

رأي استشاري بشأن السياسة
حول برنامج **Meta** للتحقق
الشامل

2.....	I. الملخص التنفيذي
5.....	II. طلب مقدم من Meta
6.....	III. نظام Meta للتحقق الشامل
6.....	تفسير Meta لأسباب استخدام نظام التحقق الشامل
7.....	آلية عمل نظام التحقق الشامل
8.....	المراجعة الثانوية للاستجابة المبكرة (ERSR)
12.....	المراجعة الثانوية العامة (GSR)
13.....	التحقق الشامل وتقارير الإغفاء من الإنفاذ
14.....	IV. إطار عمل تحليل المجلس
14.....	المعايير الدولية لحقوق الإنسان
16.....	قيم شركة Meta
16.....	V. تقييم نظام التحقق الشامل
17.....	اتساع نطاق العمل لخدمة أهداف متعددة ومتناقضة تسمح بظهور المحتوى المخالف
19.....	الوصول غير المتكافئ إلى السياسات التقديرية والإنفاذ
21.....	المنضمون إلى البرنامج يتجاوزون السعة الاستيعابية
22.....	الإخفاق في تتبع أدوات القياس الأساسية لتقييم البرنامج وإدخال التحسينات عليه
23.....	الافتقار إلى الشفافية والمراجعة بشأن البرنامج وطريقة عمله
23.....	استنتاجات حول التحقق الشامل
24.....	VI. توصيات بشأن الإنفاذ
24.....	توصيات حوكمة نظام منع الأخطاء الذي يستند إلى الكيان
25.....	المستخدمون الذين يجب تضمينهم في أنظمة منع الأخطاء التي تستند إلى الكيان
26.....	ضرورة أن يكون متخذو القرار مؤهلين ومخولين لاتخاذ القرارات التي تحترم الحقوق
26.....	إرشادات لإنشاء وحوكمة القوائم لأنظمة منع الأخطاء التي تستند إلى الكيان
27.....	إرشادات الاحتفاظ بقوائم أنظمة منع الأخطاء التي تستند إلى الكيان وتدقيقها
28.....	ضرورة تمييز بعض الكيانات التي تتلقى حماية إضافية بشكل علني
29.....	توصيات حوكمة نظام منع الأخطاء الذي يستند إلى المحتوى
29.....	المحتوى الذي ينبغي تحديده ووضع الأولوية عليه لأنظمة منع الأخطاء التي تستند إلى المحتوى
29.....	التصحیحات الفنية
30.....	توصيات عامة لحوكمة نظام منع الأخطاء
30.....	تخفيف الأضرار بعد اكتشاف المحتوى المخالف
31.....	التأكد من توفر القدرة على تقديم الطعون
31.....	التعلم والتحسين
32.....	VII. توصيات بشأن الشفافية

عقب الإعلان في شهر أكتوبر 2021 عن برنامج Meta للتحقق الشامل في صحيفة وول ستريت جورنال، وافق مجلس الإشراف على طلب مُقدم من الشركة بشأن إجراء مراجعة لبرنامج التحقق الشامل وتقديم توصيات تتعلق بكيفية إضفاء تحسينات عليه. ويعتبر هذا الرأي الاستشاري بشأن السياسة هو وسيلتنا للرد على هذا الطلب. فهو يحلّ عملية التحقق الشامل في ضوء التزامات Meta تجاه حقوق الإنسان والقيم المعلنة، كما أنه يثير تساؤلات مهمة حول كيفية تعامل Meta مع أهم مستخدميها. عندما بدأ المجلس دراسة هذا الرأي الاستشاري بشأن السياسة، شاركت Meta أنها، في ذلك الوقت، تقوم بتنفيذ نحو 100 مليون محاولة إنفاذ على المحتوى كل يوم. وفي ظل وجود هذا الكم الهائل، حتى لو استطاعت Meta اتخاذ قرارات دقيقة بنسبة 99٪ تتعلق بالمحتوى، فإنها لا تزال ترتكب مليون خطأ يوميًا. وفي هذا الصدد، وفي الوقت الذي يتعين على نظام مراجعة المحتوى أن يعامل جميع المستخدمين معاملة منصفة، يستجيب برنامج التحقق الشامل لتحديات أوسع نطاقًا تتعلق بالإشراف على كميات هائلة من المحتوى.

ووفقًا لشركة Meta، فإن اتخاذ قرارات بشأن المحتوى على هذا النطاق يعني في بعض الأحيان إزالة محتوى لا ينتهك سياساتها عن طريق الخطأ. ويهدف برنامج التحقق الشامل إلى معالجة هذا الأمر من خلال توفير مراحل إضافية من المراجعة البشرية لمنشورات معينة تُقرر بشكل مبدئي أنها تخالف قواعد الشركة. عندما ينشر المستخدمون المدرجون في قوائم Meta للتحقق الشامل مثل هذا المحتوى، لا تتم إزالة المحتوى على الفور كما هو الحال بالنسبة إلى معظم الأشخاص، ولكن يتم تركه، في انتظار الخضوع لمراجعة بشرية إضافية. وتشير Meta إلى هذا النوع من التحقق الشامل باسم "المراجعة الثانوية للاستجابة المبكرة" (ERSR). في أواخر عام 2021، شرعت Meta في توسيع نطاق برنامج التحقق الشامل ليشمل منشورات معينة تم وضع علامة عليها لتخضع لمراجعة إضافية استنادًا إلى المحتوى نفسه، بدلاً من هوية الشخص الذي قام بنشرها. تشير Meta إلى هذا النوع من التحقق الشامل باسم "المراجعة الثانوية العامة" (GSR).

في إطار المراجعة التي نجريها، وجدنا العديد من العيوب في برنامج Meta للتحقق الشامل. في الوقت الذي أُخبرت فيه Meta المجلس أن الهدف من برنامج التحقق الشامل هو الوفاء بالتزامات Meta تجاه حقوق الإنسان، وجدنا أن البرنامج قد تمت هيكلته بصورة أكبر لتبديد مخاوف الأنشطة التجارية. ويدرك المجلس أن Meta تعتبر نشاطًا تجاريًا، ولكن مع توفير حماية إضافية لبعض المستخدمين الذين يتم اختيارهم بناءً على اهتمامات النشاط التجاري، فإن برنامج التحقق الشامل يتيح إبقاء بعض عناصر المحتوى لفترة أطول، رغم أنها كانت ستزال لولا هذا البرنامج، وهو الأمر الذي قد يؤدي إلى حدوث ضرر. وقد اكتشفنا كذلك إخفاق Meta في تتبع البيانات المتعلقة بما إذا كان نظام التحقق الشامل قد أسفر عن اتخاذ قرارات أكثر دقة أم لا، وأعربنا عن قلقنا إزاء الافتقار إلى الشفافية فيما يتعلق بالبرنامج.

ولمواجهة هذا الأمر، بادر المجلس بتقديم عدة توصيات إلى Meta. يتعين على أي نظام تم تصميمه لمنع الأخطاء أن يعطي الأولوية للتعبير نظرًا لأهميته بالنسبة لمجال حقوق الإنسان، بما في ذلك التعبير الذي يحظى باهتمام عام. بينما تتحرك Meta نحو تحسين عملياتها لجميع المستخدمين، يتعين على الشركة اتخاذ خطوات لتخفيف الضرر الناجم عن المحتوى الذي تم تركه لحين إجراء المراجعة الإضافية، وزيادة مستوى الشفافية فيما يتعلق بأنظمتها بشكل جذري.

أهم النتائج

يدرك المجلس أن حجم ومستوى تعقيد المحتوى المنشور على فيسبوك و Instagram ينطوي على تحديات تتعلق بإنشاء أنظمة تدعم Meta في الوفاء بالتزاماتها المتعلقة بحقوق الإنسان، ومع ذلك، فإن نظام التحقق الشامل بشكله الحالي، تشوبه العيوب في مجالات رئيسية يتعين على الشركة معالجتها:

عدم المساواة في معاملة المستخدمين. يمنح برنامج التحقق الشامل مستخدمين محددين قدرًا أكبر من الحماية عن غيرهم. إذا تبين أن منشورًا لمستخدم مدرج ضمن قوائم Meta للتحقق الشامل ينتهك قواعد الشركة، فإن المنشور يظل على المنصة في انتظار الخضوع لمراجعة إضافية. وبعد ذلك، تقوم Meta بتطبيق سياساتها بالكامل، بما في ذلك الاستثناءات والأحكام التي تستند إلى السياق، على المنشور، مما يزيد من فرصة بقاءه على المنصة. وعلى العكس من ذلك، تقل فرص وصول محتوى المستخدمين العاديين إلى المراجعين الذين يمكنهم تطبيق قواعد Meta بالكامل. إن هذا النوع من عدم المساواة في المعاملة يثير القلق بشكل خاص في ظل عدم

وجود معايير شفافة بشأن قوائم Meta للتحقق الشامل. وعلى الرغم من وجود معايير واضحة تحكم ضم شركاء الأعمال وقادة الحكومات، فإن المستخدمين الذين ينشرون محتوى يُرجح أن يكون مهمًا من منظور حقوق الإنسان، مثل الصحفيين ومنظمات المجتمع المدني، لديهم مسارات أقل وضوحًا للوصول إلى البرنامج.

الإزالة المتأخرة للمحتوى المخالف. عند تحديد محتوى وارد من مستخدمين مدرجين في قوائم التحقق الشامل على أنه ينتهك قواعد Meta، يظل المحتوى متاحًا بشكل كامل على المنصة أثناء خضوعه لعملية المراجعة الإضافية. أخبرت Meta المجلس بأن عملية المراجعة قد تستغرق، في المتوسط، أكثر من خمسة أيام للتوصل إلى قرار بشأن المحتوى الوارد من مستخدمين مدرجين في قوائم التحقق الشامل. ويعني هذا أن برنامج التحقق الشامل يتسبب في ترك المحتوى الذي تم تحديده على أنه ينتهك قواعد Meta على فيسبوك و Instagram أثناء تحقيقه لانتشار كبير ويمكن أن يتسبب في حدوث ضرر. ونظرًا لأن حجم المحتوى المحدد للتحقق الشامل قد يتجاوز قدرة Meta على المراجعة، فقد أصبحت لدى البرنامج حالات متراكمة مما يؤخر اتخاذ القرارات.

الإخفاق في تتبع أدوات القياس الأساسية. لا ترصد أدوات القياس التي تستخدمها Meta حاليًا لقياس مدى فعالية برنامج التحقق الشامل لجميع المخاوف الأساسية. على سبيل المثال، لم تُقدم Meta للمجلس معلومات توضح قيامها بتتبع ما إذا كانت القرارات التي اتخذتها من خلال برنامج التحقق الشامل تتسم بدقة أكثر أم أقل مقارنة بالياتها العادية لمراقبة الجودة. وبدون ذلك، يصعب معرفة ما إذا كان البرنامج يحقق أهدافه الأساسية المتمثلة في إصدار قرارات صحيحة بشأن الإشراف على المحتوى، أو قياس ما إذا كان برنامج التحقق الشامل يمنح Meta وسيلة لتحديد عن سياساتها.

الافتقار إلى الشفافية بشأن كيفية عمل برنامج التحقق الشامل. يشعر المجلس بالقلق أيضًا إزاء المعلومات المحدودة التي قدمتها Meta إلى العامة وللمستخدمين بشأن برنامج التحقق الشامل. لا تُبلغ Meta حاليًا المستخدمين بأنهم مدرجون في قوائم التحقق الشامل ولا تقوم بمشاركة إجراءات البرنامج المتعلقة بإنشاء هذه القوائم وتدقيقها بشكل علني. فمن غير الواضح، على سبيل المثال، ما إذا كان يتم الاحتفاظ في قوائم التحقق الشامل بالكيانات التي تكرر نشر محتوى مخالف على أساس ملفها الشخصي. ويحول هذا الافتقار إلى الشفافية دون فهم المجلس والعامة للنتائج الكاملة المترتبة على البرنامج.

توصيات مجلس الإشراف

لوفاء بالتزامات Meta في مجال حقوق الإنسان ومعالجة هذه المشاكل، يجب اتباع نهج مختلف تمامًا لهيكل البرنامج الذي يصحح الأخطاء الأعلى تأثيرًا على فيسبوك و Instagram. وقد قدّم المجلس 32 توصية في هذا المجال، وتم تلخيص العديد منها أدناه.

في الوقت الذي تسعى فيه Meta إلى تحسين إشرافها على المحتوى لجميع المستخدمين، ينبغي عليها إعطاء الأولوية للتعبير الذي يتسم بأهمية على صعيد حقوق الإنسان، بما في ذلك التعبير الذي يحظى بأهمية خاصة بالنسبة إلى العامة. ينبغي إعطاء الأولوية للمستخدمين الذين يُحتمل أن يقدموا هذا النوع من التعبير لإدراجهم ضمن قوائم الكيانات التي تحظى بمراجعة إضافية مقارنة بشركاء أعمال Meta. يجب مراجعة المنشورات التي يتم نشرها من قبل هؤلاء المستخدمين في دفق عمل منفصل، بحيث لا تكون في وضع منافسة مع شركاء أعمال Meta على الموارد المحدودة. على الرغم من أن عدد المتابعين يمكن أن يشير إلى اهتمام عام بتعبير المستخدم، يجب ألا تكون شهرة المستخدم أو عدد المتابعين هو المعيار الوحيد للحصول على حماية إضافية. إذا كرر المستخدمون الذين تم تضمينهم بسبب أهميتهم التجارية نشر محتوى مخالف، فلا ينبغي أن يستفيدوا من ميزة الحماية الخاصة.

تعزيز مستوى الشفافية بشأن التحقق الشامل وكيفية عمله بشكل جذري. يجب على Meta الوقوف على مستوى أداء أدوات القياس الرئيسية المتعلقة ببرنامج التحقق الشامل الخاص بها وتدقيقها ونشرها حتى تتمكن من تحديد ما إذا كان البرنامج يعمل بفعالية أم لا. يجب على الشركة وضع معايير علنية وواضحة للإدراج في قوائم التحقق الشامل الخاصة بها، ويجب أن تتيح للمستخدمين الذين يستوفون هذه المعايير إمكانية التقدم بطلب لإضافتهم إليها. ويجب أيضًا تمييز حسابات بعض فئات الكيانات التي تتمتع بالحماية من خلال برنامج التحقق الشامل، بما في ذلك الجهات الحكومية والمرشحون السياسيون وشركاء الأعمال، على أنها حسابات عامة. سيتيح

ذلك للعامّة مساءلة المستخدمين المتميزين بشأن مدى وفاء الكيانات المحمية بالتزاماتها نحو اتباع القواعد. وعلاوة على ذلك، ونظرًا لعدم إمكانية تصعيد نحو ثلث محتوى نظام Meta للتحقق الشامل إلى المجلس اعتبارًا من شهر مايو وحتى يونيو 2022، يتعين على Meta التأكد من إمكانية تقديم طعون إلى المجلس بشأن محتوى برنامج التحقق الشامل وكل أنواع المحتوى الأخرى التي تخضع للوثائق الحاكمة.

تقليل الضرر الناجم عن المحتوى الذي تم تركه على المنصة أثناء المراجعة المحسّنة. يجب إزالة المحتوى الذي حدد التقييم الأولي لشركة Meta أنه مخالف ويتسم بخطورة شديدة أو إخفائه أثناء إخضاعه لمراجعة إضافية. لا ينبغي السماح ببقاء مثل هذا المحتوى على المنصة وحصده لمشاهدات لمجرد أن الشخص الذي قام بنشره هو شريك أعمال أو أحد المشاهير. ولضمان اتخاذ القرارات في أسرع وقت ممكن، يتعين على Meta استثمار الموارد اللازمة بما يتناسب مع قدرتها على مراجعة المحتوى الذي تقرر أنه يحتاج إلى مراجعة إضافية.

II. طلب مقدم من Meta

1. **علم مجلس الإشراف لأول مرة ببرنامج التحقق الشامل عام 2021 عند اتخاذ قراره بشأن حالة تعليق حسابات الرئيس الأمريكي الأسبق دونالد ترامب.** وعلى الرغم من أن Meta لم تذكر برنامج التحقق الشامل في إحالتها الأولية أو المواد التي أرسلتها إلى المجلس، فقد وصفت برنامج التحقق الشامل في ردها على سؤال المجلس بشأن مدى حصول الحساب على أي معاملة مختلفة. وكجزء من قراره الصادر في مايو 2021، قدم المجلس توصيتين تتعلقان ببرنامج التحقق الشامل:
 - "إنتاج معلومات أكثر لمساعدة المستخدمين على فهم وتقييم العملية ومعايير تطبيق استثناء الأهمية الإخبارية، بما في ذلك كيفية تطبيقه على حسابات الشخصية المؤثرة."
 - "يجب على الشركة أيضًا أن تشرح بوضوح الأساس المنطقي والمعايير والعمليات الخاصة بمراجعة التحقق الشامل، والإبلاغ عن معدلات الخطأ النسبي للقرارات التي يتم اتخاذها من خلال التحقق الشامل مقارنة بإجراءات الإنفاذ العادية."
2. في سبتمبر 2021، كشفت صحيفة وول ستريت جورنال عن وثائق من إنتاج الموظفة السابقة وناقدة الشركة فرانسيس هوجين. وقد وصف **تقرير الصحيفة** التحقق الشامل بأنه إعفاء للمستخدمين الأكثر تأثيرًا على مستوى تقنيات Meta من عمليات الإشراف على المحتوى المعتادة. وقد ذكرت صحيفة إنديبندينت ما قالتها فرانسيس هوجين من أن الشركة "كذبت" على المجلس بشأن التحقق الشامل "بشكل متكرر" خلال حالة ترامب. وقد كشفت وثائق داخلية من شركة Meta نشرتها صحيفة وول ستريت جورنال أن بعض موظفي الشركة يعتبرون أن ممارسات "الإدراج في القائمة البيضاء" للتحقق الشامل "غير قابلة للدفاع عنها علنًا". وعلى نحو مشابه، وفقًا لصحيفة وول ستريت جورنال، فقد مُنح المستخدمين المستفيدين من نظام التحقق الشامل في ذلك الوقت مهلة "تصحيح ذاتي" لمدة 24 ساعة لتعديل المحتوى المخالف أو إزالته وبالتالي تجنب أي عقوبات تفرضها Meta.
3. في 21 سبتمبر 2021، بعد نشر مقالات وول ستريت جورنال، دعا المجلس شركة Meta إلى الالتزام بالشفافية فيما يتعلق بنظام التحقق الشامل. وفي اليوم التالي، قدمت Meta إيجازًا إلى المجلس بشأن التحقق الشامل. وقد **خُصص المجلس** إلى أن "فريق فيسبوك المكلف بتقديم المعلومات لم يكن صريحًا بشكل كامل في ردوده بشأن التحقق الشامل. وفي بعض الحالات، أخفقت فيسبوك في تقديم المعلومات ذات الصلة إلى مجلس الإشراف، بينما في حالات أخرى لم تكن المعلومات التي قدمتها للمجلس كاملة."
4. بعد فترة وجيزة من مطالبة المجلس لشركة Meta بتوفير المزيد من الشفافية بشأن التحقق الشامل، قدمت Meta طلب الرأي الاستشاري بشأن السياسة هذا. بعد تلخيص النظام بشكل موجز، وصفت Meta التحقق الشامل بأنه برنامج "يوفر مستويات إضافية من المراجعة لأنواع معينة من المحتوى التي تقرر أنظمتنا الداخلية تمييزها كعناصر محتوى مخالفة (من خلال الأتمتة أو المراجعة البشرية)، وذلك لمنع أو تقليل أخطاء الإشراف". وتعرّف Meta النتائج الإيجابية الخاطئة بأنها إزالة محتوى لا ينتهك سياسات المحتوى التي تحدد الأمور المسموح بها وغير المسموح بها على فيسبوك و Instagram عن طريق الخطأ.
5. طرحت Meta على المجلس الأسئلة الثلاثة التالية:

نظرًا للتعقيدات التي تكتنف الإشراف على المحتوى على نطاق واسع، كيف ينبغي على فيسبوك الموازنة بين رغبتها في تطبيق معايير المجتمع بعدالة وموضوعية وبين حاجتنا للمرونة والدقة واتخاذ قرارات مرتبطة بالسياق ضمن نظام التحقق الشامل؟

ما التحسينات التي يجب أن يجريها فيسبوك على الطريقة التي نحكم بها نظام المراجعة الثانوية للاستجابة المبكرة ("ER") الخاص بنا لإنفاذ معايير مجتمعنا بشكل عادل مع تقليل احتمالية الإفراط في الإنفاذ، والحفاظ على مرونة الأعمال، وتعزيز الشفافية في عملية المراجعة؟

ما المعايير التي يجب أن يستخدمها فيسبوك لتحديد من يتم تضمينه في المراجعة الثانوية للاستجابة المبكرة (ER) ومنحه الأولوية كأحد العوامل العديدة لأداة تصنيف التحقق الشامل لدينا للمساعدة على ضمان المساواة في الوصول إلى هذا النظام وتنفيذه؟

6. قبل المجلس طلب Meta في 21 أكتوبر 2021. وبعد قبول المجلس للطلب، طرح بعض الأسئلة على Meta. وبلغت الأسئلة التي طرحها المجلس على Meta 74 سؤالاً. قدّمت الشركة إجابة كاملة عن 58 سؤالاً، وإجابة جزئية عن 11 سؤالاً، ولم تجب عن خمسة أسئلة. واستغرقت Meta شهرًا للرد على بعض هذه الأسئلة.

7. تلقى المجلس أيضًا 87 تعليقًا عامًا حول هذا الرأي الاستشاري بشأن السياسة: تسعة تعليقات من دول آسيا والمحيط الهادئ والأوقيانوس، وتعليقين من جنوب ووسط آسيا، و12 تعليقًا من أوروبا، وثلاثة تعليقات من أمريكا اللاتينية والكاريبي، وثلاثة تعليقات من الشرق الأوسط وشمال أفريقيا، وثلاثة تعليقات من دول أفريقيا الواقعة جنوب الصحراء الكبرى، و55 تعليقًا من الولايات المتحدة وكندا. لقراءة التعليقات العامة المقدمة حول هذا الرأي الاستشاري بشأن السياسة، يرجى النقر [هنا](#). علاوة على ذلك، أقام المجلس أربع ورش عمل إقليمية ركزت على برنامج التحقق الشامل.

8. استنادًا إلى تحليله لهذه المعلومات، والأبحاث المستقلة، وتفاعل المساهمين، يجيب المجلس الآن عن أسئلة Meta ويقدم تقييمه لنظام التحقق الشامل. أخبرت Meta المجلس أيضًا أنها أجريت تغييرات جوهرية على برنامج التحقق الشامل خلال العام الماضي. وينظر المجلس إلى هذه التغييرات على أنها، على الأقل جزئيًا، محاولة للرد على الانتقادات العامة للبرنامج. ويعتمد تفسير المجلس للبرنامج وتحليله له على ما ذكرته Meta عن طريقة عمل البرنامج حاليًا. ومع ذلك، يشير المجلس في بعض الأحيان إلى فهمه للممارسات السابقة لأنها تحدد الجوانب المحتمل للمخاطر المتكررة.

9. استكشف المجلس ما إذا كان البرنامج يسعى، في الممارسة العملية، إلى معالجة الآثار السلبية والتخفيف من حدتها وفقًا لمسؤوليات Meta تجاه حقوق الإنسان. ويتضمن هذا التحليل، الذي يستند إلى المعايير الدولية لحقوق الإنسان والقيم والالتزامات المعلنة لشركة Meta، أسئلة مهمة حول كيفية تعامل Meta مع مستخدميها الأكثر نفوذًا وقوة، والسماح بتدفق المحتوى عبر منصاتها، وتوفير المعلومات للجمهور حول إجراءاتها.

iii. نظام Meta للتحقق الشامل

تفسير Meta لأسباب استخدام نظام التحقق الشامل

10. ينشئ المستخدمون على فيسبوك وInstagram المليارات من عناصر المحتوى كل يوم. وتقوم Meta باستمرار بالإشراف على المحتوى؛ أو فحصه، وتقييمه، واتخاذ إجراءات حياله استنادًا إلى سياسات المحتوى التي وضعتها الشركة. تُعرف هذه السياسات على فيسبوك باسم معايير المجتمع، وعلى Instagram باسم إرشادات المجتمع.

11. وفقًا لشركة Meta، يمثل الإشراف على هذا القدر الهائل من المحتوى تحديات كبيرة، ويزيل المراجعون والأنظمة الآلية في بعض الأحيان عن طريق الخطأ عناصر محتوى لا تنتهك سياسات Meta. تشير Meta إلى

هذه القرارات باسم النتائج الإيجابية الخاطئة. وتعتبر النتائج السلبية الخاطئة شكلاً من أشكال النقص في الإنفاذ وتشير إلى محتوى ينتهك سياسات Meta لكن المراجعة تقرر أنه غير مخالف. يتضمن نقص الإنفاذ أيضاً المحتوى المخالف الذي لم تكتشفه المراجعة التلقائية أو البشرية، وخيارات تصميم النظام التي تسمح ببقاء المحتوى المخالف ظاهراً بعد خضوعه لمراجعة أولى.

12. لا يعالج نظام التحقق الشامل إلا حالات فرط الإنفاذ، أو النتائج الإيجابية الزائفة. ومن خلال هذا النظام، تُوجّل Meta اتخاذ أي إجراء إنفاذ بشأن محتوى محدد تقرر مبدئياً أنه مخالف للسماح بإجراء مراجعة إضافية محتملة بهدف تجنب النتائج الإيجابية الخاطئة.

13. وصفت Meta نظام التحقق الشامل بأنه استراتيجية لمنع الأخطاء تسمح للشركة بالموازنة بين حماية حرية المستخدمين من النتائج الإيجابية الخاطئة والحاجة إلى إزالة المحتوى المخالف سريعاً. كجزء من طلب الرأي الاستشاري بشأن السياسة، سلّطت Meta الضوء على تضمين "الصحفيين الذين ينقلون الأخبار من مناطق الصراعات وقادة المجتمع الذين يعززون الوعي بحالات الكراهية أو العنف"، بالإضافة إلى الجهات الفاعلة المدنية حيث "يزيد اهتمام الأشخاص بمشاهدة ما يقوله قادتهم".

14. يتضمن النظام أيضاً مستخدمين تصفهم Meta بأنهم "شركاء أعمال". وقد خصصت Meta لهؤلاء الشركاء نقاط اتصال خاصة بهم. ووفقاً للشركة، يتضمن هؤلاء المستخدمون "الهيئات الصحفية وناشري الأخبار والفنانين الترفيهيين والموسيقيين والفنانين والمبدعين والمنظمات الخيرية". ويتفهم المجلس أن هذه الفئة تشمل مستخدمين يحتمل أن يدروا أموالاً للشركة، إما من خلال شراكات الأعمال الرسمية أو بسبب جذب المستخدمين إلى المنصة والحفاظ على تفاعلهم بها. ويعلم المجلس أن مسمى "شركاء الأعمال" يحتمل أيضاً أن يشمل الشركات الكبرى والأحزاب والحملات السياسية والمشاهير.

15. أخبرت Meta المجلس أنها تضيف "شركاء الأعمال" إلى نظام التحقق الشامل لمنع عمليات الحذف الخاطئة التي تحد من قدرة المستخدمين والمعلنين على الوصول إلى جمهورهم وعلائقهم، وما قد يترتب على تلك الأخطاء من تأثير على الاقتصاد والسمعة الخاصة بالشركة. وبالنسبة إلى هؤلاء المستخدمين، تهدف Meta إلى تجنب "التجارب السلبية لكل من شركاء أعمال فيسبوك والعدد الكبير من المستخدمين الذين يتابعونهم".

16. صرحت Meta أنها تفضل نقص الإنفاذ مقارنة بفرط الإنفاذ عندما يتعلق الأمر بالمحتوى في نظام التحقق الشامل، حيث إنه "في المشهد التجاري الحالي، يعتبر توسيع نطاق نظام التحقق الشامل (منع النتائج الإيجابية الخاطئة) بشكل عام أكثر أهمية من تقليل كلفة التحقق الشامل [أي مشاهدات المحتوى المخالف]. وذلك بسبب مفهوم الرقابة". ويفسر المجلس هذا على أنه يعني، لأسباب تجارية، أن التعامل مع "مفهوم الرقابة" قد يستحوذ على الأولوية مقارنة بمسؤوليات حقوق الإنسان الأخرى ذات الصلة بالإشراف على المحتوى.

آلية عمل نظام التحقق الشامل

17. يتم تطبيق العمليات العادية للإشراف على المحتوى في Meta على معظم المستخدمين. وعندما تتبين مخالفة المحتوى لسياسات Meta، تتخذ Meta أحد إجراءات الإنفاذ. ويتضمن ذلك حذف المحتوى ووضع شاشات تحذيرية، حسب نوع انتهاك السياسة. قد تؤدي بعض الانتهاكات أيضاً إلى فرض عقوبات على مستوى الحساب، مثل تعليق الحساب أو إنهائه. ومع ذلك، يتلقى المحتوى معاملة مختلفة في بعض الحالات، كما يحدث من خلال نظام التحقق الشامل.

18. تستخدم Meta مصطلح التحقق الشامل للإشارة إلى برنامج منع النتائج الإيجابية الخاطئة. وهو برنامج يوفر للمحتوى طبقات إضافية من المراجعة قبل اتخاذ إجراءات الإنفاذ. وقد يتم خلال هذه المراجعة المحسنة تطبيق السياسات المتبعة عند التصعيد فقط، والتي لا يتم تطبيقها إلا بواسطة فرق متخصصة في Meta. تتضمن هذه السياسات الأهمية الإخبارية واستثناءات روح السياسة وجميع القواعد التي حددت Meta أنها تتطلب سيقاً إضافياً لإنفاذها. يتم تفعيل عمليات مراجعة التحقق الشامل في مجموعتين من الظروف.

19. أولاً، يوفر التحقق الشامل مراجعة بشرية إضافية مضمونة للمحتوى المنشور بواسطة كيانات معينة محددة متى قامت تلك الكيانات بنشر محتوى تبيّن أنه يستحق اتخاذ إجراءات إنفاذ بموجب سياسات المحتوى في Meta. تطلق Meta على هذه العملية اسم المراجعة الثانوية للاستجابة المبكرة أو ERSR. يشير مسمى "كيان" إلى أي شيء على فيسبوك أو Instagram يمكنه نشر المحتوى، مثل الصفحات والملفات الشخصية على فيسبوك وحسابات Instagram. يمكن أن تمثل الكيانات أفراداً أو مجموعات أو مؤسسات. تنشئ Meta وتحفظ بقوائم للكيانات التي قررت أنها تستحق التمتع بالمزايا التي تكفلها المراجعة الثانوية للاستجابة المبكرة. وهذا يعني أنه إذا قام كيان مؤهل بنشر محتوى تحدد أنه ينتهك معايير المجتمع أو إرشادات المجتمع، فلن تتم إزالته وفق الإجراءات المطبقة على المستخدمين العاديين، وإنما سيتم إرساله ليخضع لمستويات إضافية من المراجعة. ولأن المراجعة الثانوية للاستجابة المبكرة تستند إلى القوائم، لا يتمتع بمزاياها إلا مجموعة محددة من المستخدمين المحددين سلفاً.

20. يوفر الجزء الثاني من نظام التحقق الشامل مراجعة إضافية لعناصر محتوى معينة تبيّن أنها تنتهك سياسات Meta، بغض النظر عن هوية المستخدم الذي نشرها. وتطلق Meta على هذا النوع اسم المراجعة الثانوية العامة، أو GSR. متى تقرر أن أي عنصر محتوى منشور بواسطة أي كيان على المنصة على أنه ينتهك سياسة Meta، وذلك من خلال الوسائل التلقائية أو المراجعة البشرية، تستخدم Meta عملية تلقائية تعرف باسم "أداة تصنيف التحقق الشامل" لتقوم على الفور بتحليل العوامل المتعددة وتحديد ما إذا كان ينبغي إخضاع المحتوى لمراجعة إضافية، وكيفية تحديد أولويته في قائمة الانتظار مقارنة بعناصر المحتوى الأخرى التي تنتظر نفس النوع من المراجعة. وفقاً لشركة Meta، ولأن هذا النظام يستند إلى خصائص المحتوى، يكون المحتوى الذي ينشره أي مستخدم على فيسبوك أو Instagram مؤهلاً للمراجعة الثانوية العامة. تم تطبيق المراجعة الثانوية العامة عام 2021، ويدرك المجلس أنه قد تم تطويرها وتنفيذها، إلى حد ما، على مستوى المنصة استجابة للانتقادات التي واجهت المراجعة الثانوية للاستجابة المبكرة، بما في ذلك ما كشفته هوجين.

21. قد يحدث الاكتشاف الأولي للمحتوى في كلا النوعين من التحقق الشامل على النحو الذي قد يؤهل للمراجعة إما بشكل استباقي، من خلال أنظمة Meta التلقائية بعد نشر المحتوى، أو كرد فعل استجابة لبلاغات المستخدمين. تتضمن إجراءات الإنفاذ التي قد تؤهل للمراجعة التحقق الشامل حذف المحتوى وتطبيق شاشة تحذيرية، حسب نوع انتهاك السياسة. ونظراً لأن معظم انتهاكات السياسة قد تؤدي إلى فرض عقوبات على مستوى الحساب، مثل التعليق والإنهاء، فإن هذه الأنواع من الإنفاذ تتأثر أيضاً. يتم تطبيق التحقق الشامل على فيسبوك و Instagram، باستثناء بعض أنواع المحتوى غير المؤهلة حالياً للبرنامج (على سبيل المثال، مقاطع ريلز والبودكاست). وفقاً لشركة Meta، "فإن 10% من المحتوى العادي الذي يخضع خلافاً لذلك لإجراءات إنفاذ السلامة غير مؤهلة لمراجعة التحقق الشامل حالياً".

22. خلال الوقت التالي لتحديد المحتوى المؤهل للتحقق الشامل (من خلال المراجعة الثانوية العامة أو المراجعة الثانوية للاستجابة المبكرة) للإنفاذ وقبل خضوعه لعملية مراجعة إضافية، يظل الوصول إلى هذا المحتوى متاحاً بالكامل على المنصة، حتى إذا كان التقييم الأولي ينص على أن المحتوى ينتهك معايير أو إرشادات المجتمع.

23. يتفهم المجلس أنه إذا وفرت Meta المزيد من المشرفين، فسيحصل عدد أكبر من عناصر المحتوى الموجودة في قائمة انتظار مراجعة التحقق الشامل على مراجعة بشرية إضافية. مع هذا، فقد اختارت Meta ألا تضمن المراجعة البشرية الإضافية إلا للمحتوى الذي يخضع لعملية المراجعة الثانوية للاستجابة المبكرة، وهي النظام المخصص للكيانات المؤهلة. ولم تستثمر Meta الموارد اللازمة ليحصل كل المحتوى المحدد من خلال المراجعة الثانوية العامة على مراجعة بشرية إضافية. وبينما تختلف مسارات المراجعة لهاتين الآليتين، كما هو موضح أدناه، إذا وجد أي مراجع في أي مرحلة من مراحل العملية أن المحتوى لا ينتهك سياسات Meta، فستنتهي عملية المراجعة، ويظل المحتوى على المنصة.

المراجعة الثانوية للاستجابة المبكرة (ERSR)

24. تقول Meta أنها تضيف الكيانات إلى قوائم المراجعة الثانوية للاستجابة المبكرة من خلال تخصيص "علامة" لهم ترتبط بطبيعة وحساسية الكيان. وتتوافق علامات معينة مع مختلف قوائم المراجعة الثانوية للاستجابة المبكرة. وتقول Meta أنها تضيف علامة المراجعة الثانوية للاستجابة المبكرة إلى الكيانات المقابلة للفئات

التالية: (1) الجهات المدنية والحكومية؛ (2) الأحداث العالمية المهمة؛ (3) المؤسسات الإعلامية والشركات والمجتمعات والمبدعين، بما في ذلك المعلنين؛ (4) الفئات التي تعرضت لفرط الإنفاذ تاريخياً؛ (5) الجهات القانونية والتنظيمية أو الكيانات التي قد يمثل اتخاذ إجراء خاطئ بشأنها خطورة قانونية على شركة Meta، على سبيل المثال في سياق دعوى قضائية جارية؛ (6) الكيانات التي يخضع المحتوى الخاص بها للمراجعة، ويعني ذلك الحالات التي يؤدي فيها أي قرار يتخذه المراجع إلى تقويض المداورات الجارية أو يمثل خطراً على Meta. وفقاً لشركة Meta، بخلاف العوامل التي تستخدمها لتحديد ما إذا كان الكيان يندرج تحت أي من الفئات سالفة الذكر، مثل الإنفاق الإعلاني أو سجل الإنفاذ، تحدد أهلية الخضوع للمراجعة الثانوية للاستجابة المبكرة أيضاً من خلال تقييم تأثير ارتكاب خطأ محتمل في الإنفاذ على الشركة على صعيد قيادة الشركة التي من شأنها المشاركة في إيجاد حل. بعبارة أخرى، يتمثل أحد المبادئ الأساسية للمراجعة الثانوية للاستجابة المبكرة في تجنب استفزاز الأشخاص الذين لديهم وسائل تتيح لهم إشراك كبار المديرين التنفيذيين مباشرة أو إثارة جدل عام قد يضطر هؤلاء المديرين التنفيذيين إلى علاجه.

25. أُخبرت Meta المجلس أنها تعمل حالياً على دمج قوائم المراجعة الثانوية للاستجابة المبكرة وتحديثها. في السابق، كانت قوائم Meta تتوافق مع مستوى التصعيد المطلوب لإنفاذ سياسات المحتوى ضد كيان معين. ووفقاً لشركة Meta، تخضع الآن كل الكيانات المؤهلة للمراجعة الثانوية للاستجابة المبكرة لنفس عملية المراجعة. وقد تتضمن هذه العملية التصعيد بشكل تقديري إلى أعلى مستويات الشركة.

26. أُخبرت Meta المجلس أنها وضعت خلال الربع الثاني من عام 2022 معايير عامة لإضافة الكيانات وإزالتها من قوائم المراجعة الثانوية للاستجابة المبكرة، وعمليات جديدة لعمليات التدقيق الدوري والإشراف الداخلي. ولم تقدم Meta تفاصيل عن هذه العمليات، والإجراءات التي قد تؤدي إعادة تقييم أحد الكيانات المدرجة وإزالته من القائمة. أوضحت Meta، بوجه عام، أن صلاحية العلامات التي تضع الكيان على قوائم المراجعة الثانوية للاستجابة المبكرة تنتهي بمرور عام، ويجب نظرياً تقييم الكيانات المؤهلة وإضافة العلامات إليها من جديد. وفقاً لشركة Meta، هذا المنطق يغطي نظرياً الكيانات التي تندرج تحت الفئات التالية: الجهات القانونية والتنظيمية؛ والأحداث العالمية المهمة؛ المؤسسات الإعلامية؛ والشركات، والمجتمعات، والمبدعين؛ والفئات التي تعرضت لفرط الإنفاذ تاريخياً؛ والكيانات التي تم تصعيدها لمراجعة سباقية أعلى. وأشارت Meta إلى استثناء حالتين من قاعدة انتهاء صلاحية العلامات بعد مرور عام. أولاً، لا توجد مدة صلاحية افتراضية لعلامات الكيانات المدرجة في فئة الجهات المدنية والحكومية. ثانياً، قد تحصل علامات الكيانات المدرجة في الفئات سالفة الذكر على مدة استحقاق أقصر وفقاً لتقدير شركة Meta.

27. عندما تضع المراجعة التلقائية أو البشرية علامة على عنصر محتوى لأي من الكيانات المؤهلة لاتخاذ إجراء إنفاذ بشأنه، لا يتم اتخاذ أي إجراء إنفاذ، ويتم إرسال المحتوى بدلاً من ذلك ليخضع إلى مراجعة محسنة بواسطة مُشرف بشري. يتولى تنفيذ أول مستوى من مستويات المراجعة المحسنة بواسطة فريق داخلي تطلق عليه Meta اسم "فريق السوق الإقليمي". ويضم هذا الفريق موظفين من Meta ومتعاقدين يتمتعون بمعارف سياقية ولغوية إضافية عن سوق جغرافي معين. إذا قرر أحد مراجعي فريق السوق أن المحتوى غير مخالف، تنتهي العملية، ويظل المحتوى على المنصة.

28. مع هذا، إذا وجد مراجع السوق أن المحتوى ينتهك سياسات Meta، يظل المحتوى على المنصة أثناء تصعيده إلى فريق تطلق عليه Meta اسم "فريق الاستجابة المبكرة" لإخضاعه لمراجعة أخرى. وفقاً لشركة Meta، لدى هذا الفريق "خبرة أكبر بشأن السياسة وقدرة على تضمين سياق إضافي."

29. يُسمح أيضاً لفريق الاستجابة المبكرة بسلطة تقديرية أوسع من مشرفي المحتوى الآخرين في Meta ويمكنه تطبيق سياسات المحتوى التي "تتطلب معلومات أو سياق إضافي ليتم إنفاذها". وغالباً ما تميز Meta سياسات المحتوى هذه بعلامة تعجب صفراء داخل كل معيار من معايير المجتمع، كما هو موضح أدناه. على سبيل المثال، في نهاية معيار مجتمع فيسبوك بشأن العنف والتحرير، تحظر Meta "التهديد بارتكاب أعمال عنف ضد مسؤولي إنفاذ القانون". ووفقاً لشركة Meta، فإن تحديد مدى الاحتفاظ بعناصر المحتوى التي قد تشكل انتهاكاً لبنود السياسة التي تستند إلى السياق أو إزالته لا يتم إلا بواسطة فريق يُسمح له بتضمين سياق إضافي، مثل "فريق الاستجابة المبكرة".



For the following Community Standards, we require additional information and/or context to enforce:

Do not post:

- Violent threats against law enforcement officers.
- Violent threats against people accused of a crime. We remove this content when we have reason to believe that the content is intended to cause physical harm.
- Coded statements where the method of violence or harm is not clearly articulated, but the threat is veiled or implicit. Meta looks at the below signals to determine whether there is a threat of harm in the content.

30. يمكن أيضًا لفريق الاستجابة المبكرة تطبيق ما تطلق عليه Meta اسم استثناءات "الأهمية الإخبارية" و"روح السياسة"، والتي تتيح للمحتوى المخالف البقاء على المنصة لأن Meta ترى أنه يصب في المصلحة العامة أو ترى أنه على الرغم من انتهاك نص السياسة، فإنه لا ينتهك مقاصدها. يرى المجلس أيضًا أن هذا التقدير يتسع ليشمل فرض عقوبات على مستوى الحساب. ومع هذا، كما كشفت Meta، لا يتمتع فريق الاستجابة المبكرة خبرة لغوية أو إقليمية ويعتمد على الترجمات والمعلومات السياقية المقدمة من فريق السوق الإقليمي المعني لتقييم المحتوى.

31. حين قدمت Meta إيجابًا إلى المجلس، بلغ ما تم تصعيده إلى المراجعين الذين يمكنهم تطبيق هذه السياسات والاستثناءات المرتبطة بالسياق من خلال التحقق الشامل 0.01% تقريبًا من كل المحتوى الذي يلزم اتخاذ إجراء إنفاذ بشأنه بموجب سياسة Meta. من المضمون وصول المحتوى المنشور بواسطة مستخدمين مدرجين في قوائم المراجعة الثانوية للاستجابة المبكرة إلى هؤلاء المراجعين قبل اتخاذ أي إجراء إنفاذ؛ ولا يمكن إزالة المحتوى أو وضع شاشة تحذيرية من خلال المراجعة التلقائية، أو مسؤولي المراجعة البشرية على نطاق واسع، أو مراجعي فريق السوق. طوال الوقت الذي ينتظر فيه المحتوى الخاضع لمراجعة التحقق الشامل القرار النهائي، يظل موجودًا على المنصة، حيث يمكن للمستخدمين الإعجاب به ومشاركته بحرية.

32. إذا تبين أن المحتوى مخالف بعد مراجعة فريق الاستجابة المبكرة للمحتوى، فقد تتخذ Meta إجراء الإنفاذ المناسب، مثل إزالة المحتوى أو وضع شاشة تحذيرية. ومع هذا، فقد تقوم Meta أيضًا بتصعيد القرار أكثر. يتفهم المجلس أن إجراءات التصعيد في هذه المرحلة تكون تقديرية بدرجة كبيرة. وإذا وجد فريق الاستجابة المبكرة أن المحتوى يمثل "تفسيرًا متطرفًا لسياسات [Meta]" أو أنه "يشكل خطورة ملموسة على الشركة أو المجتمع، و/أو وجود حالة من عدم الاتفاق بين الجهات المعنية الداخلية بشأن الاستجابة المناسبة"، فقد يُجري فريق الاستجابة المبكرة مراجعة أخرى بالاشتراك مع فرق Meta الأخرى. وفقًا لشركة Meta، فإن مراجعات التصعيد هذه "تشمل آراء وإسهامات الخبراء المتخصصين في سياسة المحتوى و فرق السياسة العامة المحلية والاتصالات والشؤون القانونية" وقد تشمل إسهامات من فرق أخرى. بعد هذه المراجعة، قد يتم تصعيد المحتوى أكثر إلى قيادات الشركة قبل فرض أي إجراء إنفاذ.

33. علاوة على ذلك، أُخبرت Meta المجلس أنه إذا كانت "المشكلة تنطوي على مخاطر كبيرة لحجب الخدمة، أو مخاطر قانونية أو تنظيمية أو متعلقة بالسلامة، أو إذا كان الوقت محدودًا [أمامها] لاتخاذ إجراء، فسيقوم [فريق الاستجابة المبكرة] في حالات نادرة بتصعيد القرار مباشرةً إلى القيادة العليا العالمية." ذكرت Meta أنها تجري تقييم لمخاطر المساءلة ومدى إلحاح الحالة والتأثير الجيوسياسي ومخاطر حجب الخدمة وعدم الاتفاق بين الفرق الداخلية كعوامل لتصعيد هذه القرارات.

34. باختصار، قد تخضع عناصر المحتوى المنشورة بواسطة كيان مدرج على قائمة المراجعة الثانوية للاستجابة المبكرة لما يصل إلى خمس مراجعات قبل أن يصل إلى مرحلة الإنفاذ، حتى إذا وجد المراجعون مرة بعد مرة أنه ينتهك قواعد فيسبوك أو Instagram وقاموا بتصعيده عبر مختلف مسارات التحقق الشامل:

1. المراجعة الأولية من خلال الوسائل التلقائية أو المراجعة البشرية التي تحدد المحتوى الذي يحتاج إلى الإنفاذ بموجب سياسات Meta.

2. مراجعة فريق السوق الإقليمي.
3. مراجعة فريق الاستجابة المبكرة. وهذه هي المراجعة الأولى التي يمكن أن تجيز اتخاذ إجراء إنفاذ ضد المحتوى. وقد يطلب الفريق إجراء مراجعة محسنة لفريق الاستجابة المبكرة تضم فرقاً أخرى أو تتحول مباشرة إلى مراجعة القيادة العالمية.
4. المراجعة المحسنة لفريق الاستجابة المبكرة مع الخبراء المتخصصين، وفرق السياسة العامة والاتصالات والشؤون القانونية.
5. مراجعة القيادة العالمية. وهو تصعيد تقديري من فريق الاستجابة المبكرة استناداً إلى خطورة العواقب على الشركة.

تتوقف العملية ويظل المحتوى على المنصة إذا تبين أنه غير مخالف في أي مرحلة من مراحل المراجعة.

35. قد يستغرق مسار المراجعة الثانوية للاستجابة المبكرة عدة أيام. وفقاً لشركة Meta، يتراوح الهدف الداخلي لمراجعة التحقق الشامل التي يجريها فريق السوق بين 12 إلى 120 ساعة تبعاً لخطورة الانتهاك المحتمل. عملياً، ذكرت Meta أن متوسط الوقت اللازم لصدور قرار نهائي للمراجعة الثانوية للاستجابة المبكرة أكثر من خمسة أيام. بالنسبة إلى المحتوى الذي ينشره المستخدمون في الولايات المتحدة، تقول Meta إنها تستغرق "12 يوماً تقريباً في المتوسط لاتخاذ قرار". ويزيد متوسط الوقت اللازم لاتخاذ قرار في البلدان الأخرى عن نظيره في الولايات المتحدة. على سبيل المثال، يبلغ متوسط الوقت اللازم لاتخاذ قرار بالنسبة إلى أفغانستان وسوريا نحو 17 يوماً. في المعلومات التي قدمتها Meta إلى المجلس، بلغت أطول مدة لبقاء عنصر محتوى في قائمة انتظار المراجعة الثانوية للاستجابة المبكرة 222 يوماً. قدمت Meta عدة رسومات بيانية بهذه البيانات إلى المجلس، حيث عرضت متوسط الوقت المستغرق لاتخاذ قرار بين مارس 2021 وفبراير 2022، مع معدل الإلغاء وعدد المهام، أو عناصر المحتوى التي تمت مراجعتها، والبلدان المختلفة.

36. تقول Meta إن المحتوى الذي يستغرق وقتاً أطول في انتظار المراجعة تم تصنيفه على أنه أقل خطورة بموجب "إطار عمل خطورة الانتهاك" الخاص بالشركة. ويصنّف هذا النظام المحتوى تبعاً لمعيار المجتمع الذي أشارت المراجعة الأولى إلى انتهاك المحتوى له. يصنّف إطار عمل Meta كل معيار من معايير المجتمع تبعاً للضرر المحتمل الذي قد ينتج عن انتهاكات السياسة، وهو قرار قالت Meta إنها اتخذته استناداً إلى أبحاث أجرتها الشركة. على سبيل المثال، تعتبر الشركة أن الخطاب الذي يحض على الكراهية أكثر ضرراً من المحتوى غير المهم أو الاحتمالي، حيث يتم وضع الأولوية على الخطاب المحتمل أن يحض على الكراهية مقارنة بالمحتوى غير المهم أو الاحتمالي في قائمة انتظار المراجعة الثانوية للاستجابة المبكرة.

37. مع هذا، ذكرت صحيفة وول ستريت جورنال في سبتمبر 2021 أن نجم كرة القدم البرازيلي نيمار قد نشر على حساباته في فيسبوك وInstagram صوراً حميمية لشخص آخر دون موافقته. ووفقاً لتقرير نشرته صحيفة [الجاردريان](#)، ظل الفيديو منشوراً على الإنترنت لأكثر من يوم، و"وجدت مراجعة داخلية لمنشورات نيمار أن الفيديو قد حصد 56 مليون مشاهدة على فيسبوك وInstagram قبل إزالته" على الرغم من أنه يمثل انتهاكاً سافراً لسياسات المحتوى في شركة Meta. وفقاً لشركة Meta، يرجع سبب توفر المحتوى المخالف لفترة طويلة إلى "تأخر مراجعة المحتوى لتراكم العناصر المطلوب مراجعتها في ذلك الوقت".

38. يتمثل أحد أدوات القياس الأساسية التي أدرت Meta المجلس أنها تستخدمه كمسوغ لنظام التحقق الشامل وتقييم أدائه في "معدل الإلغاء". ويشير إلى النسبة المئوية لعناصر المحتوى التي تبين أنها غير مخالفة خلال مراجعة التحقق الشامل، مما أدى إلى إلغاء القرار الأولي وعدم اتخاذ إجراء الإنفاذ الذي تفرضه قواعد ضد المحتوى. قدمت Meta إلى المجلس العديد من الأرقام بشأن معدل الإلغاء بالنسبة إلى المحتوى في المراجعة الثانوية للاستجابة المبكرة. وفقاً لشركة Meta، تراوح معدل الإلغاء لفترات زمنية مختلفة على مدار العام الماضي بين 30 و90%. عندما يكون معدل الإلغاء منخفضاً، فإن المراجعة الثانوية للاستجابة المبكرة تحتفظ بعناصر محتوى أكثر على المنصة تبين في النهاية أنها مخالفة خلال الطبقات المتعددة لمراجعة التحقق الشامل. عندما يكون معدل الإلغاء مرتفعاً، فإن المراجعة الثانوية للاستجابة المبكرة تحفظ المزيد من عناصر المحتوى غير المخالف من الإنفاذ الخاطئ.

39. وفقًا لشركة Meta، "تحدث معظم المشاهدات فور نشر المحتوى، ولذلك فإن سرعة مراجعة القرارات وإزالة المحتوى سريعًا لهما دور مهم في منع حدوث الضرر." لهذا، يظل الوصول إلى المحتوى المخالف الخاضع للمراجعة الثانوية للاستجابة المبكرة متاحًا على المنصة طوال المدة التي يحتمل أن يحظى فيها بغالبية المشاهدات.

المراجعة الثانوية العامة (GSR)

40. تتمثل الآلية الثانية من آليات نظام التحقق الشامل بحسب شركة Meta في المراجعة الثانوية العامة (GSR). بينما يتم تطبيق المراجعة الثانوية للاستجابة المبكرة على كل المحتوى المنشور بواسطة كيانات مؤهلة معينة، تنطبق المراجعة الثانوية العامة على أي محتوى منشور على المنصة، بغض النظر عن القائم بالنشر، استنادًا إلى ما تحدده الخوارزميات.

41. تعتبر المراجعة الثانوية العامة نظامًا جديدًا نسبيًا. في خريف عام 2021، عندما كشفت فرانسيس هوجين التي أبلغت عن مخالفات فيسبوك عن معلومات تتعلق بالتحقق الشامل، يدرك المجلس أنها تشير إلى ما سبقت الإشارة إليه، والذي كان يستند بالكامل إلى الكيان الذي ينشر المحتوى. تقوم الشركة حاليًا بتضمين المراجعة الثانوية العامة في التحقق الشامل أيضًا. ذكرت Meta أن المراجعة الثانوية العامة كانت متوفرة للمحتوى المنشور بواسطة بعض المستخدمين في 2021 وتم تطبيقها بالكامل على المحتوى المنشور بواسطة كل المستخدمين في أوائل 2022.

42. لتحديد المحتوى الذي سيتم إرساله عبر مسار المراجعة الثانوية العامة قبل اتخاذ إجراء إنفاذ، تستخدم Meta خوارزمية (وهي مجموعة من القواعد التي يجب اتباعها بواسطة جهاز كمبيوتر من أجل مهمة معينة) تعرف باسم أداة تصنيف التحقق الشامل. تستند هذه الخوارزمية إلى الميزات التالية: "حساسية الموضوع (مدى رواج/حساسية الموضوع)، وشدة الإنفاذ (شدة إجراء الإنفاذ المحتمل)، واحتمالية وجود نتيجة إيجابية خاطئة، والوصول المتوقع، وحساسية الكيان." وبالتالي فإن حساسية الكيان أحد العوامل المؤثرة في كلا النظامين، على الرغم من أنها في المراجعة الثانوية للاستجابة المبكرة تمثل العامل الرئيسي بينما تمثل في المراجعة الثانوية العامة عامل واحد فقط من عدة عوامل. وقد ذكرت Meta أنها نظرت في تضمين عوامل إضافية وتوقع القيام بذلك في المستقبل.

43. وفقًا لشركة Meta، يجب أن يفى المحتوى بشرطين ليكون مؤهلاً للمراجعة الثانوية العامة. أولاً، يجب أن يتم تحديد حاجته إلى الإنفاذ (أي انتهاك أحد معايير أو إرشادات المجتمع) بواسطة المراجعة التلقائية أو البشرية. ثانيًا، يجب أن تمنحه أداة تصنيف التحقق الشامل أولوية عالية. في حالة استيفاء الشرطين، لا يتم اتخاذ إجراء إنفاذ بشأن المحتوى على الفور ويوضع بدلاً من ذلك في قائمة انتظار ليخضع لمراجعة بشرية إضافية بواسطة فريق من فرق السوق المحلي. وهي نفس فرق السوق التي تنفذ أيضًا أول مراجعة محسنة للمحتوى المنشور بواسطة الكيانات المؤهلة للمراجعة الثانوية للاستجابة المبكرة.

44. يتعذر على فرق السوق مراجعة كل المحتوى الذي يستحق الخضوع للمراجعة الثانوية للاستجابة المبكرة وكل المحتوى المدرج في قائمة الانتظار لإخضاعه لمراجعة ثانوية عامة محتملة. ولأن الكيانات المؤهلة المدرجة على قائمة المراجعة الثانوية للاستجابة المبكرة تضمن الخضوع للمراجعة، يجب على فرق السوق أن تخصص السعة الاستيعابية من المراجعين إلى هذا المحتوى أولاً. وفي حالة وجود أي سعة استيعابية متبقية، يراجع فريق السوق المعني محتوى المراجعة الثانوية العامة الذي تم تحديده بواسطة الخوارزمية. تراجع فرق السوق أيضًا عناصر محتوى معينة خارج نطاق برنامج التحقق الشامل، إلى جانب مهام أخرى يجب أن يضعوا الأولوية عليها.

45. لذلك، على الرغم من أن محتوى المراجعة الثانوية العامة قد يحظى بأولوية عالية من قبل خوارزمية أداة تصنيف التحقق الشامل على أنه يستحق مراجعة إضافية لأنه ربما تم تحديده على أنه نتيجة إيجابية خاطئة محتملة، فقد لا يكون لدى فريق السوق السعة الاستيعابية لمراجعته. في بعض الحالات، في حالة عدم توفر السعة الاستيعابية للمراجعة على مستوى فريق السوق، واختارت Meta الاستعانة بمصادر خارجية، فقد يتم إرسال بعض عناصر المحتوى المؤهلة للمراجعة الثانوية العامة لتخضع لمراجعة إضافية بواسطة مراجعين خارجيين. إذا

تمت مراجعة المحتوى المؤهل للمراجعة الثانوية العامة بواسطة أحد مراجعي السوق، يكون القرار نهائيًا في معظم الحالات. وإذا تبين أن المحتوى مخالف، فإنه يخضع بشكل عام للإنفاذ. إما إذا تبين أنه غير مخالف، فإنه يظل على المنصة. ومع هذا، إذا توفرت لدى فريق الاستجابة المبكرة سعة استيعابية إضافية بعد التزاماته بمراجعة كل المحتوى المؤهل للمراجعة الثانوية للاستجابة المبكرة قبل إزالته المحتملة، قد يقوم هذا الفريق بمراجعة المحتوى ذو الأولوية العالية المؤهل للمراجعة الثانوية العامة والذي وجد مراجع فريق السوق أنه مخالف قبل أن تبدأ Meta في إجراء الإنفاذ.

46. على غرار عملية المراجعة الثانوية للاستجابة المبكرة، فإن المحتوى الذي تم تقييمه في الأصل على أنه يشكل مخالفة لمعايير المجتمع وتم وضعه في قائمة انتظار المراجعة الثانوية العامة يظل على المنصة أثناء انتظار الخضوع لمراجعة إضافية. ومع هذا، وعلى خلاف المراجعة الثانوية للاستجابة المبكرة، فإن المحتوى الموجود في قائمة انتظار المراجعة الثانوية العامة انتظارًا للمراجعة لن يظل على المنصة إلى أجل غير مسمى. و"تنتهي مهلة" المحتوى الذي لم تتم مراجعته من قائمة انتظار المراجعة الثانوية العامة في نهاية المطاف. وعند حدوث ذلك، تعود Meta إلى قرار الإنفاذ الأولي دون أي مراجعة إضافية. مما يعني أنه يتم تطبيق الإجراء الذي كان سيتم تطبيقه، مثل إزالة المحتوى أو وضع شاشة تحذيرية عليه، بشكل متأخر دون أي مراجعة إضافية. وإذا لم يصل المراجعون إلى عنصر محتوى معين في قائمة انتظار المراجعة الثانوية العامة، فسيظل على المنصة لمدة تتراوح بين يومين إلى أربعة أيام قبل أن تزيله Meta من قائمة انتظار المراجعة وتطبق عليه إجراء الإنفاذ. في الوقت ذاته، تواصل أداة تصنيف التحقق الشامل تحديد عناصر محتوى أحدث وأكثر أولوية، وتعلق اتخاذ إجراءات إنفاذ ضدها، وتضيفها إلى قائمة انتظار المراجعة الثانوية العامة.

47. تتقيد الفعالية الإجمالية للمراجعة الثانوية العامة باختيار Meta للمقدار الذي ستوفره من السعة الاستيعابية للمراجعين لهذا النوع من المراجعة في كل من أسواقها. تتم مراجعة معظم المحتوى المؤهل للمراجعة الثانوية العامة بواسطة مراجع خارجي، أو مراجع ضمن فريق السوق، أو تنتهي مهلة وجوده في النظام. يعني ذلك أن معظم المحتوى المؤهل للمراجعة الثانوية العامة لا يصل مطلقًا إلى فريق الاستجابة المبكرة وبالتالي لن يصل على الإطلاق إلى مستوى المراجعة الذي يمكن أن يتم فيه تطبيق التحليل السياقي وسياسات التصعيد فقط واستثناءات السياسة.

48. تحسب Meta أيضًا معدل الإلغاء للمحتوى الذي يحصل على مراجعة التحقق الشامل من خلال مسار المراجعة الثانوية العامة. وقد قدمت Meta للمجلس معدلات مختلفة لهذا الرقم على مدار العام الماضي. حين قدمت Meta إيجازًا إلى المجلس في فبراير 2022، بلغ معدل الإلغاء للمراجعة الثانوية العامة نحو 80%. وقد قدمت Meta لاحقًا معلومات جديدة إلى المجلس، حيث أشارت أنه بين مارس ومايو 2022، بلغ معدل الإلغاء نحو 70%. وعلى الرغم من وجود تباين أيضًا، فقد كان تباين هذه الأرقام أقل بمرور الوقت. وقد وجد أن معظم المحتوى المؤهل للمراجعة الثانوية العامة الذي تحدد أنه مخالف في البداية لم يكن مخالفًا لأي من سياسات Meta عند إخضاعه لمراجعة ثانوية. مع انتهاء مهلة وجود المحتوى في قائمة انتظار المراجعة الثانوية العامة، من المحتمل بشكل كبير أن تتخذ Meta إجراءات إنفاذ تسفر عن عدد كبير من النتائج الإيجابية الخاطئة التي حددتها أداة تصنيف التحقق الشامل.

التحقق الشامل وتقارير الإعفاء من الإنفاذ

49. وصفت تقارير صحيفة وول ستريت جورنال التحقق الشامل بأنه نظام لإعفاء "كبار الشخصيات من إجراءات الإنفاذ المعتادة المتبعة في الشركة". كشفت Meta للمجلس أن لديها نظامًا يحجب بعض إجراءات الإنفاذ خارج نظام التحقق الشامل. تشير Meta إلى هذه الممارسة باسم "التصحیحات الفنية"، ووصفتها التقارير العامة باسم "قوائم السماح" و"القوائم البيضاء".

50. "التصحیحات الفنية" هي استثناءات تلقائية لإنفاذ سياسة المحتوى. ويعني ذلك أنها تتجاوز معظم محاولات المراجعة التلقائية والبشرية لتطبيق أحد إجراءات الإنفاذ لمجموعة محددة سلفًا من انتهاكات سياسة المحتوى. يتم تلقائيًا فحص كل عنصر محتوى تم تحديد حاجته إلى الإنفاذ لتحديد مدى أهليته لتطبيق أي "تصحیحات فنية".

51. إذا كان المحتوى يتمتع بحماية أي تصحيح، فسيتم إعفاؤه من إجراء الإنفاذ المحدد. كما أوضحت Meta، ينطبق "التصحيح الفني" فقط على كيان معين لانتهاك سياسة معين ولا يهدف إلى منع الإنفاذ لانتهاكات السياسة الأخرى. حين قدمت Meta إيجازًا إلى المجلس، ذكرت الشركة أنها تطبق نحو ألف تصحيح فني كل يوم. ولم تكشف Meta عن أعداد وأنواع الكيانات التي استفادت من "التصحيات الفنية".

52. وإذا لم يكن المحتوى يتمتع بحماية أي تصحيح، يتم فحص أهليته للتحقق الشامل. عند هذه النقطة، يتم تطبيق عمليات التحقق الشامل المعتادة لدى Meta لتحديد ما إذا كان المستخدم كيانًا مؤهلاً للمراجعة الثانوية للاستجابة المبكرة أو إذا كان المحتوى يحظى بأولوية بواسطة أداة تصنيف التحقق الشامل ليُدخل قائمة انتظار المراجعة الثانوية العامة.

53. ذكرت Meta أنها لا تطبق "التصحيات الفنية" بشكل أساسي على "مجموعتين من أنواع الانتهاك (المحتوى غير المهم أو الاحتيالي/السلوك الزائف وانتحال الشخصية)". أكدت Meta لاحقًا أنه اعتبارًا من 21 سبتمبر 2022، توجد أربعة "تصحيات فنية" نشطة وأن هذا أيضًا قد يتغير بمرور الوقت.

54. أخبرت Meta المجلس أن "عددًا محدودًا من 'التصحيات الفنية' لا يزال قائمًا، وتعترف [Meta] بوجود احتياج مستمر لها". وفقًا لشركة Meta، فإن هذه "التصحيات تساعد [Meta] على منع وقوع أخطاء الإنفاذ فيما يتعلق بالمحتوى أو الكيانات التي من غير المرجح بدرجة كبيرة أن تنتهك سياساتنا وتوجيه موارد المراجعة البشرية إلى الجوانب الأشد احتياجًا لها".

55. أقرت Meta بوجود أوجه قصور في ممارساتها السابقة بشأن التصحيات الفنية. أخبرت Meta المجلس أن "الافتقار إلى الحوكمة بشأن الممارسات في السابق، [...] أدى دون قصد إلى عدم خضوع بعض الكيانات للكثير من إجراءات الإنفاذ". ذكرت Meta أنه كان "بإمكان الفرق المختلفة تطبيق تصحيحات مختلفة على نفس الكيان بطريقة قد تسفر، عند دمج تلك التصحيات، عن عدم خضوع الكيان والمحتوى الذي ينشره لمجموعة واسعة من إجراءات الإنفاذ". ذكرت Meta أنه نظرًا لأن هذه الممارسة كان "نتيجة غير مقصودة لنظام غير مركزي، فإن [Meta] تتخذ خطوات لضمان وجود هيكل حوكمة بشأن استخدام قوائم التحقق الشامل".

التحقق الشامل في سياق الطلبات الحكومية لإزالة المحتوى

56. عندما تطلب الحكومات من Meta إزالة المحتوى، فقد تزيل Meta المحتوى لأنه ينتهك سياسات المحتوى التي تقرها الشركة. وقد تزيل أيضًا المحتوى أو تفرض عليه "حجبًا جغرافيًا" لأسباب قانونية، أو تقيّد الوصول إليه في مناطق معينة. أخبرت Meta المجلس أنها تضيف الكيانات إلى قوائم المراجعة الثانوية للاستجابة المبكرة ضمن برنامج التحقق الشامل لتوفير الحماية لهم ضد الإجراءات الخاطئة التي قد تشكل خطرًا قانونيًا على Meta، على سبيل المثال في سياق دعوى قضائية قائمة.

57. وفقًا لشركة Meta، يتم التعامل مع الطلبات الحكومية لإزالة المحتوى بواسطة فرق متخصصة يمكنها اتخاذ إجراء إنفاذ فوري بشأن المحتوى، سواء تم نشره بواسطة كيان مؤهل للمراجعة الثانوية للاستجابة المبكرة أم لا وبغض النظر عن احتمال أنه كان سيحظى بأولوية عالية بواسطة أداة تصنيف التحقق الشامل. بعبارة أخرى، تُجَب عمليات الإزالة الناتجة عن طلبات حكومية الامتيازات التي يكفلها نظام التحقق الشامل.

IV. إطار عمل تحليل المجلس

المعايير الدولية لحقوق الإنسان

58. في 16 مارس، 2021، أعلنت Meta عن [سياسة حقوق الإنسان الخاصة بالشركة](#)، حيث وضحت التزامها تجاه احترام الحقوق على نحو يتفق مع مبادئ الأمم المتحدة التوجيهية بشأن الأعمال التجارية وحقوق الإنسان (UNGPs). تُرسي مبادئ الأمم المتحدة التوجيهية بشأن الأعمال التجارية وحقوق الإنسان، التي أقرها مجلس حقوق الإنسان التابع للأمم المتحدة عام 2011، إطار عمل تطوعي لمسؤوليات الأنشطة التجارية فيما يتعلق

بحقوق الإنسان. تتضمن هذه الحقوق، "كحد أدنى، [...] تلك المنصوص عليها في الشريعة الدولية لحقوق الإنسان،" (المبدأ 12).

59. كشركة عالمية ملتزمة بمبادئ الأمم المتحدة التوجيهية بشأن الأعمال التجارية وحقوق الإنسان، ينبغي على Meta احترام المعايير الدولية لحقوق الإنسان أينما مارست أعمالها ومعالجة أي آثار ضارة بحقوق الإنسان (المبدأ 11). يعني ذلك أيضًا أنه ينبغي على شركة Meta "السعي إلى تجنب التسبب في حدوث آثار ضارة تتعلق بحقوق الإنسان أو التخفيف من حدتها عندما ترتبط ارتباطًا مباشرًا بعملياتها أو منتجاتها أو الخدمات التي تقدمها في إطار علاقاتها التجارية، حتى عندما لا تسهم هي في تلك الآثار" (المبدأ 13).

60. تنص مبادئ الأمم المتحدة التوجيهية بشأن الأعمال التجارية وحقوق الإنسان أيضًا على وجوب بذل الشركات للعناية الواجبة بحقوق الإنسان لتقييم الآثار الفعلية والمحملة والتصرف بناءً على النتائج التي تتوصل إليها (المبدأ 17). وللقيام بذلك بفعالية، يجب على الشركات مراقبة المؤشرات الكمية والنوعية ودمج الملاحظات الواردة من أصحاب المصلحة المتضررين (المبدأ 20).

61. من خلال الحالات التي يدرسها، يقيم المجلس الآثار المترتبة على حقوق الإنسان نتيجة لقرارات الإنفاذ. عندما تكشف هذه الحالات أن Meta تُحدث تأثيرًا سلبيًا، أو أنها لا تتخذ الخطوات اللازمة لتحديد ومراقبة والحد من الآثار السلبية على نطاق أوسع، يُصدر المجلس التوصيات التصحيحية المناسبة. في رأي استشاري بشأن السياسة، يركز المجلس مباشرةً على خيارات السياسة لدى Meta، بما في ذلك عمليات التطوير والإنفاذ، لتقييم وفاء الشركة بالتزاماتها نحو احترام الحقوق الواردة في مبادئ الأمم المتحدة التوجيهية بشأن الأعمال التجارية وحقوق الإنسان.

62. وبتطبيقه على التحقق الشامل، استكشف المجلس ما إذا كان البرنامج يخدم، في الممارسة العملية، معالجة الآثار السلبية على حقوق الإنسان والتخفيف من حدتها وفقًا لمسؤوليات Meta أم لا. كما فحص المجلس عن كثب أدوات القياس التي تستخدمها Meta لتحديد فعالية البرنامج، وما يعكسه ذلك عن أهداف الشركة.

63. وجد المجلس في تحليله أنه قد تتأثر مجموعة واسعة من الحقوق ببرنامج التحقق الشامل. يمكن تعزيز حرية التعبير، التي تشمل الحق في التماس المعلومات وتلقيها (المادة 19، العهد الدولي الخاص بالحقوق المدنية والسياسية؛ [التعليق العام رقم 34](#)، 2011، الفقرة 11)، إلى الحد الذي يعمل فيه التحقق الشامل على الحد من اتخاذ إجراءات إنفاذ ضد المحتوى الذي لا ينتهك سياسات المنصة. وهو ما يؤثر إيجابيًا على المستخدم الذي ينشر المحتوى والأشخاص الذين يرغبون في الوصول إلى هذا المحتوى.

64. يلاحظ المجلس أيضًا أن التحقق الشامل، نظريًا، يمكن أن يساعد على ضمان استفادة الأشخاص الذين يواجهون معوقات معينة عند ممارسة حقوقهم في حرية التعبير من طبقة الحماية الإضافية التي قد يوفرها البرنامج. إذ يمكن، على سبيل المثال، منع البلاغات الجماعية الموجهة بشأن المحتوى غير المخالف من خلال نظام لمنع أخطاء النتائج الإيجابية الخاطئة.

65. مع ذلك، قد تكون هذه التأثيرات الإيجابية محدودة إذا كان النظام مصممًا بالأساس لحماية أو إعطاء أولوية لتعبير الأشخاص الأقوياء بالفعل. ويلاحظ المجلس أيضًا أن برنامج التحقق الشامل يثير مخاوف تتعلق بعدم التمييز، حيث تحظى بعض الكيانات بحماية إضافية.

66. علاوة على ذلك، قد تُسهم الحماية التي يوفرها برنامج التحقق الشامل للمحتوى المخالف في خلق بيئة تحول دون قدرة المستهدفين بالمحتوى المخالف على التعبير. وقد يؤثر نطاق المحتوى المخالف الذي يظل على المنصة لوقت إضافي بشكل كبير على مجموعة متنوعة من حقوق الإنسان، وقد تتباين العواقب تبعًا لأوضاع المستخدمين المتأثرين. ومن المرجح أن يشعر الأفراد والفئات التي تواجه التهميش والتمييز بالآثار السلبية على حقوق الإنسان بشكل أكثر حدة.

67. يأخذ تحليل المجلس هذه المعايير في الحسبان. وتقر التوصيات التي يصدرها بشأن السياسة بالقيود التي تحد من قدرة Meta على الإشراف على المحتوى على نطاق واسع. إذا كان إشراف Meta على المحتوى المنشور

بواسطة كل المستخدمين أكثر دقة، لم تكن لاحتياج إلى برامج خاصة تقوم على كيانات مؤهلة للمساعدة في تعزيز احترامها لحقوق الإنسان.

قيم شركة Meta

68. تُرسي المعايير الدولية لحقوق الإنسان المبادئ لسياسات Meta وممارساتها. ومع هذا، يمكن لشركات وسائل التواصل الاجتماعي، في ظل هذه المعايير، أن تتبنى أساليب مختلفة لاحترام الحقوق. ويجب أن تسترشد شركة Meta بقيمها عند اتخاذ القرارات التقديرية.

69. أشارت Meta إلى أنها تعتقد خمس قيم تؤثر في تطوير إنفاذ سياسات المحتوى على فيسبوك وInstagram. وهذه القيم هي "حرية الرأي" و"المصادقية" و"الخصوصية" و"السلامة" و"الكرامة". وفقاً لشركة Meta، تعتبر "حرية الرأي" هي القيمة ذات الأهمية القصوى للشركة. ويرى المجلس أن التحقق الشامل، وأي نظام لمنع أخطاء النتائج الإيجابية الخاطئة بوجه عام، يمسان في الأساس قيم "حرية الرأي" و"الخصوصية" و"السلامة" و"الكرامة".

70. يُسهّم نظام منع أخطاء النتائج الإيجابية الخاطئة الذي يُبقي المحتوى على المنصة حال عدم مخالفته لسياسات Meta في دعم فيسبوك وInstagram كأماكن لحرية التعبير. وبالعكس، ففي حالة إبقاء نظام منع أخطاء النتائج الإيجابية الخاطئة للمحتوى المخالف والضار على المنصة وتسهيل وصوله، فإنه يؤثر سلباً على قيم "حرية الرأي" و"السلامة" و"الخصوصية" و"الكرامة" الخاصة بالآخرين. وفي حالة منح النظام امتيازات لخطاب بعض المستخدمين دوناً عن غيرهم من خلال تأخير وتقليل احتمالية الإنفاذ، فإن عدم تكافؤ المعاملة هذا يمس قيمة "الكرامة" التي تعتقدتها شركة Meta، وهو ما يرتبط بتوقع أن تتعامل Meta مع جميع المستخدمين بنفس الطريقة. يجب على Meta ضمان هيكله أنظمتها على نحو يراعي جميع قيمها.

v. تقييم نظام التحقق الشامل

71. حين قدمت Meta إيجازاً إلى المجلس، بلغ العدد اليومي لمحاولات الإنفاذ على المحتوى 100 مليون محاولة. وفي ظل وجود هذا الكم الهائل، حتى لو استطاعت Meta اتخاذ قرارات إشراف على المحتوى بنسبة دقة 99٪، فإنها لا تزال ترتكب مليون خطأ كل يوم. تتضمن أخطاء الإشراف على المحتوى في Meta فرط الإنفاذ ونقص الإنفاذ، بما يعني أن Meta تزيل المحتوى غير المخالف وتتخلف عن إزالة المحتوى المخالف.

72. في هذا الصدد، يشكل استخدام Meta لنظام التحقق الشامل استجابة للتحديات الأوسع نطاقاً التي تكتنف الإشراف على كميات هائلة من المحتوى. يتفق المجلس على أنه في خضم هذا السياق الصعب، تحتاج Meta إلى آليات لمواجهة كل من النتائج الإيجابية الخاطئة والنتائج السلبية الخاطئة. وتقع على كاهل Meta مسؤولية حل هذه المشكلات الأكبر بطرق تفيد كل المستخدمين ولا تقتصر على قلة مختارة. يجب أن تتماشى أي قرارات تتعلق بتأخير إجراءات الإنفاذ أو الإعفاء منها لمستخدمين بعينهم أو عناصر محتوى معينة مع مسؤوليات Meta تجاه حقوق الإنسان وقيمتها المعلنة. ولم ينجح نظام التحقق الشامل بطريقة تشغيله السابقة أو شكله الحالي في القيام بذلك.

73. يلاحظ المجلس أن Meta أدخلت تحسينات على هذا النظام، قبل إحالة هذا الطلب إلى المجلس وأثناء تقييم المجلس لنظام التحقق الشامل. ومع هذا، لا يماشى عدد من جوانب نظام التحقق الشامل مع مسؤولية Meta تجاه تحديد الآثار السلبية على حقوق الإنسان والتخفيف من حدتها أو دعم قيم الشركة. ويتضمن ذلك:

- اتساع نطاق العمل لخدمة أهداف متعددة ومتناقضة تسمح بظهور المحتوى المخالف وانتشاره على نطاق واسع.
- عدم المساواة في الوصول إلى الإنفاذ والسياسات التقديرية.
- الإدراج في البرنامج قد يتجاوز السعة الاستيعابية.
- الإخفاق في تتبع أدوات القياس الأساسية لتقييم البرنامج وإدخال التحسينات عليه.
- الافتقار إلى الشفافية والمراجعة فيما يتعلق بطريقة عمل البرنامج.

74. على الرغم من وجود قدر كبير من القلق العام بشأن البرنامج، لم تعالج Meta بشكل فعال العناصر التي تثير المشكلات في نظامها. في هذا القسم، يسلط المجلس الضوء على عدد من هذه المشكلات. وفي الأقسام التي تليه، سنقدم مجموعة توصيات لشركة Meta لتوضيح كيف يمكن لنظام منع الأخطاء أن يتوافق بشكل أفضل مع التزامات الشركة.

اتساع نطاق العمل لخدمة أهداف ومتناقضة تسمح بظهور المحتوى المخالف.

75. أخبرت Meta المجلس أن المراجعة الثانوية للاستجابة المبكرة مصممة "لحماية حرية الرأي، [و] تحسين الشفافية والثقة لدى المجتمع." ولفنت Meta في الطلب الذي قدمته إلى المجلس الانتباه إلى قيامها بتضمين الصحفيين وقادة المجتمعات في نظام التحقق الشامل. وسلطت الشركة الضوء على أن نظام التحقق الشامل يضمن الحفاظ على حرية الرأي من خلال عدة سيناريوهات مهمة:

- "أعضاء المجتمعات المهمشة الذين يعيدون نشر الخطاب المخالف الذي يستهدفهم ويحرض على الكراهية بهدف تعزيز الوعي به أو إدانته، والذي أزيل عن طريق الخطأ لانتهاك سياساتنا بشأن الخطاب الذي يحرض على الكراهية."
- "الصحفيون الذين يقدمون تقارير إخبارية في مناطق النزاع التي تنتشر بها منظمات مصنفة، ممن أزيل المحتوى الخاص بهم عن طريق الخطأ لانتهاك سياساتنا بشأن المنظمات الخطرة والأفراد الخطرين."
- "العُري المرتبط بالصحة، مثل الصور الفوتوغرافية لعمليات التجميل بعد استئصال الثدي، أو الرضاعة الطبيعية، التي أزيلت عن طريق الخطأ لانتهاك سياساتنا بشأن العُري."

76. في اجتماع مع Meta، عند السؤال عن الآثار السلبية التي قد تترتب على استخدام المراجعة الثانوية للاستجابة المبكرة، ذكر مسؤولو Meta أن إحدى المشكلات، على سبيل المثال، تتمثل في احتمال أن تتسبب في إعاقة التواصل وتدفق المعلومات في الأزمات مثل الكوارث الطبيعية أو الاضطرابات السياسية. تتناقض نقاط التركيز هذه والتي أوردتها Meta في المبادئ التي يستند إليها النظام تناقضًا صارخًا مع آلية عمل النظام.

77. يشاطر المجلس شركة Meta القلق بشأن الإزالة عن طريق الخطأ للمحتوى غير المخالف المنشور بواسطة أشخاص يلفتون الانتباه إلى انتهاكات حقوق الإنسان، ويسعون إلى تعزيز صحة المرأة، ويقدمون تقارير أخرى تحظى باهتمام عام. في الواقع، عالجت قرارات المجلس هذه الأخطاء. تشير Meta إلى هذه الحالات على أنها "أخطاء إنفاذ" فقط بعد أن يلفت المجلس انتباه الشركة إليها. ومن الأمثلة على ذلك قرار حالة *حزام وماموم* ([2021-012-FB-UA](#))، والذي تناول الإزالة غير الصحيحة لتعبير فنان من السكان الأصليين يتصدى للكراهية بعد عدة قرارات خاطئة من المراجعة البشرية؛ وقرار حالة *نكر طالبان في التقارير الإخبارية الصادر عن المجلس* ([2022-005-FB-UA](#)) بشأن الإزالة غير الصحيحة لمنشور من منفذ إخباري يتضمن تقريرًا عن منظمة مصنفة؛ وقرار حالة *أعراض سرطان الثدي والعُري* ([2020-004-IG-UA](#))، بشأن الإزالة التلقائية غير الصحيحة لمنشور من المفترض أن يستفيد من الاستثناء المتعلق بالصحة الوارد في سياسات Meta بشأن عُرْي البالغين.

78. في حين تركز Meta في وصف البرنامج على الأصوات المعرضة للخطر التي تنشر محتوى غير مخالف، فقد ذكرت Meta أيضًا أن برنامج التحقق الشامل يخدم وظيفة تجارية أساسية، حيث يلعب "دورًا مهمًا في إدارة علاقات فيسبوك مع العديد من شركاء الأعمال [الذين تتعاون معهم]". وبشكل وثيق الصلة، فإن إطار عمل حساسية علامات التحقق الشامل، الذي يدعم كلاً من عامل "حساسية الكيان" فيما يتعلق بتصنيف المراجعة الثانوية العامة وعلامات المراجعة الثانوية للاستجابة المبكرة يرتبط ارتباطًا مباشرًا، ضمن عوامل أخرى، بدرجة رد الفعل المتوقع على صعيد السمعة والصعيد الداخلي في حالة إزالة محتوى معين عن طريق الخطأ. على سبيل المثال، تصف Meta خطر "التصعيد على أعلى المستويات (الرئيس التنفيذي، مدير التشغيل)" على أنه يقابل علامة "خطورة شديدة لأقصى درجة". يشير ربط الأولوية القصوى في التحقق الشامل بالمخاوف المتعلقة بإدارة العلاقات التجارية إلى أن Meta تريد تجنب عواقب ترتبط في الأساس بالأعمال التجارية وليس بحقوق الإنسان.

79. ولتقييم كيفية ترتيب Meta لأولوية الكيانات داخل نظام التحقق الشامل، طلب المجلس أكثر من مرة أن تشارك Meta قائمة المراجعة الثانوية للاستجابة المبكرة ليقوم المجلس بتحليلها. ولم تقدم Meta هذه القائمة للمجلس. لا يمكن للمجلس إجراء تقييم كامل لدرجة وفاء الشركة بمسؤولياتها تجاه حقوق الإنسان في ضوء البرنامج أو الملف الشخصي للكيانات التي تضمن المراجعة المحسنة إذا لم يكن على علم بكيفية تنفيذ البرنامج وتحديد من يستفيد منه بدقة. قالت Meta إن توفير قائمة المستخدمين الذين يخضعون للتحقق الشامل سيخالف الالتزامات القانونية للشركة فيما يتعلق بخصوصية المستخدمين. استنادًا إلى مشورة قانونية، يرى المجلس، وقد أوضح ذلك لشركة Meta، أنه كان من الممكن تخفيف حدة هذه المخاوف وتقديم إفصاحات أكثر شمولية.

80. بعد خمسة أشهر تقريبًا من طلب المجلس لهذه المعلومات لأول مرة، قدمت Meta للمجلس قائمة تضم بيانات مجمعة محدودة عن كل كيان مدرج في القائمة الحالية للمراجعة الثانوية للاستجابة المبكرة. على وجه التحديد، لم تفصح Meta إلا عن نوع الكيان (على سبيل المثال، مستخدم Instagram، صفحة فيسبوك)، والبلد واللغة التي حددها الكيان بنفسه، وما إذا كان الكيان "مدينًا" أو "شريكًا" وفقًا لتحديد شركة Meta. ولم يتم تقديم كل المعلومات لكل فئة من فئات الكيانات. على سبيل المثال، لم يحدد ربع كيانات Instagram المدرجة بلدًا أو لغة بعينها في إعدادات ملفاتهم الشخصية كما لا تعتبر جهات مادية أو شركاء لشركة Meta.⁵⁰ يعني ذلك أنه، بالنسبة إلى هذه الكيانات، لم تفصح Meta إلا عن وجود مجموعة من مستخدمي Instagram المستفيدين من التحقق الشامل، ولم تفصح عن هوياتهم أو سماتهم.

81. يُضعف هذا الإفصاح المحدود قدرة المجلس على الاضطلاع بمسؤوليات الإشراف المنوطة به. يتضمن وصف Meta لفئة الكيانات "المدينة"، على سبيل المثال، الجهات الحكومية والمسؤولين المنتخبين، و"الشخصيات المدنية المؤثرة" والمرشحين للمناصب العامة، وغيرها. وبالمثل، تشمل فئة "الشريك" المؤسسات الإخبارية والمشاهير والفنانين وغيرهم. لا يمكن للمجلس أن يقيم، على سبيل المثال، إلى أي درجة يحظى تعبير الصحفيين والمدافعين عن الحقوق والمعارضين في دول معينة بنفس الحماية التي تحظى بها الجهات الحكومية المدرجة في قائمة المراجعة الثانوية للاستجابة المبكرة بموجب سياسة البرنامج.

82. أخبرت Meta المجلس أنه ليس لديها نظام تقييم شامل يحدد بشكل منهجي الصحفيين أو المدافعين عن حقوق الإنسان أو شخصيات المجتمع المدني الذين ينتمون إلى منطقة جغرافية معينة ممن يجب أن يخضعوا للمراجعة الثانوية للاستجابة المبكرة. ويستند إدراج هؤلاء المستخدمين في القائمة إلى قرارات غير مركزية يتخذها بعض موظفي Meta الذين تصفهم الشركة بأنهم "خبراء داخليين يتمتعون بمعرفة واسعة بالسوق." وهذا يزيد من احتمال وجود فجوات وتناقضات كبيرة فيما يتعلق بمن يتمتع تعبيره بطبقات الحماية الإضافية التي توفرها المراجعة الثانوية للاستجابة المبكرة في نظام التحقق الشامل.

83. يشكل الصحفيون الذين ينشرون المحتوى من مناطق النزاعات، والمعارضة السياسية التي تسعى لشغل المناصب عبر الانتخابات، والمشاهير الذين ينشرون مجموعة واسعة من المحتوى، وشركاء الأعمال الذين ينشرون المحتوى بهدف بيع السلع، مستويات مخاطر مختلفة تمامًا من منظور حرية التعبير وحقوق الإنسان. ونظرًا للمشاكل التي تكتنف إشراف Meta على المحتوى على نطاق واسع، وفي ظل القيود الحالية، يجب أن يخضع المحتوى الذي يتم إنشاؤه بواسطة المستخدم إلى ترتيب مختلف للأولويات حسب الحقوق. وقد وصفت Meta نظامًا لا يتضمن استراتيجيات أو تكتيكات لضمان توفير الحماية للأفراد وأشكال التعبير الأكثر احتياجًا إليها على المدى القريب، ويهدف في نهاية المطاف إلى تحسين الإشراف على المحتوى للجميع.

84. بموجب المراجعة الثانوية للاستجابة المبكرة، إذا تم تحديد محتوى من أي كيان مؤهل باعتباره مخالفًا وتم تمييزه ليخضع لمراجعة إضافية، فإن هذا المحتوى، بغض النظر عن مستوى المخاطر، يظل على المنصة خلال فترة ذروة انتشاره والتي تلي نشره مباشرةً. وهذا أمر مهم لسببين. أولاً، ينتشر المحتوى سريع الانتشار على المنصة الواحدة ومن منصة لأخرى. ثانيًا، بمجرد أن ينشر كيان يتمتع بمعدلات وصول كبيرة أي عنصر محتوى، يتم بلا شك تسجيل المحتوى وإعادة مشاركته بشكل فردي بواسطة المستخدمين حتى إذا تم حذف المنشور الأصلي. ويعني ذلك أن الحسابات التي تستفيد من المراجعة الثانوية للاستجابة المبكرة في نظام التحقق الشامل قد تقوم بتحميل محتوى مخالف وهي تعلم أنه قد يصل إلى آفاق بعيدة حتى لو كان مخالفًا.

85. على الرغم من علم المجلس بأن **Meta** أشارت أنها تتبع نظامًا لتحديد أولويات المحتوى عالي الخطورة الخاضع للمراجعة الثانوية للاستجابة المبكرة، يظل هذا المحتوى على المنصة إلى أن تنتهي جميع المراجعات الضرورية، والتي تستغرق أحيانًا فترات طويلة. على سبيل المثال، في حالة نيمار، من الصعب فهم أسباب عدم صعود الصور الحميمية المنشورة دون موافقة أصحابها على حساب يضم أكثر من 100 مليون متابع إلى مقدمة قائمة انتظار المراجعة السريعة عالية المستوى حال اتباع أي نظام لترتيب الأولويات. ونظرًا للطبيعة الخطيرة لانتهاك السياسة والتأثير على الضحية، تسلط هذه الحالة الضوء على حاجة **Meta** إلى اعتماد مناهج مختلفة للتعامل مع المحتوى الذي ينتظر المراجعة وتقليل المدد الزمنية اللازمة للمراجعة.

86. بعد التأخر في اتخاذ إجراءات إنفاذ ضد المحتوى المخالف مصدر ضرر كبير في ظل برنامج التحقق الشامل. ووفقًا لبحث أجرته **Meta**، ترجع مشاهدات المستخدمين للمحتوى المخالف في ظل نظام التحقق الشامل إلى "عمليات الإلغاء غير الصحيحة، وتأخير اتخاذ إجراءات الإنفاذ ضد الحالات غير الملغاة التي يتم إبطاء الإنفاذ فيها بسبب عملية المراجعة الثانوية". وتقر الشركة بأن توفير حماية إضافية لمحتوى بعض المستخدمين ذوي الامتيازات قد يعرض باقي المستخدمين لمحتوى مخالف مثل الخطاب الذي يحض على الكراهية أو المنشورات المسيئة.

87. يختلف المحتوى الذي يتأهل تلقائيًا للمراجعة الثانوية للاستجابة المبكرة عن المحتوى الذي يتم اكتشافه وإرساله إلى المراجعة الثانوية العامة. من ناحية، وكما ذكر أعلاه، يبدو أن هناك تباين في النسبة المئوية لمحتوى المراجعة الثانوية للاستجابة المبكرة الذي يتبين في نهاية المطاف أنه مخالف. خلال الفترات الزمنية التي ينخفض فيها معدل الإلغاء، يتمثل العيب الرئيسي للنظام في عدم قدرته على ضمان الإزالة السريعة للمحتوى المخالف.

88. من الناحية الأخرى، يتبين باستمرار في نهاية المطاف أن معظم محتوى المراجعة الثانوية العامة غير مخالف. بالنسبة لهذا النظام، يبدو أن معدل الإلغاء يكشف عن وجود مشكلات أكبر على صعيد فرط الإنفاذ على نطاق واسع، وأن المراجعة الثانوية تسمح في الغالب باستمرار الوصول إلى المحتوى غير المخالف. وبالتالي، يلاحظ المجلس أنه بقدر ما تحافظ المراجعة الثانوية العامة على حرية التعبير بشكل أكبر، يظل تأثيرها محدودًا بقيود السعة الاستيعابية التي تفرضها **Meta**.

89. إيجازًا للقول، يرى المجلس أنه في حين تصف **Meta** برنامج التحقق الشامل على أنه يهدف إلى حماية الأصوات الضعيفة والمهمة، يبدو أنه قد تمت هيكلته ومعايرته بشكل أكثر مباشرة لمعالجة اعتبارات تتعلق بالأعمال التجارية. وفي حين يدرك المجلس أن **Meta** عبارة عن شركة ويفترض أن تصمم سياسات تعالج الاعتبارات المتعلقة بالأعمال التجارية، لا ينبغي وصف هذه السياسات نفسها على أنها تدابير للحد من المخاطر التي تؤثر في حقوق الإنسان ما لم تكن تحقق هذا الهدف. بالإضافة إلى ذلك، إذا كانت اختيارات تصميم الأعمال في **Meta** تؤثر سلبيًا في حقوق الإنسان، يجب أن تحدد تلك الآثار السلبية ثم تمنعها أو تخفف من حدتها أو توقفها من خلال إدخال تحسينات على البرنامج.

الوصول غير المتكافئ إلى السياسات التقديرية والإنفاذ

90. صُمم برنامج التحقق الشامل لإخضاع بعض عناصر المحتوى لقرارات إشراف أكثر دقة، وتحديد مدى انطباق أي استثناء أو سياسة متخصصة تمنع تطبيق إجراء الإنفاذ. وفقًا لشركة **Meta**، "إذا تم تصعيد المحتوى الخاضع لبرنامج التحقق الشامل لمراجعة إضافية، فقد يصدر قرارًا بشأنه استنادًا إلى [...] السياسات المتعلقة بالسياق." يتيح التحقق الشامل إجراء مراجعة بشرية من خلال "فريق الاستجابة المبكرة" والتي يرى المجلس أنها قد تمنح استثناءات في الإنفاذ، وذلك على صعيد المحتوى المحدد والعقوبات التي تُفرض على الكيان نفسه. ويضمن المحتوى الذي تتم مراجعته من خلال المراجعة الثانوية للاستجابة المبكرة الوصول إلى هذا الفريق قبل الإزالة المحتملة، بينما تزيد فرص وصول المحتوى الذي تتم مراجعته من خلال المراجعة الثانوية العامة إلى هذا الفريق.

91. تخبر **Meta** المجلس والعامه مرارًا وتكرارًا أنها تطبق نفس مجموعة السياسات على كل المستخدمين. مثل هذه التصريحات وسياسات المحتوى المعلنة مضللة، حيث لا يصل إلا القليل من عناصر المحتوى إلى مراجع لديه صلاحيات لتطبيق المجموعة الكاملة من السياسات.

92. بناءً على ما سبق، فإن أهلية الخضوع للمراجعة الثانوية للاستجابة المبكرة توفر ميزة كبيرة للمستخدم. ويعني ذلك أن الكثير من عناصر المحتوى التي يقرر المستخدم نشرها قد تظل على المنصة. في حالة المحتوى غير المخالف، تتم حمايته من الإزالة عن طريق الخطأ. وفي حالة المحتوى المخالف، يُسمح ببقائه على المنصة خلال ذروة المشاهدات قبل أن تتم إزالته لاحقًا.

93. يرى المجلس كذلك أنه، بالإضافة إلى تطبيق سياسات المحتوى على نحو تقديري أكبر، فقد يستفيد المحتوى الذي تتم مراجعته عند التصعيد من قرارات عدم فرض تقييدات على الحساب والتي كانت ستفرض عند اتباع الإجراءات العادية. بوجه عام، هناك "إنذارات" تفرض على الحساب مقابل انتهاكات سياسة المحتوى، والتي تسفر عن عواقب محددة. وفقًا لمركز شفافية **Meta**، تؤدي الإنذارات إلى منع الحسابات من نشر المحتوى لفترات زمنية طويلة للغاية. وتقوم **Meta** بتعطيل الحسابات في حالة الإنذارات الجسيمة أو المتكررة.

94. استفسر المجلس عن الجانب التقديري عند تطبيق السياسات وعواقب الإنفاذ. وكان رد **Meta** أنها "لا تملك بيانات ذات دلالة إحصائية تميز بين العقوبات التي تم فرضها على الكيانات التي تخضع لنظام التحقق الشامل مقارنة بتلك التي لا تخضع له" وأنها "لا تعلم ولا تجد بحثًا أو تحليلًا" يعالج هذه التناقضات المحتملة. ونظرًا لأن التحقق الشامل قد يعفي المستخدمين من العواقب على مستوى الحساب، يشعر المجلس بالقلق من أن الشركة قد اختارت إما عدم تتبع هذه المعلومات وتحليلها أو أنها لم تكشف عنها للمجلس.

95. وفقًا لتقرير عام نشرته صحيفة الجارديان، بعد أن نشر نيمار المحتوى المخالف، "لم يخضع للإجراء الطبيعي الذي تتخذه فيسبوك ضد أي شخص ينشر صورًا عارية غير مُصرح بها، وهو حذف حساب هذا الشخص". كشف بعض المبلغين عن المخالفات عن هذا المثال، ومن غير الواضح إلى أي مدى تنتشر هذه الممارسة. طلب المجلس من **Meta** أيضًا تأكيد القيود على مستوى الحساب التي فرضتها في هذه الحالة. وقد كشفت الشركة في نهاية المطاف أن العقوبة الوحيدة كانت إزالة المحتوى، في حين أن العقوبة الطبيعية كانت تعطيل الحساب. يلاحظ المجلس أن **Meta** أعلنت لاحقًا أنها أبرمت اتفاقًا اقتصاديًا مع نيمار يقوم من خلاله "ببث مقاطع فيديو حصرية أثناء اللعب على منصة Facebook Gaming ومشاركة محتوى الفيديو مع أكثر من 166 مليون معجب على Instagram".

96. يثير الوصول غير المتكافئ إلى المراجعات عند التصعيد واستثناءات السياسة المخاوف بشكل خاص نظرًا للافتقار إلى معايير موضوعية أو شفافة لتضمين الكيانات في قوائم المراجعة الثانوية للاستجابة المبكرة. وكما ذكرنا أعلاه، من غير الواضح كيف تضمن **Meta** حصول الأشخاص المرّجّح تعرضهم لفرط الإنفاذ أو مواجهة التحديات عند ممارسة حقوقهم في حرية التعبير على تلك الحماية الإضافية. يشعر المجلس بالقلق من أن الأشخاص الأكثر عرضة للخطر في أغلب الأحيان، كالصحفيين والمدافعين عن حقوق الإنسان، الذين ينشرون تقارير عن المنظمات الخطرة أو يوثقون حالات إساءة صادمة، هم الأقل احتمالًا لأن يضافوا بشكل استباقي إلى هذه القوائم نظرًا للاستثمارات اللازمة للعثور على هؤلاء الأشخاص حول العالم.

97. من الناحية الأخرى، أوضحت **Meta** للمجلس أن لديها فريقًا متخصصًا يعمل على ضمان إدراج كل الكيانات المؤهلة التي تمثل مسؤولين حكوميين أو منظمات في قوائم المراجعة الثانوية للاستجابة المبكرة. وتبدو معايير تضمين "الشركات والمؤسسات الإعلامية ومنشئي المحتوى" أيضًا أكثر وضوحًا. ووفقًا لشركة **Meta**، يتمثل أحد المعايير، على سبيل المثال، في مقدار إنفاق الكيان أو الإيرادات الناتجة عنه على مستوى "مجموعة تطبيقات" **Meta** على الرغم من احتمال تباين المبلغ بمرور الوقت.

98. ويشعر المجلس بالقلق أيضًا من أن **Meta** تولي اهتمامًا غير متناسب بالأسواق الأكثر ربحية عند تشغيل برنامج التحقق الشامل، بدلاً من التركيز على السياقات التي تنطوي على مخاطر أكبر على حقوق الإنسان، بما في ذلك حرية التعبير. حين قدمت **Meta** إيجازًا إلى المجلس، كان أصل 42٪ من المحتوى الذي تمت مراجعته من خلال مسار المراجعة الثانوية للاستجابة المبكرة يرجع إلى الولايات المتحدة أو كندا. وبالمثل،

تنتمي 20% من جميع الكيانات المدرجة في قوائم المراجعة الثانوية للاستجابة المبكرة في ذلك الوقت إلى هذين البلدين. على العكس من ذلك، ينتمي 9% فقط من "الأشخاص النشطين شهرياً" على فيسبوك إلى الولايات المتحدة وكندا وفقاً لشركة Meta. وتُظهر هذه البيانات أن المراجعة الثانوية للاستجابة المبكرة تمنح المستخدمين المقيمين في الولايات المتحدة وكندا وصولاً غير متناسب إلى مسارات المراجعة المتخصصة التي تضمن الوصول إلى المجموعة الكاملة من سياسات Meta، وتحليل السياق، واحتمال عدم فرض العقوبات القياسية على الحساب عند نشر محتوى مخالف.

99. يرتبط هذا التفاوت بحقيقة أن "متوسط دخل الفرد" في الولايات المتحدة وكندا هو الأعلى في العالم، وهو أكبر بثلاث مرات تقريباً من نظيره في أوروبا وبنحو 12 مرة تقريباً من نظيره في منطقة آسيا والمحيط الهادئ. وتسلب هذه الحقائق الضوء على الحوافز المالية التي تشكل طريقة عمل المراجعة الثانوية للاستجابة المبكرة وتعزز المخاوف المتعلقة بالعدالة. ومن خلال تصميم برنامج التحقق الشامل، يحصل المستخدمون في الأسواق المربحة التي تزيد فيها مخاطر تعرض Meta لمشكلات وتداعيات تتعلق بالعلاقات العامة على حماية أكبر لمحتواهم وتعبيرهم مقارنة بالمستخدمين في الأماكن الأخرى.

100. بالإضافة إلى ذلك، تحدد أداة تصنيف التحقق الشامل، في المراجعة الثانوية العامة، أولوية المحتوى وفقاً لعوامل مثل "حساسية الموضوع" والتي قد تتطلب إجراء تقييم ألي للغة المحتوى. ويشعر المجلس بالقلق من أن Meta لا تعطي الأولوية لتدريب عملياتها الآلية على اللغات الأقل تحدثاً والأسواق الأقل ربحاً. ويحد توجيه استثمارية كافية إلى الإشراف في هذه اللغات من قدرة الخوارزميات على تحديد الموضوعات التي يدور حولها هذا المحتوى. وهو ما يشير إلى احتمال ألا يكون المستخدمون في هذه الأسواق، بما في ذلك جنوب الكرة الأرضية، في أفضل وضع عند تقييم أهلية الخسوع للمراجعة الثانوية العامة. وبالمثل، كشفت Meta أنه "يتم إجراء المراجعة في بعض اللغات بواسطة أشخاص غير ناطقين بها من خلال استخدام الأدوات التي توفرها للترجمة وتمييز الإهانات". وهو ما يعزز من قلق المجلس تجاه عدم استفادة جميع المستخدمين بنفس الدرجة من نظام التحقق الشامل، حتى من خلال المراجعة الثانوية العامة.

المنضمون إلى البرنامج يتجاوزون السعة الاستيعابية

101. دائماً ما يتجاوز عدد الكيانات المؤهلة للمراجعة الثانوية للاستجابة المبكرة والمراجعة الثانوية العامة السعة الاستيعابية التي تخصصها Meta لبرنامج التحقق الشامل. ويعتبر عدم التوافق بين عدد عناصر المحتوى المقرر إخضاعها للمراجعة المحسنة من خلال هذه الأنظمة والموارد البشرية غير الكافية المخصصة لأداء هذه المهمة عيباً خطيراً في النظام.

102. أخبرت Meta المجلس أنها "لم تكن تهدف مطلقاً إلى وجود عدد متراكم ثابت من الحالات، وعلى الرغم من أن قيود السعة التشغيلية وتزايد أعداد عناصر المحتوى قد أدت إلى تراكم الحالات في المراجعة الثانوية للاستجابة المبكرة. فإن هذا التراكم يتألف من محتوى يعتبر وفق تقييمنا أقل خطورة." وعلى الرغم من تصريح Meta بأنها لم تكن تهدف إلى الاحتفاظ بعدد متراكم من الحالات باستمرار، فقد أخفقت الشركة في تخصيص ما يكفي من الموارد البشرية لتلبية احتياجات الإشراف على المحتوى لهذه البرامج. علاوة على ذلك، وكما ذكر أعلاه، ليس كل المحتوى الذي يؤجل اتخاذ إجراء إنفاذ بشأنه منخفض الخطورة.

103. إن محدودية السعة الاستيعابية للمراجعة البشرية لها تبعات مرتبطة تختلف في المراجعة الثانوية للاستجابة المبكرة عنها في المراجعة الثانوية العامة. بالنسبة إلى المراجعة الثانوية للاستجابة المبكرة، يعني نقص السعة الاستيعابية بقاء المحتوى على المنصة خلال الفترة الزمنية المرجح أن يحصل فيها على أكبر قدر من المشاهدات. ولأن هذا المحتوى يظل على المنصة إلى أن يخضع للمراجعة المحسنة، فإن المحتوى المخالف لسياسات Meta المنشور بواسطة مستخدمين بارزين مدرجين في قوائم المراجعة الثانوية للاستجابة المبكرة يظل على المنصة خلال الفترة الزمنية التي يحقق فيها أعلى مشاهدات. وعلى الرغم من أن Meta قد تحاول مراجعة المحتوى الذي قد يتسبب في أكبر قدر من الضرر أولاً، فإنه من غير الواضح ما إذا كانت تفعل ذلك باستمرار، ولا يزال ذلك يعكس قرار تصميم لتوفير الحماية التلقائية للكيانات التي يتم اختيارها بدرجة كبيرة استناداً إلى معايير تجارية.

104. بالنسبة إلى المراجعة الثانوية العامة، قد تسفر محدودية السعة الاستيعابية عن عاقبتين. أولاً، قد ينشغل فريق الاستجابة المبكرة بمحتوى المراجعة الثانوية للاستجابة المبكرة، حيث يجب مراجعة هذا المحتوى لتطبيق أي إجراء إنفاذ مطلوب. ولذلك لا يكون فريق الاستجابة المبكرة متاحاً في كثير من الأحيان لمراجعة محتوى المراجعة الثانوية العامة، ولا يصل محتوى المراجعة الثانوية العامة لهذا المستوى الحيوي من المراجعة الذي قد تُطبق فيه السياسات التي تتطلب سابقاً إضافياً وبعض التقدير. ثانياً، تعني محدودية السعة الاستيعابية على مستوى فريق السوق انتهاء مهلة بقاء عدد أكبر من عناصر محتوى المراجعة الثانوية العامة في قائمة الانتظار قبل المراجعة لتتم إزالتها افتراضياً. ونظراً لما يبدو دائماً من أن معظم عناصر هذا المحتوى تكون غير مخالفة، فإن ذلك يعني أن أحد أهم عواقب محدودية السعة الاستيعابية لمحتوى المراجعة الثانوية العامة يتمثل في أن Meta تزيل المزيد من المحتوى المرّجّح أنه غير مخالف.

105. تؤدي هذه العيوب إلى تفاهم الفروق في معاملة مختلف المستخدمين على المنصة. يحظى المستخدمون المتميزون المؤهلون للمراجعة الثانوية للاستجابة المبكرة بفرص أكبر لمراجعة محتوهم بواسطة مشرف يمكنه الاستعانة بالسياق للإبقاء على المحتوى، كما أن لديهم عدد أكبر من استثناءات السياسة التي يمكن تطبيقها للإبقاء على المحتوى، ويستفيدون من نظام يضمن حصول المحتوى المخالف على مشاهدات لبعض الوقت. على العكس من ذلك، يتمتع المستخدمون العاديون الذين قد يخضع المحتوى الخاص بهم للمراجعة الثانوية العامة بفرص محدودة لمراجعة محتوهم، وتزيد احتمالات خضوعهم لسياسات المحتوى دون مراجعة سياقية أو تطبيق أي استثناءات للسياسة، ومع انتهاء مهلة بقاء المحتوى في قائمة الانتظار، تزيد احتمالات إزالة المحتوى غير المخالف الخاص بهم. وهذا النظام له تداعيات خطيرة على قيم "حرية الرأي" و"الكرامة" و"الخصوصية" و"السلامة" التي تزعم Meta السعي إلى الالتزام بها.

الإخفاق في تتبع أدوات القياس الأساسية لتقييم البرنامج وإدخال التحسينات عليه

106. أجرى المجلس تقييماً لأدوات القياس التي تستخدمها Meta لتبرير برنامج التحقق الشامل وتقييمه. لا تعالج أدوات القياس التي تستخدمها Meta حالياً كل الجوانب المهمة ولا يبدو أنها أدت إلى إجراء تغييرات عند اكتشاف أوجه القصور. علاوة على ذلك، أخفقت Meta في مراقبة وتحديد الأهداف الخاصة بمجموعة كبيرة للغاية من أدوات القياس بحيث ترسم صورة كاملة لآلية عمل البرنامج وتضع أهدافاً للتحسين وفقاً لها.

107. وكما ناقشنا أعلاه، يتمثل أحد أدوات القياس التي تحسبها Meta في معدل الإلغاء، أو النسبة المئوية للمحتوى الخاضع للتحقق الشامل والذي يتبين في نهاية المطاف أنه غير مخالف، على الرغم من تحديده كمخالف في بادئ الأمر بواسطة مراجعة بشرية أو آلية. وفقاً لشركة Meta، فإن "معدل الإلغاء هو معدل كفاءة نظام التحقق الشامل". ووفقاً للمعلومات التي قدمتها إلى المجلس، فإن Meta "تريد أن تكون النسبة المئوية [لمعدل الإلغاء] عالية. في حالة عدم إلغاء أي قرارات من خلال التحقق الشامل، يعني ذلك أنها [كانت] تتحقق من المحتوى الخاطئ".

108. على الرغم من تصريح Meta بضرورة أن يكون معدل الإلغاء مرتفعاً، فإنها تواصل توفير أكبر قدر من الحماية للمستخدمين المؤهلين للمراجعة الثانوية للاستجابة المبكرة. ووفقاً للأرقام التي قدمتها Meta إلى المجلس، يتباين هذا المعدل بشكل كبير. ويشير توفير هذه الحماية للمحتوى دون تحقيق معدل إلغاء مرتفع باستمرار إلى أن Meta، وفقاً لأهدافها، تجري مراجعة التحقق الشامل للمحتوى الخاطئ.

109. يتسم الإشراف على المحتوى بأعداد كبيرة بكل من فرط الإنفاذ ونقص الإنفاذ. تركز Meta على معدل انتشار المحتوى المخالف كأداة قياس عامة رئيسية لتقييم مدى فعالية جهود الإشراف في إزالة المحتوى الضار. يتضمن ذلك المحتوى الذي يتدفق خلال نظام التحقق الشامل. تحسب Meta معدل الانتشار من خلال تقدير النسبة المئوية لكل مشاهدات المحتوى على فيسبوك أو Instagram التي شكلت مشاهدات لمحتوى مخالف. قد يشجع استخدام معدل الانتشار كأداة لقياس النجاح العام لشركة Meta على المضي قدماً في أتمتة إزالة المحتوى وتقليل الإنفاذ الذي يستند إلى السياق لضمان خفض معدل الانتشار على مستوى المنصة، دون آليات مناسبة لمنع حذف المحتوى عن طريق الخطأ على نطاق واسع. ويدعم الارتفاع المستمر لمعدلات الإلغاء في المراجعة الثانوية العامة، على سبيل المثال، هذا الاستنتاج.

110. يلاحظ المجلس أن Meta لم تقدم للمجلس معلومات توضح أنها تتتبع البيانات الخاصة بدقة القرارات التي يتم اتخاذها من خلال نظام التحقق الشامل. يعني ذلك أنه على الرغم من أن من المفترض أن يضمن البرنامج اتخاذ قرارات إشراف دقيقة بشأن المحتوى، لا يبدو أن Meta تتتبع ما إذا كانت القرارات التي تتخذها عبر مسارات التحقق الشامل أكثر أو أقل دقة من القرارات التي تتخذها من خلال آليات مراقبة الجودة العادية واسعة النطاق. قد تكون بيانات الدقة مؤشراً أساسياً لما تمثله المخاوف غير المتعلقة بسياسة المحتوى من تأثير محتمل على قرارات الإشراف المتخذة ضمن برنامج التحقق الشامل. ومن خلال الاعتماد على معدل الإلغاء فقط لقياس النجاح، فإن Meta لا تأخذ في الحسبان مدى صحة القرارات النهائية.

111. علاوة على ذلك، ذكرت Meta أن فرق السوق الإقليمية وفرق الاستجابة المبكرة هي فرق متخصصة، تتمتع بمجموعة معينة من المهارات، والتدريب وصلاحيات الوصول إلى الأدوات الداخلية التي تتيح لهم اتخاذ قرارات إشراف على مستوى التحقق الشامل. مع هذا، وكما ذكرنا أعلاه، قد يتم اتخاذ القرارات، في نقاط معينة من المراجعة الثانوية للاستجابة المبكرة والمراجعة الثانوية العامة، بواسطة مراجعين خارجيين. لا يتمتع هؤلاء المراجعون بنفس صلاحيات الوصول أو التدريب المتوفر لموظفي Meta. إذا كان هدف برنامج التحقق الشامل يتمثل في اتخاذ أكثر قرارات السياسة دقة للكيانات المؤهلة والمحتوى المهم، يجب أن يكون قياس دقة قرارات التحقق الشامل بوجه عام، وقياسها على مستوى مختلف أنواع المراجعين بوجه خاص، مبدأً أساسياً لفهم مدى عمل التصميم التشغيلي على النحو المنشود.

112. بالإضافة إلى ذلك، نظراً لأن التحقق الشامل يهدف إلى حماية المحتوى المهم الأكثر عرضة لخطر فرط الإنفاذ، يجب أن تركز Meta على طرق إضافية لاكتشاف هذا المحتوى. كشفت Meta أنها بينما تسعى جاهدة لفهم فرط الإنفاذ ونقص الإنفاذ والتخفيف من حدتهما لمجموعات محددة من السكان وجوانب المشكلات، فإنها لا تزال "بحاجة إلى تحديد السكان الذين يعانون من فرط/نقص الإنفاذ بشكل مركزي. وإلى أن تنتهي هذه الجهود، لا توجد لدينا طريقة جيدة لتحديد ذلك بشكل مسبق."

الافتقار إلى الشفافية والمراجعة بشأن البرنامج وطريقة عمله

113. أخيراً، يشعر المجلس بالقلق إزاء المعلومات المحدودة التي قدمتها Meta إلى العامة ومستخدميها بشأن هذا البرنامج. نتج هذا الرأي الاستشاري بشأن السياسة عن عدم كشف Meta للمجلس عن المعلومات المهمة حول هذا البرنامج في سياق مداولاته بشأن حالة لمستخدم بارز يخضع لبرنامج التحقق الشامل.

114. في الوقت الحالي، لا تُخبر Meta المستخدمين أنهم يخضعون للمراجعة الثانوية للاستجابة المبكرة، وهي آلية تستند إلى الكيان في برنامج التحقق الشامل. ولا تخبر المستخدمين أيضاً عند إبلاغهم عن محتوى منشور بواسطة كيان خاضع للتحقق الشامل. ولا توفر الشركة إلى قدرًا محدودًا من الشفافية بشأن عمليات المراجعة الثانوية المعقدة التي يستفيد منها المحتوى الخاضع للتحقق الشامل.

115. علاوة على ذلك، لا تشارك Meta إجراءات إنشاء قوائم المراجعة الثانوية للاستجابة المبكرة وإطار تدقيقها علناً. ولا يعلم المجلس، على سبيل المثال، ما إذا كان يتم الاحتفاظ في قوائم المراجعة الثانوية للاستجابة المبكرة بالكيانات التي تكرر نشر محتوى مخالف على أساس ملفها الشخصي. ولم تقدم Meta أي إشارة بشأن ما إذا كان سجل الانتهاكات أو تكرارها يمثل عاملاً في إنشاء قوائم المراجعة الثانوية للاستجابة المبكرة أو الاحتفاظ بها. ويحول الافتقار إلى الشفافية بشأن التدقيق دون فهم المجلس والعامة للنتائج الكاملة المترتبة على نظام التحقق الشامل.

استنتاجات حول التحقق الشامل

116. يقر المجلس بأن نظام منع الأخطاء قد يمثل ضمانة مفيدة ضد الإزالة غير الصحيحة للمحتوى المهم. ومع هذا، إذا لم يكن التحقق الشامل يستهدف مثل هذا التعبير ويسمح ببقاء المحتوى المخالف بشدة على المنصة، فإن البرنامج يترك آثاراً سلبية على حقوق الإنسان لا تراقبها Meta ولا تخفف من حدتها. وبالتالي، فقد خلص المجلس إلى أن طريقة تصميم وتنفيذ التحقق الشامل في الوقت الحالي لا تفي بمسؤوليات Meta تجاه حقوق الإنسان ولا قيم الشركة.

117. في قراراته بشأن الحالات، يأخذ المجلس في اعتباره الاختبار الثلاثي الوارد في المادة 19 من العهد الدولي الخاص بالحقوق المدنية والسياسية، ويجري تقييمًا لمدى استيفاء القيود التي تُفرض على حرية التعبير لمتطلبات الشرعية والهدف المشروع والضرورة والتناسب.
118. تشير الشرعية إلى مدى وضوح القواعد وتيسير الوصول إليها. يتسم وجود النظام وغرضه وطبيعته بعدم الوضوح بطرق لا يمكن تبريرها نظرًا لما يتركه نظام التحقق الشامل من آثار كبيرة على ممارسة الحقوق الأساسية. وتتسم سياسات المحتوى المقدمة على أنها قابلة للتطبيق عالميًا والتي لا يمكن تطبيقها إلا بالرجوع إلى السياق الإضافي عند التصعيد، بما في ذلك من خلال التحقق الشامل، بالتضليل.
119. يشير الهدف المشروع إلى مدى سعي القيود المفروضة لتحقيق الأهداف المحددة في المادة 19، بما في ذلك احترام حقوق الآخرين وحماية الأمن الوطني، والنظام العام، والصحة العامة. تشير أدوات القياس التي تقيس من خلالها الشركة فعالية نظام الإنفاذ إلى أن دوافعها تركز بشكل كبير على أسباب تجارية.
120. تشير الضرورة والتناسب إلى ما إذا كانت القيود المفروضة على حرية التعبير هي الطريقة الأقل تدخلًا لتحقيق الهدف المشروع. في هذا الصدد، يكرر المجلس مخاوفه بشأن الوصول غير المتكافئ لمزايا نظام التحقق الشامل. تتبع Meta عمليات واضحة لتقرر أن بعض مستخدميها عبارة عن كيانات مؤهلة، مثل الجهات الحكومية وشركاء الأعمال. دون وجود معايير واضحة للمستخدمين الآخرين المرجح قيامهم بنشر محتوى يمثل قيمة كبيرة لحقوق الإنسان، فإن البرنامج يفيد الآخرين بشكل أقل وضوحًا، بما في ذلك أعضاء الفئات المهمشة والمعرضة للتمييز. لا تجمع Meta كذلك أو تراقب المعلومات المتعلقة بما إذا كان البرنامج يسفر عن نتائج أكثر دقة عند الممارسة العملية. أخيرًا، من خلال التحقق الشامل، يعتبر إبقاء المحتوى المخالف على منصاتها هو الإجراء الافتراضي الذي تتخذه Meta. كأمر يتعلق بالسياسة، تتخلى Meta عن الاستجابة التي ترى أنها متناسبة على نطاق واسع لبعض عناصر المحتوى، وهو ما يستند غالبًا إلى الاعتبارات الاقتصادية أو العلاقات العامة فقط.
121. للامتثال إلى مسؤوليات Meta تجاه حقوق الإنسان والقيم التي تتبناها الشركة، يجب هيكلة نظام لمنع فرط الإنفاذ بطريقة مختلفة تمامًا عن الوضع الراهن.

.VI توصيات بشأن الإنفاذ

122. استجابة للأسئلة التي طرحتها Meta، يقدم المجلس هنا توصيات بشأن أنظمة منع الأخطاء التي تستند إلى الكيان وأنظمة منع الأخطاء التي تستند إلى المحتوى ويتم تحديدها ديناميكيًا. وتقع على كاهل Meta مسؤولية معالجة تحديات الإشراف على المحتوى بطرق تفيد جميع المستخدمين ولا تقتصر على فئة مختارة. مع هذا، ونظرًا لتركيز هذا الرأي الاستشاري بشأن السياسة، يركز المجلس هنا على أنظمة منع الأخطاء ذات النطاق المحدود.

توصيات حوكمة نظام منع الأخطاء الذي يستند إلى الكيان

123. يجب تصميم أي نظام يستند إلى أهلية الكيانات، مثل المراجعة الثانوية للاستجابة المبكرة، بعناية وأن يخضع للإشراف والمراقبة المستمرة. وهو ما يفترض أن يضمن تحقيق هذا النظام لأغراضه المعلنة وتقييمه للعوامل الخارجية والعواقب غير المقصودة التي قد يتسبب بها. يجب أن يحمي هذا النظام المستخدمين المرجح أن ينشروا تعبيرًا يحظى بأهمية خاصة من منظور حقوق الإنسان.
124. ومن المهم للغاية أن توضح Meta أهدافها وأن تكيّف أنظمتها لتحقيق هذه الأهداف بدقة. ويجب عليها أيضًا تجنب توفير الحماية للتعبير الذي ينتهك سياسات المحتوى لديها أو التزاماتها تجاه حقوق الإنسان. بالإضافة إلى ذلك، نظرًا لاحتمال استفادة بعض المستخدمين من وسائل الحماية الإضافية وسبل التعبير، يجب على الشركة أن

توفر للجمهور معلومات قوية عن هذه العمليات حتى يتمكنوا من تقييم المعلومات والآراء التي يرونها على المنصة بشكل مناسب.

المستخدمون الذين يجب تضمينهم في أنظمة منع الأخطاء التي تستند إلى الكيان

125. تذكر **Meta** أن فئات التضمين لنظام منع الأخطاء الذي يستند إلى الكيان تغطي "الجهات المدنية والحكومية"، و"الأحداث العالمية المهمة"، و"وسائل الإعلام"، و"الفئات التي تعرضت لفرط الإنفاذ تاريخياً"، و"المجتمعات المهمشة"، و"الشركات" و"منشئي المحتوى"، و"الكيانات التي يتم تصعيدها للمراجعة" و"الجهات القانونية والتنظيمية".
126. تتطلب هذه الفئات واسعة النطاق فرزاً وتصنيفاً إضافياً. في ضوء التزامات **Meta** تجاه حقوق الإنسان وقيمتها المعلنة، إذا قررت الشركة تشغيل نظام يستند إلى الكيان لمنع النتائج الإيجابية الخاطئة، هناك فئات مستخدمين ينبغي توفير تلك الحماية لها، ومستخدمين يمكن توفير تلك الحماية لهم، ومستخدمين لا ينبغي توفير هذه الحماية لهم تبعاً للمخاطر التي قد يمثلونها على حقوق الإنسان.
127. أولاً، الكيانات التي ينبغي تضمينها هي الكيانات المرجح أن تقدم تعبيراً يحظى بأهمية من منظور حقوق الإنسان، بما في ذلك الأمور التي تحظى باهتمام عام. ولا يستفيد من ذلك هؤلاء المستخدمين فسخب، بل تصل الاستفادة أيضاً إلى الأشخاص الذين يرغبون في الوصول إلى المعلومات التي ينشرونها.
128. يجب أن يشتمل هؤلاء المستخدمون، على سبيل المثال، على الأشخاص الذين يتعرض محتوهم لخطر كبير من فرط الإنفاذ، والصحفيين والمؤسسات الإعلامية، والمسؤولين العموميين والمرشحين للمناصب، وغيرهم من الجهات المدنية بما في ذلك المدافعين عن حقوق الإنسان ومؤيدي المجتمعات المهمشة. في هذا الصدد، ينظر المجلس إلى النظام الذي يستند إلى القوائم على أنه وسيلة لتوفير سبل حماية إضافية للتعبير النقدي، وليس الحماية فقط استناداً إلى هوية المتحدث. يدرك المجلس أن **Meta** تحتفظ بقوائم متعددة للكيانات التي توفر لها حماية أكبر، بما في ذلك سجل الصحفيين وقائمة "الشركاء الموثوق بهم" من المجتمع المدني. قد تشكل هذه الكيانات الحالية التي تم فحصها مصدرًا تستند إليه الشركة في إنشاء نظام موضوعي وعالمي يستند إلى معايير حقوق الإنسان ويتوفر لجميع الأشخاص الذين يلبي التعبير الخاص بهم معايير الانضمام إلى النظام.
129. ثانيًا، الكيانات التي يمكن تضمينها يمكن أن تستند إلى أولويات الشركة وقد تضم مستخدمين لهم أهمية تجارية وشركاء أعمال. وقد يشتمل ذلك على المعلنين والشركات التي لديها صفحات أو مجموعات تعاني من خطر فرط الإنفاذ، أو المستخدمين الذين يشكلون خطورة خاصة على صعيد سمعة الشركة، أو غيرهم من المستخدمين الذين تربطهم علاقة تجارية مع **Meta**.
130. ثالثًا، هناك كيانات لا ينبغي تضمينها في أي نظام منع أخطاء يستند إلى الكيان من شأنه تأخير كل إجراءات الإنفاذ. ويشتمل ذلك على الكيانات والمستخدمين الذين يكررون إنشاء أو مشاركة محتوى ينتهك سياسات **Meta** أو شروط الخدمة. ويمكن الاستفادة من نظام إنفاذ **Meta** الحالي على مستوى الحساب، والذي يستند إلى الإنذارات والعقوبات، لأغراض تطبيق هذه القاعدة. وإذا كرر المستخدمون الذين تم تضمينهم بسبب أهميتهم التجارية نشر المحتوى المخالف، فلا ينبغي لهم الاستمرار في الاستفادة من نظام يؤخر الإنفاذ. تضطلع **Meta** بمسؤولية اكتشاف هؤلاء المستخدمين واستبعادهم من الأنظمة التي توفر لمحتوهم المخالف فرص ظهور إضافية. وفي حين قد يعتبر عدد المتابعين دليلاً مشروعاً على درجة الاهتمام العام بتعبير المستخدم، يجب ألا تكون شهرة المستخدم أو عدد متابعيه المعيار الوحيد لنظام منع الأخطاء الذي يستند إلى الكيان.
131. إن قيام **Meta** بتضمين جميع الكيانات في النظام ذاته يجعلها في منافسة مباشرة على موارد المراجعة المحدودة. ويجب أن تضع **Meta** الأولوية على توفير الموارد الكافية لأنظمة منع الأخطاء التي تخفف من حدة الأضرار التي تؤثر في حقوق الإنسان. وفي هذا السياق، يجب أن تضمن **Meta** إخضاع المحتوى الذي يؤثر في حقوق الإنسان أو المصلحة العامة للمراجعة في الوقت المناسب بواسطة مراجعين مهرة لديهم القدرة على أخذ السياق الإضافي في الاعتبار، بغض النظر عما إذا كان مصدر المحتوى هو المسارات التي تستند إلى الكيان أو إلى المحتوى.

132. يوصي المجلس **Meta** باتخاذ خطوات إما لاستخدام مسارات منفصلة أو إنشاء آليات تحديد أولويات للتمييز بين المستخدمين الذين يجب تضمينهم بموجب مسؤوليات **Meta** تجاه حقوق الإنسان والمستخدمين الذين يتم تضمينهم بموجب الأولويات التجارية، مع عدم إغفال اختلاف مستويات المخاطر الخاصة بكل منهم. تميل الشركات، على سبيل المثال، إلى تحديد المحتوى باعتباره مخالفًا لقواعد المحتوى غير المهم أو الاحتمالي، لأنها قد تنشر المحتوى التجاري سريعًا. قد يقوم المستخدمون الذين لديهم أعداد كبيرة من المتابعين بتناول القضايا المهمة التي تتعلق بالمصلحة العامة في منشوراتهم، لكنهم قد ينشرون المحتوى المخالف أيضًا.

ضرورة أن يكون متخذو القرار مؤهلين ومخولين لاتخاذ القرارات التي تحترم الحقوق

133. تماشياً مع توصيات المجلس بوجه عام، يجب على **Meta** ترتيب أولوية دفع المراجعة الثانوية لمنع الأخطاء وفقاً لمستوى المخاطر وقيمة حقوق الإنسان.

134. يجب أن تتم مراجعة المحتوى المنشور بواسطة الكيانات التي ينبغي على **Meta** تضمينها استناداً إلى المخاوف المتعلقة بحقوق الإنسان بواسطة فرق تتمتع بخبرة في السياق واللغة. ويجب أن يخلو مسار المراجعة هذا، بما في ذلك مسارات التصعيد الخاصة به، من الاعتبارات التجارية. ويجب على **Meta** اتخاذ الخطوات اللازمة للتأكد من أن هذا الفريق لا يتبع فرق السياسة العامة أو العلاقات الحكومية أو الفرق المسؤولة عن إدارة العلاقات مع أي مستخدمين متأثرين.

135. ويمكن للمسار المخصص لحل المشكلات التي ترتبط صراحة بأولويات الأعمال التجارية لدى **Meta** أن يعالج، على سبيل المثال، الإنفاذ فيما يتعلق بالإعلانات، وقواعد المحتوى غير المهم أو الاحتمالي، وتقييد الميزات والمشكلات السلوكية. ومن أمثلة المشكلات السلوكية فرض عقوبة على صفحة أعمال عن طريق الخطأ بسبب تحميل الصور بوثيرة أسرع من الملف الشخصي العادي. ويجب ألا تحل هذه المراجعات محل الموارد التي تستهدف تخفيف الآثار السلبية على حقوق الإنسان، سواء من خلال إعطاء أولوية أقل أو الفصل إلى دفع عمل مختلف.

136. يلاحظ المجلس أن فريق الاستجابة المبكرة، الذي يحق له تطبيق استثناءات السياسة وتفسير السياق، لا يشترط أن يتمتع مراجعوه بخبرة ثقافية أو لغوية. وفقاً لشركة **Meta**، يتخذ هذا الفريق قراراته استناداً إلى الملاحظات المقدمة من فرق السوق الإقليمية. وقد أقرت **Meta** ذاتها بأن "الاعتماد على الترجمة ليس مثاليًا". في هذا السياق، يحث المجلس **Meta** على توفير الخبرة الثقافية واللغوية عند هذه المستويات من المراجعة. ينبغي على **Meta** النظر في ضم موظفين يتمتعون بخبرة ثقافية ولغوية من المناطق المعرضة للخطر إلى هذه الفرق ووضع إجراءات للاستعانة بموظفين من ذوي الخبرة عند اتخاذ القرارات.

إرشادات لإنشاء وحوكمة القوائم لأنظمة منع الأخطاء التي تستند إلى الكيان

137. ينبغي على **Meta** وضع معايير واضحة وعلنية لأهلية الانضمام إلى نظام منع الأخطاء الذي يستند إلى الكيان. ويجب أن تفرق هذه المعايير بين المستخدمين الذين يتأهل تعبيرهم لحماية إضافية من منظور حقوق الإنسان، بما في ذلك المعلومات التي تصب في المصلحة العامة، والمستخدمين الذين يتم تضمينهم لأسباب تجارية. على سبيل المثال، تعرّف **Meta** حالياً إحدى فئات التحقق الشامل على أنها "المؤسسات الإعلامية والشركات والمجمعات ومنشئي المحتوى". تضم هذه الفئة "الهيئات الصحية وناشري الأخبار والفنانين الترفيهيين والموسيقيين والفنانين والمبدعين والمنظمات الخيرية". وهذه المعايير واسعة بدرجة تجعلها غير كافية. ويجب على **Meta** أيضاً وضع المعايير استناداً إلى أنماط السلوك المخالف أو غير المرغوب فيه على المنصة لتجنب منح الحماية للمستخدمين المؤذنين.

138. يجب ألا تضيف **Meta** الكيانات إلى أنظمة منع الأخطاء إلا عندما تكون العملية موضوعية ومحكمة وشفافة. ويجب أن تكون جميع الكيانات المقترحة إضافتها إلى القوائم على علم بهذه الإمكانية وأن يكون بمقدورها رفض هذه الإضافة إن رغبت في ذلك. ويجب مطالبة الكيانات التي اختارت الانضمام بمراجعة قواعد المحتوى التي

تفرضها Meta وأن تعيد إلزام نفسها باتتباع تلك القواعد. وبينما يرى المجلس أن برنامج التحقق الشامل يوفر مزايا للمستخدمين المنضمين إليه، يجب أن تأخذ Meta موافقة المستخدمين قبل ضمهم.

139. يجب أن توفر المعايير العامة الواضحة أساسًا للمستخدمين المؤهلين ليسعوا بشكل استباقي للإدراج في هذه القوائم. يجب على Meta وضع عملية يتمكن المستخدمون بموجبها من تقديم طلبات للتمتع بالحماية التي يكفلها نظام منع أخطاء فرط الإنفاذ حال استيفائهم للمعايير التي تعلنها الشركة. يجب أن تكون الجهات الحكومية مؤهلة للإضافة أو تقديم الطلبات بناءً على هذه المعايير والشروط دون أن تحصل على أي تفضيلات أخرى.

140. بالإضافة إلى تلبية المعايير العامة، يجب أن تشمل عملية التضمين على ما يلي، بغض النظر عما إذا كان بدء العملية قد تم بواسطة Meta أو المستخدم: (1) متطلب لمراجعة سياسات Meta بشأن المحتوى مع تعهد إضافي صريح باتتباعها و(2) إقرار بالقواعد الخاصة بالبرنامج و(3) نظام لإخبار المستخدمين استباقيًا بالتغييرات التي تجريها Meta على سياسات المحتوى ليسهل الوعي بها والامتثال إليها.

141. تتعاون Meta في بعض الأحيان مع المجتمع المدني من خلال برنامج "الشركاء الموثوق بهم" ومبادرات إشراك أصحاب المصلحة الآخرين بهدف تجميع معلومات عن الكيانات التي يجب أخذها في الاعتبار لتوفير الحماية لها. يوصي المجلس Meta بتعزيز تفاعلها مع المجتمع المدني لأغراض إنشاء القوائم. ويجب أن يتمكن المستخدمون من ترشيح آخرين يستوفون المعايير العامة، طالما كان يحق للمرشحين رفض الانضمام. يكون هذا الأمر مُلحًا بشكل خاص في البلدان التي يكون وجود الشركة فيها محدودًا بدرجة لا تسمح بتحديد المرشحين للإدراج في القوائم بشكل مستقل.

142. يجب أن تدار عملية إنشاء القوائم، وتحديدًا هذه المشاركة، بواسطة فرق متخصصة، مستقلة عن الفرق التي قد تؤدي تفويضاتها إلى تضارب في المصالح، مثل فرق السياسة العامة لدى Meta. لضمان استيفاء المعايير، يجب أن يتأكد الموظفون المتخصصون، عبر الاستفادة من المعطيات المحلية، من تطبيق معايير الإدراج بموضوعية. غالبًا ما تتفاعل فرق السياسة العامة مع الجهات الحكومية وتؤثر فيها، مما يخلق حوافز متضاربة لا يمكن تجنبها. وعلى الرغم من إمكانية اقتراحهم للمرشحين، لا ينبغي أن يكونوا صناع القرار.

143. أخبرت Meta المجلس أنه في الوقت الحالي يمكن لموظف واحد بالشركة أن يتخذ قرارًا بإضافة كيانات إلى قائمة تحقق شامل معينة، ولا يشترط خضوع هذه القرارات للمراجعة. من الآن فصاعدًا، ينبغي على الشركة اتباع عملية ثابتة لإجراء مراجعة موضوعية تستند إلى المعايير لكل الكيانات التي تتلقى مزايا إضافية. إذ ينبغي مشاركة ما لا يقل عن عضوين من فرق مختلفة لإنهاء الانضمام إلى أي حماية تستند إلى القوائم، وينبغي ألا يكون أي شخص على علاقة شخصية أو تجارية مع الكيانات المرشحة من صناع القرار.

إرشادات الاحتفاظ بقوائم أنظمة منع الأخطاء التي تستند إلى الكيان وتدقيقها

144. بالإضافة إلى وضع معايير واضحة للانضمام إلى برنامج حماية لمنع الأخطاء، يجب أن تضع Meta معايير وعمليات واضحة للتدقيق والإزالة. إذا لم تعد الكيانات تفي بمعايير الأهلية، يجب إزالتها.

145. أخبرت Meta المجلس أن هيكل الحوكمة المقترح الجديد يتضمن قواعد لإضافة الكيانات وإزالتها من القوائم؛ وقواعد لعلامات انتهاء الصلاحية؛ وإجراءات للتدقيق الدوري؛ وهيكل للإشراف. ومع ذلك، فقد كشفت Meta أيضًا عن وجود استثناءات لبعض هذه القواعد، مثل عدم وجود مدد افتراضية لانتهاء صلاحية الكيانات "المدنية والحكومية". شاركت Meta أيضًا أنها تجري حاليًا تدقيقًا لمجموعة محدودة من الكيانات في قائمة المراجعة الثانوية للاستجابة المبكرة مع انتقالها إلى هيكل قوائم أكثر بساطة.

146. يوصي المجلس Meta بطلب إجراء مراجعة سنوية على الأقل لكل الكيانات المدرجة في أي نظام لمنع الأخطاء يوفر مزايا لهذه الكيانات. يجب كذلك وجود بروتوكولات واضحة لتقليل هذه الفترة الزمنية عند الضرورة. وعلى غرار توصياته بشأن الإدراج الأولي في أي نظام يستند إلى القوائم، يوصي المجلس بمشاركة ما لا يقل عن شخصين من انتماءات مختلفة في عمليات التدقيق الداخلية.

147. يجب أن تتأكد Meta أيضًا من وجود معايير إزالة واضحة من برامج الحماية التي تستند إلى القوائم. يجب أن يتمثل أحد المعايير في حجم المحتوى المخالف المنشور بواسطة الكيان. وقد يستند ذلك، على سبيل المثال، إلى سياسة تتألف من "ثلاثة إنذارات"، ما لم تفرض Meta عقوبة أكثر قسوة بسبب الانتهاك أو الانتهاكات ذات الصلة (على سبيل المثال، إزالة الحساب لمشاركة صور حميمية دون موافقة أصحابها). يجب أن يحذر هذا النظام الكيانات قبل أن يزيلها من التحقق الشامل عند بلوغها الإنذار الأخير، بغض النظر عما إذا كان الانتهاك يستوجب الإزالة من المنصة ككل أم لا. ويجب أن تتاح للكيانات إمكانية الطعن ضد قرار الإزالة وإعادة التقديم للانضمام إلى البرنامج في المستقبل.

148. أخيرًا، يؤكد المجلس أنه، في حين تمثل إجراءات التدقيق الداخلي خطوة في الاتجاه الصحيح، فإن التدقيق الداخلي دون إشراف خارجي يعد إجراءً قصيرًا. يجب إجراء تدقيقات خارجية، بواسطة المجلس أو جهة خارجية أخرى (على سبيل المثال، الباحثين أو المجتمع المدني) لتقييم ما إذا كان نظام منع الأخطاء يخفف من الآثار السلبية على حقوق الإنسان أم لا. وفي حين يقر المجلس بجديّة مخاوف الخصوصية والسلامة التي تكتنف التدقيق الخارجي، يرى المجلس أن بإمكان Meta اتخاذ خطوات من شأنها إخفاء هوية البيانات وتجميعها لمعالجة هذه المخاوف.

ضرورة تمييز بعض الكيانات التي تتلقى حماية إضافية بشكل علني

149. دعا المجلس Meta مرارًا وتكرارًا إلى إبلاغ المستخدمين وعامة الجمهور بسياساتها وممارساتها. يجب لأي نظام منع أخطاء يستند إلى الكيان أن يوضح لكل مستخدم المنصة الطريقة التي تتبعها Meta لتطبيق قواعدها. في الوقت الحالي، لا يعلم المستخدمون بما إذا كانوا مدرجين في قوائم المراجعة الثانوية للاستجابة المبكرة أم لا. بالإضافة إلى ذلك، لا يتم إبلاغ المستخدمين الذين يشاهدون ويبلغون عن المحتوى المنشور بواسطة مستخدمين مدرجين في قوائم المراجعة الثانوية للاستجابة المبكرة بأن المحتوى قد يخضع لإجراءات مراجعة معينة.

150. يوصي المجلس بتمييز حسابات بعض فئات الكيانات التي تتمتع بالحماية من خلال النظام بشكل علني. تتضمن هذه الفئات كل الجهات الحكومية والمرشحين السياسيين وكل شركاء الأعمال وكل الكيانات الإعلامية وكل الشخصيات العامة الأخرى المدرجة نظرًا لفائدتها التجارية للشركة في تجنب النتائج الإيجابية الخاطئة. سيتيح ذلك للعامة مساهمة المستخدمين المتميزين بشأن مدى وفاء الكيانات المحمية بالتزاماتها نحو اتباع القواعد ومساهمة Meta عن الالتزام بالمعايير المعلنة للبرنامج.

151. يحدد المجلس العديد من المخاطر المترتبة على الإفصاح عن المستخدمين المسجلين في برنامج منع أخطاء النتائج الإيجابية الخاطئة علنًا. أولاً، قد يؤدي ذلك إلى زيادة المخاطر العدائية من المستخدمين الذين يحاولون السيطرة على الحسابات التي تتمتع بحماية خاصة، نظرًا لعلمهم بأن المحتوى المخالف سيبقى على المنصة لفترة من الوقت. ثانيًا، قد تواجه بعض فئات المستخدمين مضايقات أو هجمات أخرى إذا تصوّر الآخرون أنهم على علاقة مع الشركة أو يحظون بحماية خاصة منها.

152. مع هذا، يرى المجلس أنه يمكن التخفيف من حدة هذه المخاطر، وأن الفوائد تفوق هذه الأضرار المحتملة. أولاً، يجب على Meta استثمار الموارد اللازمة لتعزيز حماية حسابات المستخدمين الخاضعين لنظام منع الأخطاء. وتتمتع Meta بخبرة في توفير طبقات حماية إضافية للصحفيين والفئات الأخرى من المستخدمين. ويمكن موازنة هذه الإجراءات لاستخدامها في أي نظام مستقبلي لمنع الأخطاء يستند إلى الكيان. وعلى الرغم من أن المخاطر العدائية حقيقية، فإنها لا تستعصي على الحل في هذا السياق. وعلى الرغم من أن قائمة المراجعة الثانوية للاستجابة المبكرة ليست علنية حاليًا، يفترض العديد من المستخدمين بالفعل أن الحسابات البارزة مدرجة في برنامج التحقق الشامل.

153. يجب ألا تكشف Meta هوية المستفيدين من البرنامج من المدافعين عن حقوق الإنسان، والكيانات المدرجة بسبب تعرضها لفرط الإنفاذ تاريخيًا، أو لأنها معرضة لخطر الضرر، على الرغم من أنه ينبغي أن يكون بإمكانها اختيار أو عدم اختيار كشف هويتها. وسيؤدي اتباع معايير واضحة للإدراج في البرنامج أو إلغاء المشاركة فيه مع اختلاف الأهداف إلى تسهيل هذه العملية.

154. أخيرًا، عندما يبلغ المستخدمون عن المحتوى المنشور بواسطة كيان محدد علنًا كمستفيد من المراجعة الإضافية، يجب أن توضح لغة البلاغ بشكل صريح أنه سيتم تطبيق إجراءات خاصة، مع توضيح الخطوات واحتمالية طول المدة اللازمة لاتخاذ القرار.

توصيات حوكمة نظام منع الأخطاء الذي يستند إلى المحتوى

155. في حين يفترض أن يضم النظام الذي يستند إلى الكيان المستخدمين المرجح أن ينتجوا أوجه تعبير تستحق حماية إضافية من منظور حقوق الإنسان والمستخدمين الذين قد يكونوا معرضين بشكل خاص لمخاطر عالية من فرط الإنفاذ، يسعى النظام الذي يستند إلى المحتوى لحماية هذا المحتوى مباشرةً بغض النظر عن ناشره.

المحتوى الذي ينبغي تحديده ووضع الأولوية عليه لأنظمة منع الأخطاء التي تستند إلى المحتوى

156. وفقًا لشركة Meta، يقوم نظام الاستجابة العامة لديها "بتصنيف المحتوى استنادًا إلى مخاطر النتائج الإيجابية الخاطئة باستخدام معايير مثل حساسية الموضوع (مدى رواج/حساسية الموضوع)، وخطورة الإنفاذ (خطورة إجراء الإنفاذ المحتمل)، واحتمالية حدوث نتيجة إيجابية خاطئة، والوصول المتوقع، وحساسية الكيان (يتم الاستناد بدرجة كبيرة عند تحديد الخطورة إلى القوائم المجمعة المذكورة أعلاه)".

157. تتمثل العوامل الأكثر ترجيحًا لدى خوارزمية التصنيف في حساسية الموضوع وحساسية الكيان. وكما ناقشنا أعلاه، ترتبط حساسية الكيان، من بين عوامل أخرى، ارتباطًا مباشرًا بدرجة التصعيد الداخلي الذي قد ينتج عن الخطأ. وفي هذا الصدد، تعطي أداة تصنيف التحقق الشامل في Meta الأولوية أيضًا للمحتوى الذي قد يتسبب في ضرر اقتصادي أو يضر بسمعة الشركة، وهو هدف تخدمه بالفعل المراجعة الثانوية للاستجابة المبكرة. وعلى الرغم من أن تصميم المراجعة الثانوية العامة ربما كان استجابة لبعض الانتقادات الموجهة لنظام التحقق الشامل في صورته السابقة التي كان يستند فيها إلى الكيان فقط، فإن ذلك يشير إلى استمرار الشركة في إعطاء الأولوية للتعبير استنادًا إلى المتحدث وليس إلى أهمية التعبير.

158. يتفق المجلس على أن الأهلية الشاملة لنظام منع أخطاء النتائج الإيجابية الخاطئة هو خطوة إيجابية. ومع هذا، يجب لهذا النظام أن يضع الأولوية على اكتشاف المحتوى غير المستهدف أيضًا بنظام يستند إلى الكيان. ويجب أن يوفر حماية محسنة استنادًا إلى منطق حقوق الإنسان. وفي حين قد تمنح Meta بعض الحماية الإضافية عندما يهدد فرط الإنفاذ مصالحها التجارية، على نحو مشابه للأنظمة التي تستند إلى القوائم، فإنه لا ينبغي لها القيام بذلك على حساب التزاماتها تجاه حقوق الإنسان.

159. يمكن لأداة تصنيف خوارزمية نظام منع النتائج الإيجابية الخاطئة، على سبيل المثال، إعطاء الأولوية للمحتوى استنادًا إلى أنواع القرارات التي يصعب على الوسائل التلقائية والمشرفين الذين يعملون على نطاق واسع اتخاذها (على سبيل المثال، الخطاب الذي تعرض تاريخيًا لفرط الإنفاذ أو خطاب المجتمعات المهمشة). في الوقت ذاته، يمكن أن تنظم الخوارزمية أولوية ترتيب مراجعة هذا المحتوى حسب شدة الانتهاك المحتمل، واحتمال أن يمثل نتيجة إيجابية خاطئة، واحتمال انتشاره بشكل واسع وسريع.

160. يوصي المجلس أنه لزيادة تأثير نظام منع النتائج الإيجابية الخاطئة استنادًا إلى المحتوى، يجب على Meta النظر في إمكانية تخصيص حد أدنى من السعة الاستيعابية للمراجعة بواسطة الفرق التي يمكنها تطبيق كل سياسات المحتوى (على سبيل المثال، فريق الاستجابة المبكرة). ويجب عليها أيضًا تحليل المحتوى الذي يخضع لمراجعة إضافية لتحديد الحالات التي تؤدي فيها أنظمة Meta إلى حدوث أكبر عدد من الأخطاء عالية التأثير، لترتيب أولوية موارد المراجعة وفقًا لذلك.

التصحیحات الفنية

161. أوضحت Meta أن "التصحیحات الفنية" تحظر تمامًا اتخاذ أي إجراء إنفاذ بشأن انتهاكات سياسة محددة لكيان معين. قد توجد أسباب تجارية لتوفير مثل هذه الحماية لمجموعة منتقاة للغاية من الكيانات، ولكن يحتمل أن يخلق أي نظام من هذا القبيل خطرًا كبيرًا يتمثل في استثناء الكيانات التي تنشر المحتوى المخالف من إجراءات إنفاذ

الإشراف على المحتوى. وفي حالة استخدام مثل هذا النظام، يجب أن يخضع لأعلى مستويات التدقيق الداخلي والخارجي. تستثني "التصحیحات الفنية" كيانات معينة من بعض إجراءات الإنفاذ وتُفهم بطريقة صحيحة على أنها "قائمة سماح" أو "قائمة بيضاء"، مهما كان نطاقها محدودًا.

162. تنطبق هنا كل التوصيات بشأن البرامج التي تستند إلى القوائم، مثل وضوح المعايير وثباتها، وإجراء عمليات المراجعة من فرق مختلفة لمنح أي إعفاء، وعمليات التدقيق للتأكد من أهلية الكيان للإعفاء. بالإضافة إلى ذلك، يجب حظر منح الإعفاء للمحتوى الذي تصنفه Meta على أنه يمثل انتهاكًا شديد الخطورة. ويجب على Meta إجراء عمليات تدقيق دورية لكل إجراءات الإنفاذ التي حجبتها هذه الإعفاءات. إذا كان العدد، كما ذكرت Meta، نحو ألف إجراء يوميًا، يجب أن توفر السعة الاستيعابية لإجراء هذا التدقيق. يجب تضمين هذا التدقيق، مع المعلومات المتعلقة بنطاق البرنامج ودقته، في تقارير الشفافية ربع السنوية الصادرة عن Meta.

163. أخيرًا، يجب على الشركة أن تبحث بشكل استباقي ودوري عن أي إعفاءات غير متوقعة أو غير مقصودة ربما بقيت من التكرارات السابقة لهذا البرنامج. في قراراته، لاحظ المجلس أكثر من مرة الحالات التي فشلت فيها Meta دون قصد في تحديث الأنظمة أو صيانتها، وقد تكون العواقب المترتبة على هذه الفجوات في حوكمة نظام الإعفاءات شديدة الخطورة.

توصيات عامة لحوكمة نظام منع الأخطاء

164. بالإضافة إلى تغييرات الحوكمة العامة بشأن كيفية إنشاء وتدقيق أنظمة منع الأخطاء التي تستند إلى القوائم أو المحتوى، يوصي المجلس أيضًا بأن تركز الإجراءات داخل تلك الأنظمة على تخفيف الأضرار وأن تخضع للمراقبة المستمرة من أجل التعلم والتحسين.

تخفيف الأضرار بعد اكتشاف المحتوى المخالف

165. كما تدرك الشركة ذاتها، يرجع أحد الأسباب الأساسية للضرر في نظام منع أخطاء النتائج الإيجابية الخاطئة الذي تتبعه Meta إلى تأخير اتخاذ إجراءات إنفاذ ضد المحتوى المخالف خلال الفترة المرجح أن يحظى فيها بأكثر قدر من المشاهدات. وكما أشرنا أعلاه، تحدد Meta ذاتها أن أكبر محررات مشاهدة المستخدمين للمحتوى المخالف المنشور بواسطة المستخدمين أو المحتوى الخاضع للتحقق الشامل على منصات Meta هي "عمليات الإلغاء غير الصحيحة، وتأخير اتخاذ إجراءات الإنفاذ ضد الحالات غير الملغاة التي يتم إبطاء الإنفاذ فيها بسبب عملية المراجعة الثانوية". يحث المجلس Meta على اتخاذ خطوات للتخفيف من حدة هذه الأضرار.

166. أولاً، يجب على Meta اتخاذ تدابير للتأكد من اتخاذ قرارات المراجعة الإضافية بأسرع ما يمكن. ويجب ضخ الاستثمارات وإجراء التغييرات الهيكلية لزيادة أعداد فرق المراجعة بحيث يتوفر ويعمل المراجعون في المناطق الزمنية المعنية متى تم تمييز المحتوى على أنه يحتاج إلى مراجعة بشرية محسنة.

167. ثانيًا، يوصي المجلس Meta باتباع طرق أخرى بخلاف الطريقة الافتراضية المتمثلة في عدم اتخاذ إجراء إنفاذ ضد عناصر المحتوى التي تخضع لمراجعة محسنة. وقد تتضمن تلك الطرق استخدام أقل الطرق تدخلًا، على سبيل المثال، خفض تصنيف المحتوى، أو إبطاء سرعة انتشاره، أو إخفاء المحتوى أو إزالته مؤقتًا. ويجب أن يؤدي إنشاء أولويات أو مسارات مختلفة للمحتوى والكيانات ذات الطبيعة المختلفة إلى إضفاء سهولة على قيام Meta بتطبيق العواقب المختلفة على أنواع المحتوى المختلفة.

168. يجب إزالة المحتوى المخالف الذي تبين خلال التقييم الأولي لشركة Meta أنه شديد الخطورة، على سبيل المثال وفقًا لإطار عمل Meta، أو إخفائه إلى أن تنتهي المراجعة، وعدم السماح ببقائه على المنصة وحصوله على المشاهدات لمجرد أن المستخدم الذي نشره هو شريك أعمال أو شخصية مشهورة. ويجب أن يستند الاختلاف بين خيارات الإنفاذ، مثل الإزالة والإخفاء وخفض التصنيف، إلى شدة الانتهاك. وقد تم تصميم إطار عمل Meta، نظرًا، لياخذ في الحسبان احتمال حدوث ضرر على المدى القريب، وما إذا كان قد تم تحديد المحتوى على أنه من المحتمل بشكل خاص أن يشكل خطأ إنفاذ. في حالة إخفاء المحتوى لهذه الأسباب، يجب تقديم إشعار للمستخدمين في مكان المحتوى يشير إلى أنه قيد المراجعة.

169. ثالثاً، يجب ألا تقوم Meta بتشغيل هذه البرامج على نحو يسمح بوجود تراكمات. إن الاحتفاظ بقائمة انتظار مراجعة محتوى تتجاوز السعة الاستيعابية يعني بقاء المحتوى الذي قد يكون مخالفاً على المنصة لفترات زمنية طويلة. ويؤدي تأخير إي إنفاذ مع استغراق أسابيع للوصول إلى قرار إلى استثناء الكيانات المؤهلة وظيفياً من القواعد.

170. ويتعين على Meta استثمار الموارد اللازمة بما يتناسب مع قدرتها على مراجعة المحتوى الذي تقرر أنه يحتاج إلى طبقات مراجعة إضافية. ومع ذلك، لا يعني هذا أن تجعل الخوارزمية تحدد عناصر محتوى أقل. يجب ألا يؤدي فشل Meta في توفير السعة الاستيعابية الكافية للمراجعة إلى تأخير الإنفاذ ضد المحتوى أو عمليات الحذف الخاطئة تماماً المتخذة بواسطة الأنظمة أو المراجعين على نطاق واسع. ابتكرت Meta عمليات لترتيب أولويات المراجعة والتأكد من أن فرق عملها لديها تدفق مستمر من المحتوى لمراجعتها. ونظراً لأن المراجعة الثانوية العامة تسفر حالياً عن معدل إلغاء مرتفع، يرى المجلس أن المزيد من المحتوى سيستفيد من هذه المراجعة.

171. رابعاً، يجب ألا ترتب Meta أولوية المراجعة الثانوية التي تستند إلى الكيان تلقائياً وأن تجعل قدرًا كبيراً من المراجعة التي تستند إلى المحتوى المحددة خوارزمية تعتمد على سعة المراجعة الإضافية.

التأكد من توفر القدرة على تقديم الطعون

172. أخبرت Meta المجلس أنها لا توفر فرص الطعن أو المراجعة بطريقة متسقة على مستوى كل أنواع المحتوى. ويبدو أن طعون المحتوى الخاضع لبرنامج التحقق الشامل تعاني من عدم الاتساق ذاته.

173. يدرك المجلس أن توفير إمكانية تقديم طعون بشأن المحتوى الذي وصل بالفعل إلى أعلى مستويات التحليل داخل الشركة قد يكون غير ضروري، نظراً لأن الاستئناف من شأنه تكرار نفس هذه المسارات. مع هذا، يشعر المجلس بالقلق من أن بعض عناصر المحتوى قد لا تحصل على الأهلية للطعن، على الرغم من عدم وصولها لهذه المستويات العالية. يرى المجلس أنه كان بإمكان Meta وكان ينبغي عليها توفير مزيد من الوضوح بشأن هذه النقطة عند سؤال المجلس عنها مراراً وتكراراً.

174. بالإضافة إلى ذلك، يشعر المجلس بالقلق بشكل خاص بشأن هذا الالتباس لتعلقه بأهلية تقديم الطعون التي تصل من خلالها الحالات إلى المجلس، بالنسبة إلى المستخدمين لاستعادة المحتوى الخاص بهم والذي خضع لمراجعة التحقق الشامل وكذلك للإبلاغ عن محتوى المستخدمين الآخرين الذي يستفيد من التحقق الشامل. في الواقع، وفقاً لشركة Meta، "بالنسبة إلى شهري مايو ويونيو 2022، لم يكن بالإمكان تصعيد ما متوسطه 35% من المحتوى الخاضع للتحقق الشامل [...] إلى مجلس الإشراف". ويُعد المستخدمون المدرجون في قوائم المراجعة الثانوية للاستجابة المبكرة من بين المستخدمين الأعلى وصولاً على المنصة. وقد يحرم هذا الوضع المجلس من بعض أكثر حالات الإشراف على المحتوى أهمية على فيسبوك وInstagram.

175. كخطوة أولى، يجب على Meta توضيح الأهلية لتقديم الطعون بوجه عام، والتأكد من إمكانية الطعن داخلياً بشأن المحتوى الذي لا يصل إلى أعلى مستويات المراجعة. ثانياً، يجب على Meta ضمان توفير فرصة التقدم بطعن إلى المجلس لكل المحتوى الذي يحق للمجلس مراجعته بموجب وثائقه الحاكمة، بغض النظر عن مدى وصول المحتوى إلى أعلى مستويات المراجعة داخل Meta.

التعلم والتحسين

176. للوفاء بمسؤولياتها تجاه حقوق الإنسان، يجب على Meta إجراء مراقبة دورية لأنشطتها التي تؤثر في الحقوق. ويجب أن توجه نتائج هذه المراجعات شركة Meta أثناء تحسين سياساتها وممارساتها لتقليل الأضرار على

حقوق الإنسان. في هذه الحالة، لدى Meta مجموعة من أدوات القياس المتعلقة ببرنامج التحقق الشامل والتي توضح للشركة بالفعل الجوانب التي ينبغي تحسينها. يرى المجلس أنه ينبغي على شركة Meta أيضاً تزويد عامة الجمهور بمعلومات عن طريقة عمل هذا النظام، للوفاء بمسؤولياتها تجاه الشفافية وتحميل نفسها المسؤولية تجاه إجراء التحسينات.

177. أولاً، تحتفظ Meta بالفعل بمعدل إلغاء لنظامها الذي يستند إلى الكيان (المراجعة الثانوية للاستجابة المبكرة) ونظامها الذي يستند إلى المحتوى (المراجعة الثانوية العامة). ينبغي على Meta استخدام الاتجاهات فيما يتعلق بمعدلات الإلغاء لتحديد هل ستتم العودة افتراضياً لإجراء الإنفاذ الأصلي خلال إطار زمني أقصر أو تحديد إجراءات الإنفاذ الأخرى التي سيتم تطبيقها إلى أن تنتهي المراجعة. إذا كان معدل الإلغاء منخفض باستمرار لمجموعة معينة من انتهاكات السياسة أو المحتوى المنشور بلغات معينة، على سبيل المثال، يجب أن تعابير Meta باستمرار مدى سرعة وتدخل إجراء الإنفاذ الذي ينبغي عليها تطبيقه.

178. ثانياً، أخبرت Meta المجلس أنها أجرت تمارين تحليلية تالية بعد أن حدد فريق "تقييم المخاطر" لدى Meta جوانب الخطر أو حالات وجود حدث ترى الشركة أنه بمثابة إخفاق. ويوصي المجلس بإخضاع نظام التحقق الشامل إلى هذه المراجعات وغيرها بانتظام، استناداً إلى تقييمات المخاطر الداخلية التي تختبر الضغط على النظام في النقاط الرئيسية الموضحة في هذا الرأي الاستشاري بشأن السياسة.

179. ثالثاً، كشفت Meta أن إحدى الفئات التي تستخدمها للمراجعة الثانوية للاستجابة المبكرة هي "الكيانات التي تعرضت لفرط الإنفاذ تاريخياً". يعني ذلك أن الشركة قد حددت بالفعل كيانات تقر Meta أنها غير قادرة على إنفاذ سياساتها بطريقة متسقة وفعالة. بالإضافة إلى تزويد هذه الكيانات بإمكانية الوصول إلى برامج منع أخطاء فرط الإنفاذ، ينبغي على Meta استخدام هذه البيانات لتحسين ممارسات الإنفاذ على نطاق واسع. يجب أن تقيس Meta فرط الإنفاذ لهذه الكيانات وينبغي عليها استخدام هذه البيانات للمساعدة على تحديد الكيانات الأخرى التي تتعرض لفرط الإنفاذ. ويجب أن يكون تقليل أداة القياس هذه هدفاً واضحاً وذو أولوية كبيرة بالنسبة إلى الشركة.

180. هناك أدوات قياس أخرى يجب على Meta تطويرها ومراقبتها لمواءمة استراتيجيات منع الأخطاء بشكل أفضل مع معايير حقوق الإنسان. على سبيل المثال، يجب على Meta إنشاء أدوات قياس جديدة لتحديد تأثير ترك المحتوى المخالف على المنصة. ويجب على الشركة تحديداً حساب عدد المشاهدات التي تحققها عناصر المحتوى التي تمت إزالتها في نهاية المطاف أثناء انتظارها للمراجعة بسبب آليات منع الأخطاء. ويجب على Meta تحديد أساس لأداة القياس تلك وأن تُصدر تقارير عن مستهدفات تقليلها.

181. كشفت Meta أيضاً أنها تتخذ خطوات أيضاً لمعالجة بعض المشكلات المتعلقة بنقص الإنفاذ. ويتضمن ذلك "أدوات تصنيف لاكتشاف المحتوى المحتمل أن ينتهك سياساتنا؛ وبلاغات المستخدمين التي تحدد المحتوى المحتمل أن يكون مخالفاً؛ وعمليات المسح التي تتم من خلال المراجعة البشرية حيث تراجع فرقنا المحتوى المحتمل أن يكون مخالفاً؛ وعمليات المراجعة المبكرة عالية الخطورة، وهو نظام يراجع فيه مراجعون بشريون المحتوى المتوقع أن ينتشر سريعاً؛ وطعون المبلغين، حيث يمكن للمستخدمين الذين يبلغون عن المحتوى المخالف التقدم بطعن على القرار [الذي تتخذه Meta]".

182. يلاحظ المجلس أن الجهود المبذولة بخلاف الطعون وعمليات الإنفاذ التلقائية على نطاق واسع لها نطاق ضيق. بالإضافة إلى ذلك، تتنافس بعض هذه المبادرات على الموارد مع نظام التحقق الشامل. على سبيل المثال، يتم تنفيذ عمليات المراجعة المبكرة عالية الخطورة بواسطة فرق السوق، التي يجب أن تخصص أيضاً سعة استيعابية لنظام التحقق الشامل. توفر عمليات المراجعة المبكرة عالية الخطورة المراجعة أيضاً للمحتوى المتوقع أن ينتشر سريعاً. يتفق المجلس على أن المحتوى عالي الوصول قد يسبب ضرراً أكبر لكنه يرى ضرورة أن يكون مصحوباً بجهود لتحسين الإشراف بشكل شامل. ويجب أن تستمر Meta في توجيه الاستثمارات إلى أنظمة الاكتشاف والتحذير المبكر؛ وتوظيف وتضمين أشخاص يتمتعون بخبرات محلية ولغوية في جهود الثقة والسلامة وعمليات مراجعة المحتوى وإنشاء قوائم نظام منع الأخطاء.

VII. توصيات بشأن الشفافية

183. أصدر المجلس سلسلة من التوصيات تتناول كيف ينبغي على Meta تصميم وحوكمة أي برنامج لمنع أخطاء النتائج الإيجابية الخاطئة. تعني مسؤوليات Meta تجاه حقوق الإنسان أيضًا أن عليها توفير الشفافية للجمهور حول هذه البرامج. ويجب أن تتضمن تقارير الشفافية بيانات شاملة حتى يتمكن المستخدمون والجمهور من فهم طريقة عمل البرنامج وتبعاته بالنسبة إلى الخطاب العام. وبالإضافة إلى أدوات القياس التي سبق وصفها، يوصي المجلس بـ Meta بتضمين:

- a. معدلات الإلغاء لأنظمة منع أخطاء النتائج الإيجابية الخاطئة، مصنفة حسب خيارات التصميم وفرق الإنفاذ (على سبيل المثال، الأسواق والاستجابة المبكرة والمتعاقدين، وغير ذلك). على سبيل المثال، أوصى المجلس Meta بإنشاء تدفقات مختلفة لمختلف فئات الكيانات المختلفة أو المحتوى استنادًا إلى التعبير ومستوى المخاطر. يجب الإبلاغ عن معدل الإلغاء لكل الأنظمة التي تستند إلى الكيان والتي تستند إلى المحتوى، وفئات الكيانات والمحتوى المضمنة.
- b. العدد الإجمالي والنسبة المئوية لسياسات التصعيد فقط التي تم تطبيقها نتيجة لبرامج منع أخطاء النتائج الإيجابية الخاطئة نسبة إلى إجمالي قرارات الإنفاذ.
- c. وسط ومتوسط الوقت اللازم للتوصل إلى القرار النهائي بشأن المحتوى الخاضع لبرامج منع أخطاء النتائج الإيجابية الخاطئة، مع تصنيفها حسب البلد واللغة.
- d. بيانات مجمعة عن أي قوائم مستخدمة لبرامج منع الأخطاء، بما في ذلك نوع الكيان والمنطقة.
- e. معدل عمليات الإزالة الخاطئة (النتائج الإيجابية الخاطئة) لكل المحتوى الذي تمت مراجعته، بما في ذلك إجمالي الضرر الناتج عن هذه النتائج الإيجابية الخاطئة ويقاس كإجمالي المشاهدات المتوقعة للمحتوى (أي فرط الإنفاذ).
- f. معدل القرارات الخاطئة للإبقاء على المحتوى (النتائج السلبية الخاطئة)، بما في ذلك إجمالي الضرر الناتج عن هذه النتائج السلبية الخاطئة ويقاس كمجموع المشاهدات التي حققها المحتوى (أي نقص الإنفاذ).

184. أوصى المجلس سابقًا بأن تكشف Meta عن معدلات الخطأ إجمالاً، ولكن يجب عليها أيضًا "الإبلاغ عن معدلات الخطأ النسبية للقرارات المتخذة من خلال التحقق الشامل مقارنة بإجراءات الإنفاذ العادية". ويعتقد المجلس أن تركيز Meta على معدل الانتشار، رغم أنه مفيد في سياقات محددة معينة، لا يوفر الحوافز الصحيحة للشركة أو الأدوات المناسبة للجمهور لفهم طريقة عمل منظومة الإشراف على المحتوى في Meta.

185. أخبرت Meta المجلس أنها "تستثمر حاليًا في قياس أداة قياس مجمعة تسمح لنا بفهم النتائج الإيجابية الخاطئة خلال النظام بالكامل وتعمل على بناء أداة القياس هذه والتي نأمل أن نشاركها خارجيًا في تقارير الشفافية التي نصدرها. وستكون أداة القياس هذه هي الأداة المقابلة لقياس النتائج السلبية الخاطئة التي ننشرها في التقارير حاليًا من خلال أداة قياس معدل الانتشار." ويمثل ذلك خطوة في الاتجاه الصحيح ويحث المجلس Meta على إكمال هذا العمل بأسرع ما يمكن.

186. بالإضافة إلى أدوات القياس التي تم التأكيد عليها في الأقسام السابقة، التي تعمل على تحسين قياس الأداء وتوفير المعلومات، يجب على Meta أيضًا توفير معلومات أساسية في مركز الشفافية بشأن طريقة عمل أي نظام لمنع الأخطاء تستخدمه من شأنه تحديد الكيانات أو المستخدمين الذين يحصلون على حماية إضافية. ويدرك المجلس احتمال وجود محاولات عدائية من المستخدمين لتجاوز الإنفاذ، وقد تختار Meta تلخيص بعض نقاط ممارسات الإنفاذ التي تتبعها. ويعتبر المستوى الحالي للشفافية غير كافٍ ولا يبرره الخوف من المخاطر العدائية.

187. وبصورة أشمل، يرى المجلس أن توفير مزيد من الشفافية للباحثين الخارجيين، لا سيما إمكانية الوصول إلى البيانات، هو عنصر أساسي في الإشراف على أنظمة منع الأخطاء. خلال تفاعل المساهمين الذي تم إجراؤه من أجل هذا التحليل، سمع المجلس مخاوف بشأن سعي Meta للحد من برامجها الحالية للوصول إلى البيانات للأطراف الخارجية. وبالنظر إلى أن أنظمة مثل التحقق الشامل تتطلب إجراء مفاضلات معقدة، يمكن للباحثين المستقلين تزويد Meta برؤى قيمة حول آثار خياراتها. ويرى المجلس أنه يجب على Meta إنشاء مسار للباحثين الخارجيين للوصول إلى البيانات غير العلنية حول نظام التحقق الشامل وهو ما من شأنه أن يسمح لهم بفهم البرنامج بشكل كامل من خلال التحقيقات التي تصب في المصلحة العامة وتقديم توصياتهم بشأن التحسين. بينما يجب اتخاذ تدابير تخفيفية لحماية خصوصية المستخدمين، بإمكان Meta وينبغي عليها السماح بفهم أكبر لكيفية عمل منصاتها.

VIII. ملحق توصيات وإجراءات التنفيذ

قدّم المجلس عدة توصيات لشركة Meta في رأيه الاستشاري بشأن السياسة. يجمع هذا الملحق بين توصيات وإجراءات التنفيذ لمراقبة التقدم الذي تحرزه Meta. ينبغي على Meta توفير معلومات عن أعمال التنفيذ في تقاريرها ربع السنوية المقدمة إلى المجلس. علاوة على ذلك، ينبغي على Meta عقد اجتماع نصف سنوي للمسؤولين المختصين رفيعي المستوى لإطلاع المجلس على أعمال تنفيذ توصيات الرأي الاستشاري بشأن السياسة.

#	التوصية	إجراءات التنفيذ
حوكمة منع الأخطاء التي تستند إلى الكيان		
1	يجب أن تقسّم Meta أي برنامج من برامج منع أخطاء فرط الإنفاذ يستند إلى القوائم إلى أنظمة منفصلة، إما حسب المسارات المنفصلة أو ترتيب الأولويات: نظام لحماية التعبير تماشيًا مع مسؤوليات Meta تجاه حقوق الإنسان، ونظام لحماية التعبير الذي ترى Meta أنه يمثل أولوية تجارية ويندرج خارج هذه الفئة.	توفر Meta للمجلس معلومات تفصيلية عن كيفية تقسيم كل من التضمين والتشغيل لفئات الكيانات هذه. وتنتشر Meta تفاصيل هذه الأنظمة في مركز الشفافية. <i>الإنفاذ</i>
2	يجب أن تتأكد Meta من أن مسار المراجعة وهيكل اتخاذ القرارات بشأن المحتوى الذي ينطوي على آثار تتعلق بحقوق الإنسان أو المصلحة العامة بما في ذلك مسارات التصعيد، يخلو من اعتبارات الأعمال. ويجب على Meta اتخاذ الخطوات اللازمة للتأكد من أن الفريق المسؤول عن هذا النظام لا يتبع فرق السياسة العامة أو العلاقات الحكومية أو الفرق المسؤولة عن إدارة العلاقات مع أي مستخدمين متأثرين.	توفر Meta للمجلس معلومات تفصيلية عن مسارات اتخاذ القرارات والفرق المشاركة في الإشراف على المحتوى الذي ينطوي على آثار تتعلق بحقوق الإنسان. <i>الإنفاذ</i>
3	يجب على Meta تحسين الطريقة التي يقوم من خلالها دفع العمل المخصص للوفاء بمسؤوليات Meta تجاه حقوق الإنسان بتضمين الخبرة السياقية واللغوية في المراجعة المحسنة، وتحديدًا في مستويات اتخاذ القرار.	توفر Meta للمجلس معلومات تفصيلية عن التحسن الذي أحرزته في العملية الحالية لتضمين الخبرة السياقية واللغوية عند النظر في القرارات التي تستند إلى السياق واستثناءات السياسة. <i>الإنفاذ</i>
4	ينبغي على Meta وضع معايير واضحة وعلنية لأهلية الانضمام إلى نظام منع الأخطاء الذي يستند إلى القوائم. ويجب أن تفرق هذه المعايير بين المستخدمين الذين يستحقون حماية إضافية من منظور حقوق الإنسان، والمستخدمين الذين يتم تضمينهم لأسباب تجارية.	تُصدر Meta تقريرًا أو تحديثًا لمركز الشفافية يعرض تفاصيل معايير الأهلية للمراجعة المحسنة التي تستند إلى القوائم لمختلف فئات المستخدمين الذين سيضمهم البرنامج. <i>الشفافية</i>
5	يجب على Meta وضع عملية تتيح للمستخدمين تقديم طلبات للتمتع بالحماية التي يكفلها نظام منع أخطاء فرط الإنفاذ حال استيفائهم للمعايير التي تحددها الشركة علنيًا. يجب أن تكون الجهات الحكومية مؤهلة للإضافة أو تقديم الطلبات بناءً على هذه	تتخذ Meta نظام تقديم علني يسهل الوصول إليه ويتسم بالشفافية لأي حماية من فرط الإنفاذ تستند إلى القوائم، يحدد بالتفصيل أغراض النظام وكيفية تقييم الشركة لطلبات الانضمام. تقوم Meta بتضمين عدد الكيانات التي انضمت بنجاح إلى برنامج منع الأخطاء من خلال

<p>تقديم طلب، مع ذكر البلد والفئة كل عام في مركز الشفافية.</p> <p>الإنفاد</p>	<p>المعايير والشروط دون أن تحصل على أي تفضيلات أخرى.</p>	
<p>توفر Meta للمجلس التجربة الكاملة التي يخوضها المستخدم للانضمام إلى أي نظام يستند إلى القوائم، بما في ذلك كيفية التزام المستخدمين باتباع سياسة المحتوى وكيفية إخطارهم بالتغييرات التي تطرأ عليها.</p> <p>الإنفاد</p>	<p>يجب على Meta التأكد من اشمال عملية الإدراج في الأنظمة التي تستند إلى القوائم، بغض النظر عن القائم ببدء العملية (الكيان نفسه أو Meta) على ما يلي بحد أدنى: (1) التزام إضافي صريح يعبر فيه المستخدم عن التزامه بسياسات المحتوى التي تقرها Meta؛ و(2) إقرار بالقواعد الخاصة للبرنامج؛ و(3) نظام تتم من خلاله مشاركة التغييرات التي تطرأ على سياسات المحتوى مع المستخدمين بشكل استباقي.</p>	6
<p>توفر Meta للمجلس معلومات عن كيفية تفاعل الشركة مع المجتمع المدني لتحديد الأهلية التي تستند إلى القوائم. توفر Meta بيانات في مركز الشفافية، مصنفة حسب البلد، حول كيفية إضافة الكيانات نتيجة لتفاعل المجتمع المدني مقارنة بتحديد الكيانات استباقيًا بواسطة Meta.</p> <p>الإنفاد</p>	<p>يجب على Meta تعزيز تفاعلها مع المجتمع المدني لأغراض إنشاء القوائم والترشيحات. يجب أن يتمكن المستخدمون ومنظمات المجتمع المدني الموثوق بها من ترشيح الآخرين الذين يستوفون المعايير. يكون هذا الأمر مُلغًا بشكل خاص في البلدان التي يكون وجود الشركة فيها محدودًا بدرجة لا تسمح بتحديد المرشحين للإدراج في القوائم بشكل مستقل.</p>	7
<p>توفر Meta للمجلس وثائق داخلية توضح بالتفصيل الفرق التي تتولى إنشاء القوائم وموقعها داخل المنظمة.</p> <p>الإنفاد</p>	<p>يجب على Meta الاستعانة بفرق متخصصة، ومستقلة عن التأثير السياسي أو الاقتصادي، بما في ذلك فرق السياسة العامة في Meta، لتقييم الكيانات التي سيتم إدراجها في القوائم. ولضمان استيفاء المعايير، يجب أن يتأكد الموظفون المتخصصون، عبر الاستفادة من المعطيات المحلية، من تطبيق معايير الإدراج بموضوعية.</p>	8
<p>توفر Meta للمجلس معلومات تفصيلية عن عملية إضافة الكيانات الجديدة إلى القوائم، بما في ذلك عدد الموظفين الذين يجب أن يوافقوا على الإدراج والفرق التي ينتمون إليها.</p> <p>الإنفاد</p>	<p>يجب على Meta اشتراط مشاركة أكثر من موظف في العملية النهائية لإضافة الكيانات الجديدة إلى أي قوائم لأنظمة منع أخطاء النتائج الإيجابية الخاطئة. يجب أن ينتمي هؤلاء الأشخاص لفرق مختلفة لكنها مرتبطة.</p>	9
<p>توفر Meta للمجلس معلومات تفصيلية عن حد إجراءات الإنفاذ ضد الكيانات التي تم إلغاء الحماية المكفولة لها بموجب نظام يستند إلى القوائم، بما في ذلك الإشعارات المرسلة إلى المستخدمين عند حصولهم على إنذارات ضد أهليتهم، وعند تجريدهم من الأهلية، والخيارات المتاحة للطعن. ويجب أيضًا أن تقدم للمجلس بيانات عن عدد الكيانات التي تتم إزالتها كل عام بسبب نشر المحتوى المخالف.</p> <p>الإنفاد</p>	<p>يجب على Meta وضع معايير واضحة للإزالة. يجب أن يتمثل أحد المعايير في حجم المحتوى المخالف المنشور بواسطة الكيان. يجب أن يستند التجريد من الأهلية إلى نظام إنذارات شفاف، يتم فيه تحذير المستخدمين من أن تكرار المخالفات قد يؤدي إلى الإزالة من النظام و/أو من منصات Meta. ويجب أن يحصل المستخدمون على فرصة للطعن على تلك الإنذارات من خلال عملية عادلة يسهل الوصول إليها.</p>	10
<p>توفر Meta للمجلس بيانات عن مقدار ونوع الكيان وسبب الإزالة من قوائم الكيانات نتيجة لعمليات التدقيق،</p>	<p>يجب على Meta وضع معايير وعمليات واضحة للتدقيق. إذا لم تعد الكيانات تفي بمعايير الأهلية، يجب إزالتها على الفور من</p>	11

<p>بالإضافة إلى جدول زمني لإجراء عمليات التدقيق بشكل دوري.</p> <p>الإنفاد</p>	<p>النظام. يجب على Meta مراجعة كل الكيانات المنضمة إلى أي برنامج من برامج منع الأخطاء مرة واحدة على الأقل سنويًا. يجب كذلك وجود بروتوكولات واضحة لتقليل هذه الفترة الزمنية عند الضرورة.</p>	
شفافية القوائم		
<p>تميّز Meta جميع الكيانات في هذه الفئات ككيانات مستفيدة من برنامج منع أخطاء يستند إلى القوائم وتعلن التغيير في مركز الشفافية.</p> <p>الشفافية</p>	<p>يجب أن تميّز Meta بشكل علني صفحات وحسابات الكيانات التي تتلقى الحماية استنادًا إلى القوائم في الفئات التالية: جميع الجهات الحكومية والمرشحين السياسيين وجميع شركاء الأعمال وجميع الجهات الإعلامية وجميع الشخصيات العامة الأخرى المدرجة بسبب الفائدة التجارية التي تعود على الشركة من تجنب النتائج الإيجابية الخاطئة. قد تختار فئات أخرى من المستخدمين أن يتم تحديد هويتها.</p>	12
<p>توفر Meta للمجلس الإشعارات المرسله للمستخدمين الذين يبلغون عن محتوى منشور بواسطة مستخدمين يستفيدون من المراجعة الإضافية مع تأكيد التطبيق العالمي والبيانات التي توضح ظهور هذه الإشعارات باستمرار للمستخدمين.</p> <p>الإنفاد</p>	<p>يجب على Meta إخطار المستخدمين الذين يبلغون عن المحتوى المنشور بواسطة كيان محدد علنًا كمستفيد من المراجعة الإضافية أنه سيتم تطبيق إجراءات خاصة، مع توضيح الخطوات واحتمالية طول المدة اللازمة لاتخاذ القرار.</p>	13
<p>توفر Meta للمجلس (1) الإشعارات المرسله إلى المستخدمين لإبلاغهم بالإدراج في برنامج المراجعة المحسنة التي تستند إلى القوائم مع توفير خيار للرفض؛ وتعمل Meta على (2) تقديم تقارير علنية في مركز الشفافية عن الأعداد السنوية للكيانات التي رفضت الإدراج، مصنفة حسب البلد.</p> <p>الإنفاد</p>	<p>يجب على Meta إبلاغ كل الكيانات التي تدرجها في القوائم لتحصل على مراجعة محسنة وأن تمنحها فرصة لرفض الإدراج.</p>	14
المراجعة المحسنة وترتيب الأولويات		
<p>تقدم Meta إلى المجلس وثائق توضح عملية دراسة هذه التوصية وحيثيات قرار الشركة بشأن مدى تنفيذها وأن تنشر مبررات ذلك في مركز الشفافية.</p> <p>الإنفاد</p>	<p>يجب على Meta النظر في إمكانية تخصيص حد أدنى من السعة الاستيعابية للمراجعة بواسطة الفرق التي يمكنها تطبيق كل سياسات المحتوى (على سبيل المثال، فريق الاستجابة المبكرة) لمراجعة المحتوى الذي تم تمييزه من خلال أنظمة منع الأخطاء التي تستند إلى المحتوى.</p>	15
<p>توفر Meta للمجلس بيانات توضح الانخفاض ربع السنوي في المدة اللازمة للتوصل إلى قرار بشأن كل عناصر المحتوى التي تحصل على مراجعة محسنة، مصنفة حسب فئة الإدراج والبلد.</p> <p>الإنفاد</p>	<p>يجب على Meta اتخاذ تدابير للتأكد من اتخاذ قرارات المراجعة الإضافية لأنظمة منع الأخطاء التي تؤخر الإنفاذ بأسرع ما يمكن. ويجب ضخ الاستثمارات وإجراء التغييرات الهيكلية لزيادة أعداد فرق المراجعة بحيث يتوفر ويعمل المراجعون في المناطق الزمنية المعنية متى تم تمييز المحتوى على أنه يحتاج إلى مراجعة بشرية محسنة.</p>	16

<p>تقوم Meta بتحديث مركز الشفافية بتوضيح أسلوبها الجديد بشأن إجراء الإنفاذ أثناء خضوع المحتوى للمراجعة المحسنة وتقديم للمجلس معلومات تفصيلية عن تبعات إجراء الإنفاذ الذي ستطبقه استنادًا إلى المعايير الخاصة بالمحتوى. وتشارك Meta مع المجلس بيانات عن تطبيق هذه التدابير وتأثيرها</p> <p style="text-align: right;">/الإنفاذ</p>	<p>يجب ألا تؤخر Meta كل الإجراءات ضد المحتوى الذي يحتمل أن يمثل مخالفة شديدة ويجب عليها استكشاف وضع شاشات بيئية أو إزالة ذلك المحتوى أثناء انتظار المراجعة المحسنة. ويجب أن يستند الفرق بين الإزالة أو الإخفاء أو خفض الترتيب إلى تقييم الضرر، وقد يستند، على سبيل المثال، إلى سياسة المحتوى المحتمل أنه قد تم انتهاكها. في حالة إخفاء المحتوى لهذه الأسباب، يجب تقديم إشعار للمستخدمين في مكان المحتوى يشير إلى أنه قيد المراجعة.</p>	17
الموارد		
<p>توفر Meta للمجلس بيانات توضح الانخفاض ربع السنوي في إجمالي المحتوى المتراكم وعدد الأيام التي بها محتوى متراكم في قوائم انتظار مراجعة التحقق الشامل.</p> <p style="text-align: right;">/الإنفاذ</p>	<p>يجب ألا تقوم Meta بتشغيل هذه البرامج على نحو يسمح بوجود تراكمات. ومع هذا، يجب ألا تحقق Meta تقدمًا فيما يتعلق بالسعة النسبية للمراجعة من خلال الرفع المصطنع لحد أداة التصنيف أو توجيه خوارزمياتها لتحديد عناصر محتوى أقل.</p>	18
<p>توفر Meta للمجلس وثائق داخلية توضح بالتفصيل توزيع وقت وحجم المراجعة بين الأنظمة التي تستند إلى الكيان وتلك التي تستند إلى المحتوى.</p> <p style="text-align: right;">/الإنفاذ</p>	<p>يجب ألا ترتب Meta أولوية المراجعة الثانوية التي تستند إلى الكيان تلقائيًا وأن تجعل قدرًا كبيرًا من المراجعة التي تستند إلى المحتوى المحددة خوارزميًا تعتمد على سعة المراجعة الإضافية.</p>	19
<p>توفر Meta للمجلس معلومات توضح النسبة المئوية للمحتوى الذي تتم مراجعته بواسطة فرق يمكنها تطبيق الاستثناءات والنظر في السياق لأنه قد تمت مشاركته بواسطة كيان مؤهل أو نظرًا لتحديده من خلال الخوارزميات على أنه يستحق المراجعة المحسنة، على أن تصنف هذه المعلومات حسب نظام منع الأخطاء (على سبيل المثال: المراجعة الثانوية العامة مقابل المراجعة الثانوية للاستجابة المبكرة).</p> <p style="text-align: right;">/الإنفاذ</p>	<p>يجب على Meta التأكد من أن المحتوى الذي يتلقى أي نوع من المراجعة المحسنة نظرًا لأهميته من منظور حقوق الإنسان، بما في ذلك المحتوى الذي يحظى بأهمية عامة، يحظى بالمراجعة من خلال فرق يمكنها تطبيق كل الاستثناءات والنظر في السياق.</p>	20
الوسائل التلقائية لإيقاف الإنفاذ ("التصحيات الفنية")		
<p>تنشر Meta عدد الكيانات التي تستفيد حاليًا من "التصحيات الفنية" على أساس سنوي، مع توضيح سياسات المحتوى التي تم إيقاف الإنفاذ بشأنها.</p> <p style="text-align: right;">/الإنفاذ</p>	<p>يجب على Meta وضع معايير واضحة لتطبيق أي وسائل تلقائية لإيقاف الإنفاذ ("التصحيات الفنية")، وألا تسمح لهذه الوسائل بإيقاف الإنفاذ عند وجود انتهاكات شديدة الخطورة لسياسة المحتوى. ويجب مشاركة ما لا يقل عن فريقين لهما تبعات مختلفة في منح التصحيحات الفنية لتوفير التدقيق عبر فرق مختلفة.</p>	21
<p>توفر Meta معلومات للمجلس عن عملياتها الدورية لتدقيق القوائم.</p>	<p>يجب على Meta إجراء تدقيقات دورية للتأكد من أن الكيانات التي تستفيد من</p>	22

الإنفاد	الوسائل التلقائية لإيقاف الإنفاذ ("التصحيات الفنية") تفي بكل معايير الإدراج. ويجب مشاركة ما لا يقل عن فريقين لهما تبعيات مختلفة في عمليات التدقيق تلك لتوفير التدقيق عبر فرق مختلفة.	
الإنفاد	يجب على Meta إجراء عمليات تدقيق دورية متعددة الفرق للبحث بشكل استباقي ودوري عن المعوقات غير المتوقعة وغير المقصودة التي تمنع الإنفاذ والتي قد تنتج عن خطأ بالنظام.	23
العدالة الإجرائية		
الإنفاد	يجب على Meta التأكد من إمكانية تقديم طعون إليها بشأن كل عناصر المحتوى التي لا تصل إلى أعلى مستويات المراجعة الداخلية.	24
الإنفاد	يجب على Meta ضمان توفير فرصة التقدم بطعن إلى المجلس لكل المحتوى الذي يحق للمجلس مراجعته بموجب وثائقه الحاكمة، بغض النظر عن مدى وصول المحتوى إلى أعلى مستويات المراجعة داخل Meta.	25
التعلم والتحسين		
الإنفاد	ينبغي على Meta استخدام البيانات التي تجمعها لتحديد "الكيانات التي تعرضت لفرط الإنفاذ تاريخياً" والاستعانة بها في تحسين ممارسات الإنفاذ على نطاق واسع. يجب أن تقيس Meta فرط الإنفاذ لهذه الكيانات وينبغي عليها استخدام هذه البيانات للمساعدة على تحديد الكيانات الأخرى التي تتعرض لفرط الإنفاذ. ويجب أن يكون تقليل فرط الإنفاذ هدفاً واضحاً وذو أولوية كبيرة بالنسبة إلى الشركة.	26
الإنفاد	ينبغي على Meta استخدام الاتجاهات فيما يتعلق بمعدلات الإلغاء لتحديد هل ستتم العودة افتراضياً لإجراء الإنفاذ الأصلي خلال إطار زمني أقصر أو تحديد إجراءات الإنفاذ الأخرى التي سيتم تطبيقها إلى أن تنتهي المراجعة. إذا كان معدل الإلغاء منخفض باستمرار لمجموعة معينة من انتهاكات السياسة أو المحتوى المنشور بلغات	27

	معينة، على سبيل المثال، يجب أن تعابير Meta باستمرار مدى سرعة وتدخل إجراء الإنفاذ الذي ينبغي عليها تطبيقه.	
تحسين المساءلة الخاصة بالبرنامج		
28	يجب على Meta إجراء مراجعات دورية لمختلف جوانب نظام المراجعة المحسنة، بما في ذلك المحتوى الذي يأخذ أطول وقت للتوصل إلى قرار بشأنه والمحتوى المخالف البارز المتروك على المنصة.	تنشر Meta نتائج مراجعات نظام التحقق الشامل على أساس سنوي، بما في ذلك ملخصات التغييرات التي تمت نتيجة لهذه المراجعات. <i>الشفافية</i>
29	يجب أن تنشر Meta بشكل علني أدوات القياس التي تحسب الآثار السلبية لتأخير الإنفاذ نتيجة لأنظمة المراجعة المحسنة، مثل المشاهدات التي حصل عليها المحتوى الذي تم الاحتفاظ به على المنصة نتيجة لأنظمة منع الأخطاء ثم تبين لاحقاً أنه مخالف. وكجزء من النشر العلني، يجب على Meta تحديد أساس لأدوات القياس هذه وأن تعلن مستهدفات خفضها.	تقوم Meta بتضمين أداة قياس أساسية أو أكثر توضح الآثار السلبية لتأخير الإنفاذ انتظاراً لآليات المراجعة المحسنة في تقرير إنفاذ معايير المجتمع، مع مستهدفات خفض أدوات القياس هذه والتقدم المحرز نحو تحقيق تلك المستهدفات. <i>الشفافية</i>
30	يجب على Meta نشر تقارير شفافية دورية تركز تحديداً على أنظمة منع النتائج الإيجابية الخاطئة التي تؤخر الإنفاذ. يجب أن تحتوي تلك التقارير على بيانات تسمح للمستخدمين والعامة بفهم طريقة عمل تلك البرامج وما قد تمثله من تبعات على الخطاب العام. بحد أدنى، يوصي المجلس Meta بتضمين: أ. معدلات الإلغاء لأنظمة منع أخطاء النتائج الإيجابية الخاطئة، مصنفة وفقاً للعوامل المختلفة. على سبيل المثال، أوصى المجلس Meta بإنشاء تدفقات مختلفة لمختلف فئات الكيانات المختلفة أو المحتوى استناداً إلى التعبير ومستوى المخاطر. يجب الإبلاغ عن معدل الإلغاء لكل الأنظمة التي تستند إلى الكيان والتي تستند إلى المحتوى، وفئات الكيانات والمحتوى المضمنة. ب. العدد الإجمالي والنسبة المئوية لسياسات التصعيد فقط التي تم تطبيقها نتيجة لبرامج منع أخطاء النتائج الإيجابية الخاطئة نسبة إلى إجمالي قرارات الإنفاذ. ج. وسط ومتوسط الوقت اللازم للتوصل إلى القرار النهائي بشأن المحتوى الخاضع لبرامج منع أخطاء النتائج الإيجابية الخاطئة، مع تصنيفها حسب البلد واللغة. د. بيانات مجمعة عن أي قوائم مستخدمة لبرامج منع الأخطاء، بما في ذلك نوع الكيان والمنطقة.	تصدر Meta تقارير شفافية سنوية تتضمن أدوات القياس هذه. <i>الشفافية</i>

	<p>هـ. معدل عمليات الإزالة الخاطئة (النتائج الإيجابية الخاطئة) مقابل كل المحتوى الذي تمت مراجعته، بما في ذلك إجمالي الضرر الناتج عن هذه النتائج الإيجابية الخاطئة ويقاس كإجمالي المشاهدات المتوقعة للمحتوى (أي فرط الإنفاذ)</p> <p>و. معدل القرارات الخاطئة للإبقاء على المحتوى (النتائج السلبية الخاطئة)، بما في ذلك إجمالي الضرر الناتج عن هذه النتائج الإيجابية الخاطئة ويقاس كمجموع المشاهدات التي حققها المحتوى (أي نقص الإنفاذ)</p>	
<p>إضافة قسم إلى مركز الشفافية يوضح مصفوفة أنظمة منع الأخطاء (يدرك المجلس احتمال وجود محاولات عدائية من المستخدمين لتجاوز الإنفاذ، وقد تختار Meta تلخيص بعض نقاط ممارسات الإنفاذ التي تتبعها).</p> <p><i>الشفافية</i></p>	<p>يجب على Meta توفير معلومات أساسية في مركز الشفافية عن طريقة عمل أي نظام تستخدمه لمنع الأخطاء من شأنه تحديد هوية الكيانات أو المستخدمين من أجل حماية إضافية.</p>	31
<p>تكشف Meta عن مسار للباحثين الخارجيين للحصول من خلاله على بيانات غير علنية عن برامج منع أخطاء النتائج الإيجابية الخاطئة.</p> <p><i>الشفافية</i></p>	<p>يجب على Meta إنشاء مسار للباحثين الخارجيين للوصول إلى البيانات غير العلنية عن برامج منع أخطاء النتائج الإيجابية الخاطئة من شأنه أن يسمح لهم بفهم البرنامج بشكل كامل من خلال التحقيقات التي تصب في المصلحة العامة وتقديم توصياتهم بشأن التحسين. يتفهم المجلس أن المخاوف المتعلقة بخصوصية البيانات يجب أن تتطلب تدقيقًا صارمًا وتجميعًا للبيانات.</p>	32