

Rapport sur les droits humains



Sommaire

À propos de ce rapport	03
Résumé analytique	06
Gestion des risques en matière de droits humains	09
1. Liberté d'opinion et d'expression	10
2. Confidentialité	13
3. Égalité et non-discrimination	14
4. Vie, liberté et sécurité de la personne	15
5. Intérêt supérieur de l'enfant	16
6. Participation publique, processus de vote et éligibilité	16
7. Liberté de réunion et d'association	17
8. Droit à la santé	17
Principales problématiques	18
Innové en respectant les droits humains	18
Favoriser l'accessibilité grâce à la technologie wearable	19
Continuer à innover dans l'IA de manière responsable	19
Fraude et arnaques	22
Démanteler les réseaux criminels spécialisés dans les arnaques en Asie du Sud-Est	22
Comment se préparer aux crises et y répondre	24
République démocratique du Congo	25



Implication des parties prenantes	26
Partenaires de confiance	27
Étude de cas : Comment prévenir le recrutement de jeunes en Amérique latine	28
Organisations internationales	29
Transparence et recours	30
Annexe	33
Comment les droits humains sont régis et gérés chez Meta	34
Formation du personnel de Meta aux droits humains	34
Liens vers les rapports référencés	34





À propos de ce rapport

Ce Rapport sur les droits humains annuel dresse le bilan des actions menées entre le 1er janvier et le 31 décembre 2025. Nous rendons compte des services et produits Meta, notamment Facebook, Messenger, Instagram, WhatsApp, Threads et Reality Labs.

Il s'appuie sur les efforts de Meta pour respecter les droits humains et reflète les progrès réalisés dans le cadre de nos engagements vis-à-vis des [Principes directeurs relatifs aux entreprises et aux droits de l'homme des Nations Unies](#) et de notre [Politique relative aux droits humains au sein de l'entreprise](#). Le rapport montre comment nous avons appliqué ces principes dans l'entreprise en 2025 et fournit des conseils permettant de trouver des informations approfondies.

Le contenu de ce rapport se base sur notre [Évaluation globale des risques principaux en matière de droits humains \(EGRP\)](#), réalisée en 2022. L'objectif de cette évaluation était d'identifier nos incidences négatives potentielles sur les droits humains¹ des personnes qui utilisent nos produits et d'autres personnes susceptibles d'être concernées par nos actions. Ce rapport décrit ces risques principaux ainsi que des exemples de nos actions et de nos mesures d'atténuation en 2025.

Les droits humains sont restés un sujet d'une importance cruciale pour notre entreprise et nos parties prenantes en 2025. Nous avons raccourci et simplifié le rapport tout en nous efforçant de fournir une représentation fidèle de notre travail au sein de plusieurs équipes et de l'engagement des parties prenantes à travers le monde.

Politiques et progrès

Outre ce Rapport sur les droits humains, Meta rend compte chaque année de ses politiques et de ses progrès grâce aux mécanismes suivants :



-  Rapport annuel
-  Déclaration de procuration
-  Rapport sur les pratiques professionnelles responsables
-  Espace modération
-  Rapport sur la durabilité
-  Rapport CDP sur le changement climatique
-  Pacte mondial des Nations Unies

Ce rapport complète le [Rapport sur les pratiques professionnelles responsables de Meta](#). Nous [rapportons](#) séparément nos efforts pour identifier et atténuer les risques d'esclavage moderne et de traite d'êtres humains dans nos activités commerciales et nos chaînes d'approvisionnement. De plus, nous nous conformons aux rapports obligatoires au niveau national et de l'Union européenne, disponibles dans notre [Espace modération](#). Des liens vers d'autres publications de Meta peuvent être trouvés dans l'[Annexe](#) du présent rapport.

[Aller à l'annexe](#)

¹ L'expression « incidence négative sur les droits humains » est tirée des Principes directeurs relatifs aux entreprises et aux droits de l'homme des Nations Unies et désigne une incidence qui se produit lorsqu'une action supprime ou réduit la capacité d'une personne à jouir de ses droits humains.



Notre [Politique relative aux droits humains au sein de l'entreprise](#) s'applique à l'ensemble de l'entreprise. Chaque service et entité de Meta dispose de ses propres politiques et procédures, qui peuvent avoir des incidences différentes sur les droits humains. Le présent rapport fait référence à des actions de Meta en tant qu'entreprise, qui peuvent concerner une ou plusieurs entités de Meta. Les déclarations n'ont pas pour objectif de suggérer que Meta a entrepris les mêmes actions pour toutes les entités et/ou dans toutes les circonstances².



² La discussion sur la modération du contenu et les actions connexes entreprises sur Facebook et Instagram présentée dans ce rapport ne concerne pas WhatsApp, à moins qu'il ne soit spécifiquement indiqué qu'une politique ou action donnée s'y applique. En outre, bien que de nombreuses actions décrites dans ce rapport s'appliquent à Facebook et Instagram, il existe des distinctions intentionnelles dans les politiques et les procédures entre les services. Si une politique est qualifiée de politique « Facebook », elle peut ne pas s'appliquer à Instagram. Aucune déclaration du présent rapport n'est destinée à créer, ou ne doit être interprétée comme créant, de nouvelles obligations (juridiques ou autres) concernant l'application d'une politique ou d'une procédure à d'autres services ou entités.



Résumé analytique



Ce rapport est le cinquième Rapport sur les droits humains annuel de Meta. Il fait le point sur le travail réalisé par Meta en 2025 pour gérer les risques en matière de droits humains à grande échelle et respecter nos engagements envers les [Principes directeurs relatifs aux entreprises et aux droits de l'homme des Nations Unies](#).



Évaluation des risques principaux

En 2025, nos priorités reflétaient les risques principaux identifiés dans l'[Évaluation globale des risques principaux en matière de droits humains \(EGRP\)](#) de 2022 : liberté d'opinion et d'expression, confidentialité, égalité et non-discrimination, vie, liberté et sécurité de la personne, intérêt supérieur de l'enfant, participation publique, processus de vote et éligibilité, liberté de réunion et d'association, et droit à la santé.

→ Lire la suite

Innovation

Meta a cherché à repousser les limites de l'intelligence artificielle (IA), en tenant toujours compte des droits humains dans notre processus de développement. Nous avons lancé de nouveaux modèles, cadres de sécurité et produits, tels que [les lunettes Meta Ray-Ban Display et le Meta Neural Band](#). Ce bracelet utilise les signaux générés par les muscles pour un contrôle mains libres des lunettes IA. Nous avons développé ces produits de manière responsable, en tenant compte de l'[accessibilité](#) et conformément au droit à l'égalité et à la non-discrimination.

→ Lire la suite

Lutte contre la fraude et les arnaques

Nous avons également pris davantage de mesures pour lutter contre la fraude et les arnaques. [Ces mesures](#) reposaient principalement sur la mise en application des politiques de la plateforme ainsi que sur des partenariats interrégionaux associant des acteurs du secteur, les forces de l'ordre et des organisations de la société civile locales. De nombreux centres d'arnaques ont recours à la traite d'êtres humains et au travail forcé pour mener à bien leurs activités. Nos mesures d'atténuation ont non seulement contribué au respect du droit à la vie, à la liberté et à la sécurité, mais aussi à l'interdiction de l'esclavage, de la servitude et du travail forcé.

→ Lire la suite

Sécurité des enfants et des jeunes

Nous avons également poursuivi notre engagement en faveur de la [sécurité des enfants et des jeunes](#). Entre autres initiatives, nous avons annoncé qu'Instagram allait [mettre à jour](#) l'expérience de contenu des comptes Ado en s'inspirant des critères de classification des films pour les 13 ans et plus et des commentaires des parents*. Nous nous sommes pour cela appuyés sur les conseils de spécialistes et sur le principe de développement des capacités de l'enfant décrit dans la [Convention relative aux droits de l'enfant des Nations Unies](#).

→ Lire la suite

Réponse aux crises

Notre Rapport sur les droits humains 2024 comprenait des informations détaillées sur la façon dont Meta intègre les principes des droits humains à [notre préparation et notre réponse aux crises](#), en nous appuyant notamment sur notre [protocole de politique de crise](#). Ce travail s'est poursuivi en 2025. Nous avons appliqué le protocole de politique de crise pour répondre à différentes situations, de celles observées en Indonésie et dans la région de Cachemire en Inde aux manifestations de la génération Z à Madagascar, en Serbie et dans d'autres pays. Dans ce rapport, nous partageons une étude de cas sur notre réponse à la crise en République démocratique du Congo.

→ Lire la suite



Implication des parties prenantes Conseil de surveillance

Notre [Politique relative aux droits humains au sein de l'entreprise](#) soutient notre engagement proactif envers les parties prenantes. En 2025, nous avons interagi avec un large éventail de groupes dédiés aux droits humains, de communautés vulnérables, de membres de la société civile, d'universitaires, de groupes de réflexion et d'organismes de réglementation. Les principaux sujets abordés ont été l'approche de Meta en matière d'IA et de modération du contenu, l'impact des [changements apportés à nos politiques et produits en janvier](#), et nos initiatives de lutte contre la fraude et les arnaques.

Nous avons continué de faire appel à nos [partenaires de confiance](#) dans le monde entier pour identifier des tendances et mieux comprendre l'impact du contenu et des comportements en ligne sur les communautés locales. Ils ont apporté leur contribution et identifié des contenus potentiellement en infraction dans plus de 20 pays et régions.

→ Lire la suite

Nous tenons également à faire le point sur le travail du [Conseil de surveillance](#). Le Conseil de surveillance est un organisme indépendant qui émet des décisions contraignantes et des recommandations non contraignantes. Pendant la période considérée, le Conseil de surveillance a évalué plusieurs cas en lien avec nos efforts pour respecter les droits humains, notamment la liberté d'expression, le droit à la santé et le droit à l'égalité et à la non-discrimination, entre autres.

Par exemple, en réponse à une recommandation, Meta a [augmenté](#) la granularité des notifications existantes afin de fournir aux utilisateurs des informations plus détaillées lorsque leur contenu est restreint à la suite d'un signalement officiel des autorités indiquant qu'il enfreint les lois locales.

→ Lire la suite

Gestion des demandes gouvernementales

En 2025, nous avons maintenu notre engagement à respecter la liberté d'expression et la confidentialité vis-à-vis de la [Global Network Initiative](#), notamment en répondant aux demandes gouvernementales de restriction de contenu. Le nombre de demandes gouvernementales de ce type a augmenté. En plus de nos rapports réguliers sur l'intégrité, nous avons publié des [études de cas](#) au sujet de demandes gouvernementales de restriction du discours politique en Autriche, en Colombie, en Côte d'Ivoire, en Allemagne, en Inde, en Irak, en Italie, au Kazakhstan, en Malaisie, en Nouvelle-Zélande, en Pologne, en Roumanie, à Singapour, en Afrique du Sud, en Tanzanie, en Turquie et au Vietnam. En 2025, Meta a également fait l'objet de sa quatrième évaluation par la GNI (couvrant la période d'avril 2022 à décembre 2024). Le Conseil d'administration de la GNI a décidé à l'unanimité que Meta avait fait « des efforts de bonne foi pour mettre en œuvre les principes de la GNI, avec une amélioration au fil du temps ».

Voir les études de cas

Gestion des risques en matière de droits humains

Les [Principes directeurs relatifs aux entreprises et aux droits de l'homme des Nations Unies](#) mettent en lumière le fait que les entreprises doivent identifier leurs incidences négatives sur les droits humains afin de les prévenir ou de les atténuer efficacement.

Étant donné l'ampleur des opérations de Meta et l'éventail des droits qu'elles peuvent impliquer, il est important d'anticiper et de gérer nos [risques principaux](#), bien que cela soit complexe. Nous gérons deux types de risques inhérents : ceux provenant de nos propres activités et ceux provenant des activités de tiers, notamment les personnes qui utilisent nos plateformes.



Le tableau ci-dessous présente nos risques principaux inhérents en matière de droits humains³, tels qu'ils sont définis dans notre Évaluation globale des risques principaux en matière de droits humains (EGRP) de 2022, que nous avons publiée dans notre [Rapport sur les droits humains 2022](#). Ce tableau fournit des exemples qui illustrent la manière dont nous avons fait face aux risques potentiels en 2025. Plus loin dans ce rapport, nous analyserons plus en détail certains de ces exemples ainsi que la manière dont nous avons géré les risques potentiels liés à l'intelligence artificielle (IA), aux élections et aux conflits.

1. Liberté d'opinion et d'expression

Le [droit à la liberté d'opinion et d'expression](#) comprend le droit de rechercher, de recevoir et de partager des informations et des idées en tout genre. Il s'agit d'un droit fondamental, essentiel à la protection de la dignité humaine, de l'autonomie individuelle et de la démocratie. La liberté d'expression fait partie intégrante de notre mission, en accord avec l'une de nos valeurs, qui est de donner la parole à tout le monde.

Exemples de risques principaux inhérents potentiels en matière de droits humains identifiés lors de l'EGRP

Les politiques de modération du contenu de Meta et leur mise en application peuvent limiter la liberté d'expression

Exemples de mesures prises par Meta pour remédier aux risques potentiels en 2025

En janvier 2025, nous avons mis à jour notre [politique en matière de conduite haineuse](#) pour permettre [plus de discussion](#) sur certains sujets qui font partie du débat public, comme l'immigration et le genre.

Nous avons axé nos systèmes automatisés sur la détection des infractions illégales et graves, comme le terrorisme, l'exploitation sexuelle d'enfants, la drogue, la fraude et les arnaques, tandis que pour les infractions moins graves, nous nous sommes appuyés sur les signalements d'utilisateurs. Nous avons également ajusté nos systèmes de façon à exiger un degré de fiabilité beaucoup plus élevé avant qu'un contenu ne soit retiré. Plus tard en 2025, nous avons fait état de l'amélioration de la précision de notre mise en application dans notre Rapport d'application des Standards de la communauté, qui fait partie de nos [Rapports sur l'intégrité](#).

Nous avons mis fin aux changements effectués en 2021 pour réduire la quantité de [contenu civique](#) que le public voit et avons commencé à recommander plus de contenu politique en fonction de signaux personnalisés, comme les contenus aimés.

³ Pour comprendre la différence entre un risque « inhérent » et « résiduel », consultez la page 11 du [Rapport sur les droits humains 2024 de Meta](#).



Exemples de risques principaux inhérents potentiels en matière de droits humains identifiés lors de l'EGRP

Les politiques de modération du contenu de Meta et leur mise en application peuvent limiter la liberté d'expression (suite)

Exemples de mesures prises par Meta pour remédier aux risques potentiels en 2025

Nous avons mis fin à notre programme de vérification tierce aux États-Unis et amorcé un programme de [Notes de la communauté](#). Dans le reste du monde, nous avons continué de faire appel à des [médias de vérification](#) indépendants de Meta et certifiés par une organisation neutre appelée l'[International Fact-Checking Network \(IFCN\)](#) ou, en Europe, par l'[European Fact-Checking Standards Network \(EFCSN\)](#), afin de lutter contre les fausses informations sur Facebook, Instagram et Threads.

→ Lire la suite

Limites gouvernementales excessives en matière de contenu

Nous nous efforçons de respecter nos engagements vis-à-vis de la [Global Network Initiative](#). En 2025, Meta a fait l'objet de sa quatrième évaluation par la GNI (couvrant la période d'avril 2022 à décembre 2024). Le Conseil d'administration de la GNI a décidé à l'unanimité que Meta avait fait « des efforts de bonne foi pour mettre en œuvre les principes de la GNI, avec une amélioration au fil du temps ». Nous avons également rendu compte de nos [réponses](#) aux demandes gouvernementales de données ou de restrictions de contenu ([ici](#) et [ici](#)). En 2025, dans un souci de transparence, nous avons publié des [études de cas](#) impliquant des discours politiques en Autriche, en Colombie, en Côte d'Ivoire, en Allemagne, en Inde, en Irak, en Italie, au Kazakhstan, en Malaisie, en Nouvelle-Zélande, en Pologne, en Roumanie, à Singapour, en Afrique du Sud, en Tanzanie, en Turquie et au Vietnam.



Lancement des Notes de la communauté

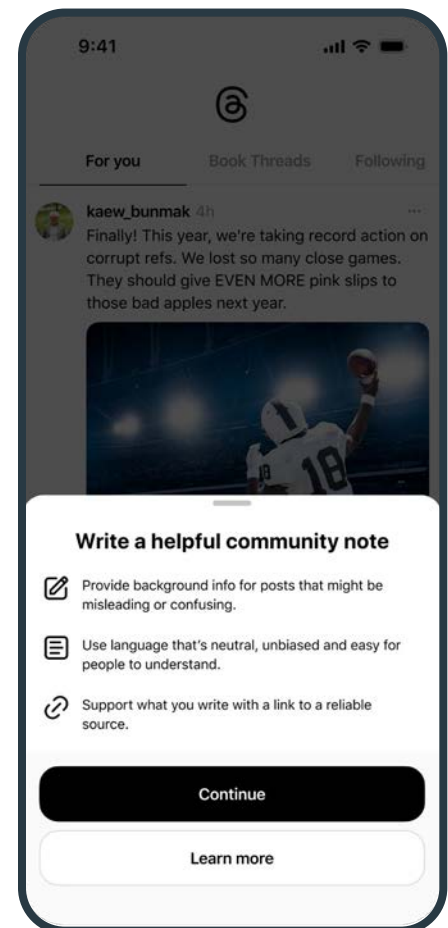
En 2025, Meta a lancé les [Notes de la communauté](#) aux États-Unis. Les Notes de la communauté sont un nouvel outil qui permet à la communauté d'ajouter du contexte au contenu pouvant être trompeur ou déroutant. Notre approche était axée sur les principes des droits humains et visait à améliorer l'accès à des informations considérées comme fiables par notre communauté d'utilisateurs, au moyen d'interventions proportionnées et respectueuses de la liberté d'expression.

Les notes sont rédigées et évaluées par des utilisateurs éligibles, à savoir des personnes âgées de 18 ans ou plus disposant d'un compte en règle et ayant effectué le contrôle de sécurité. Une note ne s'affiche que lorsqu'elle atteint un « consensus utile », ce qui signifie qu'un nombre suffisant de contributeurs qui étaient auparavant en désaccord conviennent désormais que la note est utile. Les noms des auteurs sont masqués afin de réduire la pression sociale et le risque de représailles, ainsi que pour encourager la participation, y compris sur des sujets sensibles.

Les notes de la communauté peuvent favoriser la liberté d'expression en apportant du contexte plutôt qu'en supprimant ou en rétrogradant des propos qui n'enfreignent aucune règle, ainsi qu'en élargissant la participation au-delà d'un groupe restreint de médias de vérification professionnels. Elles comprennent également des mécanismes de sécurité intégrés visant à prévenir les abus, notamment des exigences relatives à l'intégrité des comptes, des protections contre le spam et les URL et la mise en application des Standards de la communauté.

Meta a sollicité l'avis du Conseil de surveillance sur les facteurs que nous devrions prendre en considération pour décider d'introduire ou non les Notes de la communauté dans un pays, comme le faible niveau de liberté d'expression, l'absence de liberté de la presse et le taux d'alphabétisation numérique.

En novembre 2025, le Conseil de surveillance a accepté [la demande d'avis consultatif en matière de politique de Meta](#) afin de nous aider à prendre de meilleures décisions dans le développement du programme Notes de la communauté en dehors des États-Unis. Le Conseil de surveillance a publié son [avis](#) en mars 2026.





2. Confidentialité

Le [droit à la confidentialité](#) constitue une condition nécessaire au respect d'autres droits humains, tels que la liberté d'expression, la liberté de réunion et d'association, et la liberté de croyance et de religion. L'un des principes de base décrits dans notre [Politique relative aux droits humains au sein de l'entreprise](#) consiste à garantir la sécurité des utilisateurs et à protéger leur confidentialité.

Exemples de risques principaux inhérents potentiels en matière de droits humains identifiés lors de l'EGRP

Le contenu ou les comportements sur les applications Meta peuvent avoir une incidence négative sur les droits à la confidentialité et à la protection des données

Exemples de mesures prises par Meta pour remédier aux risques potentiels en 2025

En 2025, nous avons développé le [traitement privé pour Meta AI sur WhatsApp](#), qui repose sur un environnement d'exécution de confiance. Le traitement privé permet aux utilisateurs de demander à Meta AI de traiter leurs demandes dans un environnement cloud sécurisé et privé où personne, pas même Meta ou WhatsApp, ne peut accéder à leurs messages personnels. Cette technologie alimente désormais certaines fonctionnalités facultatives de Meta AI sur WhatsApp, comme les résumés de messages et l'aide à la rédaction. Une fois que le traitement privé a terminé de répondre à une demande, les messages ne sont plus stockés. Pour en savoir plus, consultez les [informations sur nos avancées en matière de confidentialité](#).

Nous avons engagé une [action en justice](#) contre l'entité à l'origine de CrushAI ou d'applications de « nudification » qui permettent de créer, à l'aide de l'IA, des images intimes ou sexuellement explicites représentant des personnes sans leur consentement. Nous avons également [annoncé](#) les étapes supplémentaires que nous avons prises pour lutter contre ces applications. Ces mesures sont venues compléter nos efforts de longue date visant à empêcher la diffusion en ligne d'images intimes non consensuelles, qu'elles soient réelles ou générées par l'IA, à l'aide d'outils tels que [StopNCII.org](#) et le [service Take It Down du National Center for Missing & Exploited Children](#).



3. Égalité et non-discrimination

Le [droit à l'égalité et à la non-discrimination](#) prévoit une protection égale contre la discrimination. En vertu de ce droit, nous n'autorisons pas les conduites haineuses sur nos plateformes, tel que défini dans notre [politique](#).

Exemples de risques principaux inhérents potentiels en matière de droits humains identifiés lors de l'EGRP

Certaines langues et certains dialectes peuvent être plus difficiles à modérer que d'autres

Exemples de mesures prises par Meta pour remédier aux risques potentiels en 2025

Nous avons [lancé](#) et rendu open source notre suite de modèles [Meta Omnilingual Automatic Speech Recognition \(ASR\)](#), qui offre des fonctionnalités de reconnaissance automatique de la parole dans plus de 1 600 langues, dont 500 langues à faibles ressources. En permettant la transcription de contenus oraux dans des langues qui étaient encore très peu couvertes, cette technologie a ouvert de nouvelles perspectives pour l'éducation et renforcé la liberté d'expression, chaque personne pouvant accéder au savoir et partager du contenu dans la langue qu'elle maîtrise le mieux.

Contenu ayant une incidence négative sur l'égalité et la non-discrimination (p. ex., conduite haineuse)

Nous avons évalué les risques associés aux [modifications de politiques et de produits](#) annoncées en janvier 2025 et réfléchi aux mesures d'atténuation pertinentes, telles que le recours à d'autres politiques et les signalements d'utilisateurs, pour aborder ce type de contenu. Dans les pays traversant des conflits ou des crises, nous pourrions appliquer la détection proactive par défaut pour certains types d'infractions, tels que les conduites haineuses. Par exemple, nous avons utilisé la détection proactive en Libye après avoir déployé le protocole de politique de crise.



4. Vie, liberté et sécurité de la personne

Le [droit à la vie, à la liberté et à la sécurité de la personne](#) vise à protéger les individus contre les blessures physiques et la privation de liberté. Pour Meta, le respect de ce droit humain implique l'atténuation du risque que le contenu provoque des dommages, notamment les risques de violence et de traite d'êtres humains, les menaces en ligne soutenues par des États, et les groupes non étatiques qui se livrent à la violence ou à la haine ou qui en font l'apologie.

Exemples de risques principaux inhérents potentiels en matière de droits humains identifiés lors de l'EGRP

Les acteurs malveillants qui :

- utilisent les services et applications Meta pour coordonner des préjudices en ligne et hors ligne ;
- abusent des services et applications pour mettre en place des cyberattaques ou de l'hameçonnage ;
- menacent ou harcèlent des défenseurs des droits humains, des militants ou d'autres groupes vulnérables.

Exemples de mesures prises par Meta pour remédier aux risques potentiels en 2025

En menant des recherches approfondies et en consultant des partenaires, Meta a identifié un risque potentiel accru de recrutement d'enfants soldats au Soudan et élaboré des messages de sécurité destinés aux jeunes de la région, disponibles en arabe, en français et en anglais. En novembre 2025, Meta a lancé deux campagnes complémentaires : une promotion rapide diffusée auprès d'environ un million de jeunes Soudanais (âgés de 13 à 18 ans) sur Facebook et Instagram, et un message personnalisé interactif s'affichant sous forme de fenêtre pop-up auprès d'environ 40 000 jeunes impliqués dans des [démantèlements stratégiques de réseaux](#). Ces messages mettaient en garde contre les risques de recrutement et orientaient les jeunes vers des ressources de sécurité externes.

→ Lire la suite

Nous avons continué de soutenir le Human Rights Defenders Fund ainsi que notre [programme de partenaires de confiance](#), afin d'améliorer la réponse d'urgence pour les défenseurs des droits humains et d'autres personnes vulnérables.

5. Intérêt supérieur de l'enfant

La Convention internationale des droits de l'enfant (CIDE) énonce que « l'intérêt supérieur de l'enfant doit être une considération primordiale » dans le cadre de toutes les actions concernant les enfants. Le [Cadre de l'intérêt supérieur de l'enfant](#) de Meta est conforme aux valeurs fondamentales de la CIDE. La protection de l'enfant en ligne est une priorité de Meta. Nous proposons des outils aux ados, aux parents et aux représentants légaux sous la forme de protections intégrées afin de garantir leur sécurité tout en leur permettant d'exercer leurs droits à la liberté d'expression et à l'accès à l'information.

Exemples de risques principaux inhérents potentiels en matière de droits humains identifiés lors de l'EGRP

Les enfants peuvent être exposés à des contenus inappropriés et non sollicités, voire à des comportements de prédateurs

Exemples de mesures prises par Meta pour remédier aux risques potentiels en 2025

En ce qui concerne les [comptes Ado](#), nous avons déployé de [nouvelles protections](#) et [mis à jour les paramètres de contenu](#) en nous inspirant des critères de classification des films pour les 13 ans et plus et des commentaires des parents. L'objectif est que les ados voient un contenu similaire à celui des films adaptés à leur âge par défaut sur Instagram. Nous avons également lancé un nouveau paramètre, plus strict, pour les parents qui souhaitent que l'expérience de leur ado soit plus restrictive. Nous avons annoncé de nouveaux contrôles visant à aider les parents à gérer la façon dont leurs [ados interagissent avec l'IA](#) sur nos applications, [engagé une action en justice](#) contre des entreprises à l'origine d'applications de « nudification » et [travaillé avec des écoles pour lutter contre le cyberharcèlement](#).

6. Participation publique, processus de vote et éligibilité

Le [droit à la participation publique, au processus de vote et à l'éligibilité](#) lors d'élections libres et équitables est l'un des fondements de la démocratie. Protéger l'intégrité des élections sur nos services et applications est l'une de nos priorités. Nous mettons tout en œuvre pour protéger les élections en ligne avant, pendant et après la période des élections.

Exemples de risques principaux inhérents potentiels en matière de droits humains identifiés lors de l'EGRP

Contenu en infraction pouvant avoir une incidence négative sur la participation publique, le vote ou la candidature à une fonction publique. Cela peut venir d'activités comprenant, mais sans s'y limiter :

- les acteurs malveillants coordonnés qui interfèrent avec les élections ;
- la menace de préjudice hors ligne et de violence envers les candidats ;
- les tentatives individuelles d'empêcher des personnes de voter, l'augmentation du spam, l'ingérence étrangère ou les signalements de contenus qui enfreignent nos politiques.

Exemples de mesures prises par Meta pour remédier aux risques potentiels en 2025

Nous avons continué de travailler sur notre approche des élections à grande échelle et à haut risque, que nous avons détaillée dans notre [Rapport sur les droits humains 2024](#). En 2025, nous avons préparé des élections en Australie, au Cameroun, au Chili, en Côte d'Ivoire, en Équateur, en Égypte, en Allemagne, en Irak, au Japon, au Myanmar, aux Philippines, en Pologne, en Roumanie et en Tanzanie.

7. Liberté de réunion et d'association

Le [droit à la liberté de réunion et d'association](#) est essentiel à la démocratie et interdépendant de nombreux autres droits garantis par le droit international relatif aux droits humains, notamment le droit à la liberté d'expression et le droit de participation aux affaires publiques. Pour Meta, ce droit est lié à deux de nos valeurs fondamentales : donner la parole à tout le monde et créer des liens et une communauté.

Exemples de risques principaux inhérents potentiels en matière de droits humains identifiés lors de l'EGRP

Certains types de contenu sur les plateformes Meta peuvent donner l'impression à certaines personnes qu'elles ne peuvent pas interagir librement sur les applications Meta ou hors ligne

Exemples de mesures prises par Meta pour remédier aux risques potentiels en 2025

Nous avons déployé notre [protocole de politique de crise](#) pour traiter les contenus en infraction liés à des manifestations de masse, notamment celles de la génération Z, par exemple en Iran, à Madagascar, au Népal et en Serbie.

8. Droit à la santé

Le [droit à la santé](#) désigne le droit de jouir du meilleur état de santé physique et mentale possible. Meta respecte ce droit en améliorant l'accès aux informations de santé, en permettant aux personnes ayant des problèmes de santé similaires d'entrer en contact les unes avec les autres et en donnant à chacun les moyens de prendre des décisions éclairées concernant sa santé et son bien-être.

Exemples de risques principaux inhérents potentiels en matière de droits humains identifiés lors de l'EGRP

Contenu en infraction qui incite ou vise à causer un préjudice hors ligne

Exemples de mesures prises par Meta pour remédier aux risques potentiels en 2025

Nous avons annoncé des [protections pour les ados dans nos offres d'IA](#), en les concevant pour répondre de façon sûre aux demandes concernant l'automutilation, le suicide et les troubles du comportement alimentaire. Nous avons également [annoncé](#) de nouveaux contrôles, notamment la mise à disposition pour les parents d'informations sur les sujets dont leurs ados discutent avec l'IA.

Aux États-Unis, nous avons intensifié notre lutte contre le harcèlement en ligne dans les écoles en mettant à la disposition du corps enseignant et du personnel de direction un [système de signalements prioritaires](#) pour les cas impliquant la sécurité des ados.

Principales problématiques

Innover en respectant les droits humains

Les innovations dans le domaine de la technologie wearable et de l'intelligence artificielle (IA) ont permis de faire progresser les droits humains tout en contribuant à atténuer les risques de manière responsable.

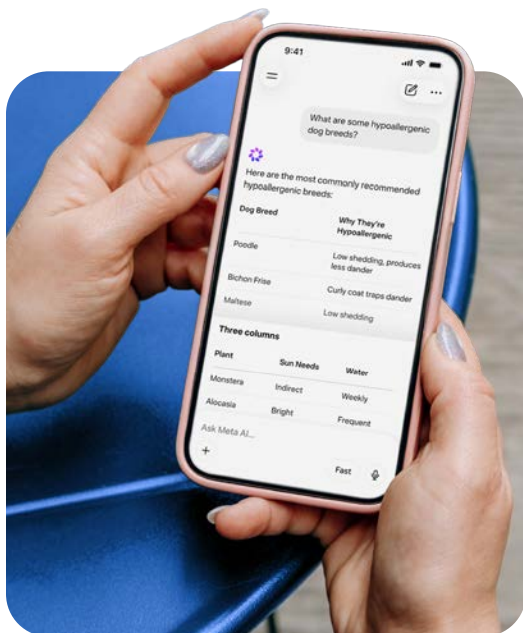
Favoriser l'accessibilité grâce à la technologie wearable

En 2025, nous [avons lancé les Meta Ray-Ban Display et le Meta Neural Band](#), une combinaison de lunettes IA nouvelle génération et d'un bracelet qui utilise l'[électromyographie \(EMG\) de surface](#) pour convertir les signaux électriques générés par de légers mouvements musculaires en commandes permettant un contrôle mains libres.

Les prototypes de ce bracelet ont été développés de manière responsable, en tenant compte de l'[accessibilité](#) et conformément au droit à l'égalité et à la non-discrimination. Basé sur des recherches menées auprès de 200 000 participants, ce bracelet offre de [nouvelles possibilités](#) aux personnes dont la mobilité est réduite en raison de lésions de la moelle épinière, d'accidents vasculaires cérébraux ou d'autres maladies invalidantes, ainsi qu'à celles qui souffrent de tremblements ou qui ont moins de cinq doigts. Dans un souci de confidentialité, les signaux EMG sont [traités directement sur l'appareil](#) par un système de machine learning qui les convertit en commandes numériques.



Associée à des lunettes IA, notre technologie [permet aux personnes en situation de handicap](#) de bénéficier d'une plus grande autonomie au quotidien. Les appels mains libres, la traduction en temps réel et les sous-titres en direct éliminent les obstacles à la communication, tandis que notre fonctionnalité « [Appeler un bénévole](#) », créée en partenariat avec [Be My Eyes](#), permet aux personnes aveugles ou malvoyantes d'entrer en contact avec des bénévoles pour obtenir une description en temps réel de leur environnement. Nous avons collaboré directement avec des organisations spécialisées dans le handicap, dont la [Blinded Veterans Association](#) et les [Blind Rehabilitation Centers](#) du département des Anciens combattants des États-Unis, afin de développer des guides de formation et de tenir compte de l'accessibilité dans la conception de nos outils.



Continuer à innover dans l'IA de manière responsable

Nous avons déployé l'[assistant Meta AI autonome](#) en même temps que notre grand modèle de langage open source mis à jour, [Llama 4](#). Nous avons également lancé le [traitement privé](#), une nouvelle [technologie qui repose sur un environnement d'exécution de confiance](#), pour que l'assistant Meta AI puisse être déployé sur WhatsApp tout en préservant la confidentialité, notamment en chiffrant les communications. Pour plus de détails, veuillez consulter notre analyse préalable sur les [avantages du chiffrement de bout en bout pour les droits humains](#).

Afin de limiter les risques d'utilisation détournée de nos produits et modèles d'IA et de résultats potentiellement nuisibles, le développement de Llama 4 a compris un filtrage des données pré-entraînement, des mesures d'atténuation post-entraînement et des mesures d'atténuation au niveau du système. Tout cela est expliqué dans notre [guide sur les mesures de protection liées à l'IA pour les développeurs](#).

Nous avons également publié notre [cadre sur l'IA de pointe](#), qui se concentre sur les risques les plus critiques en matière de cybersécurité ainsi que sur ceux liés aux armes chimiques et biologiques. Ce cadre met en avant les différents processus que nous suivons pour anticiper et atténuer les risques lors du développement de systèmes d'IA de pointe, notamment :

Identifier les conséquences catastrophiques à prévenir



Notre cadre a identifié des conséquences potentiellement catastrophiques liées aux risques cybernétiques, chimiques et biologiques, que nous nous efforçons de prévenir. La priorité était de déterminer si ces conséquences catastrophiques avaient été provoquées par les progrès technologiques et, le cas échéant, d'identifier des moyens d'atténuer ces risques.

Exercices de modélisation des menaces



Nous avons réalisé des exercices de modélisation des menaces afin d'anticiper la manière dont différents acteurs pourraient chercher à utiliser l'IA de pointe de manière détournée dans le but de provoquer ces conséquences catastrophiques, en collaborant avec des spécialistes externes lorsque cela était nécessaire. Ces exercices ont été fondamentaux pour notre approche axée sur les résultats.

Établissement de seuils de risque

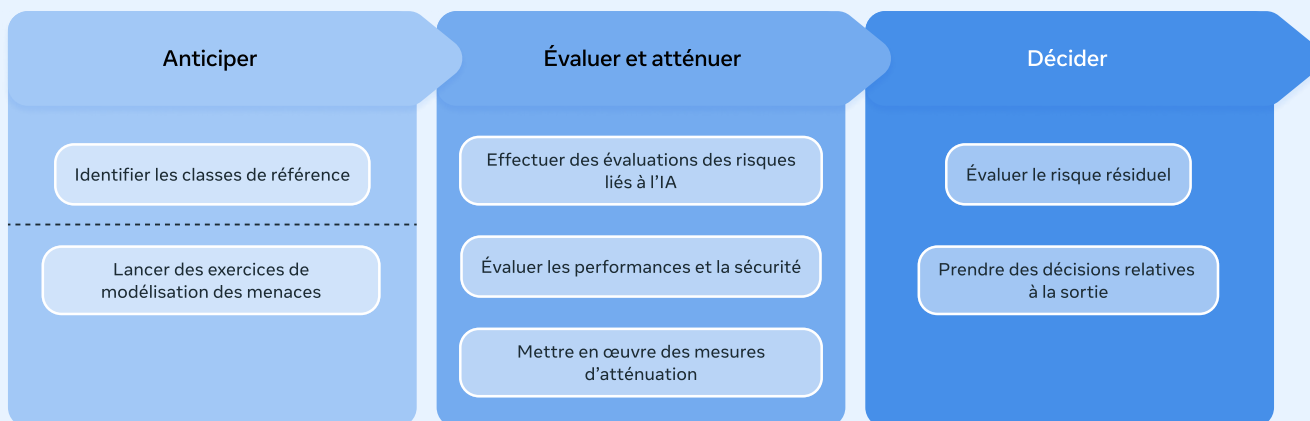


Nous avons défini des seuils de risque basés sur la mesure dans laquelle nos modèles pouvaient faciliter certaines menaces. Nous avons mis en place des processus pour maintenir les risques à des niveaux acceptables, notamment en appliquant des mesures d'atténuation.

Nous avons également continué de soutenir un vaste écosystème pour un développement et un déploiement responsables de l'IA. Afin de permettre aux développeurs de créer des applications sûres, nous avons renforcé nos investissements dans des [mécanismes de sécurité au niveau du système](#), notamment [Llama Guard 4](#) pour une protection unifiée des modalités texte et image, et [Llama Prompt Guard 2](#) pour une détection améliorée des tentatives de contournement et d'injection de prompts.

Nos travaux se sont appuyés sur divers échanges et interactions avec la société civile, les organisations internationales, le milieu universitaire et le grand public au sujet des aspects de l'IA impliquant des droits humains. En partenariat avec le Deliberative Democracy Lab de l'université Stanford, nous avons organisé des [forums communautaires](#) qui ont permis de recueillir des points de vue variés sur l'avenir des agents IA et leur gouvernance.

Cadre pour l'IA de pointe 2025





Outils de bien-être open source pour l'examen de contenu

Pour que nos examinateurs puissent appliquer nos politiques avec précision, ils doivent également se concentrer sur [leur résilience et leur bien-être](#).

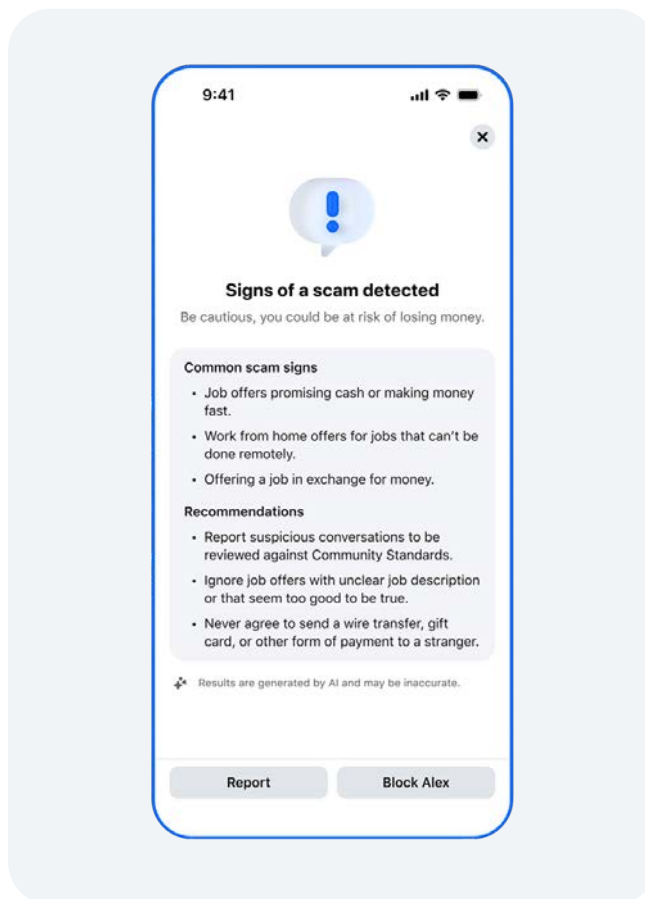
En 2025, nous avons [rendu open source une suite de filtres d'examen de contenu](#) conçus pour préserver le bien-être des modérateurs afin que d'autres développeurs puissent tirer parti de notre expérience dans ce domaine. Ces filtres image et vidéo personnalisables, qui comprennent des fonctionnalités telles que le floutage, la conversion en niveaux de gris ou en tons sépia, le lissage des détails et des écrans d'avertissement prédictifs alimentés par l'IA, peuvent contribuer à réduire les effets néfastes potentiels sur la santé mentale de l'exposition à des contenus explicites ou perturbants.

Fraude et arnaques

Au cours de la dernière décennie, les réseaux de fraude et d'arnaques se sont considérablement développés et ont eu une incidence négative sur les droits humains. De nombreux centres d'arnaques ont recours à la traite d'êtres humains et au travail forcé pour mener à bien leurs activités, et les victimes de ces arnaques peuvent subir des pertes financières considérables ainsi que d'autres préjudices. En 2025, nous avons [supprimé](#) plus de 159 millions de publicités frauduleuses pour infraction à nos politiques, et nous avons [désactivé](#) 10,9 millions de comptes Facebook et Instagram liés à des centres d'arnaques criminels.

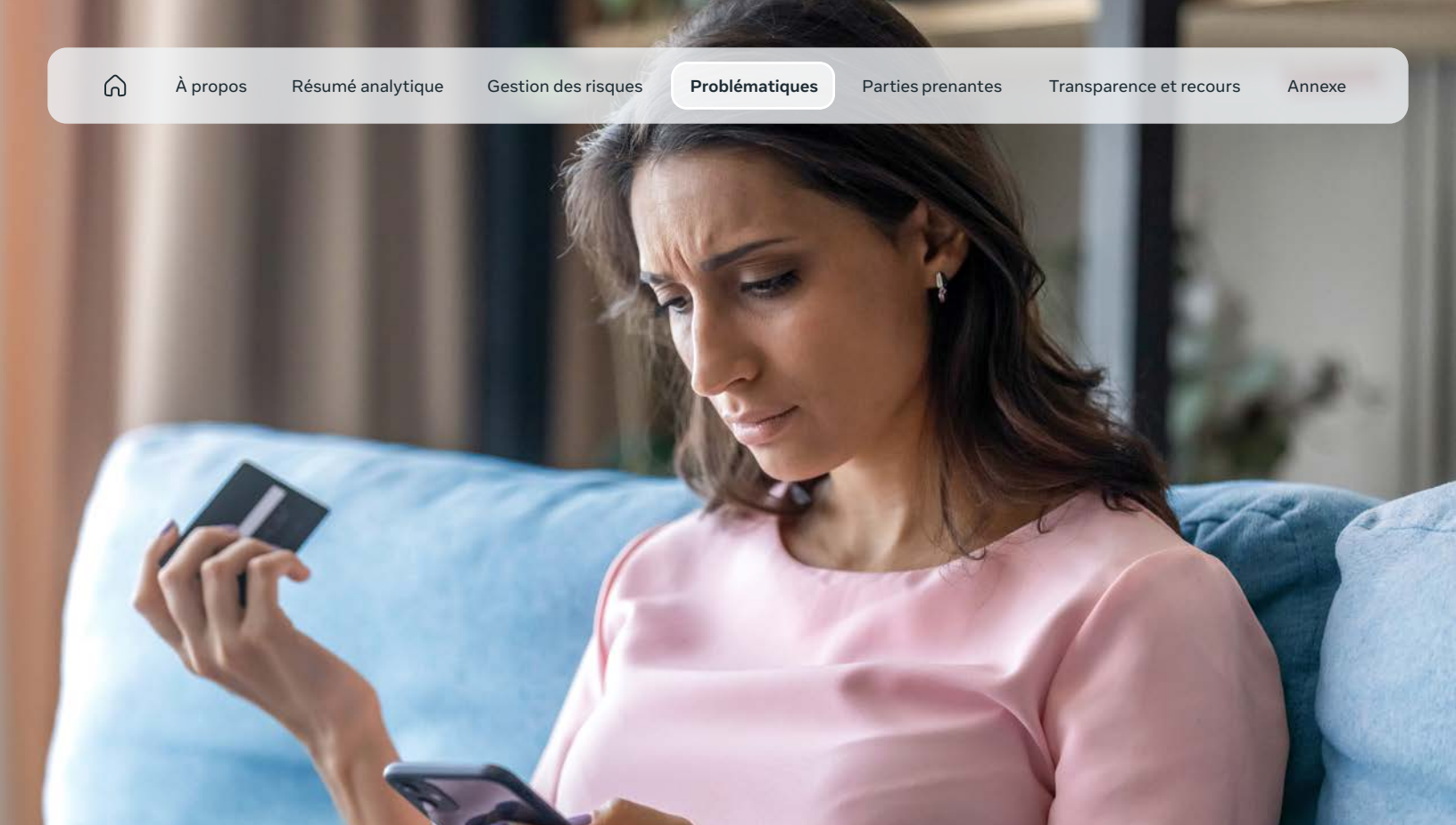
Nous avons donné la priorité à la sécurité en mettant en place des [mécanismes de protection intégrés à nos plateformes](#), tels que la [reconnaissance faciale](#) pour lutter contre l'usurpation d'identité, les [alertes de détection d'arnaques](#) sur Messenger et les [avertissements de partage d'écran](#) sur WhatsApp. Nos [efforts](#) de lutte contre les arnaques ont également compris des procédures judiciaires et un soutien aux forces de l'ordre afin de sanctionner leurs auteurs. Trouver un équilibre entre la sécurité et le bien-être de nos communautés et des mesures d'atténuation proportionnées et respectueuses des droits peut s'avérer difficile. Nous avons agi en tenant compte de la confidentialité et conformément aux normes internationales relatives aux droits humains, y compris les [principes de la Global Network Initiative](#).

La fraude et les arnaques sont des défis mondiaux. Le partage de signaux, la collaboration entre les communautés de défenseurs et une approche impliquant l'ensemble de la société sont nécessaires pour une action efficace. En 2025, Meta a établi plusieurs partenariats axés sur le partage de signaux, notamment avec le [Global Signal Exchange](#), la [Global Anti-Scam Alliance](#) et le [Fraud Intelligence Reciprocal Exchange](#). Nous avons également [plaidé en faveur](#) de la publication de listes de pays ne prenant pas les mesures adéquates ainsi que de l'instauration de sanctions internationales, en tenant toujours compte des droits humains et des conséquences non voulues sur les populations vulnérables, notamment les personnes victimes de traite d'êtres humains qui ont été forcées de participer à des arnaques.



Démanteler les réseaux criminels spécialisés dans les arnaques en Asie du Sud-Est

Un grand nombre de réseaux et de complexes d'arnaques sophistiqués opérant à l'échelle industrielle ont [vu le jour](#) au cours de la dernière décennie. En [Asie-Pacifique](#), ces complexes emploient des centaines de milliers de personnes et opèrent en toute impunité, ou presque. Les centres d'arnaques criminels attirent des demandeurs d'emploi au moyen d'offres trompeuses et les obligent à se livrer à des arnaques en ligne, souvent sous la menace de violences physiques, ce qui soulève la question de l'interdiction de l'esclavage, de la servitude et du travail forcé en vertu des normes internationales relatives aux droits humains.



Nous avons [supprimé](#) plus de 6 400 comptes et Pages Facebook entre janvier et octobre 2025 pour [infraction](#) à nos politiques en matière [de fraude, d'arnaques et de pratiques trompeuses](#), ainsi que [d'organisations et de personnes dangereuses](#). Nous avons démantelé un réseau basé au Cambodge [se faisant passer](#) pour des forces de l'ordre internationales afin de proposer des services frauduleux de récupération de fonds. Pour y parvenir, nous avons eu recours à une collaboration internationale et avons suivi de près l'impact de changements géopolitiques tels que le conflit au Myanmar et la situation le long de la frontière dans le nord du Cambodge. Nous discutons de ces interventions plus en détail dans notre [Rapport des menaces conflictuelles](#) du premier semestre 2026.



En 2025, notre travail dans ce domaine a abouti à notre participation à un projet pilote appelé [Joint Disruption Week](#) à Bangkok, une opération de démantèlement conjointe dans le cadre de laquelle Meta a collaboré avec le FBI, la Police royale thaïlandaise et d'autres autorités. Ce projet pilote [a abouti](#) à la suppression de 59 000 comptes, Pages et groupes des plateformes de Meta, ainsi qu'à l'émission de six mandats d'arrestation.



En plus de cela, nous avons mis à jour nos [Standards de la communauté en matière d'exploitation d'êtres humains](#) afin de détecter davantage de contenus de recrutement émanant de réseaux criminels spécialisés dans les arnaques. Cette mise à jour restreint la publication d'offres d'emploi dans les régions présentant un risque élevé d'exploitation selon les forces de l'ordre et les organisations non gouvernementales locales.

Meta a également collaboré avec l'Office des Nations Unies contre la drogue et le crime dans le cadre de la campagne [#TrappedInScamCrime](#), qui visait à lutter contre le recrutement en ligne et la traite d'êtres humains à des fins d'activités criminelles forcées liées à des centres d'arnaques en Asie du Sud-Est. La campagne a mis en avant des messages de prévention fondés sur l'expérience de victimes afin d'aider le public à repérer les offres d'emploi trompeuses et à accéder à des ressources de soutien.

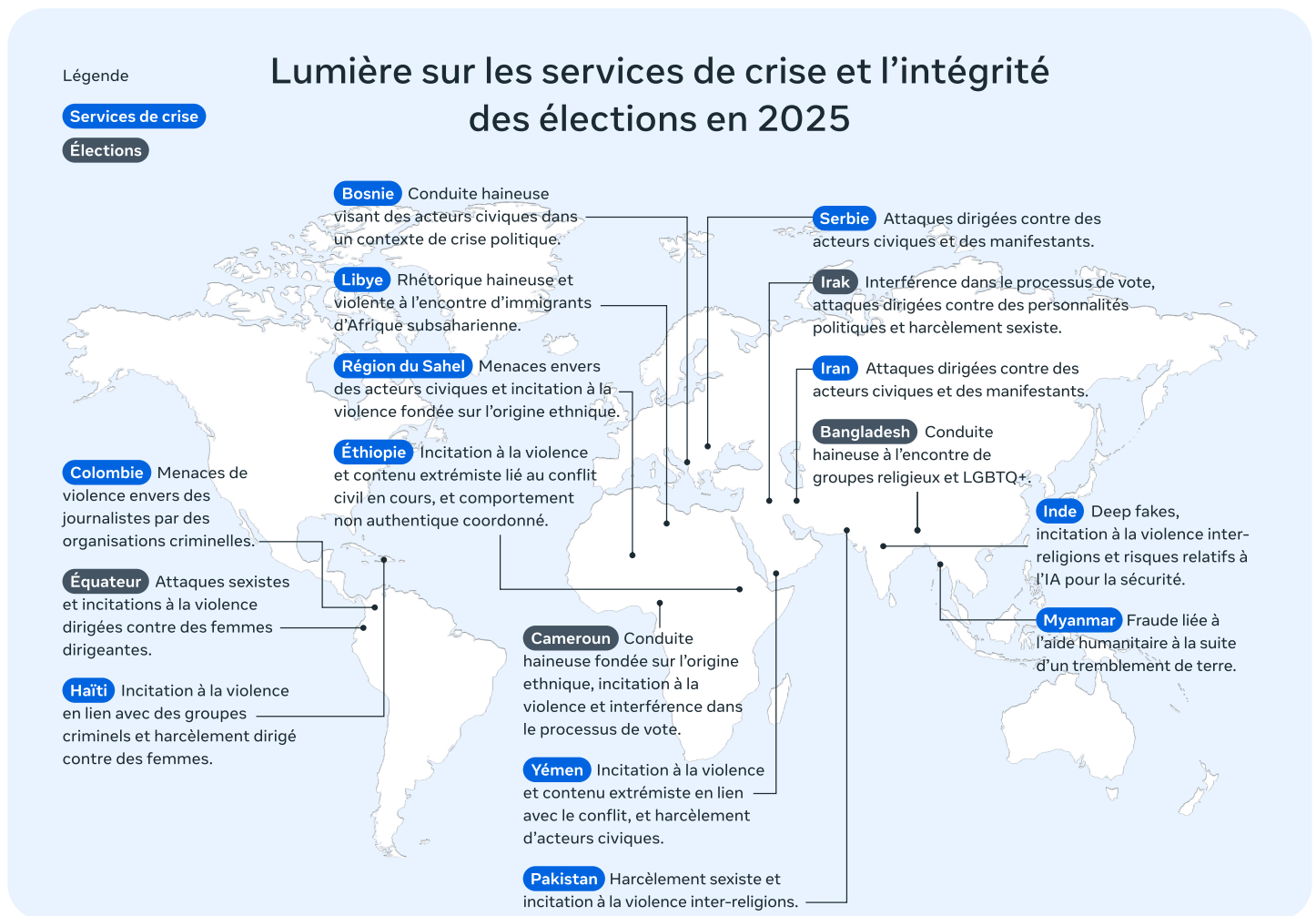
Nous avons également collaboré avec l'Association des nations de l'Asie du Sud-Est pour élaborer des [principes](#) visant à lutter contre les arnaques en ligne. Ces mesures sont venues s'ajouter à nos efforts de [lutte contre la traite d'êtres humains](#) au sens large.

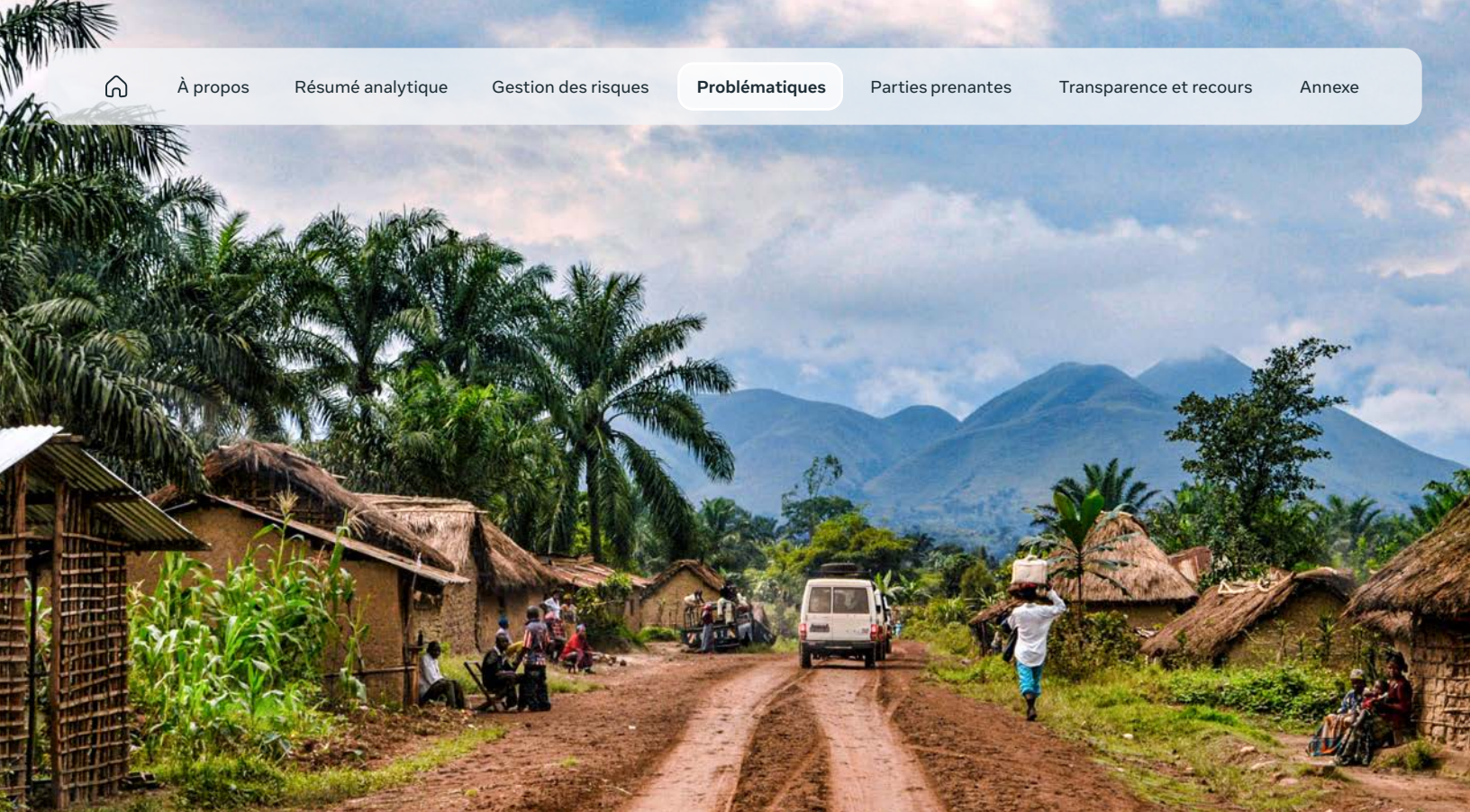
Comment se préparer aux crises et y répondre

Nous nous préparons à intervenir dans de nombreuses situations de crise à travers le monde, notamment des conflits, des violences intercommunautaires, des troubles civils, des manifestations de masse et des catastrophes environnementales, mais aussi des attentats terroristes et des fusillades. En 2025, nous avons lancé et coordonné des réponses aux crises au Cameroun, en Inde (Cachemire), en Indonésie, en Libye, à Madagascar, au Népal, en Serbie, en Tanzanie et à la frontière entre la Thaïlande et le Cambodge, entre autres pays et territoires. Nous avons continué de répondre aux crises soumises au [protocole de politique de crise](#) pour les conflits en Ukraine, au Soudan et au Moyen-Orient. Nous abordons ces crises dans nos [précédents Rapports sur les droits humains](#).

En période de crise, le protocole de politique de crise est un outil fondamental. Il nous permet de mettre rapidement en place les mesures nécessaires pour limiter les risques de préjudice, tel que nous l'expliquons dans notre [Rapport sur les droits humains 2024](#). Les mesures déployées dépendent des risques observés. Elles s'appuient sur les interventions menées lors de crises passées, les principes des droits humains et le droit des conflits armés.

À la page suivante, nous illustrons la manière dont nous nous préparons aux crises et y répondons dans les zones touchées par des conflits et à haut risque, en nous appuyant sur l'exemple d'une région ne figurant pas dans nos précédents rapports.





République démocratique du Congo

En 2025, les combats dans l'est de la République démocratique du Congo (RDC) se sont intensifiés lorsque le groupe armé M23, qui aurait été [soutenu](#) par le Rwanda, s'est emparé de territoires du Nord-Kivu, provoquant des déplacements massifs de populations civiles et de [graves violations des droits humains](#). Des partenaires de confiance nous ont avertis de l'escalade de la violence dans l'est du pays ainsi que du risque croissant d'affrontements intercommunautaires, ce qui a conduit Meta à soumettre cette crise à notre protocole de politique de crise.

Une équipe pluridisciplinaire de spécialistes de Meta couvrant les domaines des politiques, des opérations, de la sécurité et des affaires publiques surveillait la situation sur le terrain en RDC depuis la reprise du conflit fin 2021, en prenant des mesures à l'encontre des contenus qui enfreignaient nos politiques en matière [de fausses informations, de violence et d'incitation, d'attaques coordonnées et de promotion d'actions criminelles](#) et [d'organisations et de personnes dangereuses](#). Nous avons déployé des mesures d'atténuation pour gérer les risques potentiels liés au contenu en ligne, notamment l'identification précoce d'affirmations susceptibles de causer un préjudice en vertu de notre politique en matière de fausses informations, et l'utilisation de notre [politique relative aux lieux à haut risque temporaire](#). Nous avons également identifié des groupes à risque, tels que les Rwandaphones, les Rwandais, les Banyamulenge et les Tutsis, afin de prévenir les risques de préjudice hors ligne.

En 2025, à mesure que le conflit s'intensifiait, nous avons évalué et atténué les risques liés aux contenus connexes. Nous avons clarifié et expliqué nos politiques, passé en revue les mots-clés afin de garantir une détection efficace du contenu en infraction et une mise en application précise de nos politiques, et appliqué notre politique relative aux lieux à haut risque temporaire.

Meta a également consulté des partenaires de confiance pour éclairer la désignation potentielle de nouvelles déclarations considérées comme nuisibles, à savoir les déclarations fausses ou trompeuses susceptibles de contribuer directement à un risque de blessures physiques imminentes, en vertu de notre politique en matière de fausses informations. Les contributions de nos partenaires de confiance nous ont permis de détecter et de supprimer plus efficacement les contenus en infraction, notamment ceux susceptibles d'exacerber les divisions locales.

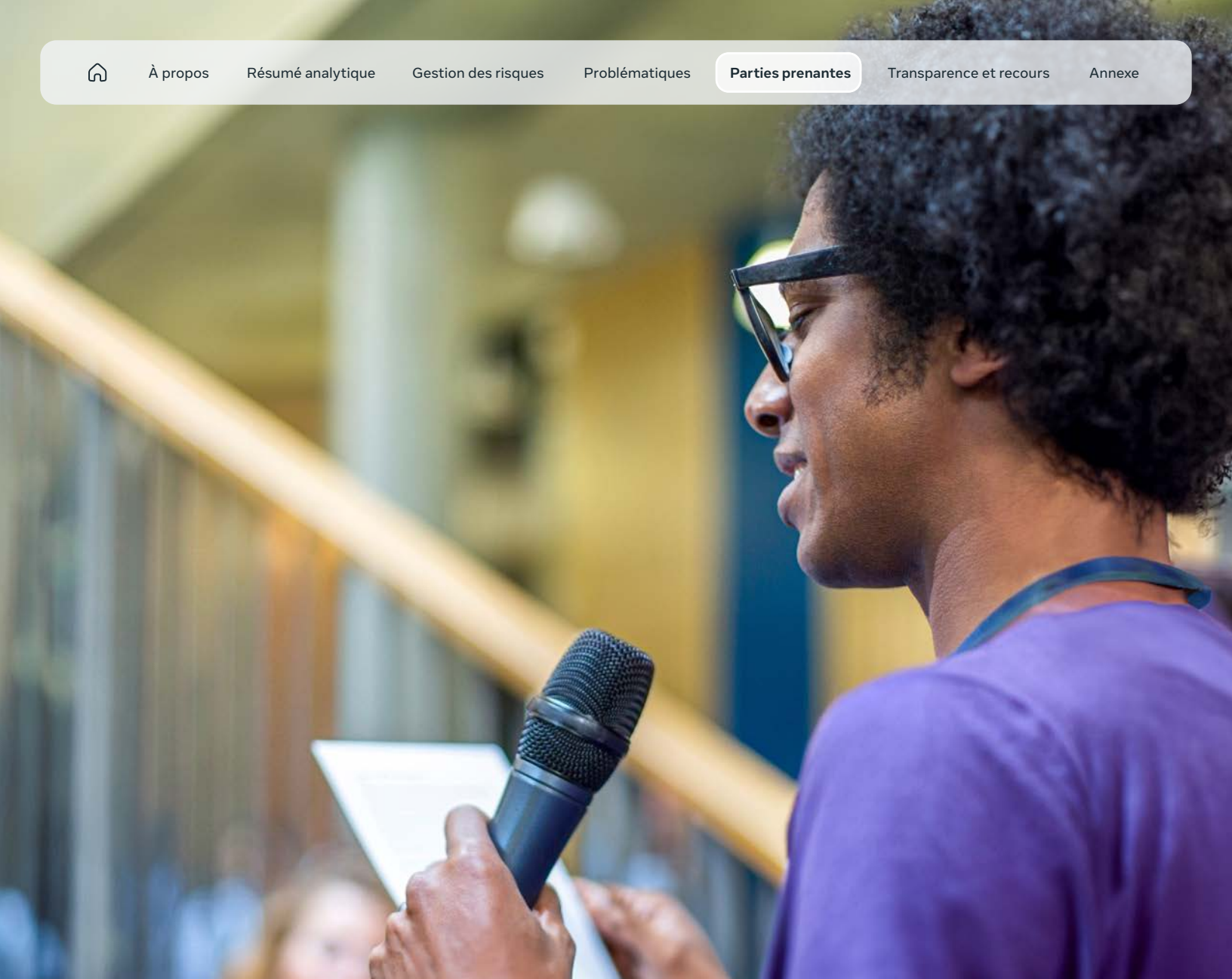
Implication des parties prenantes

Les [interactions](#) structurées et proactives avec notre communauté mondiale d'utilisateurs contribuent à façonner les politiques de Meta et sont essentielles à notre gestion des risques en matière de droits humains.

Nous avons interagi avec un vaste éventail de parties prenantes en 2025, notamment des membres de la société civile, des universitaires, des groupes de réflexion, des spécialistes des droits humains et des organismes de réglementation. Parmi les principales problématiques figuraient notre approche de l'intelligence artificielle (IA) et de la modération du contenu, l'impact des changements apportés à nos politiques et produits en janvier, notamment les [Notes de la communauté](#), et la fraude et les arnaques.

Nous avons interagi avec plus de 615 parties prenantes externes de 65 pays pour éclairer l'élaboration de politiques dans plusieurs domaines, notamment les organisations et personnes dangereuses, ainsi que les règles publicitaires liées à la santé et au bien-être. Nous en avons consulté des centaines d'autres dans le cadre de champs de travail relatifs à la prévision des risques, à l'analyse des tendances, à la mise en application de nos politiques et à la prévention.

De plus, nous avons interagi avec la société civile au sujet de différentes problématiques relatives à la liberté d'expression et à la réglementation, notamment en Argentine et au Brésil, mais aussi en Malaisie, dans le cadre de la [Digital Rights Asia Pacific Assembly](#). Nous avons également formé la société civile sur nos politiques. En Afrique du Nord, par exemple, nous avons formé des groupes de la société civile sur nos politiques en matière de [fausses informations](#) et de [conduites haineuses](#), afin de les aider à identifier les contenus susceptibles de les enfreindre. La formation a également donné aux défenseurs des droits humains et aux groupes vulnérables des outils pour protéger leur présence en ligne.



Partenaires de confiance

Nos partenaires de confiance sont des alliés importants dans l'identification des infractions graves à nos Standards de la communauté. Nous avons continué de faire appel à nos [partenaires de confiance](#) pour identifier des tendances, mieux comprendre l'impact du contenu et des comportements en ligne sur les communautés locales et renforcer nos canaux de signalement dédiés à la société civile.

En 2025, nos partenaires de confiance ont apporté leur contribution et identifié des contenus nuisibles dans des pays et territoires ayant connu une montée des tensions ou tenu des élections. Cela comprenait l'Afghanistan, le Cambodge, la Côte d'Ivoire, la République démocratique du Congo, l'Inde, l'Indonésie, Israël, le Kenya, le Liban, Madagascar, le Mexique, le Nigeria, la Palestine, la Somalie, le Soudan, la Syrie, la Thaïlande, la Tanzanie, l'Ukraine et le Venezuela.

Nous avons supprimé plus de 60 000 contenus en infraction dans le cadre de notre programme des partenaires de confiance. Vous trouverez un exemple de l'impact de nos partenaires de confiance en Amérique latine à la page suivante.

ÉTUDE DE CAS

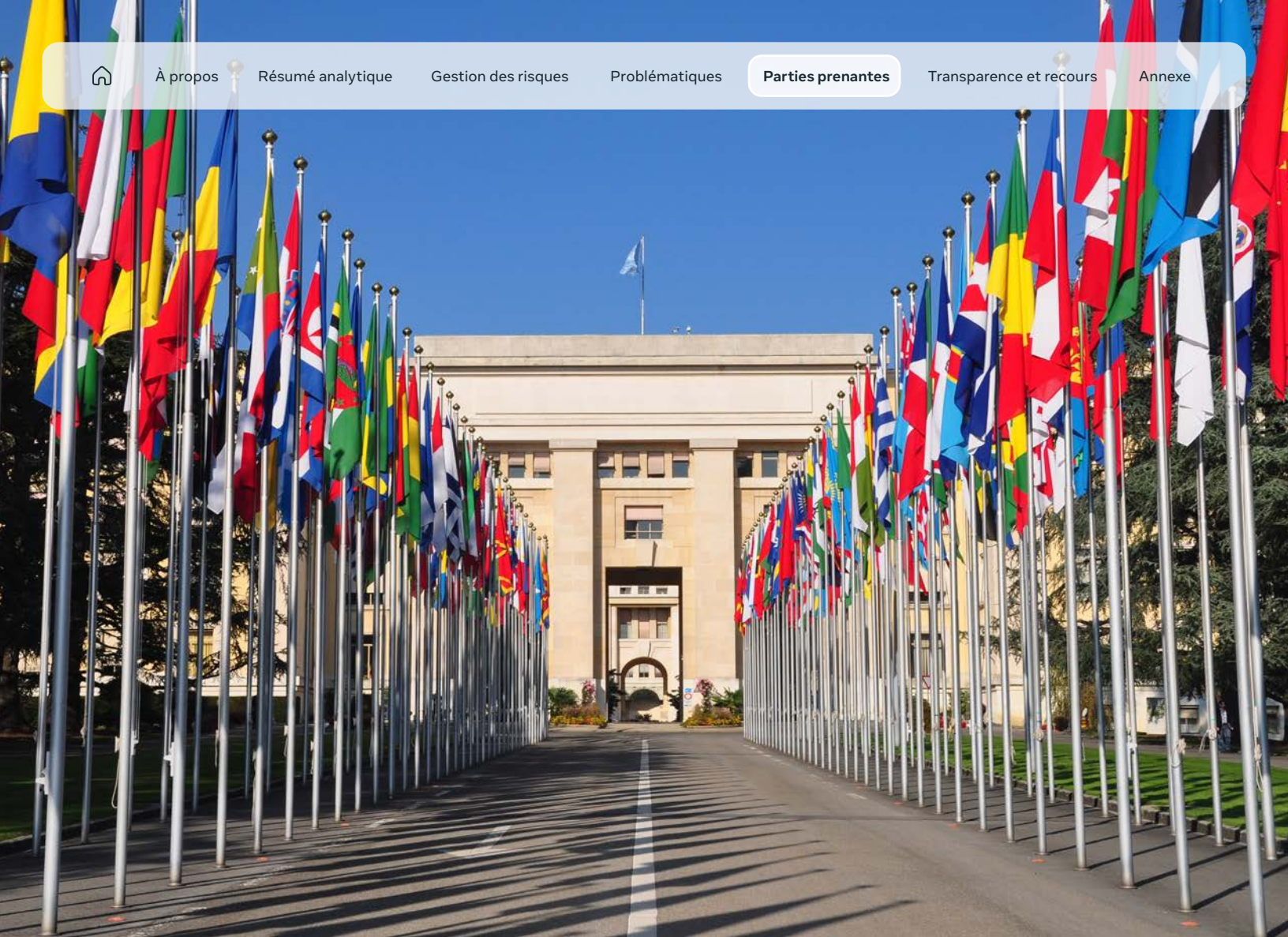
Comment prévenir le recrutement de jeunes en Amérique latine



En Amérique latine, nos partenaires de confiance ont joué un rôle crucial dans le signalement des contenus glorifiant, soutenant ou représentant [des organisations ou des personnes dangereuses](#). Ils ont également fourni des observations sur les initiatives de prévention du recrutement des jeunes.

Au Brésil, nos partenaires de confiance ont signalé un contenu Instagram présentant des membres d'une organisation criminelle désignée, notamment des représentations d'armes à feu interdites, d'actes de violence et de drogues. En Colombie, nos partenaires de confiance ont signalé des profils associés à des groupes criminels à l'origine de menaces et d'attaques violentes à l'encontre de journalistes dans le département du Guaviare. En Haïti, le contenu signalé glorifiait des actes criminels commis par des individus désignés associés à des gangs. Ces contributions nous ont aidés à supprimer les contenus en infraction, à limiter la présence d'organisations et de personnes dangereuses sur nos plateformes et à réduire le risque d'attaques physiques.

Meta a collaboré avec des partenaires de confiance et des experts régionaux pour lancer des campagnes de messages ciblant les jeunes en Amérique latine. Diffusées auprès de 80 000 utilisateurs âgés de 13 à 25 ans au Mexique et de 55 000 autres en Colombie, en Haïti et au Brésil, ces campagnes visaient à informer les utilisateurs des infractions sur nos plateformes, à les sensibiliser aux risques liés aux interactions avec des organisations et des personnes dangereuses, et à les aider à faire des choix plus judicieux en ligne. La campagne a entraîné une augmentation mesurable de la désactivation de comptes et de la suppression d'amis.



Organisations internationales

Meta a continué de collaborer avec les Nations Unies tout au long de l'année. Nous avons pris part aux discussions sur les politiques qui ont eu lieu lors de la 80e session de l'Assemblée générale des Nations Unies (AGNU). Les sujets abordés comprenaient la gouvernance mondiale du numérique et de l'IA, l'utilisation de l'IA pour la [sécurité nationale](#) et la réduction de la fracture de l'IA. Nous avons également participé au lancement du Dialogue mondial sur la gouvernance de l'IA, en marge de la Semaine de haut niveau de l'Assemblée générale des Nations Unies.

Meta a continué de travailler en étroite collaboration avec le [Haut-Commissariat des Nations Unies aux droits de l'homme \(HCDH\)](#), notamment sur le [projet B-Tech](#). Nous avons répondu aux appels à soumissions et consulté la [Commission africaine des droits de l'homme et des peuples](#) et les procédures spéciales des Nations Unies en matière de droits humains (des spécialistes indépendants en droits humains), notamment le Rapporteur spécial sur la liberté d'opinion et d'expression. Nous avons également collaboré avec d'autres organisations humanitaires et agences des Nations Unies pour améliorer notre réponse aux crises.

De plus, nous avons participé au [Forum sur la gouvernance de l'Internet](#) 2025, au [Forum sur les entreprises et les droits de l'homme des Nations Unies](#) et au [RightsX Summit](#), où nous avons évoqué des cas d'utilisation positifs de l'IA pour les défenseurs des droits humains, entre autres. Nous avons continué de participer et d'interagir avec les acteurs de plusieurs autres forums multipartites, tels que le [Global Internet Forum to Counter Terrorism](#), [RightsCon](#) et [Tech Against Trafficking](#).

Transparence et recours



Le [Conseil de surveillance](#), en tant qu'organisme indépendant, nous aide à résoudre certaines des questions les plus complexes qui soient en matière de liberté d'expression en ligne : quels contenus retirer, quels contenus conserver et pourquoi. Il examine des cas soumis par Meta ou ayant fait l'objet d'un appel sur Facebook, Instagram ou Threads qui sont en désaccord avec notre modération du contenu, et il émet des décisions contraignantes indiquant si le contenu doit être retiré ou conservé. Le Conseil de surveillance fournit également des recommandations pour améliorer nos pratiques de modération de contenu et offre des avis consultatifs en matière de politique sur demande.



En 2025, le Conseil de surveillance a évalué des cas concernant la mise en application de nos politiques au regard du cadre international des droits humains, notamment la liberté d'expression, le droit à la santé et le droit à l'égalité et à la non-discrimination, entre autres.

Exemples de décisions du Conseil de surveillance :



Le Conseil de surveillance a [annulé](#) la décision de Meta de ne pas signaler un extrait audio vraisemblablement manipulé dans lequel deux politiciens kurdes irakiens parlaient de la falsification des élections législatives. Le Conseil a estimé que, pour se conformer à ses responsabilités en matière de droits humains, Meta aurait dû appliquer la mention « Haut risque » à l'extrait audio en question ainsi qu'à tous les contenus dans lequel il apparaissait.



La majorité du Conseil de surveillance a [validé](#) la décision de Meta de laisser en ligne deux vidéos, l'une montrant une femme transgenre prise à partie parce qu'elle utilisait les toilettes des dames et l'autre, une athlète transgenre ayant remporté une course sur piste.

Exemples d'actions entreprises à la suite de recommandations du Conseil de surveillance :



À la suite des [recommandations](#) sur le cas d'une [femme iranienne prise à partie dans la rue](#) et afin d'éviter une modération excessive, nous avons décidé d'adopter une approche plus réactive de la mise en application de nos politiques. Cela a permis de réduire le risque que des plaisanteries ou des propos figurés potentiellement inoffensifs ne soient sanctionnés.



En réponse aux [recommandations](#) de fournir plus de détails sur ses décisions relatives au contenu, Meta a [augmenté](#) la granularité des notifications existantes afin de fournir aux utilisateurs des informations plus détaillées lorsque leur contenu est restreint à la suite d'un signalement officiel des autorités indiquant qu'il enfreint les lois locales.

Vous trouverez plus de détails dans les [Rapports semestriels de Meta sur le Conseil de surveillance](#) et sur notre [page sur l'impact du Conseil de surveillance](#).

Augmenter la transparence en notifiant les utilisateurs

9 %

T1 2024-T1 2025

Augmentation des taux de notification pour les cas éligibles de mise en application des Standards de la communauté sur Facebook

16 %

T1 2024-T1 2025

Augmentation des taux de notification pour les cas éligibles de mise en application des Standards de la communauté sur Instagram

~100 %

au T1 2025

Taux de notification pour les contenus restreints pour infraction aux lois locales sur Instagram et Facebook



Nous avons également soutenu les [travaux de réflexion](#) du Conseil de surveillance en matière de liberté d'expression et d'évaluation des risques systémiques, afin que les organismes de réglementation puissent tirer parti de l'expérience du Conseil. En 2025, le Conseil a pris part aux discussions sur les politiques qui ont eu lieu lors de l'Assemblée générale des Nations Unies et collaboré avec la Commission africaine des droits de l'homme et des peuples, entre autres.

***Avertissement :** il existe de nombreuses différences entre les réseaux sociaux et les films. Nous n'avons pas collaboré avec la MPA lors de la mise à jour de nos paramètres de contenu. Par ailleurs, elle n'évalue aucun contenu sur Instagram et ne soutient ni n'approuve en rien nos paramètres de contenu. Nous nous sommes plutôt inspirés des règles publiques de la MPA, que les parents connaissent déjà, ainsi que des commentaires des parents. Nos systèmes de modération du contenu sont différents de ceux des organismes de classification des films. L'expérience peut donc ne pas être exactement la même.



Annexe



Comment les droits humains sont régis et gérés chez Meta

Nos spécialistes en droits humains guident la mise en œuvre de notre [Politique relative aux droits humains au sein de l'entreprise](#), qui est supervisée par notre responsable des affaires internationales et notre directeur juridique.

Les tâches des spécialistes en droits humains consistent notamment à promouvoir l'intégration de la politique dans les politiques, programmes et services existants et en cours d'élaboration, à faire preuve de diligence raisonnable et à soutenir la formation du personnel sur la politique. La politique donne des orientations pour créer des produits respectueux des droits, répondre aux crises émergentes et travailler avec rapidité et agilité pour intégrer les droits humains à grande échelle.

Formation du personnel de Meta aux droits humains

Chez Meta, la manière dont nous agissons est aussi importante que ce que nous faisons. Notre formation aux droits humains met en évidence les incidences potentielles et réelles de nos services, politiques et décisions commerciales sur les droits humains. Elle vise à promouvoir un état d'esprit axé sur les droits humains dans notre travail quotidien, en encourageant le respect des droits humains dans l'intérêt de toutes les personnes qui utilisent nos services.

Nous avons lancé notre formation *Bigger than Meta: Human Rights* (Au-delà de Meta : les droits humains) en 2022, et elle s'est poursuivie tout au long de l'année 2025. Notre formation à la confidentialité contribue également aux objectifs de notre formation aux droits humains en se concentrant sur le développement notre capacité collective à protéger les individus, en particulier les groupes marginalisés, contre les préjudices liés au traitement de leurs données.

Liens vers les rapports référencés

[Rapports sur les pratiques professionnelles responsables](#)

[Rapports sur la durabilité](#)

[Rapport sur les droits humains 2024](#), [Rapport sur les droits humains 2023](#), [Rapport sur les droits humains 2022](#), [Rapport sur les droits humains 2021](#)

[Déclaration contre l'esclavage et la traite d'êtres humains 2026](#)

[Rapport sur les minerais de conflit 2025](#)

[Rapports de transparence Meta](#)

[Rapports réglementaires et autres rapports de transparence](#)

[Rapports des menaces conflictuelles](#)

[Rapports semestriels de Meta sur le Conseil de surveillance](#)

