムは人間のモデレーターに対して、重要な文脈をもとに潜在的なルール違反に関する判断を下すためのコンテンツを提示する能力を持っています。この取り組みは、24時間体制で日本語も含め多言語に対応可能な国際的で機能横断的なチームが主導しています。また、発生しうるあらゆるエラーに対応するための苦情処理のプロセスも整えています。
Q 5.2(c)	日本語通報対応にファクトチェック機関やマスメディア(新聞・放送)を関与させている場合は、当該機関等の名称及び関与手順(どのようなケースで関与させており、その意見等を日本語通報対応へどのように反映させているのかなど)
A	ファクトチェック団体やマスメディアは、報告への対応には関与していません。
Q 5.2(d)	上記(c)以外の外部有識者・団体を日本語通報対応に関与させている場合は、当該外部有識者・団体の名称及び関与手順(どのようなケースで関与させており、その意見等を日本語通報対応へどのように反映させているのかなど)
A	そのような特別なプロセスは持っていません。
Q 5.2(e)	日本語通報対応の開始から完了までの目標期間を定めている場合は、当該目標期間(「開始」及び「完了」の定義を含む)
A	報告される問題やトピックの複雑さによって異なります。
Q 5.2(f)	日本語通報対応の結果を通報者に通知し、又は通報者の知り得る状態に置いている場合は、当該通報の方法、言語及び内容
A	基本的に、ユーザーに対しメールで通知します。

Q 5.2(g)	日本語通報対応の有無・内容に関し、通報者からの不服申立てや問合せがあったときの専用窓口や処理手順を定めている場合は、その概要(日本語による不服申立て・問合せを受け付けているか否か、及び外部有識者や第三者機関を関与させている場合はその詳細を含む)
A	政府機関や法執行機関は、自身のケースをXに報告し、エスカレートすることができます。また、私たちはユーザーのための異議申し立て手続きも用意しています。 Help Center
Q 5.2(h)	特定の通報主体(※)からの日本語による通報を優先的に取り扱うこととしている場合は、当該通報主体の類型及び具体的な優先的取扱いの内容 ※ 例えば、国会、裁判所、行政機関、地方公共団体、国際機関、専門機関、ファクトチェック機関など
A	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。 https://legalrequests.twitter.com/forms/landing_disclaimer
5.3	対象偽・誤情報の流通・拡散について、第三者通報を待たず自社で検知・対応(以下「検知・対応」)(※1)することがある場合、次の5点を御教示下さい(※2)。 ※1 検知内容の合理性やモデレーション等の必要性を調査・判断し、場合によってはモデレーション等を実施するなどの対応を含みます。 ※2 対象偽・誤情報の類型(Q3.1(a)参照)ごとに異なる場合は当該類型別に御教示下さい。
Q 5.3(a)	検知・対応に向けた調査を実施する頻度・ケース
A	報告された際に、X上の投稿をレビューします。 同時に、Xはプラットフォーム上の不正確性や操作を特定し、対処するために、人間のエージェントと自動検出の組み合わせを使用しています。これは、悪意ある行為者がプラットフォームを操作するコストを増加させ、それに伴う害をXが緩和することを目的としています。
Q 5.3(b)	検知・対応にAIその他の機械的手段を利用している場合は、当該手段の概要及び利用手順(どのようなケース・頻度で用いるのかなど)
A	自動コンテンツモデレーション Xは、プラットフォーム上で施行されるXのルールとポリシーに違反するコンテンツを自動的に検出するために、ヒューリスティックと機械学習アルゴリズムの組み合わせを使用しています。 機械学習モデル 私たちは、自然言語処理モデル、画像処理モデル、その他の高度な機械学習手法の組み合わせを使用して、違反の可能性があるコンテンツを検出します。これらのモデルは複雑さと生成する出力において異なります。例えば、プラットフォーム上での虐待を検出するために使用されるモデルは、過去に検出された虐待違反に基づいて訓練されています。これらの機械学習モデルによってフラグが立てられたコンテンツは、アクションが取られる前に人間のコンテンツレビュアーによってレビューされるか、場合によってはモデルの出力に基づいて自動的に処理されます。

	<p>ヒューリスティックモデル</p> <p>ヒューリスティックは、プラットフォーム上で新しい形式の違反が出現した際に、Xが迅速に反応できるようにするために一般的に利用されます。ヒューリスティックは、特定のカテゴリーの違反に典型的なテキストやキーワードの一般的なパターンです。ヒューリスティックによって検出されたコンテンツも、コンテンツに対して行動が取られる前に人間のコンテンツレビュアーによってレビューされる場合があります。これらのヒューリスティックは、コンテンツを人間のエージェントによるレビューのためにフラグを立て、そのようなコンテンツがレビューされる順序を優先するために使用されます。</p>
Q 5.3(c)	<p>検知・対応にファクトチェック機関やマスメディア(新聞・放送)を関与させている場合は、当該機関等の名称及び関与手順(どのようなケースで関与させており、その意見等を検知・対応へどのように反映させているのかなど)</p>
A	<p>ファクトチェック団体やマスメディアは、アカウントにサインアップしてX上で発言することができます。</p> <p>また、どの団体や個人もXのコミュニティノートにサインアップすることができます。</p> <p>いずれにしましても、当該機関は検知・対応プロセスに関与しておりません。</p>
Q 5.3(d)	<p>上記(c)以外の外部有識者・団体を検知・対応に関与させている場合は、当該外部有識者・団体の名称及び関与手順(どのようなケースで関与させており、その意見等を検知・対応へどのように反映させているのかなど)</p>
A	<p>該当するものはございません。</p>
Q 5.3(e)	<p>検知してからモデレーション等を実施するまで(モデレーション等を実施しない場合は実施しない旨決定するまで)の目標期間を定めている場合は、当該目標期間</p>
A	<p>該当するものはございません。</p>
Q5.4	<p>Q5.2及び5.3で御回答頂いた内容を公開している場合は、その方法(※)を御教示下さい。 ※ オンライン上で公開している場合はURLを付記下さい。</p>
A	<p>該当するものはございません。</p>

	<p>6. 偽・誤情報の流通・拡散への対応状況について</p>
6	<p>2022年中及び2023年中次の20点それぞれに関して把握している数字があれば、御教示下さい。</p> <p>※1 別の期間の数字のみ回答可能な場合は、期間を明示した上で当該数字を御教示下さい。 ※2 可能であれば、対象偽・誤情報の類型別(Q3.1(a)参照)、モデレーション等の方法別(同(c)参照)の数字も併せて御教示下さい。また、可能であれば、(i)~(l)については通報主体別(Q5.1(b)参照)、(r)・(s)については発信者(投稿者)の類型別の数字も併せて御教示下さい。</p>
Q 6.1(a)	<p>対象偽・誤情報(※)についてモデレーション等を実施した件数</p> <p>※ Q4.1のとおり、対象偽・誤情報は「日本国内において流通するもの」に限りますが、対象偽・誤情報に関する数字を把握していない場合で、何らかの推計値や、対象偽・誤情報に関する数字を含むものと思われる別の数</p>

	字を把握しているときは、どのような数字であるか明示頂いた上で当該数字を御回答下さい。以下同じ。
A	<p>2022 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: 5,303 プラットフォームの操作とスパムに関するポリシー: 874,068 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>-----</p> <p>2023 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: 4,829 プラットフォームの操作とスパムに関するポリシー: 4,451,797 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>※複数のポリシーに違反するコンテンツを一つのポリシーで対応するケース等もあることから、上記データをそれぞれのポリシー違反に絞って抽出することはできません。</p>
Q 6.1(b)	上記(a)のうち、AIその他の機械的手段のみによって検知・対応した件数
A	<p>2022 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: 3 プラットフォームの操作とスパムに関するポリシー: 5,300 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>-----</p> <p>2023 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: 900 プラットフォームの操作とスパムに関するポリシー: 2,973,405 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>※複数のポリシーに違反するコンテンツを一つのポリシーで対応するケース等もあることから、上記データをそれぞれのポリシー違反に絞って抽出することはできません。</p>
Q 6.1(c)	上記(a)のうち、AIその他の機械的手段と人間の手を組み合わせてモデレーション等を実施した(第三者通報への対応過程でAIその他の機械的手段を利用した場合を含む)件数
A	ご用意できる情報がございません。
Q 6.1(d)	上記(b)(c)のうち、AIその他の機械的手段による判断結果に誤りがあることが事後的に判明した(上記(c)において人間の手によって補正された場合を含む)件数
A	<p>2022 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: 0</p>

	<p>プラットフォームの操作とスパムに関するポリシー: 1,596 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>-----</p> <p>2023 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: 31 プラットフォームの操作とスパムに関するポリシー: 17,226 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>※複数のポリシーに違反するコンテンツを一つのポリシーで対応するケース等もあることから、上記データをそれぞれのポリシー違反に絞って抽出することはできません。</p>
Q 6.1(e)	上記(a)のうち、ファクトチェック機関の意見等を反映してモデレーション等を実施した件数
A	該当するものはございません。
Q 6.1(f)	上記(a)のうち、マスメディアの意見等を反映してモデレーション等を実施した件数
A	該当するものはございません。
Q 6.1(g)	上記(a)のうち、上記(e)(f)以外の外部有識者・団体の意見等を反映してモデレーション等を実施した件数
A	該当するものはございません。
Q 6.1(h)	上記(e)～(g)のうち、ファクトチェック機関、マスメディアその他の外部有識者・団体の意見等に誤りがあることが事後的に判明した件数
A	該当するものはございません。
Q 6.1(i)	<p>対象偽・誤情報に関する日本語による(※)第三者通報の総数 ※ 日本語による第三者通報の件数を把握していない場合で、言語以外の基準により日本国内における同様の数字として把握している数字があるときは、当該基準を明示頂いた上で当該数字を御回答下さい(推計でもかまいません)。また、そのような数字も把握していない場合で、日本国内における数字を含むものと思われる別の数字を把握しているときは、どの範囲の数字であるか明示頂いた上で当該数字を御回答下さい。以下同じ。</p>
A	現時点でご提示できる情報はございません。
Q 6.1(j)	上記(i)の第三者通報を契機として対象偽・誤情報についてモデレーション等を実施した件数
A	現時点でご提示できる情報はございません。
Q 6.1(k)	上記(i)の第三者通報を受け付けてからモデレーション等を実施するまで(モデレーション等を実施しない場合は実施しない旨決定するまで)に要した期間の平均値

A	現時点でご提示できる情報はございません。
Q 6.1(l)	上記(i)のうち、モデレーション等の有無・内容に関し、通報者からの不服申立てや問合せがあった件数
A	現時点でご提示できる情報はございません。
Q 6.1(m)	Q4.2の窓口への対象偽・誤情報に関する発信者(投稿者)からの日本語による(※)苦情・不服申立て件数 ※ 日本語による苦情・不服申立て件数を把握していない場合で、言語以外の基準により日本国内における同様の数字として把握している数字があるときは、当該基準を明示頂いた上で当該数字を御回答下さい(推計でもかまいません)。また、そのような数字も把握していない場合で、日本国内における数字を含むものと思われる別の数字を把握しているときは、どの範囲の数字であるか明示頂いた上で当該数字を御回答下さい。以下同じ。
A	<p>2022 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: 1,609 プラットフォームの操作とスパムに関するポリシー: 281,410 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>-----</p> <p>2023 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: 1,450 プラットフォームの操作とスパムに関するポリシー: 694,969 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>※複数のポリシーに違反するコンテンツを一つのポリシーで対応するケース等もあることから、上記データをそれぞれのポリシー違反に絞って抽出することはできません。</p>
Q 6.1(n)	Q4.2の窓口以外を通じた対象偽・誤情報に関する発信者(投稿者)からの日本語による苦情・不服申立ての件数
A	現時点でご提示できる情報はございません。
Q 6.1(o)	上記(m)(n)のうち、当該苦情・不服申立てに基づき、問題となったモデレーション等を撤回した件数
A	<p>2022 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: 105 プラットフォームの操作とスパムに関するポリシー: 6,525 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>-----</p> <p>2023 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: 143 プラットフォームの操作とスパムに関するポリシー: 37,519 市民活動の阻害に関するポリシー: 0 合成または操作されたメディアに関するポリシー: 0</p> <p>※複数のポリシーに違反するコンテンツを一つのポリシーで対応するケース等もあることから、上記データをそれぞれのポリシー違反に絞って抽出することはできません。</p>
Q	上記(m)(n)の苦情・不服申立てを受け付けてから、問題となったモデレーション等を撤回するまで(撤回しない場

6.1(p)	合は撤回しない旨決定するまで)に要した期間の平均値
A	現時点でご提示できる情報はございません。
Q 6.1(q)	上記(m)(n)のうち、Q4.3(g)の再審査要求があった件数
A	現時点でご提示できる情報はございません。
Q 6.1(r)	上記(m)(n)のうち、Q4.3(h)の類型に該当する特定の発信者(投稿者)による苦情・不服申立ての件数
A	該当するものはございません。
Q 6.1(s)	上記(r)の苦情・不服申立てを受け付けてから、問題となったモデレーション等を撤回するまで(撤回しない場合は撤回しない旨決定するまで)に要した期間の平均値
A	現時点でご提示できる情報はございません。
Q 6.1(t)	モデレーション等の有無・内容に関し、発信者(投稿者)、通報者との間で、訴訟・仲裁・調停など裁判所等の外部紛争解決機関が関与する手続に発展した件数
A	現時点でご提示できる情報はございません。
Q6.2	Q6.1で御回答頂いた数字について、いわゆる透明性レポート等の形で公開している場合は、その方法(※)と頻度(例えば、年1回など)を御教示下さい。 ※ オンライン上で公開している場合はURLを付記下さい。
A	該当する回答情報は公開しておりません。

7. 令和6年能登半島地震関連の偽・誤情報の流通・拡散への対応状況について	
Q 7.1	2024年1月1日に発生した令和6年能登半島地震に関連する偽・誤情報の流通・拡散に対する取組状況を御教示下さい。特に、現時点までにモデレーション等を実施した主なコンテンツとモデレーション等の方法別(Q3.1(c)参照)件数を御教示下さい。
A	1月1日に発生した地震に関連する詳細な数字はありません。
Q 7.2	令和6年能登半島地震に関連するコンテンツのうち、その内容が誤りであるとするファクトチェック機関のファクトチェック記事が存在するコンテンツについて、Q7.1で御回答頂いた件数に含まれる場合は、当該件数を御教示下さい。
A	現時点でご提示できる情報はございません。
Q 7.3	令和6年能登半島地震に関連する偽・誤情報の流通・拡散への対応として、上記4及び5にて御回答頂いた頂いた回答と比べて体制を強化した場合は、強化した内容を御教示下さい。 ※ 例えば、対行政機関窓口の設置、モデレーション等を担当する人員の増強など。
A	地震に関連する状況について、積極的に監視しました。

Q 7.4	令和6年能登半島地震に関連する偽・誤情報の流通・拡散への対応に関し、他の事業者、ファクトチェック機関その他のファクトチェック関連団体又は日本の行政機関(※)との間で情報共有等の連携を実施している場合は、その詳細を御教示下さい。 ※ 例えば、内閣府(防災担当)、警察庁など。
A	地震期間中、内閣府、総務省、警察庁、および多数の地方自治体など、さまざまな政府機関や代表者と連携しました。その際、人工地震説に関する誤情報やデマ情報に対処するコミュニティノートなど積極的に情報共有しました。
Q 7.5	地震その他の災害に関連する偽・誤情報の流通・拡散への対応に関し、今後(さらに)体制を強化する予定・余地があれば、強化する内容を御教示下さい。
A	今後は、この変化しつづける課題に対応するためにNPOや行政との連携し、それぞれの緊急コミュニケーションのサポートを検討しております。
Q 7.6	地震その他の災害に関連する偽・誤情報の流通・拡散への対応に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。
A	私たちは、表現の自由を保護するとともに情報の正確性を確保するバランスを取る中で、重要な問題についてユーザー自身が判断を下すことが重要であると考えています。これが、X上の人々がコンテキストを追加する集団的なプロセスに貢献することを可能にするコミュニティノートというプロダクトの背後にある哲学です。これは、X自身のコンテンツモデレーション作業に加えて行われます。

	8. レコメンドや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について
Q 8.1	対象偽・誤情報を含む対象サービス上で表示される日本語のコンテンツについて、アルゴリズムを用いてモデレーション等の要否・内容を自動的に決定している場合、当該決定に当たって用いられる主なパラメータ及び当該決定の過程(当該パラメータが当該決定にどのように寄与しているのかなど)を御教示下さい。
A	<p>自動コンテンツモデレーション</p> <p>Xは、プラットフォーム上で施行されるXのルールとポリシーに違反するコンテンツを自動的に検出するために、ヒューリスティックと機械学習アルゴリズムの組み合わせを使用しています。</p> <p>機械学習モデル</p> <p>私たちは、自然言語処理モデル、画像処理モデル、その他の高度な機械学習手法の組み合わせを使用して、違反の可能性があるコンテンツを検出します。これらのモデルは複雑さと生成する出力において異なります。例えば、プラットフォーム上での虐待を検出するために使用されるモデルは、過去に検出された虐待違反に基づいて訓練されています。これらの機械学習モデルによってフラグが立てられたコンテンツは、アクションが取られる前に人間のコンテンツレビュアーによってレビューされるか、場合によってはモデルの出力に基づいて自動的に処理されます。</p> <p>ヒューリスティックモデル</p> <p>ヒューリスティックは、プラットフォーム上で新しい形式の違反が出現した際に、Xが迅速に反応できるようにするために一般的に利用されます。ヒューリスティックは、特定のカテゴリーの違反に典型的なテキストやキーワードの一般的なパターンです。ヒューリスティックによって検出されたコンテンツも、コンテンツに対して行動が取られる前に人間のコンテンツレビュアーによってレビューされる場合があります。これらのヒューリスティックは、コンテンツを人間のエージェントによるレビューのためにフラグを立て、そのようなコンテンツがレビューされる順序を優先するために使用されます。</p>
Q 8.2	Q8.1で御回答頂いたパラメータを用いてモデレーション等の要否・内容を自動的に決定していることについて、既に公開している規約等の中で説明している場合は、当該規約等のタイトル及びURLを付記して、該当箇所を抜粋下さい。
A	現時点でご提示できる情報はございません。
Q 8.3	対象サービス上で表示される日本語のコンテンツ(広告を含みます)について、アルゴリズムを用いて表示先や表示順位等のレコメンドの要否・内容を自動的に決定している場合、当該決定に当たって用いられる主なパラメータ及び当該決定の過程(当該パラメータが当該決定にどのように寄与しているのかなど)を御教示下さい。
A	Xは おすすめ の タイムラインのアルゴリズム を開示しています。また、例えば こちら のように、今後も様々なアップデートをご提供します。
Q 8.4	Q8.3で御回答頂いたパラメータを用いてコンテンツの表示先や表示順位等のレコメンドの要否・内容を自動的に決定していることにつき、既に公開している規約等で説明している場合は、当該規約等のタイトル及びURLを付記し、該当箇所を抜粋下さい。

A	Twitter's Recommendation Algorithm
Q 8.5	<p>Q8.1及び8.3のアルゴリズムについて、透明性・アカウントビリティ確保に向けた取組の一環として、特定の第三者(※1)に対して特定の目的(※2)で開示している場合、当該第三者の名称又は類型その他の詳細(※3)を御教示下さい。</p> <p>※1 例えば、秘密保持契約を締結した研究機関や研究者など。</p> <p>※2 例えば、透明性・アカウントビリティ確保状況に関する監視・研究目的など。</p> <p>※3 例えば、開示目的、開示に至った経緯など。</p>
A	<p>Twitter's Recommendation Algorithm</p> <p>こちらを公開しており、研究者を含む誰でも私たちのブログやGitHubを通じて確認することができます。</p>
Q 8.6	<p>Q8.4で既に関示しているか否かにかかわらず、Q8.1及び8.3のアルゴリズムについて、今後、特定の第三者に対して特定の目的で開示することの可否・継続性に関するお考えがあれば、お聞かせ下さい。</p>
A	<p>Xは今後も、この分野における革新的な取り組みについて情報を共有します。</p>

9. 広告エコシステムの透明性・アカウントビリティ確保に向けた取組について	
Q 9.1	<p>対象サービスのうち、広告収入を主たる運営基盤としているものについて、日本に関するコンテンツの発信者(投稿者)へ広告収入の分配を行っている場合、分配の基準を御教示下さい。特に、コンテンツのPV数やいいね、リポスト等の拡散行為の対象となった数という定量的な基準以外の定性的な基準(※)や、災害発生時を含む有事において適用される特別な基準が存在する場合は、当該定性的な基準や特別な基準を具体的に御教示下さい。</p> <p>※ 例えば、コンテンツに偽・誤情報が含まれる場合には基準となるPV数等から除外する、そのようなコンテンツには広告を掲載しないなど。</p>
A	<p>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。</p> <p>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。</p> <p>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：</p> <ul style="list-style-type: none"> ・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。 <p>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。</p> <p>Blog comment on Ads Revenue Sharing</p>
Q 9.2	<p>Q9.1で御回答頂いた内容を公開したり、普及啓発に向けた取組を実施したりしている場合は、その方法・内容(※)を御教示下さい。</p> <p>※ オンライン上で公開・普及啓発等している場合はURLを付記下さい。</p>
A	<p>以下のようにXは情報を公開しております。</p>

[Ads Revenue Sharing](#)
[Ads Revenue Sharing Terms](#)
[Blog comment on Ads Revenue Sharing](#)

10. AI・ディープフェイク技術への対応状況について

Q 10.1	上記4で御回答頂いたもののほか、対象サービスに関連してAIシステムを開発し、又は対象サービスにAIシステムを組み込んで第三者に提供し、若しくは自ら利用している事例がある場合は、その概要を御教示下さい。
A	現時点でご提供できる情報はございませんが、今後アップデートの際は共有させていただきます。なお、A 10.4もご参照ください。
Q 10.2	Q10.1で御回答頂いたAIシステムの開発・提供・利用に関し、総務省及び経済産業省が策定・公表した「AI事業者ガイドライン案」(※1)において、「生成AIによって、内容が真実・公平であるかのように装った情報を誰でも作ることができるようになり、AIが生成した偽情報・誤情報・偏向情報が社会を不安定化・混乱させるリスクが高まっていることを認識した上で、必要な対策を講じる」ことが求められていることを受けて、何らかの対策(※2)を講じている場合は、その概要を御教示下さい。 ※1 AI戦略会議第7回(2023年12月21日)資料1.3 https://www8.cao.go.jp/cstp/ai/ai_senryaku/7kai/13gaidorain.pdf ※2 例えば、生成AIが事実と異なることをもっともらしく回答する(ハルシネーション)リスクへの対策としての開発上の工夫、AIシステム提供時におけるハルシネーションリスクに関する透明性・アカウントビリティ確保の取組(ユーザーへの情報開示等)、AIの学習過程を含む開発メンバーや提供先・利用者のリテラシー向上に向けた取組など。
A	Xの利用規約とルールを通じ、ポリシー運用を行なっています。 なお、現在私たちはユーザーがAIを使用してビデオ、画像、および音声を作成する機能は提供しておりません。
Q 10.3	対象サービス上でAIが生成した日本に関するコンテンツが投稿された場合に、これを検知するツールや、投稿時のラベリングを投稿者に義務付ける規約等を導入している場合は、その詳細及び実効性担保に向けた工夫(もしあれば)を御教示下さい。
A	私たちは、合成メディアや操作されたメディアなど今後予測されるリスクに対抗するための ポリシー を持っています。 また、メディアマッチングによりコミュニティノートがはるかに多くの投稿に表示されるようになっていきます。最近では、「メディアに対するノート」を導入・強化しており、写真やビデオにノートが追加されると、マッチングするメディアを含む他の投稿に自動的に表示されるようになっていきます。場合によっては、個別のノートが数百または数千の投稿と一致することもあります。
Q 10.4	以上御回答頂いたほか、対象サービス上でAIが生成した日本に関するコンテンツが投稿された場合に対応するため、他のAI関連事業者(開発者・提供者・利用者)との間で連携・協力に向けた取組(※)を行っている場合は、その詳細を御教示下さい。 ※ 例えば、情報共有、資金援助、共同開発その他の技術提携・実装など。
A	私たちはAI選挙協定の共同署名者です。これらの問題は一つの企業によって解決されることはなく、産業間の協力が重要であると私たちは考えています。

<https://www.aielectionsaccord.com/>
<https://twitter.com/GlobalAffairs/status/1758546745015808421>

11. ファクトチェックの推進に向けた取組について

Q11.1 上記4及び5で御回答頂いたもののほか、日本国内に所在するファクトチェック機関その他のファクトチェック関連団体(※1)との間で連携・協力に向けた取組(※2)を行っている場合は、その詳細を御教示下さい。
※1 日本国内に向けたサービスを提供する等の活動をしているものを含みます。
※2 例えば、情報共有、資金援助、ファクトチェックを容易にするツールの開発や技術提携・実装、人材育成・交流、対象サービス上でのコンテンツの優先的表示、イベントの共催・後援など。

A 特にございませぬ。

Q11.2 ファクトチェック機関その他のファクトチェック関連団体との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

A 特にございませぬ。

Q11.3 その他、日本国内におけるファクトチェックの推進の観点から行っている具体的な取組(※)があれば、詳細御教示下さい。
※ 例えば、ファクトチェックを容易にするツールの開発・提供、ファクトチェックを実施する人材の育成など。

A 特にございませぬ。

12. マスメディア(新聞・放送)との連携状況について

Q 12.1 上記4及び5で御回答頂いたもののほか、日本国内に所在するマスメディア(新聞・放送)(※1)との間で連携・協力に向けた取組(※2)を行っている場合は、その詳細を御教示下さい。
※1 日本国内で活動しているものを含みます。
※2 例えば、情報・ノウハウの共有、資金援助、人材交流、対象サービス上でのコンテンツの優先的表示、イベントの共催・後援など。

A 特にございませぬ。

Q 12.2 マスメディアとの連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

A 特にございませぬ。

13. 利用者のICTリテラシー向上に向けた取組について

Q13.1 日本国内に所在する教育機関・普及啓発機関や消費者団体・利用者団体(※1)との間で連携・協力に向けた取組(※2)を行っている場合は、その詳細を御教示下さい。
※1 日本国内で活動しているものを含みます。
※2 例えば、対象サービスの利用者のICTリテラシー向上に関する定期的な意見交換、教育教材の共同開発、教育・普及啓発イベントの共催・後援など。

A Xは最近においては、総務省の「ICT活用のためのリテラシー向上に関する検討会」へ参加しており、今後も引き続き、日本でのさらなる取り組みを検討していきます。

Q 13.2	教育機関・普及啓発機関や消費者団体・利用者団体との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。
A	特にございません。
Q 13.3	その他、日本国内における対象サービスの利用者のICTリテラシー向上の観点から行っている具体的な取組(※)があれば、詳細御教示下さい。 ※ 例えば、教材作成、普及啓発活動など。
A	特にございません。
14. 研究開発の推進に向けた取組について	
Q 14.1	上記10で御回答頂いたもののほか、対象サービスに関連して、日本国内に所在する研究機関・研究者・アカデミア(※1)との間で、特に認知科学、心理学、情報工学、計算社会科学、法学などの分野において連携・協力に向けた取組(※2)を行っている場合は、その詳細を御教示下さい。 ※1 日本国内で活動しているものを含みます。 ※2 例えば、情報共有、資金援助、共同開発その他の技術提携・実装、イベントの共催・後援など。
A	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
Q 14.2	研究機関・研究者・アカデミアとの連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。
A	特にございません。
Q 14.3	研究機関や研究者等向けに、対象サービス上の情報流通の状況に関する実データをAPI等を通じて提供している場合は、提供先その他の詳細(※)を御教示下さい。 ※ 例えば、提供開始(予定)時期、対価の額等の提供条件など。
A	機密保持等の理由から、APIの提供先等に関する情報はご共有できません。
Q 14.4	Q14.3で提供していない場合において、過去に提供していたことがある場合は、提供先、提供期間及び提供を終了した理由を御教示下さい。
A	Xが公にかつオープンに伝えている通り、当社は破産の危機に直面しており、事業を安定させる必要がありました。 また、無料APIを商業目的で使用している人も多数いました。 それでもなお、私たちはこのコミュニティに引き続きサービスを提供する新しい方法を探っています。 https://x.com/XDevelopers/status/1641222788911624192?s=20 なお、機密保持等の理由から、APIの提供先等に関する情報はご共有できません。
Q	Q14.3で提供していない場合(Q14.4の場合を含みます。)、将来的に提供を開始(再開)することの可否について

14.5	てお考えをお聞かせ下さい。
A	現時点でご提示できる情報はございません。
Q 14.6	その他、対象サービスに関連した研究開発の推進の観点からの具体的な取組(※)があれば、詳細御教示下さい。 ※ 例えば、電子透かし技術の開発、ディープフェイク検知・対抗技術の開発、自動ファクトチェック技術の開発など。
A	Xは、サービスの特定の側面をオープンソースにし、レビュー可能にいたしました。例えば、Xは おすすめのタイムラインのアルゴリズム を公開しています。また、コミュニティノートのアロリズムと データ も公開しています。
15. サイバーセキュリティ関係機関等との連携について	
Q 15.1	サイバーセキュリティ関係機関(※1)との間で、対象サービスに関連して連携・協力に向けた取組(※2)を行っている場合は、その詳細を御教示下さい。 ※1 日本国内に向けたサービスを提供する等の活動をしているものを含まます。 ※2 例えば、情報共有、資金援助、共同開発その他の技術提携・実装、イベントの共催・後援など。
A	私たちは、 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まれます。 さらに、 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しています。
Q 15.2	サイバーセキュリティ関係機関との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。
A	特にございません。
Q 15.3	その他、対象サービス上の対象偽・誤情報の流通・拡散への対応として、サイバーセキュリティ対策との連携の観点から行っている具体的な取組(ISACへの参加など)があれば、詳細御教示下さい。
A	現時点でご提示できる情報はございません。
16. 行政機関や地方公共団体等との連携について	
Q 16.1	日本の行政機関や地方公共団体等との間で連携・協力に向けた取組(※)を行っている場合は、その詳細を御教示下さい。 ※ 例えば、情報共有、対象サービス上でのコンテンツの優先的表示、イベントの共催・後援、官民連絡会議(米国におけるGlobal Internet Forum to Counter Terrorism (GIFCT)や選挙のintegrityに関するワーキンググループ(https://perma.cc/JC5V-MARG)に相当するものなど)の設置など。
A	警察庁や各都道府県警、またその他省庁と随時連携しております。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が含まれます。

	<p>また、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を提供しています。 https://x.com/BizJP/status/1691314449817473026?s=20</p>
Q 16.2	<p>行政機関や地方公共団体等との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。</p>
A	<p>特にございません。</p>
17. 国際機関等との連携について	
Q 17.1	<p>国際機関や専門機関等との間で連携・協力に向けた取組(※)を行っている場合は、その詳細を御教示下さい。 ※ 例えば、情報共有、対象サービス上でのコンテンツの優先的表示、イベントの共催・後援など。</p>
A	<p>国際機関との連携の一端としては、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があります。</p>
Q 17.2	<p>国際機関や専門機関等との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。</p>
A	<p>特にございません。</p>
Q 17.3	<p>その他、国際連携の観点から行っている具体的な取組(国際会議への出席など)があれば、詳細御教示下さい。</p>
A	<p>17.2の情報に加えて、私たちがGIFCTおよびクライストチャーチ・コールのメンバーであるなど、他の協力例もご紹介します。</p> <p>詳細は以下の通りです。</p> <p>テロ関連の違反コンテンツを特定するため、私たちは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、さまざまな方法を取っています。私たちは、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します。これには、デジタルフィンガープリント技術やAI技術の発展や導入が含まれます。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しています。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私たちの取り組みや実施したアクションについては、こちらから詳しく確認できます。私たちの技術への投資は続けられ、この種のコンテンツを報告する必要性を減らしています。</p>
18. その他のステークホルダーとの連携状況について	
Q 18.1	<p>日本国内に所在するその他のコンテンツ制作主体(※)との間で連携・協力に向けた取組を行っている場合は、その詳細を御教示下さい。</p> <p>※1 ジャーナリスト、クリエイター、ネットメディアなど。日本国内で活動しているものを含みます。 ※2 例えば、情報・ノウハウの共有、資金援助、人材交流、対象サービス上でのコンテンツの優先的表示、イベ</p>

	ントの共催・後援など。
A	現時点でご提示できる情報はございません。
Q18.2	日本国内に所在する他の電気通信事業者(※)との間で、ゼロレーティングサービスの享受などを通じた連携を行っている場合は、その詳細を御教示下さい。 ※ 日本国内で事業を運営しているものを含みます。
A	該当するものはございません。
Q 18.3	日本国内に所在する他のプラットフォーム(※1)事業者やメタバース関連事業者(※2)と連携・協力に向けた取組(※3)を行っている場合は、その詳細を御教示下さい。 ※1 SNS、検索サービス、動画共有サービス、ブログ・掲示板、ニュースポータル、キュレーションなど。 ※2 日本国内で事業を運営しているものを含みます。 ※3 例えば、情報共有、技術提携・実装、人材育成・交流、イベントの共催・後援など。
A	一般社団法人ソーシャルメディア利用環境整備機構(SMAJ)における偽・誤情報の課題検討等に参加しています。
Q 18.4	Q18.1～18.3のステークホルダーとの連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。
A	特にございません。
19. 特に紹介したい取組について	
	Q19.1 以上御回答頂いたほか、偽・誤情報の流通・拡散への対策をはじめ、デジタル空間における情報流通の健全性確保に向けた貴社の取組として共有可能なものがございましたら、当該取組の詳細(※1)を御教示下さい(※2)。 ※1 例えば、当該取組を実施するに至った背景・課題、当該取組による効果など。 ※2 この質問への御回答内容は、前記【御回答に当たっての留意事項】(c)のとおり、本検討会において「偽・誤情報対策に係る取組事例集(仮称)」としてとりまとめ、公開させて頂く予定です。
A	これらのトピックと問題に対処するために、最も重要な対策・開発の一つであるコミュニティノートをご紹介します。 コミュニティノート は、多くのユーザーが協力して、役に立つ背景情報をポストに追加し、他のユーザーへ十分な情報を提供するためのプログラムです。 コミュニティノートは、誤解を招く可能性があるポストに、Xユーザーが協力して役に立つノートを追加できるようにすることで、より正確な情報を入手できるようにすることを目指しています。 協力者はノートを作成/評価します 協力者は、ノートの作成と評価を志願してコミュニティノートに 登録 した、通常のXユーザーが務めています。参加者が増えるほど、このプログラムはより優れたものになります。 人々が多様な視点から「役に立つ」と評価したノートだけがポストに表示されます

	<p>コミュニティノートは多数決に基づきません。幅広いユーザーにとって役に立つノートを特定するため、過去の評価において、評価が相違することのあった協力者の間で「役に立つ」という評価が一致することがノートには求められます。これにより評価の偏りを防止できます。</p> <p>コミュニティノートでの多様な視点の取り入れ方の詳細についてはこちらをご覧ください。</p> <p>表示されるノートをXが決めることはありません。決めるのはユーザーです Xがノートの作成や評価を行ったり、ノートに介入したりすることはありません(Xルールに違反している場合を除きます)。Xでは、表示されるノートをユーザーの意見によって決めることは、正確な情報を入手するのに役立つ背景情報を追加するための公平で効果的な方法であると考えています。</p> <p>オープンソースと透明性 利用者がコミュニティノートの仕組みについて理解したり、コミュニティノートの成り立ちに貢献できたりすることは重要です。このプログラムでは透明性が重視されているため、すべての協力者は毎日公表され、ランキングアルゴリズムはどなたでも調査できます。この仕組みの詳細についてはこちらをご覧ください。</p>
20. その他	
	<p>Q20.1 以上御回答頂いたほか、デジタル空間における情報流通の健全性確保に関する課題や今後の対応にあたっての基本的な考え方、具体的な方策などについて、御意見や補足等ございましたら御教示下さい。</p>
A	<p>Xは、不正行為や操作から私たちのプラットフォームの信頼性を守ることを目指しています。悪意ある操作を難しくするため、私たちは多角的な取り組みを展開しており、これには新しい認証システムやサブスクリプションサービスの導入、政府関連アカウントに特別なグレーバッジを提供すること、またコードを公開してX上で情報がどのように流通するかをユーザーが直接確認できるようにする努力が含まれます。</p> <p>私たちのポリシーは、なりすましや誤解を招くアカウント、偽造アカウントの禁止、スパムや不正な操作への対抗措置など、情報の真実性を守るための幅広い分野を網羅しています。</p> <p>自由な表現を保護しつつ情報の真実性を保つため、ユーザーが自ら重要な問題に関して判断を下せるよう重視しています。このアプローチの一環として、特定の状況では合成メディアや操作されたメディアにはラベルをつけることを選択しています。コミュニティノートはこの考え方に基いており、Xのコンテンツモデレーションに加え、ユーザーが情報に文脈を加える共同作業に貢献することを可能にします。</p> <p>コミュニティノートは、使用方法とプログラムコードを完全に公開しており、その仕組みを深く理解することができます。</p> <p>さらに、Xはデジタルリテラシーやメディアリテラシーの向上にも取り組んでいきます。情報を得るためのサービスはXだけに限らず多岐にわたるため、包括的なリテラシー向上の取り組みは非常に重要なものと考えています。</p>