

マイクロソフト・コーポレーション・LinkedIn

1 前提：対象サービスの範囲について

【想定対象サービス】

SNS	LinkedIn
検索	Bing 検索
動画共有	

ニュースポータル・ キュレーション	Microsoft Start
ブログ・掲示板	

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。

パート1:マイクロソフト関係

対象サービスについて:Bing と Start の概要

1. Bing について

Bing 検索は、Web 上の最も関連性の高い検索結果をユーザーに提供することを主な目的としているオンライン検索エンジンです。ユーザーは、特定の項目について調べるために Bing にアクセスすると、検索キーワードに対応した、インターネット上の最も関連性が高い、権威のあるサードパーティの Web サイトへのリンクが提供されることを期待しています。

Bing は、サードパーティ コンテンツの検索エンジンおよびインデクサーであり、コンテンツをホストしません。Bing は、ユーザー向けのニュース フィード機能を持たず、ユーザーがコンテンツを投稿して共有することを認めておらず、コンテンツを「バズらせる」こともありません。そのため、オーガニック検索結果の誤情報に対処するに

は、他のタイプのオンライン サービスに適したものと異なるアプローチが必要になる場合があります。

Bing 検索では、オーガニック検索の結果がランク付けされる際に、信頼性が高い権威あるニュースや情報が最初に表示されるようにするとともに、ユーザーが特定のサイトの信頼性を評価するために役立つツールも提供しています。同時にマイクロソフトは、ユーザーが検索エンジンを通じてあらゆる種類の情報を見つけられるようにすることは、重要な公共の利益をもたらすとも考えています。さらに、Bing ユーザーは、さまざまな正当な理由（調査目的など）により、他の文脈では有害または攻撃的とされ得るコンテンツを検索で探す場合もあります。情報に自由にアクセスする権利をはじめとする、ユーザーの権利のバランスを保つために、マイクロソフトでは以下の原則を採用しています。

1) マイクロソフトは、ユーザーからのクエリに関連する、信頼性が高い権威ある結果を提供します。

- マイクロソフトは、ユーザーの検索キーワードに関連する、最高品質の権威あるコンテンツを提供します。
- マイクロソフトは、公平で、バランスの取れた、包括的なコンテンツを常に提供することを目標としています。複数の信頼できる視点がある場合は、それらを有益な方法で表示するように努めます。権威ある情報源がない場合には、偏った情報や誤解を招く可能性のある情報を広めないようにします。
- マイクロソフトは、ユーザーの意図を尊重します。ユーザーが特定の情報にアクセスしたいという明確な意図を表明した場合は、たとえ信頼性が低くても、関連性の高い結果を提供し、（以下に詳述するように）ユーザーがそのような検索結果に惑わされないように努めます。

2) マイクロソフトは、表現の自由、プライバシー、公共安全などに関わる現地の法律やその他の基本的権利を尊重しながら、法律の範囲内で、自由にオープンな情報へのアクセスを推進します。

- マイクロソフトは、可能な限り多くの Web サイト へのオープンなアクセスを提供しますが、コンテンツが現地の法律やマイクロソフトのポリシーに違反しているなどの限られたケースにおいては、特定の介入（Web サイトの検索結果からの削除やランキングの降格など）を行うことがあります。
- マイクロソフトがコンテンツへのアクセスを制限する場合には、マイクロソフトの措置を特定の範囲に限定し、表現の自由、情報へのオープンなアクセス、メディアの多元性などの重要な利益を不当に制限することのないよう努め、透明性を確保します。

3) マイクロソフトは、有害なコンテンツや予期せぬ攻撃的なコンテンツからユーザーを保護するための措置を講じています。

- マイクロソフトは、さまざまな理由から、有害なコンテンツや物議を醸すようなコンテンツについて調査または確認したいと考えるユーザーがいることを認識していますが、そのようなコンテンツによってユーザーが惑わされることがないようにすることの重要性も認識しています。
- 検索結果に有害な情報や誤解を招く情報が含まれている可能性があるとしてマイクロソフトが判断した特定の種類のコンテンツについては、警告または公共サービス広告などにより補足情報を提供し、潜在的なリスクについてユーザーに通知する場合があります。マイクロソフトは、ユーザーが Bing で見つけるコンテンツの種類を、SafeSearch や Family Safety などの機能を通じてコントロールできるようにしています。

- 特定のコンテンツにアクセスしたいというユーザーの明確な意図が認められない限り、マイクロソフトではユーザーが権威ある結果を見つけたいと考えているものと見なします。

4) マイクロソフトの原則とプラクティス、そして意思決定と行動については、高い透明性を保っています。

- マイクロソフトでは、検索結果のランキングと関連性に関するマイクロソフトの原則、およびマイクロソフトのモデレーション ポリシーに関する情報をユーザーに提供しています。
- マイクロソフトがコンテンツへのアクセスを制限する場合は、コンテンツが検索結果から削除されたことを適宜ユーザーに通知しています。
- マイクロソフトが受け取ったクレームと実施した措置に関する情報をまとめた透明性レポートも、定期的に発行しています。

Bing の検索アルゴリズムによる検索結果でのページのランク付け方法

Bing では、多くのパラメータを使用し、検索結果の各ページをランク付けしています。以下に、情報の完全性を維持するための主なパラメータの概要を示しますが、すべてのパラメータを取り上げているわけではありません。これらのパラメータは、一般的に重要性が高いと思われるものから順に記載されています。Bing の複雑なシステムは、これらの基準に従って検索結果を提供しています。以下の各パラメータの相対的な重要性は、検索ごとに異なり、時間の経過とともに変化する場合があります。

- **関連性:** 関連性とは、ランディング ページ内のコンテンツが、検索クエリを実行したユーザーの意図にどれだけ近いかを意味します。それには、ページ上で直接マッチするキーワードや、ページを参照するリンクで使われている語句などが含まれます。Bing はまた、クエリのキーワードと完全には一致していなくても同じ意味だと考えられる、同義語や略語などの単語も考慮に入れます。多くのクエリには、2 つ以上の意図が込められている場合があります。Bing では、すべての考えられる意図を反映した包括的な結果セットを提供するよう努めています。ユーザーが低品質のコンテンツを調査したいという意図を明確に示さない限り、ユーザーは高品質な権威あるコンテンツを求めているものと見なします。
- **品質と信頼性:** Web サイトの品質と信頼性 (QC) を判定する際には、サイトが明確な目的のもとに作成されており、使いやすく見栄え (Presentation) が良いかどうか評価されます。また、QC に関しては、以下のようページの「オーソリティ (権威)」も評価されます。
 - **評判:** 他のどのような Web サイトがそのサイトにリンクしているか? 有名なニュース サイトは、まったく新しいブログよりも高い評価を得ていると見なされます。
 - **議論のレベル:** そのコンテンツは、個人や集団に危害を加えることのみを目的としているか? たとえば、暴力を助長したり、悪口やいじめの温床となったりしているサイトは、バランスの取れたニュース記事と比べて議論のレベルが低く、したがって権威も低いと見なされます。
 - **歪曲のレベル:** そのサイトは事実と意見をどれだけ区別しているか? 風刺やパロディであることが明確に表示されているサイトは、その意図を曖昧にしようとしているサイトよりも権威があります。
 - **情報所有者の起源と透明性:** そのサイトは、直接入手した情報を掲載しているか、それとも他者のコンテンツを要約または再掲載しているか? オリジナルのコンテンツを掲載していない場合は、情報の出典を明記しているか? 個人のブログに直接入手した情報が掲載されている場合は、出典が明記されていないコンテンツよりも権威があると見なされることがあります。

違法コンテンツへの対応 - 児童の性的搾取・暴行画像および関連コンテンツ

児童の性的虐待に関するコンテンツの作成、配布、およびそれらへのアクセスは、例外なく非難され、ほとんどの法域で違法とされています。また、Bing およびマイクロソフトのポリシーにも違反しています。Bing は、サードパーティのテクノロジーおよび業界団体、法執行機関、政府および非政府組織と協力し、こうした忌まわしいコンテンツがオンラインで拡散されるのを防ぐために取り組んでいます。その 1 つの方法として、信頼できる機関による審査を実施して（またはマイクロソフトの PhotoDNA コンテンツ検出ツールで特定して）、児童の性的搾取や虐待に直接または間接的に関連するコンテンツを含んでいることが判明したページを検索結果から削除しています。

特に、Internet Watch Foundation (英国)、NCMEC (米国)、FSM (ドイツ) が、その誠実な判断により、児童の性的虐待のコンテンツをホストしているか、それらへのアクセスを提供していると認定したページは、検索結果から削除しています。表示される検索結果からこれらのリンクを削除しても、Web 上でそうしたコンテンツにアクセスしたり、Bing 以外の手段で見つけたりすることはできますが、そのようなページを探し出したり、そこから利益を得たりしようとする人々がページを利用しづらくなることは確かです。

Bing はまた、PhotoDNA のハッシュマッチング技術を使用し、ユーザーがアップロードした画像をビジュアル検索機能でスキャンして、搾取・虐待の可能性がある画像かどうかを確認しています。Bing におけるユーザーの活動は、Microsoft サービス規約によって管理されます。マイクロソフトおよび Bing による PhotoDNA の使用と、コンテンツ モデレーションの実施に関する詳細については、[Digital Safety | Moderation and enforcement](#) をご覧ください。

マイクロソフトは、特定のカテゴリでマイクロソフトが受領し、該当する期間中に承認した削除リクエストの数などに関する透明性レポートを 6 か月ごとに公表しています。リクエストの詳細については、[Reports Hub | Microsoft CSR](#)をご覧ください。

情報の完全性の維持

Bing では、独立した研究機関や非営利団体と提携し、脅威インテリジェンスを維持して潜在的なアルゴリズム介入に関する情報を提供するための取り組みを定期的実施しています。

また、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検索キーワードを削除しています。Bing 検索ではさらに、検索されたトピックに関する権威ある情報にユーザーを誘導するための回答や公共サービス広告を検索結果の上部に表示したり、有害な情報を含むことが知られている特定の URL（認定されていないオンライン薬局やマルウェアを含むサイトなど）に関する警告を表示したりすることもあります。

検索結果ページの右下隅からアクセスできる Bing 検索の「フィードバック」ツールでは、ユーザーが検索結果に関するフィードバック（検索結果ページのスクリーンショットなど）を Bing 検索に提供することができます。フィードバックの性質に応じて、Bing 検索では、防御的な検索介入、Web サイトの掲載中止、自動サジェストのキーワードの削除などの適切な措置を講じる場合があります。

2. Start について

Microsoft Start は、パブリッシャーからのニュースや情報コンテンツを表示する、パーソナライズされたフィードであり、独立した Web サイト (MSN.com)、Android および iOS のモバイル アプリ、Windows 10 のタスク バーの [ニュースと関心事項] エクスペリエンス、Windows 11 のウィジェット エクスペリエンス、Microsoft Edge の新規タブ ページなど、マイクロソフトのさまざまな製品や機能で利用できます。

Start は、マイクロソフトだけでなく、増え続けるシンジケーション パートナーのためにも、Web やモバイルのエクスペリエンスを通じて高品質なニュースを配信しています。Microsoft Start のモデルは、権威の低いコンテンツが伝播されるリスクを低減します。マイクロソフトのコンテンツ プロバイダーは精査されており、偽情報、プロパガンダ、意図的な誤報を禁止する一連の厳格な基準を遵守しなければなりません。Microsoft Start は無料でダウンロードでき、ユーザーが閲覧できる記事やビデオの数に制限はありません。

Microsoft Start の全ページ (ランディング ページと各記事、下記参照) の下部にフィードバック機能があり、ドロップダウン メニューのオプションの 1 つとして [Content Quality] があります。このフィードバック機能は、[設定] メニューにも含まれています。

マイクロソフトの全製品に適用される情報の完全性に関する原則

2022 年 6 月、マイクロソフトは最初の [情報の完全性に関する原則](#) を発表しました。これらの原則は、対象となるマイクロソフトのすべての製品とチームに適用され、社内全体で情報の完全性を確保するための取り組みを進めながら、幅広い多様性を受け入れられるよう認識を高めるためのものです。これらの原則では、各チームがポリシー、製品開発およびリスク評価作業に関する詳細を伝えるために使用できる、基盤となる一連のコミットメントを定めています。情報の完全性に関する 4 つの原則は、以下のとおりです。

- **表現の自由:** マイクロソフトは、表現の自由を尊重するとともに、マイクロソフトのプラットフォーム、製品、サービスでお客様が常に情報を作成、公開、検索できるようにします。
- **権威あるコンテンツ:** マイクロソフトの製品では、内部データおよび信頼できるサードパーティのデータを使用した、他国のサイバー影響工作に影響されないコンテンツの表示を優先します。
- **利益の不獲得:** マイクロソフトは、他国のサイバー影響工作に関連するコンテンツやそのアクターから意図的に利益を得ることがないようにします。
- **プロアクティブな取り組み:** マイクロソフトのプラットフォームおよび製品が、他国のサイバー影響工作に関連するサイトやコンテンツの拡散に使用されないようにプロアクティブに取り組めます。

Microsoft Start コミュニティは、2020 年 9 月に導入され、日本では 2021 年 12 月に開始されました。このコミュニティは、各種の問題や出来事に関する多様な信頼できる議論とコンテンツをサポートします。このコミュニティは、安全かつ包摂的な、互いを尊重し合うフォーラムであり、各参加者は、自分の投稿内容や他の参加者への対応に責任を持ちます。[コミュニティ ガイドライン](#)は、このような価値観を守るために定められており、マイクロソフトはその遵守方法について、透明性を確保し、明確なガイダンスを提供するよう努めています。

- 特定の投稿にフラグが立てられた場合、その投稿は審査されます。その投稿がコミュニティ ガイドラインに反する場合は削除されます。
- コメントが削除された場合は、その旨がユーザーのアクティビティ フィードに表示され、ユーザーはその決定に異議を申し立てることができます。

必要に応じて、Microsoft Start は、特定のユーザーのコメント機能を停止する場合があります。ユーザーが基

準に従うことを拒否し続けると、永久追放処分が下される場合もありますが、ユーザーは異議を申し立てることができます。Start コミュニティ ユーザーは、MS Start の自身のユーザー プロフィール ページにアクセスすれば、どのコメントがコミュニティ ガイドラインに違反して削除されたかを確認し、異議を申し立てることができます。現在のところ、メールによる通知はサポートしていません。

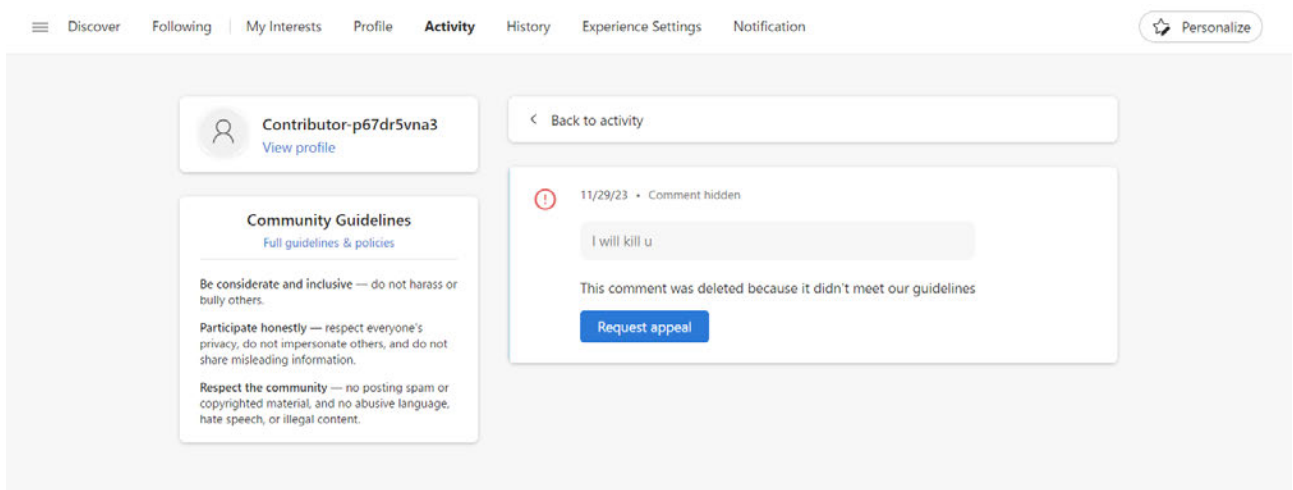
コミュニティ投稿のユーザー レポートからコメントが削除された場合、コミュニティ ユーザーは異議を申し立てることができます。異議申し立ては、MS Start のコメント モデレーション チームによって審査され、そのコメントがコミュニティ ガイドラインに違反していないと判断された場合には、コメントが復活します。

コミュニティ ユ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、日本語でもご利用いただけます。[Trusted Community Engagement and Contributions | Profile Activity \(msn.com\)](https://www.msn.com/en-ca/community/activity?profileStatus=true)



ユーザーは、MS Start のプロフィール ページにある [異議申し立てを要求する] ボタンを押すことで、プラットフォーム投稿のユーザー報告により削除されたコメントに対する異議申し立てを行うことができます。例:

<https://www.msn.com/en-ca/community/activity?profileStatus=true>



ユーザーは、MS Start に掲載されているコミュニティ ガイドラインで異議申し立てのプロセスに関する詳細を確認できます。[信頼できるコミュニティの関与と貢献 \(msn.com\)](#)。

このコンテンツは日本語でも提供されています。

コミュニティのガイドライン

基本 ガイドライン **適用**

ポリシーの概要

Microsoft の強制ポリシーは、互いを尊重しあう、健全でアクティブなコミュニティを奨励することを目的としています。警告とアカウントの制限はすべてに適合するわけではありません。Microsoft のシステムでは、これらの決定に到達するためにいくつかの要因を考慮に入れています。正確な基準は公開されていませんが、違反の重大度と頻度に基づいて警告とアカウント制限が表示されます。アカウントのアクティビティが制限されている間は、アカウント制限期間中はコンテンツにコメントしたり、コンテンツを公開したりすることはできません。特定の重大な違反が発生すると、アカウントの即時および永続的な制限につながる可能性があります。

違反

投稿またはコメントにフラグが設定されている場合は、コミュニティガイドラインに違反しているかどうかを判断するためにレビューされます。これらのガイドラインを満たしていない場合は削除され、決定に異議を申し立てる方法に関する情報と、コミュニティ標準のリマインダー (該当する場合) が通知されます。

違反が繰り返し発生した場合、アカウントの取り消しが行われる可能性があります。これにより、コンテンツを一定の時間、コミュニティにコメントまたは公開する機能が制限されます。これらの取り消しは、無期限にアカウントに残ります。警告と違反が繰り返された後も、Microsoft のガイドラインへの違反を続けられた場合、または重大な違反を投稿した場合は、お客様のアカウントを無期限に制限する場合があります。

異議申し立て

異議申し立てを要求すると、モデレーターのチームがコメントを確認します。コメントが誤って削除されたと判断した場合は、直ちに公開されます。すべての異議申し立てに対する決定は最終的なものです。

アカウントが取り消しを受け取ると、その取り消しの原因となった違反に異議を申し立てることができなくなり、アカウント制限期間中はコミュニティへの参加が禁止されます。

アカウント履歴

アカウントの状態に関連するイベントのタイムラインは、[アカウントの状態ページ](#) から入手できます。ここでは、すべての違反、警告、アカウント制限、およびすべての異議申し立てと結果の決定を確認できます。

[ガイドライン](#)

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（※1）最近の月間アクティブユーザー数（※2）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Q 2 - 4 参照。

※2 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。

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Q 2 - 4 参照。

Q 2 - 3 Q 2 - 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Q 2 - 4 参照。

【構成員限り】

【構成員限り】

Q 2 - 4 Q 2 - 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

3 偽・誤情報の流通・拡散への対応方針について

Bing では品質、安全性、ユーザー価値に関連する特定の状況において、特定の結果の削除を決定し、ユーザーに対する警告または啓発を実施し、検索結果をカスタマイズするオプションを提供することがあります。

スパム

Bing インデックスでキャプチャされた特定のページは、ユーザーにとってほとんど、あるいはまったく価値がないことが判明したり、より関連性の高い情報を提供するページへの関連性を歪曲するために、検索と広告システムの動作を意図的に操作するという特性を持っていたりする場合があります。こうしたページには、広告だけのもの、または検索の対象に関連するコンテンツが存在しないか見せかけのコンテンツしかないような、ほとんどが広告だけで構成されている他の Web サイトへのリンクが含まれている場合があります。ユーザーの検索エクスペリエンスを向上させ、より関連性の高いコンテンツを提供するために、Bing ではこのような検索結果を削除したり、検索結果内のより有用で関連性の高いページを優先するよう、Bing のアルゴリズムを調整したりする場合があります。

機密性の高い個人情報（同意を得ていない私的な画像の配布を含む）

一般公開されている Web ページには、特定可能な個人の同意なしに投稿された機密性の高い個人情報、またはセキュリティまたはプライバシーのリスクが生じる状況で投稿された機密性の高い個人情報が意図的に、または不注意により含まれる場合があります。たとえば、個人情報、電話番号、識別番号などが不注意によって投稿されてしまう場合や、不正アクセスやハッキングに使用するために、メール パスワード、ログイン資格情報、クレジットカード番号などのデータが、意図的かつ悪意を持って投稿される場合があります。Bing ではそうした検

索結果を検証してから削除します。

別の例として、他人のアダルト画像または露骨な性的表現を含む画像が、それが実際のものであるか AI や他の画像編集ツールを使って作成された“ディープフェイク”であるかにかかわらず、同意なしにオンラインで共有される場合があります。これは一般的に、同意を得ていない性的な画像、または「リベンジ ポルノ」と呼ばれます。マイクロソフトはこれを重大なプライバシーの侵害行為と見なしています。被害者が自身の画像とプライバシーを取り戻せるようにするため、Bing ではそのような写真や動画へのリンクを検索結果から削除することがあります。被害者がディープフェイクなどのオンラインで共有された無許可の写真や動画を報告するには、Bing の問題報告 Web ページのフォームに記入する必要があります。

このフォームにも記載されているように、Bing の検索結果からリンクが削除されても、情報は元の Web 上に残ったままであることに留意することが重要です。Web サイトの所有者は、公開する情報のプライバシーの懸念に最も対処しやすい立場にあります。Web からコンテンツを削除するには、そのコンテンツが掲載されている Web サイトの所有者に被害者から連絡する必要があります。

成人向けコンテンツ

Bing ではセーフサーチ機能が用意されており、ほとんどのユーザーは、検索結果に適用する成人向けコンテンツのフィルター処理の種類を設定できます。Bing では、意図的に要求された場合以外は不快感を与えたり害を与えたりする可能性のあるコンテンツを提供しないよう努めており、ユーザーが自身の検索エクスペリエンスをカスタマイズできるセーフサーチ設定を提供しています。ほとんどの国または地域では、セーフサーチは中程度に初期設定されています。これにより、視覚的に露骨な検索結果は制限されますが、露骨なテキストは制限されません。ユーザーはいつでもセーフサーチの設定を変更できます。

国や地域によって、成人向けコンテンツの表示（または成人コンテンツにアクセスする検索結果）に関する現地の慣習、宗教的・文化的な規範、または現地の法律が異なる場合があります。これは、一部の国における Bing のセーフサーチの既定の設定に影響を与える可能性があります。

違法医薬品

Bing ではユーザーに信頼できる検索結果を提供する取り組みを推進すべく、オンラインでの違法な医薬品の販売によってユーザーが被害を受ける可能性を最小限に抑えるための措置を講じています。米国の場合、米国食品医薬品局（FDA）のインターネット医薬品販売者に対する警告書サイトに記載されているオンライン医薬品販売者は Bing でブロックされています。米国医薬品業界の連邦取締機関である FDA は、違法行為および危険行為に関与したオンライン医薬品販売者のリストを一般公開しています。また、FDA は BeSafeRx: Know Your Online Pharmacy にも情報を掲載しています。

Q 3 - 1 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。

（1）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

（2）上記（1）の類型ごとに想定される典型事例

Bing: 前述の Bing 概要の説明中の「 Bing での検索結果の提供方法」をご参照ください。Bing はコンテンツをホストしていないため、Bing のポリシーでは、WWW 上の情報の優先順位を上げたり下げたりする方法に重点を置いています。

Start: マイクロソフトは誤解を招く虚偽のコンテンツ、有害なコンテンツ、その他の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、論争を引き起こすことを主な目的としたコンテンツを禁止しています。たとえば、以下のようなコンテンツです。根拠のない主張、詐欺まがいの無料サービスまたは価格請求、扇情的なテキストまたは画像、販売促進されている製品/サービスに関連しないコンテンツ、不当表示、サードパーティ製品およびサービスの許可されていないプロモーション、情報影響操作、外国からの干渉、公益を害する可能性のある不正な、または誤解を招くコンテンツ、ならびにその他の同様の行動（偽情報）。

（3）上記（1）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

該当せず

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。

マイクロソフトの基本的な考え方と価値観を示した本文書冒頭の概要のセクションをご参照ください。

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、（1）当該見直し・レビューの頻度・タイミング、及び（2）レビューの実施主体を御教示下さい。

該当せず

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

マイクロソフトは表現の自由および情報にアクセスする権利を尊重しています。その一方で、Bing ではマイクロソフトの責任ある AI、プライバシー、デジタル セーフティ、情報の完全性、およびその他の重要な問題に関する原則に従い、安全な検索エクスペリエンスをユーザーに提供するための、コンテンツ フィルタリング、運用監視、悪用検知をはじめとする安全対策を開発しました。

マイクロソフトが個人、企業、政府からコンテンツの削除依頼を受けた場合、品質、安全性、ユーザーの要求、関連法令、公共政策に関する懸念が存在する場合に限り、Bing は検索結果を削除し、特定のリスクについてユーザーに知らせ、コンテンツを修正するための選択肢をユーザーに提供することがあります。Bing では、ユーザーによる関連情報へのアクセスが制限されることのないよう、限られた状況または条件においてのみ検索結果を削除できるようになっています。

Bing の向上した機能や新機能をユーザーに公開するまでに、マイクロソフトは多大な労力を投じています。マイクロソフトの研究者は多くの場合、新機能を試してフィードバックを提供してくれるユーザーを募っています。それらのユーザーに満足してもらえた場合、少数の実際のユーザーに機能をリリースし、機能がどのように利用されるかを確認します。問題がないことが確認されたらすべてのユーザーに対してリリースしますが、これで終わり

はありません。リリースされたすべての機能の使用状況とフィードバックを定期的に確認して、問題がないか、改善の余地はないかを評価しています。

Q 4 - 1 Q 3 - 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、Q 3 - 1（3）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。

※1 例えば、書面、電子メール、対象サービス上での表示など。

※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、Q 4 - 3の苦情・不服申立て受付窓口に関する情報など。

該当せず

Q 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。

(1) 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。

(2) 日本語による苦情・不服申立てを受け付けているか否か

(3) 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。

該当せず

Q 4 - 3 Q 4 - 2で窓口を設置しているか否かににかかわらず、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。

(1) 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

(2) 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）

(3) 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）

(4) 上記（3）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）

(5) 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）

(6) 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

- (7) 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）
- (8) 特定の発信者（投稿者）（※）からの日本語による苦情・不服申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
 ※ 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

該当せず

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。

該当せず

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

Q5-1 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。

- (1) 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。
- (2) 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
- (3) 日本語による通報を受け付けているか否か
- (4) 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。

該当せず

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（1）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。
- (1) 日本語通報対応が可能な人員として常用している人数
- (2) 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）
- (3) 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）
- (4) 上記（3）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）
- (5) 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）
- (6) 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

- (7) 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）
- (8) 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

該当せず

Q 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（Q 3 - 1（1）参照）ごとに異なる場合は当該類型別に御教示下さい。

- (1) 検知・対応に向けた調査を実施する頻度・ケース
- (2) 検知・対応にA I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）
- (3) 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）
- (4) 上記（3）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）
- (5) 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

該当せず

Q 5 - 4 Q 5 - 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。

該当せず

6 偽・誤情報の流通・拡散への対応状況について

Q 6 - 1 2022年中及び2023年中それぞれに関する（※1）次の20点）として把握している数字があれば、御教示下さい（※2）。

※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（Q 3 - 1（1）参照）、モデレーション等の方法別（同（3）参照）の数字も併せて御教示下さい。また、可能であれば、（9）～（12）については通報主体別（Q 5 - 1（2）参照）、（18）・（19）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。

- (1) 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
 - ※ Q 4 - 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。
- (2) 上記（1）のうち、A Iその他の機械的手段のみによって検知・対応した件数
- (3) 上記（1）のうち、A I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A Iその他の機械的手段を利用した場合を含

む) 件数

- (4) 上記(2)(3)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(3)において人間の手によって補正された場合を含む)件数
- (5) 上記(1)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
- (6) 上記(1)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
- (7) 上記(1)のうち、上記(5)(6)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
- (8) 上記(5)～(7)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
- (9) 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。
- (10) 上記(9)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
- (11) 上記(9)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
- (12) 上記(9)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
- (13) 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。
- (14) 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
- (15) 上記(13)(14)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
- (16) 上記(13)(14)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
- (17) 上記(13)(14)のうち、Q4-3(7)の再審査要求があった件数
- (18) 上記(13)(14)のうち、Q4-3(8)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
- (19) 上記(18)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
- (20) 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

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。

該当せず

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

Q7-1 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（3）参照）件数を御教示下さい。

Q7-2 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。

Q7-3 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。

Q7-4 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。

Q7-5 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。

Q7-6 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。

該当せず

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

Q8-1 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。

該当せず

Q8-2 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。

該当せず

Q8-3 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。

該当せず

Q 8 - 4 Q 8 - 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の可否・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。

該当せず

Q 8 - 5 Q 8 - 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。

※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。

※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。

※3 例えば、開示目的、開示に至った経緯など。

該当せず

Q 8 - 5 Q 8 - 4 で既に開示しているか否かにかかわらず、Q 8 - 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。

該当せず

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 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。

※ 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。

該当せず

Q 9 - 2 Q 9 - 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。

※ 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。

該当せず

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。

[Copilot in Bing: Our approach to Responsible AI – Microsoft Support](#)

マイクロソフトは 2023 年 2 月、AI で強化された Web 検索エクスペリエンスである新しい Bing をリリースし

ました。新しい Bing は、Web の検索結果を要約し、チャット エクスペリエンスを提供することで、ユーザーを支援します。ユーザーは詩やジョーク、物語といったクリエイティブなコンテンツを生成したり、Bing Image Creator で画像を生成したりできます。AI で強化された新しい Bing は、マイクロソフトと OpenAI 社のさまざまな先進テクノロジーを活用しています。これらテクノロジーには、最先端の大規模言語モデル (LLM) である GPT-4、自然言語記述からデジタル画像を生成するディープ ラーニング モデルの DALL-E など(共に OpenAI 社の技術)が含まれます。マイクロソフトは、両モデルが一般公開される前の数か月間で、この最先端の AI テクノロジーと新しい Bing の Web 検索機能を連携させるための一連のカスタマイズされた機能とテクノロジーを開発しました。2023 年 11 月、マイクロソフトは新しい Bing の名称を Copilot in Bing に変更しました。

マイクロソフトは、責任ある AI へのコミットメントに真摯に取り組んでいます。Copilot in Bing のエクスペリエンスは、マイクロソフトの AI の原則とマイクロソフトの責任ある AI の基準に従って開発されました。開発には、マイクロソフトの責任ある AI オフィス、エンジニアリング チーム、Microsoft Research、Aether をはじめとする全社の責任ある AI のエキスパートたちが協力しました。マイクロソフトの責任ある AI の詳細については、こちらをご参照ください。

Copilot in Bing では、情報を求めているユーザーに対応する際、Web 検索結果に根拠を置いています (質問 1 の回答をご参照ください)。つまり、ユーザーのクエリやプロンプトに対して応答を提供するときは、Web 上の上位のコンテンツを中心とします。そして、ユーザーが詳細を知ることができるように Web サイトへのリンクも提供します。Bing は、関連性、品質および信頼性、鮮度といった特徴に強い重み付けをして Web 検索コンテンツをランク付けしています。Copilot in Bing からの応答は、出力文がクエリまたはプロンプトによる Web 検索結果や、Bing のファクトチェック済み情報のナレッジ ベース、あるいはチャットの最近の会話履歴 (チャット エクスペリエンスの場合) といった入力ソースに含まれる情報で裏付けられていることにより、根拠のある応答と見なされます。根拠のない応答とは、出力文がそのような入力ソースに根拠付けられていない応答のことです。

Q10-2 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。

マイクロソフトは 悪意ある人物が Bing の生成 AI エクスペリエンスを使用して虚偽の情報を作り出すリスクを最小限に抑えるための保護策を講じています。こうした取り組みとしては、分類 (Classifiers) やメタプロンプトの使用、来歴ツール、報告機能の強化、堅牢な運用とインシデント対応などがあります。また、Bing の生成 AI 機能は、不正な情報や虚偽の情報、または誤解を招く情報を作成または共有するためにこれらのサービスを使用することを禁止しています。

生成 AI モデルが生み出すコンテンツの品質が向上しているため、AI が生成したコンテンツの出所に関する透明性を高める必要性が高まっています。現在、Azure OpenAI サービスが提供するすべての AI 生成画像には、改ざんを防止するためにコンテンツの出所と来歴を開示するコンテンツ クレデンシャルが含まれています。コンテンツ クレデンシャルは、Joint Development Foundation のプロジェクトである Coalition for Content Provenance and Authenticity (C2PA) のオープン技術仕様に基づいています。

ディープフェイクに対する、決め手となるような解決策は存在しません。最も有効なアプローチは、従来の徹底的なセキュリティ計画を実行することです。具体的には、ディープフェイクの脅威の大部分を軽減することが期待される4つの軽減策があります。

- **生成 AI コンテンツのウォーターマーク** - マイクロソフトが Bing Image Creator で行っているように、生成 AI 画像の作成者を明らかにし、生成 AI により作成されたコンテンツであることを明示することが極めて重要です。
- **来歴** - C2PA 基準またはマイクロソフトの新しい“サービスとしてのコンテンツ クレデンシャル”製品を通じて提供される来歴により、個人または組織は、どのコンテンツが自らの出所から提供されているのかを主張できます。
- **検知** - 画像や動画が AI によって作成または操作されているかどうかを判断できることには大きな価値があります。どちらのテクノロジーも進化し続けるため、検知機能を最終的なソリューションと見なすべきではありませんが、多くの状況において有益なツールです。
- **リテラシー** - 消費者に対する啓発は、AI によるメディアの操作に惑わされないようにするために不可欠なことです。人々がコンテンツ クレデンシャルおよびメディアの明確なラベル付けを要求し、これらに対応していないメディアに疑問を呈することが、より強靱な、そうした策略に負けない社会の実現につながります。

Q10-3 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。

該当せず

Q10-4 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。

該当せず

1.1 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。

マイクロソフトは日本ファクトチェックセンターやセイファーインターネット協会等と、双方の取り組みや問題意識の共有等を目的とした意見交換を行っています。たとえば、日本での持続可能なファクトチェックエコシステムの構築方、テクノロジー企業の貢献方法、ファクトチェック機関との連携方法等について話し合っています。

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。

日本における持続可能なファクトチェック・エコシステムの構築をどのように支援していくかについて、ファクトチェック機関などのステークホルダーと継続して協議していきたいと考えていますが、マルチステークホルダーのアプローチと取り組みが必要だと考えています。

Q 1 1 - 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。

特記事項なし

1 2 マスメディア（新聞・放送）との連携状況について

Q 1 2 - 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。

特記事項なし

Q 1 2 - 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。

特記事項なし

1 3 利用者の ICTリテラシー向上に向けた取組について

Q 1 3 - 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。

マイクロソフトは安心ネットづくり促進協議会（JISPA）の会員であり、安心協の会合に定期的に参加しています。また、マイクロソフトは、毎年開催されている JISPA 主催の「高校生 ICT Conference」のスポンサーも務めており、ワークショップにマイクロソフトからスピーカーを派遣し、ディスカッションに参加するなどの支援を行っています。

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）とも連携し、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しています。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しました。作成されたコンテンツは、経済産業省のSTEM Libraryに掲載されています。

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しています。

Q13-2 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。

特記事項なし

Q13-3 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。

マイクロソフトは、ウェブサイトではリテラシー学習教材を提供しています。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しています。「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できます。

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しています。

- [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行います。
- [The Investigators](#) は教育版マイン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しています。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です。

そのほか、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」などを提供しています。

また、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しています。

14 研究開発の推進に向けた取組について

Q14-1 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。

2023年8月、マイクロソフトと東京大学は、日本の社会と未来に貢献する学生の育成を目的として、グリーントランスフォーメーション、ダイバーシティ & インクルージョン、人工知能の各分野の研究推進に向けて連携するための基本合意書を締結しました。

Q14-2 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。

特記事項なし

Q14-3 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。

特記事項なし

Q14-4 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。

特記事項なし

Q14-5 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お考えをお聞かせ下さい。

特記事項なし

Q14-6 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。

特記事項なし

15 サイバーセキュリティ関係機関等との連携について

Q15-1 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。

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めています。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。

Q15-2 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。

特記事項なし

Q15-3 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。

特記事項なし

16 行政機関や地方公共団体等との連携について

Q16-1 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 (GIFST) 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。

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めています。以下は、こうした協力の例です。

- 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しています。

Q 1 6 - 2 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。

特記事項なし

1 7 国際機関等との連携について

Q 1 7 - 1 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。

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めています。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からです。たとえば、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してきました。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、支援しています。

Q 1 7 - 2 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。

特記事項なし

Q 1 7 - 3 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。

特記事項なし

18 その他のステークホルダーとの連携状況について

Q18-1 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。

※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、イベントの共催・後援など。

特記事項なし

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。

特記事項なし

Q18-3 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。

特記事項なし

Q18-4 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。

特記事項なし

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】（3）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。

特記事項なし

20 その他

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。

特記事項なし

パート2: LinkedIn関係

1 前提：対象サービスの範囲について

【想定対象サービス】

SNS	
検索	Bing 検索
動画共有	

ニュースポータル・ キュレーション	Microsoft Start
ブログ・掲示板	

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。

概要

LinkedIn は実名制のプラットフォームです。つまり、LinkedIn ではユーザーに実名を使うよう求めています。パロディ アカウントや匿名アカウント、ペットの猫や犬のアカウントは許可しておらず、実在する人物でなければなりません。第2に、LinkedIn のプラットフォームはビジネスに特化しています。実際、LinkedIn は世界最大のプロフェッショナル ネットワークであり、200 を超える国で 10 億を超えるユーザーが利用しています。LinkedIn のビジョンは、世界で働くすべての人のために、経済的なチャンスを作り出すことです。

LinkedIn の基盤は、メンバー プロフィールとして知られているものです。メンバー プロフィールは、履歴書のようにユーザーの職務経歴を伝えるデジタル ポートフォリオです。プラットフォーム上でのメンバーの活動やメンバーが共有しているコンテンツはこのプロフィールに関連付けられ、現在および将来の雇用主、同僚、潜在的なビジネス パートナー、人材紹介会社などが参照できます。また、プラットフォーム上でのメンバーの活動は、その人の実際の職業に直接関連しているため、多くのメンバーは自身の活動を関心のある専門分野に限定しています。

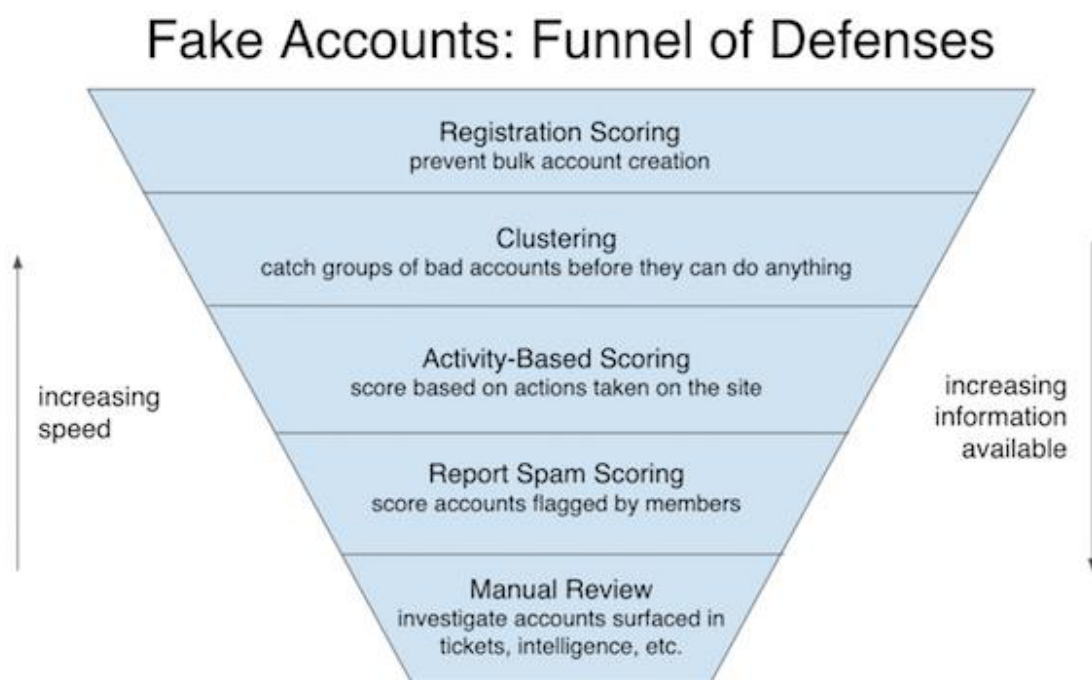
同じ理由から、メンバーは表示されるコンテンツがビジネスにふさわしい性質のものであることを期待しています。LinkedIn は、職業倫理に反するコンテンツをメンバーが簡単に報告できる方法を複数用意しています。

メンバーが LinkedIn にアクセスする目的は、仕事を見つけたり、スキルを学んだり、人脈を作ったりすることです。それには、メンバーがサイトに安心し、自信を持って、主体的に参加できる必要があることを私たちは認識しています。そのため LinkedIn では、安全な環境を構築することを基本的な優先事項としており、これは私たちのビジネスにとって絶対不可欠なものです。

偽アカウント

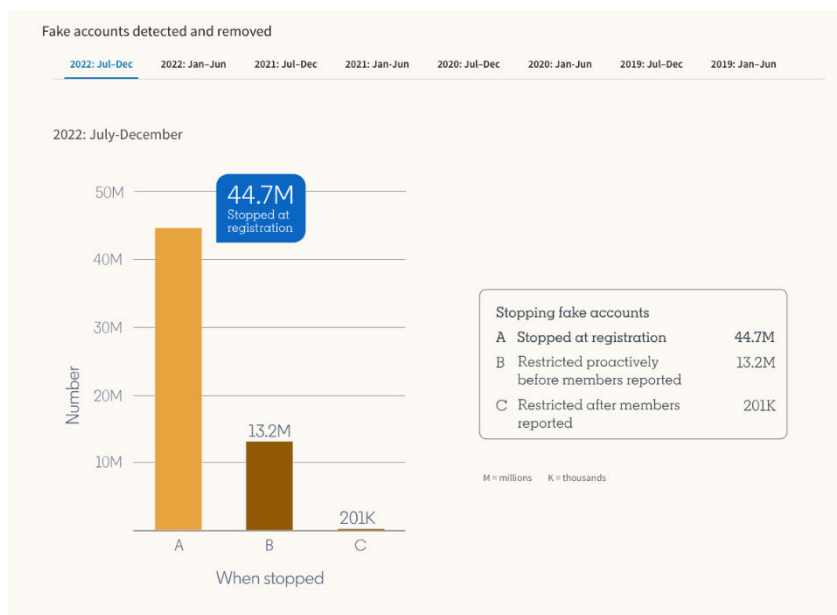
インターネットやソーシャル メディアにはさまざまな脅威が存在することを、私たちは理解しています。私たちにとって唯一最大の脅威ベクトルは、偽アカウントです。偽アカウントを使用したオンライン上での行動（誤情報や偽情報の投稿など）から、それが実際にどういった職業の人物によるものを推察することは不可能だからです。そのため、LinkedIn では偽アカウント対策に積極的に多額の投資をしています。人工知能などのテクノロジーと専門家チームが協力して、偽アカウントが使用される前に阻止して削除するよう努めています。こうした取り組みの結果、昨年 LinkedIn では全世界で 8,000 万件超の偽アカウントをブロックしました。

偽アカウントが大量に作成されるのを防ぐため、LinkedIn ではデータ サイエンティスト、ソフトウェア エンジニア、機械学習エンジニア、調査員で構成される専任の Trust チームを設置しており、プラットフォーム上の不正行為を常に分析し、その対策に使用する技術の向上を図っています。Trust チームが使用する分析の枠組みは、ファネルで表すことができます。



第 1 に、大量の偽アカウントをあらかじめ登録時にブロックしてしまえるよう、機械学習モデルを使って、似ていると思われる、あるいは同じように振る舞う一連のアカウントを検出します。第 2 に、共通の属性に基づいてグループ化した、アカウントのクラスターを作成します。その後、統計的に異常な分布を示すアカウントのクラスターを見つけ、教師あり機械学習モデルでクラスターをスコアリングしてから、そのクラスター内のアカウントにアクションを実行します。第 3 に、アクティビティ ベースのモデルを使用し、不正アカウントに典型的な特定の種類の悪質な行動や異常な行動を探します。第 4 に、メンバーから提供される貴重な情報、すなわち報告される新たに検出されたアカウントを使用して、モデルのさらなるスコアリングとレビューを実行できます。

LinkedIn の透明性レポートにあるとおり、自動化された防御機能により、毎年数百万件の偽アカウントが検出されています。



LinkedIn 上のコンテンツが偽物ではない信頼できるものであることを保証するために、テクノロジーと専門家チームを活用し、詐欺や偽アカウントを事前に見つけて削除しています。2023 年には、検出されたスパムや詐欺の 99.6% を自動防御機能で削除し、検出された偽アカウントの 99.7% はメンバーから報告を受ける前にブロックしました。また、人工知能などのテクノロジーと専門家チームを活用して詐欺行為を阻止しています。さらに、ヘルプセンターでは、メンバーが不審なメッセージを見分けるのに役立つ情報も提供しています。

コンテンツ モデレーション

LinkedIn のミッションは、プロフェッショナルに特化し、その生産性を高めて成功に導くことです。そのため、実名制のネットワークであるだけでなく、メンバーが投稿するコンテンツについても、ビジネスに関連するものに限定しており、そうしたコンテンツを管理するための基準を策定しています。

これらの基準を、将来を見据えたプロフェッショナル コミュニティ ポリシーで具体的に定めています。コミュニティポリシーでは、LinkedIn での職業倫理に反する行為は一切許されないと明言されています。たとえば、ハラスメ

ント、誤情報、ヘイトスピーチ、またはいじめを決して許さないという方針が明確に定められています。さらに、LinkedIn では、いかなる有料の政治広告も許可していません。また、メンバーによるコンテンツの収益化や広告の掲載は許可しておらず、広告収益共有プログラムも提供していません。

この取り組みをサポートするために、政策研究者、ライター、実装の専門家によるグローバル チームを結成し、プラットフォーム全体のコンテンツ基準を定めて実施しています。同業他社、政策立案者、法執行機関、政府機関とも連携しています。また、6 か月ごとに独自の透明性レポートを公開し、プロフェッショナル コミュニティ ポリシーに違反したコンテンツに対して講じた措置について説明するとともに、偽アカウントやコンテンツ違反についても取り上げています。

LinkedIn のプロフェッショナル コミュニティ ポリシーは、最良のプロフェッショナル ライフ、すなわち互いを尊重し合い、互いの成功を支援するコミュニティを LinkedIn で体験できるよう意図されています。これらのポリシーでは、LinkedIn で誤情報や偽情報が許されないことを明言しています。これらのポリシーに繰り返し違反すると、アカウントが永久に削除される可能性があります。プロフェッショナル コミュニティ ポリシーに違反していると思われるコンテンツがあれば、メンバーは誰でも LinkedIn に報告することができます。その裏付けとして、Transparency Center (透明性センター) を通じ、メンバーを保護するために講じている措置、メンバーのデータに関する質問やコンテンツの削除要求への対応など、透明性に関する最新情報をメンバーに定期的に提供しています。

LinkedIn のチームは、自動化された複数の手法を用い、人間によるレビューやメンバーからの報告に基づいて、利用規約やプロフェッショナル コミュニティ ポリシーに違反しているコンテンツを特定します。また、世界の報道機関と提携してファクトチェックを支援しています。メンバーは、LinkedIn によるコンテンツ削除の決定に対して異議申し立てを行う場合、LinkedIn の決定が間違っていると思う理由を提示して再審査を求めることができます。

LinkedIn はまた、ニュース編集者からなるグローバル チームを通じて、メンバーに民間で行われている議論、選挙のプロセス、公共の安全に関する情報を提供することを目指しています。これらの編集者は、信頼できる情報源から入手したタイムリーな関連情報を各メンバーに提供し、世界の選挙などの実情やニュースを LinkedIn のメンバーに伝えます。また、コンテンツ モデレーション チームは、それに関連してさまざまな言語で交わされているやり取りを注意深く監視しています。

結論

LinkedIn は、実名制、ユーザー プライバシー、データ セキュリティ、責任あるコンテンツ モデレーションへのコミットメントを果たすことで、メンバーのロイヤルティとエンゲージメントを高めるよう取り組んでいます。誤情報や偽情報などの不適切なコンテンツに対してプラットフォームを安全に保つことは、お客様やメンバーのためになるだけでなく、私たちのビジネスにとっても不可欠なことです。

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 (※1) 最近の月間アクティブユーザー数 (※2) 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Q 2 - 4 参照。

※2 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。

【構成員限り】

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Q 2 - 4 参照。

該当なし

Q 2 - 3 Q 2 - 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Q 2 - 4 参照。

該当なし

Q 2 - 4 Q 2 - 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

ユーザーの LinkedIn プロフィールに記載されている“場所”です。

3 偽・誤情報の流通・拡散への対応方針について

Q 3 - 1 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。

(1) 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

LinkedIn の虚偽または誤解を招くコンテンツ ポリシー

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

虚偽または誤解を招く内容を投稿することは、[LinkedIn のプロフェッショナル コミュニティ ポリシー](#)に違反する行為です。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、これを削除するものとします。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ません。

LinkedIn の[広告掲載ポリシー](#)には上記の規定が盛り込まれており、同様に誤情報や偽情報を禁止しています。さらに、LinkedIn の[広告掲載ポリシー](#)では詐欺広告および虚偽広告も禁止しており、広告の宣伝文句には事実の裏付けがなければなりません。

- 詐欺広告および虚偽広告: 詐欺広告や虚偽広告は許容されません。製品またはサービスは、広

告の内容と正確に一致している必要があります。広告の宣伝文句はすべて事実に基づいている必要があります。競合する製品またはサービスに関する虚偽の、または不正確な記載があってはなりません。貴社が第三者と連携しているか、貴社の製品を第三者が推奨していることを、当事者の許可なくほのめかすことはできません。不正確な価格や特典を宣伝することはできません。割引、特典、価格は、広告内のリンクから簡単に確認できる必要があります。

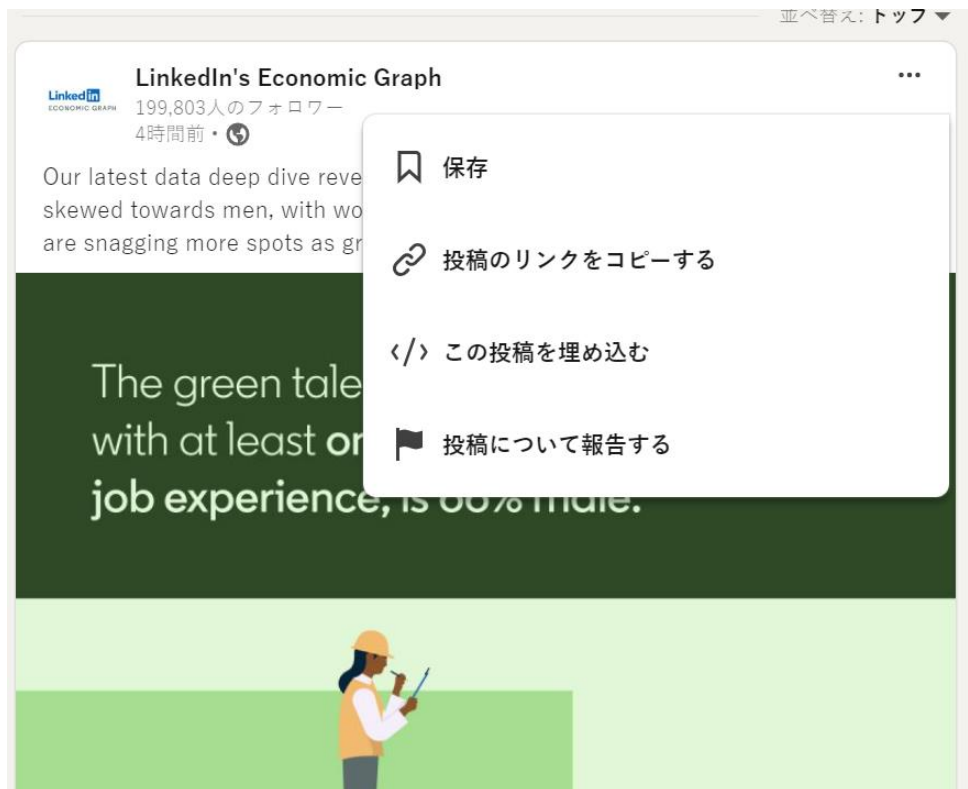
LinkedIn は、LinkedIn 上のコンテンツに対してグローバルなアプローチを採用しています。LinkedIn の安全性、信頼性、専門性を維持するために、LinkedIn で許可されない、人を不快にさせる有害なコンテンツの範囲を プロフェッショナル コミュニティ ポリシー で明確に規定しています。偽アカウントや誤情報、不正なコンテンツは許可されておらず、LinkedIn はプラットフォームからそれらを削除するために積極的な措置を講じています。

LinkedIn は、スパム、フィッシングや詐欺、重複アカウント、偽アカウント、誤情報などの不正行為を含む悪用を特定して防止するための自動防御機能を備えています。LinkedIn の Trust & Safety チームは、不正行為を特定して制限するために日々取り組んでいます。機械学習モデルなどのスケーラブルな テクノロジー を頻繁に導入し、プラットフォームの安全性を維持しています。

LinkedIn のメンバーは、以下で説明するプロセスを使用して、誤情報、不正なコンテンツ、偽アカウントなどの、プロフェッショナル コミュニティ ポリシー に違反していると思われるコンテンツを報告することもできます。報告されたコンテンツやフラグが付けられたコンテンツが プロフェッショナル コミュニティ ポリシー に違反している場合、そのコンテンツはプラットフォームから削除されます。重大な違反の場合や、過去に不正利用を行ったことがある場合は、違反したメンバーの LinkedIn アカウントを削除する場合があります。

違反の報告

LinkedIn のメンバーには、LinkedIn の プロフェッショナル コミュニティ ポリシー に違反していると思われるコンテンツを発見した場合に、LinkedIn で投稿の右上隅の 3 つの点（下図）をクリックすると表示される製品内報告メカニズムを使用して報告することを推奨しています。



誤情報は報告の選択肢の 1 つとして特に強調されています。

この投稿を報告

該当する理由を選択

ハラスメント + 詐欺 + スпам + 誤情報 + ヘイトスピーチ +

脅迫または暴力 + 自傷 + 生々しいコンテンツ +

危険な組織または過激派組織 + 性的コンテンツ + フェイクアカウント +

児童虐待 + 違法な商品やサービス + 侵害 +

次へ

この投稿を報告
×

該当する理由を選択

ハラスメント +

詐欺 +

スパム +

誤情報 ✓

ヘイトスピーチ +

脅迫または暴力 +

自傷 +

生々しいコンテンツ +

危険な組織または過激派組織 +

性的コンテンツ +

フェイクアカウント +

児童虐待 +

違法な商品やサービス +

侵害 +

次へ

この投稿を報告
×

次の理由を選択しました

誤情報

虚偽の事実や出来事を、真実である、または真実であり得るかのように示す虚偽のコンテンツや情報 (ニュースストーリーを含む)

戻る

報告を送信

通常、報告されたコンテンツは訓練を受けたコンテンツ レビューアーが審査しています。さらに、LinkedIn では、個人情報の開示、スパムや悪意のあるページ、違法な素材のコンテンツなどの潜在的な違反を、自動的に社内のコンテンツ モデレーション チームに通報しています。報告または通報されたコンテンツがプロフェッショナル コミュニティ ポリシーに違反していると判明した場合、そのコンテンツはプラットフォームから削除されます。

(2) 上記 (1) の類型ごとに想定される典型事例

LinkedIn では、LinkedIn のポリシーに違反している虚偽の、または誤解を招くコンテンツの具体例も、ヘルプセンターの記事「[虚偽または誤解を招くコンテンツ](https://www.linkedin.com/help/linkedin/answer/a1340752/)」(https://www.linkedin.com/help/linkedin/answer/a1340752/)で示しています。

このポリシーに従い削除される虚偽の、または誤解を招くコンテンツの例

- 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
- 次のような、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
- 次のような、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

（3）上記（1）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

虚偽または誤解を招く情報は削除されます。

LinkedIn のプロフェッショナル コミュニティ ポリシーへの準拠について
[\(https://www.linkedin.com/help/linkedin/answer/a1342754/\)](https://www.linkedin.com/help/linkedin/answer/a1342754/)

LinkedIn の 利用規約 および プロフェッショナル コミュニティ ポリシー に違反すると、アカウントまたはコンテンツに対するしかるべき措置が取られる可能性があります。違反の重大度に応じて、特定のコンテンツが表示制限、ラベル付け、または完全に削除される場合があります。その際、通常は、コンテンツが LinkedIn のポリシーに違

反していること、違反の内容、および LinkedIn が取る措置について通知します。コンテンツが誤って削除されたと思われる場合は、異議申し立てをすることができます。

違反を繰り返すと、アカウントが削除される場合があります。LinkedIn はアカウントの削除について異議を申し立てる機会を提供しており、メンバーがプロフェッショナル コミュニティ ポリシーを遵守することに同意する場合は、削除されたアカウントを復活させることができます。違反が続くと、LinkedIn プラットフォームからアカウントが永久に削除されます。

LinkedIn のプロフェッショナル コミュニティ ポリシーに対する特定の重大な違反（例：児童の性的虐待コンテンツ、テロ行為、極めて暴力的なコンテンツ、悪質なセクハラ）については、1 回の違反でもアカウントが永久に削除される場合があります。

通常はポリシーの文言に違反するようなコンテンツでも、意識の向上や批判を目的としてシェアさせる場合には、許容されることがあります。そのような場合、LinkedIn は当該コンテンツがセンシティブまたは不快だと感じる可能性のあるメンバー、またはその他の理由で閲覧を望まないメンバーのために、当該コンテンツにラベル付けして、表示されないようにすることがあります。ただし、コンテンツを削除したり、投稿者にペナルティを科したりすることはありません。

また、LinkedIn ブログでも、[「New LinkedIn profile features help verify identity, detect and remove fake accounts, boost authenticity」](#)(新たな LinkedIn プロフィール機能による本人確認、偽アカウントの検出と削除、真正性の向上)、[「How We're Protecting Members From Fake Profiles」](#)(偽プロフィールからメンバーを保護する方法)、[「Automated Fake Account Detection」](#)(偽アカウントの自動検出)、[「An Update on How We Keep Members Safe」](#)(メンバーの安全を守る方法に関する最新情報) など、虚偽情報に対処するための取り組みを説明しています。

Q 3 - 2 Q 3 - 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。

上記の質問 1 に対する回答をご参照ください。

Q 3 - 3 Q 3 - 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、(1) 当該見直し・レビューの頻度・タイミング、及び (2) レビューの実施主体を御教示下さい。

該当なし

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








Q 4 - 1 Q 3 - 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、Q 3 - 1 (3) 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。


※1 例えば、書面、電子メール、対象サービス上での表示など。


※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。

LinkedInの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（報告されたものか自動で検出されたものかを問わず）LinkedIn から削除しています。メンバーは、LinkedIn でそのコンテンツに対する措置が講じられた理由について報告を受け、その決定に異議を申し立てることができます。LinkedIn では、コンテンツ審査のプロセスと決定事項の質を保つための選任の品質保証チームを擁しています。たとえば、品質保証チームは、過去に行われたコンテンツ モデレーションに関する決定事項の質を定期的に確認します。それにより、プロセスを改善し、信頼できる情報源としての LinkedIn のプラットフォームをさらに強化することも可能となります。

投稿に問題があります

 LinkedIn Trust & Safety チーム <messages-noreply@>
To :
  Reply  Reply All  Forward  

 If there are problems with how this message is displayed, click here to view it in a web browser.

 Translate message to: English | Never translate from: Japanese | [Translation preferences](#)



LinkedIn

投稿に問題があります

様

投稿は、プロフェッショナルコミュニティポリシーに準拠していません。この投稿はLinkedInから削除されているため、今はあなたのみがアクセスできます。

詳細およびその他のリソースについては[通知ページ](#)をご覧ください。

[通知ページを表示](#)



弊社のポリシーの詳細については、プロフェッショナルコミュニティポリシーをご覧ください。

[詳細はこちら](#)

お客様の投稿はLinkedInのポリシーに準拠していません

状況は次のとおりです

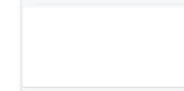
お客様の投稿は、誤情報に関する弊社の**プロフェッショナルコミュニティポリシー**に準拠していないことが判明しました。このコンテンツはLinkedInから削除されており、ご本人のみがアクセスできます。

弊社の**プロフェッショナルコミュニティポリシー**に準拠しないコンテンツを繰り返し作成すると、お客様の**LinkedInアカウントの使用が制限される**可能性があります。



お客様の投稿

🚫 誤情報のため、投稿を削除しました



可能なアクション

お客様の投稿が弊社の**プロフェッショナルコミュニティポリシー**に準拠しているとお考えの場合は、異議を申し立てて弊社に決定の見直しを要求できます。

EU諸国にお住まいの場合は、補償を求めるその他の権利について[こちら](#)をご覧ください。

異議申し立てを送信

異議申し立てのプロセス:

[← 通知ページに戻る](#)

異議申し立てのプロセス

お客様の投稿が弊社の**プロフェッショナルコミュニティポリシー**に準拠しているとお考えの場合は、次の手順に従って異議を申し立てることができます。

手順1 - LinkedInのポリシーを確認する

下のボタンをクリックし、誤情報に関する弊社のポリシーをご確認ください。

[LinkedInのポリシーを確認する](#)

手順2 - 削除された投稿を確認する

投稿を見直して、**プロフェッショナルコミュニティポリシー**に準拠していることを確認してください。

◆ 誤情報のため、投稿を削除しました

[投稿を見直す](#)

続行するには、すべてのチェックボックスをオンにしてください。

- 私は誤情報に関するLinkedInのポリシーを読みました。
- 投稿を確認しましたが、ポリシーに準拠していると信じます。
- LinkedInに、投稿の再審査を希望します。

[異議申し立てを送信](#)

手順 1 - LinkedIn のポリシーを確認する

メンバーは、決定に対する異議申し立てを行う前に、プロフェッショナル コミュニティ ポリシーを確認する必要があります。メンバーがこのリンクをクリックすると、どのポリシーに違反したのかが枠内に表示されます。

LinkedInでは身元を偽らず、自分や所属組織に関する正確な情報を提供し、真実かつ確実な情報だけを共有してください。



虚偽または誤解を招くコンテンツを共有しないでください：虚偽または誤解を招くコンテンツ、あるいは意図的に欺くためのコンテンツを共有しないでください。選挙やその他の公民活動を妨害したり、それらに不適切な影響を与えたりするコンテンツをシェアしないでください。世界の主要な保健機関や公衆衛生当局からのガイダンスと正反対の内容のコンテンツをシェアしないでください。これには、ワクチンの安全性または効果や、治療に関する虚偽の情報が含まれます。個人的利益（個人的な関係や家族関係、金銭の支払、無料の製品やサービス、またはその他の価値を含む）と引き換えにコンテンツを共有したり、人や物を推奨したりしないでください。ただし、受け取った個人的利益についてはっきりと目立つように示し、それ以外の点でLinkedInの[広告ポリシー](#)に準拠している場合は、その限りではありません。

虚偽または誤解を招くコンテンツに関する弊社のポリシーについては[こちら](#)をご覧ください。

偽のプロフィールを作成したり、自分に関する情報を偽ったりしないでください：偽のプロフィールや団体は許可されません。自分や、自分のビジネス、資格、実務経験、所属、成果について、誤解を招くような情報や虚偽の情報を投稿しないでください。[プロフィール写真](#)には、他人の画像や、自分の外見を表さない画像を使用しないでください。LinkedIn上で、実際には関連している職業上の関わりがないビジネスや組織に自分自分を関連付けしないでください。他人のLinkedInアカウントを使用したり、使用を試みたり、自分以外の人のメンバープロフィールを作成したりしないでください。また、LinkedInアカウントを他の人と共有しないでください。

詐欺や詐取を行ったり、他人を欺いたりしないでください。ロマンス詐欺、ネズミ講、その他の手段でメンバーを搾取するためにLinkedInを利用しないでください。メンバー、プラットフォーム、またはサービスをリスクにさらすような悪意のあるソフトウェアをシェアしないでください。フィッシング詐欺の試みは許容されません。

詐欺や不正のコンテンツに関する弊社のポリシーについては[こちら](#)をご覧ください。

専門性



LinkedInは、世界の人々をつなげることで個人と組織の生産性を高め、さらなる成功に結びつけることをミッションに掲げ、活動しています。仕事に関するものであれば幅広い交流を許容しますが、プロフェッショナルな表現を使用する必要があります。プロフェッショナルなプラットフォームを維持するために、メンバーは互いに敬意を払い、礼儀正しいやりとりを心がけてください。非礼な、不適切な、あるいは無礼な方法で他人とやり取りしたり、コンテンツをシェアしたりしないでください。

他人に憎悪を向けしないでください。実際の、またはそうと信じられている人種、民族、出身国、コース

手順 2 - 削除された投稿を確認する

メンバーがこのリンクをクリックして、そのポリシーを確認すると、「手順 2」に進んで、自分の投稿を確認するように促されます。



自分の投稿をクリックして確認すると、3つのチェックボックスが表示されます。異議申し立てを「送信」するには、これらのチェックボックスにチェックを入れる必要があります。

[← 通知ページに戻る](#)

異議申し立てのプロセス

お客様の投稿が弊社の**プロフェッショナルコミュニティポリシー**に準拠しているとお考えの場合は、次の手順に従って異議を申し立てることができます。

手順1 - LinkedInのポリシーを確認する

下のボタンをクリックし、誤情報に関する弊社のポリシーをご確認ください。

[LinkedInのポリシーを確認する](#)

手順2 - 削除された投稿を確認する

投稿を見直して、**プロフェッショナルコミュニティポリシー**に準拠していることを確認してください。

❗ 誤情報のため、投稿を削除しました

[投稿を見直す](#)

続行するには、すべてのチェックボックスをオンにしてください。

- 私は誤情報に関するLinkedInのポリシーを読みました。
- 投稿を確認しましたが、ポリシーに準拠していると信じます。
- LinkedInに、投稿の再審査を希望します。

[異議申し立てを送信](#)

送信すると、次の確認ページが表示されます。



[LinkedIn.comに戻る](#)

申し立てを受け付けました

今後の対応

こちらでpostを再審査いたします。弊社の**プロフェッショナルコミュニティポリシー**に準拠していることが判明した場合は、再度掲載させていただきます。

審査が完了するまでしばらくお待ちください。



異議申し立てが提出されると、メールによる通知が送信されます。

申し立てを受け付けました

LinkedIn Trust & Safety チーム <messages-noreply@>
To ○

☺ Reply ↶ Reply All → Forward 📧 ⋮

📘 If there are problems with how this message is displayed, click here to view it in a web browser.

🌐 Translate message to: English | Never translate from: Japanese | Translation preferences



Q4-2 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し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。

(1) 苦情・不服申し立ての受付方法（※）

※ 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。

コンテンツの誤情報を通報するプロセスは、Q3-1 に対する回答で示したスクリーンショットに記載されています。

通報された違反に対する異議申し立てのプロセスは、Q4-1 に対する回答で示したスクリーンショットに記載されています。

違反の報告

LinkedIn のメンバーには、LinkedIn の [プロフェッショナル コミュニティ ポリシー](#) に違反していると思われるコンテンツを発見した場合に、LinkedIn で投稿の右上隅の 3 つの点（下図）をクリックすると表示される製品内報告メカニズムを使用して報告することを推奨しています。

誤情報は報告の選択肢の 1 つとして特に強調されています。

通常、報告されたコンテンツは訓練を受けたコンテンツ レビューアーが審査しています。さらに、LinkedIn では、個人情報の開示、スパムや悪意のあるページ、違法な素材のコンテンツなどの潜在的な違反を、自動的に社内のコンテンツ モデレーション チームに通報しています。報告または通報されたコンテンツがプロフェッショナル コミュニティ ポリシーに違反していると判明した場合、そのコンテンツはプラットフォームから削除されます。

メンバーが上記の報告プロセスを使用し、最新情報を受け取ることを選択すると、報告したメンバーに対してその報告を受領したことがメールで連絡され、その後の決定事項に関する最新情報が提供されます。また、通常は、メンバーのコンテンツが LinkedIn から削除されると、そのメンバーに通知が届きます。

メンバーは、LinkedIn の決定事項に対して異議申し立てを行う場合、LinkedIn が間違った決定を下していると考え理由を提示して再審査を求めることができます。この異議申し立てのプロセスを開始するには、メンバーが自身のアカウントにログインし、画面のメッセージに従って操作するか、コンテンツが削除されたことを知らせるメッセージが送られてきたときに返信します。

審査が完了すると、現在の最新の状況（以下のいずれか）が投稿者に知らされます。

- お客様の投稿が LinkedIn のプロフェッショナル コミュニティ ポリシーに違反していない場合は、LinkedIn で公開されます。
- 投稿がプロフェッショナル コミュニティ ポリシーに違反していることがわかった場合、その投稿にアクセスできるのは投稿者のみになります。

（2）日本語による苦情・不服申し立てを受け付けているか否か

ユーザーの LinkedIn で言語が日本語に設定されている場合は、日本語での苦情・不服申し立てを受け付けています。

（3）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。

上記のプロセスは、すべての LinkedIn ユーザーが利用できます。

Q 4 - 3 Q 4 - 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し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・不服申し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。

（1）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

メンバーの問題を適時に解決するための十分なスタッフを配置しています。

- (2) 日本語苦情等対応にA I 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）

該当なし

- (3) 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）

該当なし

- (4) 上記（3）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）

該当なし

- (5) 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）

LinkedIn はメンバーに対して可能な限り迅速に対応することを心がけています。

- (6) 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

上記のスクリーンショットをご参照ください。

- (7) 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に関与させている場合はその詳細を含む）

該当なし

- (8) 特定の発信者（投稿者）(※)からの日本語による苦情・不服申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

※ 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

該当なし

Q 4 - 4 Q 4 - 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。

該当なし

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

Q5-1 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。

（1）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。

LinkedIn のメンバーは、いかなるコンテンツについても、LinkedIn による審査を要求するために通報することができます。

違反の報告

LinkedIn のメンバーには、LinkedIn のプロフェッショナル コミュニティ ポリシーに違反していると思われるコンテンツを発見した場合に、LinkedIn で投稿の右上隅の3つの点（下図）をクリックすると表示される製品内報告メカニズムを使用して報告することを推奨しています。

誤情報は報告の選択肢の1つとして特に強調されています。

通常、報告されたコンテンツは訓練を受けたコンテンツ レビューアーが審査しています。さらに、LinkedIn では、個人情報の開示、スパムや悪意のあるページ、違法な素材のコンテンツなどの潜在的な違反を、自動的に社内コンテンツ モデレーション チームに通報しています。報告または通報されたコンテンツがプロフェッショナル コミュニティ ポリシーに違反していると判明した場合、そのコンテンツはプラットフォームから削除されます。

メンバーが上記の報告プロセスを使用し、最新情報を受け取ることを選択すると、報告したメンバーに対してその報告を受領したことがメールで連絡され、その後の決定事項に関する最新情報が提供されます。また、通常は、メンバーのコンテンツが LinkedIn から削除されると、そのメンバーに通知が届きます。

メンバーは、LinkedIn の決定事項に対して異議申し立てを行う場合、LinkedIn が間違った決定を下していると考え理由を提示して再審査を求めることができます。この異議申し立てのプロセスを開始するには、メンバーが自身のアカウントにログインし、画面のメッセージに従って操作するか、コンテンツが削除されたことを知らせるメッセージが送られてきたときに返信します。

（2）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

LinkedIn ユーザー

（3）日本語による通報を受け付けているか否か

ユーザーの LinkedIn で言語が日本語に設定されている場合は、日本語での苦情・不服申し立てを受け付けています。

(4) 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。

このプロセスは、すべての LinkedIn ユーザーが利用できます。

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 (1) 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。

(1) 日本語通報対応が可能な人員として常用している人数

質問4-3の回答と同じ

(2) 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)

質問4-3の回答と同じ

(3) 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)

質問4-3の回答と同じ

(4) 上記 (3) 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)

質問4-3の回答と同じ

(5) 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)

質問4-3の回答と同じ

(6) 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

質問4-3の回答と同じ

(7) 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に関与させている場合はその詳細を含む)

質問4-3の回答と同じ

(8) 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

質問4-3の回答と同じ

Q5-3 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(1)参照)ごとに異なる場合は当該類型別に御教示下さい。

(1) 検知・対応に向けた調査を実施する頻度・ケース

該当なし

(2) 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)

該当なし

(3) 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)

該当なし

(4) 上記(3)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)

該当なし

(5) 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

該当なし

Q5-4 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。

該当なし

6 偽・誤情報の流通・拡散への対応状況について

Q6-1 2022年中及び2023年中それぞれに関する(※1)次の20点)として把握し

ている数字があれば、御教示下さい（※2）。

※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（1）参照）、モデレーション等の方法別（同（3）参照）の数字も併せて御教示下さい。また、可能であれば、（9）～（12）については通報主体別（Q5-1（2）参照）、（18）・（19）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。

（1）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

※ 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。

2024年1月:0件

（2）上記（1）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

（3）上記（1）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

（4）上記（2）（3）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（3）において人間の手によって補正された場合を含む）件数

（5）上記（1）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

（6）上記（1）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

（7）上記（1）のうち、上記（5）（6）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

（8）上記（5）～（7）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

（9）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。

報告は製品内プロセスを通じて行われます。2024年1月は、日本のメンバーに関して、対象となる偽・誤情報に関する通知は行われていません。

（10）上記（9）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

（11）上記（9）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

LinkedIn はメンバーに対して可能な限り迅速に対応するよう努めています。

（12）上記（9）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

（13）Q4-2の窓口へ対象偽・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。

（14）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

- (15) 上記(13)(14)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
- (16) 上記(13)(14)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
- (17) 上記(13)(14)のうち、Q4-3(7)の再審査要求があった件数
- (18) 上記(13)(14)のうち、Q4-3(8)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
- (19) 上記(18)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
- (20) 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

該当なし

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。

LinkedInの透明性レポートは、年に2回、次のURLで公開しています:

<https://about.linkedin.com/transparency/community-report>

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

Q7-1 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(3)参照）件数を御教示下さい。

2024年1月中、令和6年能登半島地震に関連する偽・誤情報はプラットフォーム上では見つかりませんでした。

Q7-2 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。

該当なし

Q7-3 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。

該当なし

Q7-4 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。

該当なし

Q7-5 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。

該当なし

Q7-6 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。

該当なし

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

Q8-1 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。

前述のとおり、誤情報は削除されます。

Q8-2 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。

該当なし

Q8-3 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。

前述のとおり、誤情報は削除されます。

Q8-4 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。

該当なし

Q8-5 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。

※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。

※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。

※3 例えば、開示目的、開示に至った経緯など。

該当なし

Q8-6 Q8-4で既に開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。

該当なし

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

Q9-1 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。

※ 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。

LinkedIn では、他のいくつかのプラットフォームと異なり、メンバーによるコンテンツの収益化や広告の掲載を許可していません。また、広告収益共有プログラムも提供していません。したがって、LinkedIn で誤情報を公開しているメンバーは、その誤情報を収益化することはできず、また LinkedIn を通じて広告収入を得ることもできません。

LinkedIn では広告収益共有プログラムを提供しておらず、第三者が LinkedIn に投稿するコンテンツに広告を掲載して収益化することを許可していません。

Q9-2 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。

※ 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。

該当なし

10 AI・ディープフェイク技術への対応状況について

Q10-1 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。

LinkedIn が現在ユーザーに提供しているAI機能のほとんどは、Microsoft が運用する Azure AI サービスを通じて OpenAI GPT モデルを利用しています。たとえば、LinkedIn では AI を活用したテイクアウェイとアドバイスを提供しています。これは、Azure AI サービスによる生成 AI を使用してメンバーにパーソナライズされたアドバイスと洞察を提供し、情報を入手し、スキルを伸ばし、適切な仕事をより簡単かつ効率的に獲得できるようにする一連の機能です。もう1つの例として、LinkedIn Learning サービスに AI を活用したコーチング機能があります。これは、チャットボットであり、Azure AI サービスによる生成 AI を使用して、LinkedIn Learning のコンテンツ ライブラリの中を案内し、リーダーシップやマネジメントなどの重要なスキルを身に付けることを支援するも

のです。

また、LinkedIn では InBart という GAI モデルを開発しており、LinkedIn 上のリクルーターが求職者候補への InMail メッセージを書くのに役立っています。

Q10-2 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。

LinkedIn では Azure AI サービスが展開する安全対策を導入しています。また、LinkedIn のすべての AI 製品およびサービスについて「責任ある AI の原則」に取り組んでいます。詳細については <https://www.linkedin.com/blog/member/trust-and-safety/responsible-ai-principles> をご確認ください。違法なコンテンツや有害なコンテンツへの対応の問題に特化した「Embrace Accountability」(アカウントビリティを受け入れる)という原則は、LinkedIn が潜在的な有害性や目的への適合性を評価し、それに対処することや、人間による監督とアカウントビリティを確保することを含め、堅牢な AI ガバナンスを展開することを意味しています。LinkedIn は、AI のベスト プラクティス、規範、法律が変化する中で、他者から学び、他者を支援することに取り組んでいます。

Q10-3 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。

該当なし

Q10-4 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。

該当なし

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。

該当なし

Q 1 1 - 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。

該当なし

Q 1 1 - 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。

該当なし

1 2 マスメディア（新聞・放送）との連携状況について

Q 1 2 - 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。

LinkedIn は、日本の多くのメディアと協力して、日本向けに LinkedIn のサービスを紹介しています。

Q 1 2 - 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。

該当なし

1 3 利用者の ICTリテラシー向上に向けた取組について

Q 1 3 - 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。

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しています。

Q 1 3 - 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。

該当なし

Q 1 3 - 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。

該当なし

14 研究開発の推進に向けた取組について

Q14-1 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。

該当なし

Q14-2 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。

該当なし

Q14-3 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。

該当なし

Q14-4 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。

該当なし

Q14-5 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お考えをお聞かせ下さい。

該当なし

Q14-6 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。

該当なし

15 サイバーセキュリティ関係機関等との連携について

Q15-1 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。

該当なし

Q15-2 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。

該当なし

Q15-3 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。

該当なし

16 行政機関や地方公共団体等との連携について

Q16-1 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 (GIFST) 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。

該当なし

Q16-2 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。

該当なし

17 国際機関等との連携について

Q17-1 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。

該当なし

Q17-2 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。

該当なし

Q17-3 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。

該当なし

18 その他のステークホルダーとの連携状況について

Q18-1 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。

※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、イベントの共催・後援など。

該当なし

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。

該当なし

Q18-3 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。

該当なし

Q18-4 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。

該当なし

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】（3）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。

該当なし

20 その他

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。

該当なし