

## SNS 等におけるなりすまし型「偽広告」への対応に関するヒアリングシート

## 1 前提:対象サービス・サービス規模について

Q1 貴社又は貴社の関連会社が直近で把握している日本国内における平均月間アクティブユーザー数が 1,000 万人以上であるソーシャルネットワーキングサービスその他交流型のプラットフォームサービス(以下「SNS 等」という。)(他のサービスに付随して提供される SNS 等を除く。)のサービス名(以下「対象サービス」という。)について、御教示下さい。

- ◆ 以下の質問については、Q1で御回答頂いたサービス(以下「対象サービス」という。)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(実施予定時期を含む。)について、御回答内容に含めて記載下さい。なお、その際は、既に実施済の取組なのか、実施予定の取組なのかが明確になるよう記載下さい。

以下二つとなります。

Facebook 平均月間アクティブユーザー数 2600 万 (2019 年 3 月時点)

Instagram 平均月間アクティブアカウント数 3300 万 (2019 年 3 月時点)

## 2 広告出稿時の事前審査等について(総務省要請 記1関係)

## Q2-1

総務省要請の記1①を踏まえ、広告の事前審査基準を新たに策定した場合、次の3点を御教示下さい。  
なお、当該要請前に、既に当該基準を策定していた場合は、Q2-2を御回答下さい。

## (1) 広告の事前審査基準の名称及び当該基準の公表 URL

電話認証

この度、新規広告主を対象に電話認証を導入しました。新規広告主は、広告を公開する前に、広告アカウントに関連付けられた認証済みの電話番号が必要になる場合があります。この手順では、広告を公開する前に、Meta が SMS、音声、WhatsApp のいずれかの方法で送信したランダムに生成されたコードを確認することで、アカウント管理者が電話番号を認証することが求められます。電話認証は、広告主が当社のプラットフォームで広告を公開する前に、当社が広告主に対して実施している既存の措置の基盤を強化する追加要件です。詳細については、[こちら](#)をご覧ください。

Meta は、コミュニティの安全性を最適化するために、ポリシーと取り組みを継続的に見直し、改善していきます。この新しい対策の追加的な効果を評価し、その有効性を高める方法を模索していきます。

## (2) 広告の事前審査基準の内容の概要(当該基準により広告の掲載を認めない内容等)

当社の「[アイデンティティの表明における真正性](#)」ポリシーは、Facebook と Instagram の両方に適用されます。ユーザーは、他者に成りすましたり、アイデンティティを偽って他人を欺いたり、当社のポリシーに違反したり、ポリシーの施行を回避したりすることを禁じられています。

さらに、Meta において広告を展開する広告主は、当社の[コミュニティ規定](#)および[広告規定](#)に従う必要があります。Instagram の場合、[Instagram コミュニティガイドライン](#)にも従う必要があります。

コミュニティ規定では人々や企業を保護するための[詐欺行為](#)および[詐欺的手法](#)に関するポリシーも設けています。このポリシーに基づき、故意に他人を欺いたり、故意に虚偽の情報を伝えたり、あるいは金銭や財産をだまし取ったり、搾取したりするコンテンツは削除されます。

また当社の[広告規定](#)では、[許可されないビジネス手法](#)を含む、欺瞞的行為や誤解を招く行為を厳格に禁止しています。さらに、コミュニティの安全を守るため、それらの関連性についても調査しています。広告主は、[他の不正な事業資産に関連する事業資産を管理すること、すでに問題があるとして削除された事業資産と同様の行為](#)することを禁止されています。

詐欺の敵対的な性質を考慮し、弊社ではその対策のため、さらに何ができるかについて常に検討しています。弊社では、特に日本の消費者に対する詐欺に対抗するため、総務省による行政指導といった指導に感謝しています。

### (3) 広告の事前審査基準の見直しの時期

悪質な犯罪者は検知を逃れるために常に戦術を変えているため、私たちはこの問題に対するニーズや新たな課題を考慮しながら、常に取締りとポリシーの更新と見直しを行っています。

Q2-2総務省要請前に、既に広告の事前審査基準を策定していた場合、当該要請の記1①を踏まえた対応について、次の3点を御教示下さい。

#### (1) 広告の事前審査基準の名称及び当該基準の公表 URL

##### 広告審査

広告主に対する弊社の事前審査基準は広告審査システムに依拠しています。これは主に自動化されたツールを使用して広告をポリシーに照らしてチェックするもので、すべての広告が公開される前に審査を行うよう設計されています。ポリシー違反が検出された場合、広告が公開される前に削除されます。

広告審査システムは、ポリシー違反の広告を審査します。この審査プロセスでは、広告の特定の要素(画像、動画、テキスト、ターゲット情報など)や、広告に関連するランディングページやその他のリンク先など、さまざまな情報が審査の対象となります。

審査の大部分は自動化されていますが、これらのシステムの構築とトレーニングは当社のチームが担当しており、場合によっては広告を手動で審査することもあります。

審査結果に基づいて、広告の拒否または掲載許可が判断されます。広告が拒否された場合、広告主は新しい広告を作成することができます。新しい広告クリエイティブを使用するか、拒否された広告を修正していただきます。広告が誤って拒否されたと思われる場合は、再審査を依頼することもできます。広告審査システムに関する詳細情報は「[審査中の広告について](#)」をご参照ください。

個々の広告の審査以外にも、[広告主の行動\(過去に広告が拒否された回数や、広告審査を回避しようとした試みなど、違反の深刻度を含む\)](#)を審査および調査する場合があります。

## 認証:

広告主には、電話認証以外の追加の認証手続きが必要となる場合があります。一部のビジネスツールでは、アクセスに認証が必要となる場合があります。例えば、特定の製品にアクセスするにはビジネス認証が必要です。さらに、広告アカウントから何らかの被害が疑われる場合、お客様のビジネス資産が不正に利用された場合や、当社の広告規定に違反している疑いがある場合、Meta は認証を求める場合があります。

### (2) 既存の広告の事前審査基準の実効性を確認し、対応が不十分な点を踏まえた当該基準の改訂内容

悪質な行為者が常に検知を逃れるための戦術を変えているため、弊社ではポリシーに違反するアカウントを特定し調査するためのさまざまな方法を模索し続けています。

2-2-1 で説明されている広告審査システムの既存の基盤を基に構築された、弊社の新しい電話番号認証(上記 2-1-1 で説明)と強化された削除の取り組みは、新しい広告主と悪質な行為者のネットワークに焦点を当てることで、悪質な行為者への対策をより効果的に補完することを目的としています。

Meta は、コミュニティの安全性を最適化するためのポリシーと取り組みを継続的に見直し、改善していきます。この新しい対策の追加的な効果を評価し、その有効性を向上させる方法を引き続き模索していきます。

### (3) 広告の事前審査基準の見直しの時期

当社のポリシーと実施は絶え間なく繰り返され、常に進化しています。これは、悪意のある犯罪者が当社のシステムを回避するために展開する新たな手法に対応するためです。

## Q2-3

総務省要請の記1②を踏まえた対応について、次の6点を御教示下さい。

### (1) なりすまし型「偽広告」を端緒とした詐欺の手口・実態の把握方法

Meta は、このような詐欺広告に対処するために多大なリソースを投入しています。当社のアプローチは、次の4つの分野に重点を置いています。

- 1) 詐欺広告の一般的な特徴を理解するためのより詳細な分析と、施行および自動検出システムの改善
  - 当社は、詐欺広告の最新の傾向に関する理解を絶えず深め、その学習結果を検出テクノロジーと審査担当者の精度向上に役立てています。日本語と日本文化に精通した審査担当者は、これらの取り組みとプロセスに深く関与しています。
  - また、Meta の機械学習技術を活用して、さらに精度の高い検知モデルを構築し、自動検知システムを改善することで、取締りの強化を図っています。
  - すべての広告はシステムによって審査され、当社の取締りを回避しようとする広告が見つかった場合は、広告を拒否する以上の措置を取ります。
  - 著名人を装った広告は、乗っ取られたアカウントや偽アカウントから投稿されることがあります。当社のコミュニティ規定施行レポート(2024年第1四半期)で共有したように、当社の検知テクノロジーは、偽アカウント作成の試みを毎日6億3100万件ブロックし、作成後数分以内にさらに数百万件を検知しています。当社のシステムは、数千ものアカウント属性を調査し、当社のプラットフォーム上の他のユーザーとのつながりなど、悪意のあるユーザーが偽装しにくい行動の検知に重点的に取り組んでいます。
- 2) 著名人や著名人のイメージを悪用し、ユーザーを欺いて広告への関与を促す行為に対するより厳格な措置

- 悪質な行為者をより迅速に特定し、ポリシー違反の広告が実際に表示される前に防止するための手段を活用しています。例えば、疑わしい広告主である兆候が見られる場合や、ポリシー違反の履歴が見られる場合には、広告を拒否したり、ポリシーに違反する広告アカウントを制限または無効化し、将来的な広告掲載能力を削除するなどの措置を取ります。
  - 複数の Facebook または Instagram アカウントを連携させて不正行為を行うなど、組織的な不正行為や詐欺行為への参加や関与を企てる行為は許可されません。例えば、4月に実施した取り組みでは、さまざまな日本の著名人の画像を使用した詐欺アカウントやページのネットワークを削除しました。当社のチームは、これらの広告を作成するために使用されていた約100万の広告と5,000のFacebookアカウントを特定し、削除しました。
  - 悪意のある行為者のネットワーク全体を一挙に排除することは、現在進行中の自動化された施行システムや安全ツール、機能の補完となります。これらの取り組みを併せて行うことで、悪意のある行為者が活動を継続することを可能な限り困難にしています。
- 3) 現実社会での影響の確率、悪意のある行為者に対しては、法の下で責任を問う
- すでに、私たちはサイト上のインテグリティを確保するために多大なリソースを投入しており、これはアカウントやページ、広告の停止や削除にとどまらず、場合によっては、Metaのルールに違反した人物に対してさらなる法的措置を検討することを意味します。
  - 例えば、複数の国において、著名人を偽装したページにユーザーを欺いて訪問させ、連絡先情報を提供させたり、その情報を収集してオンラインマーケティング担当者や広告主に横流していた企業に対して、Facebook および Instagram の利用規約およびポリシーに違反したとして訴訟を起こしています。
- 4) ユーザーツール/透明性:
- 当社のアプリ上でユーザーの安全を確保できるよう、当社はユーザー向けの報告ツールと透明性の提供を継続しています。例えば、
  - なりすましアカウントの報告ツール: なりすましアカウントをアプリ内で報告できるよう、なりすましアカウントのプロフィールにある3つのドットをクリックして報告することが簡単にできるようになっています。
  - Instagram の偽装アカウントに関する警告: Instagram では、他者になりすましたアカウントなど、自動システムが不正と判断したアカウントを削除しています。しかし、犯罪者がすぐに不正にアカウントを使用するとは限らないため、現在、なりすましの疑いがあるアカウントがフォローリクエストを送った場合、警告を送信するテストを行っています。また、企業になりすました疑いがあるアカウントがダイレクトメッセージを送った場合にも、警告を送信します。
  - Instagram 認証バッジの可視性を向上: 認証バッジが Instagram 上のより多くの場所に表示されるように拡大することで、ユーザーがやりとりしているアカウントが本物で認証済みであることを確認しやすくします。
  - また、Instagram 上でもアカウントの透明性に関する情報を表示しています。アカウントのプロフィール、フィード投稿、ストーリーの「その他」3つのドットメニューのオプションで、ユーザー名をタップするとアクセスできる「このアカウントについて」機能で表示されます。アカウントの透明性に関する情報には、広告主全員に対して、アカウントが Instagram に参加した日付、以前にユーザー名を変更したことがある場合はその回数、アカウントの所在地となる国が表示されます。これにより、ユーザーは交流している相手についてより多くの情報を得ることができ、広告主と交流するかどうかを選択する際に、より確信を持って決定することができます。これは、Facebook ページにも適用され、ページとページを管理する人々に関するより多くの情報を提供しています。
  - また、Facebook と Instagram のすべてのアクティブな広告を広告ライブラリで公開しています。
  - 公開されている広告ライブラリは、Facebook のアカウントを持たない方でも、数十カ国で閲覧することができます。Facebook と Instagram のすべてのアクティブな広告が掲載されています。
  - 広告ライブラリから、ユーザーは広告を報告することができます。ページごとに検索して、そのページが現在掲載している広告を確認し、国別にフィルタリングすることもできます。

- 広告の報告、非表示、ブロックなど、フィード内およびライブラリ内の広告に対する否定的なフィードバックのシグナルを組み込んでいます。これにより、公開前にこれらの広告を自動的に検出するシステムが改善されます。
- 広告主に対するより厳格な認証
- 弊社は、アクセス、プライバシー、潜在的な被害への対処のバランスを取るリスクベースの認証アプローチを採用しています。この考え方にに基づき、新たに作成された広告アカウントの管理者は、広告を公開するために電話番号の認証が必要になる場合があります(2-1-1のセクションで説明されている通り)。詳細は[こちら](#)をご覧ください。
- 広告主は、電話番号の確認以外にも追加の確認手続きが必要になる場合があります。一部のビジネスツールでは、アクセスに確認が必要になる場合があります。例えば、特定の製品にアクセスするには[ビジネス認証](#)が必要になります。
- さらに、お客様の広告アカウントから何らかの被害が疑われる場合、お客様のビジネス資産が不正に使用された疑いや、当社の広告規定に違反している疑いがある場合には確認を求める場合があります。

## (2) (1)により把握した、なりすまし型「偽広告」を端緒とした詐欺の手口・実態の概要

私たちは、多くの場合、これらの広告は、[クローキング](#)と呼ばれる技術を使用して、広告審査システムとユーザーに異なるコンテンツを表示していることを知りました。

多くの広告には、[クリックを誘うような特徴](#)があり、広告をクリックすると、予期しないウェブサイトに誘導されることがあります(例えば、ニュース記事のように見える広告が、実際には美容製品を宣伝している場合など)。

多くの場合、これらの広告の一部は、直接的に詐欺的な広告にリンクしているわけではなく、最初の広告から詐欺的な広告にたどり着くには、数回のクリックが必要となっています。

しかし、この問題は新しいものではなく、オンライン広告の世界では長年悩みの種とされてきたものです。この問題は今後も継続されることが懸念されます。これらの詐欺広告を行う犯罪者はインターネット上で、さまざまな方法やチャネルを使って被害者を狙っていますが、弊社ではこの問題に取り組むために多額の投資を続けています。

(3) なりすまし型「偽広告」を端緒とした詐欺の手口・実態を踏まえ、当該手口・実態に応じた審査を実施するための対応の内容(例えば、「クローズドチャットを遷移先としたなりすまし型「偽広告」など特定の手口が多発している場合は、審査において取り扱わないこととしている」といった内容等)

※ 詐欺の手口・実態等を踏まえた広告の事前審査について、要請において例示した「クローズドチャットを遷移先として設定した広告を原則として取り扱わない」方法を採用しなかった場合には、その理由を併せて御教示下さい。

前述の通り、広告審査には自動システムと手動システムの両方を使用しています。

特に著名人を装った広告については、2-3-2で説明したように、広告で使用される戦術の変化について学んだことを自動的に組み込む追加の検出モデルを構築し、改善しています。

また、ユーザーに詐欺広告の疑いがある広告を当社に報告するようお願いして、調査を行い、適切な措置を講じています。

場合によっては、悪意のある業者がプライベートチャットに誘導するリンクを使用することがあり、当社ではそのリンクを監視することができず、それが詐欺であるかどうかを正確に判断できないことがあります。しかし、当社はこれま

でも、ユーザーからの報告や信頼のおけるパートナーからの報告など、さまざまなシグナルを基に、広告および広告の掲載ページを検証してきました。

強調すべきは、プライベートチャットに誘導する広告がすべて信頼できないわけではないということです。プライベートチャットに誘導する広告は、広告主や企業がユーザーとつながり、1対1の会話を通じてユーザーが興味を持っている製品やサービスについて話し合う方法として、業界で広く使用されています。

ビジネスメッセージングは、何百万もの企業の成長を支援しています。BCGとの2022における共同調査では、以下の点が強調されています。

- 少なくとも3人に1人のアジア太平洋地域の消費者が、週に1回以上企業とチャットしている
- 調査対象の大企業の70%が、ビジネスメッセージングを自社のビジネス全体にとって非常に重要または極めて重要と評価
- 50%以上の企業が、ビジネスメッセージングをアフターサービスに利用していると回答。一方、小売およびeコマース企業の61%が、販売前の問い合わせ処理に利用していると回答
- 平均して50%の企業が、メッセージングを利用して注文や取引を処理していると回答

この機能を広告に制限することは、活気のある日本のデジタル経済にとって機会損失となる可能性があります。言うまでもなく、この機能や仕組みは新聞広告やチラシ広告など、広告主がフォローアップのための1対1の会話用に連絡先(電話番号やメールアドレス)を共有する、他の伝統的な広告にも見られるものです。

(4) なりすまし型「偽広告」では、著名人の画像が用いられているものが多いところ、審査の中で、肖像使用の許諾を確認している場合は、その内容

当社のポリシーでは、広告に含まれるコンテンツを使用するには、必要な許可を取得することが求められています。広告審査プロセスの一環として、当社のポリシーおよび利用規約に従って、イメージに著名人が使用されている場合も含め、審査と適切な措置を実施しています。

(5) なりすまし型「偽広告」を端緒とした詐欺の手口・実態を踏まえた審査により、掲載を認めなかった広告の件数(件数の回答に当たっては、例えば、「クローズドチャットが遷移先として設定されていたため、●件の広告の掲載を認めなかった」など、可能な限り、理由とともに件数を御回答ください。)

前述の通り、悪質な犯罪者はしばしば組織化され、意欲的で、資金力も豊富です。そのため、悪質な犯罪者が弊社のプラットフォームや広告審査システムを悪用して被害をもたらそうとした場合、そのネットワークを停止します。最近では、日本に影響を及ぼすネットワークに対してより絞り込んだ対策を講じるため、日本を対象とする削除を行いました。

この取り組みにより、約100万の広告と、それらの広告を作成するために使用されていた5,000件のFacebookアカウントを特定し、削除しました。

削除は、現在運用中の自動化された対応策や安全ツールの補完となります。これらの取り組みを併せて行うことで、悪意のある犯罪者の活動を可能な限り妨げます。

このように当社が実施している複数の対策を組み合わせることで、2024年3月5日から2024年6月1日までの間に、日本を標的とした約5,275,000件の詐欺広告と5,400件の広告アカウントを削除しました。

(6) 今後、なりすまし型「偽広告」を端緒とした詐欺の手口・実態を踏まえ、審査方法の見直しを行う際の契機や基本的な考え方

## Same response with Section 2-3-1

Q2-3の(1)を参照ください

Q2-4 総務省要請の記1③を踏まえた対応について、次の2点を御教示下さい。

(1) 総務省要請の記1④の対応を確実に実施するための事前審査体制の整備状況(日本語並びに日本の社会、文化及び法令を理解する者の配置状況やAI等技術の利用状況など)とその公表 URL

弊社では、ポリシー違反の特定に多角的なアプローチを採用しており、AI 検出テクノロジー、人間のレビュープロセス、ユーザーからの報告を組み合わせ活用しています。

人工知能(AI)検出テクノロジー:

弊社では、機械学習とメディア照合技術を活用して、ポリシー違反の可能性のあるコンテンツを積極的に特定しています。これらの技術により、コンテンツの審査業務の範囲を拡大し、可能な限り自動的に判断を下すことが可能になります。現在、弊社ではポリシー違反のコンテンツを特定し削除するにあたり、主に AI に依存しており、多くの場合、コンテンツが誰かの目に触れる前に処理されています。さらに、AI により、コンテンツ審査担当者が優先的に審査すべき案件を特定することができ、最も有害で緊急性の高いコンテンツを優先的に処理することが可能になります。

人間の目によるレビュー:

特定のケースでは、人間の目によるレビューが不可欠です。例えば、投稿がポリシーに違反しているかどうかを当社のテクノロジーが判断できないような場合です。この人間の目によるレビューチームには日本人・日本語の専門レビュアーも数多く在籍しています。

(2) 上記(1)の体制において、なりすまし型「偽広告」など問題のある特定の種類の広告が今以上に増加した場合にも適正な対応が可能となるよう実施している工夫(例えば、対応人員を迅速に増加させられるような人事制度上の工夫、AI等技術の迅速なチューニングを可能にするような技術上の工夫など)があれば、その具体的内容

広告詐欺対策は、多くのチームにまたがる取り組みです。認証メカニズムを設計するセキュリティエンジニア、脅威の主体を追跡する脅威インテリジェンスチーム、不正なアカウントやコンテンツを検出するために機械学習を調整し使用するインテグリティチーム、問題の改善を支援するサポートチームなどです。私たちは、変化する戦術から学んだことを取り入れ、執行の改善に役立てるために、新しいアプローチをテストし導入することで、既存の施行システムの強化を常に繰り返しています。

当社のレビューチームはグローバルに 24 時間 365 日体制で対応しています。世界中に 20 以上のサイトを展開しており、これらのチームは 80 以上の言語でコンテンツをレビューすることができます。当社のグローバルな体制により、特定の国で需要が急増した場合、あるいは危機的状況が発生した場合や予測不可能な事態が発生した場合でも、迅速にリソースを再配置することができます。このような事態が発生した場合、当社のリソースを適切に適応させることが可能となります。

Meta は、重大な出来事が発生した際の差し迫った被害のリスクを常に調査して、人々の安全を確保するための、期間限定の的を絞ったポリシーや製品に関する対応策を講じることもできるようにしています。このアプローチには、地域、地域、国家レベルの出来事の前後における潜在的なリスクの調査が含まれます。

Q2-5

総務省要請の記1④を踏まえた対応について、次の3点を御教示下さい。

(1) 当該要請以降に、なりすまし型「偽広告」によってなりすまされた被害者(その法定代理人を含む。以下同じ。)から通報があった場合における当該被害者の人数

このデータの詳細は取得しておりません。

(2) 当該通報以降、なりすまし型「偽広告」の事前審査で強化した内容

なりすましによる虚偽広告詐欺が急増していることを受け、私たちはタスクフォースを結成し、著名人の広告を積極的に監視し、確認と対応を行う体制を整えました。

私たちはまた、当社のポリシーを更新しました。

コミュニティ規定において、複数の Facebook または Instagram アカウントを連携して使用するなど、組織的な詐欺や詐欺行為への関与を企てる、または関与していると偽る行為を禁止しています。さらに、コミュニティの安全性を保護するために、関連するネットワークも調査しています。広告主は、他の不正なビジネス資産に繋がるビジネス資産を管理したり、過去に削除されたビジネス資産と同様の行為を禁止しています。また、信頼のおけるパートナーから報告された著名人を装った広告に関連する違反については、ビジネス資産(広告アカウント、ユーザー、ページ)に対して厳しい処罰を科しています。

さらに、検出と実施をより効率的に行うために、さまざまなツールの適用を開始しました。

日本に影響を与えるネットワークに対して、よりの絞った対策を行うために、削除を行いました。例えば、4月に実施した作戦では、さまざまな日本の著名人の肖像を無断で広告に使用した、日本を標的とした詐欺アカウントとページのネットワークを削除しました。当社のチームは、広告に使用されていた約 100 万の広告と 5,000 件の Facebook アカウントを特定し、削除しました。

また、既知の詐欺広告で使用されたキーワードや画像に基づいて、詐欺広告を自動的に検出して削除するシステムも導入しました。これらのツールにより自動検出と削除が可能になるため、この問題に対処する上で役立つ自動化ツールが改善されています。

今年から日本に関連するものも含めた新たな詐欺、被害、戦術を組み込むことで、AI 対応の機械学習による検出および施行モデルを改善しています。

違反が確認されたアカウントや広告に関する情報を活用し、積極的な検知と削除の取り組みを拡大することも行っています。違反広告に関する情報を自動的に機械学習システムにフィードバックすること新たな攻撃パターンの積極的な検知を強化します。

詐欺対策の継続的な取り組みの一環として、最近、新規広告主の確認と削除の強化を実施しました。詳細は 2-1-1 を参照してください。

(3) 当該通報以降、当該被害者に関するなりすまし型「偽広告」として掲載しなかった広告の件数

このデータの詳細は取得しておりません。

Q2-6

総務省要請の記1⑤を踏まえた対応について、次の3点を御教示下さい。



(1) 総務省要請前に実施していた事前審査における広告主(出稿する広告の内容を最終的に決定する権限を有する主体をいう。以下同じ。)の本人確認(本人確認書類の確認等を通じた身元確認と、当該身元確認済みの本人による広告出稿であることについての本人認証の双方を含む。以下同じ。)の方法(広告主の属性や出稿内容により、通常よりも厳格な本人確認を行う場合があったときは、当該厳格な本人確認の方法及びその適用条件を含む。また、身元確認の際に本人確認書類の提出を求めている場合は、その内容を含む。)なお、広告主以外の第三者(広告代理店など)による広告出稿に関し、広告主の本人確認とは別途、当該第三者の本人確認を行っていたときは、当該第三者の本人確認の方法もご回答ください。また、広告主の本人確認を当該第三者に委ね、貴社として独自に広告主の本人確認を行っていなかった場合は、当該第三者による広告主の本人確認が十分であることの確認方法を御回答ください。

広告主の認証は、詐欺が非常に敵対的な性質をもつことを踏まえると、当社のプラットフォームにおける不正行為や詐欺行為を防止するための万能策ではありません。しかし、Meta はリスクに基づく認証アプローチを採用し、広告がプラットフォーム上で承認されて実行される前に、自動化されたプロセスと人間によるレビュープロセスの組み合わせにより広告を審査しています。

これは、ユーザー教育キャンペーンの実施や、他企業、法執行機関、その他の利害関係者との協力・提携など、当社のプラットフォームにおける詐欺のリスクを軽減するための数ある対策のひとつです。このアプローチは、アクセス、プライバシー、潜在的な被害への対処のバランスを考慮したものです。

犯罪者は常にその手口を変化させてシステムを回避しようとするため、より強固なセキュリティとインテグリティの要件を課すことで、より効果的に犯罪者に対処することができます。弊社は、疑わしい行為の検出と正当なアクセス試行の認証を支援するために、さまざまなシグナルと検証課題を使用しています。この原則に基づき、弊社は以下の方法で検証を行います。

#### 新規広告主に対する検証要件

新規広告主は、広告を公開する前に、広告アカウントに関連付けられた確認済みの電話番号を所有していることが求められる場合があります。必要な手順では、広告を公開する前に、Meta が SMS、音声、WhatsApp のいずれかの方法で送信したランダムに生成されたコードを確認することで、アカウント管理者が電話番号の確認を行うよう求められます。新規広告主とは、(a) 既存の広告アカウントを持たない、または広告アカウントを開設してから 90 日未満で広告費を一切使用していない広告主、および (b) 管理されたビジネスアカウントまたは以前に確認済みのビジネスアカウントに関連付けられていない広告主を指します。

電話番号認証は、広告主が当社のプラットフォームに広告を掲載する前に実施している既存の措置を補完する追加要件です。

Meta は、コミュニティの安全性を最大限に確保するために、ポリシーおよび要件の見直しと改善を継続的に実施しています。当社は、この新しい措置の追加的な利点を評価し、その有効性を向上させる方法を引き続き模索していきます。詳細については、[こちら](#)をご覧ください。

#### 追加の認証手順

広告主には、電話認証に加えて、追加の認証手続きを行っていただく場合があります。追加の認証手続きには、以下のものが含まれます。

#### 広告カテゴリーまたはターゲットパラメータ

Meta の [広告規定](#) に従い、一部の広告コンテンツが制限され、関連する規制当局からのライセンスが必要となる場合があります。その場合、広告を掲載する前に、Meta が広告主の権限を認証します。さらに、Meta は広告主の情報を確認する場合があります。

社会問題に関する広告、選挙または政治に関する広告については、すべての広告主が本人確認を含む承認プロセスを完了する必要があります。

2段階認証をオンにしているかどうか

Meta はすべての広告主がアカウントの2段階認証および [Meta Business Manager](#) での2段階認証を有効にして、不正アクセスから保護することをお勧めします。

使用する広告機能の種類

一部のビジネスツールでは、アクセスに認証が必要とされる場合があります。たとえば、特定の製品にアクセスするには[ビジネス認証](#)が必要です。

そして、お客様の広告アカウントに何らかの問題が発生している可能性があるとして判断された場合、お客様のビジネス資産が危険にさらされておらず、弊社の広告規定に違反していないことを確認するために、Meta が認証を求める場合があります。

確認を求められる可能性がある事項

確認要件は状況によって異なりますが、一般的には以下のカテゴリーのいずれか、または複数に該当します。

所在地、電話番号、メールアドレス、納税者番号などの属性を含む事業の詳細

ライセンス番号などの規制当局からの承認

広告主の個人識別情報

事業またはその資産と貴社との関係

(2) 総務省要請後、上記(1)の方法やその実効性を検証し、対応が不十分な点を踏まえた改善を行った場合、当該改善の具体的内容

コミュニティの安全を維持するために、ビジネスや身元の詳細を確認するためのポリシーや要件を含む規定を設けています。新規の広告主は、広告を公開する前に、広告アカウントに関連付けられた確認済みの電話番号を所有していることが求められる場合があります。これには、広告を公開する前に、Meta が SMS または音声通話で送信したコードをアカウント管理者が確認し、電話番号を認証することが含まれます。これらの取り組みは、ビジネスと人々の双方に利益をもたらすために、Meta の広告プラットフォーム内の透明性と説明責任を高めることを目的としています。

(3) 上記(2)の改善の結果、改善後に掲載を認めなかった広告のうち、改善前の本人確認方法であれば掲載を認めていたと思われるものの件数

電話認証ツールが最近になって導入されたばかりであるため、現時点では共有できるデータがございません。

3 なりすまし型「偽広告」の削除等について(総務省要請 記2関係)

総務省要請の記2を踏まえた対応について、次の点を御教示下さい。

Q3-1 自らが提供する SNS 等において流通するおけるなりすまし型「偽広告」に関する、利用規約等を踏まえた対応状況

私たちは、[コミュニティ規定](#)、[広告規定](#)、[Instagram コミュニティガイドライン](#)に違反するコンテンツを、発見次第削除しています。

一部の問題のあるコンテンツは、Facebook や Instagram を利用する人々にネガティブな体験をもたらす可能性があります。私たちは、ポリシーで定める削除規定に完全に合致していない場合でも、このようなコンテンツの配信を減少させることがあります。

さらに、詐欺に関連する特定の検索語句の使用をブロックしています。これらの措置は、問題のあるコンテンツや低品質なコンテンツの配信を制限する詳細を規定した[コンテンツ配信ガイドライン](#)に沿ったものです。

広告審査システムを回避して弊社のプラットフォームを悪用し、被害をもたらそうとする悪意のある行為者のネットワークを調査し、排除しています。

悪質な犯罪者は、よく組織化され、高い動機を持ち、資金力も備えていることがよくあります。従って、当社のプラットフォームを悪用し、広告審査システムを回避して被害をもたらそうとする悪質な犯罪者のネットワークを調査、排除することに、現在より多くの労力を費やしています。例えば、4月に実施した取り組みでは、さまざまな日本の著名人の肖像を無断で広告に使用し、日本をターゲットとした詐欺アカウントとページのネットワークを排除する削除を実施し、約 100 万の広告と、それらの広告を作成するために使用されていた 5,000 件の Facebook アカウントを特定し、削除しました。悪意のある行為者のネットワーク全体を一斉に停止させることは、すでに実施中の自動化された施行システムや安全ツール、機能の向上につながります。これらの取り組みを併せて行うことで、悪意のある犯罪者が活動を継続することが困難になるようにしています。

これらの当社が実施している複数の対策を組み合わせにより、2024 年 3 月 5 日から 2024 年 6 月 1 日までの間に、日本を標的とした約 527 万 5 千件の詐欺広告と約 5400 件の広告アカウントを削除しました。

Q3-2 なりすまされた被害者からの削除申出を受け付ける方法の具体的内容及び当該方法を公表しているページの URL

Facebook および Instagram における偽アカウントの削除に全力で取り組んでいます。当社のアプリを通じて、なりすましアカウントを作成している場合は削除します。[なりすましアカウントの報告方法](#)については、Facebook 用は[こちら](#)、Instagram 用は[こちら](#)をご覧ください

また、詐欺的なコンテンツを報告するための専用窓口を政府および法執行機関向けに用意しています。日本では、警察庁および金融庁と緊密に連携しています。報告されたコンテンツは、当社のポリシーに従って確認し、適切な措置を講じます。

Q3-3 なりすまし型「偽広告」について、なりすまされた被害者から削除申出を受けた際の対応状況（申出を受けてから判断を行うまでの期間や、判断を行った申出者に対する通知の実施状況を含む。）

私たちは、ユーザーからの報告(なりすましによる虚偽広告など)を主に機械学習を用いてポリシーに照らし合わせています。また、当社のテクノロジーは、被害を可能な限り防ぐために、悪質性の高いコンテンツや、急速に広がり多くのユーザーに影響を及ぼす可能性のある拡散性の高いコンテンツを優先的に特定するのにも利用されています。

ほとんどの投稿やコメントは、投稿された時点、閲覧数が増加した時点、または報告された時点のいずれかで、複数の機械学習によるクラシファイヤによって事前に分類されます。

投稿された時点または閲覧数が増加する時点でクラシファイヤが対応しなかった投稿やコメント、およびその後報告された投稿やコメントについては、クラシファイヤが、その報告がコミュニティ基準に違反しているか否かを十分に判断できるかどうかを評価し、その後、対応します。

判別が難しいと分類された報告については、同じくクラシファイヤによってランク付けされ、違反の可能性と予測される悪質性の度合いを考慮した上で、人間の手によって確認されます。クラシファイヤと人間のハイブリッドシステムは、現在のところ最も効果的であることが分かっていますが、クラシフィケーション技術の向上に伴い、将来的には変更される可能性もあります。

私たちは、コンテンツの削除精度の向上や、報告を行うユーザーとのより良いコミュニケーションなど、ユーザーからの報告への対応を改善する方法を引き続き検討してまいります。

また、コンテンツを報告した後にユーザーが受け取るメッセージを更新し、審査状況をより正確に反映させるよう取り組んでいます。例えば、実際にはそうではないのに、その報告に否定的な見解を表明したかのように誤解される可能性を減らすため、コンテンツを報告した日本のユーザーへの対応に使用される文言の修正を検討しています。

Q3-4 削除等の実施に関する基準のうち、なりすまし型「偽広告」に係る基準及び当該基準を公表しているページの URL

前述の通り、Facebook と Instagram の両方に適用される「[アイデンティティの表明における真正性](#)」ポリシーでは、ユーザーが他人になりすましたり、本人確認情報を偽って他人を欺いたりすることを禁じています。

さらに、広告主が Meta の技術を使用して広告を配信する場合は、[Facebook のコミュニティ規定](#)および[広告規定](#)に従う必要があります。Instagram の広告主は、[Instagram のコミュニティガイドライン](#)にも従う必要があります。

コミュニティ規定に基づき、弊社では、[詐欺行為および詐欺的手法](#)に関するポリシーを定めています。このポリシーに基づき、故意に他人を欺いたり、故意に虚偽の情報を流したり、あるいは金品をだまし取ったり、搾取したりするコンテンツは削除されます。これには、弊社のサービスを利用して、これらの行為を調整したり、宣伝したりするコンテンツも対象となります。また、当社の[広告規定](#)では、不正行為や誤解を招く行為を厳格に禁止しており、[許可されないビジネス手法](#)も含まれます。さらに、コミュニティの安全を守るため、関連するつながりについても調査しています。広告主は、[他の不正な事業資産に関連する事業資産を管理したり、当社がすでに削除した事業資産と同様の行為をしたりしてはなりません。](#)

Q3-5 なりすまし型「偽広告」に係る削除の申出件数及び実施件数

例えば、4月に実施した取り組みでは、さまざまな日本の著名人の肖像を無断で広告に使用した、日本をターゲットにした詐欺アカウントとページのネットワークを削除しました。このチームは、100万近くの広告と、その広告を作成するために使用された5,000件のFacebookアカウントを特定し、削除しました。悪意のある犯罪者たちのネットワーク全体を一挙に排除することは、現在進行中の自動化された施行システムと、Facebookの安全ツールや機能の強化につながります。これらの取り組みを併せて実施することで、悪意のある犯罪者による活動の継続を可能な限り困難にしています。

全体として、当社が実施している複数の対策を組み合わせることで、日本を標的とした約 5,275,000 件の詐欺広告と 5,400 件の広告アカウントを削除しました。より詳細な情報は[こちら](#)をご覧ください。  
世界全体では、2024 年第 1 四半期に 6 億 3,090 万、第 2 四半期に 12 億の偽アカウントをプロアクティブに削除しました。プロアクティブな削除率は、第 1 四半期が 99.4%、第 2 四半期が 99.7%です。  
また、スパムコンテンツのプロアクティブな削除数は、第 1 四半期が 436,000,000 件、第 2 四半期が 322,300,000 件で、それぞれ 98.2%、97.4%の削除率でした。より詳細な情報は[こちら](#)をご覧ください。

Q3-6 なりすまし型「偽広告」に係るアカウント停止の申出件数及び実施件数

3-5を参照ください

Q3-7 削除等の対応に当たる人的・技術的体制(特に、削除等の対応に当たる人材のうち日本語を理解する者の人数)

ポリシー違反の特定にあたっては、AI 検知テクノロジー、人間の目による確認、ユーザーからの報告を組み合わせで使用しています。

AI 検出テクノロジー:

機械学習とメディア照合テクノロジーを導入し、ポリシー違反の可能性のあるコンテンツを積極的に検出しています。これらのテクノロジーにより、コンテンツの審査作業の規模を拡大し、可能な場合は自動的に判断を下すことができます。現在、主に AI を利用して違反コンテンツを特定し、削除しています。多くの場合、誰かが目にする前に削除されます。また、AI は、最も有害で緊急性の高いコンテンツを優先的に処理できるよう、コンテンツ審査担当者が案件を優先的に処理するのにも役立っています。

人間による確認:

人の目による確認が不可欠なケースでは、人的な確認作業が行われます。例えば、当社のテクノロジーでは違反投稿かどうかを判断できないような重大な状況下では、人の目による確認が不可欠です。  
現在、世界中で 15,000 人以上のコンテンツ確認者が、80 以上の言語で 24 時間 365 日体制でコンテンツの確認を行っています。この人的確認チームの一員として、当社には日本人や日本語を理解する担当者も配置されています。

4 その他

Q4-1 以上の御回答のほか、なりすまし型「偽広告」の流通を防止するために、貴社が実施している取組や今後実施予定の取組等(例えば、詐欺と疑わしいリンクが表示されている場合は、ポップアップで警告を表示させるようにした等)がございましたら、その内容をご回答ください。

以上