



# 誹謗中傷対策に関するInstagramの取組み

小俣栄一郎

Facebook Japan株式会社

公共政策本部

「コミュニティづくりを応援し、  
人と人がより身近になる世界を実現する」

FacebookとInstagramを安全でポジティブな自己表現の場として保つことは最優先事項。  
専門家、保護者等と協力しながら、いじめや嫌がらせを防止するための新しい技術を開発し、  
業界をリードしていきたい。

いじめや嫌がらせは非常に個別・属人的なものであり、文脈に左右されることも多い。  
プロアクティブな検出技術に加えて利用者からの報告にも頼っている。  
訓練された審査チームが24時間体制で対応。

# 機能紹介

# いじめ防止や誹謗中傷に対する機能

## 16歳未満の新規アカウントをデフォルトで非公開に設定

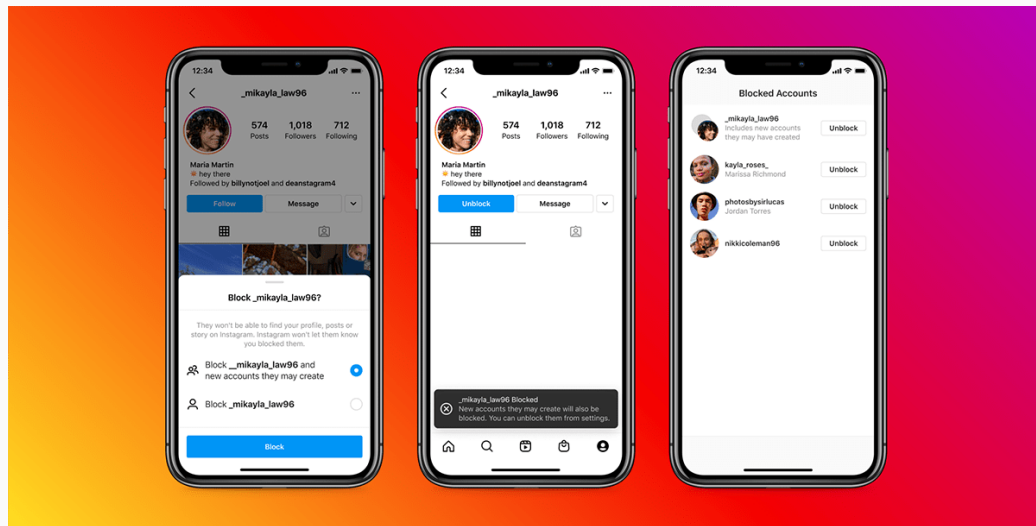
既にアカウントを公開に設定している若者層の利用者には、非公開アカウントのメリット及び公開から非公開アカウントに切り替える方法を説明する通知を表示



# いじめ防止や誹謗中傷に対する機能

## ブロック

すでにブロックした相手が新しいアカウントを通じて連絡しようとすることも困難に。他の利用者をブロックする際、現在のアカウントだけでなく、同じ相手が作成した可能性がある新しいアカウントも先手を打ってブロック可能。



# いじめ防止や誹謗中傷に対する機能

## ミュート

フォローを外すのがためられるアカウントがある場合、ミュートにすると、そのアカウントの投稿やストーリーズがフィードに表示されなくなる。

## 望まないやりとりからアカウントを保護する「制限」機能

Instagram上で相手がどのような態度をとっているのか把握しながらも、相手に知られることなく望まないやり取りからアカウントを守る ことが可能に





# いじめ防止や誹謗中傷に対する機能

## メッセージコントロール

誰からメッセージを受け取るのかをコントロールすることが可能。



# いじめ防止や誹謗中傷に対する機能

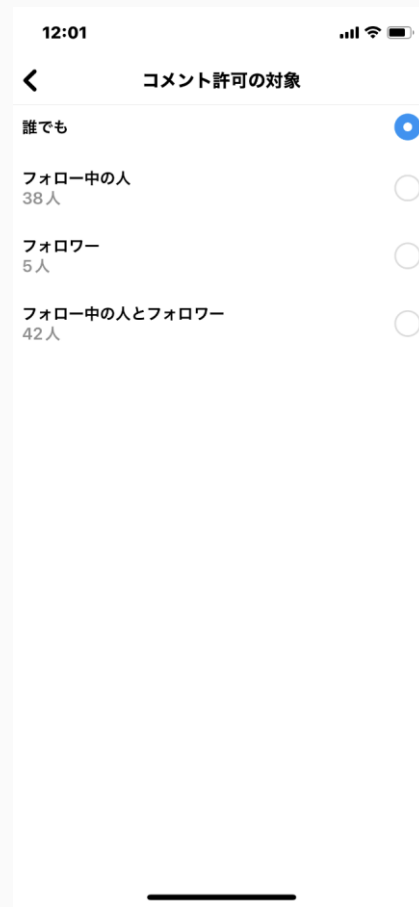
## コメントコントロール機能

フィード投稿にコメントできる対象を管理できる機能を追加。プロフィール画面の歯車アイコンをタップし、「プライバシー設定」下の[コメント]を選択すると設定を変更可。

公開アカウントの利用者は「誰でも」「フォロー中の人とフォロワー」「フォロー中の人」「フォロワー」のいずれかを選び、自身のフィード投稿にコメントできる対象を管理可能。

フィード投稿へのコメントをブロックするアカウントも指定可能。

非公開アカウントの利用者も同様に、フィード投稿へのコメントをブロックするアカウントを指定可能。



# いじめ防止や誹謗中傷に対する機能

## いじめコメントを自動でフィルターし、機械学習によるコメントフィルターを改善、いじめコメントを自動的に非表示に

いじめコメントを自動的にブロックし、非表示にするフィルターを導入。コメントフィルターは、機械学習（マシンラーニング）を応用し、利用者が嫌がらせやハラスメントを目的としたコメントを目にする必要がないよう、自動的にブロック。



# いじめ防止や誹謗中傷に対する機能

## ポジティブなコメントを固定してハイライト

ネガティブなコメントの削除だけでなく、ポジティブなやり取りを広めたり、促したりすることも容易にできるようにしたいと考えています。この機能により、選択したコメントをコメントスレッドのトップに固定することで、利用者がアカウントの雰囲気作りやコミュニティとの交流を行うことが可能に。



# いじめ防止や誹謗中傷に対する機能

## タグ付けとメンションを許可する相手を選択

タグ付けやメンションの機能を使用して、他人を標的にしたいじめが行われる可能性があることがわかったため、Instagram上で利用者がタグ付けやメンションを許可する相手を管理できる新しい管理機能を追加。

コメント、キャプション、ストーリーズでのタグ付けやメンションを許可する人の設定を、[すべての人]、[フォロー中の人のみ]、[誰にも許可しない]のいずれかから選択可能に。



# いじめ防止や誹謗中傷に対する機能

## 誹謗中傷目的のコメントやDMから利用者を守る「抑制」機能

自身をフォローしていないアカウントや、最近フォローしたばかりのアカウントから届くコメントを一時的に非表示にしたり、DMを「リクエスト」に自動で分類することが可能に。



## 不適切なコンテンツの報告

いじめや不快なコンテンツなど、Instagramに存在するべきではないものを見つけた利用者は、それをInstagramに報告可。Instagramはコンテンツを24時間体制で速やかに審査し、「コミュニティガイドライン」に違反しているコンテンツを削除します。Instagramでの報告はいつでも匿名ですので、報告されたコンテンツの投稿者には報告者が誰なのかがわかりません。報告を送信する方法については、Instagramのヘルプセンターを参照。

以下のいずれかの理由に該当する投稿を報告することができます。

- スпам
- ヌードや性的行為に関するコンテンツ
- ヘイトスピーチや差別的なシンボル
- 人種差別的な言葉または行為
- 暴力または危険な団体
- いじめや嫌がらせ
- 違法または規制対象商品の販売
- 知的財産権の侵害
- 自殺や自傷行為

キャンペーン



# キャンペーン

## 安心安全に使うための24種類のチュートリアル動画を公開

2022年2月より、Instagramを安心安全に使うための24種類の機能について、詳しい設定方法を紹介するチュートリアル動画を「みんなのデジタル教室」サイト内のリソースページで公開。

いじめや嫌がらせを防ぐための機能、アカウントのセキュリティを強化して乗っ取りなどのトラブルを避けるための機能、投稿やアカウントを報告してポジティブなコミュニティを作るための機能などを紹介。



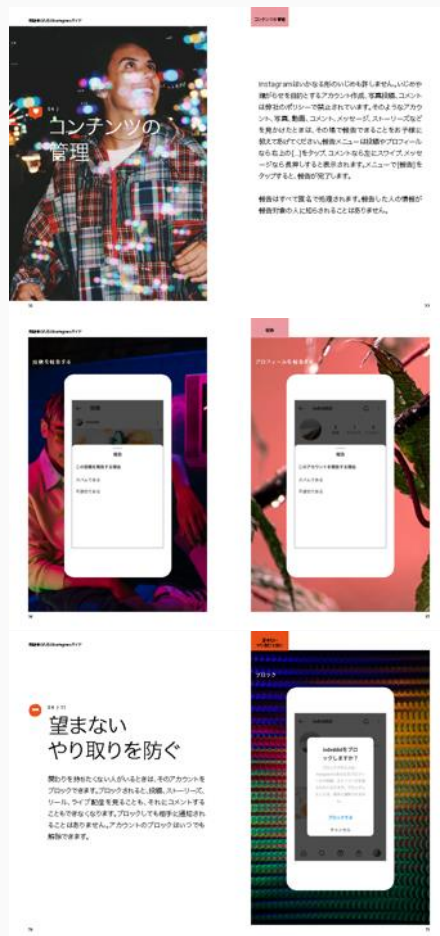
# キャンペーン

## 「保護者のためのInstagramガイド」

Instagramを利用している子どもを持つ保護者に向け、「保護者のためのInstagramガイド」を公開。2018年8月に本ガイドの初版をローンチし、2021年3月に機能のアップデートを反映、セキュリティやサポート情報を新規追加した第二版に一新。本ガイドは、6つの国内NPOの協力を得て作成し多くの画像を用いて分かりやすく紹介。

協力団体：

認定NPO法人カタリバ、特定非営利活動法人キッズドア、特定非営利活動法人ストップいじめ！ナビ、認定NPO法人 3keys、認定NPO法人育て上げネット、TDU・雫穿大学



# キャンペーン

## クリエイターと一緒に考える「#インスタANZENカイギ」プロジェクト

クリエイターと彼らのファンである利用者が、Instagramの安心安全な使い方について一緒に考える場を提供することを目的に2020年から「#インスタANZENカイギ」プロジェクトを推進。

2022年2月8日には、オンラインの安全に関する教育・啓発活動が世界中で展開されるセーフターインターネットデーに合わせ、若年層に人気のクリエイター5名が「リールで知ろう！インスタANZENヒント」を公開。



# キャンペーン

## 「SAVE YOUR LOVE 好きっていいね。」

Instagramでは、利用者が自分の”LOVE（好き）”を安心して楽しむことができる環境づくりに注力し、様々な機能の開発や改良を日々努力。特にSNSの使い方について悩みを抱えていることが多いが若年層の利用者に対して、日頃からSNSとの付き合い方について積極的に発信している藤田ニコルさん、EXITりんたろー。さん、青山テルマさんが出演するキャンペーンを展開。

2021年12月に、渋谷駅構内にて特設ブースでの限定ステッカーを配布などを通じて、安心安全なInstagramの使い方について呼びかけ。

