پست‌های طرفدار دولت ایران، شناسه پرونده: **2026-41/42-IG-UA**

تصمیم کامل پرونده

شرح و پس‌زمینه پرونده

این پرونده به این موضوع می‌پردازد که Meta چگونه باید با محتوایی برخورد کند که ممکن است بخشی از عملیات نفوذ پنهانی در پلتفرم‌هایش باشد؛ عملیاتی که دولت‌ها یا سازمان‌های دیگر با هدف تأثیرگذاری فریبکارانه بر کاربران انجام می‌دهند. هدف از چنین فعالیت‌هایی ایجاد این تصور است که موضعی خاص از حمایت گسترده و فزاینده‌ای برخوردار است. این تصمیم با بررسی دو محتوای منتشرشده در حمایت از دولت ایران و نیروهای امنیتی آن در دوره‌های تنش و بحران ملی، چالش شناسایی چنین کارزارهای پنهانی را تحلیل می‌کند.

هر یک از این محتواها را حساب‌هایی منتشر کرده‌اند که نام کاربری آن‌ها شامل عددی است که با یک گروه مجرم سایبری طرفدار حکومت ایران مرتبط دانسته شده است. این حساب‌ها نیز کمتر از دو ماه پیش از انتشار پست‌ها ایجاد شده بودند. این نشانه‌ها، در کنار ماهیت محتوا، حاکی از آن بود که پست‌ها ممکن است بخشی از یک عملیات سازمان‌یافته نفوذ آنلاین باشند.

نخستین پست که در ۱۲ ژوئن ۲۰۲۵ در اینستاگرام منتشر شد، تصویری از مردی با پوشش به سبک نظامی است و شرحی به زبان فارسی دارد که می‌گوید: «فقط کمی صبر کن، ایران.» متن روی تصویر چنین است: «صدای اتحاد مردم ایران تن و بدن تون را داره میلرزونه آره؟ به امام حسین دنیا را به شوک میبریم. ما با لباس رزم میخوابیدیم منتظرتون بودیم.» این محتوا پیش از آنکه حساب بعداً غیرفعال شود، بیش از ۱۴۰۰۰۰ بار مشاهده شد، حدود ۵۰۰۰ پسند دریافت کرد و تعداد نظرهای آن اندکی کمتر از ۱۰۰۰ مورد بود.

این تصویر پس از چند هفته افزایش فشار دیپلماتیک بین‌المللی بر ایران و در همان روزی منتشر شد که آژانس بین‌المللی انرژی اتمی اعلام کرد این کشور به تعهدات خود در زمینه منع گسترش سلاح‌های هسته‌ای پایبند نیست. روز بعد از انتشار پست، اسرائیل مجموعه‌ای از حملات نظامی را علیه ایران آغاز کرد که در پی آن چند تن از فرماندهان ارشد نظامی حکومت و دانشمندان هسته‌ای کشته شدند. تهران در واکنش دست به اقدام تلافی‌جویانه زد و درگیری تشدید شد؛ ایالات متحده نیز در ۲۲ ژوئن ۲۰۲۵ حملاتی را علیه ایران آغاز کرد. در ۲۴ ژوئن ۲۰۲۵، آتش‌بسی که دونالد ترامپ، رئیس‌جمهور ایالات متحده، اعلام کرده بود به اجرا درآمد؛ اما در ماه‌های بعد، همه طرف‌ها فعالیت نظامی را از سر گرفتند. تحریم‌های پیشین علیه حکومت ایران دوباره اعمال شد و وضعیت اقتصادی کشور وخیم‌تر شد. در این درگیری نزدیک به ۱۲۰۰ ایرانی کشته شدند که یک‌سوم آنان غیرنظامی بودند؛ همچنین ۲۸ اسرائیلی کشته شدند که به‌جز یک نفر، همگی غیرنظامی بودند.

در ۲۸ دسامبر ۲۰۲۵، در واکنش به وخامت اوضاع کشور، از جمله شرایط اقتصادی، اعتراضات گسترده‌ای در ایران آغاز شد. این اعتراضات به‌سرعت به تظاهرات گسترده سراسری تبدیل شد که خواستار پایان نظام جمهوری اسلامی و گذار به شکل جدیدی از حکومت بود. از اوایل ژانویه ۲۰۲۶، بنا بر گزارش‌ها نیروهای امنیتی کشتن و بازداشت معترضان را آغاز کردند.

در ۶ ژانویه ۲۰۲۶، کاربر دیگری در اینستاگرام ویدئویی کوتاه منتشر کرد که ظاهراً کاروانی از موتورسیکلت‌های حامل نیروهای امنیتی حکومت ایران را در حال عبور از خیابان‌های شهری نشان می‌داد. شرح فارسی ویدئو چنین بود: «آماده‌ایم تا آخرین نفس، جان خود را برای این کشور و نظام آن فدا کنیم.» هشتگ‌های آن نیز از جمله شامل #Iran، #Special_Unit، #Protest و #Unrest بود. روی این ویدیو موسیقی‌ای قرار داده شده است که متن آواز آن مضمونی مشابه با شرح ویدیو داشت و ادعا می‌کرد

که آن‌ها جان خود را فدای ایران و حکومت خواهند کرد. این محتوا بیش از ۹۰۰۰ بار مشاهده شد، بیش از ۳۰۰ پسند دریافت کرد و تعداد نظرهای آن کمتر از ۵۰ مورد بود. در روزهای ۷ و ۸ ژانویه، اعتراضات به اوج رسید و حکومت دسترسی به اینترنت را در کشور قطع کرد. نیروهای امنیتی کشتار گسترده‌ای از معترضان به راه انداختند؛ به سوی جمعیت تیراندازی کردند و هزاران نفر را کشتند و شمار بسیار بیشتری را زخمی کردند. اعتراضات تا ۲۸ فوریه ادامه یافت؛ در این روز، ایالات متحده و اسرائیل بار دیگر به ایران حمله کردند.

چندین کاربر هر دو پست را گزارش کردند. پس از آنکه سامانه‌های خودکار Meta تشخیص دادند هیچ‌یک از پست‌ها ناقض خطمشی‌های این شرکت نیستند، یک کاربر نسبت به هر دو تصمیم درخواست تجدیدنظر کرد. هیچ‌یک از دو درخواست تجدیدنظر برای بررسی انسانی در اولویت قرار نگرفت. پس از پایان مهلت رسیدگی، کاربر برای هر دو پرونده به «هیئت نظارت» درخواست تجدیدنظر داد.

پس از انتخاب این پرونده‌ها از سوی «هیئت»، Meta تأیید کرد که هیچ‌یک از دو پست به‌تنهایی ناقض خطمشی‌های این شرکت نیست. سپس شرکت حساب‌ها را بررسی کرد و به این نتیجه رسید که ظاهراً بخشی از یک عملیات سازمان‌یافته نبوده‌اند. بااین‌حال، «هیئت» تصمیم گرفت رسیدگی به پرونده را ادامه دهد تا دشواری شناسایی عملیات نفوذ و دیگر فعالیت‌های فریبکارانه هماهنگ و نیز تأثیر آن‌ها بر آزادی بیان کاربران را برجسته کند. پس از انتخاب این پرونده‌ها از سوی «هیئت»، Meta افشا کرد که حساب منتشرکننده محتوای پرونده نخست، به‌تازگی بر اساس خطمشی یکپارچگی حساب (Account Integrity Policy) غیرفعال شده بود؛ در نتیجه، پست نیز از پلتفرم حذف شده بود. در زمان نگارش این تصمیم، حسابی که ویدئوی کاروان موتورسیکلت‌ها را منتشر کرده بود همچنان فعال بود.

مطالب ارسالی کاربران

کاربری که درباره هر دو محتوا به «هیئت» درخواست تجدیدنظر داده بود، اظهار داشت که عدد به‌کاررفته در نام هر دو حساب ظاهراً نشانه‌ای رمزی بوده است که به حساب‌های هماهنگ امکان می‌داد یکدیگر را شناسایی و محتوای یکدیگر را تقویت کنند. او اظهار داشت که محتوا از گفتار سیاسی متعارف فراتر می‌رود و در واقع از زبان رمزی و تعابیر تلویحی تشکیل شده است که هدفشان ارعاب و تحریک است. این کاربر همچنین توضیح داد که زبان مربوط به آمادگی و وفاداری، تهدیدی ضمنی علیه مخالفان و منتقدان دولت ایران بوده است. او همچنین هر دو پست را با خطر آسیب در دنیای واقعی مرتبط دانست و خواستار اجرای پیش‌دستانه‌تر و زمینه‌محورتر خطمشی‌ها در قبال شبکه‌های هماهنگ شد.

خطمشی‌ها و مطالب ارسالی **Meta**

I. خطمشی‌های محتوا

Meta تشخیص داد که خطمشی «سازمان‌ها و افراد خطرناک» (Dangerous Organizations and Individuals Policy) و خطمشی «خشونت و تحریک به خشونت» (Violence and Incitement Policy)

ممکن است با این پرونده‌ها مرتبط باشند. شرکت هر یک از محتواها را بر اساس این خطمشی‌ها ارزیابی کرد.

خطمشی «سازمان‌ها و افراد خطرناک» حضور «سازمان‌ها یا افرادِ تعیین‌شده‌ای را که مأموریتی خشونت‌آمیز اعلام می‌کنند یا به خشونت دست می‌زنند» در پلتفرم‌های Meta ممنوع می‌کند. Meta به «هیئت» اعلام کرد که هیچ‌یک از دو پست به نهادهای تعیین‌شده اشاره نمی‌کند و بنابراین ناقض خطمشی نیست. در پرونده نخست به ایران و نیروهای نظامی آن اشاره شده است، اما نام هیچ سازمان یا فرد مشخصی مطرح نمی‌شود. مردی که با لباس به سبک نظامی در تصویر دیده می‌شود، هیچ نشان یا نمادی متعلق به نهادهای تعیین‌شده را بر لباس خود ندارد. در پرونده دوم، مردانی سوار بر موتورسیکلت‌های پلیس و با لباس رزم دیده می‌شوند. هشتگ به‌کاررفته، #Special_Unit، به یگان ویژه اشاره دارد. به باور Meta، این هشتگ به نیروی انتظامی ملی ایران اشاره می‌کند که اغلب با سرواژه فارسی «فراجا» شناخته می‌شود و طبق این خطمشی در شمار نهادهای تعیین‌شده قرار ندارد. این محتوا به هیچ سازمان یا فرد دیگری اشاره نمی‌کند.

هدف خطمشی «خشونت و تحریک به خشونت» Meta، پیشگیری از آسیب احتمالی در دنیای واقعی است که ممکن است با محتوای منتشرشده در پلتفرم‌های این شرکت مرتبط باشد؛ این هدف از طریق حذف «زبانی که تهدید به خشونت می‌کند، به آن تحریک می‌کند یا آن را تسهیل می‌کند» دنبال می‌شود. Meta به «هیئت» گفت که از نظر این شرکت، هیچ‌یک از دو پست حاوی تهدید به خشونت، تحریک به خشونت یا تسهیل خشونت علیه هدفی قابل‌شناسایی نبوده است.

Meta ارزیابی کرد که پرونده نخست حاوی «ادبیات ملی‌گرایانه اغراق‌آمیز» است، اما هیچ اقدام خشونت‌آمیز مشخص، روش آسیب‌رسانی یا قصد ارتکاب خشونت علیه فرد یا گروهی را توصیف نمی‌کند. شرکت توضیح داد که زبان به‌کاررفته در پست مبهم است، مخاطب مشخصی ندارد و به نظر می‌رسد بیانیه‌ای در حمایت از ایران باشد که در دوره فشار شدید بین‌المللی منتشر شده است.

درباره پست دوم، Meta توضیح داد که هرچند زبان آن میهن‌پرستانه است، قصد، آرزو یا فراخوانی برای ارتکاب اقدام خشونت‌آمیز مشخصی را بیان نمی‌کند. شرکت خاطرنشان کرد که حتی اگر بتوان محتوا را حاوی زبانی تهدیدآمیز دانست، این زبان متوجه هیچ هدف قابل‌شناسایی‌ای نیست. Meta تشخیص داد که هشتگ‌های اشاره‌کننده به اعتراضاتِ هم‌زمان با انتشار پست، تهدیدی به خشونت علیه معترضان محسوب نمی‌شوند و به خشونت علیه آنان نیز تحریک یا آن را تسهیل نمی‌کنند.

خطمشی «خشونت و تحریک به خشونت» در بخش فرعی از مواردی که برای اجرای آن‌ها «اطلاعات و/یا زمینه بیشتر» لازم است، بیان می‌کند که Meta «اظهارات رمزی‌ای را حذف می‌کند که در آن‌ها روش ارتکاب خشونت به‌روشنی بیان نشده، اما تهدید با توجه به ترکیب نشانه تهدید و نشانه زمینه‌ای، پوشیده یا ضمنی است.» نشانه تهدید می‌تواند عبارتی رمزی باشد که «در بستری تلافی‌جویانه منتشر شده» یا «نقش فراخوانی تهدیدآمیز به اقدام را ایفا می‌کند». نشانه‌های زمینه‌ای نیز شامل تأیید کارشناسان محلی، اطلاعاتی مبنی بر اینکه آن عبارت ممکن است به خشونت قریب‌الوقوع منجر شود، یا گزارش محتوا از سوی فردی است که هدف تهدید قرار گرفته است.

II. خطمشی‌های مربوط به نقش‌آفرین و رفتار

بیشتر استانداردهای جامعه Meta درباره تک‌تک محتواها اعمال می‌شوند. این استانداردها چارچوبی فراهم می‌کنند تا هر پست جداگانه با قواعد مربوط سنجیده شود و اگر ناقض آن‌ها باشد، حذف شود یا اقدام اجرایی دیگری درباره آن صورت گیرد؛ برای مثال، صفحه هشدار روی آن نمایش داده شود. از جمله این خطمشی‌های محتوامحور، خطمشی «خشونت و تحریک به خشونت» است که Meta در این پرونده‌ها درباره محتوا اعمال کرد.

بااین‌حال، برخی دیگر از خطمشی‌های شرکت دامنه گسترده‌تری دارند: این خطمشی‌ها بر شخص یا نقش‌آفرین پشت محتوای موردنظر یا بر الگوهای رفتاری در سراسر پلتفرم‌های شرکت تمرکز می‌کنند. این تصمیم به سه مورد از این خطمشی‌ها می‌پردازد.

یکی از این خطمشی‌ها، بازنمایی هویت واقعی (Authentic Identity Representation Policy) است. این خطمشی، حساب‌هایی را در فیس‌بوک یا اینستاگرام ممنوع می‌کند که عمداً هویت خود را نادرست جلوه می‌دهند تا دیگران را گمراه یا فریب دهند، از اقدامات اجرایی بگریزند یا خطمشی‌های دیگر را نقض کنند. این ممنوعیتِ مبتنی بر نقش‌آفرین، بر هویت شخصِ پشت حساب و محتوایی تمرکز دارد که آن حساب منتشر می‌کند.

دومین خطمشی از این نوع، استاندارد جامعه درباره رفتار فریبکارانه (Inauthentic Behavior Community Standard) است که به رفتار در سراسر پلتفرم‌های Meta می‌پردازد. در توضیح مبنای این خطمشی آمده است: «رفتار فریبکارانه به انواع گوناگون و پیچیده فریب اشاره دارد که شبکه‌ای از دارایی‌های فریبکارانه تحت کنترل یک یا چند فرد انجام می‌دهد؛ هدف از آن فریب دادن Meta یا جامعه ما یا دور زدن اقدامات اجرایی ذیل استانداردهای جامعه است.» Meta در پاسخ به پرسش‌های «هیئت» توضیح داد که از این خطمشی برای بررسی فعالیت‌های فریبکارانه هماهنگ میان حساب‌های مرتبط استفاده می‌کند و شبکه‌هایی را هدف می‌گیرد که آگاهانه هویت خود را نادرست جلوه می‌دهند و به اقدامات فریبکارانه دست می‌زنند.

عملیات نفوذ تحت حمایت دولت در بخشی از این خطمشی قرار می‌گیرد که به پیچیده‌ترین شکل رفتار فریبکارانه می‌پردازد؛ Meta این شکل را «رفتار فریبکارانه هماهنگ» می‌نامد. این رفتار سه ویژگی دارد:

(۱) هماهنگی میان شبکه‌ای از حساب‌ها برای انجام شکل‌های به‌ویژه پیچیده رفتار فریبکارانه

(۲) استفاده از هویت‌های جعلی به‌عنوان بخش محوری عملیات

(۳) استفاده از روش‌های خصمانه برای گریز از شناسایی یا معتبر جلوه دادن خود.

اقدام اجرایی در این زمینه بر مختل‌سازی شبکه متمرکز است و به حذف کامل شبکه‌های هماهنگ می‌انجامد.

Meta با بررسی طیف گسترده‌ای از نشانه‌های رفتاری، بر شناسایی الگوهای رفتار فریبکارانه در شبکه‌های کاربران تمرکز می‌کند. این رویکرد جامع و رفتارمحور، جزئیاتی مانند نحوه ایجاد و استفاده از حساب‌ها، اطلاعات جغرافیایی، زمان‌بندی و هماهنگی میان حساب‌های مختلف و الگوهای انتشار محتوا در سراسر شبکه مشکوک را بررسی می‌کند.

سومین خطمشی، استاندارد جامعه درباره یکپارچگی حساب (Account Integrity) است که آن نیز هم مبتنی بر نقش‌آفرین و هم مبتنی بر رفتار است. این خطمشی شرایطی را مشخص می‌کند که در آن‌ها Meta حساب‌ها را محدود یا غیرفعال می‌کند. این شرایط، از جمله، شامل نقض‌های شدید یا مکرر خطمشی‌ها و ایجاد یا استفاده از حساب‌هایی برای دور زدن حذف حساب‌های دیگر است. این خطمشی بیان می‌کند که Meta می‌تواند حساب‌هایی را غیرفعال کند که یکی از وضعیت‌های زیر را نشان می‌دهند:

«ارتباط نزدیک با شبکه‌ای از حساب‌ها یا نهادهای دیگر که خطمشی‌های [Meta] را نقض می‌کنند یا از آن‌ها می‌گریزند»،

«هماهنگی در شبکه‌ای از حساب‌ها یا نهادهای دیگر که به‌طور مداوم یا فاحش خطمشی‌های [Meta] را نقض می‌کنند» و «فعالیت یا رفتاری که از طریق شبکه‌ای از حساب‌ها، هدف آشکار نقض خطمشی‌ها را نشان می‌دهد».

این سه خطمشی از نظر موضوع تمرکز و ماهیت آسیبی که هر نوع نقض ایجاد می‌کند با یکدیگر تفاوت دارند. خطمشی «بازنمایی هویت واقعی» به مواردی می‌پردازد که در آن‌ها یک حساب منفرد می‌کوشد کاربران دیگر را درباره هویت شخص یا نهادی که آن

حساب نمایندگی می‌کند، گمراه کند. برای شمول این خطمشی، وجود ارتباط با حساب‌های دیگر یا انجام فعالیت هماهنگ ضروری نیست.

ممنوعیت رفتار فریبکارانه هماهنگ در خطمشی «رفتار فریبکارانه» ناظر بر موقعیت‌هایی است که در آن‌ها چند نقش‌آفرین، با استفاده از چندین حساب متمایز و دارایی‌های دیگر، فعالیت فریبکارانه و فریبکارانهی را به‌صورت شبکه‌ای هماهنگ می‌کنند.

ممنوعیت‌های شبکه‌ایِ خطمشی «یکپارچگی حساب» برای موقعیت‌هایی طراحی شده‌اند که در آن‌ها یک نقش‌آفرین چندین حساب را اداره می‌کند تا از اقدامات اجرایی بگریزد. برای شمول این خطمشی لازم نیست نقش‌آفرین هویت خود را در هیچ‌یک از حساب‌های منفرد نادرست جلوه داده باشد.

III. حساب اول

پیش از آنکه «هیئت» این پرونده را برای رسیدگی انتخاب کند، Meta حساب منتشرکننده محتوای پرونده نخست را غیرفعال کرده بود. Meta تشخیص داد که این حساب هویت خود را نادرست جلوه می‌دهد و نماینده کاربری واقعی نیست؛ بنابراین آن را بر اساس خطمشی «بازنمایی هویت واقعی» غیرفعال کرد. بااین‌حال، پس از آنکه «هیئت» درباره شرایط غیرفعال‌سازی حساب پرس‌وجو کرد، Meta تشخیص داد که حساب باید به‌جای آن، بر اساس خطمشی «یکپارچگی حساب» غیرفعال می‌شد؛ زیرا این حساب ارتباط نزدیکی با چندین حساب دیگر داشت که هر یک به‌طور مستقل، به‌دلیل نقض خطمشی‌های Meta ــ از جمله خطمشی‌های منع‌کننده رفتار فریبکارانه و کلاهبرداری ــ غیرفعال شده بودند.

هرچند این حساب به شبکه‌ای از حساب‌های دیگر مرتبط بود، با کاربران دیگری که بخشی از شبکه هماهنگ گسترده‌تر یا عملیات نفوذ باشند ارتباط نداشت. بنابراین، Meta تشخیص داد که این حساب خطمشی «رفتار فریبکارانه» را نقض نکرده است.

«هیئت» درباره تلاش‌های Meta برای نظارت بر محتوا و اجرای خطمشی‌ها در ایران، نحوه شناسایی عملیات نفوذ و دیگر اشکال رفتار فریبکارانه و اقدام اجرایی علیه آن‌ها، و نیز نحوه تعامل و اجرای خطمشی‌های مرتبط Meta پرسش‌هایی مطرح کرد. Meta به همه پرسش‌ها پاسخ داد.

نظرات عمومی

«هیئت» پنج نظر عمومی دریافت کرد که شرایط ارسال را برآورده می‌کردند. چهار نظر از ایالات متحده و کانادا و یک نظر از خاورمیانه ارسال شده بود. برای مطالعه نظرات عمومی‌ای که با رضایت صاحبان آن‌ها برای انتشار ارائه شده‌اند، اینجا کلیک کنید.

این نظرات، از جمله، به موضوعات زیر می‌پرداختند: ضرورت استفاده از رویکردهای گسترده و رفتارمحور، به‌جای رویکردهای صرفاً متمرکز بر محتوا، برای مقابله با عملیات نفوذ؛ نحوه عملکرد این عملیات، از جمله عملیات انجام‌شده توسط دولت ایران یا به نمایندگی از آن؛ و ضرورت اطمینان از اینکه تلاش‌ها برای مقابله با عملیات نفوذ، بیان مشروع، به‌ویژه گفتار سیاسی، را محدود نکنند.

تحلیل هیئت نظارت

«هیئت» دو پست را هم بر اساس خطمشی‌های محتوایی مرتبط Meta و هم با توجه به تعهدات حقوق بشری این شرکت ارزیابی کرد. ارزیابی حقوق بشری بر آزمون سه‌بخشیِ قانونی بودن، هدف مشروع و ضرورت و تناسب استوار است. این آزمون در بخش‌های ۵(III) تا ۵(V) در ادامه با جزئیات بیشتری شرح داده می‌شود؛ در آن بخش‌ها «هیئت» غیرفعال‌سازی حساب در این پرونده و رویکرد Meta به عملیات نفوذ را در پرتو تعهدات حقوق بشری شرکت بررسی کرده است.

I. *تصمیم‌گیری درباره محتوا*

«هیئت» تشخیص می‌دهد که هیچ‌یک از دو پست خطمشی «سازمان‌ها و افراد خطرناک» را نقض نمی‌کند. اگرچه مرد به‌تصویرکشیده‌شده در پست نخست لباسی به سبک نظامی به تن دارد، هیچ نشان، درجه یا علامت دیگری وجود ندارد که دلالت بر عضویت یا حمایت او از سپاه پاسداران انقلاب اسلامی (IRGC) یا هر نهاد دیگری داشته باشد که از سوی Meta تعیین شده است. شخصیت مذهبیِ مورد اشاره در پست نخست به هیچ‌یک از نهادهای تعیین‌شده وابسته نیست. در پست دوم، موتورسواران هیچ وابستگیِ قابل‌مشاهده‌ای به گروهی تعیین‌شده ندارند. اگرچه هر دو پست با لحنی حمایتی و میهن‌پرستانه به کشور ایران اشاره می‌کنند، هیچ نشانه‌ای وجود ندارد که این حمایت را بتوان به حمایت از سپاه پاسداران انقلاب اسلامی ایران (IRGC) یا نهاد تعیین‌شده دیگری تعمیم داد.

پست نخست نیز خطمشی «خشونت و تحریک به خشونت» را نقض نمی‌کند؛ زیرا حاوی تهدید صریح یا ضمنی یا فراخوان به خشونت نیست. این پست در دوره‌ای از تنش شدید منطقه‌ای منتشر شده و می‌گوید ایران «جهان» را به شوک می‌برد، اما مشخص نمی‌کند چگونه یا دقیقاً علیه چه کسی. این عبارت را می‌توان به چند شکل تفسیر کرد، اما محتمل‌ترین برداشت آن است که ملت ایران متحد است و در برابر حمله خارجی مقابله خواهد کرد. عبارت «ما با لباس رزم میخوابیدیم منتظرتون بودیم» را می‌توان چنین تفسیر کرد که کاربر ــ که ظاهراً قصد دارد از جانب ایرانیان ملی‌گرا سخن بگوید ــ آماده پاسخ دادن به تجاوز است؛ اما این عبارت نه تهدید به خشونت می‌کند و نه خواستار اجرای آن می‌شود. «هیئت» خاطرنشان می‌کند که در دوره‌های بحران و تنش بین‌المللی، اظهارات حمایتی یا مخالفانه درباره دولت‌ها و نیروهای مسلح آن‌ها و نیز فراخوان‌های مرتبط به خشونت، احتمالاً افزایش می‌یابد. پلتفرم‌های رسانه‌های اجتماعی باید اطمینان حاصل کنند که گفتار سیاسی مشروع را در مواردی که هیچ خطمشی‌ای نقض نشده است حذف نمی‌کنند.

بااین‌حال، اکثریت «هیئت» تشخیص می‌دهد که با توجه به شرایطی که پست دوم در آن منتشر شده، این پست خطمشی «خشونت و تحریک به خشونت» را نقض می‌کند و باید حذف می‌شد. با توجه به اعتراضاتِ در جریان در آن زمان، جنایات فجیعی که در واکنش به اعتراضات صورت می‌گرفت و سابقه نیروهای امنیتی ایران در نقض فاحش حقوق بشر، این پست با تعریف Meta از «تهدید تلویحی (Veiled Threat)» مطابقت دارد؛ این تعریف، پست‌های بالقوه مبهم را در صورتی تهدید محسوب می‌کند که هم «نشانه تهدید» و هم «نشانه زمینه‌ای» داشته باشند و این دو در کنار هم تهدیدی ضمنی یا تلویحی را شکل دهند (برای نمونه، بنگرید به محتوای هدف‌گیرنده مدافع حقوق بشر در پرو).

نشانه تهدید: این پست مشخصاً شرط وجود «نشانه تهدید» را از طریق «فراخوانی تهدیدآمیز به اقدام» برآورده می‌کند. در ویدئو، نیروهای امنیتی سوار بر موتورسیکلت در حال عبور از شهری در ایران دیده می‌شوند و این شرح روی آن آمده است: «آماده‌ایم تا آخرین نفس، جان خود را برای این کشور و نظام آن فدا کنیم.» با توجه به شرایط آن زمان، باید این زبان حمایتی را فراخوانی از نیروهای امنیتی برای جنگیدن تا پای مرگ علیه معترضان دانست، نه علیه نیرویی خارجی. این برداشت با استفاده از هشتگ‌های #protest و #unrest در پست تقویت می‌شود. این عبارت، در کنار تصاویری که برای برجسته کردن قدرت نیروهای امنیتی به‌کار رفته‌اند، می‌تواند و باید فراخوانی تهدیدآمیز به خشونت علیه معترضان تفسیر شود.

نشانه زمینه‌ای: این پست پس از آن منتشر شد که نیروهای امنیتی ایران کشتن معترضان را آغاز کرده بودند. این زمینه، در کنار سابقه مستند نیروهای امنیتی ایران در واکنش‌های خشونت‌آمیز به اعتراضات، از جمله کشتن و اعدام معترضان، برای این نتیجه‌گیری کافی است که این محتوا می‌تواند به خشونتی قریب‌الوقوع منجر شود. کشتارهای گسترده‌ای که نیروهای امنیتی در روزهای بلافاصله پس از انتشار این پست انجام دادند و طی آن هزاران غیرنظامی کشته شدند، بعدها مؤید این ارزیابی بود.

ارزیابی این موارد دشوار است و در محیطی که اطلاعات به‌سرعت تغییر می‌کند و قابل‌اتکا نیست، به اطلاعات زمینه‌ای و شناخت منطقه‌ای نیاز دارد. Meta تهدیدهای تلویحی را فقط پس از ارجاع پرونده به بازبینان متخصص داخلی ارزیابی می‌کند. بازبینانی که در مقیاس وسیع محتوا را بررسی می‌کنند، اختیار شناسایی محتوای ناقض خط‌مشی‌هایی را ندارند که فقط پس از ارجاع به سطح تخصصی قابل اجرا هستند. «هیئت» پیش‌تر درباره تمایز میان تهدیدهایی که برای اقدام اجرایی به بررسی زمینه نیاز دارند و تهدیدهایی که چنین نیازی ندارند، ابراز نگرانی کرده است (برای نمونه، بنگرید به محتوای هدف‌گیرنده مدافع حقوق بشر در پرو).

اکثریت «هیئت» تشخیص می‌دهد که حذف محتوای دوم از پلتفرم، مطابق با تفسیر صحیح خط‌مشی‌های محتوایی Meta، با مسئولیت‌های حقوق بشری این شرکت سازگار است. «هیئت» پیش‌تر تشخیص داده است که ممنوعیت‌های خط‌مشی «خشونت و تحریک به خشونت» درباره «تهدید به خشونت علیه اهداف گوناگون» و «اظهارات رمزی‌ای که در آن‌ها روش خشونت به‌روشنی بیان نشده، اما تهدید تلویحی یا ضمنی است» به‌اندازه کافی روشن‌اند و هدفی مشروع را دنبال می‌کنند؛ بنابراین، دو شرط نخست آزمون سه‌بخشی را برآورده می‌سازند.

با توجه به تشدید خشونت دولتی علیه معترضان در زمان انتشار پست، «هیئت» تشخیص می‌دهد که حذف محتوا محدودیتی ضروری و متناسب بر آزادی بیان است تا از حق حیات معترضان و نیز حق آنان برای تجمع مسالمت‌آمیز حمایت کند (ماده ۲۱ میثاق بین‌المللی حقوق مدنی و سیاسی (ICCPR)). هیچ اقدام دیگری جز حذف پست نمی‌توانست خطرهای ناشی از آن را به‌طور کافی کاهش دهد. این دو پست را می‌توان بر اساس اینکه ظاهراً چه کسانی را هدف قرار می‌دهند، از یکدیگر متمایز کرد. به نظر می‌رسد پست نخست، در آستانه یک درگیری مسلحانه بین‌المللی، دولت‌های خارجیِ متخاصم را مخاطب قرار می‌دهد. پست دوم، از طریق هشتگ‌ها و زمان انتشارش، جمعیتی داخلی را هدف قرار می‌دهد که علیه نیروهای نمایش‌داده‌شده در ویدئو اعتراض می‌کنند.

بااین‌حال، اقلیتی از اعضای «هیئت» تشخیص می‌دهند که پست دوم حاوی تهدید یا فراخوانی به خشونت نیست که خط‌مشی «خشونت و تحریک به خشونت» را نقض کند. این گفته که کاربر با میل و رغبت حاضر است برای کشورش، ایران، جان بدهد، در کنار تصویری از افرادی که ظاهراً از نیروهای امنیتی هستند، به‌طور روشن مصداق تهدیدی ــ صریح یا رمزی ــ علیه هدفی قابل‌شناسایی نیست. افزودن هشتگ‌های #protest و #unrest به شرح پست، به‌خودی‌خود نمی‌تواند به‌منزله معرفی معترضان به‌عنوان هدف تلقی شود؛ زیرا تهدیدی متوجه آنان نشده است. افزون بر این، با وجود زمینه‌ای که در بالا به آن اشاره شد، در محتوا هیچ نشانه‌ای وجود ندارد که نشان دهد این پست می‌تواند موجب آسیب احتمالی در دنیای واقعی به اهداف قابل‌شناسایی شود. از نظر اقلیت اعضای «هیئت»، حذف این پست صرفاً بر پایه زمینه انتشار آن، در حالی که محتوا حتی تهدیدی ضمنی در بر ندارد، نه ضروری است و نه متناسب. بنابراین، اقلیت نیز تشخیص می‌دهد که باقی گذاشتن پست در پلتفرم Meta با خط‌مشی‌ها و تعهدات حقوق بشری این شرکت سازگار بوده است.

اظهارنظر در حمایت از دولت، حتی اگر آن دولت در سطح بین‌المللی محکوم شده باشد، همچنان از گفتار سیاسیِ مورد حمایت برخوردار است. بااین‌حال، هیچ‌کس ــ از جمله دولت‌ها، مأموران و حامیان آن‌ها ــ به استناد حق آزادی بیان مجاز نیست معترضانِ پایبند به قانون را به خشونت تهدید کند.

II. *عملیات نفوذ و تصمیم‌گیری درباره حساب‌ها*

شناسایی عملیات نفوذ در رسانه‌های اجتماعی صرفاً با بررسی پست‌های منفرد یا حتی حساب‌های منفرد دشوار است. از آنجا که این عملیات اغلب به موضوعات اجتماعی و سیاسیِ تفرقه‌برانگیز می‌پردازند، تشخیص تفاوت میان کاربری که باورهای خود را بیان می‌کند و ناخواسته اطلاعات نادرست منتشر می‌سازد، با فردی که بخشی از تلاش هماهنگ برای انتشار اطلاعات گمراه‌کننده با هدفی مشخص است، دشوار است. سنجش تأثیر عملیات نفوذ اغلب دشوار است، اما این عملیات می‌توانند اعتماد را در نظام‌های دموکراتیک، گفتمان عمومی و یکپارچگی اطلاعات تضعیف کنند و اغلب طراحی شده‌اند تا شکاف‌های اجتماعی را تشدید کنند. این عملیات همچنین ممکن است با تشدید تنش و تحریک خشونت در برخی بسترها آسیب‌های شدیدی ایجاد کنند. بااین‌حال، هرچند پلتفرم‌ها باید برای مقابله با انتشار سازمان‌یافته اطلاعات گمراه‌کننده تلاش کنند، باید احتیاط لازم را به‌کار برد تا اقدامات نامتناسب و غیرضروری آزادی بیان را محدود نکنند.

«هیئت» این پرونده‌ها را بستری برای پرداختن به عملیات نفوذ انتخاب کرد؛ زیرا محتوا و نشانه‌های پیرامون آن این احتمال را مطرح می‌کردند که هر یک از پست‌ها بخشی از چنین عملیاتی باشد. پس از اعلام پرونده‌ها، Meta تحقیق کرد و به این نتیجه رسید که هیچ‌یک از دو پست بخشی از عملیات نفوذ نبوده است؛ «هیئت» نیز این یافته را پذیرفت.

بخش باقی‌مانده این تصمیم، غیرفعال‌سازی حساب بر اساس خطمشی‌های «رفتار فریبکارانه» و «یکپارچگی حساب» را در پرتو تعهدات حقوق بشری Meta ارزیابی می‌کند.

///. قانون‌مندی (شفافیت و دسترسی به قوانین)

اصل قانون‌مندی ایجاب می‌کند قواعد محدودکننده آزادی بیان در دسترس و روشن باشند و با دقت کافی تدوین شوند تا افراد بتوانند رفتار خود را بر اساس آن‌ها تنظیم کنند (تفسیر عمومی شماره ۳۴، بند ۲۵). همچنین، این قوانین «نمی‌توانند اختیار بی قید و شرط برای محدود کردن آزادی بیان برای افراد مسئول اجرای آنها اعطا کنند» و باید «راهنمایی کافی برای کسانی که مسئول اجرا هستند فراهم کنند تا بتوانند تشخیص دهند چه نوع ابراز بیان به درستی محدود می‌شود و چه نوع نمی‌شود» (*همان منبع*). گزارشگر ویژه سازمان ملل در امور آزادی بیان اعلام کرده است که قواعدی که بر حکمرانی بازیگران خصوصی بر گفتار آنلاین اعمال می‌شوند باید روشن و مشخص باشند (A/HRC/38/35، بند ۴۶). افرادی که از پلتفرم‌های Meta استفاده می‌کنند باید بتوانند به قوانین دسترسی داشته باشند و آن‌ها را بفهمند و بازبینی‌کنندگان محتوا باید در رابطه با اجرای قانون‌هایشان رهنمودهای روشنی داشته باشند.

Meta در اقدامات اجرایی خود علیه عملیات نفوذ از سامانه‌های هوش مصنوعی و تخصص انسانی بهره می‌گیرد و برای شناسایی و حذف این عملیات، طیف گسترده‌ای از نشانه‌ها را تحلیل می‌کند. خطمشی رفتار فریبکارانه، قواعدی را که در این زمینه به‌صورت عمومی اعلام شده‌اند، تعیین می‌کند. بااین‌حال، این خطمشی سطحی و فاقد جزئیات کافی است. این خطمشی بیان می‌کند که ایجاد یا استفاده از حساب‌ها و دارایی‌های فریبکارانه برای فریب دادن Meta یا کاربران آن، گریز از اجرای خطمشی‌های Meta یا سوءاستفاده از سامانه‌های این شرکت برای آزار، ارعاب یا خاموش کردن صدای دیگران ممنوع است؛ بااین‌حال، نمونه‌های مشخصی از انواع رفتارهای ممنوع ارائه نمی‌کند.

«هیئت» خاطرنشان می‌کند که عملیات نفوذ را نهادهایی بسیار پیچیده، برخوردار از منابع قابل‌توجه و سازمان‌یافته انجام می‌دهند. هرگونه جزئیاتی که Meta و دیگر پلتفرم‌های رسانه‌های اجتماعی درباره شیوه مقابله با این نهادها ارائه کنند، از سوی همان نهادها تحلیل و برای تقویت عملیاتشان به‌کار گرفته خواهد شد. «هیئت» پیش‌تر در زمینه‌های مختلف، استفاده از این استدلال برای مخالفت با شفافیت را به چالش کشیده است (از جمله بنگرید به نمادهای مورد استفاده سازمان‌های خطرناک). بااین‌حال، با توجه به ماهیتِ به‌سرعت در حال تحول عملیات نفوذ، از جمله استفاده روزافزون آن‌ها از هوش مصنوعی، تدوین قواعد مشخص و مفصلِ در دسترس عموم مستلزم تغییرات مداوم خواهد بود و احتمالاً همان نهادهایی از این قواعد سوءاستفاده خواهند کرد که پلتفرم‌ها می‌کوشند کاربران خود را در برابر آن‌ها محافظت کنند. هرچند حقوق بین‌الملل بشر اقتضا می‌کند قواعد برای کاربران و نیز مجریان آن‌ها

روشن و در دسترس باشد، این اصل پلتفرم‌ها را ملزم نمی‌کند همه جزئیات روش‌های شناساییِ پشتیبان این قواعد را به‌صورت عمومی منتشر کنند؛ به‌ویژه اگر افشای آن جزئیات کارآمدی این سازوکارها را تضعیف کند.

بر این اساس، «هیئت» تشخیص می‌دهد که خطمشی «رفتار فریبکارانه» به‌اندازه کافی روشن است و با هنجارهای بین‌المللی حقوق بشر مطابقت دارد. بااین‌حال، می‌توان این خطمشی را با ارائه توضیحاتی درباره دسته‌های رفتاریِ ممنوع و نیز روشن کردن این نکته بهبود داد که اقدامات اجرایی بر مبنای شبکه انجام می‌شوند، نه بر مبنای محتوا.

خطمشی یکپارچگی حساب نیز نمونه‌ها یا جزئیات کافی درباره ممنوعیت‌های شبکه‌ای آن ارائه نمی‌کند. هرچند در اینجا نیز باید از ارائه اطلاعات مربوط به اقدامات اجرایی به بازیگران تهدیدگر خودداری کرد، «هیئت» بر نگرانی‌های پیشین خود تأکید می‌کند که قواعد عمومی Meta درباره غیرفعال‌سازی حساب پراکنده و متناقض‌اند (بنگرید به مسدودسازی حساب به‌دلیل هدف قرار دادن چهره‌های عمومی).

IV. هدف مشروع

هرگونه محدودیت بر آزادی بیان نیز باید یک یا چند مورد از اهداف مشروعِ ذکرشده در میثاق بین‌المللی حقوق مدنی و سیاسی را دنبال کند؛ از جمله حمایت از حقوق و حیثیت دیگران.

این نخستین بار است که «هیئت» بررسی می‌کند آیا خطمشی «رفتار فریبکارانه» هدفی مشروع را دنبال می‌کند یا خیر. در توضیح مبنای خطمشی «رفتار فریبکارانه» آمده است که این خطمشی رفتاری را هدف می‌گیرد که «با هدف فریب دادن Meta یا جامعه ما یا دور زدن اقدامات اجرایی ذیل استانداردهای جامعه» انجام می‌شود. ممنوع کردن چنین فعالیتی هدفی مشروع را دنبال می‌کند؛ زیرا می‌کوشد از حق کاربران برای دریافت اطلاعات حمایت کند و همزمان آنان را در برابر کلاهبرداری محافظت کند. اگرچه هدف مجاز حمایت از «امنیت ملی» که در ماده ۱۹(۳) آمده است مستقیماً درباره اهداف Meta صدق نمی‌کند، زیرا Meta یک دولت نیست، «هیئت» می‌پذیرد که این شرکت خصوصی منفعتی مشروع و مشابه دارد تا مانع استفاده بازیگران بدخواه از پلتفرم‌هایش برای حمله به نهادهای عمومی شود (ماده ۱۹(۳) میثاق بین‌المللی حقوق مدنی و سیاسی (ICCPR)).

خطمشی «یکپارچگی حساب» شرایطی را مشخص می‌کند که در آن Meta حساب کاربران را در صورت نقض دیگر خطمشی‌های شرکت و تلاش برای گریز از مجازات‌های ناشی از آن نقض‌ها محدود یا غیرفعال می‌کند. بنابراین، هرگونه ارزیابی درباره هدف مشروع باید بر نقض‌های اصلی خطمشی متمرکز باشد که از نظر Meta غیرفعال‌سازی حساب را توجیه کرده‌اند. پس از آن می‌توان مجازاتِ غیرفعال‌سازی حساب را از نظر ضرورت و تناسب ارزیابی کرد.

V. ضرورت و تناسب

بر اساس ماده ICCPR 19(3)، ضرورت و تناسب ایجاب می‌کند که محدودیت های بیان «باید برای تحقق کار حفاظت‌شان مناسب باشند؛ آنها باید ابزاری با کمترین مداخله در میان ابزارهایی باشند که ممکن است کار محافظت‌شان را انجام دهند؛ آنها باید متناسب با منافعِ مورد حفاظت باشند» (نظر عمومی‌شماره 34، پاراگراف 34).

ممنوعیت رفتار فریبکارانه هماهنگ در خطمشی «رفتار فریبکارانه» Meta، شبکه‌های پیچیده‌ای از بازیگران سازمان‌یافته و برخوردار از منابع قابل‌توجه را هدف می‌گیرد. Meta در پاسخ به پرسش‌های «هیئت» توضیح داد که برای غیرفعال‌سازی یک شبکه به‌دلیل رفتار فریبکارانه هماهنگ، معیار اثباتی بالایی در نظر گرفته می‌شود؛ بر اساس گزارش‌های تهدید Meta، از سال

۲۰۱۷ تا زمان اعلام این پرونده در سال ۲۰۲۶، کمتر از ۳۰۰ شبکه حذف شده‌اند. پس از آنکه کارکنان یا الگوریتم‌ها نشانه‌هایی از رفتار فریبکارانه هماهنگ احتمالی را علامت‌گذاری کردند، تیم‌های متخصص طیف گسترده‌ای از نشانه‌های فنی، رفتار شبکه مشکوک حساب‌ها و کاربران و محتوای منتشرشده را ارزیابی می‌کنند و سپس درباره نقض شدن یا نشدن خطمشی تصمیم می‌گیرند. اکثریت «هیئت» پس از بررسی مطالبی که Meta به اشتراک گذاشته است، قانع شده‌اند که از نظر نظری، اگر این آستانه مطابق توضیحات احراز شود، غیرفعال‌سازی حساب‌های درگیر احتمالاً ضروری و متناسب خواهد بود و برای حمایت از حقوق دیگران، راهکار کم‌مداخله‌تر دیگری وجود نخواهد داشت.

بااین‌حال، این امر لزوماً به این معنا نیست که خطمشی در تک‌تک موارد به‌شیوه‌ای متناسب اعمال خواهد شد. از آنجا که حساب مورد بحث در این پرونده بر اساس خطمشی «رفتار فریبکارانه» غیرفعال نشده است، «هیئت» نمی‌تواند ضرورت و تناسب این غیرفعال‌سازی مشخص را بر اساس آن خطمشی ارزیابی کند. افزون بر این، بسیاری از نشانه‌های رفتاری‌ای که Meta برای شناسایی رفتار فریبکارانه هماهنگ به آن‌ها اتکا می‌کند، منحصر به بازیگرانی نیستند که این خطمشی هدف می‌گیرد. حساب‌های تازه‌ساخته‌شده، هویت‌های مستعار، نشانه‌های جغرافیایی مخدوش‌شده با ابزارهای دور زدن محدودیت‌ها و انتشار هماهنگ و فشرده، از ویژگی‌های فعالان، روزنامه‌نگاران و مدافعان حقوق بشر نیز هستند که زیر سرکوب دولتی فعالیت می‌کنند و برای آنان ناشناس ماندن پیش‌شرط امنیت است (A/HRC/29/32). هرگاه سامانه‌های اقدام اجرایی بر اساس رفتار شبکه‌های خصمانه تنظیم شوند، خطر جدی طبقه‌بندی نادرست اعتراض سازمان‌یافته به‌عنوان هماهنگی ساختگی وجود دارد؛ این خطر بیش از همه متوجه کاربرانی است که در عمل کمترین دسترسی را به راهکارهای جبرانی دارند.

اگر رفتار فریبکارانه هماهنگ احراز شود، مجازات‌های خفیف‌تر و اقدامات میانی، به‌جای غیرفعال‌سازی حساب، برای رسیدگی به آن کافی نخواهند بود. تعلیق حساب و صدور هشدار، بازیگران تهدیدگر سازمان‌یافته‌ای را که منابع قابل‌توجهی برای اجرای عملیات نفوذ صرف کرده‌اند، از ادامه فعالیت بازنخواهد داشت. اگرچه اجرای این خطمشی به‌ضرورت کندتر و زمان‌برتر از نظارت صرفاً محتوامحور است، اقدامات میانی در این زمینه مناسب نیستند. اگر پیش از انجام تحقیق، به‌طور خودکار برچسب هشداری روی محتوا قرار گیرد و به کاربران اعلام کند که محتوا ممکن است بخشی از عملیات نفوذ باشد، مرتکبان فرصت خواهند یافت اجزای شبکه خود را تغییر کاربری دهند. افزون بر این، اقدام میانی علیه حساب‌ها پیش از آنکه تیم‌های متخصص بتوانند به‌طور کامل بررسی کنند آیا معیارهای نقض خطمشی درباره آن‌ها صدق می‌کند یا خیر، ممکن است به شمار نامتناسبی از نتایج مثبت کاذب بینجامد و پیامدهای جدی‌ای برای کاربرانی داشته باشد که به‌اشتباه هدف اقدام قرار گرفته‌اند.

از نظر اقلیت اعضای «هیئت»، اگرچه رویکردی که Meta می‌گوید در پیش گرفته، معقول به نظر می‌رسد، نگرانی‌های عملی‌ای وجود دارد. اثربخشی رویکرد اعلام‌شده شرکت به کافی بودن این آستانه، وجود تضمین‌هایی برای اجرای غیرتبعیض‌آمیز و منابع اختصاص‌یافته به چنین تحقیقاتی بستگی دارد. این در حالی است که گزارش‌هایی از افزایش فشار بر چندین حوزه کاری مرتبط که در این فعالیت نقش دارند، از جمله «برنامه شرکای مورد اعتماد»، منتشر شده است. بنابراین، Meta باید اطلاعات بیشتری را به‌صورت عمومی منتشر کند (یا این اطلاعات را در اختیار «هیئت» و پژوهشگران قرار دهد) تا اثربخشی رویکرد کنونی آن در اجرای خطمشی‌های مبتنی بر نقش‌آفرین یا رفتار، به تفکیک زبان، کشور و منطقه، به‌طور مستقل ارزیابی شود.

«هیئت» یافته Meta را می‌پذیرد که هیچ‌یک از حساب‌های این پرونده بخشی از شبکه رفتار فریبکارانه هماهنگ نبوده‌اند؛ تشکیل چنین شبکه‌ای مستلزم ارتباط با کاربران دیگری است که به‌صورت هماهنگ فعالیت می‌کنند. بااین‌حال، مشخص شد حسابی که محتوای نخست را منتشر کرده بود، هم فریبکارانه بوده و هم با حساب‌های دیگری ارتباط داشته است که همان کاربر اداره می‌کرد و به‌دلیل نقض خطمشی‌های Meta غیرفعال شده بودند. Meta این موضوع را تنها پس از طی چند مرحله کشف کرد: کاربری که محتوا را گزارش داده بود، به تصمیم Meta برای باقی گذاشتن آن در پلتفرم اعتراض کرد؛ سپس همان کاربر به «هیئت» درخواست تجدیدنظر داد؛ «هیئت» پرونده را انتخاب کرد و از Meta خواست توضیح دهد چرا حساب غیرفعال شده است. این روند نشان می‌دهد که با وجود تخصص بالای پژوهشگران رفتار فریبکارانه هماهنگ و فناوری پیشرفته‌ای که در اختیار دارند، ممکن است در شیوه کشف چنین محتواهایی توسط Meta و ارجاع آن‌ها برای بررسی تخصصی، کاستی‌هایی وجود داشته باشد. برخی نشانه‌ها یا الگوهایی که انسان‌ها قادر به تشخیص آن‌ها هستند، ممکن است بعداً برای تحقیقات مربوط به اصالت حساب‌ها مفید باشند. بنابراین، «هیئت» پیشنهاد می‌کند مسیرهای ارتباطی بهبود یابند تا محتوای بالقوه مشکوک شناسایی و برای بررسی مطرح شود.

خطمشی «رفتار فریبکارانه» همچنین ایجاد، استفاده یا ادعای استفاده از حساب‌های فریبکارانه را برای سوءاستفاده از سامانه‌های گزارش‌دهی Meta به‌منظور آزار، ارعاب یا خاموش کردن صدای دیگران ممنوع می‌کند. «هیئت» خاطرنشان می‌کند که شبکه‌های هماهنگ حساب‌ها یکی از مجاری اصلی آزار طیفی از اهداف، از جمله فعالان، روزنامه‌نگاران، زنان و گروه‌های اقلیت هستند. اگرچه خطمشی «قلدری و آزار» بخش عمده‌ای از رفتارهای نگران‌کننده را ممنوع می‌کند، ممنوعیت خطمشی «رفتار فریبکارانه» به سوءاستفاده از سامانه گزارش‌دهی محدود است و دیگر شکل‌های آزار هماهنگ را دربر نمی‌گیرد. «هیئت» از ارجاع پرونده یا درخواست نظر مشورتی درباره خطمشی استقبال می‌کند تا بتواند بررسی کند Meta چگونه از چنین کارزارهای آزار هماهنگی پیشگیری می‌کند و به آن‌ها پاسخ می‌دهد.

در مورد خطمشی «یکپارچگی حساب»، ضرورت و تناسب هر مورد غیرفعال‌سازی باید با توجه به نقض اصلی خطمشی که موجب آن شده است ارزیابی شود. بخش مربوط از خطمشی که Meta در این پرونده به آن استناد کرده، حساب‌هایی را هدف می‌گیرد که هدف انحصاری یا اصلی آن‌ها نقض خطمشی‌های Meta است. اگر این شرط احراز شود، غیرفعال‌سازی احتمالاً اقدامی ضروری و متناسب خواهد بود، مشروط بر آنکه اجرای خطمشی یکسان، عینی و بدون تبعیض باشد. همچنین، درباره حساب‌هایی که برای جایگزینی حساب‌های غیرفعال‌شده به‌دلیل نقض خطمشی‌ها ایجاد شده‌اند، می‌توان فرض کرد که غیرفعال‌سازی‌های پیشین نیز ضروری و متناسب بوده‌اند و امکان درخواست تجدیدنظر درباره آن‌ها وجود داشته است. بنابراین، قواعد این خطمشی که ایجاد حساب جایگزین را ممنوع می‌کنند، به‌عنوان اقدامی برای حفظ یکپارچگی عمل می‌کنند و مانع بی‌اثر شدن تصمیم‌های پیشین می‌شوند. «هیئت» درباره این پرسش اصلی که آیا حذف حساب‌های پیشین موجه بوده است یا خیر، داوری یا اظهارنظری نمی‌کند؛ بلکه صرفاً این اصل را می‌پذیرد که نباید به افراد اجازه داده شود با ایجاد حسابی جدید، حذف حساب قبلی را دور بزنند.

VI. *شفافیت*

Meta باید انتشار منظم آمار تعداد حساب‌هایی را بررسی کند که به‌دلیل انتشار اطلاعات گمراه‌کننده حذف شده‌اند؛ این گزارش‌دهی باید حساب‌هایی را نیز دربر گیرد که عضو هیچ شبکه‌ای نبوده‌اند. گزارش‌های تهدید خصمانه Meta در حال حاضر فقط به حساب‌هایی می‌پردازند که به‌دلیل نقض‌های مبتنی بر رفتار شبکه‌ای حذف شده‌اند. به نظر نمی‌رسد گزارش عمومی و منظمی درباره غیرفعال‌سازی حساب‌ها به‌دلیل انتشار محتوای نادرست یا گمراه‌کننده خارج از شبکه‌ها وجود داشته باشد. برای پژوهشگران و جامعه مدنی مفید است بدانند چه تعداد حساب به‌دلیل انتشار عمدی محتوای گمراه‌کننده حذف می‌شوند؛ ارائه این اطلاعات کمک قابل‌توجهی به نهادهای دخیل در عملیات نفوذ نخواهد کرد.

۶. تصمیم هیئت نظارت

«هیئت» تصمیم Meta مبنی بر باقی گذاشتن پست نخست را تأیید می‌کند، اما تصمیم Meta مبنی بر باقی گذاشتن پست دوم را لغو می‌کند.

توصیه‌ها

اقدامات اجرایی

۱. برای افزایش تبادل اطلاعات و بهبود تحقیقات درباره عملیات نفوذ، Meta باید سامانه‌ای برای تبادل نشانه‌ها میان تیم‌های متخصص در رفتار فریبکارانه هماهنگ و بازبینان رسیدگی‌کننده به حجم بالای محتوا ایجاد کند. این سامانه باید:

- در چارچوب هرگونه اعلام وضعیت بر اساس «پروتکل خطمشی بحران» فعال شود؛ این امر کشورهایی را دربرمی‌گیرد که در معرض خطر ویژه تأثیرپذیری از عملیات نفوذ خارجی یا داخلی قرار دارند.
- به تیم‌های بررسی رفتار فریبکارانه هماهنگ امکان دهد نشانه‌های محتوایی را علامت‌گذاری کنند تا بازبینان بتوانند محتوای مفید برای تحقیقات این تیم‌ها را به سطوح بالاتر ارجاع دهند. محتوایی که در چارچوب این سامانه علامت‌گذاری شده است نباید حذف شود، مگر آنکه مشخص شود بخشی از کارزار رفتار فریبکارانه هماهنگ است یا خطمشی دیگری را نقض می‌کند.
- به بازبینانِ رسیدگی‌کننده به حجم بالای محتوا اجازه دهد الگوهایی را که در بررسی‌های موردی با آن‌ها روبه‌رو می‌شوند و ممکن است به بررسی جامع‌تری نیاز داشته باشند، به‌صورت پیش‌دستانه گزارش کنند.

«هیئت» زمانی این توصیه را اجراشده تلقی می‌کند که Meta اسنادی ارائه دهد که نحوه فعال شدن این سامانه جدید را با جزئیات توضیح دهند و جریان‌های جدید تبادل اطلاعات را امکان‌پذیر سازند.

***یادداشت فرآیند اجرایی:**

- تصمیم‌های «هیئت نظارت» توسط گروهی مشتمل بر پنج «عضو» اخذ می‌شوند و بر اساس رأی اکثریت کل «هیئت» تصویب می‌گردند. تصمیم‌های «هیئت» الزاماً بیان‌کننده نظرات همه «اعضا» نیست.

- بر اساس منشور «هیئت نظارت» می‌تواند به درخواست‌های تجدیدنظر کاربرانی رسیدگی کند که Meta محتوایشان را حذف کرده است؛ به درخواست‌های تجدیدنظر کاربرانی که محتوای باقی‌مانده در پلتفرم را گزارش کرده‌اند؛ و نیز به تصمیم‌هایی که Meta به آن ارجاع می‌دهد (ماده ۲، بخش ۱ منشور). «هیئت» دارای اختیارات الزام‌آور برای تأیید یا لغو تصمیم‌های محتوایی Meta است (ماده ۳ منشور، بخش ۵؛ ماده ۴ منشور). «هیئت» می‌تواند توصیه‌های غیرالزام‌آوری ارائه کند که Meta باید به آن‌ها پاسخ دهد (ماده ۳ منشور، بخش ۴؛ ماده ۴). وقتی Meta برای اقدامی بر پایه توصیه‌ها متعهد شود، «هیئت» بر پیاده‌سازی آن نظارت می‌کند.

- درخصوص تصمیم این پرونده، تحقیق مستقلی ازطرف «هیئت» انجام گرفت. Duco Advisors، شرکتی مشاوره‌ای که بر تعاملات ژئوپولیتیک، اعتماد و امنیت و فناوری تمرکز دارد، در این زمینه به «هیئت» کمک کرد. خدمات تخصصی زبانی این پروژه را شرکت Lionbridge Technologies, LLC ارائه کرده است. متخصصان این شرکت به بیش از ۳۵۰ زبان تسلط دارند و در ۵۰۰۰ شهر در سراسر جهان فعالیت می‌کنند.