

1 権利侵害情報等への対応について

Q1-1 権利侵害情報等への対応に係る利用規約等の内容・実態について以下の点をご教示ください。

Q1-1(1)

権利侵害情報等への対応に関して、利用規約等において、削除以外のモデレーション（表示順位の低下等）としてどのような対応を行うことと定めているのか。行為や情報の類型ごとに取得する対応が異なる場合、その旨を分かりやすく整理しご回答ください。

【回答】

Metaが公表しているコンテンツに関するエンフォースメント措置の枠組みは、transparency.meta.com/enforcement/taking-action に記載されています。

Metaは、コンテンツに対するエンフォースメントについて「削除(remove)」「抑制(reduce)」「情報提供(inform)」という3つの柱からなるアプローチを採用しており、このうち後二者は、コンテンツの削除を伴うことなく機能するものです。

Metaのサービスにおけるコンテンツのランキングに関するアプローチは、transparency.meta.com/features/explaining-ranking に記載されています。

Metaのシステムが、あるコンテンツを低品質または問題のあるものと評価した場合、Metaは、すべての利用者に対するフィードにおける当該コンテンツの拡散を抑制することがあります。こうした抑制(デモーションとも呼ばれます)に関するMetaのアプローチは、transparency.meta.com/features/approach-to-ranking/types-of-content-we-demote に記載されています。

本回答においては、エンフォースメント措置は次の7つの類型に分類されます。

1. コンテンツ単位の削除。
2. 削除に至らないコンテンツ単位の措置。表示範囲の制限および拡散の抑制(ランキング)、ならびに、センシティブまたは誤解を招くおそれのあるコンテンツに対する警告ラベルその他の文脈情報ラベルの付与を含みます。
3. 地域(地理的)単位の制限。本来的に地域性を有する問題に対して適用されます(例えば、特定の法域の法令の下で違法となるコンテンツ。裁判所の命令または有効な法的要請により制限が求められる場合を含みます)。
4. アカウントおよびエンティティ単位の措置。
5. コンテンツの収益化機能に関する措置。
6. 広告およびコマース領域に固有の措置。
7. コンテンツに対する措置を伴わない、救済およびユーザー保護のための措置。

Metaは段階的なエンフォースメントの枠組みを運用しており、適用される措置は、当該コンテンツの性質および侵害行為の類型に応じた比例的なものとなります。

上記4.の類型に関しては、この段階的な枠組みが具体的な形で公表されています(<https://transparency.meta.com/enforcement/taking-action/counting-strikes/>)。Metaはストライク制度を運用しており、アカウントに記録されたストライクの数に応じて、適用される制限が段階的に強化されます。当該コンテンツがMetaのより重大なポリシーに違反する場合には、追加的かつより長期の機能単位の制限が適用されることがあります。ほとんどの場合、アカウントの無効化は、繰り返しの警告および制限を受けた後もなお違反コンテンツの投稿を継続する場合に限って行われます。もっとも、児童の性的搾取に関するコンテンツの投稿

など、違反が十分に重大である場合には、Metaは、1回の違反をもってアカウント、FacebookページもしくはグループまたはInstagramアカウントを無効化することがあります。

Metaのポリシー違反を繰り返すFacebookページおよびグループについては、おすすめ表示からの除外や拡散の抑制が行われることがあります。ページについてはさらに一定の収益化機能の利用が制限されることがあり、グループについては管理者による投稿の事前承認が必須とされることがあります。

Metaは、コンテンツを削除した場合、またはアカウント、FacebookページもしくはグループAに制限を適用した場合には、その旨を利用者に通知します。ほとんどの場合、利用者は、FacebookおよびInstagramの「アカウントのステータス」を通じて、自己の違反履歴、アカウントに適用されている可能性のある一部の制限およびその期間を確認することができます。Metaの措置が誤りであると利用者が考える場合には、原則として異議申立てを行うことができます。当該コンテンツがコミュニティ規定に適合していると判断された場合には、当該コンテンツは復元され、これに伴うストライクおよび制限も取り消されます。

本回答に記載した措置は、Metaの利用規約に基づくものです。同規約は、

(i) 他者の権利(知的財産権を含みます)を侵害し、またはこれに違反するコンテンツの投稿を禁止し、

(ii) コンテンツが同規約もしくはMetaのポリシーに違反するとMetaが判断する場合、または法令上許容され、もしくは義務付けられる場合に、当該コンテンツ、サービスもしくは情報を削除し、またはこれらへのアクセスを制限する権利をMetaに留保し、

(iii) 利用者が同規約またはポリシーに明白に、重大に、もしくは繰り返し違反した場合に、当該アカウントへのアクセスを停止し、または恒久的に無効化する権利をMetaに留保するものです。

Q1-1(2)

権利侵害情報にあたると考えられるもののうち、利用規約等に基づく対応が困難と考えるものはどのようなものか。また、そうした情報の削除を行うのはどのような場合があり得るか。特に名誉毀損への対応、侮蔑的表現への対応に関する考え方については必ずご回答ください。また、裁判による判決が無ければ削除しないものがあれば具体的にご回答ください。

【回答】

コンテンツは、まずMetaのコミュニティ規定に照らして評価されます。コンテンツがコミュニティ規定に違反する場合、Metaは、法的な請求がなされているか否か、または裁判所の命令が取得されているか否かにかかわらず、コミュニティ規定違反を根拠として措置を講じます。

Metaは、現地法令上違法であると主張されているコンテンツについて、有効な申出または要請を受領した時点で審査を行います。当該コンテンツがMetaのコミュニティ規定に違反しない場合であっても、法的義務が存在するとき、すなわち裁判所の命令がある場合、権限を有する規制当局もしくは法執行機関からの有効な要請がある場合、または情報流通プラットフォーム対処法に基づく権利侵害情報のように、利用者からの申出を規律する現地法上の枠組みが適用される場合には、現地法令違反を理由として制限されることがあります。権利侵害の申出がなされたコンテンツのうち、Metaのコミュニティ規定に違反しないと評価された

ものについては、情報流通プラットフォーム対処法に基づく有効な申出がなされた場合、Metaは、同法第26条ガイドラインに公表された基準に照らして当該コンテンツを評価します。当該ガイドラインに基づいて現地法令に違反するか否かを判断できない場合、Metaは、当該申出の取扱いについて、指定した専門員に意見を求めます。

名誉毀損は法的概念であり、コミュニティ規定違反とは別個のものです。名誉毀損の成立を認定するには、事実関係に即した、かつ法域ごとの個別的な判断が必要となりますが、基礎となる事実関係にアクセスできない媒介者であるMetaにとって、そのような判断を行うことは困難です。第26条ガイドラインが適切な判断を示していない事案において、Metaは、申出のあったコンテンツが名誉毀損として違法であるか否かを判断し得る立場にはありません。

侮辱的または相手を貶める表現は、Metaのコミュニティ規定、主として「いじめと嫌がらせ」および「ヘイト行為」に関する規定において取り扱われます。コンテンツがMetaのコミュニティ規定に違反する場合、Metaが措置を講じるにあたり、法的な請求も裁判所の命令も必要ではありません。公表されているポリシーは次のとおりです。

<https://transparency.meta.com/policies/community-standards/bullying-harassment/>
<https://transparency.meta.com/policies/community-standards/hateful-conduct/>

この点についても、情報流通プラットフォーム対処法に基づき申出がなされた権利侵害情報が、これらのポリシーの対象外となる場合であっても、現地法令上の名誉権または名誉感情を侵害すると主張されているときは、上記の情報流通プラットフォーム対処法に基づくプロセスを通じて評価されます。

Q1-1(3)

申立ての対象となった投稿について、申立て理由と投稿内容が一致していないが、当該投稿内容が他の利用規約上の禁止事項に該当する場合は、当該他の利用規約に基づく対応は行われるか。

(例: 誹謗中傷を受けたとして申立てがあった場合であって、当該投稿は誹謗中傷には該当しないと考えられるが、個人を特定する情報が記載されているような場合はプライバシーに関する規約等に基づいて対処されるか)

【回答】

然り。申立てを行った者が選択した理由は、当該申立ての振分けおよび優先順位付けを決定するものであり、評価の範囲や、Metaが措置を講じる根拠となるポリシーを必ずしも限定するものではありません。Metaが情報流通プラットフォーム対処法に基づく年次報告書において述べたとおり、「Metaはコミュニティ規定のすべての規定に照らして通報されたコンテンツを審査するため、措置が講じられた際に問題とされた違反の類型が、もともとの通報理由とは異なる場合があります。」。

Q1-1(4)

権利侵害情報について、削除以外のモデレーションを含めた対応の実施の判断基準及び意思決定プロセス等、貴社のガバナンス体制について。また、公表されている場合は、その公表場所も併せてご回答ください。

【回答】

権利侵害コンテンツに関するMetaのガバナンスは、コミュニティ規定および利用規約といった公表されたポリシー、文書化されたエンフォースメントのプロセス、および独立した監督に

基づいています。Metaは、テクノロジーと人によるレビューチームを組み合わせでポリシーを執行しています。

Metaの[コミュニティ規定](#)は、Metaのサービス上で何が許容され、何が許容されないか、および利用可能なエンフォースメントの選択肢を定めるものです。コミュニティ規定は、利用者の皆様からのご意見、ならびに公共安全、テクノロジー、人権といった分野の専門家の助言に基づいて策定されています。同規定は全文が公表され、定期的に更新されており、日本語でもご覧いただけます。情報流通プラットフォーム対処法に基づく権利侵害情報の申出については、Metaは、コミュニティ規定および現地法令に照らしてこれを評価します。

Metaは、ポリシーの適用状況を明らかにするため、定期的に[グローバルな透明性レポート](#)を公表しています。日本においては、[情報流通プラットフォーム対処法に基づく年次透明性レポート](#)も公表しています。

Metaは、コミュニティ規定違反の種類の大部分について異議申立ての機会を提供しており、グローバルな異議申立件数および判断の取消率について、コミュニティ規定施行レポート(Community Standards Enforcement Report)において公表しています。同レポートには、異議申立てがなされたコンテンツ、復元されたコンテンツ、および異議申立てを経て当初の判断が覆された割合に関するデータが含まれています。また、対象となるコンテンツに関する判断については、監督委員会(Oversight Board)に付託することができ、同委員会はMetaから独立してこれを審査します。

削除に至らない措置に関する追加情報については、上記質問1-1(1)に対する回答をご参照ください。

Q1-1(5)

利用者が知るべき利用規約等の情報について、利用者が容易に認識・理解できるように実施している工夫について。

【回答】

Metaのアプローチは以下のとおりです。

- ポリシーは全文が、公開の形で、一箇所に集約して公表されています。Metaのコミュニティ規定は、Facebook、Instagram、MessengerおよびThreads上で何が許容され、何が許容されないかを一元的に確認できる窓口である「Meta透明性センター(Meta Transparency Center)」において公表されています。コミュニティ規定は、世界中のすべての利用者に適用され、AIによって生成されたコンテンツを含むあらゆる種類のコンテンツに適用されます。
- ポリシーは、単に読まれるためではなく、適用されるために書かれています。コミュニティ規定の各セクションは、当該ポリシーの目的を説明する「ポリシーの根拠(Policy Rationale)」から始まり、これに続いて、(i) 許可されないコンテンツ、(ii) Metaがエンフォースメントを行うために追加の情報または文脈を要するコンテンツ、(iii) 警告画面付きで許可されるコンテンツ、(iv) 許可されるものの18歳以上の成人のみが閲覧できるコンテンツを、それぞれ個別に定める具体的なポリシー項目が置かれています。これにより利用者は、抽象的な原則の表明ではなく、禁止の内容に加えて、その境界線とMetaの対応の段階性を確認することができます。同じことは、ポリシーに違反した場合の帰結についても当てはまります。例えばMetaは、コンテンツの重大性、文脈および時期に応じてストライク制度がどのように適用されるかを公表しています(上記質問1-1(1)に対する回答をご参照ください)。さらにMetaは、ストライク制度の

公平性および比例性を確保するための限界についても公表しています。すなわち、ほとんどの違反については90日より前に、より重大な違反については4年より前に投稿された違反コンテンツに係るストライクは算定に含まれず、また、すべてのストライクは1年で失効します。

- ポリシーの変更は公表され、公表が利用者に対する通知となります。Metaは、必要に応じてポリシーを変更する権利を留保しています。変更はいずれも透明性センターに反映され、また、長年にわたる具体的な変更点を示す詳細な変更履歴も提供されています。透明性センターは、ポリシーの現行の条文を掲載するものであると同時に、改定を伝達するチャンネルでもあります。
- 公表は、プロダクト内での提示と組み合わせられています。Metaは、公表された規定文のみに依拠しているわけではありません。利用者は、自らに関係する場面において、ポリシーの実質的な内容に触れることになります。例えば、Instagramにおいて攻撃的なコメントが投稿される前に表示される警告は、再度投稿が試みられた際にコミュニティ規定を改めて提示し、当該コメントが削除または非表示となり得ることを示します。Metaのポリシーに違反した利用者は、違反したポリシーを示し、コミュニティ規定へのリンクを付した通知を受け取ります。これは、利用者がMetaのポリシーを理解し、再度の違反を避けられるようにするためにMetaが講じている教育的な措置です。

Q1-2 権利侵害情報等の発信・拡散への事前の対応について、次の点をご教示下さい。

Q1-2(1)

権利侵害情報等を発信・拡散する行為を事前に抑止するための貴社のサービス設計・機能上の工夫等。(ご発表いただいたもの以外にある場合)

【回答】

質問1-2(1)から1-2(3)まで及び1-2(5)は、権利侵害のおそれのあるコンテンツの投稿および拡散を防止するために、プラットフォームのサービスがどのように設計・運用されているか、また、その設計がどのように評価・改善されているかに関するものですので、以下にまとめて回答します。以下に記載する措置は、Meta自身のコミュニティ規定のエンフォースメントに向けられたものです。

1. 作成・アップロードの時点における措置

(1) 能動的な検知およびエンフォースメント

Metaは、利用者その他の者からの申出を待つことなく、コミュニティ規定に違反するおそれのあるコンテンツを検知し、能動的に措置を講じています。Metaは、ポリシーに違反するおそれのあるコンテンツを能動的に検知するためのAIおよび機械学習システムに投資しています。これらのシステムの有効性は、[Metaコミュニティ規定施行レポート](#)における能動的検知率の指標を通じて公表されています。最新のレポートに記載のとおり、児童の性的搾取、危険な団体およびテロリズム関連コンテンツを含む多くのポリシー領域において、Metaのグローバルな能動的検知率は90%を超えています。

Metaの情報流通プラットフォーム対処法に基づく年次透明性レポートにおいて公表しているとおり、報告対象期間(2025年7月30日～2026年3月31日)において、Metaは日本国内において約287万件のコンテンツを、いかなる申出を受ける前に能動的に検知し、送信防止措置を講じました(Facebook: 822,163件、Instagram: 660,882件、Threads: 1,390,718件)。

(略)

(3) 投稿前の警告

Instagramでは、攻撃的とみなされるおそれのあるコメントを投稿しようとした利用者に対し、当該コメントが公開される前に警告が表示され、コミュニティ規定が改めて示されるとともに、当該コメントが削除または非表示となる可能性があることが示されます。

2021年8月時点で、これらの警告は世界全体で1日あたり約100万回表示されており、そのうち約50%のケースにおいて、利用者が自らコメントを修正または削除しました。本措置は投稿を禁止するものではなく、判断を利用者に委ねつつ、有害となり得るコンテンツを実効的に減少させるものです。

2. 拡散の防止およびやり取りに関するコントロール

(4) 拡散の抑制およびおすすめ表示の対象適格性

Metaは、コンテンツが低品質または問題のあるものである場合、フィードにおけるその拡散を抑制することがあります。上記質問1-1(1)に対する回答をご参照ください。

Metaはまた、低品質、不適切、センシティブ、または若年の閲覧者にとってふさわしくないおそれのあるコンテンツをおすすめ表示しないよう努めています。これらの基準に適合しないコンテンツは、削除されるか否かにかかわらず、おすすめ表示面に表示されません。

<https://transparency.meta.com/enforcement/taking-action/lowering-distribution-of-problematic-content/>

これらの拡散およびおすすめ表示に関するコントロールは、コンテンツの削除とは独立して機能します。したがって、Metaがコミュニティ規定に違反しないと評価したコンテンツについても、削除されることなく、フィードにおける拡散が抑制され、またはおすすめ表示面から除外されることがあります。

利用者はまた、投稿を非表示にすること、および類似のコンテンツをより多く／より少なく見たい旨を示すことなどにより、自らが目にするコンテンツをカスタマイズすることができます。

<https://transparency.meta.com/features/explaining-ranking/ig-feed-recommendations/>

(5) 拡散およびやり取りに関する利用者単位のコントロール

利用者による自らの利用体験のコントロール

Metaは、サービスの動作が利用者自身の選好を反映したものとなるよう、実質的なコントロールを利用者自身の手になています。すなわち利用者は、おすすめの投稿について「興味がある」または「興味がない」と示すこと、投稿を非表示にすることによっておすすめの内容を学習させること、InstagramおよびThreadsにおいて「Your Algorithm」を用いて、より多く見たいものおよびより少なく見たいものをMetaに伝えること、時系列順、「お気に入り」、「友達」またはカスタムのフィードに切り替えること、おすすめの投稿をスヌーズすること、ならびに、センシティブな内容の表示設定および政治的なコンテンツの表示設定を調整することができます。特定のアカウントとのやり取りが問題である場合には、利用者は、ミュート、ブロック、制限 (Restrict) または「Limits」を適用することができます。さらに利用者は、コミュニティ規定に違反すると考えるコンテンツを報告し、誤りであるとする判断について異議を申し立てることができます。

攻撃的なコメントの自動非表示 (Instagram)

Instagramは、Metaのコミュニティ規定に違反しない場合であっても、侮辱、性的な言及、個人または集団に対する人格攻撃など、侮辱的または相手を傷つけるおそれのある語句や絵文字を含むコメントを自動的に非表示にする設定を提供しています。

また利用者は、非表示にする語句、数字および絵文字を自ら設定することもできます（「非表示ワード」）。ティーンアカウントについては、これらの設定は自動的に有効化され、16歳未満のティーンは、親または保護者の許可なくこれを変更することはできません。

注目が急増した際の保護 (Instagram)

Instagramは「Limits」機能を提供しています。これは、注目が急激に高まった際に、自分をフォローしていない人、または最近フォローしたばかりの人からのコメントおよびメッセージリクエストを一時的に非表示にできるものです。この設計により、利用者は、自分を標的としている可能性のあるアカウントからの接触を制限しつつ、既存のフォロワーからの声には引き続き接することができます。

3. システミックリスクの管理、評価および改善

Metaは、権利侵害コンテンツに個別の事案ごとにのみ対応しているわけではありません。Metaは、その製品およびインフラ全体に適用される、体系的かつ標準化されたリスク管理プロセスを運用しています。同プロセスにおいては、対象となるリスクを特定し、体系的なアセスメントおよびステークホルダーとの対話を通じて証拠を収集し、固有リスクと既存のコントロールの有効性の双方を測定し、当該分野の専門家により知見の妥当性を検証し、比例的な軽減措置を特定した上で、これを文書化し、Metaのガバナンス・プロセスを通じて追跡します。Metaは、三つのディフェンスラインによる説明責任のモデルを採用しています。すなわち、インテグリティおよびプロダクトの各機能がリスクを保有し管理し、リスクおよびコンプライアンスの専門機能が監督を行いコントロールの有効性を評価し、内部監査が独立した保証を提供します。このサイクル全体を通じて得られた知見は、優先順位付け、是正および改善に活用されます。

Metaの[情報流通プラットフォーム対処法に基づく年次透明性レポート](#)には、権利侵害情報に関するMetaのコンテンツモデレーションのシステムおよびポリシーについて、同法に定める7つの事項のそれぞれに係る自己評価も記載されています。Metaは、情報流通プラットフォーム対処法に関する業務フローにおける潜在的な体系的課題を特定し是正するための社内のインシデント監視プロセスを維持しており、そこから得られた知見を、自社のシステム、社内プロセスおよび利用者向けのコミュニケーションの改善に活用しています。同レポートに記載のとおり、Metaは、日本の利用者にとって安全で透明性の高いオンライン環境を推進することに取り組んでいます。これらの優先課題には、Metaのステークホルダーからの意見を踏まえた継続的な改善が必要です。

異議申立件数および判断の取消率を含むグローバルなエンフォースメントに関するデータは、[コミュニティ規定施行レポート](#)において公表されています。

Q1-2(2)

誹謗中傷を含むネガティブな情報がトレンドに掲載される等、プラットフォーム側の仕組みを一つの要因として拡散される場合、それを抑制するための取組や問題意識。

【回答】

上記Q1-2(1)ご参照

Q1-2(3)

(2)も含む、権利侵害情報の発信・拡散に係るシステミックリスクへの対応に関する取組や問題意識。

【回答】

上記Q1-2(1)ご参照

Q1-2(4)

権利侵害情報の発信・拡散に係るシステミックリスク等の実態把握のための、国、研究者等へのデータアクセス提供による協力に関する現在の取組。

【回答】

Metaは、ポリシーの施行状況について一般に情報を提供するため、定期的に[透明性レポート](#)を公表しています。これには以下が含まれます。

- 「[コミュニティ規定施行レポート](#)」は、FacebookおよびInstagram全体において、Metaがコミュニティ規定をどのように施行しているかに関する包括的なグローバルデータを提供するものです。
- 「[コンテンツ制限レポート](#)」は、政府からの要請に応じ、または現地法に準拠して制限されたコンテンツに関する国別のデータを提供するものです。

さらに、Metaは「[責任ある事業活動報告書](#)」および「[人権報告書](#)」を公表しており、これらはいずれも一般に公開されています。これらの報告書では、責任ある事業活動および人権デューデリジェンスに対するMetaの幅広い取組みについて詳述しています。

Metaは、データセット、調査、マップおよびAPIを構築することにより、研究者が自社プラットフォームの政治的、経済的および社会的な影響を研究できるよう支援しています。Metaのツールおよびプロセスは、研究者がその研究を支える情報および分析機能にアクセスできるようにするものです。詳細については、Meta透明性センターの[リサーチツールおよびデータセットのページ](#)をご参照ください。

Metaは、その他の取組みにも参画しています。例えば[2022年3月](#)、Metaは、ハーバード大学バークマン・クライン・インターネット&社会研究センター(Berkman Klein Center for Internet & Society)が運営する独立した研究プロジェクトである「Lumen」への参加を表明し、2023年11月以降、同プロジェクトへの案件の提出を行っています。同プロジェクトは、オンラインコンテンツに関して政府および裁判所から出される削除要請について、研究者が調査・分析を行うことを可能にするものです。MetaがLumenに提出した削除要請は、[同プロジェクトのデータベース](#)において公開されています。これにより、インターネット利用者のデジタル権利の分析、報告および擁護に向けた国際的なコミュニティの取組みが、一層促進されることとなります。

Q1-2(5)

上記(1)～(3)で記載いただいた取組を含めた、①その効果の評価・分析 ② ①を踏まえた改善の仕組み。また、これらの公表状況。

【回答】

上記Q1-2(1)ご参照

Q1-2(6)

一般的監視（網羅的監視）とはどのようなものと考えているか。

【回答】

Metaは、「一般的監視」とは、プラットフォームに対し、継続的かつ網羅的に、自己のサービス上で伝達または保存されるすべてのコンテンツを監視すること、または違法な活動を示す事実もしくは状況を能動的に探索することを求め、それにより特定のコンテンツが違法であるか否かを自ら判断させる義務を意味するものと理解しています。

その本質的な特徴は、技術の利用や特定の検知手段の用い方にあるのではなく、また、措置が大規模に行われるという事実にあるのでもありません。本質は、媒介者に課される義務の範囲および性質にあります。すなわち、プラットフォームが、具体的に指摘を受けていないコンテンツを自ら精査し、当該コンテンツが違法であるか否かについて自ら判断することを求められる点にあります。

この種の義務を課すことの禁止は、国際的な媒介者責任の枠組みにおける重要かつ確立した原則です。このような義務は、利用者のプライバシー、データ保護および表現の自由に対して重大な影響を及ぼすものです。また、本来裁判所および権限ある当局に属する適法性の判断をプラットフォームに行わせることとなり、判断が不確かな場合には適法な表現であってもこれを削除する強い誘因を生じさせます。

Q1-2(7)

上記(1)～(3)で記載いただいた貴社における取組も含めて、権利侵害情報の発信・拡散を未然に防ぐ取組についてどのような場合が一般的監視（網羅的監視）にあたると思われるか。また、その理由。

【回答】

プラットフォームが自らのコンテンツポリシーを適用すること（当該ポリシーが大規模に能動的に執行される場合を含みます）は、それ自体で一般的監視に該当するものではありません。予防的措置が一般的監視に該当するか否かは、用いられる技術や検知の手段よりも、むしろプラットフォームに課される義務の範囲および性質によって決まります。具体的には、プラットフォームが、自らに対して個別に特定されて示されたものではないコンテンツを精査し、当該コンテンツが違法であるか否かについて自ら判断することを求められているか否かが問題となります。

プラットフォームは、自らのポリシーを能動的に、かつサービス全体にわたって適用することができますが、そのことによって、自らが事業を行うすべての法域の法令の下で利用者のコンテンツのすべてが適法であるか否かを判断することを求められるものではありません。上記(1)から(3)までに記載したMetaの措置は、この類型に属します。これらは、Metaのサービス上で許容されないコンテンツの類型を定めるMeta自身のグローバルなコンテンツポリシー（例えばコミュニティ規定）を執行するために、大規模に適用されている自主的な措置です。Metaのグローバルなコンテンツポリシーは、個別の事案において法令上の禁止と重なり得るものの、いずれの法域の法令とも範囲を同じくするものではなく、また、Metaに対して利用者のコンテンツのすべての適法性を判断することを求めるものでもありません。

これに対し、プラットフォームが、継続的かつ組織的に、自らに対して個別に特定されて示されたものではないコンテンツを精査し、当該コンテンツが適用される法令または権利を侵害するか否かを自ら判断するための措置を実施することを法律上義務付けられる場合には、

状況は実質的に異なることとなります。措置が予防的であること、自動化されていること、または大規模に適用されていることは、Metaの見解では、決定的な要素ではありません。

Metaが、日本法の下でコンテンツが権利または利益を侵害するか否かを評価する場合、その評価は、個別に特定されたコンテンツに関する有効な申出または要請に応じて行われます。当該評価は、Metaに対して個別に示されたコンテンツに向けられたものであり、利用者のコンテンツの違法性を監視する一般的な義務を構成するものではありません。

2 災害発生時の対応について

Q2-1 災害時の対応について以下の点をご教示ください。

Q2-1(1)

利用規約等において、特に災害時の情報流通に関して定めている事項。

【回答】

Facebook、InstagramおよびThreadsに関するMetaの利用規約は、利用者の行為、禁止されるコンテンツの種類（詳細はコミュニティ規定に定めています）、知的財産および利用者の権利に関する一般的な規定から構成されています。同規約には、災害時の情報流通について特に定めた規定は含まれていません。

コミュニティ規定は、Metaのプラットフォーム上で何が許容され、何が許容されないかを定めるものであり、災害時を含め、常に適用されます。災害時の情報流通に対するMetaの具体的な支援は、以下に記載するプロダクト機能および運用上の枠組みを通じて提供されています。

Q2-1(2)

信頼できる情報源からの情報の優先表示等、災害時に実施している、もしくは実施を検討している取組。（ご発表いただいたもの以外にある場合）

【回答】

（Facebook、InstagramおよびThreadsに共通の回答）

Metaは、利用規約上の定めとは別に、専用のプロダクト機能および確立された運用上の枠組みを通じて、災害時の情報流通を支援しています。

プロダクト機能（危機対応ツール）

危機対応（**Crisis Response**）は、特定の危機に関する情報およびツールを、信頼できる情報源による当該事象に関する権威ある情報を含めて、Metaが一元的に集約する画面です。その中に、次の機能が含まれます。

- 災害時の安否確認（**Safety Check**） — 被災地域の人々が自らの無事を知らせ、友達の安否を確認することができる機能です。地震など人々が危険にさらされるおそれのある事象が発生した場合、世界各地の危機情報を扱う機関からMetaに通知が行われ、被災地域の人々が当該事象について投稿している場合に、災害時の安否確認が有効化されます。災害時の安否確認は、もともと2011年の東日本大震災を受け

て開発されたものです。日本における近時の有効化事例としては、2026年7月28日の熊本地震、および2026年8月の千葉県における大雨があります。

支援情報の共有(**Community Help**)は、危機の発生中および発生後において、避難所、物資、移動手段といったリソースの調整を含め、人々が支援を求め、また支援を提供することを可能にする機能です。

(略)

Q2-1(3)

災害に備えて平時より実施している取組。

(利用者向けの啓発活動、災害時の情報確認方法に関する案内等)

【回答】

(Facebook、InstagramおよびThreadsに共通の回答)

以下は、質問2-1(2)に対する回答に記載した内容に加えて実施している取組です。

A. 利用者のリテラシー向上に向けた啓発活動

Metaは、インターネットおよびソーシャルメディアの利用者のICTリテラシー向上のため、総務省が推進する官民連携プロジェクト「DIGITAL POSITIVE ACTION」に参加しています。同プロジェクトの取組には、官民の施策を集約したウェブサイトの運営、セミナーおよびシンポジウムの開催、教育・啓発用資材の制作、ならびに各種広告媒体を通じた普及啓発キャンペーンの実施が含まれます。Metaは、同プロジェクトの趣旨に賛同し、参加団体の一つとして参画しています。

B. 情報の信頼性を評価するための手段の利用者への提供

偽誤情報に対するMetaのアプローチは、平時から一貫して適用される次の3つの戦略を組み合わせたものです。(i) コミュニティ規定に違反するコンテンツを削除すること、(ii) 削除の基準には至らないものの問題のあるコンテンツについて、その拡散を抑制すること、(iii) 利用者が十分な情報に基づいて判断できるよう、ラベルの付与、信頼できる正確な情報への誘導、追加的な文脈情報の提供、およびリテラシー教育の機会の提供を通じて、利用者に情報を提供することです。また、Metaは、独立した専門家から、特定のコンテンツが虚偽であり、かつ差し迫った身体的危害を助長するおそれがあるとの指摘を受けた場合には、当該偽誤情報を削除します。

(略)

さらに、2024年9月、Metaは、IFCNの認証を受けた機関であるLitmusと連携し、日本において第三者によるファクトチェック・プログラムを開始しました。「虚偽(False)」、「改変(Altered)」または「一部虚偽(Partly False)」と評価されたコンテンツについては、拡散が抑制されることがあります。また、Metaは、当該コンテンツに対し、ファクトチェック機関による記事へのリンクを付した警告ラベルを付与し、利用者が追加的な文脈情報を確認できるようにしています。あわせて、当該コンテンツを共有しようとする利用者、または過去に共有した利用者に対し、当該情報がファクトチェック機関による評価を受けたことを通知します。さらに、ファクトチェック機関による評価を受けたコンテンツを含む広告は、これを拒否します。

C. 関係当局との恒常的な協力体制

Metaと総務省との間では、大規模地震等の大規模災害の発生時に、総務省から偽誤情報の拡散のおそれに関する注意喚起が共有され得る枠組みが設けられています。Metaは、こうした注意喚起が社内のインテグリティ部門に速やかに伝達される社内プロセスを整備しており、これにより、警戒水準を高め、迅速に対応することが可能となっています。また、Metaの第三者ファクトチェック・パートナーも、拡散傾向にあり、時宜性があり、かつ影響の大きい偽誤情報について能動的な監視を行っており、これには大規模災害に関するものが含まれます。

D. 更なる対話

Metaは、プラットフォームが災害時に生じ得る情報に関する課題に対して実効的に貢献し得る方途について、日本の防災関係機関との連携を含め、継続的に議論を行っていく用意があります。

Q2-2 その他

事業者様によっては、地震等の災害関連への対応として、日本にローカライズした独自の取組をされていると伺いました。災害に限らず、このように日本にローカライズしたお取組みがございましたらご教示ください。また、ローカライズに関する貴社のお考えをご教示ください。

【回答】

(Facebook、InstagramおよびThreadsに共通の回答)

1. 災害に関連する取組

Metaの危機対応ツールの機能の一つである「災害時の安否確認(Safety Check)」（災害時に有効化され、利用者が自らの無事を知らせ、つながりのある人々の安否を確認できる機能）は、2011年の東日本大震災を契機として開発されたものです。日本における最近の有効化事例としては、2026年7月28日の熊本県における地震および2026年8月の千葉県における大雨があります。

2. 日本向けにローカライズされた運用上の措置

[Metaの情報流通プラットフォーム対処法に基づく年次透明性レポート](#)に記載のとおり、Metaは、同法に基づく申出のための専用の受付窓口を設置しています。Metaの同法への対応に関する詳細な情報は、同レポートに記載されています。

3. 日本向けにローカライズされたりテラシーおよび啓発に関する取組

(1) 総務省「DIGITAL POSITIVE ACTION」への参加

Metaは、デジタルリテラシーに関する日本の中核的なマルチステークホルダー型官民連携プログラムである、総務省の「DIGITAL POSITIVE ACTION」に継続的に参加しています。直近では、2026年7月16日に開催された第4回会合において、2025年後半の前回会合で報告した内容を踏まえ、青少年の安全、偽誤情報対策、詐欺の防止、ならびに——とりわけスポーツの文脈に着目した——いじめおよび嫌がらせへの対応にわたる最新の取組を紹介しました。

(2) 「みんなのデジタル教室」(#WeThinkDigital)によるデジタルリテラシー教育

Metaは、対面型の教育ワークショップおよびオンラインキャンペーンを通じて、若年層および保護者双方のデジタルリテラシーの向上に投資しています。日本のNPO法人ACEとの連携

により実施している「#WeThinkDigital」プログラムを通じて、Metaは2020年以降、日本全国の中学生および高校生48,000名以上にデジタルリテラシー教育を提供してきました。同プログラムはその後、総務省との連携により「#WeThinkDigital 2.0」としてリニューアルされ、デジタルリテラシーに関するテーマに特化した新たなコースが追加されました。

(3) 日本の関係当局および市民社会との連携による啓発キャンペーン

Metaは、総務省、こども家庭庁、警察庁およびNPOと連携し、若年層およびその保護者に向けた啓発キャンペーンを継続的に実施しています。これらのキャンペーンでは、10代のアカウント(Teen Accounts)などの安全機能に関する周知にとどまらず、性的な画像を用いた脅迫(セクストーション)や違法な求人(いわゆる「闇バイト」)など、日本において特に顕在化している被害についても取り上げてきました。とりわけ違法な求人に関するキャンペーンについては、警察庁から直接提供を受けたキーワードを対応に取り入れました。

基準のグローバルな一貫性

Metaのコミュニティ規定は、世界中のすべての利用者に対し、Metaのすべてのプラットフォームにおいて等しく適用されます。コミュニティ規定は、一般に公開され、透明性が確保され、かつ継続的に更新される、グローバルに一貫した単一の枠組みを提供するものです。単一のグローバルな枠組みであることは、レビューアおよびシステムが同一の基準に従って運用されることを意味し、これにより不整合および誤りの低減に資します。

基準の一貫性は、運用の一律性を意味するものではありません。例えばMetaは、日本語および日本の文化に関する体制も維持しています。またMetaは、コミュニティ規定および関連情報を各市場において参照可能なものとするための取組を行っており(例えば、コミュニティ規定は日本語で公表されています)、さらに、啓発・教育活動およびパートナーシップについては、現地の事情に応じた形としています。

これまでの回答に記載したとおり、Metaは、日本において、利用者および利用者以外の方が情報流通プラットフォーム対処法に基づきコンテンツを申出できる専用の受付窓口を維持しています。この受付窓口およびMetaの日本におけるコンテンツに関するエンフォースメントのその他の側面に関する詳細な情報は、[Metaの情報流通プラットフォーム対処法に基づく年次透明性レポート](#)に記載されています。

Metaは、日本の関係当局およびステークホルダーとの継続的な対話を重視しています。

3 その他

ADR手続や事業者の判断に対する不服申立て等の事業者と利用者間の紛争処理の手続について、貴社の所感をご教示ください。

【回答】

Metaとしては、生じ得る紛争は、当初の判断を行ったプラットフォーム自身による速やかな再考によって最も実効的に解決されると考えています。社内における再審査は、最も迅速であり、利用者にとって無償であり、かつ、当該コンテンツおよび当初の評価の文脈に直接アクセスできる当事者によって行われるものです。

またMetaは、異議申立ての結果に関する透明性が、説明責任の中核をなすものと考えています。Metaは、異議申立ての件数、および違反類型ごとの判断の取消率について、グローバルなデータを公表しています。あわせて、異議申立ておよびコンテンツの復元に関する詳

細な情報を、情報流通プラットフォーム対処法に基づく年次透明性レポートにおいて公表しています。

Metaは、Facebook、InstagramおよびThreadsにおけるコミュニティ規定違反の種類の大部分について、異議申立ての機会を提供しています。例えば、次のとおりです。

- コンテンツがコミュニティ規定に違反しないと判断された場合、当該申出を行った個人または団体は、その判断についての再審査を求めることができます。
- コンテンツがコミュニティ規定に違反するとして措置の対象となった場合、当該コンテンツを投稿した利用者は、措置の理由について通知を受け、その判断についての再審査を求めることができます。

Metaは、異議申立ての結果を当事者に通知します。異議申立ての結果として、当初の判断が覆されることがあります。

利用者は、Metaの社内における異議申立ての審査手続を尽くした後、コンテンツに関するエンフォースメントの判断について [監督委員会 \(Oversight Board\)](#) に対して異議を申し立てることもできます。監督委員会は、Metaとは別個の独立した機関であり、Metaが行ったコンテンツに関するエンフォースメントの判断のうち選定されたものを審査するために設立されたものです。同委員会の使命は、Metaの当初のコンテンツに関する判断が、公表されているコミュニティ規定、Metaの基本的価値観および人権に関するコミットメントと整合しているか否かを検討することにあります。

- 独立した外部審査：監督委員会は、Metaのコンテンツに関する判断についての利用者からの異議申立て、またはMetaから直接付託された事案の中から、独自に事案を選定して審査します。事案が選定された場合、同委員会は、Metaの当初の判断を覆すか、または維持するかを決定します。事案に関する決定は次のページで公開されています。 <https://www.oversightboard.com/decision/>

監督委員会の個別事案に関する決定は、Metaに対して拘束力を有します。

- 拘束力のない政策勧告：監督委員会は、個別のコンテンツに関する拘束力のある決定に加えて、Metaのより広範なコンテンツモデレーションの実務について、拘束力のない勧告をMetaに対して発することがあります。これらの勧告は、個別事案において、または特定の論点に関する独立した政策勧告意見 (policy advisory opinion) の形で行われ得ます。Metaは、これらの勧告に対して公に回答することにコミットしています。同委員会の勧告、Metaの公開回答、および勧告への対応状況は、次のページでご覧いただけます。
<https://transparency.meta.com/oversight/oversight-board-recommendations/>
- 透明性に関する報告：Metaおよび監督委員会は、事案件数、決定の結果、および勧告に対するMetaの実施状況を詳述した透明性レポートも公表しています。これらのレポートは次のページでご覧いただけます。
<https://transparency.meta.com/oversight/meta-biannual-updates-on-the-oversight-board/> および <https://www.oversightboard.com/news/>