

プラットフォームサービス事業者におけるフェイクニュースや偽情報への対策の取組事例のヒアリング

社名	Meta Platforms, Inc.	
1. 前提（サービス概要）		
①	対象サービス名	Facebook
	サービス分類	ソーシャルメディア
②	加入者数、月間アクティブユーザ数又は書き込み数	【日本の数値】
		【グローバルの数値】 29億6000万人(2022年12月時点)
		【(質問に答えられない場合) 参考となる数値】
2. 個別の取組事例		
取組事例1 一般的な考え方及び対応策		
①	<ul style="list-style-type: none"> ✓ 取組を行った契機や、背景にある課題意識（特定の事案が契機となっている場合には、当該事案の概要及び貴社の一次的な対応） ✓ その課題がどの程度貴社や貴社のユーザに悪影響を及ぼしていたか 	<p>【取組の契機や背景にある課題意識】</p> <p>Misinformation は、包括的な禁止事項を明示する方法がないため、弊社のコミュニティ規定で取り扱う他の種類の発言とは異なります。例えば、過度な暴力描写やヘイトスピーチについては、弊社が禁じている発言はポリシーに規定されているため、そのポリシーに賛成しない人に対しても施行することができます。しかし misinformation については、そのような方針を提供することができません。世界は絶え間なく変化し続けているため、ある時点では真実であったことが、次の瞬間には真実でなくなることがあります。また、人によって、自身の周りの世界について有する情報のレベルが異なるため、真実でない情報を真実だと信じてしまうこともあります。弊社はこういった情報にアクセスできず、「misinformation」を単純に禁止するポリシーを置いたとしても、弊社のサービスを利用する人々に対する有用な告知とはなりませんし、またそのようなポリシーに強制力を持たせることもできないと考えてい</p>

		<p>ます。</p> <p>代わりに、弊社のポリシーでは、misinformation についてさまざまなカテゴリーを明確にし、対象となる発言を弊社が見つけたときの対処法を示した明確なガイダンスを設けるよう努めています。それぞれのカテゴリーにおいて、弊社のアプローチには、表現、安全、尊厳、真正性、プライバシーに対する価値観のバランスを取る弊社の試みが反映されています。</p> <p>弊社は、差し迫った物理的な危害のリスクを直接助長する可能性が高い misinformation は削除します。また、政治過程の機能の妨害を直接助長する可能性のあるコンテンツや、人を欺くよう加工された特定のメディアも削除します。このようなカテゴリーにおける misinformation の要素を判断するにあたり、弊社は、知識と専門性を有する独立した専門家と連携し、コンテンツの真実性や、差し迫った危害のリスクを直接助長する可能性が高いかどうかを評価します。例えば、当事国で活動する人権団体と連携して内戦に関する噂の真偽を判断したり、新型コロナウイルス感染症の世界的流行時に保健機関と連携したりすることが挙げられます。</p> <p>その他のすべての misinformation については、弊社は、その情報の表示頻度の抑制や、生産的な対話を促す環境作りに力を入れています。私たちは、例えば、大げさなことを言うときに(例:「このチームの成績はスポーツ史上最悪だ!」)や、ユーモアや風刺を伝えるとき(例:「私の夫は『ハズバンド・オブ・ザ・イヤー』を受賞した!」)、害のない方法で「misinformation」を利用することが多々あります。また、不正確な情報を含んだストーリーを交えて体験談をシェアすることもあります。場合によっては、自分では深く信じていても、他人にとっては虚偽だと思える意見や、自分では真実だと信じているが、他人には不完全であったり誤解を</p>
--	--	--

	<p>まねいたりすると捉えられる情報を伝えることもあります。</p> <p>弊社は、このような発言が一般的なものであると認識しているため、大げさな話や拡散しやすい misinformation が広がることを抑え、利用者に信頼できる情報を提供することに力を入れています。その取り組みとして、弊社は第三者のファクトチェック団体と連携し、弊社のプラットフォームで最も拡散しやすいコンテンツの正確性についてレビューおよび評価を行っています（弊社のファクトチェックプログラムの仕組みについては下記 URL をご覧ください）。また、利用者が閲覧、信頼、共有するコンテンツを自分自身で決められるよう、メディアとデジタルのリテラシー向上のための リソース（下記 URL 参照）も提供しています。</p> <p>https://www.facebook.com/formedia/blog/third-party-fact-checking-industry-investments （英語のみ）</p> <p>https://transparency.fb.com/en-gb/features/how-fact-checking-works/ （英語のみ）</p> <p>最後に、弊社は、misinformation の拡散と重複することが多い他の分野でのコンテンツや言動も禁止しています。例えば、弊社のコミュニティ規定により、偽アカウント、不正行為、組織的な偽装行為は禁じられています。</p> <p>https://transparency.fb.com/ja-jp/policies/community-standards/account-integrity-and-authentic-identity/</p> <p>https://transparency.fb.com/ja-jp/policies/community-standards/fraud-deception/</p> <p>https://transparency.fb.com/ja-jp/policies/community-standards/inauthentic-behavior/</p>
--	---

		<p>オンライン・オフライン環境の変化や進化に伴い、弊社のポリシーも引き続き進化を続けていきます。以下に掲げる misinformation を繰り返し発信するページやグループ、プロフィール、Instagram アカウントは、コンテンツの削除に加え、配信数の削減や広告機能の制限、弊社のプラットフォームからの排除といった措置を受ける場合があります。Facebook がコンテンツを削除した場合の詳細は、下記 URL でご覧いただくことができます。</p> <p>https://www.facebook.com/help/260743102021762</p> <p>https://transparency.fb.com/ja-jp/features/approach-to-misinformation/</p>
②	✓ 解決手段の目的・ねらい	<p>【解決手段の目的・ねらい】</p> <p>Misinformation は、オフラインとオンラインの様々な行動を伴う複雑な社会現象であり、また、単一のプラットフォームを超えるものです。オンライン上の他の有害コンテンツとは異なり、前述のとおり、何が禁止されるべきかを明確に示す方法はありません。政府、政策立案者、市民団体、学者、ジャーナリスト、そして一般の人々が、misinformation とは何かについて同意しているわけでもありません。ある人が虚偽または misinformation であると考え、別の人にとっては単なる意見でしかないかもしれません。何が misinformation なのかを定義することは非常に困難ですが、さらに困難なのは、何かが真実であるか否かを決定するのが誰なのか、つまり、何が真実の源なのかを決定することであり、これにはしばしば異なる見解が付きまといます。</p> <p>弊社の misinformation に対するアプローチは、以下に示す 3 つの戦略 - 削除、抑制、情報提供 - に従っていますが、インターネット上に存在するコンテンツの量や、それが共有されるプラットフォームの多様性を考慮すると、オンラインで misinformation コンテンツを個々に</p>

		<p>禁止または削除することは、大規模に実施可能な解決策ではないことも認識しています。また、何が真実か嘘かを判断し、何が風刺、意見、議論であるかを判断するとなると、コンテンツの審査は極めて複雑になります。</p> <p>したがって、人々が情報を批判的に評価し、十分な情報を踏まえた上で意思決定を行い、自ら誤りを正すことができるようなデジタルにリジリエントな（回復力の高い）社会を構築することが必要だと考えています。</p>
③	<ul style="list-style-type: none"> ✓ 課題解決策の導入について障害になった事項 ✓ どのように障害を乗り越えたか 	<p>【課題解決策の導入に際しての障害】</p> <p>Misinformation コンテンツに対処し、コミュニティ規定を施行するためのポリシーを策定しようとする一方で、最大の課題は、異なる基本的権利間のバランスを取ることです。このバランスへの配慮は、オープンなインターネットを守るためには必要ですが、次第に脅かされつつあり、一部の国においてはオープンなインターネットが「壁で囲まれている」という状況にさえあります。</p> <p>上記の観点から、弊社の misinformation に対するアプローチは進化しており、表現、安全、尊厳、真正性、プライバシーなどの基本的権利の価値をバランスさせる試みを反映しています。</p>
④	<ul style="list-style-type: none"> ✓ 解決手段の詳細 	<p>【解決手段の詳細】</p> <p>弊社の misinformation に対する取り組みは、表現の自由とのバランスを取りながら、正確で有益なコンテンツを利用者に提供するという原則に導かれています（下記「その他」参照）。利用者は弊社のプラットフォームで質の高いコンテンツを見たいと考えており、弊社としてもその思いは同じです。そのため、弊社の misinformation 対策に係る戦略は、上述のとおり、削除、抑制、情報提供の3つの部分から構成されています。</p>

<https://about.fb.com/ja/news/2019/05/remove-reduce-inform-new-steps-remove/>

この戦略では、弊社のポリシーに違反するコンテンツを削除し、弊社のポリシーに違反しないものの、プラットフォームの信頼性を損なう問題のあるコンテンツの拡散を抑え、クリック、閲覧、共有するものを利用者が選択できるように、追加の情報を提供しています。

1. 弊社は、偽アカウントや偽装行為に関与したアカウント、暴力と扇動、危害を加えるための計画および犯罪の助長、ヘイトスピーチ、いじめや嫌がらせ、誤解を招くように操作された動画、暴力を誘発する共謀ネットワークに寄与する可能性のある misinformation、そして差し迫った身体的危害に寄与する misinformation、といった、コミュニティ規定に違反するコンテンツを削除します。また、第三者のファクトチェッカーや特定の状況下で権威ある団体によって否定された主張を含む広告や、広告ポリシー及びコミュニティ基準に違反する広告も削除されます。

<https://transparency.fb.com/ja-jp/policies/community-standards/>

<https://transparency.fb.com/ja-jp/policies/community-standards/inauthentic-behavior/>

<https://transparency.fb.com/ja-jp/policies/community-standards/violence-incitement/>

<https://transparency.fb.com/ja-jp/policies/community-standards/coordinating-harm-publicizing-crime/>

<https://transparency.fb.com/ja-jp/policies/community-standards/hate-speech/>

<https://transparency.fb.com/ja-jp/policies/community-standards/bullying-harassment/>

<https://about.fb.com/news/2020/01/enforcing-against-manipulated-media/> (英語のみ)

		<p>https://transparency.fb.com/ja-jp/policies/community-standards/dangerous-individuals-organizations/</p> <p>https://transparency.fb.com/ja-jp/policies/ad-standards/</p> <p>https://transparency.fb.com/ja-jp/policies/ad-standards/unacceptable-content/misinformation/</p> <p>2. 弊社は、「コミュニティ規定」に基づく削除の基準には当たらないものの、弊社のプラットフォームの信頼性と完全性を損なうような問題のあるコンテンツの配信を抑制しています。これには、クリックベイトのような低品質のコンテンツや、第三者のファクトチェック機関によって否定されたコンテンツが含まれます。弊社は、Facebook のニュースフィードでコンテンツの表示を抑制し、Facebook と Instagram でそのコンテンツを目にする人の数を大幅に減らしています。</p> <p>https://about.fb.com/news/2016/08/news-feed-fyi-further-reducing-clickbait-in-feed/（英語のみ）</p> <p>https://www.facebook.com/formedia/mjp/programs/third-party-fact-checking/partner-map（英語のみ）</p> <p>a. <u>第三者ファクトチェック</u> 弊社は、60 以上の言語をカバーする、世界中の 80 以上の第三者ファクトチェック機関と提携しています。これらのファクトチェック機関は、超党派の International Fact-Checking Network (IFCN) により認定され、弊社のプラットフォーム上のコンテンツの正確性を評価・審査しています。ファクトチェック</p>
--	--	---

パートナーによって虚偽と評価された投稿は、配信を減らし、その投稿を目にすることが少なくなるようにしています。

<https://www.facebook.com/business/help/2593586717571940>

3. 弊社は、利用者に情報を提供し、misinformation の拡散を防ぐために、さらなる文脈情報を提供し、利用者が十分な情報を踏まえた上で意思決定できるよう、正確な情報につなぐようにしています。例えば；

a. Misinformation ラベル

Facebook と Instagram のコンテンツにおいて、弊社のファクトチェック機関によって虚偽または一部虚偽と評価されたものは、目立つようにラベル付けされ、利用者が何を読み、信頼し、共有するかを自分で判断しやすくしています。

b. 正確で権威ある情報と利用者をつなぐ

弊社は、「新型コロナウイルス感染症情報センター」などの機能を立ち上げし、トピックに関する信頼性の高い正確な情報と利用者をつないでいます。

https://www.facebook.com/coronavirus_info/

c. 共有されるコンテンツに関する追加的な文脈情報

共有しようとするニュース記事が 90 日以上前のものであることを知らせる新しい通知画面を導入しました。

		<p>d. オフラインでも、さまざまなプログラムに投資し、市民団体や業界のパートナーと提携して、デジタルとニュースのリテラシーという根本的な問題に対処するためのさまざまな取り組みを行っています。例えば、「みんなのデジタル教室」は、利用者がオンラインで批判的に考え、思慮深く共有できるようにするための短いレッスンや学習モジュールを集めたもので、弊社の主力プログラムの一つです。すぐに使えるオンライン・モジュールは、アジア太平洋地域のオンライン・セーフティ専門家と共同で開発されました。現在、18カ国で利用可能で、日本語を含む15カ国語以上で提供しています。</p> <p>https://wethinkdigital.fb.com/jp/ja-jp/</p>
⑤	✓ 解決手段による効果	【解決手段による効果】
	✓ 他の取組への応用の可能性	【他の取組への応用の可能性】
取組事例2 影響工作、組織的偽装行為への対策		
①	<p>✓ 取組を行った契機や、背景にある課題意識（特定の事案が契機となっている場合には、当該事案の概要及び貴社の一次的な対応）</p> <p>✓ その課題がどの程度貴社や貴社のユーザに悪影響を及ぼしていたか</p>	<p>【取組の契機や背景にある課題意識】</p> <p>弊社は、misinformationに係る活動を探知、阻止するために、弊社プラットフォームをまたいで積極的に活動しています。</p> <p>しかし、「misinformation」という言葉に近いものとして、「disinformation」という言葉がありますが、この言葉はしばしば不正確に、misinformationと同じ意味で用いられ、外国からの干渉／選挙妨害、影響力工作、情報操作、さらには名誉毀損といった意味で使用されることがあります。</p> <p>本資料においては、「disinformation」とは、戦略的目標のために公的な議論を操作するための</p>

		<p>組織的な取り組みであり、騙すことを意図し、偽装的な行動を伴うものを指すこととします。これは、虚偽または誤解を招く内容を指す「misinformation」とは明確に異なるものであることを付け加えておきます。</p> <p>また弊社では、戦略的目標のために公共の議論を操作したり誤らせたりすることを目的とした組織的な取り組みを表す「影響工作」(Influence Operations) や「組織的偽装行為」(Coordinated Inauthentic Behaviour) といった、より具体的な用語を用いて、disinformationの議論を明確にしようとしています。</p>
②	✓ 解決手段の目的・ねらい	<p>【解決手段の目的・ねらい】</p> <p>「影響工作」(Influence Operations: IO) の主要な指標は、「偽装的であること」と「組織的に行われたものであること」の2つです。影響工作に加担する者は、必ずしも事実と異なる情報を使う必要はありません。また、そうした者たちにより拡散されるコンテンツのほとんどは、明らかに誤った情報ではなく、権威ある者によって広められた場合には政治的な発言として受け入れられるようなものです。</p> <p>真の問題は、こうした工作活動の背後にいる行為者が、その背後にある組織の身元を隠したり、組織やその活動を実際よりも人気があったり信頼できたりするように見せたり、弊社のコミュニティ規定の施行を回避するなどの、人を欺くような行動を採っていることです。</p> <p>弊社の組織的偽装行為に係るポリシーは、このような欺瞞的な行為に対処することを目的としています。真正性へのコミットメントに基づき、Facebook上で自己を偽ったり、偽のアカウントを使用したり、コンテンツの人気を人為的に高めたり、コミュニティ基準に基づくその他の違反を可能にするような行為に関与することは許可されていません。</p>
③	✓ 課題解決策の導入につ	<p>【課題解決策の導入に際しての障害】</p>

	<p>いて障害になった事項 ✓ どのように障害を乗り越えたか</p>	
④	<p>✓ 解決手段の詳細</p>	<p>【解決手段の詳細】 弊社は、1) 介入・妨害の抑止、2) misinformation との戦い、3) 透明性の向上という3つのアプローチで、disinformation に対して取り組んでいます。</p> <p>弊社は、影響工作のネットワークの破壊に焦点を当てたチームを拡大し、法執行機関、国家安全保障、調査報道、サイバーセキュリティ、法律、エンジニアリングのバックグラウンドを持つ専門家を社内全体に集めました。そして、弊社は、不正なアカウントや行動の検出と拡散防止に役立つ大規模なソリューションの構築を続けています。</p> <p><u>- 介入・妨害の抑止</u> 弊社の目標は、影響工作員が弊社のプラットフォームを操作するのを完全に抑止または阻止することです。</p> <ul style="list-style-type: none"> - 弊社は影響工作の脅威を阻止するために、政府関係当局、法執行機関、セキュリティ専門家、市民団体、他のハイテク企業と協力し、直接的なコミュニケーションラインを確立し、知識を共有し、協力の機会を見出すようにしています。 - 弊社は、不正アカウント、ページ、グループなどの影響工作ネットワークを削除し、新たな脅威をより迅速に特定するために、人材とツールによる調査業務の規模を拡大し続けています。これまでに、組織的偽装行為行為に関与した数万件のページ、グループ、アカウントを削除してきました。

		<p>- また、新しい戦術に対抗する能力を向上させ、弊社のプラットフォームで見られるさまざまな欺瞞的行為（外国か国内か、国家か非国家かを問わず）に対してより迅速に行動するため、組織的偽装行為に係るポリシーの更新を続けています。</p> <p><u>- 透明性の向上</u></p> <p>弊社は、透明性の向上が監視の強化につながり、利用者が目にするコンテンツの理由やその背後にいる人物をより理解するのに役立つと考えています。弊社は、政治的な広告、ページ、投稿の透明性を高めるツールや製品を導入し、利用者が自分たちに影響を与えようとしているのは誰なのかを知ることができるようにしています。また、組織的偽装行為の取り組みに関する報告書を定期的に発行しています。</p> <p>https://about.fb.com/news/tag/coordinated-inauthentic-behavior/ （英語のみ）</p> <p>組織的偽装行為に係る弊社のポリシー（CIB ポリシー）は、影響工作に直接対処することを目的としています。このポリシーは、「複数の Facebook または Instagram のアセットを使用し、連携して不正な行動（弊社のポリシーで定義）を行い、偽アカウントの使用がその活動の中心となっている場合」と定義されており、弊社のプラットフォーム上で影響工作ネットワークを探知、特定、削除する方法を示しています。</p> <p>CIB ポリシーと並行して、アカウントの安全性と実名の使用に係るポリシーがあり、これによって毎日数百万の偽アカウントを削除しています。弊社の目標は、Facebook 上の偽アカウントをできるだけ多く削除し、影響工作を画策する者が弊社のプラットフォームで活動する機会を最小限に抑えることです。偽アカウントには、弊社のポリシーに違反する悪意あるア</p>
--	--	---

		<p>アカウント、スパムキャンペーンに利用され金銭的な動機があるアカウント、企業や組織、ペットなど人間以外の存在を表すために作られた個人プロフィールなどの悪意のないアカウントが含まれます。</p> <p>https://transparency.fb.com/ja-jp/policies/community-standards/account-integrity-and-authentic-identity/</p> <p>弊社は、以下の3つの方法で、プラットフォームを悪用する偽アカウントを阻止しようとしています。</p> <ul style="list-style-type: none">- <u>アカウント作成のブロック</u> 弊社のシステムは、アカウントが一つの場所から大量に作成されているかどうかを示す、さまざまなシグナルを探します。例えば、特定のIPアドレスを完全にブロックすることで、そのIPアドレスが弊社のシステムにアクセスできないようにし、アカウントを作成できないようにします。 - <u>サインアップ時にアカウントを削除</u> 弊社は、疑わしいメールアドレスの使用パターンや疑わしい行動、あるいは過去に削除した他の偽アカウントに関連するシグナルなどを組み合わせて、悪意のある行動の兆候を発見するようにしています。現在、弊社が削除しているアカウントのほとんどは、作成から数分以内にブロックされ、害を及ぼすことができないようになっています。
--	--	---

		<p>- <u>既存アカウントの削除</u></p> <p>上記の2つの防御策を突破したとしても、プラットフォームに侵入してくるアカウントがあります。多くの場合、最初は偽アカウントや悪意のあるものであることを示すシグナルがなかなか現れないからです。こうしたアカウントは、弊社の検知システムが不正な行動を特定したときや、利用者から報告があったときに発見されることになります。弊社は、アカウントがどのように作成され、どのように使用されているかに関する多くのシグナルを利用して、偽アカウントの可能性が高いかどうかを判断し、偽物の場合は無効化しています。</p>
⑤	<p>✓ 解決手段による効果</p> <p>✓ 他の取組への応用の可能性</p>	<p>【解決手段による効果】</p> <p>CIB ポリシーに違反したアカウントやページに対して取られた措置は、ニュースルームで報告されます。また、四半期ごとに「敵対的脅威レポート」を発行しています。</p> <p>https://about.fb.com/news/tag/coordinated-inauthentic-behavior/ (英語のみ)</p> <p>https://about.fb.com/?s=Adversarial+Threat+Report (英語のみ)</p> <p>【他の取組への応用の可能性】</p>
取組事例3 透明性とコントロールにより利用者をエンパワーする試み		
①	<p>✓ 取組を行った契機や、背景にある課題意識(特定の事案が契機となっている場合には、当該事案の概要及び貴社の一次的な対応)</p> <p>✓ その課題がどの程度貴</p>	<p>【取組の契機や背景にある課題意識】</p> <p>弊社は、利用者がオンライン上で目にするコンテンツやその体験、交流について、利用者がより大きなコントロールを持ち、より良い情報を得られるようにすることを約束しています。</p> <p>弊社は、オンラインの安全性と、misinformation や disinformation を含む有害なコンテンツに対処する最も効果的な方法は、利用者が何を読み、信頼し、共有するかを自ら批判的に決定できるようにするツールとリソースを提供することで、デジタルにレジリエントな(回復力</p>

	社や貴社のユーザに悪影響を及ぼしていたか	の高い) 社会を構築することだと信じています。弊社は、利用者に対してより高い透明性とコントロールを提供し、十分な情報を踏まえた上で意思決定に役立つ情報を提供し、メディアとデジタルに対するリテラシーを向上させることによって、これを実現します。
②	✓ 解決手段の目的・ねらい	<p>【解決手段の目的・ねらい】</p> <p>弊社は、安全性と保全性に係るポリシー、プロセス、システムの透明性を担保、すなわち、いかなる状況であれば安全性とセキュリティのリスクをもたらさないかを明らかにすることに努めています。透明性は、プラットフォームの努力を外部の目にさらすことで説明責任を促進し、ひいては弊社が下す決定に対して責任を負うことにつながると考えています。</p>
③	<p>✓ 課題解決策の導入について障害になった事項</p> <p>✓ どのように障害を乗り越えたか</p>	<p>【課題解決策の導入に際しての障害】</p>
④	✓ 解決手段の詳細	<p>【解決手段の詳細】</p> <p>これらの戦略や例に挙げた施策とは別に、弊社の戦略は、利用者が目にする投稿について追加の文脈や情報を提供し、権威ある情報とつなげることに重点を置いています。弊社が導入しているツールや製品には以下のものがあります。</p> <ul style="list-style-type: none"> - <u>正確で信頼性の高い情報と利用者をつなぐこと</u> 弊社は、正確で信頼性の高い、権威ある情報と利用者をつなげる新しい方法を探し続けています。これは、利用者が何を読み、何を信頼し、何を共有するかを決めるための手段を利用者に提供できるようにするものであり、弊社が misinformation と戦うための戦略の中核をなす要素です。

		<ul style="list-style-type: none"> - <u>利用者が共有するコンテンツについて、より多くの文脈の提供</u> 弊社の目標は、タイムリーで信頼性の高い、利用者にとって最も価値のあるコンテンツを、利用者自身が簡単に識別できるようにすることです。そのため、利用者が何を読み、何を信頼し、何を共有するかを決めるために、より多くのコンテキストを提供するツールを導入しました。 <ul style="list-style-type: none"> - 共有しようとしているニュース記事が90日以上前のものである場合、通知画面によりお知らせします。共有しようとしている記事が90日以上前のものである場合、通知画面が表示され、その記事が適切であると判断されれば、共有を継続することができます。特に新聞社からは、古い記事が最新のニュースとしてソーシャルメディア上で共有され、時事問題の状況を誤解させる可能性があるという懸念が表明されています。 - COVID-19に関連するリンクを共有しようとする際に、通知画面によってより詳細な状況を伝えることができます。また、信頼できる情報にアクセスできるよう、新型コロナウイルス感染症情報センターへの誘導も行っています。 - <u>ニュース記事に関する詳細な情報提供</u> ニュース記事の投稿にボタンを設け、それをクリックすることでその出版社のウィキペディアの項目からの情報などの文脈的情報を得られるようにします。また、その出版社のページをフォローしたり、その出版社の他の記事を見られたり、そのニュース記事がFacebook上でどのように共有されているかについての情報など、Facebookや他のソースから引き出された文脈情報をボタンをクリックすることにより得ることができます。場合によっては、それらの情報が利用できないときは、利用者にその旨を知らせますが、それもまた有益な文脈となりえます。
--	--	---

		<ul style="list-style-type: none"> - <u>虚偽の主張を繰り返し共有するページへの警告</u> 誤った情報をシェアしているページに「いいね！」を押そうとしている利用者には、そのページをフォローするかどうかを利用者が十分に判断できるように、警告を表示します。警告には、そのページが誤った情報を共有していることを示すメッセージなど、文脈に応じた追加情報へのリンクが含まれています。有害な虚偽の主張を繰り返し共有する場合、最終的に弊社のプラットフォームからそのページが完全に削除される可能性があります。 - <u>コミュニティ規定に繰り返し違反するグループに対する警告</u> コミュニティ規定に違反するグループに参加しようとしている利用者には、参加する前に十分な情報を得た上で判断できるように、警告を表示します。このようなグループへの招待通知を制限し、参加する可能性を低くします。既存のメンバーに対しては、そのグループのコンテンツの配信を減らし、フィードの下位に表示されるようにします。 <p>正確で権威のある情報と、より多くの文脈を利用者に伝えることは、単にコンテンツを削除するよりもインパクトのある方法です。例えば、ファクトチェック機関によって虚偽と判定された投稿をすべて削除しても、そのコンテンツはインターネットや他のソーシャルメディアプラットフォームには残り、食卓で話題に上がることもあるでしょう。このようなコンテンツを残し、ファクトチェック機関による調査結果を表示したり、権威ある情報を紹介したりすることで、利用者に重要な情報と文脈を提供しています。</p> <p>権威ある情報の提供に係る弊社の戦略は、「新型コロナウイルス感染症情報センター」や「気候学センター」などのプロダクトを投入することに重点を置いています。我々のプラットフ</p>
--	--	---

		<p>オーム上におけるコミュニティが特定の脅威（新型コロナウイルス感染症に伴う健康危機など）に直面し、その脅威について誤った情報が広まり、利用者が混乱するリスクが高く、信頼できる情報源や情報が広く合意されているときに、権威ある情報をより見やすく、アクセスしやすくすることで、これらの脅威に対する利用者の行動、態度、知識を変えたいと考えています。</p> <p>そうすることで、misinformation の拡散を抑え、不確実性に乗じて世間における言説を操作しようとする悪意のあるネットワークの効力を弱めようとするものです。</p> <p>国際的な人権専門家集団が指摘するように（下記 URL 参照）、 "政府やインターネット企業は、まず信頼できる情報を自ら提供することで disinformation に対処することが不可欠です...コンテンツの取り下げや検閲といった他の手段に頼ることは、公衆衛生にとって重要な情報へのアクセスを制限する結果になりかねず、必要性和比例性の基準を満たしている場合にのみ実施されるべきです"</p> <p>https://www.ohchr.org/en/press-releases/2020/03/covid-19-governments-must-promote-and-protect-access-and-free-flow</p>
⑤	<ul style="list-style-type: none"> ✓ 解決手段による効果 ✓ 他の取組への応用の可能性 	<p>【解決手段による効果】</p> <p>【他の取組への応用の可能性】</p>
取組事例 4 みんなのデジタル教室・デジタルシチズンシップ		
①	<ul style="list-style-type: none"> ✓ 取組を行った契機や、背景にある課題意識（特定 	<p>【取組の契機や背景にある課題意識】 取組事例 1 (1)に同じ。</p>

	<p>の事案が契機となっている場合には、当該事案の概要及び貴社の一次的な対応)</p> <p>✓ その課題がどの程度貴社や貴社のユーザに悪影響を及ぼしていたか</p>	
②	<p>✓ 解決手段の目的・ねらい</p>	<p>【解決手段の目的・ねらい】</p> <p>人々が情報を批判的に評価し、十分な情報を踏まえた上で意思決定を行い、自ら誤りを正すことができるようなデジタルにレジリエントな（回復力の高い）社会を構築することが必要だと考えています。</p> <p>ネット上で遭遇する脅威が多数あることを考えると、デジタル環境を安全かつ知的に渡っていくためには、デジタルリテラシーやメディアリテラシーの枠を超えた幅広いスキルと能力が必要とされます。必要なスキルを身につけるには、批判的思考やメディア・リテラシーだけでなく、安全を保ちながら、他人に敬意を持ちつつ、責任を持ってデジタル社会に参加する方法を学ぶ、総合的なアプローチが必要です。</p> <p>デジタルシチズンシップとは、この新しい空間であるネット社会で私たちが持つこととなった権利と義務を活用して渡っていくための手段ことです。また、最も重要なことは、他の人々とどのように関わりを持っていくのかということです。</p> <p>デジタルシチズンシップの育成は社会全体の責任であり、特定の利害関係者だけで達成できるものではありません。政府から産業界、市民団体、教育者、市民自身まで、あらゆるステ</p>

		ークホルダーを巻き込んだ学際的な戦略が必要だと考えています。
③	<ul style="list-style-type: none"> ✓ 課題解決策の導入について障害になった事項 ✓ どのように障害を乗り越えたか 	<p>【課題解決策の導入に際しての障害】</p> <p>デジタルシチズンシップの概念を普及させるためには、資金やインセンティブといった形で、より多くの制度的支援が必要です。これは、正式な教育だけに適用されるのではなく、長期的な戦略や教育プログラムの一部であるべきです。Meta は、オンラインプラットフォームができる限り、このようなデジタル教育の取り組みに協力し、支援したいと思います。</p>
④	<ul style="list-style-type: none"> ✓ 解決手段の詳細 	<p>【解決手段の詳細】</p> <p>日本での提供開始にあたり、Facebook Japan は、日本の教育現場のニーズに合った授業を実施するため、企業と連携した授業づくりを専門とする「NPO 法人企業教育研究会」をパートナー団体として迎え、コンテンツを共同制作しています。</p> <p>例えば、誰かを騙したり、冗談のつもりで投稿されたりした情報は、あっという間に拡散され、多くの人に影響を与えています。こういった「偽ニュース」が発信される動機や、情報を受け取る側の視点を考えます。「偽ニュースの見分け方」では、偽ニュースを見分けるための様々なポイントについて、アクティビティを通じて学びます。また、拡散しないためのコツを学びます。</p>
⑤	<ul style="list-style-type: none"> ✓ 解決手段による効果 ✓ 他の取組への応用の可能性 	<p>【解決手段による効果】</p> <p>みんなのデジタル教室の授業はこれまでに、日本でのべ 240000 名の生徒が受講しております。「偽ニュースの見分け方」の授業を受講した生徒の 93%がインターネットやアプリ、SNS への関心が高まったと回答しています。また、受講生徒の 90%がインターネットやアプリ、SNS の利用の仕方について考えが変わったと回答しました。</p> <p>https://wethinkdigital.fb.com/jp/ja-jp/</p>

		【他の取組への応用の可能性】
取組事例5 公正な選挙のための取り組み		
①	<ul style="list-style-type: none"> ✓ 取組を行った契機や、背景にある課題意識（特定の事案が契機となっている場合には、当該事案の概要及び貴社の一次的な対応） ✓ その課題がどの程度貴社や貴社のユーザに悪影響を及ぼしていたか 	<p>【取組の契機や背景にある課題意識】</p> <p>表現の自由を守り、選挙の保全性を保護することは、Facebookにとって最優先事項です。過去からの経験や、専門家や政策立案者からの情報に基づいて、弊社は選挙の安全性を高めるためのチームとテクノロジーに多大な投資を行い、最大の効果を発揮する場所に配備しています。</p> <p>誰もが弊社のテクノロジーで意見を述べ、有権者が自信を持って投票できるようにするため、Metaでは、選挙を守り、広告主の真正性、透明性、説明責任を高めることに力を注いでいます。</p> <p>表現の自由のためのテクノロジーを作るという約束は、弊社のすべての活動の基本です。弊社は、個人や組織がコミュニティを作り、自らの意見を持ち、重要だと考える問題に対して安心して関わられるようにすることが大切だと考えています。利用者一人ひとりが自らの考えを発信し、声を届け、多様な意見や情報をやり取りできるようにテクノロジーの安全・安心を確保することは、弊社の責任です。</p>
②	<ul style="list-style-type: none"> ✓ 解決手段の目的・ねらい 	<p>【解決手段の目的・ねらい】</p> <p>デジタルプラットフォームがどのように選挙を守るべきか、広く合意された規範や法律がない中、弊社のようなプラットフォームは、政治広告の不透明性や国内の市民イベントへの外国からの干渉といった問題について対応しています。</p> <p>弊社の目標のひとつは、弊社のサービスが健全な市民生活に貢献する一助となることです。。</p>

		<p>なお、政治広告主に対して明確な業界全体の基準を設ける規制は、選挙期間中にオンラインプラットフォーム全体でより公正で透明性の高い環境を作るための鍵になると考えています。</p> <p>弊社は、政治広告の透明性に関して業界をリードする立場を採っており、多くの国で現在法律で求められている以上の新しいツールを導入し、政治広告を許可する他のどのプラットフォームよりもさらに進んで対応しています。</p> <p>ここ数年、弊社では、選挙の保護を強化し、有権者に投票を促し、そして多様な考え方をシェアできる安全なテクノロジーを開発するため、そうしたチームとテクノロジーに多額の資金を投じてきました。</p> <p>他方、例えば、社会問題、選挙または政治に関連する広告のポリシーで定められた重要なトピックについての議論、討論、支持／反対の表明を含む広告など世論や投票行動に非常に大きな影響を与える種類の表現があることがあります。</p> <p>https://transparency.fb.com/ja-jp/policies/ad-standards/SIEP-advertising/SIEP/</p> <p>弊社は今でも、基本的には利用者と広告主に表現の自由を与えるべきだと考えていますが、その一方で、利用者や行動、何らかの結果（選挙など）に影響を与えかねない広告については一定のガイドラインが必要であることも認識しています。</p>
③	<ul style="list-style-type: none"> ✓ 課題解決策の導入について障害になった事項 ✓ どのように障害を乗り越 	【課題解決策の導入に際しての障害】

	越えたか	
④	✓ 解決手段の詳細	<p>【解決手段の詳細】 選挙を守るための弊社の戦略は、重要な期間だけでなく、年間を通じて適用され、3つの分野が中心となっています。</p> <ul style="list-style-type: none"> - 干渉の防止 - 有害なコンテンツを削除し、misinformationを減らす - 透明性の向上 <p><u>i. 干渉の防止</u> 上記取組事例2でも説明したように、干渉を防ぐための弊社の戦略の重要な部分は、政府関係当局、法執行機関、セキュリティ専門家、市民団体、他のハイテク企業と協力して、新たな脅威を阻止し、操作仕事を阻止するために弊社のセキュリティ業務を前進させることです。</p> <p>https://about.fb.com/news/2018/11/investigating-threats/#working-with-partners</p> <p><u>ii. 有害コンテンツの削除とmisinformationの削減</u> 上記取組事例1で説明したように、弊社は3つの戦略 - 削除、抑制、情報提供 - を適用して、弊社アプリ全体で問題のあるコンテンツを管理しています。</p> <p><u>iii. 透明性の向上</u> 弊社は、選挙を保護し、信頼性と透明性を高め、説明責任を果たすことに尽力しています。政治的な広告やページに関して、業界をリードするレベルの透明性を提供し、利用者が誰が</p>

	<p>自分たちに影響を与えようとしているのかを知ることができるようにしています。</p> <p><u>選挙・政治的な広告主の検証</u> 選挙、政治広告を掲載するには、その広告主が誰であるか、どこに住んでいるかを証明するなどの承認プロセスを経なければなりません。</p> <p><u>選挙・政治的広告</u> 選挙、政治に関する広告には、免責事項が表示され、広告の背後に誰がいるかが示されます。これらの広告は、一般に検索可能な広告ライブラリに格納されます。</p> <p><u>ページの透明性</u> Facebook では、ページの作成日、名前の変更、ページ管理者の所在地など、ページに関する情報を表示します。また、全体的または部分的に政府の編集統制下にあると考えられるメディアを、国家統制メディアとしてラベル付けすることを開始します。</p> <p><u>ニュースフィードのランキング</u> 投稿や広告にある「この広告・投稿が表示される理由」をクリックすると、その広告等が表示される理由を理解したり、ニュースフィードで友人やページ、グループから何を見るかをコントロールしたりすることができるようになります。これは、ランキングの仕組みに関する情報をアプリに直接組み込んだ初めての試みです。</p> <p>Facebook と Instagram では、広告の透明性と選挙の公正性を高めるための取り組みを継続的に進めています。そのため、選挙または政治に関連する広告の掲載をする個人または団体</p>
--	--

	<p>の真正性と正当性を確認するプロセスを実施しています。</p> <p>これらの広告の掲載を予定している広告主は、広告の認証プロセスを完了する必要があります。このプロセスは、広告掲載時に対象国に居住および所在している広告主向けのものです。認証にあたっては、広告主の所在する国または州・都道府県が発行した、失効していない書類が必要になります。また、公証済みフォームの利用が認められている国もあります。</p> <p>https://www.facebook.com/business/help/208949576550051?id=288762101909005</p> <p>このような広告には、免責情報を含め、広告費用を支払った広告主個人または団体の名称を記載する必要があります。免責情報が表示されていない広告は配信が停止され、広告主が認証プロセスを完了するまで、承認されていない広告として広告ライブラリに収録されます。</p> <p>広告の透明性への継続的な取り組みの一環として、社会問題、選挙または政治に関連する広告のパーソナライズに関するより多くの情報を提供する、透明性ツールのアップデートを公開しています。これらの情報は、広告認証や免責情報のツールが利用可能な全ての国において適用されます。</p> <p>2022年5月末には、Facebook Open Research and Transparency (FORT) を通じて、審査に合格した学術研究者に対して、社会問題、選挙または政治に関連する広告のパーソナライズに関する詳細な情報を公開しています。これらの情報は、それぞれの広告ごとに提供され、広告主が選択した興味・関心のカテゴリなどの情報が含まれます。FORT の研究者プラットフォームは、資格を持った学術研究者による、ソーシャルメディアが社会に与える影響に関</p>
--	---

		<p>する研究のために設立されたもので、その対象には利用者のプライバシーを保護する施策も含まれています。</p> <p>公開されている広告ライブラリにも、公開後に配信された社会問題、選挙または政治に関連する広告のターゲティング情報の概要が追加しています。今回のアップデートでは、位置情報や人口統計、興味・関心などの各広告のパーソナライズを使用して配信された社会問題、選挙または政治に関連する広告の総数と、それらの広告のパーソナライズに使用した社会問題、選挙または政治に関連する広告費の割合に関するデータを掲載します。また、カスタムオーディエンスや類似オーディエンスが使用されたかどうかの情報も含まれます。</p> <p>https://fort.fb.com/</p> <p>https://research.facebook.com/blog/2022/1/introducing-the-researcher-platform-empowering-independent-research-analyzing-large-scale-data-from-meta/</p>
⑤	✓ 解決手段による効果	【解決手段による効果】
	✓ 他の取組への応用の可能性	【他の取組への応用の可能性】
3. その他		
①	● その他	<p>【その他】</p> <p>有害コンテンツへの対応の枠組みは、うまく設計されれば、政府、企業、市民団体がどのように責任を分担し、協力していくことができるかについて明確な方法を示すことを通じて、インターネットの継続的な成功に寄与することができます。また、企業にとっては、有害コンテンツの拡散と閲覧を抑制する最善のシステムを構築しようとするインセンティブが確保されることとなります。他方、このような取り組みは、設計を誤れば、予期せぬ結果を招き、利用者のオンライン上の安全性を低下させ、オンライン上の言論や表現を抑制し、イノベー</p>

		<p>ションを遅らせる危険性があります。</p> <p>オンラインコンテンツへの対応の試みは、表現の自由や知る権利という基本的権利のバランスやオンライン上の有害コンテンツへの対応の必要性と大規模なコンテンツモデレーションを行うに当たっての実務的な限界という難しい問題を提起します。</p> <p>オンライン上のコンテンツへの対応の枠組みは、それが misinformation、ヘイトスピーチ、テロコンテンツなどに対するものであろうと、言論・表現の自由、安全の確保、プライバシーへの配慮、その他の基本的権利のバランスを取る必要があります。このバランスへの配慮は、このバランスは、一部の国においては脅かされつつあり、「壁で囲まれてる」という状況にある国さえいるというのが現状ですが、オープンなインターネットを守るためにより一層必要なものであると考えています。</p> <p>また、対応の枠組みは、さまざまなデジタルサービス、製品、(多様なコミュニティにサービスを提供している)プラットフォームが、それぞれのリスクプロファイルに最も適した方法でコンテンツを革新し、拡張させ、調整し、さらに、オンラインで進化する脅威にリアルタイムで対応するために戦術を再調整、転換できるような柔軟性を提供する必要がある。</p> <p>有害コンテンツへ対応するための様々な懸念とトレードオフのバランスを取るために、以下の原則を考慮する必要があります。</p> <ul style="list-style-type: none"> - <u>表現の自由</u> インターネットは、世界中の人々が自由に自己表現する力を劇的に増大させました。規制当局は、その決定が表現の自由やその他の基本的人権に与える影響を考慮する必要があります。
--	--	---

		<ul style="list-style-type: none"> - <u>インターネットのグローバルかつクロスボーダーな性質</u> 有害コンテンツに対処するためのアプローチは、インターネットのグローバルな規模と国境を越えたコミュニケーションの価値を尊重する必要があります。また、規制当局と規制の間の相互運用性を高めることを目指すべきです。グローバルでオープンなインターネットは、デジタル・コミュニケーションと情報の自由な流れを可能にする調和されたグローバルなルールによって、最もよく機能することになります。 - <u>柔軟性</u> 対応の枠組みは、コンテンツモデレーション技術の能力と限界を考慮し、インターネット企業が自社のリスクプロファイルに最も適した方法で革新と対応を行い、また、オンラインで進化する脅威をリアルタイムで反復、改善、対処するために戦術を再調整、転換する柔軟性が認められる必要があります。ある特定のプラットフォームやコンテンツの種類に有効なアプローチでも、他の場所に適用すると効果がなくなる（あるいは逆効果になる）可能性があります。 - <u>比例性と必要性</u> 規制当局は、問題となる有害コンテンツの重大性と表示頻度、法律上の位置付け、およびリスクに対処するために既に行われている他の取り組みを考慮する必要があります。 - <u>説明責任へのインセンティブ</u> 企業のコンテンツモデレーションシステムおよび手順における説明責任を確保することは、企業が安全、プライバシー、表現の自由のような価値のバランスを責任を持って取るためのインセンティブを生み出す最善の方法となります。セーフハーバーとグッドサマリタンの保護は、法的責任に直面するおそれがなく、企業がオンラインコミュニティの安全を確保するために自発的かつ積極的な措置を講じるインセンティブを
--	--	--

		<p>与えることとなります。</p> <ul style="list-style-type: none">- <u>社会全体から考える視点</u> 複雑なオンライン上のコンテンツの課題に対処するためには、社会全体の取り組みが必要です。公的機関、市民団体、報道機関、政党、学界、教育者、保護者、コンテンツ制作者、そして日々インターネットを利用する人々など、幅広いステークホルダーが役割を担っています。利用者の権利拡大、安全性の促進、オンライン上の有害コンテンツの削減は、様々なステークホルダー集団による協調的な努力と協力が必要となる共有の責任であると考えます。 <p>この原則は、オンライン上のコンテンツへの対応を発展させる際の指針となる一連の価値観を提供し、インターネットの性質と利点、ならびに国際的な人権原則、ベストプラクティスおよび基準への考慮を確保することになると考えています。</p>
--	--	--