

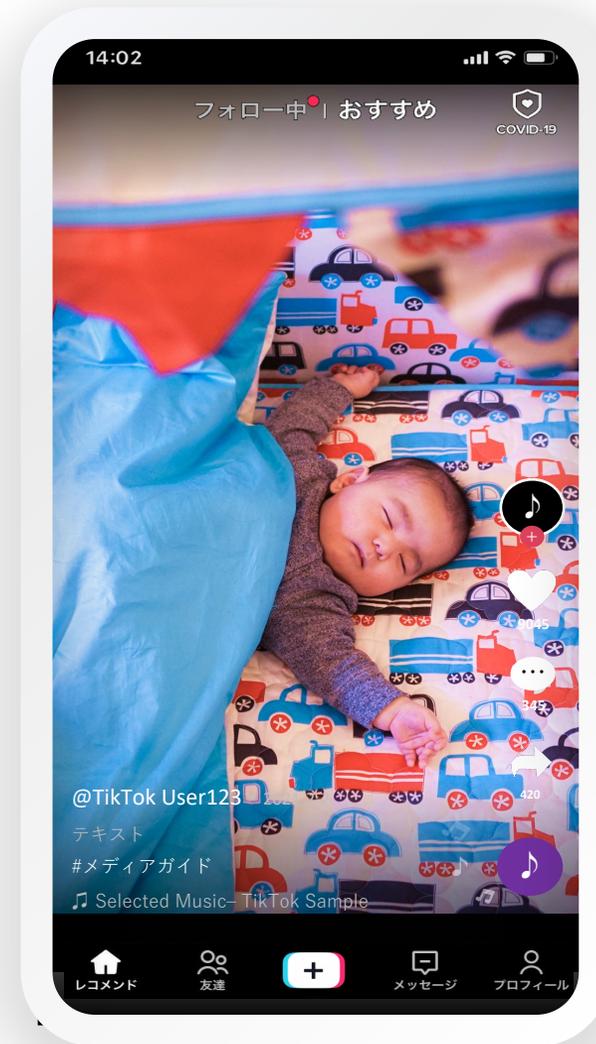
 TikTok

偽・誤情報防止の取り組み

TikTok Japan

TikTokとは？

- モバイル向けショートムービープラットフォーム
- 150の国・地域で展開
- ユーザーが投稿する動画コンテンツは多様化傾向にあり、ライブ配信も増加
- 大学・美術館、地方自治体や報道機関などの情報発信にも活用されている



TikTokのルール

コミュニティガイドライン

コミュニティガイドライン

- コミュニティに参加するユーザーが守るべきルール、モデレーションの基準を明示
- プラットフォームも遵守するもの

動画の審査・削除

- **すべての動画を24時間365日体制**で審査し、コミュニティガイドラインに違反するコンテンツは削除・アカウント停止等の対応
- 生命身体の危険がある場合、関連当局に連絡

コミュニティガイドライン

- はじめに
- コミュニティ原則
- 未成年者の安全とウェルビーイング
- 安全性と礼節
- 精神的健康と行動の健康
- 慎重に扱うべきテーマと成人向けテーマ
- 誠実性と信頼性
- 規制対象品と商業活動
- プライバシーとセキュリティ
- おすすめフィードの適確性基準
- アカウントと機能
- 執行

コンテンツ違反



コミュニティガイドラインに違反しているため、動画は削除されました

動画の詳細
06-26 12:25の投稿



コンテンツ違反

コミュニティガイドライン

誤りでしたか？ 異議申し立てを送信

偽・誤情報の審査基準

コミュニティガイドラインにおいて、削除等の対象となる偽・誤情報について定めています。

(1) 禁止されるもの

- 発信者の意図に関わらず
- 個人や社会に重大な危害を及ぼし得る、不正確な、誤解を招く、または虚偽の情報

(2) おすすめフィードの対象外となるもの

- 一般的な陰謀論や緊急事態に関連する未確認の情報が含まれるコンテンツなど

ファクトチェック機関との連携

2020年以來、TikTokは50か国語以上をサポートする、世界中の18のファクトチェックパートナーと連携しています。



削除等の基準・措置の公正性・透明性の確保

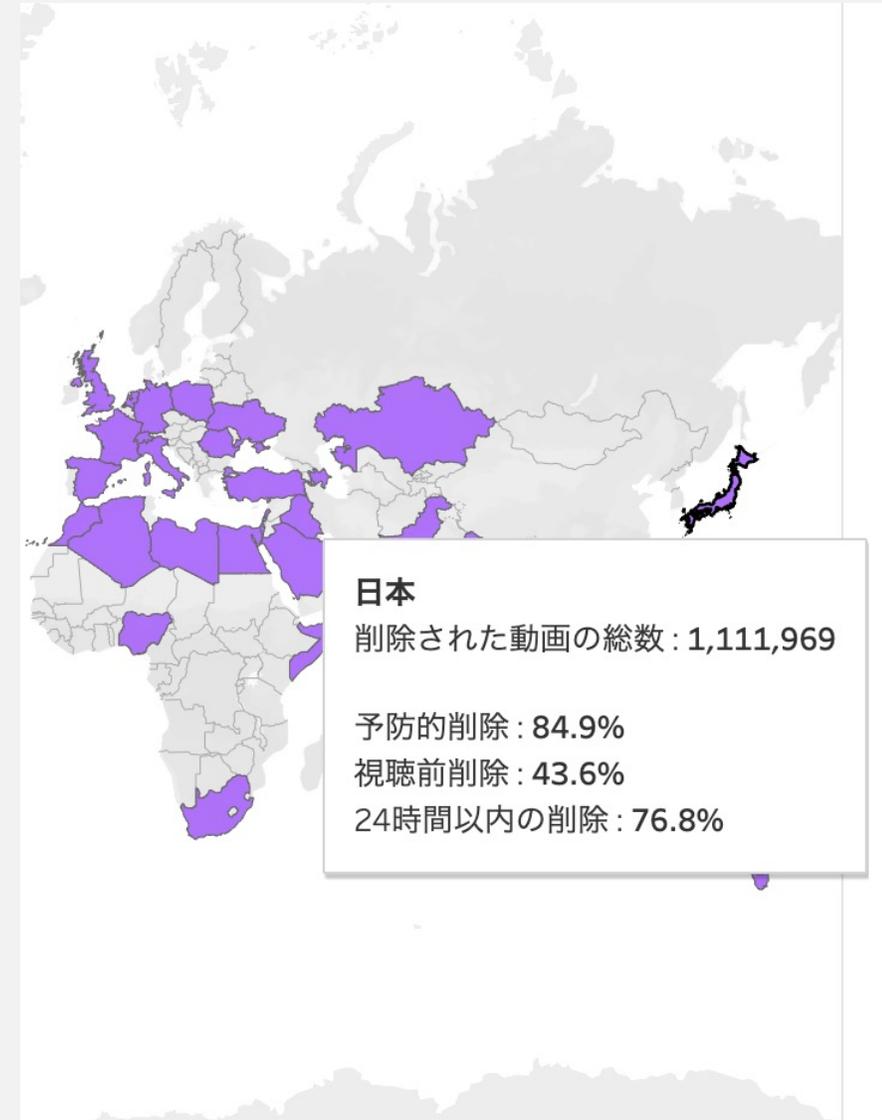
- コミュニティガイドラインの公表
- 投稿者への削除の通知、理由の明示
 - コミュニティガイドラインに違反して動画が削除された場合、どのコミュニティガイドラインに違反したかをユーザーに通知
- 異議申し立ての機会の確保
 - 削除の理由を通知する画面や当該動画の画面から、異議申し立てボタンをクリックして、申し立てできる
 - 再審査の結果も通知



削除等の基準や措置の公正性・透明性の確保

透明性レポートの公表

- 「地域・言語別のモデレーション」セクションにおいて、上位50位の地域別数値も公表
- 上位50位の地域について、データのダウンロードが可能
- モデレーターの主要言語の割合も公表
- 悪意をもって人の意思決定に影響を与えようとする活動の検知を具体的に公表



偽・誤情報の防止

信頼できる情報源へのアクセスと、慎重なアクションを呼びかけるガイド

紛争に関連する用語を検索すると、検索結果にガイドが表示される。

- ・ 情報が必ずしも正確でない可能性があることを伝える
- ・ 慎重なリアクションを呼びかける
- ・ 公式の情報源を確認することを促す

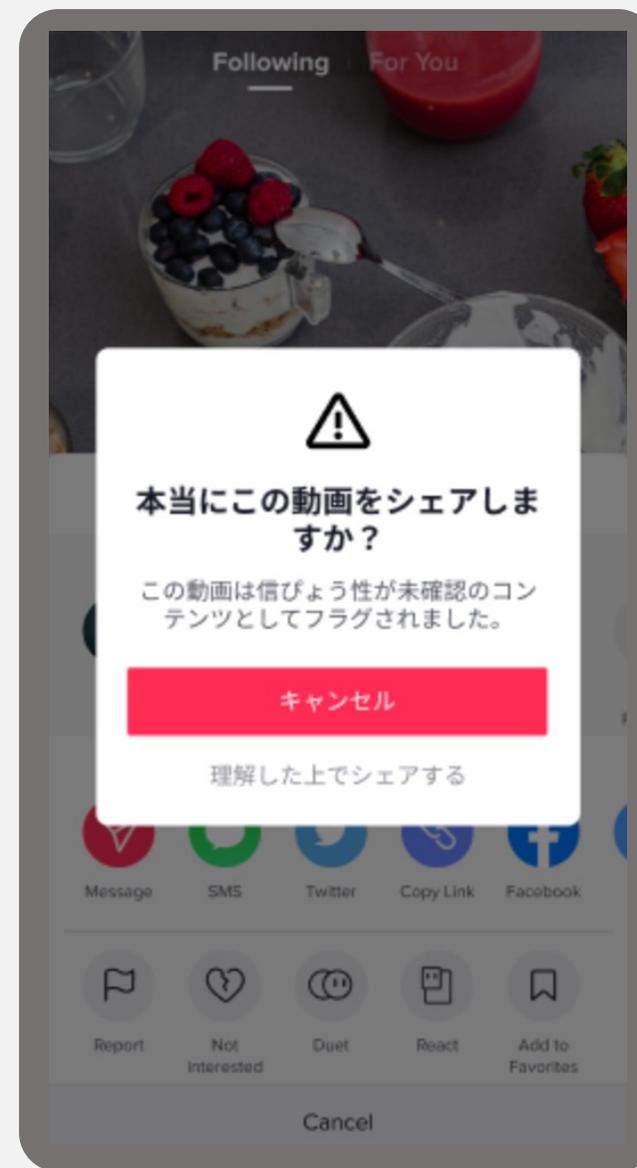
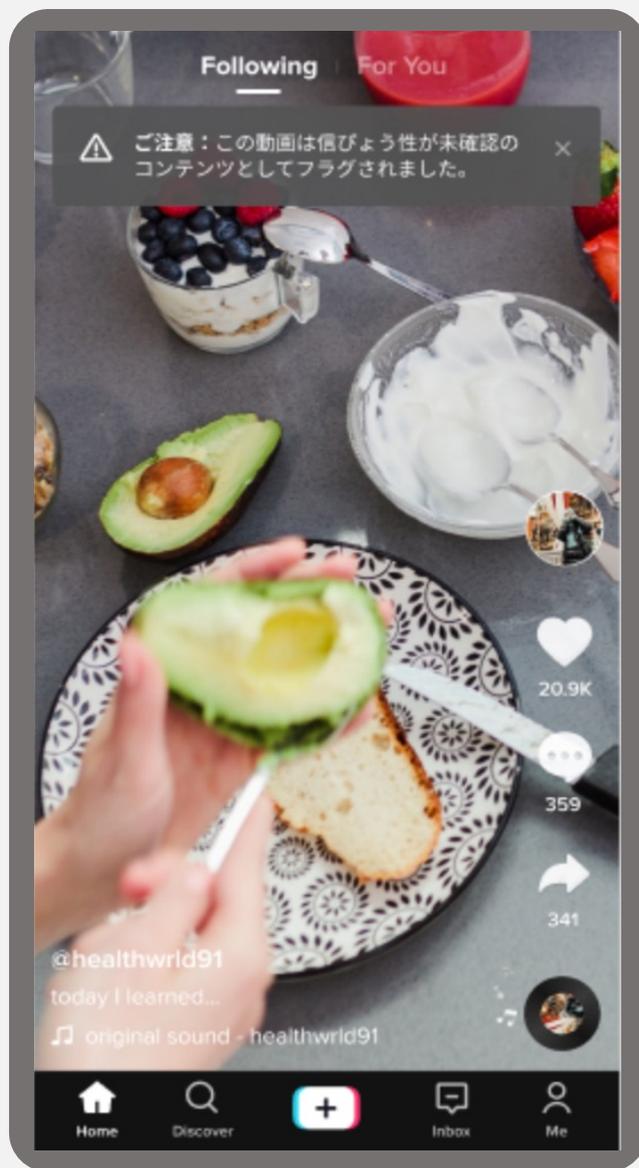


変化の激しい出来事

事態が急速に進展している出来事を取り扱うコンテンツは、視聴時点では正確ではなくなっていることがあります。配慮が必要なコンテンツにリアクションするときには注意し、公式の情報源で情報を確認してください。TikTokのコミュニティガイドラインに違反していると思われるものを見つけた場合は報告してください。

- ・ オンラインであなた自身とあなたのコミュニティを大切に扱う方法について、TikTokの**心の健康ガイド**で参考となる情報を確認してください。

偽・誤情報の防止 信ぴょう性未確認ラベル



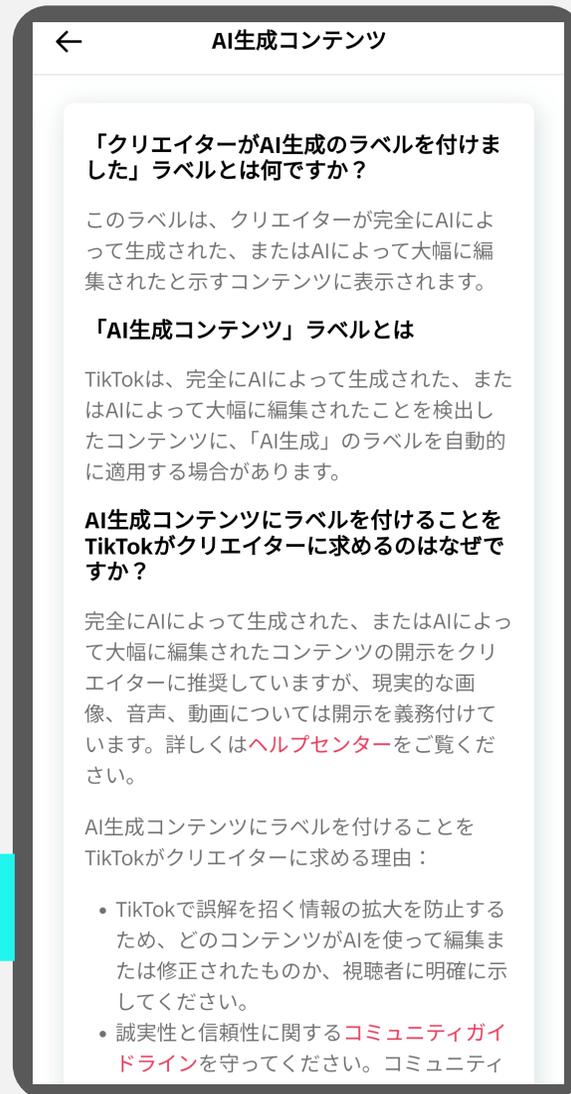
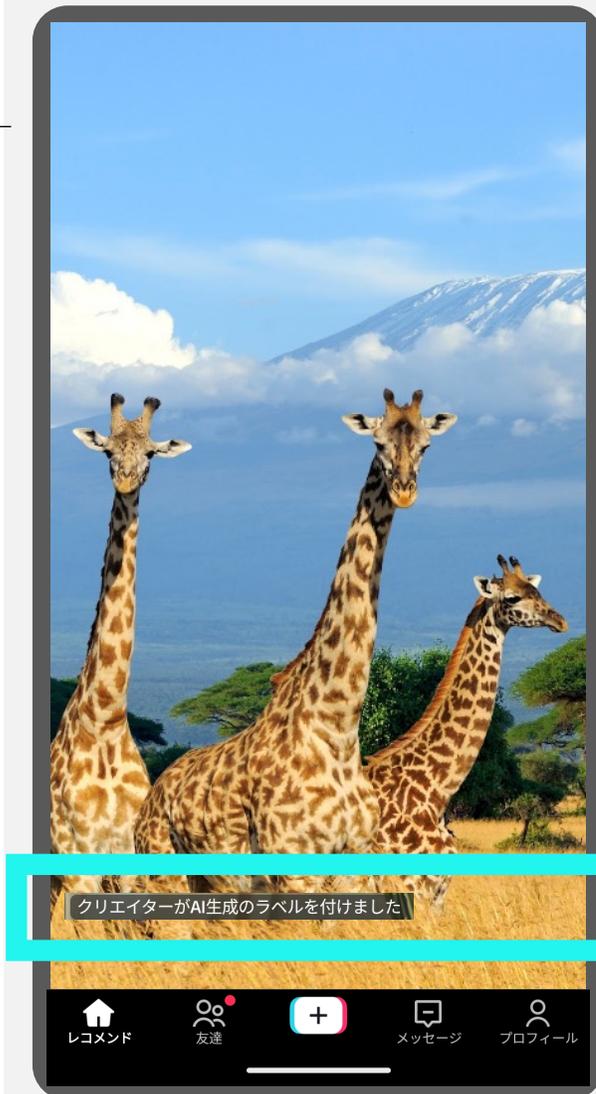
生成AIへの対応

- **実在する人物の映像または音声を含むAI生成コンテンツの制限**
 - 実在する人物の映像または音声を含むAI生成コンテンツを、コミュニティガイドラインで制限
- **AIで生成したコンテンツにラベルを表示**
 - AI生成コンテンツを投稿する際には「AI生成」ラベルをつけることを義務付け
- **透明性と責任ある共同行動のためのフレームワークに参加**
 - AIの透明性と責任あるイノベーションのためのフレームワークである「Partnership on AI」の Responsible Practices for Synthetic Media に参加

AI生成ラベル表示の義務化

AIで生成したコンテンツにユーザーがラベルを表示できる機能を開発

- AI生成コンテンツを投稿する際にはラベルをつけることを義務付け
- ラベルをクリックすると、**関連するコミュニティガイドラインへのリンク**や、**ラベルの趣旨を説明するガイド**に**遷移**



多様性あるレコメンドシステム、フィルターバブルの防止

• 多様なコンテンツを表示するレコメンドシステム

- ユーザーから提供された指標（いいね等）により、ユーザーにとって有用なコンテンツを予測する。予測スコアが高いものから順番にランキングが作成されるが、それらの類似性をチェックし、類似性が高い場合は、低いコンテンツと入れ替えることで、多様性を確保している

• おすすめフィードのリセット

- ユーザーがコンテンツが自分に合っていない、あるいはテーマの多様性が十分でないと感じたら、おすすめフィードをリセットできる

レコメンドシステムの透明性

- **レコメンドされた理由の表示**
 - なぜその動画がおすすめされたのか、動画ごとにユーザーが確認できる
- **レコメンドシステム/モデレーションの仕組みについて、積極的な情報公開**
 - 専門家が、動画審査の実践や、レコメンドシステムに関する情報にアクセスできる「透明性・説明責任情報公開センター」を開設
 - レコメンドシステムの仕組みを解説するWebページを設置し、積極的に情報を公開

