

# Rapport sur les droits humains

Statistiques et mesures  
2022



# Table des matières

<b>1. Résumé analytique</b>	3	<b>6. Mises à jour sur la diligence raisonnable</b>	34
<b>2. À propos de ce rapport de progression</b>	6	Philippines	34
<b>3. Avant-propos</b>	7	Israël et Palestine	35
<b>4. Gestions et administration du travail sur les droits humains par Meta</b>	10	Inde	36
Formation des employés de Meta	11	<b>7. Engagement des parties prenantes : renforcement de notre travail par des parties externes</b>	38
Appels et recours	11	Contribuer à la création d'espaces en ligne sûrs et soutenir l'intégrité des élections au Nigeria	39
Conseil de surveillance	11	Engager des communautés marginalisées et sous-représentées dans la politique de contenu	40
Global Network Initiative	12	Perspectives de communautés indigènes en matière de drogues à usage non médical	41
<b>5. Identification et gestion des principaux risques en matière de droits humains</b>	13	Engagement auprès d'organisations internationales	41
Nos principaux risques	14	Déploiement de l'identification et escalation des thèmes avec les partenaires de confiance	42
Liberté d'opinion et d'expression	15	Innovation par le biais de forums communautaires	43
Confidentialité	16	<b>8. Langue</b>	44
Égalité et non-discrimination	19	<b>9. Problématiques clés</b>	46
Discours haineux	20	Protocole de politique de crise	46
Accessibilité	20	Russie et Ukraine	47
Vie, liberté et sécurité de la personne	22	Iran	48
Nos nombreuses façons de respecter le droit à la vie, à la liberté et à la sécurité de la personne	23	<b>10. Notre approche des droits humains et le développement responsable de nos produits</b>	49
Protection des DDH et des journalistes	23	Métavers	49
Traite d'êtres humains et exploitation	23	Intelligence artificielle	50
Surveillance à la demande	24	<b>11. Nous tourner vers l'avenir : considérations futures</b>	53
Images intimes non consenties (IINC)	24		
Intérêt supérieur de l'enfant	25		
Participation publique, processus de vote et éligibilité	26		
Plus d'informations sur notre travail en matière d'élections nationales	28		
Élections de mi-mandat en 2022 aux États-Unis	28		
Brésil	28		
Kenya	29		
Liberté de réunion et d'association	31		
Droit à la santé	33		



Voici le second rapport annuel sur les droits humains de Meta, relatif à l'année civile 2022. Ce rapport montre dans quelle mesure nous respectons les engagements définis dans notre [Politique d'entreprise en matière de droits humains](#) qui se fonde sur les [Principes directeurs des Nations unies relatifs aux entreprises et aux droits de l'homme](#) (United Nations Guiding Principles on Business and Human Rights, UNGP). Il présente la manière dont nous adaptons notre travail pour respecter les droits humains, la manière dont nous évaluons les incidences potentielles de nos services et ce que nous faisons pour atténuer ces risques.

Au fil des ans, les services et les applications de Meta ont aidé les mouvements populaires à se développer rapidement, à remettre en question les doctrines, à défendre les droits et, ce faisant, à changer le monde. 2022 n'a pas fait exception. Alors que de nouveaux défis voient le jour, notre travail en matière de droits humains évolue et aborde désormais des thèmes tels que l'utilisation de nos services dans des pays en conflit, les menaces qui pèsent sur un Internet ouvert et l'essor de l'intelligence artificielle.

Nous avons renforcé nos systèmes de gouvernance afin de faire avancer notre travail en faveur du respect des droits humains sur l'ensemble de nos produits et services. Cela implique de continuer à valoriser le [Conseil de surveillance](#), qui a [publié](#) 12 décisions de modération de contenu et formulé 91 recommandations en 2022. Nos efforts pour renforcer la gouvernance passent également par la formation de nos employés. Au cours de l'année, nous avons lancé dans toute l'entreprise la formation sur les droits humains intitulée *Au-delà de Meta : les droits humains*.

En 2022, Meta a réussi son [évaluation](#) Global Network Initiative, qui comprend un examen de nos politiques et processus, ainsi que huit cas d'étude spécifiques exemplatifs impliquant des demandes de retrait et des demandes de données du gouvernement.

La diligence raisonnable en matière de droits humains est essentielle à notre Politique d'entreprise en matière de droits humains. Le présent rapport fait état de découvertes essentielles issues de notre [Évaluation globale des risques principaux en matière de droits humains](#) (EGRP) et fournit une mise à jour quant à notre application des recommandations issues du travail de diligence raisonnable.

En 2022, nous avons réalisé l'EGRP en partenariat avec [Business for Social Responsibility](#) (BSR), une organisation indépendante experte en droits humains. L'analyse a décelé huit risques principaux hautement prioritaires sur la base des critères des UNGP. Il s'agit, 1.) de la liberté d'opinion et d'expression ; 2.) de la confidentialité ; 3.) de l'égalité et de la non-discrimination ; 4.) de la vie, de la liberté et de la sécurité de la personne ; 5.) de l'intérêt supérieur de l'enfant ; 6.) de la participation publique, du processus de vote et de l'éligibilité ; 7.) de la liberté de réunion et d'association ; et 8.) du droit à la santé. L'évaluation a révélé que les risques liés à la manière dont Meta fournit ses services et à ses modèles de profit sont variés et nuancés. Nous décrivons également la manière dont nous avons donné suite aux recommandations formulées lors des contrôles préalables en matière de droits humains concernant les [Philippines](#), [Israël et la Palestine](#), ainsi que l'Inde.

## Alors que de nouveaux défis voient le jour, notre travail en matière de droits humains évolue et aborde désormais des thèmes tels que l'utilisation de nos services dans des pays en conflit, les menaces qui pèsent sur un Internet ouvert et l'essor de l'intelligence artificielle.

Nous continuons par ailleurs à donner suite aux recommandations formulées dans l'[audit de Meta sur les droits civils](#) (2021). En 2022, 32 recommandations et actions supplémentaires sur 117 ont été mises en œuvre ou étaient en cours de réalisation, ce qui nous confère un pourcentage de conformité de 84 %.

Nous écoutons ce que les parties prenantes externes ont à nous dire et nous apprenons afin d'endosser nos responsabilités en matière de droits humains. L'engagement des parties prenantes est ancré à tous les niveaux de notre Politique d'entreprise en matière de droits humains et nous nous efforçons de nous engager davantage auprès des communautés marginalisées. Le développement de nos services, nos politiques de contenu et de modération et les [Standards de la communauté](#) sont fondés sur des informations issues d'organisations de la société civile, de [défenseurs des droits humains](#), de groupes marginalisés, d'organisations internationales, de [partenaires de confiance](#), d'investisseurs, d'annonceurs et d'utilisateurs.

Lorsque nous nous trouvons face aux risques principaux en matière de droits humains, nous donnons la priorité aux partenariats. Notre réseau de partenaires de confiance, composé de plus de 400 organisations non gouvernementales, à but non lucratif, nationales et internationales de 113 pays, signale le contenu, les comptes et les comportements que nous pouvons alors examiner en contexte. En 2022, nous avons étendu notre réseau de partenaires de confiance à 36 nouveaux pays.

Le contexte local est essentiel à notre compréhension des risques en matière de droits humains, et nous prenons en charge de plus en plus de langues pour rendre cela possible. Nous savons que les droits humains sont particulièrement mis à mal en temps de crise. Meta évalue les risques de danger imminent, même avant une crise, et cherche à y réagir en mettant en place des politiques spécifiques, des services et des actions opérationnelles en matière de droits humains. Le présent rapport présente les progrès de notre [Protocole de politique de crise](#), qui permet de mettre en place une réponse politique réfléchie et calibrée en cas de crise. Nous fournissons des informations relatives à notre travail en Russie, en Ukraine et en Iran, ainsi qu'en matière d'élections au Brésil, au Kenya et au Nigeria.

Nous poursuivons notre travail pour que les droits humains restent au cœur de la manière dont nous comprenons et mettons en œuvre la technologie de l'intelligence artificielle (IA), qui progresse rapidement. Les modèles de l'IA peuvent soutenir les capacités des personnes à exercer leurs droits, mais peuvent aussi présenter des biais problématiques ou des effets discriminatoires et générer des contenus problématiques. Nous nous engageons à aborder ces questions de manière ouverte et collaborative et nous sommes fiers des efforts que nous déployons pour résoudre certains des problèmes les plus difficiles posés par les réseaux sociaux et l'IA.

Pleinement conscients de nos responsabilités, nous nous efforçons de faire progresser nos principes en matière de droits humains à grande échelle, au bénéfice de nos utilisateurs et des sociétés dans leur ensemble, en donnant aux individus le pouvoir de bâtir une communauté.



## 02 À propos de ce rapport de progression

Le deuxième rapport annuel de Meta<sup>1</sup> sur les droits humains reflète les progrès accomplis dans la mise en œuvre des [Principes directeurs des Nations unies relatifs aux entreprises et aux droits humains](#) et de notre [Politique d'entreprise en matière de droits humains](#). Nous expliquons comment nous agissons en matière de droits humains. Il s'agit notamment d'un examen de notre récente évaluation des risques principaux, des enseignements tirés de notre diligence raisonnable en matière de droits humains et des mesures que nous prenons face à ces risques.



Notre [premier Rapport annuel sur les droits humains](#) a été publié en juillet 2022 et témoignait de nos apprentissages et des nos progrès entre le 1er janvier 2020 et le 31 décembre 2021.

Ce second Rapport annuel sur les droits humains témoigne de nos apprentissages et des nos progrès entre le 1er janvier 2022 et le 31 décembre 2022. Il repose sur les bases du premier rapport. Nous rendons compte des services Meta, notamment Facebook, Instagram, WhatsApp, Messenger et Reality Labs.

Nos rapports sont étayés par le [premier rapport de Meta sur les pratiques commerciales responsables](#), qui est conforme aux [normes de la Global Reporting Initiative \(GRI\)](#) et du [Sustainable Accounting Standards Board \(SASB\)](#). Comme indiqué dans le [rapport de Meta sur les pratiques commerciales responsables de 2023](#), les droits humains sont considérés comme un sujet prioritaire dans le cadre de l'évaluation actualisée des sujets prioritaires de Meta.

La politique de Meta en matière de droits humains s'applique à l'ensemble de l'entreprise. Chaque service et entité de Meta dispose de ses propres politiques et procédures avec, parfois, des répercussions différentes sur les droits humains. Ce rapport fait référence à des actions prises par Meta en tant qu'entreprise et concernant une ou plusieurs entités de Meta. Les déclarations n'ont pas pour objectif de suggérer que Meta a pris les mêmes mesures pour toutes les entités.<sup>2</sup>

Lors de la rédaction de ce rapport, nous nous sommes engagés auprès de nombreuses parties prenantes internes et externes et nous avons cherché à concilier concision et exhaustivité.



<sup>1</sup> Le 28 octobre 2021, Facebook, Inc. a changé de nom en faveur de Meta Platforms, Inc. Pour une question de cohérence dans ce rapport, nous utilisons « Meta » pour désigner l'entreprise, que ce soit avant ou après le 28 octobre 2021. Les références à « Facebook » ne concernent que les services et applications de réseaux sociaux, et pas l'entreprise dans son ensemble.

<sup>2</sup> Par exemple, WhatsApp, une application d'appels et de messages chiffrés de bout en bout, dispose d'un point de contact distinct pour les questions de droits humains. L'analyse de la modération des contenus et des actions connexes sur Facebook et Instagram présentée dans ce rapport ne s'applique pas à WhatsApp et, à moins qu'une politique ou une action ne soit spécifiée comme s'appliquant à WhatsApp, elle ne s'applique pas à WhatsApp. En outre, bien que de nombreuses actions décrites dans ce rapport s'appliquent à Instagram et Facebook, 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.



La mission de Meta est de donner à toutes et à tous la possibilité de créer une communauté et de rapprocher le monde entier. Nos services et nos applications permettent aux organisations locales de s'épanouir et de remettre en question l'autorité et la doctrine établies. Un manifestant en Iran, un militant LGBTQIA+ en Ouganda, un journaliste à Kiev... c'est grâce à eux, et à tous ceux qui utilisent nos services, que nous nous efforçons de respecter les droits humains.

Étant donné que des milliards de personnes utilisent nos applications chaque jour, nous nous efforçons de prendre les responsabilités qui s'imposent à une telle échelle. Compte tenu de la rapidité des changements technologiques et politiques, la compréhension et la prise en compte de l'incidence de nos services sur les droits humains constituent un défi en constante évolution. Meta est un pionnier de la recherche et du développement en matière d'IA depuis plus de dix ans, et nous adaptons continuellement notre approche pour nous acquitter de nos responsabilités en matière de droits humains et rester vigilants face aux risques émergents à mesure que de nouveaux outils puissants tels que l'IA générative évoluent rapidement.

Les enjeux en matière de droits humains auxquels nous sommes confrontés sont complexes, varient selon les pays et les régions, et sont influencés par l'évolution des réglementations

gouvernementales dans le monde entier. Par exemple, la [loi sur les services numériques de l'Union européenne](#), adoptée en 2022, crée d'importantes obligations de diligence raisonnable et de déclaration pour Meta, mais elle est explicitement liée au cadre des droits fondamentaux de la région. Cette situation s'inscrit dans le contexte de la menace croissante qui pèse sur l'Internet ouvert et les droits fondamentaux de nos utilisateurs du fait de la propagation d'un modèle d'Internet autoritaire, dans lequel les citoyens sont séparés du reste de l'Internet mondial.

**« Plus de trois milliards et demi de personnes utilisent les applications de Meta – Facebook, Instagram, WhatsApp ou Messenger – chaque mois. Cela signifie qu'entre un tiers et la moitié de tous les êtres humains de la planète les utilisent. Ce faisant, ils ont accès à un univers interconnecté de personnes, d'idées, d'informations, de communautés et de commerce qui ne connaît pas de frontières locales ou nationales. Ce niveau d'interconnectivité est sans précédent dans l'histoire de l'humanité. »**

---

Nick Clegg, Président, Affaires Mondiales

Dans le présent rapport, nous décrivons les progrès réalisés en matière de respect des droits humains dans l'ensemble de nos programmes, de nos services et de nos politiques. Nous continuons à explorer de nouvelles voies pour intégrer les droits humains à nos services et à nos pratiques. Cela implique notamment de [faire progresser la démocratie](#), de [garantir un métavers ouvert et interopérable](#) et de trouver des moyens de garantir que [l'IA réponde aux besoins des communautés de manière équitable](#).

Tous nos efforts sont fondés sur notre engagement à respecter notre [Politique d'entreprise en matière de droits humains](#) et les [Principes directeurs des Nations unies relatifs aux entreprises et aux droits de l'homme](#) (United Nations Guiding Principles on Business and Human Rights, UNGP). Nous ne sommes pas seuls dans cet effort. Le [Conseil de surveillance](#) joue un rôle essentiel, en tant qu'arbitre indépendant et expert, en examinant les décisions relatives au contenu et en proposant des recommandations et des avis consultatifs en matière de politique. Conformément à sa charte, le Conseil de surveillance accorde une attention particulière aux normes relatives aux droits humains qui protègent la liberté d'expression.

Notre travail s'appuie également sur les principes et les pratiques des organisations multilatérales spécialisées et des Nations unies, ainsi que sur les alliances multipartites. Nous sommes membres du [Pacte mondial des Nations unies](#) et nous soutenons ses principes. Nous continuons également à participer au [projet B-Tech](#) du Haut-Commissariat des Nations unies aux droits de l'homme, qui vise à élaborer des directives et des ressources faisant autorité pour la mise en œuvre des UNGP dans le secteur de la technologie. En tant que membre de la [Global Network Initiative](#) (GNI), nous nous engageons à respecter ses principes et ses lignes directrices de mise en œuvre. En 2022, nous avons réussi l'évaluation triennale de la GNI.

Nous continuons également à explorer de nouvelles formes de gouvernance. Par exemple, nous avons lancé un projet pilote de forums communautaires afin de réunir divers groupes du monde entier pour discuter de thèmes difficiles et partager leurs points de vue sur une série de recommandations.

Ce rapport vise à donner une meilleure idée du travail accompli par Meta pour identifier, atténuer et prévenir les risques liés aux droits humains en 2022. Comme toujours, il reste du travail à accomplir et nous nous réjouissons à l'idée de présenter d'autres rapports sur nos progrès dans les années à venir.

#### SIGNÉ :



**Nick Clegg,**  
Président des affaires  
mondiales



**Jennifer Newstead,**  
Responsable des  
affaires juridiques

## 04 Gestions et administration du travail sur les droits humains par Meta



Des structures de gestion et de gouvernance claires nous permettent de faire progresser notre mission de respect des droits humains dans l'ensemble de nos programmes, de nos services et de nos politiques. Notre équipe des droits humains chez Meta guide la mise en œuvre de la [Politique d'entreprise en matière de droits humains](#), supervisée par le président des affaires mondiales et la responsable des affaires juridiques.

Comme décrit dans notre [premier Rapport annuel sur les droits humains](#), les tâches de l'équipe comprennent la promotion de l'intégration de la politique dans les politiques, programmes et services existants et en cours d'élaboration, la mise en œuvre d'une diligence raisonnable et le soutien à la formation des employés concernant la politique. Cette politique aide les équipes à concevoir des produits respectueux des droits, à répondre aux crises émergentes et à travailler avec rapidité et souplesse pour intégrer les droits humains à grande échelle.

Notre Politique d'entreprise en matière de droits humains nous impose de présenter des rapports périodiques au Conseil d'administration sur les principales questions relatives aux droits humains. Le Comité d'audit et de surveillance des risques du Conseil d'administration surveille les différents risques auxquels l'entreprise est exposée, y compris les risques liés aux droits humains, ainsi que les mesures prises par la direction pour contrôler ou atténuer ces risques. Le comité est périodiquement informé des travaux existants et en cours menés par l'équipe chargée des droits humains.

## Formation des employés de Meta

Chez Meta, *la manière* dont nous agissons est aussi importante que *ce que nous faisons*. Nos formations en matière de droits humains permettent à nos employés de mieux comprendre leurs responsabilités et les connaissances et compétences nécessaires pour les remplir.

*Au-delà de Meta : les droits humains*, notre formation lancée en 2022, met en lumière les incidences potentielles et réelles des services, des politiques et des décisions commerciales de Meta 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 tous ceux qui utilisent nos services.

Cette formation complète la formation de Meta sur les droits civils, lancée en juillet 2022 et axée sur les principes de non-discrimination, de justice et d'équité.

## Appels et recours

Chez Meta, nous nous efforçons de fournir des pistes permettant aux parties prenantes de signaler leurs inquiétudes, à Meta de les examiner et de proposer des solutions et des mesures correctives conformes au UNGP 31. Meta dispose de plusieurs voies de recours, identifiées dans le [Code de conduite](#) et sur les pages d'aide des plateformes et des applications, y compris une procédure d'appel auprès du [Conseil de surveillance](#), le premier en son genre.

Lorsqu'un contenu est supprimé, nous en informons les utilisateurs et identifions clairement les voies de [recours](#) contre les décisions de suppression de contenu. Les données relatives aux décisions de suppression et à d'autres aspects de l'application de la politique sont publiées dans [l'Espace modération](#). Pour en savoir plus, consultez notre [premier Rapport annuel sur les droits humains](#).

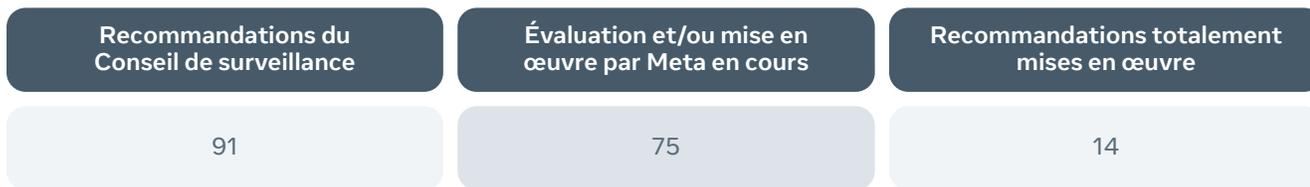
## Conseil de surveillance

Le Conseil de surveillance est un outil unique, à la pointe du secteur, qui vise à [respecter le droit à la liberté d'expression](#). Il fournit au public des mécanismes permettant de faire appel à un organe externe. En 2022, Meta a confirmé son engagement en faveur du Conseil de surveillance en approuvant une nouvelle subvention de 150 millions de dollars sur trois ans en faveur du [Fonds du Conseil de surveillance](#). Cela lui permet de continuer à fonctionner comme une entité distincte et indépendante.

Le Conseil de surveillance prend des [décisions de modération de contenu totalement indépendantes et formule des recommandations](#) concernant la politique de contenu, les services et les opérations. En 2022, le Conseil de surveillance a publié 12 décisions, neuf d'entre elles allant à l'encontre de la décision de Meta. Ces décisions concernaient du contenu d'un grand nombre de pays, notamment la Colombie, l'Éthiopie, l'Afghanistan, la Croatie, le Soudan, l'Inde et les États-Unis. Les thèmes abordés comprenaient la représentation de la violence graphique dans le contenu relatif à un coup d'État militaire, la mention des talibans dans les reportages d'actualité et la description de la violence sexuelle à l'encontre de mineurs. Ces décisions ont toutes fourni aux utilisateurs de la plateforme la possibilité d'un recours.

Outre ses décisions, le Conseil de surveillance a formulé 91 recommandations en 2022, dont 89 sont à l'étude, en cours de mise en œuvre ou totalement mises en œuvre.

#### Recommandations du Conseil de surveillance en 2022



## Global Network Initiative

Meta est membre de la [Global Network Initiative](#) (GNI) depuis 2013. La GNI est une alliance croissante de sociétés Internet et de télécommunications et de divers groupes de la société civile, y compris des groupes de défense des droits humains et de la liberté de la presse du monde entier. Nous nous engageons à respecter les principes de la GNI sur la liberté d'expression et la protection de la vie privée ([principes de la GNI](#) ou GNI Principles), qui reposent sur des lois et des normes reconnues à l'international. Ces engagements ont été confirmés à nouveau dans notre Politique d'entreprise en matière de droits humains en 2021.

Toutes les sociétés de la GNI font l'objet d'une [évaluation](#) indépendante de leur mise en œuvre des principes tous les deux ou trois ans.

En 2022, Meta a réussi son évaluation GNI, qui comprenait un examen de nos politiques et processus, ainsi que huit cas d'étude spécifiques exemplatifs impliquant des demandes de retrait et des demandes de données du gouvernement. 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 », le [critère](#) sur la base duquel les entreprises membres sont évaluées. La GNI publiera un rapport d'évaluation public distinct couvrant toutes les évaluations des entreprises pour le cycle 2021-2022.

## 05 Identification et gestion des principaux risques en matière de droits humains



En 2022, nous avons procédé à une évaluation globale des risques principaux en matière de droits humains (EGRP), un outil fondamental du programme permettant d’orienter la gestion des risques à l’avenir. L’objectif de l’EGRP était d’identifier et de hiérarchiser les principaux risques inhérents aux droits humains de Meta dans l’ensemble de l’entreprise, définis comme les principaux effets négatifs potentiels de Meta<sup>3</sup> en matière de droits humains sur les personnes, y compris les utilisateurs et les autres personnes susceptibles d’être concernées par les actions de Meta. Les risques « inhérents » sont les risques qui existaient avant la mise en place des processus ou leur modification, à la différence des risques « résiduels » qui persistent après la mise en œuvre de processus visant à réduire les risques inhérents.

**Réalisation de l’évaluation :** Business for Social Responsibility (BSR) a mené des recherches documentaires et s’est largement engagé auprès de parties prenantes internes et externes, ce qui a donné lieu à plus d’une douzaine de réunions avec des groupes de la société civile, des représentants des Nations unies et des investisseurs dans de multiples régions et dimensions de la diversité. L’évaluation s’est également appuyée sur les

<sup>3</sup> Le terme « effet négatif en matière de droits humains » est conforme aux UNGP et évoque un effet qui se produit lorsqu’une action supprime ou réduit la capacité d’une personne à jouir de ses droits humains. Le présent rapport ne doit pas être interprété comme suggérant qu’une personne ou une communauté particulière a subi un effet négatif, sauf indication contraire.

enseignements tirés de la diligence raisonnable en matière de droits humains et de l'engagement des parties prenantes. L'évaluation donnait la priorité aux risques en fonction des définitions de la « gravité » (portée, échelle, caractère remédiable) et de la « probabilité », reprises dans les UNGP et dans le cadre de signalement des UNGP :

1. **Portée** : Combien de personnes ont pu être concernées par les effets négatifs potentiels ?
2. **Échelle** : Quelle serait la gravité des effets négatifs potentiels en matière de droits humains pour les personnes concernées ?
3. **Caractère remédiable/irremédiable** : Une mesure corrective permettrait-elle aux personnes potentiellement concernées de retrouver une situation identique ou équivalente ?
4. **Probabilité** : Quelle est la probabilité que ces effets négatifs en matière de droits humains se produisent ?

Conformément aux UNGP, l'EGRP a accordé une attention particulière aux droits, aux besoins et aux difficultés des personnes susceptibles de courir un risque accru de devenir vulnérables ou marginalisées.

**Ce que nous avons appris** : L'analyse a permis d'identifier les huit droits humains les plus importants dans l'ensemble des opérations mondiales :

1. **Liberté d'opinion et d'expression**
2. **Confidentialité**
3. **Égalité et non-discrimination**
4. **Vie, liberté et sécurité de la personne**
5. **Intérêt supérieur de l'enfant**
6. **Participation publique, processus de vote et éligibilité**
7. **Liberté de réunion et d'association**
8. **Droit à la santé**

## Nos principaux risques

L'EGRP a analysé les principaux risques inhérents à tous les droits humains reconnus au niveau international dans l'ensemble de l'entreprise et a hiérarchisé les huit risques les plus importants en utilisant les critères des UNGP et le cadre de signalement des UNGP. Ces risques sont examinés ci-dessous, accompagnés d'exemples illustratifs de certains des facteurs de risque et d'une sélection des efforts de prévention et de limitation des risques déployés par Meta.



## Liberté d'opinion et d'expression

Le [droit à la liberté d'opinion et d'expression](#) est au cœur des convictions de Meta, qui œuvre à le protéger.

### Exemples de principaux risques inhérents aux droits humains identifiés lors de l'EGRP

Des perturbations sur Internet et des blocages des réseaux sociaux empêchent les citoyens de jouir de leur liberté d'expression et les privent de la possibilité de recevoir et d'envoyer des nouvelles et des informations vitales.

Limites gouvernementales abusives en matière de contenu

La mise en application de la modération du contenu par Meta peut limiter la liberté d'expression.

### Exemples de mesures prises par Meta pour remédier au risque

Nous avons récemment [lancé](#) une fonctionnalité qui permet aux utilisateurs de se connecter à WhatsApp par l'intermédiaire d'un serveur proxy lorsque leur connexion Internet est interrompue ou que l'accès à WhatsApp est bloqué.

Nous évaluons les demandes de retrait (DR) gouvernementales légales et valides sur la base des principes de la GNI et notre travail est régulièrement évalué par un organisme indépendant.

Nous effectuons des contrôles opérationnels approfondis pour vérifier la validité des DR et fournissons des [informations sur les DR gouvernementales dans l'Espace modération](#).

Nous procédons à des révisions régulières des politiques de contenu, en sollicitant l'avis de diverses parties prenantes, au regard des normes internationales en matière de droits humains.

Les modèles d'IA prédisent si un contenu constitue un discours haineux ou est violent et explicite. Des systèmes distincts, nos [solutions de mise en application](#), déterminent s'il faut prendre une mesure, telle que la suppression, la rétrogradation ou l'envoi du contenu à une équipe d'examen manuel pour un examen plus approfondi. En mettant nos politiques en application, nous souhaitons atténuer les risques de discours haineux, de violence et d'incitation et d'autres risques liés aux droits humains tout en maintenant la liberté d'expression.



« Toute notre entreprise fait des efforts continus pour prendre les bonnes décisions en matière de respect de la vie privée ; chacun doit contribuer à la réalisation de cet objectif chez Meta. »

Michel Protti, responsable de la confidentialité des données pour les produits

Le droit à la confidentialité est au cœur de la mission de Meta et constitue une condition nécessaire au respect d'autres droits humains, tels que la liberté d'expression, la liberté de réunion et d'association, ainsi que la liberté de croyance et de religion.

Nous nous engageons à améliorer et à adapter nos pratiques et nos politiques afin d'anticiper et de relever les défis liés à la protection de la confidentialité des personnes sur nos services et nos applications.

### Exemples de principaux risques inhérents aux droits humains identifiés lors de l'EGRP

Demandes gouvernementales abusives ou inutiles de données d'utilisateurs

### Exemples de mesures prises par Meta pour remédier au risque

Nous protégeons les utilisateurs des demandes de données du gouvernement illégales ou abusives par le biais de notre équipe d'intervention et d'application de la loi.

Nous rejetons les demandes gouvernementales ne respectant pas les normes reconnues à l'international, y compris celles de notre Politique d'entreprise en matière de droits humains et nous les contestons devant les tribunaux si nécessaire.

Nous répondons aux demandes légales valables lorsque nous croyons de bonne foi que la loi nous y oblige. Nous répondons également à certaines demandes légales sans que la loi ne nous y oblige, lorsque nous jugeons en toute bonne foi que notre réponse :

- est requise par la loi dans une autre juridiction dans laquelle un utilisateur peut se trouver ;
- concerne les utilisateurs de ce territoire ; et
- répond aux standards internationalement reconnus, y compris notre Politique d'entreprise en matière de droits humains.

Lorsque nous nous conformons aux demandes de données gouvernementales, nous nous

(suite)

Le contenu ou le comportement sur les applications Meta peut avoir un effet négatif sur la confidentialité et les droits en matière de protection des données (par exemple, doxxing, récupération).

Meta ou des tiers peuvent utiliser des données d'utilisateur sensibles ou à caractère hautement personnel, y compris celles liées à des caractéristiques protégées, à des fins de publicité ciblée d'une manière qui a des conséquences négatives sur les droits à la confidentialité et à la protection des données.

(suite)

efforçons d'informer les personnes lorsque leurs données ont été demandées et nous publions un rapport de transparence sur les demandes gouvernementales que nous recevons.

Meta dispose de Standards de la communauté et Règles de la communauté solides qui régissent la manière dont les utilisateurs peuvent utiliser nos applications. Nous n'autorisons pas les contenus qui partagent, offrent ou sollicitent des informations personnelles identifiables ou d'autres informations privées susceptibles de causer un préjudice physique ou financier. Nous supprimons les contenus qui vont à l'encontre de nos politiques dès que nous en prenons connaissance.

L'utilisation de l'automatisation pour accéder ou recueillir des données à partir des services et applications Meta sans notre autorisation (« récupération ») constitue une violation de nos conditions de service. Nous disposons d'une équipe dédiée, de mesures techniques et de processus qui se concentrent sur la détection, l'investigation et le blocage des tentatives de récupération non autorisées.

Les informations que nous recueillons et les processus dépendent du service que nous offrons. Pour nous aider dans cette tâche, nous développons des technologies protectrices de la confidentialité (PETs) qui permettent de limiter la collecte et l'utilisation des données. Notre équipe se consacre à l'élaboration de PETs pour les équipes de Meta, en se concentrant sur des domaines clés tels que la dépersonnalisation des données lors de la collecte et en permettant aux équipes de mettre en œuvre le chiffrement de bout en bout dans nos applications et nos services.

En consultation avec des experts en droits civils, des décideurs politiques et d'autres parties prenantes, nous avons mis en œuvre des changements visant à améliorer le contrôle des utilisateurs sur les publicités, et nous avons supprimé les options de ciblage des publicités liées à des sujets pouvant être perçus comme sensibles.

(suite)

(suite)

Nous avons collaboré avec le ministère de la Justice des États-Unis pour développer et lancer le système de réduction des variations (Variance Reduction System, VRS), une nouvelle technologie qui permettra de distribuer certaines publicités sur nos applications de manière plus équitable. Nous avons lancé le VRS aux États-Unis pour les annonces de logement et nous allons l'étendre aux annonces relatives aux emplois et aux crédits.

L'EGRP a identifié plusieurs pistes de risques liés à la publicité ciblée tout en soulignant la nature nuancée de ces risques. Par exemple, si les modèles commerciaux basés sur la publicité peuvent entraîner des risques pour le droit à la confidentialité, il est possible d'établir des mesures visant à atténuer ces risques tout en préservant la capacité générale des personnes du monde entier à exprimer leurs idées et à communiquer. De même, la publicité est un moteur essentiel pour les petites et moyennes entreprises qui, sans elle, n'auraient pas d'accès viable à des clients potentiels.



## Égalité et non-discrimination

Le [droit à l'égalité et à la non-discrimination](#) prévoit une protection égale contre toute discrimination. Alors que la démocratisation de la technologie continue de rendre l'Internet plus accessible à un plus grand nombre de personnes, toutes les communautés ne partagent pas les mêmes expériences en ligne, en particulier les communautés et individus marginalisés.

Nous aspirons à créer un environnement en ligne plus inclusif et plus équitable pour tous les utilisateurs grâce à nos Standards de la communauté et à nos Règles de la communauté, qui interdisent les discours haineux, l'intimidation et le harcèlement. Nos équipes chargées des droits humains et des droits civils s'efforcent ensemble de respecter l'égalité des droits des personnes et de renforcer l'équité sur toutes nos technologies.

### Exemples de principaux risques inhérents aux droits humains identifiés lors de l'EGRP

Contenu ayant un effet négatif sur l'égalité et la non-discrimination (par exemple, discours haineux)

### Exemples de mesures prises par Meta pour remédier au risque

Nos Standards de la communauté Facebook, y compris les politiques relatives aux [discours haineux](#) et aux caractéristiques protégées, définissent ce qui est autorisé ou non sur Facebook. Les [Règles de la communauté Instagram](#) définissent ce qui est autorisé ou non sur Instagram. Nous prenons des mesures contre les contenus qui enfreignent nos politiques. Nos [Standards publicitaires](#) fournissent des précisions et des orientations sur les types de contenus publicitaires que nous autorisons et sur les types de contenus et de ciblage publicitaires que nous interdisons.

Nous avons poursuivi [notre travail de création de systèmes d'IA responsables](#) afin de réduire les biais et d'améliorer l'équité, notamment dans des domaines tels que la diffusion des publicités et la réduction de la prévalence du langage préjudiciable dans les applications d'IA générative. Nous avons largement progressé pour aider les chercheurs à mesurer l'équité de nos modèles d'IA. Un manque de données diversifiées, ou de données représentant des personnes et des expériences très variées, peut conduire à des résultats alimentés par l'IA qui

(suite)

(suite)

reflètent des stéréotypes problématiques ou ne fonctionnent pas de la même manière pour tout le monde. C'est pourquoi nous avons travaillé à l'élaboration de [Casual Conversations v2](#), un ensemble de données public axé sur le consentement, grâce auquel les chercheurs peuvent évaluer plus précisément l'équité et la robustesse de certains types de modèles IA, afin de les rendre plus inclusifs.

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

Nous avons fait appel à des modérateurs de contenu humains et à des IA pour nous aider à supprimer le contenu en infraction avec nos politiques de nos applications.

Nous utilisons des [traductions humaines et automatiques](#) pour améliorer la modération du contenu. Nous faisons constamment évoluer nos flux de rapports et nos classificateurs d'intégrité dans les langues que nous prenons en charge.



#### **Discours haineux**

Nous avons supprimé environ 50 millions de contenus haineux de Facebook en 2022, dont 91 % ont été détectés et ont fait l'objet d'une action de manière proactive. Nous avons supprimé 16 millions de contenus haineux d'Instagram, dont 92 % ont été détectés et ont fait l'objet d'une action de manière proactive. Les discours haineux ont représenté moins de 0,02 % du contenu vu sur Facebook et Instagram.

Pour en savoir plus sur la manière dont l'équipe chargée des droits civils de Meta traite la question de l'égalité et de la non-discrimination, voir ce [rapport de progression](#) et cette [mise à jour](#).

## Accessibilité

À mesure que la technologie progresse, il est important que les personnes en situation de handicap ne soient pas laissées pour compte. Nous veillons à améliorer continuellement nos efforts pour garantir l'accessibilité de nos services et de nos applications. Les rendre accessibles aux personnes en situation de handicap est un engagement clé de notre Politique d'entreprise en matière de droits humains, dans laquelle nous citons la [Convention relative aux droits des personnes handicapées](#) en tant que norme internationale de base.



des 50 millions de contenus haineux supprimés de Facebook en 2022 ont été détectés et ont fait l'objet d'une action de manière proactive.

**Nos mesures : développer les efforts intersectoriels pour améliorer l'accessibilité**

Les technologies de réalité virtuelle, augmentée et mixte sont particulièrement prometteuses pour les personnes en situation de handicap, qu'il s'agisse de technologies d'assistance, de solutions de réadaptation ou de formation, ou d'expériences favorisant la prise de conscience et la compréhension. Meta poursuit sa collaboration avec la [XR Association](#) pour des projets visant à améliorer les connaissances liées à l'accessibilité et les efforts du secteur entier comme le lancement d'un nouveau site consacré à l'accessibilité, [xraccessibility.github.io](#), en collaboration avec [XR Access](#). Parmi les récentes innovations, citons les lunettes intelligentes [Ray-Ban Stories](#) alimentées par Meta qui permettent d'envoyer des messages et de passer des appels sur WhatsApp et Messenger en ayant les mains libres, ainsi que de réaliser des captures d'image et de vidéo à l'aide de commandes vocales et tactiles.

En plus d'intégrer des fonctionnalités dans ses services, Meta a développé et partagé des recommandations techniques pour les développeurs afin de créer des applications de réalité virtuelle accessibles. Nous avons également collaboré avec la communauté des personnes en situation de handicap et avec des organisations du secteur pour élaborer et lancer des [règles](#) d'accessibilité XR à l'intention des développeurs, en partenariat avec la XR Association.





## Vie, liberté et sécurité de la personne

Le droit à la vie, à la liberté et à la sécurité de la personne concerne le fait d'être à l'abri de l'enfermement physique ainsi que la protection contre les atteintes au corps et à l'esprit. Pour Meta, le respect de ce droit humain implique le risque que le contenu provoque des dommages, notamment la 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 principaux risques inhérents aux droits humains identifiés lors de l'EGRP

Personnes malveillantes :

- Se servir des services et applications Meta pour coordonner des préjudices en ligne et hors ligne
- Abuser des services et applications pour mettre en place des cyberattaques ou de l'hameçonnage
- Menacer ou harceler des défenseurs des droits humains (DDH), des militants ou d'autres groupes vulnérables

### Exemples de mesures prises par Meta pour remédier au risque

Nos politiques en matière de sécurité sont conçues pour identifier les menaces extérieures et s'en défendre. Meta dispose d'outils et de ressources pour les DDH et les journalistes, élaborés en tenant compte des avis des DDH.

Nous avons élargi l'accès à Facebook Protect à un plus grand nombre de personnes, en particulier aux DDH, leur permettant ainsi de mieux contrôler leurs comptes, notamment les personnes autorisées à voir leurs publications.

Le Guide de sécurité de notre Centre de confidentialité centralise une série de ressources destinées à aider les utilisateurs à gérer leurs paramètres de confidentialité afin de se protéger contre l'espionnage en ligne ou l'accès non autorisé ou forcé. Il fournit également de l'aide pour trouver un soutien en cas d'intimidation et de harcèlement. Nous avons collaboré avec des experts et des DDH sur toute une série de questions afin de créer des politiques, des outils et des ressources, notamment en matière d'alphabétisation numérique, de protection de l'enfance et de conseils aux parents.

Meta investit pour lutter contre l'utilisation abusive de ses services par des gouvernements autoritaires, des groupes terroristes ou d'autres acteurs malveillants qui tentent de surveiller les critiques du régime, les figures d'opposition et les DDH. Nous prenons des mesures lorsque nous constatons une telle activité, nous bloquons l'infrastructure du domaine pour qu'il ne soit pas partagé sur nos services et, lorsque cela est possible et approprié, nous informons les personnes qui, selon nous, ont été ciblées par ces opérations malveillantes.

## Davantage de manières de respecter le droit à la vie, à la liberté et à la sécurité de la personne

Les initiatives ci-après mettent en lumière nos approches pour aider les groupes et les individus qui courent un risque.

### Protection des DDH et des journalistes

Nous avons développé des outils de ressources pour les DDH, les journalistes et d'autres personnes vulnérables dans le centre de sécurité Meta. Notre [Guide de la sécurité pour les journalistes](#) permet d'assurer que les journalistes sont équipés pour gérer leur sécurité en ligne grâce à des outils et des ressources permettant de protéger leurs sources. Nous soutenons aussi la communauté internationale de l'information en finançant des programmes par le biais de notre [Meta Journalism Project](#).

En 2022, nous avons entrepris les actions suivantes :

- Nous avons renforcé la protection des militants et des journalistes contre les menaces de violence, notamment en supprimant les contenus qui présentent des personnes comme des militants dans des situations susceptibles de les mettre en danger.
- Nous avons sécurisé les comptes des DDH arrêtés ou détenus dans des États autoritaires ou dans des situations de conflit contre tout accès non autorisé par les forces de l'ordre locales, les services de sécurité et les acteurs du conflit.
- Nous avons fait progresser le [Fonds Meta pour les défenseurs des droits humains et l'Initiative pour la sécurité des journalistes en Asie-Pacifique](#), grâce auxquels le groupe international de défense des droits humains Civil Rights Defenders a fourni une évacuation urgente, une relocalisation temporaire, des mesures de sécurité numérique, des formations à la sécurité et une aide juridique à plus de 500 DDH, principalement en Asie du Sud et du Sud-Est. Le groupe Civil Rights Defenders a également procédé à des évaluations de la sécurité, facilité la mise en œuvre de mesures de sécurité et fourni d'autres soutiens liés à la sécurité aux DDH et aux organisations dans les régions concernées.
- Nous avons renforcé notre engagement auprès des DDH afin de leur fournir un avis régulier sur les conséquences des menaces régionales et mondiales qui pèsent sur eux, et nous avons formé les DDH au Guide de la sécurité du Centre de confidentialité.
- Nous avons mis en place une formation gratuite en ligne sur la sécurité numérique pour les journalistes et les DDH, créée et gérée par l'[International Center for Journalists](#), en partenariat avec le Border Center for Journalists and Bloggers.

En 2022, nous avons présenté le guide de la sécurité du Centre de confidentialité au bureau du rapporteur spécial des Nations unies sur les DDH.

### Traite d'êtres humains et exploitation

Pour empêcher et réduire ces pratiques, nous supprimons le contenu qui contribue à l'exploitation des personnes ou la coordonne, notamment la traite d'êtres humains.<sup>4</sup>

En 2022, nous avons mis à jour notre [Politique en matière d'exploitation humaine](#) en vue d'interdire le contenu qui invite à recourir à des services de passeurs. Nous fournissons également des [ressources d'information](#) sur Facebook concernant les risques liés à la collaboration avec les passeurs, les signes d'exploitation potentielle et les moyens de demander une immigration légale, y compris l'asile. Nous avons élaboré ces ressources en consultation avec des experts, notamment l'[Organisation internationale pour les migrations](#). Nous autorisons toujours les contenus qui sollicitent ou partagent des

<sup>4</sup> Pour en savoir plus sur la traite d'êtres humains et l'exploitation, consultez la page 32 de notre [premier Rapport annuel sur les droits humains](#).

informations sur la sécurité personnelle et le franchissement des frontières, ainsi que sur la manière de quitter un pays ou de demander l'asile. Nous autorisons le contenu condamnant la traite d'êtres humains ou sensibilisant le public à ce sujet.



#### Lignes d'écoute dans l'État du Bihar en Inde, pour les personnes victimes de trafic

La traite d'êtres humains est particulièrement répandue dans l'État du Bihar en Inde. Pour aider les victimes, Meta a soutenu la [mise en place d'une ligne d'écoute](#) sur WhatsApp, qui aide le système de traçage du ministère du Travail de l'État du Bihar. Les utilisateurs peuvent signaler les cas de travail des enfants à un numéro WhatsApp spécifique, disponible en anglais, en hindi et en magahi. La ligne d'écoute permet également de surveiller les progrès des anciens enfants travailleurs réinsérés dans cet État.

#### Surveillance à la demande

La surveillance à la demande permet à certains gouvernements et à certaines personnes malveillantes de faire taire certaines voix en ciblant et en intimidant les activistes, les dissidents, les journalistes et les chercheurs au moyen de logiciels espions abusifs. Ces logiciels espions ciblent des personnes à leur insu, sans égard pour les droits humains, les libertés civiles et l'État de droit.

Les pratiques de surveillance à la demande exploitent nos services comme un vecteur pour implanter des logiciels malveillants sur les appareils mobiles de nos utilisateurs. Nous avons interrompu et signalé un grand nombre de ces opérations. En 2022, nous avons collaboré avec des partenaires du secteur pour proposer des principes et des pratiques visant à prévenir et à atténuer cette menace. Nous avons révélé ces principes lors du [Sommet pour la démocratie](#), sous l'égide de l'[Accord en matière de cybersécurité et de technologie](#). Dans nos [recommandations politiques](#) de décembre 2022, nous proposons une feuille de route pour que les autres parties prenantes, y compris les gouvernements, le secteur de la surveillance à la demande lui-même et la société civile, unissent leurs forces pour promouvoir les droits et les libertés en ligne.

#### Images intimes non consentuelles (IINC)

Partager des images ou des vidéos intimes d'une personne, que ce soit en ligne ou hors ligne, sans son consentement, constitue une violation des droits de cette personne à la confidentialité et à la sécurité. Les images sont souvent diffusées sur plusieurs plateformes et applications. Le partage d'IINC est souvent appelé « revenge porn », car le partage de contenus intimes peut être perçu comme un acte de « vengeance » à l'égard d'une personne après la fin d'une relation.

Meta a collaboré avec des organisations à but non lucratif et d'autres entreprises pour créer un outil gratuit, [StopNCII.org](#), qui permet aux victimes potentielles de signaler des images afin que les entreprises participantes puissent les détecter et les supprimer pour éviter qu'elles ne soient partagées en ligne. En une seule opération, les individus peuvent demander la suppression des IINC sur plusieurs plateformes et applications. StopNCII.org, qui s'appuie sur une technologie développée par Facebook et Instagram, est disponible en 22 langues et fonctionne sur Facebook, Instagram et les plateformes des autres [entreprises participantes](#), notamment TikTok et Bumble.



## Intérêt supérieur de l'enfant

L'[intérêt supérieur de l'enfant](#) donne aux enfants le droit de voir leur intérêt supérieur pris en compte dans toutes les décisions et actions qui les concernent, y compris la protection contre la [traite d'êtres humains](#). Le [cadre](#) des intérêts supérieurs de l'enfant de Meta est conforme aux valeurs fondamentales de la [Convention des Nations unies relative aux droits de l'enfant](#).

Nous souhaitons que les jeunes puissent avoir des expériences en ligne positives et sûres. La [Protection des enfants en ligne](#) est une des priorités de Meta, c'est pourquoi nous proposons des outils permettant aux adolescents de donner plus de sens à leur temps passé en ligne, notamment des moyens supplémentaires de se concentrer et de fixer des limites, tels que des rappels de pause, des conseils d'experts, un mode silencieux, des encouragements à changer de sujet et l'option de masquer le nombre de « j'aime » sur les publications dans le fil. Notre approche de l'élaboration de ces technologies est fondée sur la recherche et les avis directs de parents, d'adolescents et d'experts de la santé mentale et de la psychologie de l'enfant.

### Exemples de principaux risques inhérents aux droits humains identifiés lors de l'EGRP

Exposition d'enfants à du contenu inapproprié

### Exemples de mesures prises par Meta pour remédier au risque

Nos politiques interdisent aux enfants de moins de 13 ans de créer un compte sur Facebook et Instagram. Nous rédigeons nos politiques afin de favoriser une expérience sûre pour les personnes à partir de 13 ans, et nous consultons des experts de la sécurité des enfants lors de l'élaboration de ces politiques. Nous avons mis en place une restriction liée à l'âge pour certains contenus (p. ex. certains contenus montrant de la violence graphique) ; nous nous efforçons d'éviter de recommander certains contenus non violents aux mineurs (p. ex. l'alcool, le tabac) et nous disposons d'un [contrôle de contenu sensible](#) configuré par défaut sur les paramètres les plus stricts pour les mineurs. Nous interdisons les contenus qui sexualisent les mineurs et, lorsque nous avons connaissance d'une exploitation sexuelle apparente d'enfants, nous la signalons au Centre national des États-Unis pour les enfants disparus et exploités ([National Center for Missing and Exploited Children](#), NCMEC).

Nous appliquons également des politiques publicitaires plus restrictives pour les mineurs (p. ex., en plus d'interdire les publicités pour les biens et services restreints, comme l'alcool ou le tabac, nous n'autorisons pas non plus les publicités pour des produits de perte de poids).

(suite)

(suite)

Nos politiques interdisent, et Facebook et Instagram cherchent à prévenir, les contenus qui encouragent le suicide, l'automutilation, les troubles alimentaires ou l'intimidation et le harcèlement. Nous supprimons ce type de contenu lorsque nous le trouvons et continuons à améliorer la technologie utilisée pour le détecter et le supprimer.

Exploitation sexuelle et abus d'enfants

La section consacrée à la protection de l'enfance de notre [Centre de sécurité](#) met en évidence notre approche globale de la sécurité des enfants, qui comprend des politiques interdisant l'exploitation des enfants et une technologie de pointe permettant de prévenir, de détecter, de supprimer et de signaler les violations des politiques.

Nous avons conçu plus de 30 outils pour aider les enfants, les adolescents et leurs parents, y compris, sur Instagram, des [outils de supervision dans notre Centre familial](#) qui aident les parents et leurs enfants à naviguer ensemble sur Internet.

Nous utilisons une technologie de détection sophistiquée sur Facebook et Instagram visant à détecter, supprimer et prévenir le partage d'images et de vidéos qui exploitent les enfants. Nous signalons au NCMEC et aux autres agences dans le monde les cas apparents d'exploitation d'enfants figurant sur nos sites. En 2022, nous avons soutenu le NCMEC dans la mise en place de la plateforme [Take It Down](#) pour les mineurs qui s'inquiètent de voir leurs images intimes partagées sur des plateformes et des applications en ligne.



## 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 un des fondements de la démocratie. Il s'agit également d'un droit essentiel à l'État de droit, à l'inclusion sociale, au développement économique et à la progression de tous les droits humains.

Protéger l'intégrité des élections sur nos services et applications est l'une de nos priorités. Meta [investit](#) des efforts et des ressources considérables pour protéger les élections sur Internet avant, pendant et après les périodes électorales. Notre objectif est de donner la parole aux citoyens et de les aider à participer au processus civique tout en luttant contre les discours haineux, l'ingérence et l'influence étrangère.

## Exemples de principaux risques inhérents aux droits humains identifiés lors de l'EGRP

Contenu ayant une incidence négative sur la participation publique, le vote ou la candidature à une fonction publique

## Exemples de mesures prises par Meta pour remédier au risque

Il est interdit de tenter de perturber ou d'empêcher le déroulement du scrutin. Nous nous efforçons d'identifier et de supprimer ces contenus, quelle que soit leur origine.

Nous offrons un niveau de transparence inégalé dans le secteur en ce qui concerne la publicité et les pages politiques, afin que les utilisateurs puissent voir qui tente d'influencer les votes. Sur certains marchés, nous interdisons les publicités qui découragent le vote, revendiquent prématurément la victoire lors d'une élection, tentent de discréditer l'élection ou sont incompatibles avec les recommandations des autorités sanitaires concernant la sécurité du vote.

Les discours haineux peuvent avoir des conséquences sur les électeurs. Nos [Standards de la communauté](#) comprennent des politiques relatives au harcèlement, à l'incitation, aux discours haineux et violents, et interdisent les attaques de personnes fondées sur des caractéristiques telles que l'appartenance ethnique ou la religion. Nous supprimons le contenu qui enfreint ces règles dès que nous en prenons connaissance.

Tentatives d'empêcher des personnes de voter, augmentation du nombre de spams, ingérence étrangère ou signalements de contenus contraires à nos politiques

Nous recherchons activement dans les contenus les organisations que Meta a interdites car elles tentaient d'enfreindre nos politiques ou de causer des préjudices hors ligne.

En 2022, nous avons mis à jour [Facebook Protect](#), qui propose des outils de sécurité et des protections supplémentaires aux candidats, à leurs campagnes et aux élus.

Les personnes malveillantes qui interfèrent avec les élections

Nos opérations de sécurité avancées contribuent à la sécurité des élections en bloquant les campagnes de manipulation, en identifiant les menaces émergentes et en déployant notre réseau mondial de vérification des informations de pointe.

Nous utilisons l'IA pour identifier les faux comptes et supprimer la grande majorité d'entre eux avant qu'ils ne soient signalés. Nous avons supprimé plus de 150 réseaux de comportement non authentique coordonné depuis 2017.

## Plus d'informations sur notre travail en matière d'élections nationales

Pour les élections présentant un risque plus élevé de contenu abusif, nous réunissons une équipe spécialisée et interfonctionnelle ayant une expertise en matière d'intégrité des élections, de fausses informations, de sécurité, de droits humains et de cybersécurité, et disposant d'une expertise linguistique pertinente sur le marché concerné. L'équipe surveille les risques émergents et y réagit en temps réel. Elle surveille également la couverture médiatique et les activités liées aux élections sur les autres réseaux sociaux et sur les médias traditionnels. Ces efforts nous donnent une vue d'ensemble et nous aident à déterminer quel type de contenu peut devenir viral.

En 2022, nous avons constitué et renforcé des équipes afin d'être prêts pour les élections dans le monde entier, notamment au Brésil, au Kenya et au Nigeria, ainsi que pour les élections de mi-mandat aux États-Unis. Nous avons réalisé des investissements, engagé des parties prenantes et des experts, établi des partenariats et amélioré les technologies.



### Élections de mi-mandat en 2022 aux États-Unis

Au cours de la période précédant ces élections, nous avons collaboré avec les autorités électorales locales et des États pour envoyer des alertes de vote via le fil, ce qui a donné lieu à plus de 80 millions de notifications liées aux élections. Nous avons combattu des opérations tant étrangères que nationales et nous avons exposé et interrompu des dizaines de réseaux qui tentaient d'interférer avec les élections. Nous avons banni plus d'un millier de mouvements sociaux militarisés et supprimé des dizaines de milliers de pages, groupes et comptes QAnon de nos applications. Nous avons également supprimé le groupe original Stop the Steal et banni plus de 270 organisations suprémacistes blanches.

### **Brésil**

Des élections législatives ont eu lieu au Brésil au début du mois d'octobre 2022, avec un second tour de l'élection présidentielle plus tard le même mois. Les élections se sont déroulées dans un contexte de crise économique et d'inquiétude quant aux allégations de perte de légitimité des élections. Notre travail sur l'intégrité des élections est constant, et nous avons intensifié nos efforts pour les élections brésiliennes de 2022 avec un an d'avance. Nous nous sommes concentrés sur quatre domaines clés pour préparer les élections : la prévention des interférences, la collaboration avec les autorités électorales, la lutte contre les fausses informations et la désinformation, et l'amélioration de la transparence dans les publicités.

Nous avons lancé notre centre d'opérations électorales pour permettre aux spécialistes de l'intégrité au Brésil et dans le monde entier d'identifier en temps réel les menaces potentielles pesant sur nos services et applications. Entre août et octobre, nous avons eu recours à une évaluation humaine et à une IA pour nous aider à détecter et à supprimer de manière proactive plus de 310 000 contenus de Facebook et d'Instagram provenant d'utilisateurs au Brésil pour violation de notre [politique en](#)

matière de violence et d'incitation, et plus de 290 000 contenus pour violation de notre politique en matière de discours haineux. Par ailleurs, nous avons réduit la diffusion des contenus que notre technologie a identifiés comme susceptibles d'enfreindre ces politiques, afin d'éviter qu'ils ne se propagent rapidement. Nous avons publié sur Facebook et Instagram des rappels concernant le jour de l'élection, ainsi que des rappels concernant la mise à jour des cartes électorales.

Du 1er janvier 2022 au 2 octobre 2022, plus de 30 millions de personnes ont cliqué sur les vignettes d'information électorale ajoutées aux messages concernant les élections diffusés sur Facebook au Brésil et ont été dirigées vers les informations officielles sur le site web du Tribunal Superior Electoral (TSE). Plus de 4,7 millions de personnes se sont abonnées au service de messagerie instantanée du TSE sur WhatsApp pour recevoir des informations fiables. Pour lutter contre la violence en ligne à l'encontre des femmes en politique, nous avons lancé un guide avec le soutien du tribunal TSE et du réseau Women's Democracy Network au Brésil.

Pour lutter contre les fausses informations, nous avons travaillé (et continuons à travailler) avec des agences indépendantes de vérification des informations pour vérifier la véracité des publications, des reels, des stories, des vidéos, des images et des commentaires sur Facebook et Instagram en portugais. En 2022, le nombre de partenaires de notre programme de vérification des informations au Brésil est passé de quatre à six.

Pour lutter contre la désinformation, nos équipes internes ont repéré et supprimé les réseaux qui enfreignaient notre politique en matière de comportement inauthentique coordonné.

Pour améliorer la transparence des publicités, nous avons étendu les politiques de transparence, au-delà des publicités politiques, aux publicités sur Facebook et Instagram au Brésil sur des questions sociales telles que l'économie, la sécurité et l'éducation. Nous avons rejeté 135 000 soumissions publicitaires provenant de sources nationales et étrangères qui incluaient le Brésil dans leur ciblage.

Retrouvez davantage d'informations sur nos efforts lors des élections au Brésil [ici](#) et [ici](#).

### Kenya

Le Kenya a tenu ses élections législatives en août 2022. Les processus électoraux précédents avaient été entachés de violations des droits humains. Avant les élections de 2022, nous nous sommes engagés avec des défenseurs des droits humains (DDH), des agences de l'ONU, des journalistes et des membres de la société civile, en prenant un certain nombre de mesures pour renforcer nos dispositifs de sécurité et d'intégrité, notamment en élargissant notre programme de partenaires de confiance.

Comme pour des élections similaires, nous avons créé une équipe interfonctionnelle dotée d'une expertise linguistique pour guider notre travail autour du scrutin. Nos efforts se sont concentrés sur



les groupes vulnérables. Nous avons inscrit des dirigeantes à Facebook Protect pour les aider à protéger des comptes très ciblés, adopté des mesures de sécurité plus strictes (y compris la surveillance des menaces potentielles de piratage) et mené des opérations de ratissage de Facebook et d'Instagram pour lutter contre le harcèlement. Nous avons également mis au point des systèmes de détection et d'action pour contribuer à la protection des journalistes et des DDH.

Nous avons organisé une série de séances de formation pour les DDH et les femmes journalistes indépendantes. Nous avons également lutté contre la diffusion de fausses informations, à la fois sur la plateforme et en dehors, en supprimant le contenu destiné à empêcher le vote, notamment en supprimant de fausses allégations d'actes de violence hors contexte. Nos actions se sont fondées, en partie, sur les conseils des partenaires locaux.

Nous avons augmenté les investissements pour nos médias de vérification tiers au Kenya, comme l'AFP, Africa Check et PesaCheck. Nous avons également accordé des subventions à des rédactions locales au Kenya pour renforcer la vérification des informations en langue locale. Pour sensibiliser aux fausses informations et aux discours haineux, nous avons développé des campagnes de radio locales et sur la plateforme dans plusieurs langues et lancé deux services de soutien, Trending Event et CrowdTangle Live Display, pour aider les médias de vérification tiers à identifier les fausses informations et à agir en conséquence.





## Liberté de réunion et d'association

Le [droit à la liberté de réunion et d'association](#) est essentiel à la démocratie et constitue le fondement de nombreux autres droits garantis par le droit international, notamment le droit à la liberté d'expression et le droit de participation à des affaires publiques.

### Exemples de principaux risques inhérents aux droits humains identifiés lors de l'EGRP

Certains utilisateurs peuvent penser que le contenu des plateformes Meta ne leur permet pas de se rassembler librement sur les applications Meta.

Des limites de contenu abusives imposées par le gouvernement qui restreignent le droit de réunion et d'association

### Exemples de mesures prises par Meta pour remédier au risque

Les discours haineux, l'intimidation et d'autres formes de harcèlement peuvent avoir une incidence sur le droit de réunion et d'association des individus. Nous supprimons ce type de contenu lorsqu'il est signalé par la victime, car il empêche les individus de se sentir en sécurité et respectés sur les applications Meta et hors ligne. Meta dispose de politiques étendues contre le [harcèlement et l'intimidation](#) et [la violence et l'incitation](#), y compris le contenu ciblant des particuliers par le biais de pages, de groupes et d'événements non désirés.

Nous cherchons à supprimer le harcèlement de masse parce qu'il porte atteinte à de nombreux droits, y compris le droit de réunion. Il s'agit notamment du ciblage de personnes présentant un risque élevé de préjudice hors ligne, notamment les DDH, les mineurs, les victimes d'événements violents ou de tragédies, et les figures de l'opposition dans les pays à risque pendant les périodes électorales.

La fonctionnalité de chiffrement de bout en bout de WhatsApp protège le droit à la confidentialité et donc la capacité des utilisateurs à exercer leur droit de réunion et d'association.

Nous évaluons les demandes de retrait (DR) gouvernementales légales et valides sur la base des principes de la Global Network Initiative et notre travail est régulièrement évalué par un organisme indépendant.

Nous effectuons des contrôles opérationnels approfondis pour vérifier la validité des DR et fournissons des [informations sur les DR gouvernementales dans l'Espace modération](#).

(suite)

(suite)

Nous procédons à des révisions régulières des politiques de contenu, en sollicitant l'avis de diverses parties prenantes, au regard des normes internationales en matière de droits humains.

Meta respecte le droit des employés de nos fournisseurs d'organiser des syndicats. L'organisation syndicale par les employés du fournisseur ne remet pas en cause notre décision de travailler avec ce fournisseur ou de l'engager. Nous pensons qu'il est important que les entreprises avec lesquelles nous travaillons ne s'opposent pas ou n'entravent pas le droit de leurs employés à se syndiquer.





## Droit à la santé

Selon l'[Organisation mondiale de la santé](#), « jouir du meilleur état de santé possible est l'un des droits fondamentaux de tout être humain, quelles que soient son origine ethnique, sa religion, ses opinions politiques, sa condition économique ou sociale ».

Meta respecte le [droit à la santé](#) en améliorant l'accès à des informations crédibles en matière de santé, en permettant aux personnes ayant des problèmes de santé similaires d'entrer en contact les unes avec les autres et en leur donnant les moyens de prendre des décisions éclairées concernant leur santé et leur bien-être.

### Exemples de principaux risques inhérents aux droits humains identifiés lors de l'EGRP

Fausses informations médicales ou désinformation sur les plateformes Meta

Contenu allant à l'encontre des politiques et incitant ou encourageant les préjudices hors ligne

### Exemples de mesures prises par Meta pour remédier au risque

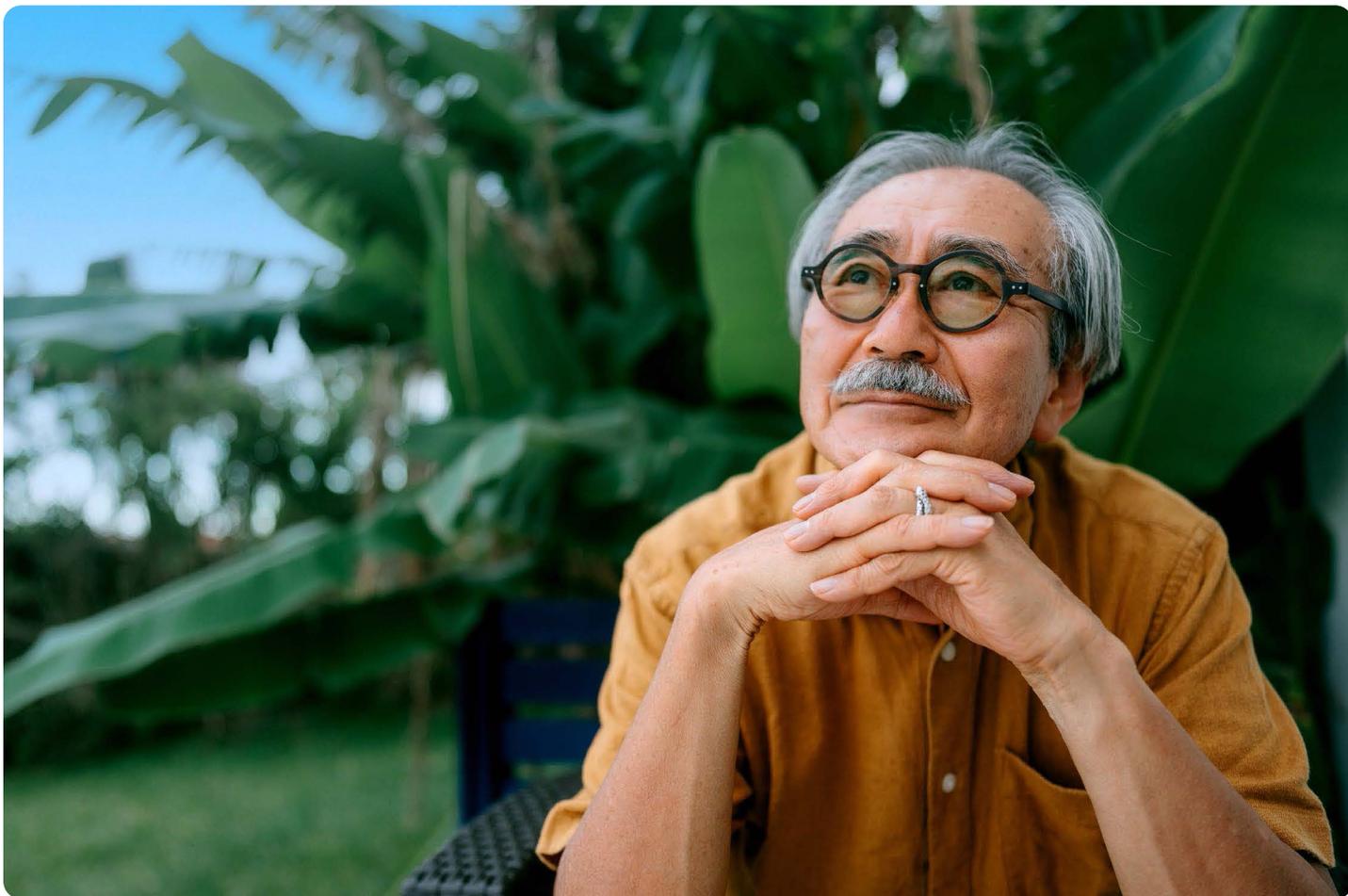
En 2020, Meta a adopté une [politique](#) visant à supprimer les fausses informations médicales si :

- il existe une urgence relative à la santé publique ;
- une autorité sanitaire majeure, au niveau mondial ou local, nous dit qu'une allégation particulière est fausse ;
- ces organisations ou autorités nous disent que l'allégation peut contribuer directement à un risque de blessure physique imminente.

Meta continue d'utiliser des outils basés sur l'IA pour renforcer le travail des médias de vérification dans la détection d'informations fausses et trompeuses en matière de santé, notamment en classant une image en fonction de son contenu visuel. Cela nous permet de faire respecter les interdictions de publicités et d'annonces commerciales pour certains produits médicaux en infraction.

Le centre de sécurité Meta reprend nos [ressources de prévention contre le suicide](#) et d'autres informations visant à aider les personnes ayant des pensées suicidaires.

Nos politiques interdisent, et Facebook et Instagram cherchent à prévenir, les contenus qui encouragent le suicide, l'automutilation, les troubles alimentaires ou l'intimidation et le harcèlement. Nous supprimons ce type de contenu dès que nous le trouvons et continuons à améliorer la technologie utilisée pour le détecter et le supprimer.



Nous nous efforçons d'évaluer les incidences sur les droits humains en faisant preuve de diligence raisonnable en matière de droits humains, comme le prévoit notre [Politique d'entreprise en matière de droits humains](#) et conformément aux [UNGP 17 et 21](#). Il s'agit notamment de prendre les mesures appropriées et de faire appliquer les conclusions, d'assurer le suivi de la mise en œuvre et de faire un rapport annuel des statistiques et des actions.

En 2022, nous avons publié les résultats des évaluations de l'impact sur les droits humains (EIDH) et des exercices de diligence raisonnable en matière de droits humains concernant [Israël et la Palestine](#), l'Inde et le [chiffrement de bout en bout](#). Nous avons poursuivi la mise en œuvre des recommandations issues de ces évaluations et des évaluations précédentes. Nous avons également fait progresser les efforts de diligence raisonnable concernant les incidences potentielles du métavers sur les droits humains.

### Philippines

Nous avons [publié les résultats](#) et notre réponse à l'EIDH indépendante que nous avons commandée aux Philippines en décembre 2021. L'EIDH a émis des recommandations

portant sur l'incitation à la violence, la surveillance, l'exploitation sexuelle en ligne, le trafic, les activités extrémistes et la responsabilité des entreprises.

Voici quelques-unes des mesures que nous avons prises par la suite :

- Nous avons mis en place de [nouveaux services et élaboré des politiques](#) plus solides en collaboration avec la Commission sur les élections, les organismes de surveillance des élections, les médias de vérification indépendants et les organisations de la société civile pendant les élections de 2022 aux Philippines. Ces efforts ont permis de renforcer notre capacité à supprimer les contenus et les réseaux en infraction, à fournir des informations crédibles sur les élections à un plus grand nombre de personnes, à promouvoir l'alphabétisation numérique et l'engagement civique, ainsi qu'à assurer une plus grande transparence des publicités politiques.
- Nous avons mis la [bibliothèque publicitaire Facebook](#) à la disposition des Philippines et exigé des annonceurs qu'ils suivent notre processus d'autorisation des publicités et qu'ils incluent des limitations de responsabilité « Financé par » sur les publicités relatives aux élections, à la politique et à certaines catégories de questions sociales.
- Nous avons procédé à plusieurs [mises à jour de nos Standards de la communauté](#), notamment en renforçant la [protection des personnalités publiques](#) telles que les journalistes et les défenseurs des droits humains (DDH).
- Nous avons supprimé d'autres types de contenus en infraction et nos politiques assurent maintenant une [protection plus solide contre le harcèlement fondé sur le genre](#). Nous avons mis en place de nouvelles politiques contre [le harcèlement de masse et les avis publiés en masse](#) et supprimé les efforts coordonnés de harcèlement de masse qui ciblent des personnes présentant un risque accru de préjudice hors ligne. En ce compris les attaques contre les dissidents.
- En partenariat avec l'[International Center for Journalists](#) et le Border Center for Journalists and Bloggers, nous avons lancé un [programme gratuit de sécurité numérique](#) pour aider les journalistes et les DDH à protéger leurs actifs numériques et à lutter contre le harcèlement en ligne. Le [Centre de sécurité Meta pour les journalistes](#) centralise toutes les ressources et outils disponibles sur notre application.

De plus amples informations sont disponibles dans la [mise à jour de la mise en œuvre des droits humains 2023](#).

## Israël et Palestine

Nous avons [publié le rapport](#) de [Business for Social Responsibility](#) (BSR) sur sa mission indépendante de diligence raisonnable en matière de droits humains (DRDH) portant sur les effets de nos politiques et processus en Israël et en Palestine au cours du mois de mai 2021. Cette décision faisait suite à la demande du Conseil de surveillance de procéder à un examen indépendant afin de déterminer si la modération du contenu de Meta en arabe et en hébreu, y compris son recours à l'automatisation, avait été appliquée sans biais.

Le rapport a mis en évidence des défis persistants à l'échelle du secteur concernant la modération du contenu dans les régions touchées par un conflit et la nécessité de protéger la liberté d'expression tout en réduisant le risque que les services en ligne

soient utilisés pour propager la haine ou inciter à la violence. Le rapport souligne également la complexité de la gestion des questions liées au conflit en raison de la dynamique sociale et historique, de divers événements violents à évolution rapide et des actions et activités des organisations terroristes. BSR n'a relevé aucune preuve de parti pris intentionnel de la part de Meta sur ces questions.

Au cours de la période couverte par le rapport, les actions suivantes ont été réalisées :

### Politique

- Révision de notre définition du terme « éloge » dans notre politique relative aux organismes et individus dangereux (OID) à la suite d'un vaste engagement des parties prenantes avec des universitaires, des acteurs de la société civile et d'autres experts de la lutte contre le terrorisme, des droits humains et de la liberté d'expression du monde entier.

### Mise en application

- À la suite d'une évaluation interne approfondie, nous avons déterminé que la création d'un routage spécialisé pour une variété de dialectes arabes à travers nos systèmes contribuerait à une plus grande précision dans notre modération de contenu arabe pour les contenus à haute gravité. Nous examinons actuellement la mise au point de mécanismes permettant d'acheminer efficacement le contenu par dialecte arabe afin d'améliorer la précision de l'examen du contenu arabe et de mieux prévenir les problèmes de mise en œuvre insuffisante ou excessive.
- Nous avons procédé à une analyse de la construction d'un classificateur arabe spécifique des dialectes pour la détection de n'importe quel contenu dans cette langue. Ces résultats et les contributions des linguistes et des experts en modèles linguistiques nous permettront d'ajouter à nos systèmes une fonctionnalité d'identification de la langue plus étendue, capable de reconnaître le contenu dans les différents dialectes arabes.
- Nous nous engageons à mettre à jour nos classificateurs afin d'améliorer régulièrement leur précision et leurs performances.
- Nous avons lancé un classificateur pour l'hébreu qui détecte de manière proactive les discours hostiles en hébreu et prend des mesures en conséquence.

Pour une mise à jour plus détaillée des progrès réalisés dans la mise en œuvre des recommandations, voir la [mise à jour de Meta : Mise à jour sur la diligence raisonnable pour Israël et la Palestine](#).

## Inde

Dans le [premier rapport annuel de Meta sur les droits humains](#), nous avons fourni un résumé détaillé de l'évaluation de l'impact sur les droits humains en Inde, élaborée après avoir consulté le guide de l'Institut danois des droits humains sur les EIDH concernant les rapports et les évaluations des activités numériques.

Sur la base des conseils de ce groupe et d'autres experts des droits humains, nous pensons que la forme de divulgation atténuée les risques pour la sécurité, conformément à l'UNGP 21(c). Dans ce rapport, nous détaillons les mesures que nous avons déjà prises ou que nous avons l'intention de prendre. Nous continuerons à étudier ses conclusions et à considérer ses recommandations comme base de

référence afin d'identifier et d'orienter les actions correspondantes. Nous suivons régulièrement notre application de ces mesures jusqu'en 2023, et même après.

Depuis lors, nous avons réalisé les progrès suivants :

- **Engagement des parties prenantes :** À la suite des observations concernant le potentiel de Meta à approfondir l'engagement de ses parties prenantes en Inde dans son ensemble, Meta affine son approche de l'engagement et de la sensibilisation de la société civile, y compris par le biais d'un programme spécifique d'engagement auprès de la société civile et des organisations communautaires indiennes.
- **Rapports de transparence :** Nos rapports sur l'Inde sont désormais plus détaillés. En plus de notre rapport sur les [demandes gouvernementales de données d'utilisateurs](#), nous continuons à fournir chaque mois un [rapport détaillé sur l'Inde](#), qui comprend des informations sur les mesures prises contre les contenus en infraction sur Facebook et Instagram, les plaintes reçues des utilisateurs en Inde, et les ordonnances reçues des comités d'appel des plaintes récemment constitués en vertu des règles sur les technologies de l'information. Nous continuons d'affiner nos systèmes de collecte de statistiques sur les contenus supprimés à la suite de demandes gouvernementales et de rapports publics, comme en témoigne notre rapport au Conseil de surveillance en réponse à ses recommandations ([T4 2022](#) et [T1 2023](#)).
- **Partenariats étendus :** Nous avons bien progressé dans la réalisation de notre engagement à élargir les partenariats avec la société civile indienne, notamment dans les domaines de l'alphabétisation numérique, de la sécurité des femmes et des enfants et de la lutte contre l'extrémisme. Le centre de sécurité pour les femmes a été relancé sous le nom de [Centre de sécurité Meta](#) en 2023 et a été traduit en assamais, bengali, gujarati, hindi, marathi, panjabi, tamoul, télougou, kannada et malayalam. Nous avons élargi la participation des organisations indiennes à notre initiative pour la résilience et nous nous préparons à lancer notre programme de réorientation des recherches, qui redirige les utilisateurs utilisant des termes de recherche liés à l'extrémisme et à la violence vers des groupes de ressources, d'éducation et de sensibilisation.
- **Modération du contenu :** En réponse aux observations relatives aux politiques visant à prévenir davantage les contenus discriminatoires et haineux à l'encontre des groupes vulnérables, Meta a développé un prototype de test pour l'application du Plan d'action de Rabat, à utiliser lors de l'évaluation d'un discours haineux potentiel. Nous avons adapté les principes du Plan de Rabat pour en tirer des outils politiques actifs en matière de contenu, notamment des cadres basés sur l'escalation pour évaluer les discours attaquant des concepts (par opposition à des personnes) et les contenus impliquant des menaces de recours à la force de la part d'un État. Nous avons augmenté le nombre de groupes visés dans toutes les catégories d'organisations et de personnes désignées en Inde afin de nous assurer que notre plateforme n'est pas utilisée de manière abusive par des personnes malveillantes.

En outre, l'expérience de Meta en matière de protection de l'intégrité des élections lui permet de [se préparer](#) aux prochaines élections législatives et régionales.

Ces efforts comprennent l'activation de notre centre d'opérations électorales, la garantie que les examinateurs de contenu maîtrisent 20 langues indiennes, la croissance de notre groupe de partenaires de vérification indépendants, de sept à 11 (maîtrisant désormais 15 langues), l'application de la transparence des publicités politiques et le travail rapproché avec les autorités électorales et la société civile.

## 07 Engagement des parties prenantes : renforcement de notre travail par des parties externes



L'engagement avec les parties prenantes externes dans le monde entier nous aide à assumer nos responsabilités en matière de droits humains et constitue un levier important pour la responsabilité et la transparence. Nous reconnaissons en particulier l'importance d'un engagement significatif avec les parties prenantes des communautés sous-représentées et marginalisées. L'engagement des parties prenantes fait partie intégrante de notre [Politique d'entreprise en matière de droits humains](#) et de nos activités connexes. Nous nous efforçons d'écouter et de demander conseil aux experts et militants des droits humains, aux universitaires et à d'autres personnes, et de les informer des développements en cours au sein de Meta. Leurs avis sont utiles dans des domaines tels que l'élaboration de nos politiques de contenu et leur mise en application, ainsi que notre approche en matière de classement algorithmique et de recommandations.

En 2022, Meta a renforcé et élargi l'engagement constant de ses parties prenantes externes auprès des groupes suivants.

Parties prenantes	Efforts de consultation et de collaboration
Organisations de la société civile	Soutien des examens des Standards de la communauté et efforts en matière d'intégrité dans des domaines clés tels que les élections
<u>Défenseurs des droits humains</u>	Mesures visant à protéger la sécurité des comptes et à réduire la propagation de fausses informations et de contenu violent
Groupes vulnérables ou marginalisés	Mécanismes d'avis pour l'élaboration de la politique de contenu et les services de sécurité des bâtiments
Organisations internationales	Partage des perspectives sur les politiques en matière de droits humains numériques et amélioration de la communication dans les pays hautement prioritaires
<u>Partenaires de confiance</u>	Pour ce programme mondial : améliorations de l'identification et du signalement des contenus en infraction, contribution des connaissances locales à l'application de notre politique et services de sécurité pour les bâtiments
Investisseurs et annonceurs	Consultations et briefings relatifs au travail de Meta sur les droits humains, y compris la diligence raisonnable
Utilisateurs	Contribution aux questions litigieuses et à la prise de décision décentralisée

Voici quelques exemples de notre travail :

## Contribuer à la création d'espaces en ligne sûrs et soutenir l'intégrité des élections au Nigeria

Dans la perspective des élections législatives de 2023 au Nigeria, nous nous sommes engagés auprès de nombreuses parties prenantes. Ces engagements ont porté sur nos [politiques électorales](#) et sur les efforts déployés pour préserver l'intégrité et l'exactitude des informations figurant sur nos applications. Nous avons recueilli des informations sur les tendances des contenus violents, ce qui nous a permis d'améliorer leur détection et de préserver l'intégrité des informations.

**Nos actions :**

- Nous avons organisé des formations à la sécurité numérique pour les groupes vulnérables, notamment les femmes politiques et les personnalités publiques féminines, les défenseurs des droits humains (DDH) et les militants.
- Nous avons organisé une [table ronde sur la politique en matière d'organismes et d'individus dangereux \(OID\)](#) au Nigeria ainsi que deux tables rondes à Londres en présence de la diaspora nigériane.
- Nous avons collaboré avec nos partenaires de confiance dans le cadre d'ateliers et de consultations afin de les aider à identifier et à faire remonter les contenus tels que les discours haineux, les fausses informations, la violence et l'incitation, l'ingérence dans le processus électoral, le harcèlement et l'intimidation.

## Engager des communautés marginalisées et sous-représentées dans la politique de contenu

Nous avons progressé dans la réalisation de notre objectif consistant à faire entendre la voix des communautés marginalisées et sous-représentées dans l'élaboration des politiques de contenu. Nous avons élaboré un [Cadre de l'inclusivité](#) afin de garantir la prise en compte de nos diverses parties prenantes dans l'élaboration des politiques et le respect des Standards de la communauté.

**Nos actions :**

- **Demandes d'avis concernant les Standards de la communauté :** Des responsables de communautés religieuses, indigènes et ethniques, sous-représentées sur nos applications, ont participé à huit tables rondes en Europe, au Moyen-Orient et en Afrique pour discuter de l'efficacité de nos Standards de la communauté à créer un environnement inclusif dans les services et applications de Meta. Parmi ces tables rondes figuraient celle organisée au Kenya après les élections et celle consacrée à l'analyse de la crise migratoire en Pologne.
- **Renforcement de nos liens avec la communauté LGBTQIA+ :** Nous avons intensifié nos engagements vis-à-vis de groupes et de représentants en Europe, au Moyen-Orient et en Afrique en ce qui concerne les effets des politiques de contenu de Meta sur les utilisateurs, notamment en matière de discours haineux, d'intimidation et de harcèlement. Il s'agissait de DDH LGBTQIA+, d'organisations de la société civile, d'universitaires et de militants du Royaume-Uni, d'Espagne, d'Allemagne, du Kenya, du Ghana, de Namibie, d'Ouganda, d'Égypte, du Soudan, du Zimbabwe, d'Afrique du Sud, de Gambie et de Tanzanie.
- **Extension de notre portée à la société civile des Caraïbes :** Nous avons écouté les préoccupations concernant la diffusion de contenu potentiellement contraire aux politiques et avons mené des discussions sur les droits numériques, la sûreté et la sécurité en ligne, la sécurité des femmes et les droits des personnes LGBTQIA+. Des experts et des groupes de la société civile d'Haïti, de Trinité-et-Tobago, de Jamaïque, du Panama, de la République dominicaine et des îles Caïmans ont participé à la première table ronde de ce genre.
- **Renforcement de l'engagement auprès des communautés religieuses :** Nous avons étendu nos engagements politiques de manière à inclure les femmes musulmanes au Royaume-Uni, les communautés musulmanes du Sahel et la communauté sikh au Royaume-Uni, ainsi que la communauté bahá'íe dont les représentants viennent du Canada, des États-Unis, du Chili, du Brésil, d'Allemagne, d'Espagne, de Suisse, d'Afrique du Sud, des Émirats arabes unis, d'Égypte et de Tunisie.

- **Réalisation d'un guide sur l'engagement en faveur des personnes en situation de handicap :**  
En collaboration avec des organisations de défense des droits des personnes en situation de handicap qui travaillent à la croisée des politiques en matière de handicap et des technologies, nous avons élaboré un guide à l'intention de nos équipes chargées des politiques de contenu.



#### **Perspectives de communautés indigènes en matière de drogues à usage non médical**

Nous nous sommes engagés avec des parties prenantes en Amérique du Nord, au Gabon, au Cameroun et au Zimbabwe afin de faire entendre leur voix dans l'élaboration de politiques sur les drogues à usage non médical dans des contextes religieux ou traditionnels. Nous avons cherché à trouver un équilibre entre la liberté d'expression concernant les substances traditionnelles et religieuses et les risques de sécurité liés à la promotion potentielle de drogues à usage non médical, telles que l'ayahuasca. Parmi les personnes consultées, un large panel d'universitaires et d'organisations de la société civile, dont des spécialistes de la santé, de la médecine, de l'anthropologie, de la psychologie clinique, des politiques en matière de drogues, de l'ethnographie, de l'ethno-ornithologie, de la religion et de la recherche sur les drogues psychédéliques, ainsi que des experts juridiques, des régulateurs, des défenseurs de la liberté d'expression, des professionnels de la santé, des associations de guérisseurs traditionnels/religieux, des chefs de communautés et des groupes autochtones. Les contributions que nous avons reçues ont été utilisées pour réviser notre approche de la discussion positive sur les drogues à usage non médical dans des contextes traditionnels ou religieux.

## Engagement auprès d'organisations internationales

La croissance rapide et les effets de la technologie ont incité de nombreuses organisations internationales à élaborer des politiques relatives aux droits humains dans le monde numérique. Nous collaborons régulièrement avec de nombreux organismes gouvernementaux internationaux, notamment le [Haut Commissariat aux droits de l'homme](#), l'[UNESCO](#), le [Bureau de la prévention du génocide et de la responsabilité de protéger](#), l'[UNICEF](#), le [Forum économique mondial](#), le [Forum sur la gouvernance de l'Internet](#), le [Haut Commissariat des Nations unies pour les réfugiés](#), l'[OCDE](#), le [Fondation des Nations unies](#) et diverses organisations régionales. Nous donnons un aperçu de nos Standards de la Communauté, de la manière dont ils sont appliqués et de la façon dont nous cherchons à protéger les DDH. En 2022, nous avons intensifié nos engagements avec les Nations unies au niveau national, notamment en Éthiopie, au Kenya, en Ukraine, en Haïti et en Afghanistan.

#### **Voici quelques exemples de nos actions :**

- L'UNICEF et Meta ont travaillé en étroite collaboration pour s'assurer que des informations essentielles étaient disponibles pour les [Ukrainiens](#), et ont mis à disposition des crédits publicitaires pour soutenir la collecte de fonds et les messages de sensibilisation sur les principaux marchés mondiaux, en Ukraine et dans les pays voisins tels que la Pologne et la Roumanie.
- Nous nous sommes entretenus régulièrement avec des fonctionnaires du Haut-Commissariat des Nations unies aux droits de l'homme afin d'aborder des thèmes internationaux et nationaux et de participer activement au [projet B-Tech](#).

- Nous avons dispensé une formation sur les Standards de la communauté et les mécanismes de signalement aux équipes nationales des Nations unies en Éthiopie, au Nigeria et en Haïti.
- Nous nous sommes associés au [Bureau de l'Envoyé du Secrétaire général pour les technologies](#) au sujet du Pacte numérique mondial et de la nécessité de soutenir une approche harmonisée à l'échelle mondiale de la gouvernance de l'Internet, qui favorise la libre circulation des données. Cette démarche est essentielle pour protéger les droits humains en ligne, y compris les droits à la confidentialité et à la liberté d'expression.
- Nous avons organisé des réunions avec de hauts fonctionnaires des Nations unies et des diplomates autour de la semaine de haut niveau de l'Assemblée générale des Nations unies pour discuter des droits humains et des thèmes connexes.
- Nous nous sommes engagés avec un certain nombre de rapporteurs spéciaux du [Conseil des droits de l'homme des Nations unies](#), notamment ceux qui se concentrent sur les DDH, la liberté d'expression, l'Afghanistan et le Myanmar, entre autres, afin de partager des informations et de mieux comprendre leurs préoccupations en ce qui concerne le rôle des services de Meta.
- Nous avons continué à répondre aux demandes de données du [mécanisme d'enquête indépendant des Nations unies pour le Myanmar](#) afin de contribuer aux efforts de responsabilisation.

## Déploiement de l'identification et escalation des thèmes avec les partenaires de confiance

Lorsque nous nous trouvons face aux défis principaux en matière de droits humains, nous donnons la priorité aux partenariats dans certaines régions du monde.

Notre réseau de [partenaires de confiance](#), composé de plus de 400 organisations non gouvernementales, à but non lucratif, nationales et internationales de 113 pays, signale le contenu, les comptes et les comportements que nous pouvons alors examiner en contexte. Ces partenaires de confiance sont des alliés importants dans nos efforts pour comprendre les effets de nos services au niveau local. Ils nous aident à nous tenir au courant des tendances émergentes, peuvent nous signaler des contenus en infraction que nous pourrions manquer, et nous permettent d'améliorer nos politiques et leur application dans des domaines tels que les fausses informations et les préjudices. Leurs rapports fournissent des informations précieuses qui contribuent à l'élaboration des politiques et à la sécurité de nos utilisateurs.

### Nos actions :

- En 2022, nous avons étendu notre réseau de partenaires de confiance à 36 pays, dont le Yémen, l'Ukraine, l'Éthiopie, le Mali, le Kenya, le Népal, Haïti et le Cambodge.
- Nos partenaires nous ont alertés sur des évolutions et des risques que nous n'aurions peut-être pas remarqués autrement. Par exemple :
  - Nos partenaires de confiance ont signalé un pic de sextorsion sur nos applications au Yémen, ce qui a conduit à une enquête et finalement à la suppression d'un grand nombre d'utilisateurs de Facebook, de comptes business et de comptes Instagram.
  - Au Cambodge, nos partenaires de confiance ont repéré des publications Facebook visant à recruter des personnes pour des centres d'appel en Asie du Sud-Est et contenant des indicateurs clés de traite d'êtres humains et d'exploitation par le travail. Plus de 30 publications ont été retirées.

- Nos partenaires de confiance au Pakistan ont signalé de fausses informations concernant la loi sur les personnes transgenres et ont ciblé des contenus qui mettaient des individus en danger. Le contenu a été supprimé et nous avons pris des mesures pour nous assurer que nous détecterions d'autres cas de contenus similaires allant à l'encontre de la politique.

## Innovation par le biais de forums communautaires

Meta continue d'explorer les moyens de faire entendre d'autres voix que celles de la Silicon Valley dans le processus décisionnel. Nous avons testé des forums communautaires basés sur la démocratie délibérative, un mode de prise de décision utilisé par les gouvernements et les organisations dans le monde entier. Cela permet de réunir des groupes représentatifs de personnes afin de tenir des débats structurés et de prendre des décisions sur des thèmes particulièrement complexes ou litigieux.

### Nos mesures :

En 2022, nous avons organisé un modèle de délibération de groupe basé sur le [sondage délibératif](#) (Deliberative Polling®), une méthode créée au Stanford Deliberative Democracy Lab. Nous avons collaboré avec ce laboratoire ainsi qu'avec l'équipe Behavioural Insights, un groupe d'analyse de données basé au Royaume-Uni, afin de garantir l'impartialité et l'indépendance du processus.

Ce forum a été le plus grand forum de sondage délibératif à ce jour, avec plus de 6300 personnes provenant de 32 pays. Il s'est concentré sur la lutte contre l'intimidation et le harcèlement dans les espaces sociaux virtuels.

Les participants ont reçu des documents pédagogiques sur le sujet, approuvés par des experts indépendants, ont délibéré entre eux lors de multiples séances en petits groupes et ont eu l'occasion de poser des questions à des experts indépendants du monde entier. Parmi eux figuraient des spécialistes de la cybersécurité, des avocats spécialisés dans les droits humains, des militants de la lutte contre l'intimidation et le harcèlement, des experts de la réalité virtuelle immersive, d'anciens responsables brand safety pour Twitter, TikTok et Second Life, ainsi que des membres des comités consultatifs des Nations unies et du Forum économique mondial et des lauréats du prix international de la liberté de la presse.

À l'issue de ces séances, les participants ont été sondés sur leur soutien aux propositions dont ils avaient débattu pour contribuer à l'élaboration de services et d'applications pour les expériences virtuelles. Cet événement s'inscrit dans le cadre de nos efforts visant à créer des programmes qui favorisent une prise de décision plus inclusive et permettent à un plus grand nombre de personnes de s'exprimer sur le développement de nos applications et technologies.

Dans l'ensemble, les participants ont fortement soutenu l'idée que les créateurs et les propriétaires de plateformes sont responsables de la surveillance des univers qu'ils créent, qu'il s'agisse d'espaces publics ou d'espaces réservés aux membres.

Les participants ont soutenu l'utilisation d'outils tels que la reconnaissance automatique de la parole, la capture vidéo et les modérateurs visibles et invisibles. Le soutien à ces outils a été plus fort pour les espaces publics que pour les espaces réservés aux membres.

Pour davantage de détails, consultez le [rapport des résultats](#) de Stanford.



La langue est l'un des facteurs les plus importants qui influencent les effets de Meta sur les droits humains. L'importance de la langue a été soulignée lors des consultations des parties prenantes dans le cadre de [l'évaluation globale des risques principaux en matière de droits humains](#). En 2022, nous avons augmenté notre capacité à prendre en charge davantage de langues et avons permis à un plus grand nombre de personnes d'exercer leurs droits en ligne, y compris [l'accès aux voies de recours](#).

Il est important de comprendre le contexte local pour appréhender les risques potentiels en matière de droits humains. C'est pourquoi chez Meta, nous disposons à la fois de traductions humaines et de traductions automatiques. La prise en charge des langues comprend la détection de la langue et les classificateurs linguistiques, la modération du contenu et la traduction des interfaces d'utilisation, ainsi que d'autres contenus tels que les Pages d'aide ou les [Standards de la communauté](#).

Nous travaillons en permanence à la prise en charge de nouvelles langues. Entre 2021 et fin 2022, notre équipe a étendu la prise en charge professionnelle de l'interface d'utilisation de Facebook à 30 langues supplémentaires. La liste comprend des langues sous-représentées en ligne et pour lesquelles l'approvisionnement linguistique et la traduction constituent un défi, comme le javanais, le shona, le kirundi, le cinghalais, le népalais et l'assamais.

En outre, à la fin de l'année 2022, nos Standards de la communauté étaient disponibles en 77 langues.



### No Language Left Behind

Le projet de Meta [No Language Left Behind](#) (NLLB) est un projet d'IA pionnier et unique qui repose sur des modèles open source capables de fournir des traductions évaluées et de haute qualité directement entre 200 langues, y compris des langues rares comme l'asturien et le luganda. Ce projet vise à permettre à toute personne d'accéder à du contenu web dans sa langue maternelle et de le partager, c'est-à-dire de communiquer avec n'importe qui, n'importe où, peu importe les préférences linguistiques de chacun.

La technologie sur laquelle repose le modèle NLLB-200, [désormais disponible](#) via l'outil de traduction de contenu de la Fondation Wikimedia, aide les rédacteurs de Wikipédia à traduire des informations dans leur langue maternelle et dans les langues de leur choix. Cela permet de rendre les connaissances disponibles dans un plus grand nombre de langues pour les lecteurs de Wikipédia dans le monde entier.





Les droits humains sont souvent particulièrement mis à mal en temps de crise. Dans ces moments-là, Meta évalue les risques de danger imminent sur les plateformes et en dehors et cherche à y réagir en mettant en place des politiques spécifiques, des services et des actions opérationnelles en matière de droits.

En 2022, Meta a renforcé sa réponse organisationnelle pour soutenir les populations pendant les périodes de tension élevée dans n'importe quel pays, y compris pendant les conflits, les élections ou les désordres civils. Cette réponse rassemble une équipe interfonctionnelle qui se concentre sur la priorisation du soutien et des ressources afin d'anticiper les moments d'escalade des tensions partout dans le monde et d'y réagir de manière efficace. Les travaux s'articulent autour des élections mondiales et des pays présentant une réaction active, ainsi que de la capacité internationale à anticiper les futurs événements critiques et à s'y préparer. Un ensemble de standards en matière d'intégrité est développé et affiné pour permettre aux utilisateurs de signaler les contenus en infraction et est représenté par une expertise nuancée dans la langue du pays.

### Protocole de politique de crise

En août 2022, sur la base d'une [recommandation](#) du Conseil de surveillance, nous avons lancé notre [protocole interne de politique de crise](#) (PPC) afin de veiller à ce que notre réponse politique face aux crises soit réfléchie et calibrée.

Le PPC est un cadre dynamique qui nous permet d'identifier les situations de crise émergentes et d'évaluer leur gravité relative. Il guide notre utilisation rapide de leviers politiques ciblés ou

exceptionnels pour atténuer les dommages potentiels en fonction des risques observés et des interventions de crise antérieures. Par conséquent, le PPC nous aide à trouver un équilibre entre une réponse globale cohérente aux crises et la capacité de s'adapter à des conditions locales qui évoluent rapidement. Pour élaborer le PPC, nous avons consulté plus de 50 experts externes internationaux en matière de droits humains, de prévention des conflits, de discours haineux, de réponse humanitaire et de sécurité nationale. Nous avons également mené une recherche originale.

Ce protocole a été conçu pour compléter et renforcer les efforts de réponse à la crise déployés à l'échelle de l'entreprise. Le PPC veille à ce que Meta utilise une approche politique holistique et interdisciplinaire pour améliorer en permanence sa réponse aux situations de crise au sein des équipes chargées des services, des politiques et des opérations.

Voici quelques exemples notables de nos activités de réponse aux crises :

## Russie et Ukraine

L'invasion de l'Ukraine par la Russie en février 2022 apporte son lot de défis en matière de droits humains pour Meta. La propagande et la désinformation visant à influencer l'opinion publique ont proliféré en ligne, tout comme la surveillance, les perturbations des services Internet et la promotion de contenus en infraction. Nous devons intensifier nos efforts de modération du contenu tout en respectant la liberté d'expression afin d'aider les citoyens ukrainiens à coordonner l'aide humanitaire et à exprimer leur résistance aux forces d'invasion.

En Russie, Facebook et Instagram ont été bloqués et Meta Platforms, Inc. a été ajouté à la liste des organisations extrémistes par les autorités. WhatsApp était toutefois toujours disponible.

### Diligence raisonnable :

Depuis le début de la guerre, notre diligence raisonnable en matière de droits humains comprend l'interprétation du droit international humanitaire et des standards en matière de droits humains afin d'aider nos équipes à relever les défis. Ces efforts ont notamment consisté à examiner des vidéos identifiant des prisonniers de guerre ainsi que des vidéos véhiculant des discours haineux, de fausses informations et de la propagande contrôlée par l'État russe et niant l'invasion de l'Ukraine par la Russie.

Nous nous sommes entretenus avec des défenseurs des droits humains (DDH), des militants, des journalistes, des groupes de la société civile et des groupes de défense des droits humains en Russie et en Ukraine afin de comprendre les problèmes auxquels ils sont confrontés. Nous avons également collaboré avec les Nations unies et d'autres institutions internationales afin d'approfondir notre compréhension des questions en jeu et de mieux comprendre les meilleures pratiques en matière de modération de contenu en temps de guerre.

### Nos actions :

Meta a pris un certain nombre de mesures pour relever ces défis :

- **Mise en place d'un centre d'opérations spéciales** composé d'experts de tous les services de l'entreprise, 24 heures sur 24 et 7 jours sur 7, y compris en russe et en ukrainien, afin de suivre l'évolution rapide de la guerre et d'y réagir en temps réel.
- **Renforcement de nos efforts de lutte contre le contenu en infraction** comme la violence et l'incitation, les discours haineux, les comportements non authentiques coordonnés, les fausses informations et la désinformation. Nous avons notamment développé notre capacité de vérification des informations par des tiers en russe et en ukrainien dans l'ensemble de la région et apporté un soutien financier supplémentaire aux partenaires ukrainiens de vérification des informations. Nous avons également intégré cinq nouveaux partenaires de confiance, en organisant avec eux des séances d'écoute afin d'améliorer notre compréhension des risques liés au contenu et de mettre au point des mécanismes de signalement appropriés.

- **Déévaluation globale du contenu des pages Facebook et des comptes Instagram des médias contrôlés par l'État russe** afin de le rendre plus difficile à trouver sur nos applications. Nous avons également commencé à [marquer et à dévaloriser](#) les publications et les stories qui contiennent des liens vers des sites de médias contrôlés par l'État russe sur Facebook et Instagram. Dans l'UE et en Ukraine, nous avons bloqué tous les médias d'État russes à la demande expresse de ces gouvernements.
- **Déploiement de fonctionnalités de sécurité** destinées aux défenseurs des droits humains et à d'autres groupes vulnérables, notamment la possibilité de verrouiller son profil Facebook, de supprimer la possibilité de consulter et de rechercher dans les listes d'amis, et d'ajouter des outils sur Messenger pour aider à protéger les utilisateurs contre le ciblage.
- **Contribuer à garantir que les comptes des DDH** et des médias indépendants bénéficient d'une protection supplémentaire contre le harcèlement ou l'usurpation d'identité et que leurs voix ne soient pas indûment restreintes grâce à notre programme de vérification croisée. Ce travail est en cours.
- **Aide à la prévision des flux de réfugiés** grâce à notre équipe [Data for Good](#) qui partage des ensembles de données préservant la confidentialité avec les partenaires de confiance.
- **Aide aux Ukrainiens** et d'autres habitants de la région à trouver des informations fiables auprès des agences locales des Nations unies et des sociétés de la Croix-Rouge en mettant temporairement en place un service d'aide communautaire sous la forme d'une ressource centrale sur Facebook.

## Iran

En septembre 2022, Mahsa Amini, 22 ans, est arrêtée par la police de la moralité iranienne pour avoir porté son hijab de manière « inappropriée ». Elle décède pendant sa garde à vue, déclenchant de vastes manifestations en Iran et dans le monde entier, qui se poursuivent encore aujourd'hui. En réponse à ces manifestations, les autorités iraniennes ont sévèrement réprimé la liberté d'expression et la liberté de réunion, et ont limité l'utilisation d'Internet et d'applications telles qu'Instagram et WhatsApp, très répandues parmi les Iraniens.

Meta a réagi rapidement et a pris un certain nombre de mesures pour aider les utilisateurs à rester connectés et en sécurité sur nos applications. Nous avons notamment mis en place une équipe dédiée parlant le farsi et chargée de traiter les problèmes liés aux manifestations. Cette équipe veillait à ce que nous appliquions correctement nos politiques et à ce que nos utilisateurs soient en sécurité. Nous surveillons toujours la situation de près.

### Nos mesures : aider les manifestants à se faire entendre

Instagram a été massivement utilisé par les Iraniens pour faire la lumière sur les manifestations et la réalité des événements sur le terrain. Les hashtags liés aux manifestations en Iran ont été utilisés sur Instagram plus de 160 millions de fois au cours des cinq premiers mois des manifestations. Les manifestants ont partagé leurs images Instagram avec les médias internationaux, dont beaucoup n'ont pas pu faire de reportages en Iran.

- **Aider les gens à entrer en contact** : Nous avons lancé Instagram Lite en Iran en 2022 pour aider les utilisateurs à accéder à Instagram lorsque la bande passante est réduite. En outre, WhatsApp a [lancé](#) un serveur proxy qui permet aux utilisateurs de se connecter lorsque leur connexion Internet est interrompue ou bloquée, ce qui est particulièrement important pour les communautés en Iran.
- **Protéger les militants et les DDH** : Lors de manifestations, les comptes de réseaux sociaux des militants et des journalistes sont souvent la cible d'intimidations, de harcèlement, d'usurpations d'identité et de tentatives de piratage. Nous avons collaboré avec des militants iraniens pour protéger leurs comptes et les aider à sensibiliser le public aux atteintes aux droits humains.

## 10 Notre approche des droits humains et le développement responsable de nos produits



### Métavers

En 2021, Meta a créé le [Fonds de recherche et de programmes XR](#), un investissement de 50 millions de dollars sur deux ans dans des programmes et des recherches externes avec des partenaires industriels, des groupes de défense des droits humains et des droits civils, des gouvernements, des organisations à but non lucratif et des institutions universitaires. Ces investissements se concentrent sur les domaines que Meta doit maîtriser pour tirer parti des avantages potentiels du métavers. Toutes les propositions doivent s'aligner sur le travail de Meta dans un ou plusieurs des quatre domaines politiques :

- Perspectives économiques : comment donner plus de choix aux citoyens et maintenir une économie numérique prospère
- Confidentialité : comment intégrer la transparence et les contrôles à nos produits ?
- Sécurité et intégrité : comment contribuer à la sécurité des utilisateurs sur nos plateformes et leur donner les moyens d'agir ou d'obtenir de l'aide s'ils voient ou vivent quelque chose qui les met mal à l'aise ?
- Équité et inclusion : comment s'assurer que ces technologies sont conçues de manière inclusive et qu'elles sont accessibles ?

Nos partenaires dans cette initiative sont des universités et des instituts de recherche, des organisations internationales, des groupes de défense des droits civils et des ONG issues de communautés sous-représentées de toutes les régions du monde ou collaborant avec elles.

## Intelligence artificielle

Nous croyons que l'intelligence artificielle (IA), développée et déployée de manière responsable, peut être un outil puissant pour faire progresser les droits humains. L'IA contribue de manière importante à la résolution de nombreuses incidences négatives potentielles sur les droits humains dont il est question dans le présent rapport. Par exemple, les capacités de l'IA ont permis de renforcer la détection et l'examen des contenus problématiques dans un nombre croissant de langues, d'améliorer l'identification des contacts et des interactions potentiellement indésirables, d'améliorer l'accessibilité grâce à la génération automatique de sous-titres et de renforcer toute une série de nos [efforts Data for Good](#).

Parallèlement, le développement et l'utilisation de l'IA présentent des risques inédits pour les droits humains. Entre autres risques potentiels, les modèles d'IA peuvent présenter des biais problématiques ou des incidences discriminatoires et générer des contenus problématiques ou s'engager dans une application excessive de la réglementation qui porte atteinte à la liberté d'expression.

Nous nous [engageons à aborder ces thèmes](#) de manière ouverte et collaborative. Par exemple, les données démographiques sont essentielles pour évaluer si un service, une application ou un processus [traite tous les groupes de manière équitable](#). La collecte de ces données est toutefois délicate et soulève d'importantes questions quant à la manière de protéger la confidentialité des utilisateurs.

Nous plaçons la responsabilité et les droits humains au cœur de notre approche. L'IA est ancrée dans notre [Politique d'entreprise en matière de droits humains](#). Nous reconnaissons également l'importance des [Principes de l'OCDE en matière d'intelligence artificielle](#), qui font explicitement référence aux droits humains et sont soutenus et adoptés en grande majorité par les pays du G20. Nous espérons favoriser un dialogue plus collaboratif et transparent entre les disciplines et les publics sur la voie à suivre pour ces thèmes cruciaux.

Nous avons créé une équipe interdisciplinaire dédiée à l'[IA responsable](#) (Responsible AI, RAI) qui veille à ce que l'IA chez Meta profite aux personnes et à la société. L'équipe comprend des experts en éthique, en sciences sociales et politiques, en politique et en droits, des chercheurs et des ingénieurs. L'équipe RAI travaille en étroite collaboration avec l'équipe des droits humains pour s'assurer que nos systèmes d'apprentissage automatique sont conçus et utilisés de manière responsable.

Nous avons également été les premiers à adopter une approche des [cartes de systèmes d'IA](#) visant à fournir une transparence normalisée sur le fonctionnement des systèmes d'IA, et pas seulement des modèles individuels, en commençant par un projet pilote axé sur le [classement des flux Instagram](#).



### Cinq piliers pour une IA responsable

L'IA responsable chez Meta repose sur cinq piliers :

- Confidentialité et sécurité
- Équité et inclusion
- Robustesse et sécurité
- Transparence et contrôle
- Responsabilité et gouvernance

Nous avons progressé dans la résolution de bon nombre de ces problèmes, mais il reste encore beaucoup à faire. Voici les domaines auxquels nous avons œuvré en 2022 :

- **Étendre la couverture linguistique**
  - Notre initiative open source [No Language Left Behind](#) exploite l'IA pour fournir des traductions de haute qualité directement entre 200 langues.
  - Notre modèle de reconnaissance [Massively Multilingual Speech](#) (MMS) prend en charge la conversion de la parole en texte et du texte en parole pour 1107 langues et permet l'identification de plus de 4000 langues. Certaines des langues prises en charge, comme le tatuyo, ne comptent que quelques centaines de locuteurs et, pour la plupart d'entre elles, aucune autre technologie de reconnaissance vocale n'existe. Nous partageons les modèles et le code publiquement.
  - [Wikipédia collabore avec Meta AI](#) pour vérifier les informations contenues dans les citations à l'aide de [Sphere](#), le programme IA open source de Meta.
- **Traiter les biais**
  - Dans le cadre de nos efforts visant à garantir que les systèmes d'IA traitent tout le monde de manière équitable, nous avons introduit et [rendu open source plusieurs ensembles de données](#) et modèles permettant de remédier à cette lacune dans le respect de la confidentialité.
  - Nous avons entraîné un [modèle d'IA pour réduire les biais démographiques](#) dans les textes, que l'on peut appeler perturbateur démographique de texte, capable de supprimer les associations stéréotypées présentes dans les données de programmation en langage naturel (PLN).
  - Nous avons introduit et rendu open source deux nouveaux ensembles de données pour aider à mesurer l'équité et à atténuer les biais potentiels dans les modèles de PLN. Ces ensembles de données comprennent une représentation plus complète des différentes dimensions démographiques pour mesurer l'équité dans ces modèles, y compris des termes pour l'identité de genre, l'âge, l'origine ethnique et le handicap.
- **Améliorer la compréhension des associations de contenus problématiques**
  - Nous avons constitué une équipe pluridisciplinaire, composée de membres de nos équipes chargées des droits civils, des droits humains, de l'ingénierie, de la RAI, de la recherche en IA, de la politique et des produits, afin de mieux comprendre les associations de contenus problématiques dans plusieurs de nos systèmes de bout en bout. Nous avons travaillé à la mise en œuvre de mesures techniques d'atténuation afin de réduire le risque que les associations de contenus problématiques se produisent sur nos applications qui utilisent des modèles d'IA.

- **Évaluer l'équité**

- La création et la distribution d'ensembles de données plus diversifiés constituent l'un de nos moyens d'aborder la question de l'équité de l'IA. Dans le cadre de notre projet Casual Conversations v2, nous avons également mis au point une importante base de données axée sur le consentement afin de mesurer la partialité et la robustesse des algorithmes.
- Meta AI Research a commencé à explorer de nouveaux indicateurs publics d'équité afin d'évaluer quantitativement trois types bien documentés de préjudices et de biais potentiels dans les modèles de vision par ordinateur. Ces indicateurs d'équité complètent les approches existantes en matière de développement responsable de l'IA, telles que la documentation des données et des modèles. Les indicateurs sont spécifiquement conçus pour s'adapter et évoluer en fonction des progrès de la recherche et de l'émergence de nouvelles approches.

Nous avons également progressé dans le modèle SEER (SElf-SupERvised), le modèle de vision artificielle autosupervisée de Meta AI Research, qui vise à améliorer les résultats pour divers ensembles d'images, sans qu'il soit nécessaire de procéder à la sélection et à l'étiquetage minutieux des données, comme c'est le cas pour la méthode de vision artificielle classique.





Les standards et les principes en matière de droits humains restent la meilleure articulation des valeurs universelles de l'humanité et la meilleure base sur laquelle établir des institutions de gouvernance légitimes. Les modèles de gouvernance de l'Internet ont beaucoup progressé ces dernières années. Comme le montre ce rapport, nous avons travaillé tout au long de l'année 2022 à l'amélioration de nos modèles de gouvernance et au respect de notre [Politique d'entreprise en matière de droits humains](#) et de notre engagement à l'égard des [Principes directeurs des Nations unies relatifs aux entreprises et aux droits de l'homme](#).

Ce travail se poursuit, tout comme notre volonté de progresser. Nous nous réjouissons de mettre en œuvre un grand nombre des mécanismes de responsabilité avancés dont il est question dans le présent rapport. Par exemple, nous continuons à travailler pour permettre aux individus de participer aux mesures de gouvernance par le biais de forums et d'assemblées communautaires.

Nous poursuivrons également notre travail pour que les droits humains restent au cœur de la manière dont nous comprenons et mettons en œuvre la science de l'intelligence artificielle (IA), qui progresse rapidement. Bien qu'il reste de nombreux défis à relever et beaucoup de choses à accomplir, nous sommes fiers des efforts que nous déployons pour résoudre certains des principaux enjeux des réseaux sociaux et des technologies de l'IA.

Nous continuerons à nous appuyer sur la diligence raisonnable en matière de droits humains pour mieux déterminer les risques principaux en matière de droits humains auxquels sont confrontées les personnes qui utilisent nos services et pour élaborer une politique qui réponde à ces risques.

Nous évoluons dans un environnement réglementaire complexe et fragmenté. L'Internet sans frontières et essentiellement libre que beaucoup d'entre nous utilisent chaque jour est menacé. Il est essentiel que nous travaillions avec les gouvernements, les Nations unies et la société civile pour veiller à ce que les cadres réglementaires en évolution soient fondés sur les principes et les valeurs des droits humains. Pleinement conscients de nos responsabilités en matière de droits humains, nous nous efforcerons de faire progresser ces principes dans notre travail quotidien, dans l'intérêt de nos utilisateurs et des sociétés dans leur ensemble.

