



Oversight Board

**RAPPORT  
ANNUEL**

**2022**

**DU CONSEIL DE  
SURVEILLANCE**



# Sommaire

Avant-propos des Coprésidents	4
Avant-propos du Président de la Fiducie	7
Synthèse analytique	8
Présentation du Conseil	10
Présentation de nos sept priorités stratégiques	12
Comment le Conseil tient compte des appels des utilisateurs	16
Chronologie des événements clés de 2022	18

## Recommandations et impact

Vue d'ensemble	21
Des engagements à l'action : obtenir des résultats pour les utilisateurs	22
Chronologie de l'impact : expliquer aux utilisateurs ce qu'ils ont fait de mal	28
Application de nos recommandations par Meta	30

## Sélection des cas

Vue d'ensemble	33
Cas soumis au Conseil	34
Cas pris en considération par le Comité de sélection des cas	37

## Décisions et avis consultatifs en matière de politiques

Vue d'ensemble	39
Décisions et avis consultatifs en matière de politiques publiés en 2022	41
Synthèses des décisions et avis consultatifs en matière de politiques	42
Application des normes internationales des droits de l'homme à la modération du contenu : Article 19	54
Normes internationales des droits de l'homme dans le processus décisionnel du Conseil	55

## Engagement et commentaires publics

Vue d'ensemble	59
Chronologie des activités d'engagement en 2022	62

## Prochaines étapes

Faire évoluer notre travail avec Meta	63
Partager les avantages d'une surveillance indépendante	65
Aider les entreprises à s'adapter aux nouvelles réglementations	67

# Avant-propos des Coprésidents



## **Evelyn Aswad, Catalina Botero-Marino, Michael McConnell, Helle Thorning-Schmidt COPRÉSIDENTS DU CONSEIL DE SURVEILLANCE**


En 2022, un grand nombre de nos recommandations adressées à Meta se sont concrétisées, ce qui a permis d'améliorer la manière dont l'entreprise traite les personnes et les communautés dans le monde entier. Notre travail a conduit Meta à revoir ses politiques de modération du contenu, à énoncer ses règles plus clairement et à les appliquer de manière plus cohérente. Meta informe désormais un plus grand nombre d'utilisateurs de la catégorie de politique spécifique qui a été enfreinte lorsque leurs publications sont supprimées, et améliore ses pratiques de modération de contenu de manière à ce qu'elles soient plus en accord avec les principes des droits de l'homme.

En réponse à nos recommandations, Meta a introduit un protocole de politique de crise afin de renforcer la cohérence de ses réponses aux situations de crise, entrepris une révision de sa politique sur les individus et organisations dangereux, et élaboré un nouveau Standard de la communauté en matière de fausse information. En réponse à une recommandation dans le cadre de notre décision « Symptômes du cancer du sein et nudité », Meta a également amélioré ses techniques d'identification du contexte du cancer du sein dans le contenu sur Instagram, ce qui a contribué à soumettre à un examen manuel des milliers de publications supplémentaires qui auraient auparavant été automatiquement supprimées.

En 2022, nous avons formulé plus de la moitié de nos 91 recommandations de politique dans le cadre de nos premiers avis consultatifs en matière de politiques, notamment sur la manière dont Meta traite ses utilisateurs les plus influents à travers son programme de vérifications croisées. En réponse, Meta s'est engagée à renforcer la protection des personnes particulièrement exposées au risque d'une application excessive de la politique, notamment les journalistes et les défenseurs des droits de l'homme.

Nous avons également protégé la voix des utilisateurs, en particulier lors des crises et des changements politiques et sociaux. Par exemple, début 2023, dans le cadre d'une décision spécifique du Conseil, nous avons exhorté Meta à mieux protéger le discours politique en Iran, où des manifestations historiques de grande ampleur ont été violemment réprimées. En réponse, Meta a autorisé la diffusion de l'expression « Marg bar Khamenei » (qui se traduit littéralement par « Mort à Khamenei [le guide suprême de l'Iran] ») dans le contexte des manifestations en cours en Iran. Meta a également modifié son système de suspension et de sanctions afin de le rendre plus juste à l'égard des utilisateurs.

Chaque décision et avis consultatif en matière de politiques a permis d'accroître la transparence des processus de modération du contenu, par ailleurs souvent opaques, notamment en dévoilant le nombre



d'exceptions faites par l'entreprise au titre de l'actualité dans le cadre de l'administration de ses règles. Nos recommandations de politique suscitent un débat public sur la manière dont les plateformes numériques peuvent relever certains des défis les plus complexes en matière de modération du contenu.

Pour intensifier notre impact, nous avons adopté sept domaines prioritaires sur lesquels nous souhaitons travailler avec les parties prenantes afin d'améliorer l'expérience des internautes. Il s'agit des élections et de l'espace civique, des situations de crise et de conflit, du genre, des discours incitant à la haine envers des groupes marginalisés, de l'utilisation des plateformes Meta par les gouvernements, du traitement équitable des utilisateurs, de l'application automatisée des politiques et de la curation de contenu. Nous nous sommes également préparés à assumer une charge de travail plus importante et à rendre des décisions plus rapidement en 2023.

En 2022, nous avons également constaté qu'il était de plus en plus admis que les décisions relatives à la modération du contenu ne devaient pas être prises uniquement par les entreprises. Dès sa création, le Conseil a été conçu pour tester une approche indépendante de la modération du contenu, qui, si elle s'avérait concluante, pourrait être appliquée à d'autres entreprises. La surveillance indépendante consiste pour les entreprises à s'ouvrir et à convier des personnes extérieures à remettre en question leur mode de fonctionnement. Au cours des trois dernières années, nous avons acquis une grande expérience en matière de surveillance indépendante qui peut aider d'autres entreprises à prendre des décisions plus judicieuses fondées sur le respect de la liberté d'expression et d'autres droits de l'homme. À mesure que les nouvelles réglementations s'accompagnent de nouvelles exigences, nous pensons pouvoir apporter une partie de la solution dans certains domaines spécifiques, tels que la transparence et la communication avec les utilisateurs.

Nous souhaitons profiter de cette occasion pour remercier les fiduciaires du Conseil de surveillance, le personnel de l'Administration et nos collègues membres du Conseil pour leur expertise et leur soutien tout au long du chemin parcouru jusqu'à présent. Plus particulièrement, nous souhaitons saluer la contribution de Jamal Greene, qui a quitté ses fonctions de membre et de Coprésident du Conseil de surveillance en décembre 2022. Le rôle de Jamal a été déterminant dans notre capacité à responsabiliser Meta, et nous tenons à le remercier pour tout ce qu'il a fait pour développer le Conseil de surveillance et accomplir notre mission. Nous remercions également les nombreuses parties prenantes qui ont soumis des commentaires publics, participé à nos travaux et contribué à rendre possibles les résultats obtenus jusqu'à présent.

Étant donné le caractère inédit de sa démarche, le Conseil de surveillance continue de s'adapter et de trouver de nouveaux moyens de remplir sa mission. Si nous avons réalisé des progrès notables jusqu'à présent, nous sommes bien conscients de l'ampleur du défi qui nous attend. Ensemble, nous pouvons contribuer à surmonter les écueils des médias sociaux et aider les utilisateurs à se rapprocher dans un climat de confiance.





# Avant-propos du Président de la Fiducie



**Stephen Neal**  
Président de la Fiducie du  
Conseil de surveillance

En 2022, ma première année pleine en tant que Président de la Fiducie du Conseil de surveillance, j'ai été très impressionné par le travail du Conseil. Les membres du Conseil ont continué à se pencher sur les cas les plus épineux et importants et à prendre des décisions déterminantes sur de nombreux sujets. Le Conseil a notamment publié son premier avis consultatif en matière de politiques, qui portait sur le partage d'informations privées sur le lieu de résidence, et la décision relative au « poème russe », liée à l'invasion de l'Ukraine. En tant que fiduciaires, nous avons contribué à la nomination de trois nouveaux membres du Conseil, originaires d'Égypte, du Mexique et des États-Unis.

Autre aspect essentiel de notre fonction : nous supervisons l'Administration du Conseil de surveillance, le personnel à temps plein qui assiste les membres du Conseil dans leurs activités. En 2022, l'Administration a recruté dans toutes les équipes jusqu'à compter aujourd'hui environ 80 collaborateurs basés à Londres, Washington et San Francisco. Nous nous sommes entourés de nouveaux collègues très compétents, dont beaucoup possèdent une expertise unique en matière de liberté d'expression et de droits de l'homme. À l'instar des médias sociaux, l'Administration revêt un caractère mondial, et les membres de son personnel parlent 40 langues à eux tous.

En juillet 2022, nous avons annoncé un engagement supplémentaire de Meta à hauteur de 150 millions de dollars, en complément des 130 millions de dollars annoncés en 2019 lors de la création de la Fiducie. En s'engageant financièrement, Meta témoigne de sa confiance dans le travail du Conseil et de ses efforts pour appliquer les standards de contenu publié sur Facebook et Instagram, et ce de manière à faire valoir la liberté d'expression de chacun ainsi que les normes pertinentes des droits de l'homme.

En 2023, nous continuerons à superviser les activités du Conseil et à préserver son indépendance, deux conditions sine qua non de son succès. Grâce à ses décisions et à ses avis consultatifs en matière de politiques, le Conseil continuera à améliorer les produits et les politiques de Meta, ce qui permettra aux utilisateurs de Facebook et d'Instagram de bénéficier d'une meilleure expérience. En œuvrant avec des groupes de la société civile, les régulateurs et d'autres plateformes, le Conseil vise à renforcer sa légitimité. Les employés de Meta constituent un autre groupe fondamental qui exerce une influence déterminante sur l'avenir de l'entreprise et que le Conseil s'efforcera de mobiliser pour défendre son travail.



**En s'engageant  
financièrement, Meta  
témoigne de sa confiance  
dans le travail du Conseil »**

PRÉSIDENT DE LA FIDUCIE DU CONSEIL DE SURVEILLANCE,  
**Stephen Neal**

L'engagement de Meta dans la création du Conseil en tant que nouvelle forme de gouvernance des médias sociaux et ses efforts continus dans ce sens sont louables et méritent d'être reconnus. En tant que fiduciaires, nous soutiendrons le Conseil dans sa collaboration fructueuse avec Meta et l'aiderons à partager son approche avec d'autres entreprises et partenaires au sein de toute l'industrie.

# Synthèse analytique

En 2022, le Conseil de surveillance a formulé 91 recommandations à Meta.

En réponse à nos recommandations, Meta :



A commencé à indiquer aux utilisateurs, lors de la suppression de leur contenu, la **politique spécifique** que celui-ci a enfreinte



A commencé à **mesurer systématiquement la transparence** de ses messages d'application de ses politiques à l'intention des utilisateurs



A amélioré ses techniques d'identification du contexte du cancer du sein dans le contenu sur Instagram, ce qui a contribué à **soumettre à un examen manuel des milliers de publications** supplémentaires qui auraient auparavant été automatiquement supprimées



A déployé à l'échelle mondiale de nouveaux messages indiquant aux utilisateurs si la suppression de leur contenu relève d'un examen manuel ou automatique



A créé une nouvelle section sur les **fausses informations** dans les Standards de la communauté



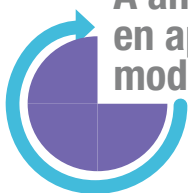
A introduit un nouveau **protocole de politique de crise**



## En 2022, le Conseil de surveillance :

A reçu près de **1,3 million** de cas d'utilisateurs du monde entier, soit environ un quart de plus qu'en 2021

A publié ses premiers avis consultatifs en matière de politiques, portant sur le partage d'informations privées sur le lieu de résidence et sur le programme de vérifications croisées de Meta



A annulé les décisions de Meta en appel dans **75 %** des 12 cas de modération du contenu :

9 infirmations et 3 confirmations

## A publié **12 décisions**

sur des sujets allant de l'invasion de l'Ukraine par la Russie à l'influence de l'application de la loi sur les suppressions de contenu

A élargi son champ d'action afin d'inclure la possibilité d'ajouter des écrans d'avertissement aux contenus admissibles



A amené Meta à revenir sur sa décision initiale dans

**32** cas examinés en vue d'une sélection,

lorsque sa décision initiale concernant une publication était inadéquate.

## En 2023, nous allons :

- ✓ publier nos premiers résumés de décisions concernant des cas où Meta a modifié sa décision initiale sur un élément de contenu ;
- ✓ rendre nos premières décisions accélérées, c'est-à-dire publier une décision concernant un cas en quelques jours ;
- ✓ atteindre notre objectif révisé de nombre total de membres du Conseil, pour un maximum d'efficacité ;
- ✓ intensifier notre engagement autour de nos sept priorités stratégiques ;
- ✓ élaborer des plans à long terme en vue d'élargir notre champ d'action ;
- ✓ contrôler la manière dont Meta applique nos recommandations et inciter l'entreprise à produire des preuves de la mise en œuvre de ces recommandations et de leur impact.

Nous sommes convaincus de l'intérêt d'une surveillance indépendante et nous étudierons la possibilité d'établir de nouveaux partenariats avec d'autres entreprises, ainsi que la manière dont notre travail peut compléter au mieux les réglementations émergentes.

# Présentation du Conseil

## Membres du Conseil de surveillance



**Afia Asantewaa Asare-Kyei**

Responsable de programme, initiative Open Society Initiative for West Africa (OSIWA)



**Evelyn Aswad**

Professeure et présidente, faculté de droit de l'université d'Oklahoma



**Endy Bayuni**

Rédacteur en chef et membre du conseil d'administration, The Jakarta Post



**Catalina Botero-Marino**

Directrice de la chaire UNESCO sur la liberté d'expression, université des Andes



**Paolo Carozza**

Professeur, université de Notre-Dame



**Katherine Chen**

Professeure, université nationale Chengchi



**Nighat Dad**

Fondatrice, Digital Rights Foundation



**Tawakkol Karman**

Lauréate du Prix Nobel de la paix



**Sudhir Krishnaswamy**

Vice-chancelier et professeur de droit, École nationale de droit de l'université de l'Inde



**Ronaldo Lemos**

Professeur, faculté de droit de l'université d'État de Rio de Janeiro



**Khaled Mansour**

Écrivain



**Michael McConnell**

Professeur et directeur du Centre de droit constitutionnel, faculté de droit de l'université de Stanford



**Suzanne Nossel**

Directrice générale, PEN America



**Julie Owono**

Directrice générale, Internet Sans Frontières



**Emi Palmor**

Avocate et maîtresse de conférence,  
Centre interdisciplinaire Herzliya,  
Israël



**Pamela San Martín**

Ancienne conseillère électorale à  
l'Institut national électoral (INE),  
Mexique



**Alan Rusbridger**

Directeur, Lady Margaret Hall Oxford



**Nicolas Suzor**

Professeur, faculté de droit de  
l'université de technologie du  
Queensland



**András Sajó**

Professeur d'université, université  
d'Europe Centrale



**Helle Thorning-Schmidt**

Ancienne Première ministre,  
Danemark



**John Samples**

Vice-président, Cato Institute



**Kenji Yoshino**

Juge en chef d'Earl Warren, professeur  
de droit constitutionnel et directeur  
du Meltzer Center for Diversity,  
Inclusion, and Belonging

## Fiduciaires du Conseil de surveillance



**Kristina Arriaga**

Fiduciaire



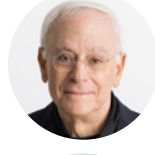
**Kate O'Regan**

Fiduciaire



**Cherine Chalaby**

Fiduciaire



**Robert Post**

Fiduciaire



**Stephen Neal**

Président de la Fiducie



**Marie Wieck**

Fiduciaire

## Administration du Conseil de surveillance



**Thomas Hughes**

Directeur

# Présentation de nos sept priorités stratégiques

**En octobre 2022, nous avons annoncé sept priorités stratégiques découlant d'une analyse approfondie des points soulevés par les utilisateurs dans leurs appels auprès du Conseil. Dans la mesure où ces priorités orientent désormais les cas que nous sélectionnons, nous encourageons les utilisateurs à en tenir compte lorsqu'ils font appel.**



## 1. Élections et espace civique

Les entreprises de médias sociaux sont confrontées à des difficultés pour appliquer de manière cohérente leurs politiques en matière d'expression politique dans de nombreuses régions du monde, y compris lors d'élections et de manifestations de grande ampleur. Nous avons souligné l'importance de la protection de l'expression politique dans notre décision relative aux manifestations de soutien à Navalny en Russie, tandis que celle concernant la mention des talibans dans les actualités abordait des questions liées à la liberté des médias. Dans le cadre de notre mandat, nous aimerions examiner les responsabilités de Meta lors d'élections, de manifestations et d'autres moments clés de la participation citoyenne.



## 2. Situations de crise et de conflit

En temps de crise, notamment en cas de conflit armé, d'attaque terroriste ou d'urgence sanitaire, les médias sociaux peuvent faciliter l'échange d'informations cruciales, favoriser le débat public et renforcer la sécurité. Pour autant, ils peuvent aussi créer un environnement propice à la propagation de fausses informations et à la haine. Nos décisions relatives à des crimes présumés à Raya Kobo et au Tigray Communication Affairs Bureau ont porté sur des publications liées au conflit en Éthiopie, tandis que celle concernant l'ancien président des États-Unis, Donald Trump, a conduit Meta à adopter un protocole de politique de crise. Dans le cadre de notre mandat, nous aimerions examiner le rôle de Meta dans la protection de la liberté d'expression dans de telles circonstances, ainsi que son degré de préparation aux éventuels préjudices que ses produits peuvent causer lors de conflits armés, de troubles civils et d'autres situations d'urgence.



## 3. Genre

Les femmes, les personnes non binaires et les personnes transgenres se heurtent à divers obstacles dans l'exercice de leur droit à la liberté d'expression sur les médias sociaux. Dans notre décision concernant les symptômes du cancer du sein et la nudité, par exemple, les systèmes automatisés de Meta n'ont pas appliqué les exceptions relatives à la sensibilisation au cancer du sein, ce qui a conduit à la suppression d'informations de santé importantes sur Instagram. Notre décision sur la nudité et l'identité de genre, publiée début 2023, a également établi que les politiques de Meta sur la nudité chez les adultes entraînaient des obstacles plus importants à l'expression des femmes, des personnes transgenres et des personnes non binaires sur Facebook et Instagram. Dans le cadre de notre mandat, nous aimerions examiner les obstacles liés au genre auxquels les femmes et les personnes LGBTQIA+ sont confrontées dans l'exercice de leurs droits à la liberté d'expression, y compris la violence et le harcèlement fondés sur le genre, ainsi que les effets des distinctions fondées sur le genre dans la politique de contenu.



## 4. Discours incitant à la haine envers des groupes marginalisés

Le discours incitant à la haine instaure un climat de discrimination et d'hostilité à l'égard des groupes marginalisés. Il est souvent propre à un contexte, codé, et ses effets néfastes s'accumulent progressivement au fil du temps. Notre décision relative à la description de Zwarte Piet a conclu que l'accumulation d'images de blackface en ligne créerait un environnement discriminatoire pour les personnes noires. En parallèle, nos décisions concernant le collier de wampum et la réappropriation des mots arabes ont examiné le « contre-discours », qui se réfère à des discours incitant à la haine pour résister à la discrimination. Dans le cadre de notre mandat, nous aimerions examiner la manière dont Meta devrait protéger les membres issus de groupes marginalisés, tout en veillant à ce que son application ne cible pas à tort les personnes qui s'opposent à la haine. Par ailleurs, nous sommes conscients que les restrictions sur le discours incitant à la haine ne doivent pas être appliquées de manière excessive ou utilisées pour limiter l'exercice légitime de la liberté d'expression, y compris l'expression de points de vue impopulaires ou controversés.



## 5. Utilisation des plateformes Meta par les gouvernements

Les gouvernements utilisent Facebook et Instagram pour relayer leurs politiques et demandent à Meta de supprimer certains contenus. En réponse à notre décision sur l'isolement d'Öcalan, Meta a accepté de fournir des informations au sujet des contenus supprimés pour violation des Standards de la communauté suite à un rapport d'un gouvernement. Notre décision relative au genre de musique UK drill a également formulé des propositions sur la manière dont Meta devrait répondre aux demandes des autorités nationales chargées de l'application de la loi. Dans le cadre de notre mandat, nous aimerions examiner la manière dont les acteurs étatiques utilisent les plateformes Meta, comment ils peuvent influencer les pratiques et les politiques de modération du contenu (parfois dans la plus grande opacité), et les conséquences de l'implication de l'État dans la modération du contenu.



## 6. Traitement équitable des utilisateurs

Lorsqu'un contenu est supprimé de Facebook et d'Instagram, son auteur n'est pas toujours informé de la règle qu'il a enfreinte. Dans d'autres cas, les utilisateurs ne sont pas traités sur un pied d'égalité ou ne bénéficient pas de garanties procédurales adéquates et de recours en cas d'erreur. Dans le cadre de notre mandat, nous aimerions examiner la manière dont Meta peut mieux traiter ses utilisateurs, en leur communiquant des notifications plus spécifiques, en veillant à ce qu'ils puissent toujours faire appel de la décision de Meta auprès de l'entreprise et en faisant preuve de plus de transparence dans certains domaines tels que les « suspensions » et les vérifications croisées.



## 7. Application automatisée des politiques et de la curation de contenu

Si les algorithmes sont fondamentaux pour modérer le contenu à grande échelle, on observe un manque de transparence et de compréhension autour du fonctionnement des systèmes automatisés de Meta et de la manière dont ils influencent le contenu que les utilisateurs voient. Notre décision concernant le dessin sur la police colombienne a montré comment l'automatisation peut amplifier les répercussions des mauvaises décisions en matière de modération du contenu. En réponse à notre décision concernant les symptômes du cancer du sein et la nudité, Meta envoie désormais de nouveaux messages indiquant aux utilisateurs si la suppression de leur contenu est d'origine automatique ou manuelle. Dans le cadre de notre mandat, nous

aimerions examiner la manière dont l'application automatisée des règles devrait être conçue et révisée, mais aussi l'exactitude et les limites des systèmes automatisés, ainsi que l'importance d'une plus grande transparence dans ce domaine.

## ||| **COLLABORER AVEC LES PARTIES PRENANTES POUR AMPLIFIER NOTRE IMPACT**

En tant que Conseil, les résultats obtenus jusqu'à présent ont été rendus possibles grâce à l'écoute et à la collaboration avec des chercheurs, des groupes de la société civile et d'autres acteurs qui ont œuvré pendant de nombreuses années sur les thèmes que nous abordons. Pour trouver des solutions pratiques à nos priorités stratégiques et aux questions épineuses qu'elles soulèvent, l'expertise en la matière et les connaissances sur le terrain de ces parties prenantes sont primordiales.

Pour chacune de ces priorités stratégiques, nous continuerons à travailler avec un large éventail de protagonistes qui témoignent de la diversité des utilisateurs des plateformes de Meta. Ainsi, nous comprendrons mieux les politiques et pratiques de mise en application qui doivent être améliorées d'urgence chez Meta, ainsi que les types de cas permettant de se pencher dessus. Pour ce faire, nous souhaitons établir des partenariats avec des organisations du monde entier, via notre procédure de commentaires publics, des tables rondes et des entretiens individuels. Pour découvrir comment votre organisation peut s'impliquer, veuillez contacter [engagement@osbadmin.com](mailto:engagement@osbadmin.com).



# Comment le Conseil tient compte des appels des utilisateurs

Ce graphique présente la procédure d'appel telle qu'appliquée aux décisions relatives à des appels d'utilisateurs rendues en 2022.







### DÉCISION

Le jury parvient à une décision quant à l'autorisation du contenu : confirmer ou renverser les décisions de Meta.



### PUBLICATION

Notre décision est publiée sur le site Web du Conseil de surveillance. Meta doit appliquer nos décisions dans un délai de sept jours après publication et répondre à toutes les recommandations dans un délai de 60 jours.



### APPROBATION

Une décision préliminaire est envoyée pour révision à tous les membres du Conseil. Une décision doit être signée par la majorité pour être publiée.

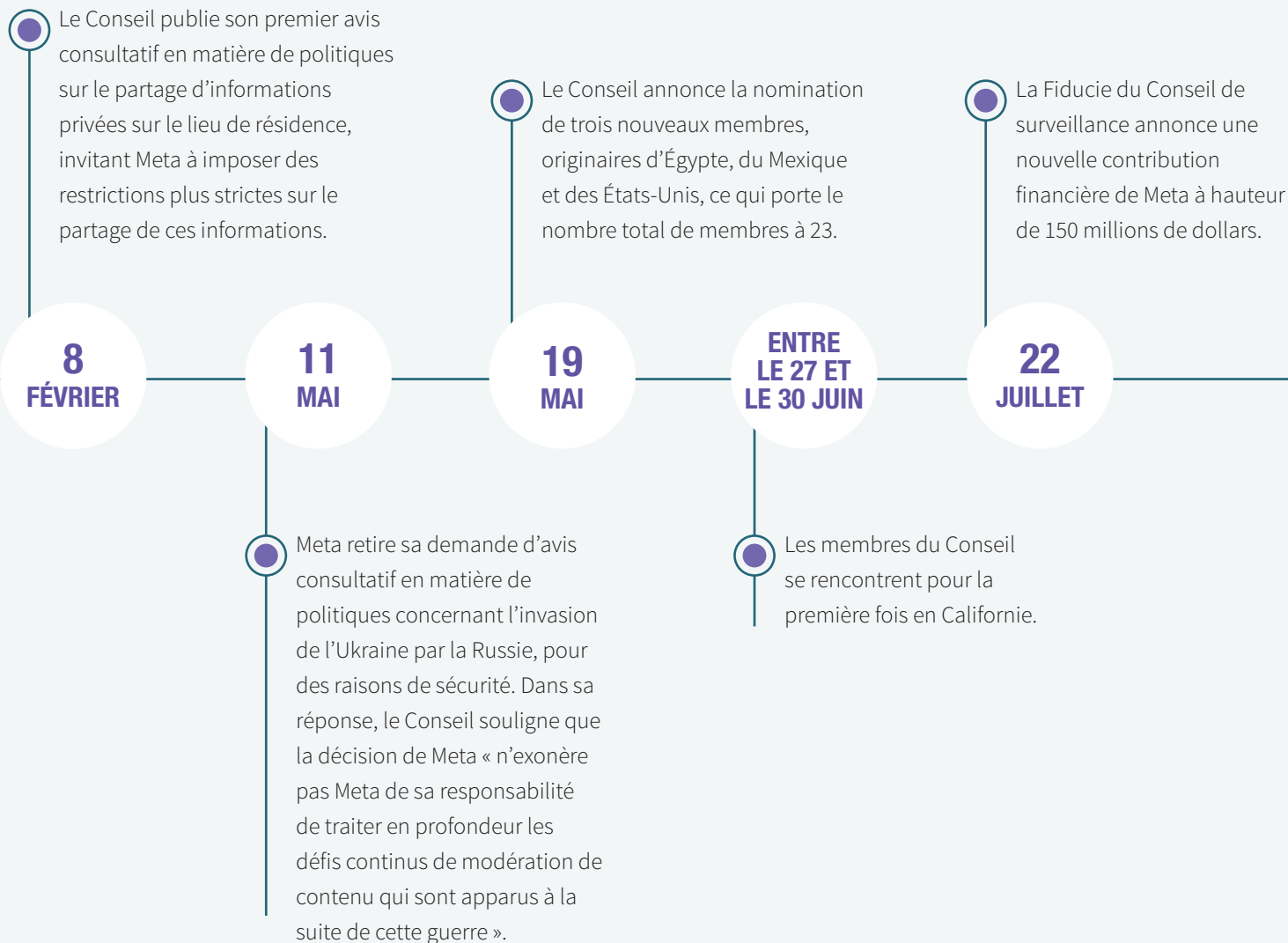


### APPLICATION

Le Conseil veille à l'application des recommandations par Meta, en faisant le point dans des rapports de transparence trimestriels.



# Chronologie des événements clés de 2022



**26  
JUILLET**

Le Conseil accepte la demande d'avis consultatif en matière de politiques formulée par Meta concernant la suppression des fausses informations sur la COVID-19.

**20  
OCTOBRE**

Le Conseil annonce sept priorités stratégiques axées sur les domaines dans lesquels il peut exercer la plus forte influence sur l'expérience des utilisateurs de Facebook et d'Instagram.

**20  
OCTOBRE**

Le Conseil obtient la possibilité d'afficher des écrans d'avertissement qui qualifient les publications de « dérangeantes » ou de « sensibles » lorsqu'il décide de restaurer ou de conserver un contenu admissible.

**22  
NOVEMBRE**

Le Conseil publie sa décision relative au genre de musique UK drill. C'est la première fois qu'il examine une publication supprimée à la suite d'une demande émanant d'autorités nationales chargées de l'application de la loi.

**6  
DÉCEMBRE**

Le Conseil publie un avis consultatif en matière de politiques concernant le programme de cross-checking de Meta. Il estime que les vérifications croisées sont défaillantes dans certains domaines clés et formule 32 propositions à l'intention de Meta.

# Recommandations et impact

Nous avons effectué **91** recommandations auprès de Meta en 2022

## En réponse à nos recommandations, Meta :



a mis en place à l'échelle mondiale de nouvelles notifications indiquant aux utilisateurs la politique spécifique qu'ils ont enfreinte dans le cadre des politiques relatives aux discours incitant à la haine, aux individus et organisations dangereux, ainsi qu'à l'intimidation et au harcèlement ;



a commencé à mesurer systématiquement la transparence de ses messages d'application de ses politiques à l'intention des utilisateurs ;



a amélioré ses techniques d'identification du contexte du cancer du sein dans le contenu sur Instagram, ce qui a contribué à soumettre à un examen manuel des milliers de publications supplémentaires qui auraient auparavant été automatiquement supprimées ;



a déployé à l'échelle mondiale de nouveaux messages indiquant aux utilisateurs si la suppression de leur contenu relève d'un examen manuel ou automatique ;



a mis en place de nouvelles notifications indiquant aux utilisateurs que leur accès à certains contenus a été restreint en raison de la législation locale, à la suite d'une demande des autorités ;



a créé une nouvelle section sur les fausses informations dans les Standards de la communauté ;



a lancé un nouveau protocole de politique de crise ;



a entamé une révision complète de sa politique sur les individus et les organisations dangereux afin de hiérarchiser les désignations en fonction du risque.



## Vue d'ensemble

Dans nos décisions et nos avis consultatifs en matière de politiques, nous formulons des recommandations spécifiques sur la manière dont Meta peut améliorer les politiques qu'elle applique aux contenus de milliards d'utilisateurs. Bien que nos recommandations ne soient pas contraignantes, Meta doit y répondre publiquement dans un délai de 60 jours. Meta a reconnu officiellement que nos recommandations modifiaient le comportement de l'entreprise. En août 2022, elle faisait une déclaration en ce sens : « [Le Conseil] ne cesse de nous pousser à réfléchir davantage à l'impact de notre modération des contenus à l'échelle mondiale et à être plus justes dans l'application de nos politiques et dans l'utilisation de nos ressources. Surtout, ces recommandations nous encouragent à faire preuve d'une plus grande transparence, car les avis extérieurs peuvent nous aider à respecter nos engagements ».

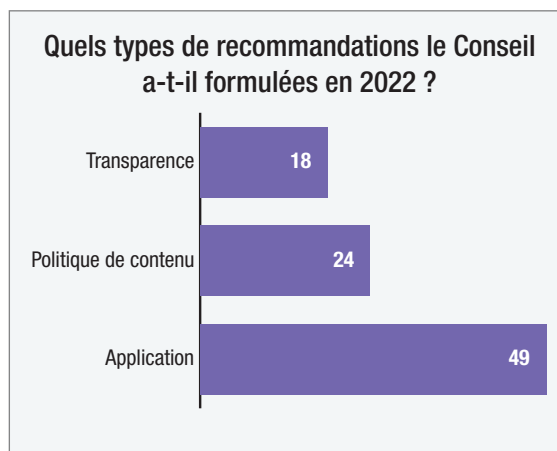
Dans le cadre de notre mandat, nous veillons à ce que Meta rende des comptes en publiant des rapports de transparence trimestriels. Ceux-ci s'appuient sur une approche rigoureuse, indépendante et fondée sur des données afin de mesurer les progrès accomplis par Meta dans l'application de nos recommandations. En formulant publiquement ces recommandations et en supervisant publiquement les réponses et la mise en œuvre de Meta, nous avons créé un espace de dialogue transparent avec l'entreprise qui n'existait pas auparavant. Cette transparence permet d'établir une relation de confiance et une image légitime auprès des utilisateurs et de la société civile.

On ne soulignera jamais assez l'importance des groupes de la société civile dans l'élaboration de nos recommandations. Dans bien des cas, ces organisations soumettent des idées concrètes de recommandations dans le cadre de notre processus de commentaires publics. Dans d'autres cas, nos propositions font écho ou s'appuient sur des appels lancés par ces groupes depuis de nombreuses années, ce qui oblige Meta à réfléchir et à répondre publiquement à des appels à l'action de longue date. Bien que nous mentionnions explicitement ces influences dans les textes de nos décisions, nous tenons à leur exprimer à nouveau notre gratitude pour le partage de leurs idées et leur expertise.

### RECOMMANDATIONS ET IMPACT EN 2022

En 2022, le Conseil a adressé 91 recommandations à Meta, soit une légère augmentation par rapport aux 86 propositions que nous lui avons présentées en 2021. Au début du mois d'avril 2023, nous avons formulé 14 recommandations supplémentaires, soit un total de 191 recommandations à Meta.

« [Les recommandations du Conseil] nous encouragent à faire preuve d'une plus grande transparence, car les avis extérieurs peuvent nous aider à respecter nos engagements »  
Mises à jour du 2e trimestre 2022 effectuées par Meta pour le Conseil de surveillance



Au total, 41 recommandations ont été classées dans les catégories « Application prouvée » ou « Application partielle prouvée », et nous avons évalué 84 autres recommandations comme « Progrès signalés ». En 2022, il a été particulièrement encourageant de constater que, lors de l'application de nos recommandations, Meta a procédé à plusieurs changements qui ont eu un impact systémique sur l'approche de l'entreprise. Il s'agit notamment d'envoyer des notifications plus spécifiques aux utilisateurs, ce que nous réclamons depuis janvier 2021. À la suite des recommandations répétées du Conseil, Meta a également actualisé la façon dont elle mesure la spécificité et la transparence des messages qu'elle envoie aux utilisateurs lorsqu'elle prend des mesures de mise en application à l'encontre d'un contenu pour violation de ses politiques. Ce changement systémique s'inscrit dans le cadre des efforts déployés par l'entreprise pour améliorer la précision de ses communications avec les utilisateurs. Meta procède également à un examen approfondi de la définition du terme « faire l'éloge » par rapport à son cadre d'éloges, de soutien substantiel et de représentation (PSR) en vertu de la politique relative aux individus et organisations dangereux. Meta s'appuie sur ce cadre pour évaluer dans quelle mesure les individus et les organisations dangereux sont favorablement représentés dans le contenu des utilisateurs.

## ENSEIGNEMENTS TIRÉS

L'accès aux données est un point sur lequel la coopération avec Meta aurait pu être améliorée en 2022. Fin 2021 et début 2022, nous avons passé huit mois à tenter d'accéder à l'outil CrowdTangle de Meta afin d'obtenir davantage d'informations lors de la sélection des cas et de l'évaluation de l'impact des recommandations. Après plusieurs obstacles, nous avons porté ce problème à l'attention de la direction de Meta au début de l'année 2022 et avons finalement obtenu l'accès à l'outil. Plus tard en 2022, Meta a pris plusieurs mesures positives concernant le partage des données, notamment le partage des recherches en cours sur les appels des utilisateurs et l'embauche d'un scientifique des données chargé de valider l'application de nos recommandations. Nous nous réjouissons de poursuivre notre partenariat avec les équipes de Meta spécialisées dans la science des données afin d'obtenir des données significatives attestant à la fois de leur application et de leur impact.

## Des engagements à l'action : obtenir des résultats pour les utilisateurs

Nous comprenons que l'application de nos recommandations prenne du temps, compte tenu de leur ambition et des adaptations techniques qu'elles requièrent souvent. En 2022, nous avons constaté des progrès dans l'application d'un grand nombre des recommandations formulées en 2021, ainsi que de nouveaux engagements en réponse à des propositions plus récentes. Il est encourageant de constater que, pour la première fois, Meta a apporté des changements systémiques à ses règles et à la manière dont elles sont appliquées, notamment en ce qui concerne les notifications aux utilisateurs et les règles relatives aux organisations dangereuses. Les exemples ci-dessous illustrent l'impact de nos recommandations sur la manière dont l'entreprise traite les utilisateurs et les communautés dans le monde entier.

## Une nouvelle étape pour mieux comprendre notre impact

Dans notre décision concernant les symptômes du cancer du sein et la nudité, publiée en janvier 2021, nous avons recommandé à Meta d’« **améliorer la détection automatique d’images avec texte superposé afin d’assurer que les publications de sensibilisation aux symptômes du cancer du sein ne soient pas signalées à tort pour examen** ». En réponse, l’équipe chargée de la mise en œuvre chez Meta a optimisé les techniques employées par Instagram pour identifier les contenus relatifs au cancer du sein à partir d’un texte et les a déployées en juillet 2021. Ces améliorations sont effectives depuis lors et, au cours des 30 jours écoulés entre le 26 février et le 27 mars 2023, elles ont contribué à l’envoi pour examen manuel de 2 500 éléments de contenu supplémentaires qui auraient été auparavant supprimés.

Bien qu’il soit ardu de contextualiser 2 500 éléments de contenu sans dénominateur, nous constatons que Meta a appliqué la recommandation en suivant scrupuleusement notre démarche et qu’elle a permis de réduire l’application excessive des règles sur la plateforme. Cette victoire en faveur d’une gouvernance indépendante marque le début d’un nouveau chapitre dans la collaboration entre Meta et le Conseil afin de mieux comprendre notre impact sur les systèmes de Meta.

### ||| CHANGEMENTS SYSTÉMIQUES DANS LES RÈGLES DE META ET LEUR APPLICATION

- ⦿ Dans le cadre de notre mandat, la recommandation la plus fréquente consiste à demander à Meta d’**expliquer aux utilisateurs ce qu’ils ont fait de mal** lorsque leur contenu est supprimé. Depuis notre première recommandation en janvier 2021, Meta s’est peu à peu rapproché de cet objectif. En réponse à cette recommandation, Meta a mis en place à l’échelle mondiale de nouveaux messages indiquant aux utilisateurs la politique spécifique qu’ils ont enfreinte dans le cadre des politiques relatives aux discours incitant à la haine, aux individus et organisations dangereux, ainsi qu’à l’intimidation et au harcèlement. Suite à notre recommandation, Meta mesure désormais systématiquement le niveau de détail de ses communications aux utilisateurs concernant toutes les suppressions de contenu.
- ⦿ En réponse à une recommandation découlant de notre décision concernant l’ancien président des États-Unis, Donald Trump, Meta a procédé à des **changements systémiques dans sa réponse aux crises et aux conflits**. En août 2022, après avoir consulté plus de 50 experts internationaux, Meta a publié son



protocole de politique de crise. Ce dernier contribuera à fournir un cadre plus cohérent et transparent sur la façon dont Meta gère les situations de crise. Le 25 janvier 2023, Meta a indiqué avoir eu recours au protocole de politique de crise pour évaluer l'environnement actuel, notamment en examinant le déroulement des élections américaines de mi-mandat de 2022, avant de prendre une décision sur les comptes de l'ancien président Donald Trump.

- ⦿ Nombre de nos propositions ont exhorté Meta à faire preuve d'une **bien plus grande transparence** à l'égard de ses règles et exceptions. En réponse, en août 2022, Meta a révélé pour la première fois le nombre d'autorisations en vertu d'un « **intérêt médiatique** » accordé aux contenus violents qu'elle considérait comme étant d'intérêt public. Du 1er juin 2021 au 1er juin 2022, Meta a documenté 68 autorisations, dont 13 accordées pour des publications de personnalités politiques. Cette transparence permet d'établir une relation de confiance et une image légitime auprès des utilisateurs et de la société civile.
- ⦿ Suite à nos recommandations concernant ses règles relatives aux **individus et organisations dangereux**, Meta a entrepris un examen approfondi de sa politique. Cet examen est axé sur l'adoption d'une approche fondée sur le risque pour désigner des personnes ou des organisations comme dangereuses, les entités considérées comme présentant le risque le plus élevé devant faire l'objet d'une application prioritaire de la législation. Dans plusieurs de nos décisions, nous avons également constaté que la définition du terme « faire l'éloge » retenue par Meta dans le cadre de cette politique limitait trop l'expression des utilisateurs. En réponse, Meta réexamine la manière dont elle évalue si un contenu s'apparente à des éloges, à un soutien substantiel ou à la représentation d'un individu ou d'une organisation désignée.

## ||| TRAITEMENT ÉQUITABLE DES UTILISATEURS

- ⦿ En 2022, Meta a **actualisé les notifications adressées aux utilisateurs pour les contenus supprimés à la suite d'une demande émanant des autorités** conformément à la législation en vigueur. Elles indiquent aux utilisateurs que le contenu a été restreint et expliquent la manière dont Meta traite ces demandes.
- ⦿ Meta a déployé à l'échelle mondiale de nouveaux messages indiquant aux utilisateurs si la suppression de leur contenu relève d'un **examen manuel ou automatique**.
- ⦿ En réponse à notre décision concernant les allégations sur un remède à la COVID **Meta a créé une nouvelle section dans les Standards de la communauté sur les fausses informations**, de manière à regrouper et à clarifier les règles en un seul endroit.



## Un rapport indépendant sur les droits de l'homme concernant l'impact de Meta en Israël et en Palestine en mai 2021

Après avoir lu, dans les commentaires publics soumis dans le cadre de notre cas de 2021 concernant la publication partagée d'Al Jazeera, des allégations selon lesquelles Meta aurait supprimé de manière disproportionnée des publications de Palestiniens, nous avons demandé à Meta d'engager une entité indépendante pour examiner si sa modération du contenu lié au conflit entre Israël et la Palestine en mai 2021 avait été biaisée.

Meta y a consenti et a chargé l'organisation à but non lucratif Business for Social Responsibility (BSR) de procéder à cet examen, qu'elle a publié en septembre 2022. Le rapport conclut que la modération du contenu de Meta pendant le conflit israélo-palestinien de mai 2021 semble avoir eu un impact négatif sur les droits des utilisateurs palestiniens à la liberté d'expression, et sur leur capacité à partager des informations et des réflexions

sur leurs expériences au fur et à mesure qu'elles se produisaient. Une grande partie des biais identifiés dans le rapport de BSR est liée à un manque de compétences linguistiques et d'instructions chez les modérateurs de contenu de Meta. Dans le cadre de notre mandat, nous avons demandé à Meta, dans trois décisions distinctes, de traduire ses instructions internes destinées aux modérateurs dans la langue du contenu qu'ils examinent. Néanmoins, Meta a toujours refusé d'appliquer cette recommandation au motif que les instructions en anglais sont suffisantes, car ses modérateurs parlent couramment l'anglais. Le rapport de BSR a mis en évidence que le manque de compétences linguistiques et de connaissance du contexte culturel des modérateurs de Meta a conduit à une application excessive des règles pour les contenus en arabe palestinien, mais insuffisante dans le cas des contenus antisémites. Comme nous l'avons déjà mentionné, les instructions rédigées exclusivement en anglais peuvent amener les examinateurs à ne pas tenir compte du contexte et des nuances entre les langues et les dialectes. En outre, dans la mesure où les données issues d'examen manuels sont utilisées pour former les classificateurs, ce biais est amplifié sur Facebook et Instagram.



**Les actions de Meta en mai 2021 semblent avoir porté atteinte aux droits de l'homme... aux droits des utilisateurs palestiniens à la liberté d'expression. »**

**Droits de l'homme et diligence raisonnable des impacts de Meta en Israël et en Palestine - rapport de Business for Social Responsibility (BSR)**

## RECOMMANDATIONS NÉCESSITANT UNE ATTENTION PARTICULIÈRE DE LA PART DE META

L'application par Meta d'un grand nombre de nos recommandations précédentes améliore déjà l'expérience des utilisateurs sur Facebook et Instagram. Toutefois, il existe trois domaines dans lesquels l'entreprise n'a pas encore appliqué les principales recommandations.

1. Harmoniser les règles de Facebook et d'Instagram.
2. Traduire les instructions internes destinées aux modérateurs dans les langues dans lesquelles ils modèrent le contenu.
3. Fournir plus d'informations sur la façon dont les publications d'intérêt médiatique sont transmises à l'échelon supérieur au sein de Meta.

**Harmoniser les règles de Facebook et d'Instagram.** Nous avons recommandé pour la première fois à Meta d'harmoniser les règles de Facebook et d'Instagram dans notre décision concernant les symptômes du cancer du sein et la nudité en janvier 2021. Si l'entreprise s'est dans un premier temps engagée à appliquer cette recommandation, elle a à plusieurs reprises repoussé son échéance. Contrairement à nos recommandations, Meta n'a pas encore annoncé aux utilisateurs qu'elle applique les Standards de la communauté Facebook sur Instagram et que si un contenu est considéré comme une violation sur Facebook, il l'est également sur Instagram.

**Traduire les instructions internes destinées aux modérateurs dans les langues dans lesquelles ils modèrent le contenu.** Dans trois décisions relatives à des contenus en punjabi, en birman et en arabe, nous avons recommandé à Meta de traduire ses instructions internes destinées aux modérateurs (standards d'implémentation internes) dans les langues des contenus visés. Pourtant, et malgré les observations similaires formulées par BSR dans son récent rapport, Meta a répété à plusieurs reprises qu'elle ne prendrait « aucune mesure supplémentaire » concernant cette recommandation, dans la mesure où ses modérateurs parlent tous couramment l'anglais.

**Fournir plus d'informations sur la façon dont les publications d'intérêt médiatique sont transmises à l'échelon supérieur au sein de Meta.** Cette année, Meta a révélé pour la première fois le nombre d'autorisations en vertu d'un intérêt médiatique qu'elle a accordé. Néanmoins, le processus par lequel elle décide si un contenu est d'intérêt médiatique reste peu connu. Dans notre décision de 2021 sur les manifestations en Colombie, nous avons demandé à Meta d'élaborer et de publier des critères clairs à l'intention des examinateurs de contenu sur la manière de faire remonter les contenus qui enfreignent les règles de Meta, mais qui pourraient relever de l'intérêt médiatique. La réponse de Meta à cette recommandation suggère une mauvaise compréhension de sa finalité. En effet, elle met l'accent sur le fait qu'un élément peut faire l'objet d'une remontée en vue de bénéficier d'une autorisation en vertu d'un intérêt médiatique, plutôt que sur ce que la recommandation demande véritablement : une description *du moment et des circonstances* dans lesquels un élément peut faire l'objet d'une remontée. Nous encourageons donc l'entreprise à privilégier le partage d'informations supplémentaires sur ce point.

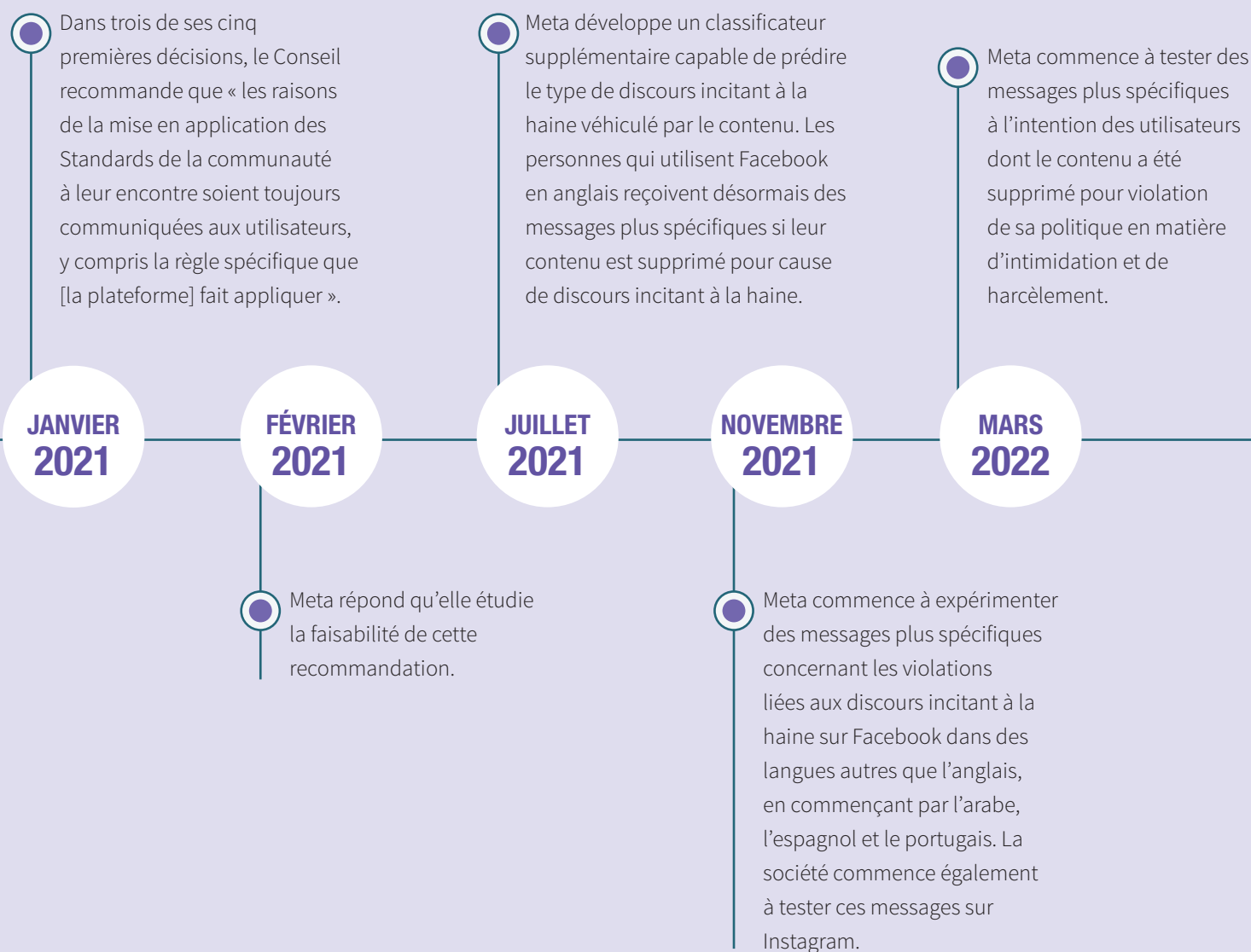
## RAPPORT SUR L'INTEMPORALITÉ DE L'IMPLÉMENTATION ET DE LA RÉPONSE À NOS RECOMMANDATIONS PAR META

- ⦿ En vertu de nos statuts, Meta doit répondre publiquement à nos recommandations dans les 60 jours suivant la publication d'un avis consultatif en matière de politiques par le Conseil.
- ⦿ En ce qui concerne les 12 décisions et les deux avis consultatifs en matière de politiques publiés en 2022, Meta a répondu à nos recommandations dans les délais impartis, sauf dans le cas de notre avis consultatif relatif aux vérifications croisées. En l'occurrence, en raison du grand nombre de recommandations que nous avons formulées (32 au total), Meta a répondu à nos propositions dans un délai de 90 jours, au lieu des 60 prévus habituellement.



# Chronologie de l'impact : expliquer aux utilisateurs ce qu'ils ont fait de mal

**En réponse à nos propositions, Meta indique désormais à un plus grand nombre d'utilisateurs du monde entier la *politique spécifique* que leur contenu a enfreinte lors de sa suppression.**



**MARS  
2022**

Meta indique que des notifications plus spécifiques ont entraîné « une augmentation statistiquement significative de la perception de sa transparence et de sa légitimité sur plusieurs marchés ».

**MAI  
2022**

Meta affirme que, d'après ses études, pour que les utilisateurs se sentent écoutés, son processus de révision des contenus doit s'apparenter davantage à un dialogue qui favorise la compréhension mutuelle.

**AOÛT  
2022**

Meta lance de nouveaux messages expliquant précisément aux utilisateurs quelle politique l'a amenée à prendre une mesure de mise en application. Ceux-ci sont disponibles dans le monde entier, en anglais, et des versions dans d'autres langues sont en cours de traduction.

**NOVEMBRE  
2022**

Meta indique que ces nouveaux messages s'appliqueront à la grande majorité des types de violation sur Facebook d'ici fin 2022, et s'étendront à Instagram en 2023. Meta évaluera également la possibilité de rendre les notifications encore plus spécifiques.

## Application de nos recommandations par Meta

Pour nous assurer que Meta respecte ses engagements, nous suivons de près les progrès réalisés dans l'application de nos recommandations. Pour ce faire, nous examinons si les critères ont été respectés pour une recommandation donnée. Nous mesurons le degré d'application par Meta à l'aide de sept catégories, en mettant à jour nos évaluations sur une base trimestrielle.



« **Application prouvée par la publication d'informations** »

Meta a fourni suffisamment de données pour que le Conseil puisse attester que la recommandation a bien été appliquée.



« **Application partielle prouvée par la publication d'informations** »

Meta a appliqué un aspect fondamental de la recommandation et a fourni suffisamment de données pour que le Conseil puisse en attester.



« **Progrès signalés** »

Meta s'est engagée à implémenter cette recommandation, mais n'a pas encore finalisé toutes les actions nécessaires.



« **Meta a signalé l'implémentation d'une action, ou décrit un travail en cours de réalisation, mais n'a publié aucune information attestant de cette implémentation** »

Meta affirme qu'elle a appliqué cette recommandation, mais elle n'a pas fourni de preuves suffisantes pour que nous puissions en attester.



« **Recommandation refusée après une étude de faisabilité** »

Meta s'est engagée à suivre la recommandation, puis a décidé de refuser son application après avoir justifié sa décision.



« **Recommandation omise, refusée ou redéfinie** »

Meta ne donnera pas suite à notre proposition.










« **En attente d'une réponse préliminaire de la part de Meta** »

Le Conseil a formulé la recommandation, mais Meta n'a pas encore répondu publiquement (elle dispose de 60 jours à compter de la publication pour le faire).

Cette approche fondée sur des données signifie que notre évaluation de l'application d'une recommandation par Meta peut différer des rapports produits par l'entreprise. Nous estimons cependant que ce type de validation menée de manière indépendante est essentiel pour demander des comptes à Meta et s'assurer que les utilisateurs ressentent l'impact de nos recommandations.

Le tableau de cette page détaille chaque catégorie d'application par Meta des 191 recommandations que nous avons formulées jusqu'au début du mois d'avril 2023, date à laquelle le présent rapport a été finalisé. Nous avons également publié un document d'accompagnement du présent rapport annuel qui reprend l'intégralité de nos 191 recommandations et présente notre évaluation de la réponse et de l'application de Meta. Ces données reflètent les rapports trimestriels de Meta pour le Conseil jusqu'au 4e trimestre 2022.

CATÉGORIE D'APPLICATION	NOMBRE DE RECOMMANDATIONS
 <b>Implémentation prouvée par la publication d'informations</b>	27
 <b>Application partielle prouvée par la publication d'informations</b>	14
 <b>Progrès signalés</b>	84
 <b>Meta a signalé l'implémentation d'une action, ou décrit un travail en cours de réalisation, mais n'a publié aucune information attestant de cette implémentation</b>	29
 <b>Recommandation refusée après une étude de faisabilité</b>	10
 <b>Recommandation omise, refusée ou redéfinie</b>	23
 <b>En attente d'une réponse préliminaire de la part de Meta</b>	4
<b>NOMBRE TOTAL DE RECOMMANDATIONS</b>	<b>191</b>

# Sélection des cas

En 2022, **1 290 942** cas ont été soumis au Conseil,

*soit une augmentation d'environ un quart par rapport à 2021.*



En moyenne, en 2022, le Conseil a reçu un cas toutes les **24 secondes.**

Plus de deux tiers des appels d'utilisateurs visant à restaurer du contenu ne concernaient que deux Standards de la communauté :



la violence et l'incitation



le discours incitant à la haine

DANS PRÈS DE DEUX TIERS DES CAS SÉLECTIONNÉS PAR LE CONSEIL EN 2022,

**Meta a reconnu que sa décision d'origine sur le contenu était mauvaise.**



## Vue d'ensemble

Chaque jour, des centaines de millions de publications sont partagées sur Facebook et Instagram, et des dizaines de milliers de modérateurs de contenu prennent des décisions en une fraction de seconde sur le contenu qui doit être conservé ou supprimé. Par conséquent, l'un de nos plus grands défis consiste à choisir les cas à examiner. En 2022, nous avons reçu près de 1,3 million de demandes d'utilisateurs en vue d'examiner de manière indépendante des décisions de modération du contenu prises par Meta. Meta a également fait remonter 21 cas au Conseil. En 2022, nous avons reçu en moyenne 3 537 cas par jour.

Nous nous concentrons sur des cas qui soulèvent des questions délicates et lourdes de conséquences, et qui sont représentatifs d'un large éventail de régions géographiques et de thèmes. En sélectionnant des cas qui nous permettent de déterminer les politiques les plus importantes pour les utilisateurs, et en procédant à un examen minutieux et raisonné, nous pouvons améliorer les politiques pour l'ensemble des utilisateurs, et pas seulement pour ceux dont les cas ont été examinés. La sélection du cas concernant le dessin sur la police colombienne, par exemple, nous a donné l'occasion de faire des recommandations sur les banques de correspondance de médias de Meta, qui peuvent automatiquement supprimer les images susceptibles d'enfreindre les règles de Meta. De même, le cas du genre de musique UK drill a été l'occasion de faire des propositions sur la manière dont Meta devrait répondre aux demandes des autorités chargées de l'application de la loi dans le monde entier.

Afin d'aborder des problèmes touchant le monde entier, nous avons continué à sélectionner des cas provenant de différentes régions du monde. Nous avons également choisi des cas qui revêtent une importance cruciale pour le débat public, tels que le cas du poème russe, qui traite des questions d'expression dans le contexte de l'invasion de l'Ukraine par la Russie, et celui de la réappropriation des mots arabes, qui porte sur l'utilisation de termes péjoratifs à l'encontre des personnes LGBTQIA+.



**Nous nous efforçons de choisir des cas présentant un réel impact et abordant des problèmes auxquels l'entreprise est régulièrement confrontée. »**

**Michael McConnell**  
COPRÉSIDENT DU CONSEIL  
DE SURVEILLANCE



## ENSEIGNEMENTS TIRÉS

Fin 2022, lors de la préparation de notre décision relative au genre de musique UK drill et de notre avis consultatif en matière de politiques sur le programme de vérifications croisées de Meta, nous avons identifié un problème lié à la procédure d'appel de l'entreprise. En effet, dans certains cas d'appel, les utilisateurs n'ont pas eu la possibilité de faire appel auprès du Conseil, bien qu'ils soient admissibles à un examen conformément à nos statuts. Cette situation a suscité de vives inquiétudes quant à la transparence et au droit de recours des utilisateurs. Dans la décision relative au genre de musique UK drill et notre avis consultatif en matière de politiques sur le programme de vérifications croisées de Meta, nous avons demandé instamment à l'entreprise de rectifier cette erreur.

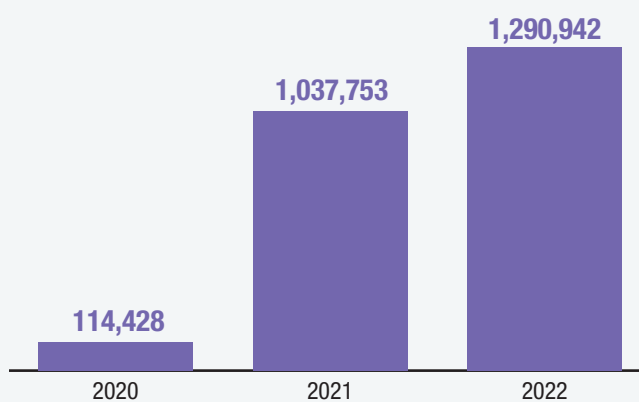
En janvier 2023, Meta a répondu en indiquant que les résidents de l'UE, du Royaume-Uni et de l'Inde pourraient bientôt faire appel des décisions relatives aux contenus éligibles prises lors de la remontée d'informations auprès de Meta et du Conseil. Les décisions prises « à la suite d'une remontée » sont prises par les équipes de spécialistes internes de Meta, plutôt que par des sous-traitants externes. Pour les utilisateurs de Facebook et d'Instagram dans d'autres pays, Meta a annoncé son intention de développer un autre mécanisme permettant aux utilisateurs de faire appel directement auprès de nous des décisions admissibles de la procédure de remontée qui ne peuvent pas faire l'objet d'un appel en interne. L'entreprise a indiqué qu'elle espérait mettre en œuvre cette solution d'ici le second semestre 2023.

## Cas soumis au Conseil

En 2022, le nombre d'appels que nous avons reçus de la part des utilisateurs a augmenté d'environ un quart, passant d'environ un million en 2021 à environ 1,3 million en 2022. Alors que les 1er et 2e trimestres 2022 ont connu une augmentation du nombre d'appels par rapport aux trimestres correspondants de 2021, la tendance s'est inversée pour les 3e et 4e trimestres 2022.

Au total, entre octobre 2020, date à laquelle nous avons commencé à accepter des cas, et décembre 2022, nous avons reçu plus de 2,4 millions d'appels, ce qui reflète une demande constante de la part des utilisateurs de faire appel des décisions de modération de contenu de Meta auprès d'un organisme indépendant. Bien que nous ne puissions examiner qu'un petit nombre de cas, nous continuons à sélectionner ceux qui soulèvent des problèmes sous-jacents auxquels sont confrontés de nombreux utilisateurs dans le monde et à formuler des recommandations pour y remédier.

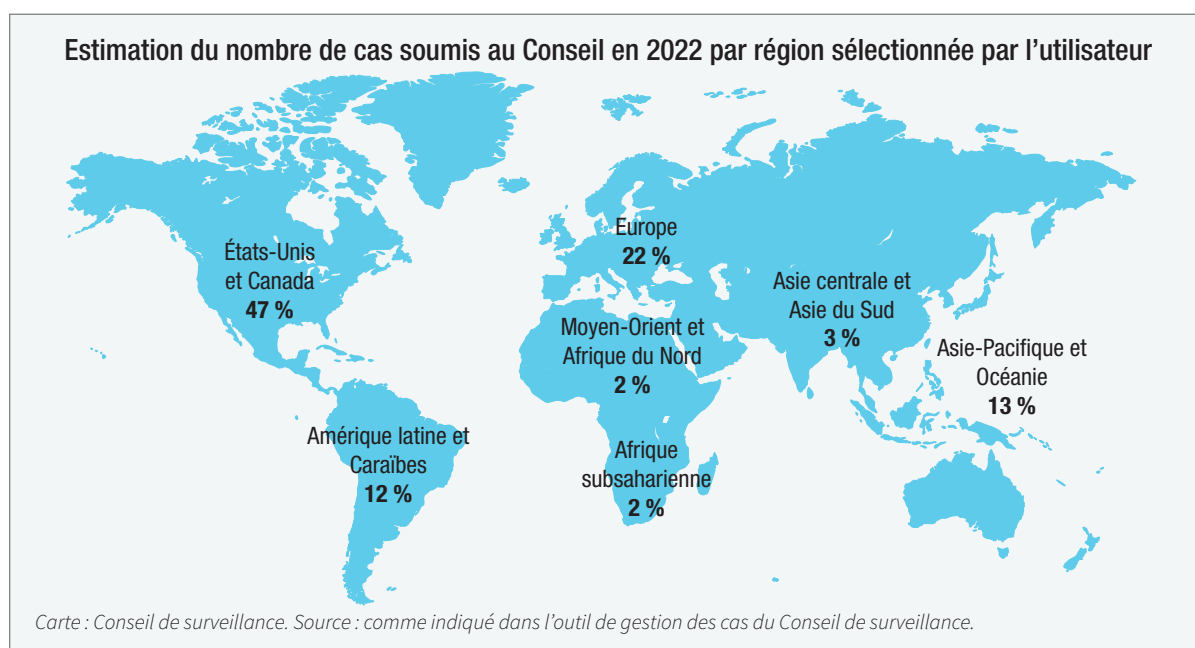
Nombre de cas soumis au Conseil de surveillance par an



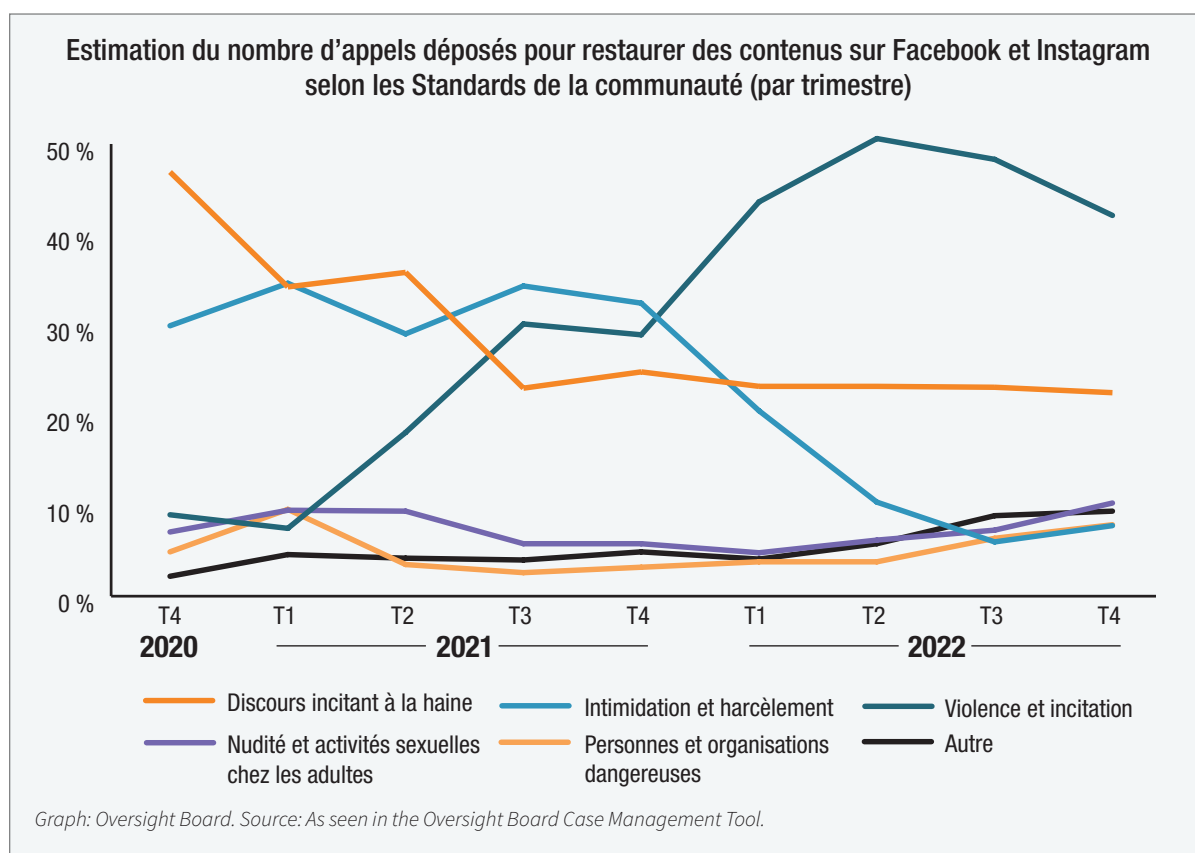
Graphique : Conseil de surveillance. Source : Comme indiqué dans l'outil de gestion des cas du Conseil de surveillance.

En 2022, la part des cas concernant des contenus sur Instagram a augmenté de manière significative. Alors qu'ils constituaient à peine 1 % des appels chaque trimestre jusqu'au 1er trimestre 2022, cette part est passée à plus de 10 % des appels au 4e trimestre.

En 2022, la répartition des cas par région sélectionnée par l'utilisateur est restée sensiblement inchangée par rapport à la période allant d'octobre 2020 à décembre 2021, avec plus de deux tiers des appels provenant des États-Unis et du Canada, ainsi que de l'Europe. La part des cas provenant des États-Unis et du Canada a légèrement diminué, passant de 49 % en 2020/2021 à 47 % en 2022. Pour ce qui est de l'Europe, elle est passée de 20 % à 22 % sur la même période. On observe également une augmentation notable de la part des cas provenant d'Asie-Pacifique et d'Océanie, passant de 9 % à 13 %. 12 % des cas provenaient d'Amérique latine et des Caraïbes en 2022, 3 % d'Asie centrale et du Sud, 2 % du Moyen-Orient et de l'Afrique du Nord, et 2 % de l'Afrique subsaharienne. Nous reconnaissons que ces chiffres ne reflètent pas la répartition des utilisateurs de Facebook et d'Instagram dans le monde ni la répartition réelle des problèmes de modération du contenu à travers le monde. En 2023, nous renforcerons nos activités de sensibilisation et d'engagement dans les régions autres que les États-Unis et l'Europe afin de mieux communiquer sur notre travail et d'encourager les utilisateurs de ces régions à soumettre des cas au Conseil.



En 2022, plus de deux tiers (69 %) des recours d'utilisateurs visant à restaurer du contenu ne concernaient que deux Standards de la communauté : violence et incitation (46 % des recours pour la restauration de contenu) et discours haineux (23 % de recours pour la restauration de contenu). Ces chiffres ne concernent que les recours d'utilisateurs pour la **restauration** de contenus que Meta considère comme étant contraire à ses règles (ce qui représente 92 % des recours en 2022) et non les recours pour **retrait** de contenus d'autres personnes encore visibles sur Facebook ou sur Instagram (ce qui représentait 8 % des recours en 2022), parce que ces derniers ne seraient pas contraires à un Standard de la communauté.



Comme le montre le graphique, en 2022, la part de recours d'utilisateurs pour une restauration de contenus retirés en vertu des Standards de la communauté de Meta en matière de violence et d'incitation a augmenté. Elle a atteint un pic de 51 % au deuxième trimestre 2022. Les recours d'utilisateurs pour la restauration de contenus supprimés en vertu de ses normes liées à l'intimidation et au harcèlement ont chuté, toutefois, d'environ 30 % au quatrième trimestre de l'année 2021 à 10 % durant la deuxième moitié de l'année 2022.

Les rapports de mise en application des Standards de la communauté de Meta soulignaient également une augmentation similaire de la quantité de contenus liés à la violence et à l'incitation traités sur Facebook en 2022, passant de 12,7 millions au quatrième trimestre de l'année 2021 à 21,7 millions au premier trimestre de l'année 2022 et à 19,3 millions au deuxième trimestre 2022. Cela peut, en partie, expliquer l'augmentation des recours passés auprès du Conseil. Le rapport de Meta montre également que, au T3 2022, le taux de contenus liés à l'intimidation et au harcèlement traités sur Facebook a chuté jusqu'à 6,6 millions, son plus bas niveau depuis près de deux ans. Cela peut expliquer en partie la réduction de la part de recours au conseil pour la restauration de contenus supprimés en vertu des Standards de la communauté.

## Cas pris en considération par le Comité de sélection des cas

Une fois que le Comité de sélection des cas du Conseil sélectionne des cas pour examen par le Conseil, Meta reconnaît parfois que sa décision d'origine sur une partie du contenu était infondée.

Dans 32 cas sur 50 sélectionnés en 2022 (64 %), Meta a considéré que sa décision initiale avait été incorrecte, contre 51 sur 130 cas (39 %) en 2020/2021. Cela représente une augmentation de 20 points de pourcentage.

Bien qu'il ne s'agisse que d'un petit échantillon et que le Conseil recherche volontairement les cas les plus complexes et difficiles, il est interpellant de constater que Meta a jugé sa décision infondée seulement dans près de deux tiers des cas. Cela soulève des questions plus larges à la fois sur la précision de la modération du contenu de Meta et sur la procédure de recours que Meta applique avant que les cas n'arrivent au Conseil. En 2023, nous commencerons à publier des décisions sommaires. Ces décisions présenteront l'examen de cas que nous n'avons pas sélectionnés pour une évaluation complète, mais qui, toutefois, entraîne l'annulation de l'action initiale de Meta.



# Décisions et avis consultatifs en matière de politiques

**12** décisions publiées en 2022

 dont 9 contraires à l'action de Meta.

 7 cas utilisateur  
5 cas de Meta

 8 cas de Facebook  
4 cas d'Instagram

Publication des premiers avis consultatifs en matière de politiques du Conseil, portant sur le partage d'informations privées et sur le programme de vérifications croisées de Meta.



Capacité à enjoindre Meta à supprimer ou à restaurer du contenu avec l'apposition d'un écran d'avertissement

**356 QUESTIONS POSÉES À META,**  
DANS LE CADRE DE NOS RÉVISIONS DE  
CAS EN 2022,



dont 86 % ont trouvé une réponse de la part de Meta

# Vue d'ensemble

Les décisions que nous prenons en réponse aux recours des utilisateurs et des recommandations de Meta sont au cœur du travail du Conseil. En 2022, nous avons aussi publié nos premiers avis consultatifs en matière de politiques, portant sur le partage d'informations privées sur le lieu de résidence et sur le programme de vérifications croisées de Meta.

## ||| DÉCISIONS

Nos décisions examinent la cohérence des choix de Meta de retrait ou de maintien de contenu avec les règles, les valeurs et les engagements établis par l'entreprise en matière des droits de l'homme. Nos décisions sont contraignantes pour Meta. Elle doit les mettre en œuvre sous 7 jours après leur publication, à moins que cela ne soit contraire à la législation. En 2022, nous avons publié 12 décisions, annulant la décision de Meta liée au contenu concerné dans 75 % des cas, soit une légère augmentation en comparaison avec le taux de 70 % en 2021.

Nous avons également apporté plusieurs modifications à notre processus de révision afin d'avoir un impact plus important. En juillet 2022, nous avons annoncé que nous tiendrions compte de deux cas concernant l'identité de genre et la nudité au sein d'une même décision, pour la première fois. Rassembler ces deux cas nous aidera à explorer une plus grande diversité d'expériences utilisateur et à comparer les publications qui ne s'intègrent pas dans une catégorie précise. En octobre 2022, nous avons étendu notre champ d'application et acquis la capacité à émettre des décisions contraignantes quant à la mise en place d'un écran d'avertissement indiquant que le contenu est sensible ou perturbant, lors de la restauration ou du maintien de contenu concerné. En novembre, notre décision sur le genre de musique « UK drill » a été publiée. Dans le cadre de cette décision, nous avons fait usage d'une demande de liberté d'information afin d'obtenir davantage de détails sur les requêtes des services de police de Londres auprès des réseaux sociaux pour le retrait de contenus.

## ||| AVIS CONSULTATIF EN MATIÈRE DE POLITIQUES

Au-delà de l'étude de cas individuels pour supprimer ou restaurer du contenu, nous pouvons aussi accepter des avis consultatifs en matière de politiques de la part de Meta. Ces avis nous permettent de fournir des recommandations détaillées à l'entreprise au regard des changements que Meta devrait apporter à ses politiques sur un sujet donné. Nous pensons que ces recommandations peuvent également être utiles à d'autres entreprises, qui rencontrent souvent des difficultés liées à la modération de contenus similaires à celles rencontrées par Meta.

En 2022, nous avons aussi publié nos premiers avis consultatifs en matière de politiques, portant sur le partage d'informations privées sur le lieu de résidence et sur le programme de vérifications croisées de Meta. Notre avis consultatif en matière de politiques sur le programme de vérifications croisées de Meta a inclus un nombre de détails encore jamais relevés sur le programme, et a effectué 32 recommandations à l'entreprise.

## COMMENT META A-T-ELLE RÉPONDU À NOS QUESTIONS

Pour nous aider à prendre des décisions et à rendre des avis consultatifs en matière de politiques, nous envoyons des questions à l'entreprise avant de délibérer sur les cas. Pour les 12 décisions et les deux avis consultatifs en matière de politiques publiés en 2022, nous avons posé à Meta 356 questions. Meta a répondu à 306 questions dans leur intégralité, à 37 questions partiellement et n'a pas répondu à 13 d'entre elles.

La part des questions que Meta a entièrement traitées en 2022, 86 %, était identique à celle traitée en 2021. La part de questions auxquelles Meta n'a pas répondu, toutefois, a chuté, passant de 6 % (19 questions sur 313) en 2021, à 4 % (13 questions sur 356) en 2022. Nous trouvons que les réponses de Meta pour l'année 2022 étaient plus complètes qu'en 2021. Lorsqu'elle n'a pas pu répondre, Meta s'est expliquée à de plus nombreuses reprises qu'en 2021.

### ENSEIGNEMENTS TIRÉS

En 2022, nos statuts prévoyaient un délai de publication des décisions sous 90 jours avant l'annonce du cas sur notre site Web. Les statuts ont également prévu des exceptions à ce délai, par exemple en cas de circonstances exceptionnelles ou d'incidents opérationnels ou techniques. En 2022, il y a eu davantage d'exceptions que de décisions ponctuelles. Dans certains cas, les négociations avec Meta sur la quantité d'informations que nous pouvions inclure dans notre décision finale (indiquée par l'entreprise de façon confidentielle) ont duré plus longtemps que prévu. Dans d'autres cas, des soucis de traduction ont retardé la publication. À d'autres occasions, les contraintes d'emploi du temps du Conseil et d'autres problèmes nous ont empêchés de publier nos décisions sous 90 jours. Nous avons toujours pour objectif de rendre nos décisions sous 90 jours (conformément à notre mise à jour de statuts de février 2023) et de nous montrer transparents en ce qui concerne les délais liés aux cas.

Le Conseil a publié deux avis consultatifs en matière de politiques en 2022. En mai 2022, nous avons également annoncé que Meta avait fait part au Conseil de son intention de retirer une précédente demande d'accompagnement concernant les soucis de modération de contenus en lien avec la guerre entre la Russie et l'Ukraine. Cette action était liée à des inquiétudes de l'entreprise portant sur la sécurité. Si nous comprenons ces inquiétudes, nous étions d'avis que la demande soulevait d'importantes problématiques et nous étions déçus par la décision de retrait de Meta. Toutefois, nous avons ultérieurement sélectionné et publié notre décision sur les poèmes russes, qui recouvrait diverses problématiques de modération de contenu en lien avec l'invasion de l'Ukraine par la Russie.



# Décisions et avis consultatifs en matière de politiques publiés en 2022

## DÉCISIONS

	PLATEFORME	SOURCE	STANDARD DE LA COMMUNAUTÉ	PAYS	DÉCISION DU CONSEIL
<b>Demande d'Adderall®</b> Cas 2021-015-FB-UA		Utilisateur	Biens et services restreints	États-Unis	<b>Décision de suppression de contenu</b> de Meta infirmée
<b>Journaliste suédois signalant des violences sexuelles sur mineurs</b> Numéro du cas : 2021-016-FB-FBR		Meta	Nudité et activités sexuelles chez les adultes	Suède	<b>Décision de suppression de contenu</b> de Meta infirmée
<b>Dessin Knin</b> Numéro du cas : 2022-001-FB-UA		Utilisateur	Discours incitant à la haine	Croatie	<b>Décision de maintien de contenu</b> de Meta infirmée
<b>Vidéo au Soudan</b> Numéro du cas : 2022-002-FB-MR		Meta	Contenu violent et explicite	Soudan	<b>Décision de maintien du contenu</b> de Meta confirmée
<b>Réappropriation de mots arabes</b> Numéro du cas : 2022-003-IG-UA		Utilisateur	Discours incitant à la haine	Maroc, Égypte, Liban	<b>Décision de suppression de contenu</b> de Meta infirmée
<b>Dessin sur la police colombienne</b> Numéro du cas : 2022-004-FB-UA		Utilisateur	Personnes et organisations dangereuses	Colombie	<b>Décision de suppression de contenu</b> de Meta infirmée
<b>Mention des talibans dans les actualités</b> Numéro du cas : 2022-005-FB-UA		Utilisateur	Personnes et organisations dangereuses	Afghanistan	<b>Décision de suppression de contenu</b> de Meta infirmée
<b>Affaire du Tigray Communication Affairs Bureau</b> Numéro du cas : 2022-006-FB-MR		Meta	Violence et incitation	Éthiopie	<b>Décision de suppression de contenu</b> de Meta confirmée
<b>Musique UK drill</b> Numéro du cas : 2022-007-IG-MR		Meta	Violence et incitation	Royaume-Uni	<b>Décision de suppression de contenu</b> de Meta infirmée
<b>Poème russe</b> Numéro du cas : 2022-008-FB-UA		Utilisateur	Discours incitant à la haine	Lettonie, Ukraine, Russie	<b>Décision de suppression de contenu</b> de Meta infirmée
<b>Vidéo après l'attaque visant une église au Nigéria</b> Numéro du cas : 2022-011-IG-UA		Utilisateur	Contenu violent et explicite	Nigéria	<b>Décision de suppression de contenu</b> de Meta infirmée
<b>Vidéo de harcèlement sexuel en Inde</b> Numéro du cas : 2022-012-IG-MR		Meta	Exploitation sexuelle d'adultes	Inde	<b>Décision</b> de Meta de garder le contenu confirmée

## AVIS CONSULTATIFS EN MATIÈRE DE POLITIQUES

**Partage d'informations confidentielles sur la résidence**

Numéro du cas : PAO-2021-01

**Programme de vérifications croisées de Meta**

Numéro du cas : PAO-2021-02

# Synthèses des décisions et avis consultatifs en matière de politiques

En octobre 2022, nous avons annoncé sept priorités stratégiques découlant d'une analyse approfondie des points soulevés par les utilisateurs dans leurs recours auprès du Conseil. Cette section regroupe les décisions et les avis consultatifs en matière de politiques publiés en 2022, conformément à ces priorités.

## Élections et espace civique



### Mention des talibans dans les actualités

### ANNULATION DE LA DÉCISION

En janvier 2022, une revue indienne en langue ourdoue a publié sur sa page Facebook une annonce d'un représentant officiel des talibans en Afghanistan, affirmant que les écoles et collèges pour les femmes et les filles rouvriraient sous deux mois. Meta a retiré cette publication, considérant qu'elle était contraire aux normes de Facebook sur les individus et organisations dangereux, qui interdit l'apologie d'entités considérées comme impliquées dans des dommages graves hors ligne, y compris les organisations terroristes. Meta a également appliqué des sanctions à l'encontre de l'administrateur de la page et limité l'accès du compte à quelques fonctionnalités Facebook uniquement.

Après qu'un modérateur a fait appel de la décision de Meta, un deuxième examen manuel a confirmé que la publication était contraire aux règles, conformément à l'annonce initiale de Facebook. Après la sélection du cas pour examen, Meta a reconnu que la décision initiale avait été une erreur de mise en application, a restauré le contenu et retiré les sanctions, puis annulé les restrictions du compte. Malgré le changement de décision de Meta, nous avons annulé la décision initiale de retrait de cette annonce positive du régime taliban. Nous avons cité une exception spécifique au Standard de la communauté relatif aux personnes et organisations dangereuses permettant la publication d'actualités utiles sur les activités de groupes terroristes. Nous avons également trouvé que le lien entre l'autorisation de diffusions dans la politique relative aux individus et aux organisations dangereux et la notion d'intérêt médiatique, trop générale, était flou.

« Le Conseil a décidé que, même si les talibans constituaient une organisation dangereuse, les actualités à leur sujet ne devraient pas être restreintes pour des raisons d'intérêt public. »

Paolo Carozza  
MEMBRE DU CONSEIL DE SURVEILLANCE



**Quels Standards de la communauté nos décisions ont-elles examinés le plus en 2022 ?**

**Discours incitant à la haine**

 3 décisions

**Personnes et organisations dangereuses**

 2 décisions

**Violence et incitation**

 2 décisions

**Contenu violent et explicite**

 2 décisions



## **Journaliste suédois signalant des violences sexuelles sur mineurs**

### **ANNULATION DE LA DÉCISION**

En août 2019, un utilisateur en Suède a publié une photo issue d'une base d'images, sur sa page Facebook, représentant une jeune fille assise, se tenant la tête dans les mains, ce qui assombrit son visage. La photo était accompagnée de descriptions détaillées des viols de deux mineures anonymes. La publication donnait les âges de deux responsables anonymes inculpés et incluait des détails explicites des conséquences douloureuses du crime sur l'une des victimes. Elle comprenait également des citations, dont l'une attribuée à l'un des responsables, se vantant du viol et évoquant l'une des victimes en des termes explicites à caractère sexuel. En septembre 2021, après avoir maintenu le contenu sur la plateforme durant deux ans, Meta a retiré la publication pour violation de la politique en matière d'exploitation sexuelle, de nudité et d'abus d'enfants.

Nous avons estimé que la décision de Meta de retrait de la publication n'était pas légitime. La fourniture d'une description clinique des conséquences d'un viol, ainsi que l'affirmation sexuellement explicite du responsable du viol à ce sujet, n'ont pas constitué un message faisant l'apologie de l'exploitation sexuelle d'enfants ou décrivant des mineurs dans un contexte sexuel. Le contexte plus large de la publication a clairement confirmé que l'intention de l'utilisateur était de signaler une problématique d'intérêt public évident tout en condamnant l'exploitation sexuelle d'une mineure. Parmi nos recommandations, dans ce cas précis, nous avons enjoint à Meta de faire une claire distinction entre le contenu qui fait la promotion de l'exploitation sexuelle d'enfant ou la cautionne, et le contenu qui sensibilise à ce sujet.

## **Situations de crise et de conflit**



## **Affaire du Tigray Communication Affairs Bureau**

### **MAINTIEN DE LA DÉCISION**

Au début de l'année 2022, Meta a recommandé un cas concernant une publication qui était visible sur la page Facebook du Communication Affairs Bureau dans la région du Tigré, un organe géré par les autorités d'une province éthiopienne. Publiée durant le conflit entre les forces fédérales et celles du Tigré dans cette région, la publication faisait référence à des pertes récentes déclarées par les forces fédérales et appelait l'armée nationale à prendre pour cible le groupe Abiy Ahmed, une référence au Premier ministre d'Éthiopie. Elle affirmait ensuite que si les forces fédérales refusaient de s'y conformer, elles trouveraient la mort.

Après un signalement par plusieurs utilisateurs et son identification par les systèmes automatisés de Meta, deux réviseurs parlant l'amharique ont estimé que la publication

n'était pas contraire aux politiques de Meta. Deux jours plus tard, après un transfert du contenu auprès d'un centre pour l'intégrité des opérations produit (Integrity Product Operations Center, IPOC) mis en place par l'entreprise pour la modération du contenu lié au conflit, afin d'obtenir une expertise, Meta a annulé sa décision initiale et retiré la publication pour violation de la politique sur la violence et l'incitation de Facebook.

Tout en ne niant pas les démarches positives effectuées par Meta, y compris la mise en place du système IPOC, afin de surveiller le contenu pour lutter contre les abus en cas de situation de conflit à haut risque, notre décision a souligné que le contenu contraire aux règles était resté sur la plateforme durant deux jours et avait été consulté plus de 300 000 fois avant son retrait. Pour le moins, cela reflétait la nature temporaire et court-termiste du recours à un IPOC en matière de conflit sur la durée. Nous avons alors recommandé un investissement dans un mécanisme plus durable de la part de Meta afin que l'entreprise puisse respecter ses responsabilités concernant les droits de la personne dans des zones de conflits et des situations de crise.



## Vidéo au Soudan

## MAINTIEN DE LA DÉCISION

Durant les dernières semaines de l'année 2021, Meta a recommandé un cas concernant une vidéo explicite publiée sur un profil Facebook après le coup d'État militaire au Soudan survenu en octobre 2021. Une personne y était filmée étendue à côté d'une voiture, gravement blessée à la tête, l'œil visiblement détaché. À l'arrière-plan, on pouvait entendre des voix disant, en arabe, qu'une personne avait été battue et laissée ainsi dans la rue. Une légende, également formulée en arabe, invitait les populations à se méfier des militaires, en mentionnant des abus récurrents et de longue date de la part des dites forces.

Après que Meta a retiré la publication pour infraction aux Standards de la communauté sur les contenus vidéo explicites publiés sur Facebook, l'entreprise a mis en place un qualificatif d'intérêt médiatique pour ce contenu. Près de cinq semaines plus tard, Meta a restauré le contenu et y a apposé un écran d'avertissement qui permettait d'en restreindre l'audience.

Notre décision validait la décision de Meta de restaurer le contenu avec l'apposition d'un écran d'avertissement. Nous avons également estimé que l'application du qualificatif d'intérêt médiatique ne constituait pas un moyen efficace de modérer les contenus explicites partagés sur Facebook à grande échelle. Pour souligner cette conclusion, nous avons cité la déclaration de Meta indiquant avoir mis en place 17 qualifications d'intérêt médiatique en lien avec la politique sur le contenu violent et explicite durant les trois premiers trimestres de l'année 2021. En comparaison, l'entreprise avait retiré plus de 90 millions de contenus pour violation dudit Standard sur la même période. Nous avons, par conséquent recommandé à Meta de revoir ses Standards de la communauté en matière de contenu explicite et violent afin de permettre le partage de vidéos explicites lorsque l'intention est de sensibiliser ou de relayer des informations utiles sur des abus.



## Discours incitant à la haine envers des groupes marginalisés



### Poème russe

### ANNULATION DE LA DÉCISION

En avril 2022, un utilisateur de Facebook en Lettonie a publié l'image d'un corps sans vie étendu dans une rue. Meta a confirmé qu'il s'agissait d'une personne tuée en Ukraine. Le texte en Russe l'accompagnant indiquait que, durant la Seconde Guerre mondiale, les atrocités qui auraient été commises par les soldats soviétiques en Allemagne avaient été légitimées en tant que vengeance pour des crimes qui auraient été commis par les soldats nazis en URSS. Faisant référence aux atrocités que les soldats russes auraient commises en Ukraine, le texte citait un poème du Soviétique Konstantin Simonov [traduction en français] : « Il faut tuer le fasciste... Tuez-le ! »

À la suite du retrait de la publication pour violation du Standard de la communauté sur les discours incitant à la haine, nous avons sélectionné le cas pour procéder à sa révision. Meta a ensuite restauré le contenu et y a apposé un écran d'avertissement. Notre décision faisait la distinction entre le contenu ayant ciblé la **nationalité** des soldats russes, ce qui aurait constitué une violation, et le contenu qui effectue un **parallèle historique** avec les actes des nazis, ce qui n'est pas contraire aux règles. La citation d'un poème de la période soviétique nous a paru être une référence artistique et culturelle utilisée pour **décrire**, et non **encourager**, une pensée. Si nous reconnaissons la complexité de l'évaluation d'un discours violent en situation de conflit, dans les cas où la loi internationale permet le ciblage de combattants, nous recommandons une révision des politiques de Meta en faveur d'une clarification et une contextualisation en lien avec les interventions militaires illégitimes.

### Dessin Knin

### ANNULATION DE LA DÉCISION

En décembre 2021, une page Facebook publique a publié une vidéo inspirée du dessin animé Disney « The Pied Piper » qui comprenait une légende, en croate, traduite par Meta comme suit : « L'artiste de Čavoglave », un village de Croatie, « et les rats de Knin », une ville en Croatie. Le narrateur indiquait que lorsqu'une population de rats décidait de vivre dans un « pays dédié aux rongeurs », elle commençait à harceler et à persécuter les populations, qui finissaient par se rendre. Après qu'un joueur de flûte de Čavoglave est apparu jouant une mélodie sur son « instrument magique », les rats ont été attirés hors de la ville, et ont disparu à bord d'un tracteur. Conclusion du narrateur : « les rats ont disparu à tout jamais de ces terres [et] tout le monde vécut heureux et eut beaucoup d'enfants. »

En dépit de près de 400 recours d'utilisateurs, Meta n'a pas supprimé le contenu. Après un recours contre la décision de Meta auprès du Conseil, Meta a effectué un deuxième examen manuel, qui a confirmé que la publication n'était pas contraire aux règles. Dès lors que nous avons sélectionné le cas pour l'examiner, Meta a annulé cette décision, déterminant que si la publication n'était pas contraire aux règles *du point de vue de la forme*, c'était bien le cas *du point de vue du fond*, au regard de la politique sur les discours incitant à la haine, et l'a supprimée. En élaborant son explication à l'attention du Conseil, Meta a de nouveau changé d'avis, estimant, cette fois, que la publication était bien contraire à cette politique *du point de vue de la forme*.

Pour appuyer notre avis selon lequel le contenu était contraire aux Standards de la communauté en matière de discours incitant à la haine et de violence et d'incitation, nous avons cité des commentaires faisant référence à une opération militaire en Croatie réalisée en 1995, l'opération « Tempête ». Cette opération a entraîné le déplacement forcé, l'exécution et la disparition de civils d'origine serbe en Croatie. Au regard de ces références, nous avons fait part de notre inquiétude quant au fait que 40 modérateurs parlant le croate avaient considéré le contenu comme conforme, car ils pensaient que les standards nécessitaient que la comparaison insultante des personnes d'origine serbe avec des rats soit *explicite*. Nous avons recommandé une clarification de la part de Meta concernant les Standards de la communauté en matière de discours incitant à la haine, qui devra indiquer que la politique interdit les références hostiles tant *implicites* qu'*explicites* concernant les groupes protégés.



## Réappropriation de mots arabes

## ANNULATION DE LA DÉCISION

En novembre 2021, un compte Instagram public a fait la description d'un espace dédié aux échanges sur les histoires queers dans la culture arabe, et a publié des images avec une légende en arabe indiquant que chaque image contenait un mot utilisé dans le monde arabophone pour dénigrer les hommes aux « manières efféminées ». L'utilisateur a affirmé que son intention était de reprendre le pouvoir en se réappropriant ces termes.

Après le retrait du contenu par Meta pour violation de sa politique en matière de discours incitant à la haine, la restauration dudit contenu après contestation de l'utilisateur, et un nouveau retrait après une autre contestation, nous avons sélectionné le cas pour l'examiner. Meta a renvoyé le contenu pour une révision supplémentaire, dont la conclusion a été que le contenu n'était pas contraire à la politique. Meta l'a ensuite restauré, indiquant que ses décisions initiales s'étaient appuyées uniquement sur un examen des images contenant les termes péjoratifs, non du contenu en soi ni des explications associées.



**Meta a des valeurs et des standards en place. La mise en application, elle, fait défaut. »**



**Endy Bayuni**  
MEMBRE DU CONSEIL  
DE SURVEILLANCE

Pour justifier notre considération du retrait du contenu comme incohérente avec la politique en matière de discours incitant à la haine, nous avons souligné que le contenu faisait l'objet d'une exception permettant le partage de tels discours lorsqu'« utilisés pour faire référence à soi-même ou pour reprendre le pouvoir face aux discriminations ». Citer des discours incitant à la haine dans l'intention de les condamner ou de sensibiliser l'audience à ces propos est autorisé. Les allers-retours dans les décisions de Meta ont, dans ce cas, reflété, selon nous, une mise en application incohérente des exceptions de la politique sur les discours incitant à la haine en ce qui concerne l'expression des personnes marginalisées. Pour faire face à cette problématique, nous avons recommandé la traduction en arabe standard moderne des Standards de mise en application interne et de la FAQ que Meta transmet à ses modérateurs.

## Utilisation des plateformes Meta par les gouvernements



### Musique UK drill

### ANNULATION DE LA DÉCISION

En janvier 2022, un compte Instagram décrit comme promouvant la musique britannique a publié l'extrait du clip d'un morceau de musique UK drill du rappeur Chinx (OS), intitulé « Secrets Not Safe ». La musique UK drill est un sous-genre régional du rap, populaire chez les jeunes personnes citadines noires au Royaume-Uni. Peu après, les services de police du Grand Londres ont envoyé un e-mail à Meta pour demander que l'entreprise examine l'ensemble des contenus associés au morceau. La police a envoyé des informations supplémentaires concernant la prévalence de la violence des gangs à Londres et fait part d'inquiétudes quant à l'influence négative de ce morceau sur le risque de représailles violentes au sein des gangs.

En réponse à cette demande, une équipe experte de chez Meta, en s'appuyant sur le contexte fourni par la police, a déterminé que la publication constituait une « menace voilée ». L'entreprise a fondé cette conclusion sur une référence à une tuerie perpétrée en 2017 qui, selon elle, augmente les potentielles répercussions que pourrait avoir le morceau en matière d'incitation à la violence. Meta a non seulement retiré le contenu du compte, considérant qu'il était contraire à sa politique en matière de violence et d'incitation, mais a également retiré plus de 50 contenus supplémentaires contenant le morceau, qui avaient été publiés par d'autres comptes, y compris celui de l'artiste. Les systèmes automatisés de Meta ont, par la suite, retiré des contenus incluant le morceau plus de 112 fois.

Après transmission du cas au Conseil par Meta, nous avons demandé à examiner la publication initiale de Chinx (OS). Meta a indiqué que les mesures adoptées pour retirer la vidéo et la chanson de la plateforme avaient entraîné la suppression du compte de l'artiste. Pour savoir si les mesures de Meta étaient cohérentes avec ses standards, ses valeurs et ses responsabilités en matière de droits de la personne, le contexte de la publication était un élément crucial.

« Les rappeurs évoquent souvent de façon granulaire les conflits violents perpétuels dans la rue, en utilisant la première personne et des images ainsi que des paroles dépeignant des actes violents. L'appropriation éventuelle d'actes violents et une forme de surprovocation sont considérées comme étant des parties intégrantes du genre : une forme d'expression artistique qui mêle faits et fiction.

Ces revendications permettent aux artistes de gagner en popularité et d'être reconnus dans ce style. La question de savoir si la musique drill entraîne une certaine violence dans le monde réel ou non est débattue, en particulier la fiabilité des preuves utilisées dans le cadre du débat. »

Dans le cadre de l'évaluation de la cohérence entre la décision de Meta et ses règles, valeurs et responsabilités en matière de droits de la personne, nous avons mentionné un manque de preuves qualifiant le contenu comme une menace crédible. Nous avons ensuite estimé qu'en l'absence de telles preuves, Meta aurait dû appuyer davantage l'aspect artistique du contenu. Notre décision indiquait que si l'autorité d'application de la

loi peut parfois fournir du contexte et une certaine expertise, ce n'est pas pour autant que chaque contenu qu'elle considère comme devant être retiré doit l'être.

Pour l'évaluation de telles demandes à l'avenir, nous avons enjoint Meta à prendre des mesures concrètes afin de faire examiner ces requêtes de façon distincte, surtout lorsqu'elles concernent l'expression artistique d'individus appartenant à des minorités ou à des groupes marginalisés pour lesquels le risque de préjugé culturel contre leur contenu est plus élevé.

Notre décision a également évoqué un manque de transparence quant aux processus de Meta pour répondre aux requêtes des autorités gouvernementales et des représentants de la loi. Nous avons repris les informations reçues en réponse à notre demande en vertu de la liberté d'information auprès de la police de Londres, révélant que l'ensemble des 286 requêtes envoyées par la police à l'attention des réseaux sociaux et des services de streaming pour le retrait de contenu musical durant l'année, jusqu'en mai 2022, était lié à la musique UK drill. Le 4 janvier 2023, toutefois, la police nous a recontactés afin de corriger des erreurs dans sa réponse, observées par la suite. Les services de police avaient, en réalité, envoyé 992, et non 286, requêtes de ce type durant l'année, et jusqu'à mai 2022, toutes liées à la musique UK drill.

## Traitement équitable des utilisateurs



### **Avis consultatif en matière de politiques sur le partage d'informations privées sur le lieu de résidence**

En février 2022, nous avons publié notre premier avis consultatif en matière de politiques. Ces avis sont utilisés par le Conseil pour étudier les politiques de Meta et émettre des recommandations en vue d'améliorations. Dans ce cas précis, Meta a demandé au Conseil une évaluation détaillée des modifications qu'il prévoit en ce qui concerne ses Standards de la communauté pour les violations de la vie privée, qui interdisent le partage d'informations d'identification sur soi-même et sur les autres, sauf exception. Selon cette politique, les informations d'identification, y compris les adresses personnelles, ne doivent pas être partagées, sauf dans les cas où elles sont partagées ou sollicitées à des fins caritatives, par exemple la recherche d'une personne, d'animaux ou d'objets disparus, ou une mise en relation avec des fournisseurs de service.

Pour protéger la vie privée des utilisateurs, la politique interdit le partage d'images représentant une vue extérieure de résidences privées, en particulier si la résidence est le foyer d'une seule famille, que son numéro est identifiable sur l'image, que le résident s'oppose à l'exposition de sa résidence privée ou si des manifestations contre le résident sont organisées. La seule exception concerne l'utilisation de la résidence comme une ambassade gouvernementale, qui pourrait faire l'objet de protestations.

Une problématique sous-jacente que Meta nous a demandé d'aborder, c'est l'avantage que peut constituer l'accès à de telles informations pour les journalistes et les militants. À l'inverse, l'exposition de telles informations sans le consentement des résidents pourrait créer un risque pour leur sécurité et entraver leur droit à la vie privée. Parmi les préjudices potentiels réels pour la sécurité et la vie privée des utilisateurs, Meta a cité la révélation d'informations confidentielles en ligne, ici, la diffusion non autorisée de documents dans



l'intention de révéler des informations personnelles aux personnes qui pourraient en abuser. Comme l'a souligné Meta, cette pratique a des conséquences négatives dans le monde réel, y compris du harcèlement, notamment la traque d'une personne, également appelée « stalking », la violence, et des meurtres.

La politique prévoyait des exceptions conformément auxquelles les informations privées sur la résidence pouvaient être publiées, y compris lorsqu'elles sont déjà publiques, car diffusées par les actualités, jugées par les tribunaux, font l'objet de communiqués de presse ou sont diffusées via d'autres sources. Dans ses directives internes à l'attention des modérateurs de contenu, Meta indique que les informations précédemment publiées par au moins cinq médias d'actualité ne sont plus considérées comme privées en vertu de la politique.

Nous corroborons l'affirmation de Meta selon laquelle des préjudices sérieux peuvent résulter de violations du droit à la confidentialité de l'adresse personnelle, et que de tels préjudices peuvent avoir un impact disproportionné sur les femmes, les enfants et les personnes LGBTQIA+. Nous approuvons également les propos de Meta en ce qui concerne la nécessité de ressources et d'un travail conséquent pour fournir un accès physique à des informations et à des données disponibles au public, et son parallèle avec la simplification du partage desdites informations grâce aux plateformes numériques, à grande échelle. Au regard de la gravité et de l'importance de tels préjudices, nous avons recommandé à Meta de limiter les circonstances autorisant le partage d'informations disponibles au public.



### **Avis consultatif en matière de politiques concernant le programme de vérifications croisées de Meta**

---

En octobre 2021, du fait de divulgations du *Wall Street Journal* concernant le programme de vérifications croisées de Meta, nous avons accepté une demande de l'entreprise concernant l'examen du programme et la communication de recommandations pour l'améliorer. Dans ce contexte, Meta a indiqué qu'elle effectuait près de 100 millions de tentatives de modération de contenu chaque jour. Avec ce taux, même si 99 % des décisions de modération étaient justes, l'entreprise ferait toujours environ 1 million d'erreurs par jour.

Le volume et la complexité des contenus publiés sur Facebook et Instagram posent de vraies difficultés pour l'élaboration d'un système qui permettrait à Meta d'observer ses engagements en matière de droits de la personne. Tout en étant conscients de cela, nous avons souligné que, tel qu'il était structuré, le système de vérifications croisées n'était pas conforme aux engagements de l'entreprise en matière de droits de la personne. Pendant des années, les vérifications croisées ont permis le maintien de contenus d'une sélection de personnalités politiques, de partenaires commerciaux, de célébrités, entre autres, sur Facebook et Instagram durant plusieurs jours, alors que ce même contenu aurait été supprimé sans délai dans d'autres circonstances. Meta nous a indiqué que, en moyenne, la prise d'une décision concernant le contenu publié par les utilisateurs sur ses listes de vérifications croisées pouvait prendre plus de 5 jours.

Meta, pour sa part, a maintenu que l'objectif initial du programme était de progresser au regard de ses engagements pour les droits de la personne. Un examen plus rigoureux nous a permis de constater que le programme semblait davantage structuré pour répondre à des problématiques commerciales. Parmi les défaillances relevées dans le cadre de notre examen, que nous avons enjoint à l'entreprise de résoudre au plus vite, nous avons constaté qu'en offrant une meilleure protection contre le retrait de contenu à certains utilisateurs plutôt qu'à d'autres, l'entreprise faisait naître des inquiétudes à grande échelle quant à un éventuel privilège accordé à des célébrités ainsi qu'à des responsables politiques et commerciaux, au détriment de millions d'autres personnes.

Si la publication d'un utilisateur présent sur les listes de vérifications croisées de Meta est identifiée comme contraire aux règles de l'entreprise, elle reste visible sur la plateforme dans l'attente d'un examen supplémentaire. Meta applique ensuite l'ensemble de ses politiques à la publication, y compris les exceptions et les clauses relatives aux éléments contextuels, cela augmentant probablement ses chances de rester visible sur la plateforme. Les utilisateurs lambdas, en comparaison, ont bien moins de chances de voir leur contenu arriver sous les yeux d'un réviseur pouvant mettre en application l'ensemble des règles de Meta.

Cette inégalité de traitement a été amplifiée par le manque de transparence relatif aux critères utilisés pour générer la liste d'entités dont les contenus bénéficiaient d'une protection supplémentaire. Si Meta dispose de critères clairs pour l'ajout de partenaires commerciaux et de responsables politiques à ses listes de vérifications croisées, les groupes dont les publications sont importantes du point de vue des droits de la personne, tels que les journalistes ou les représentants de la société civile, disposent de conditions plus floues.

Ce manque de transparence inclut également la non-divulgence de la part de Meta des indicateurs principaux utilisés pour mesurer l'efficacité de son programme. Meta n'a fourni aucune preuve de sa comparaison entre les décisions issues de vérifications croisées et celles issues de ses mécanismes habituels de contrôle qualité. Et cela va jusqu'à l'incapacité de l'entreprise à informer ses utilisateurs de leur présence sur les listes de vérifications croisées, à divulguer ses procédures de création et de vérification desdites listes, et à divulguer les cas dans lesquels les entités qui publient des contenus contraires aux règles à répétition sont consignées sur des listes à partir de leur profil.

Nous avons effectué 32 recommandations pour traiter ces problématiques. Nous avons recommandé à Meta de prioriser l'expression importante pour les droits de la personne dans ses listes de vérifications croisées, y compris les expressions d'importance publique spéciale. Nous avons également appelé l'entreprise à faire preuve de davantage de transparence concernant le programme et son fonctionnement, et avons effectué des propositions afin de réduire le préjudice causé par les contenus contraires aux règles maintenus durant leur examen renforcé. Meta a répondu à nos recommandations en mars 2023, s'engageant à faire preuve de davantage de transparence et à offrir des protections plus importantes à ceux qui connaissent un risque plus élevé d'une mise en application exagérée, tels que les journalistes et les défenseurs de droits humains.



## Demande d'Adderall®

## ANNULATION DE LA DÉCISION

En juin 2021, un utilisateur de Facebook aux États-Unis se disant atteint du trouble de déficit de l'attention avec ou sans hyperactivité (TDAH) a demandé, dans une publication sur un groupe privé, comment il devait aborder la question d'un traitement médicamenteux avec son médecin. L'utilisateur indiquait s'être fait prescrire du Xanax, mais que l'Adderall avait bien fonctionné par le passé. L'utilisateur a fait part de ses inquiétudes quant au fait de demander un changement de traitement, ce qui pourrait passer pour un comportement addictif.

En août 2021, Meta a supprimé la publication pour une prétendue violation des Standards de la communauté en matière de biens et services restreints sur Facebook. Après que nous avons sélectionné ce cas, Meta a invoqué une erreur d'application et l'a remise en ligne.

Dans notre décision, nous avons estimé que la décision initiale de Meta quant au retrait de la publication était erronée. Les Standards de la communauté de Facebook sur la restriction des biens et des services n'interdisent pas aux utilisateurs de chercher des conseils sur des produits pharmaceutiques spécifiques en cas de pathologies. Nous avons également estimé que les définitions publiques des substances indiquées dans les Standards étaient opaques pour les utilisateurs, car même selon ses définitions propres, l'Adderall et le Xanax peuvent être considérés ou non comme des produits pharmaceutiques selon les circonstances. Dans nos recommandations, nous avons enjoint à l'entreprise de vérifier les appels d'utilisateurs en temps opportun, lorsque les mesures prises à l'échelle d'un contenu entraînent des sanctions pour le compte lui-même.

## Application automatisée des politiques et de la curation de contenu



## Dessin sur la police colombienne

## ANNULATION DE LA DÉCISION

En septembre 2020, un utilisateur de Facebook en Colombie a publié une image ressemblant à l'écusson officiel de la Police nationale de Colombie. Trois personnages en uniforme y étaient représentés tenant des matraques au-dessus de leurs têtes, semblant frapper et donner des coups à un individu au sol qui saigne de la tête. La traduction de l'écusson sur Meta, en espagnol, est la suivante : « Police nationale – République de Colombie – Matraque et coups. »

Seize mois après la publication du contenu, Meta l'a supprimé après avoir établi une correspondance entre l'image et une image similaire au sein de l'une de ses bases de services pour la correspondance de médias. Ces bases peuvent identifier et supprimer automatiquement les images identifiées lors d'examen manuels comme enfreignant la politique de l'entreprise. Après que nous avons sélectionné ce cas, Meta a déterminé que la publication ne contrevenait pas à sa politique et l'a remise en ligne. L'entreprise a également republié les autres contenus contenant



**Les décisions relatives au contenu et le traitement de celui-ci par algorithme ne peuvent faire l'objet de processus distincts. »**

**Catalina Botero-Marino**  
COPRÉSIDENTE DU CONSEIL  
DE SURVEILLANCE



ledit dessin, et a reconnu que leur retrait des bases de services de mise en correspondance du contenu multimédia.

Ce cas a mis en lumière l'impact amplifié des décisions erronées de retrait de contenus par de telles bases. En dépit de 215 recours de la part des utilisateurs contre ces suppressions, dont 98 % ont abouti, Meta n'a pas retiré le dessin de sa base ni restauré les contenus similaires, jusqu'à ce que nous nous emparions du cas. Nous avons enjoint Meta « [à] améliorer au plus vite ses procédures afin de supprimer tout contenu légitime de ses bases ».



## Vidéo après l'attaque visant une église au Nigéria

### ANNULATION DE LA DÉCISION

---

Le 5 juin 2022, des terroristes ont attaqué une église catholique du Sud-ouest du Nigéria. Quelques heures plus tard, un utilisateur d'Instagram au Nigéria postait une vidéo représentant des corps ensanglantés inertes sur le sol de l'église, certains dont le visage était visible. Après que la vidéo a été identifiée par l'une des bases de services de mise en correspondance du contenu multimédia comme similaire à une vidéo considérée lors de l'examen manuel comme contraire à la politique de Meta, un dispositif de classification de contenus automatique a opté pour son maintien en ligne, avec l'ajout, toutefois, d'un écran d'avertissement.

Plusieurs jours plus tard, alors que le contenu était vérifié par une deuxième banque de correspondance de données, l'utilisateur a ajouté une légende en anglais décrivant l'attaque comme un événement « triste ». L'utilisateur a également ajouté des hashtags, dont certains faisaient référence au type de jeu vidéo d'action en temps réel « airsoft », quand d'autres faisaient la promotion d'armes à feu auprès de l'audience. Meta a considéré que si la vidéo en elle-même n'était pas contraire à ses règles, les hashtags l'étaient, car ils pouvaient être interprétés comme une « apologie de la violence et une minimisation de la souffrance des victimes ». Une majorité du Conseil était en désaccord avec cette décision, considérant que la restauration initiale du contenu avec un écran d'avertissement constituait la bonne solution. La majorité s'est appuyée sur une considération selon laquelle les hashtags ne devaient pas être vus comme étant de nature sadique, « principalement parce qu'ils sont associés à des utilisateurs d'armes à feu ». Une minorité a exprimé son désaccord avec cette analyse, indiquant que l'utilisation de hashtags liés au tir d'armes à feu pouvait être considérée comme « malintentionnée, et traumatiser les survivants ou la famille des victimes ». Cette ambiguïté nous a conduits à recommander la réalisation par Meta d'un examen de la forme employée dans la politique sur le contenu violent et explicite publié.

## Genre



### Vidéo de harcèlement sexuel en Inde

### MAINTIEN DE LA DÉCISION

En mars 2022, un compte Instagram décrit comme une « plateforme dédiée à l'angle Dalit » a publié une vidéo enregistrée en Inde dans laquelle une femme désignée comme issue d'une tribu est agressée par un groupe d'hommes. Les « Dalits », autrefois nommés les « intouchables », sont confrontés à l'oppression dans le système de castes indien. En dépit de l'absence de nudité et du fait que la femme en question n'était pas identifiable, Meta a retiré la vidéo, la considérant comme contraire à sa politique sur l'exploitation sexuelle des adultes. Plus tard, les équipes internes de Meta ont annulé cette décision et restauré le contenu après avoir appliqué le qualificatif « intérêt médiatique » et mis en place un écran d'avertissement, dans l'intention de sensibiliser le public à la discrimination de longue date contre les femmes de tribus en Inde.

La décision de Meta de restaurer le contenu après la mise en place du qualificatif d'intérêt médiatique et de l'écran d'avertissement nous a paru justifiée. Nous avons toutefois, de nouveau, soulevé l'argument selon lequel l'utilisation du qualificatif d'intérêt médiatique pour un contenu contraire aux règles ne constitue pas une bonne façon de gérer ce genre de cas à grande échelle.

« L'intérêt médiatique est un terme vague qui laisse une marge de manœuvre conséquente pour qui l'utilise, et ne peut pas garantir une mise en œuvre cohérente à grande échelle. Il n'inclut pas non plus de critères clairs pour l'évaluation du préjudice potentiel causé par un contenu contraire à la politique sur l'exploitation sexuelle des adultes. Le Conseil considère que les responsabilités incombant à Meta en matière de droits de la personne lui imposent de fournir des normes plus précises et de mettre en œuvre de façon plus efficace des processus de mise en application pour des cas tels que celui-ci. »

### RAPPORT SUR LA RAPIDITÉ D'IMPLÉMENTATION ET DE RÉPONSE DE META À LA SUITE DE NOS DÉCISIONS.

- Conformément à nos statuts, Meta est tenue d'implémenter nos décisions dans les 7 jours qui suivent la publication.
- Pour les 12 décisions que nous avons publiées en 2022, Meta a restauré ou supprimé le contenu dans les 7 jours, hormis dans des cas où le contenu avait déjà été restauré.
- Via notre Comité d'implémentation, actuellement constitué de cinq Membres du Conseil, nous continuons de presser Meta à fournir une plus grande transparence concernant ses procédures d'identification, dans des contextes parallèles, des contenus identiques à ceux ayant fait l'objet de nos décisions et à agir par rapport à ceux-ci. Cela garantirait que nos décisions soient traitées en dehors des cas particuliers et généralisées aux contenus concernés dans des contextes similaires.

# Application des normes internationales des droits de l'homme à la modération du contenu : Article 19

Un thème déterminant de la mission du Conseil réside en notre conviction que Meta prendra des décisions de modération du contenu de façon plus juste et raisonnée si elle base ces décisions sur les normes internationales des droits de l'homme qu'elle s'est elle-même engagée à appliquer. À cette fin, notre charte définit que nous devons « accorder une attention particulière à l'impact de la suppression de contenu au regard des normes relatives aux droits de l'homme protégeant la liberté d'expression ». Ces normes incluent l'article 19 du Pacte international relatif aux droits civils et politiques (PIDCP), qui stipule que tandis que « toute personne a droit à la liberté d'expression... l'exercice de [ce droit] peut... être soumis à certaines restrictions, qui doivent toutefois être expressément fixées par la loi et qui sont indispensables ». L'article 19 propose un test en trois parties pour évaluer les restrictions en matière d'expression :



## 1. La restriction est-elle conforme au principe de légalité ?

Nous examinons si les règles sur lesquelles Meta se base pour prendre sa décision sont accessibles et suffisamment claires aux yeux des utilisateurs pour garantir leur compréhension et leur application. Il est essentiel que les règles soient claires de sorte que ceux et celles en charge de les appliquer puissent prendre des décisions justes et cohérentes.

## 2. La restriction proposée a-t-elle un objectif légitime ?

Nous nous demandons si la règle sur laquelle la décision était fondée poursuit un objectif compatible avec les droits.

## 3. La restriction était-elle indispensable et proportionnelle ?

Le retrait du contenu constituait-il le moyen le moins intrusif d'atteindre l'objectif ? La restriction était-elle proportionnelle aux intérêts protégés ?



**Les entreprises doivent respecter le droit à la liberté d'expression, qui inclut le partage et l'échange d'idées de tous types, au-delà des frontières. »**

**Evelyn Aswad**  
COPRÉSIDENTE DU CONSEIL  
DE SURVEILLANCE



# Normes internationales des droits de l'homme dans le processus décisionnel du Conseil

En 2021, Meta a publié une politique d'entreprise en matière de droits de la personne qui décrivait ses engagements en la matière comme étant fondés sur les Principes directeurs des Nations Unies relatifs aux entreprises et aux droits de l'homme, dits UNGP (UN Guiding Principles). Ce cadre de l'année 2011 a fourni de nouvelles directives aux entreprises privées quant à leur responsabilité en matière de respect des droits de la personne.

Dans son premier rapport sur les droits de la personne<sup>1</sup>, publié en 2022, Meta a posé le cadre de sa décision d'inclure le Conseil dans sa démarche visant à fournir un accès à des recours en cas d'infraction aux droits de la personne. Les principes directeurs appellent les entreprises à fournir un « accès aux recours », y compris en mettant en place des « mécanismes de plaintes efficaces à l'échelle de l'organisation, destinés aux individus et aux communautés dont les droits auraient pu être négativement impactés [par ses opérations] ».

En matière de mécanisme de plaintes à l'échelle de l'organisation, le Conseil aspire à incarner les critères d'efficacité définis dans les principes directeurs, y compris en se montrant légitime, accessible, équitable, transparent, et compatible avec les droits, et en se présentant comme une source d'apprentissage en continu. Nous sélectionnons des cas, rendons des décisions et fournissons des recommandations pour favoriser chez Meta le respect des droits de l'ensemble des personnes. Dans chaque décision, nous fournissons une analyse détaillée des implications et des inquiétudes liées aux droits de la personne concernant le cas.

## LES SOUMISSIONS DU CONSEIL AUX NATIONS UNIES

En nous appuyant sur notre analyse des responsabilités de Meta en ce qui concerne les droits de la personne dans le cadre des cas et des avis consultatifs en matière de politiques, nous avons fourni des soumissions détaillées à deux organes des Nations Unies en 2022, décrivant l'évolution de notre approche visant à mettre en pratique les principes sur les droits de la personne dans le cadre de décisions et de politiques de contenu sur des plateformes numériques.

Dans notre soumission du mois de février au Haut-Commissariat des Nations Unies aux droits de l'homme sur « la mise en pratique des principes directeurs dans l'activité des entreprises technologiques<sup>2</sup> » et notre soumission de juillet au Rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression<sup>3</sup>, nous nous sommes concentrés sur les thèmes liés de « responsabilité » et de « recours » tels que définis dans les principes directeurs.

« Nous rappelons constamment à Meta que les principes sur les droits de la personne doivent être au centre de leurs Standards de la Communauté. »

**Nighat Dad**  
MEMBRE DU CONSEIL  
DE SURVEILLANCE



- 1 « Rapport de Meta sur les droits de l'homme : Perspectives et actions 2020 – 2021 (juillet 2022) », [https://about.fb.com/wp-content/uploads/2022/07/Meta\\_Human-Rights-Report-July-2022.pdf](https://about.fb.com/wp-content/uploads/2022/07/Meta_Human-Rights-Report-July-2022.pdf)
- 2 « Mettre en œuvre les principes directeurs des Nations Unies relatifs aux entreprises et aux droits de l'homme - Soumission au Haut-Commissariat des Nations Unies aux droits de l'homme sur l'application pratique des principes directeurs des Nations Unies aux activités des entreprises technologiques », Conseil de surveillance, février 2022
- 3 « Soumission à la Rapporteuse spéciale sur la liberté d'opinion et d'expression : Défis en temps de conflits et de troubles », Conseil de surveillance, juillet 2022.

L'ensemble des soumissions mettait en lumière les implications majeures sur les droits de la personne des décisions des réseaux sociaux relatives aux contenus en temps de conflit. Cette mise en exergue reflète l'importance attribuée aux principes directeurs pour inciter les entreprises à l'« augmentation des initiatives visant à minimiser les risques aux conséquences plus sévères en matière de droits de la personne ». Conformément aux principes directeurs, pour chaque cas assigné à un jury, nous effectuons une étude d'impact sur les droits de la personne (EIDH) qui est susceptible d'inclure une « analyse de sensibilité au conflit » afin de nous assurer que le Conseil lui-même observe ses responsabilités en matière de diligence raisonnable. Par ailleurs, notre équipe de sélection des cas dispose de processus en place visant à alerter Meta de la présence de contenu restant pouvant nécessiter une « attention immédiate » afin de limiter le risque de préjudice grave.

### ||| ÉTUDES DE CAS EN AFGHANISTAN, EN ÉTHIOPIE ET AU SOUDAN

Dans notre soumission au Rapporteur spécial des Nations Unies, nous avons pris 3 cas de l'année 2022 pour exemple : les images vidéo du Soudan, la mention des talibans dans les actualités, ainsi que l'entreprise éthiopienne « Tigray Communication Affairs Bureau ». Ces trois cas ont reflété le manque de progression de Meta dans le développement d'« un système transparent fondé sur des principes, pour la modération de contenu dans des zones de conflits ». La mise en place d'un protocole efficace de gestion de crise, comme nous l'avons appuyé, pourrait améliorer considérablement la capacité de l'entreprise à respecter les responsabilités qui lui incombent en matière de droits de la personne en temps de crise et de conflit.

Dans notre **décision sur les images vidéo du Soudan**, nous avons souligné un contexte de crise et de conflit au moment de la publication du contenu : « Les forces de sécurité... ont ciblé les journalistes et les activistes et se sont introduites dans leur foyer et leurs bureaux pour effectuer des fouilles. Les journalistes [ont été] attaqués, arrêtés et détenus... les forces armées ont pris le contrôle des médias publics et ont mis en place une





répression visant les diffuseurs et les publications du Soudan, et les réseaux sociaux sont devenus une source et un système cruciaux d'information pour faire état de la violence perpétrée par les forces armées. Les forces armées ont également coupé Internet... et arrêté les responsables civils. »

En application du test en trois parties de l'Article 19, nous avons interrogé l'incapacité de la politique sur le contenu violent et explicite publié sur Facebook à justifier la façon dont l'entreprise permet à ses utilisateurs de partager des contenus visuels afin de sensibiliser aux abus ou d'illustrer ceux-ci. Alors que Meta a, à juste titre, utilisé le qualificatif d'« intérêt médiatique » comme l'argument fondamental derrière la restauration du contenu concerné, nous avons souligné que cette règle ne permettait pas de définir clairement le terme concerné. La mesure corrective consistant à positionner un écran d'avertissement sur le contenu nous a paru constituer une restriction « indispensable et proportionnelle » de la liberté d'expression, qui a permis de protéger de façon adéquate la dignité des individus représentés et de leur famille.

Dans un deuxième cas, **la mention des talibans dans les actualités**, au sujet de la suppression d'une publication relayant l'annonce d'un représentant officiel des talibans en Afghanistan en lien avec l'ouverture prochaine d'une école pour filles, nous avons jugé que la décision initiale de Meta de la supprimer enfreignait la liberté d'expression des utilisateurs. Cette décision était fondée sur le refus opposé aux utilisateurs de Facebook en Afghanistan du droit d'« accès à l'information relative aux événements d'intérêt public... en particulier lorsqu'un groupe considéré comme dangereux renverse le gouvernement légitime en place ». Les restrictions sur la liberté de la presse imposées par le régime taliban faisaient du relais international un élément d'autant plus important, selon nous, dans la mesure où « les informations... étaient essentielles pour les personnes concernées par l'égalité des femmes et des filles en matière d'accès à l'éducation ».

Dans le troisième cas, **concernant l'entreprise éthiopienne « Tigray Communication Affairs Bureau »**, au sujet du retrait de contenus comprenant un appel d'une source officielle du gouvernement de la province du Tigré à l'attention des forces gouvernementales pour solliciter la mobilisation des forces gouvernementales centralisées, la décision de Meta relative à la suppression de la publication a été jugée cohérente avec ses responsabilités en matière de droits de la personne. Pour appuyer cette décision, nous avons mis en application un test à six facteurs du Plan d'action de Rabat, qui fournit des directives spécifiques sur la préservation de la liberté d'expression tout en protégeant les individus face à toute incitation à la discrimination, à l'hostilité ou à la violence.

Les six facteurs Rabat ont été examinés comme suit : (1) *Contexte* : le contenu a été publié dans un contexte de guerre civile continue qui gagne du terrain. (2) *Orateur* : l'orateur était un ministre du gouvernement régional associé à l'une des parties prenantes du conflit. (3) *Objet* : un appel explicite à l'assassinat de soldats qui ne se sont pas rendus. (4) *Contenu* : la publication peut être interprétée comme défendant le ciblage de combattants et de leaders politiques indépendamment de leur contribution aux hostilités. (5) *Ampleur de la diffusion* : le contenu a été publié sur la page publique d'un organe en lien avec l'une des parties prenantes au conflit, suivie par près de 260 000 abonnés, qui est restée active sur la plateforme durant 2 jours avant d'en être supprimée. (6) *Probabilité et imminence* : le contenu a été publié alors que les forces de la province du Tigré avaient déjà avancé en dehors de la province et que le Premier ministre déclarait un état d'urgence national et appelait les civils à se saisir d'armes et à combattre.



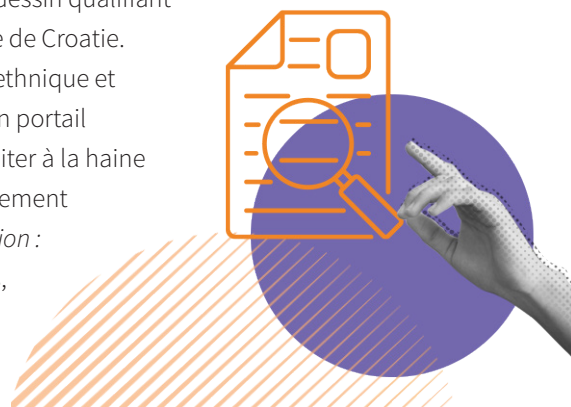
**Si la liberté d'expression est amenée à être restreinte, le motif doit être très clair. »**

**Julie Owono**  
MEMBRE DU CONSEIL  
DE SURVEILLANCE



Nous avons également recouru au test Rabat pour appuyer notre décision dans le cas du dossier sur le dessin Knin en 2022, qui concernait la publication d'un dessin qualifiant de rats des habitants aux caractéristiques ethniques serbes d'une ville de Croatie.

(1) *Contexte* : une région ayant récemment été touchée par un conflit ethnique et par des discriminations visant des minorités ethniques. (2) *Orateur* : un portail d'actualités croate connu pour ses opinions anti-Serbes. (3) *Objet* : inciter à la haine en lien avec l'ethnie. (4) *Contenu* : l'animation peut s'avérer particulièrement préjudiciable de par sa volonté de persuasion. (5) *Ampleur de la diffusion* : le contenu a été visionné plus de 380 000 fois, partagé plus de 540 fois, a enregistré plus de 2 400 réactions et plus de 1 200 commentaires. (6) *Probabilité et imminence* : le Conseil n'a pas considéré que la publication était susceptible de provoquer un préjudice imminent. Toutefois, nous avons tenu compte du fait que cette décision selon laquelle Meta « pouvait supprimer de façon légitime des publications de Facebook encourageant à la violence de façon moins directe. »

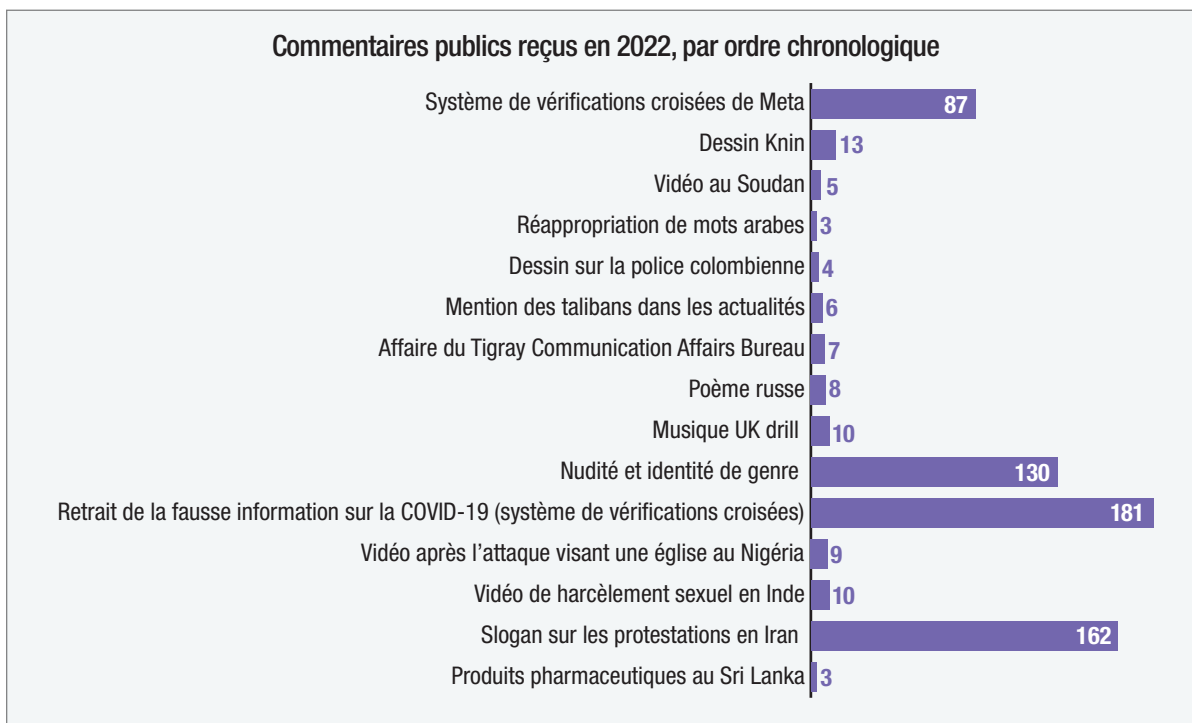


# Engagement et commentaires publics

En tant que conseil, nous savons que nous ne pouvons trouver des solutions durables aux défis que pose la modération de contenu que si nous nous montrons à l'écoute des autres organisations et travaillons en étroite collaboration avec elles. Dans cette perspective, les Membres du Conseil ainsi que nos équipes prennent part à des événements et à des débats publics, au cours desquels ils expliquent le rôle du Conseil et définissent l'importance pour nos utilisateurs d'une approche globale basée sur des principes de modération de contenu. Dans le cadre de notre processus décisionnel, les personnes et les organisations peuvent également soumettre des commentaires publics. Ces commentaires permettent aux individus de s'exprimer concernant nos décisions, et ainsi de fournir une expertise au sujet de la langue, de la culture, de la politique et des droits de l'homme. À maintes reprises, les commentaires du public ont influé sur nos décisions et nos recommandations auprès de Meta.

## ||| ENGAGEMENT ET PORTÉE

En 2022, plus de 1 000 parties prenantes ont participé à des discussions que nous avons organisées sur les problématiques liées à certains cas. Les sujets comprenaient un avis consultatif en matière de politiques de recoupement et en ce qui concerne les politiques de Meta sur les fausses informations concernant la COVID-19. Nous avons également lancé une newsletter, « Across the Board » afin de mieux informer le public sur les activités du Conseil. En 2023, nous continuerons ce travail fondamental, en convoquant au moins une table ronde par mois sur des sites en lien avec nos sept priorités stratégiques. En appuyant notre travail



d'engagement sur nos sept priorités, nous espérons également bénéficier de nos réseaux et de communautés plus larges dans l'éducation et la défense, spécialisées dans ces domaines.

## ||| COMMENTAIRES DU PUBLIC

Pour éclairer nos décisions ainsi que les avis consultatifs en matière de politiques, nous prenons très au sérieux les commentaires publics soumis par les individus et les organisations. Au cours de l'année 2022, nous avons reçu plus de 600 commentaires publics, dont le nombre a augmenté dans la deuxième moitié de l'année.

Notre décision sur le « **Tigray Communication Affairs Bureau** » a été éclairée par des analyses utiles d'experts sur la modération de contenu dans les zones de conflits. La conclusion que « l'approche actuelle de Meta en matière de modération de contenu dans les zones de conflits [est susceptible] de sembler manquer de cohérence » s'est appuyée sur des commentaires d'experts, y compris du Dr Samson Esayas, maître de conférences à la BI Norwegian Business School, qui a noté une disparité apparente entre les mesures promptes de Meta quant à la modération de contenu dans le contexte européen du conflit Russie-Ukraine et la différence apparente « de traitement pour des conflits dans d'autres régions, notamment en Éthiopie ou à Myanmar ».

Notre décision relative **au genre de musique UK drill** a trouvé un éclairage en la connaissance experte de plusieurs organisations. La Digital Rights Foundation, fondation dédiée aux droits numériques, a argué qu'une analyse de chaque phrase des paroles de la chanson supprimée en vue de constituer des preuves de méfaits passés ou de risques de préjudice futur est largement considérée comme inefficace et que la vérification

d'affirmations supposément véridiques au sein de paroles dans la musique UK drill est particulièrement complexe. La Fondation Electronic Frontier a également critiqué la surveillance, à des fins de mise en application de la loi, de la musique drill non contraire à la loi. Un commentaire public a également été émis par les services de police Metropolitan Police Service, ce qui en fait le premier commentaire d'une autorité gouvernementale. Si les services de la Metropolitan Police n'ont pas consenti à la publication du commentaire, ils ont indiqué qu'ils donneraient peut-être leur consentement ultérieurement.

« Nous nous sommes appuyés sur les précieuses informations des commentaires publics pour éclairer notre prise de décision. »

**Ronaldo Lemos**  
MEMBRE DU CONSEIL DE  
SURVEILLANCE



Notre avis consultatif en matière de politiques **vis-à-vis du programme de recoupement de Meta** a engendré un grand nombre de commentaires. Des organisations telles que le Center for Democracy and Technology, l'Institute for Technology and Society of Rio de Janeiro, PEN America et Mnemonic ont partagé leur expertise avec le Conseil. Des personnalités politiques américaines républicaines et démocrates ont également fait état de leurs inquiétudes quant au recoupement dans des commentaires publics.

Les cas ayant fait l'objet du plus grand nombre de commentaires en 2022 ont été publiés début 2023. Notre cas **sur le slogan lors des émeutes en Iran** a suscité pas moins de 162 commentaires publics. Les commentaires publics ont confirmé que l'expression « Marg bar Khamenei » (qui se traduit littéralement par « Mort à Khamenei [le guide suprême de l'Iran] ») était largement utilisée dans le contexte des manifestations récentes en Iran. Les commentaires pointaient également du doigt le retrait illégitime de contenus en farsi durant ces manifestations.

Nos cas sur la nudité et l'identité de genre ont suscité plus de 100 commentaires dont la plupart provenaient de personnes trans ou non binaires, ou de femmes cisgenres. De nombreux commentateurs ont souligné que, à l'instar des utilisateurs dans ce cas, ils avaient déjà vu leurs publications supprimées de façon illégitime. Les commentaires faisaient état d'une confusion sur le motif de retrait de la part de Meta de publications incluant des liens vers des sites Web de levées de fonds. Dans un autre commentaire, ACON, une ONG de sensibilisation au VIH en Australie, a indiqué que des contenus promouvant la prévention du VIH avaient été supprimés pour motif de sollicitation à caractère sexuel. Ces remarques ont trouvé écho dans le discours de Joanna Williams, une chercheuse ayant découvert que sur les douze organismes dédiés à la santé sexuelle qu'elle avait interrogés, neuf avaient connu des répercussions négatives de la modération de Meta dans ce domaine.

Un commentaire du centre de recherche InternetLab a fait part d'inquiétudes quant à la présupposée sexualisation du corps de femmes, ou de personnes transgenres ou non binaires, quand de telles suppositions ne sont jamais associées aux hommes cisgenres. L'impact disproportionné des retraits de contenu sur les corps de femmes a également été mentionné dans un commentaire du Dr Zahra Stardust. Enfin, les commentaires de la Gay & Lesbian Alliance Against Defamation (GLAAD) et de la The Human Rights Campaign Foundation mentionnaient des inquiétudes quant au fait que les contenus d'utilisateurs issus de groupes marginalisés avaient plus de chances d'être touchés par des signalements à répétition ou malveillants sur des contenus légitimes, aux fins de harcèlement des personnes qui les publient.

Notre **avis consultatif en matière de politiques concernant la suppression des fausses informations sur la COVID-19** a fait l'objet de 181 commentaires publics. Il s'agit du plus grand nombre de commentaires reçus pour une unique décision ou opinion en 2022. Une soumission de l'American Civil Liberties Union faisait

part de problématiques quant à la difficulté à distinguer, à grande échelle, les faits et la fiction, les opinions, les expériences et les affirmations factuelles, qui entraînera la censure par Meta de contenu légitime.

Diverses soumissions soulignaient la responsabilité de Meta de gérer les risques de sécurité publique au regard de sa portée, ainsi que le rôle que jouent ses systèmes dans l'étendue des fausses informations. Des inquiétudes ont été soulevées quant à l'adéquation des mentions et des rétrogradations en matière de gestion du risque de préjudice. Par exemple, la soumission du vice-président principal du Center of Internet and International Studies a mis en lumière des problématiques d'inefficacité des mentions dans la gestion des fausses informations diffusées par les personnalités politiques et des influenceurs majeurs.

# Chronologie des activités d'engagement en 2022

En 2022, les Membres du Conseil, les fiduciaires et l'équipe administrative ont participé à près d'une centaine d'événements dans le monde. Ci-dessous, le programme d'activités liées à nos engagements fournit quelques exemples illustrant ces événements.

MARS

Un représentant du Conseil présente l'intérêt d'une surveillance indépendante du contenu numérique lors du **Mobile World Congress** à Barcelone.

MAI

Un membre du Conseil prononce un discours devant l'**Oslo Freedom Forum** sur le rôle que peuvent jouer les médias sociaux dans le renforcement de la démocratie et la résistance à la tyrannie.

JUIN

Des membres du Conseil ainsi qu'un fiduciaire participent à la **RightsCon**, le plus grand sommet mondial sur les « droits de l'homme à l'ère du numérique ». Le jury a abordé des sujets tels que la liberté d'expression en ligne en Iran et la gouvernance pluripartite.

Les **membres du Conseil** se rencontrent pour la première fois en Californie, dans le cadre de réunions avec les **hauts responsables de Meta**, les **fiduciaires du Conseil** et diverses parties prenantes.

SEPTEMBRE

Un membre du Conseil prend la parole lors du **Concordia Annual Summit** qui se tient en même temps que l'**Assemblée générale des Nations Unies** à New York.

OCTOBRE

Des membres du Conseil se joignent à des spécialistes du secteur à New York pour discuter de l'avenir de la modération du contenu lors de la conférence **Global Freedom of Expression** organisée par l'université Columbia.

NOVEMBRE

Un membre du Conseil prend la parole lors du **Web Summit** annuel à Lisbonne.

Lors de la conférence Bread&Net à Beyrouth, un membre du Conseil participe à une table ronde réunissant différents acteurs du **Moyen-Orient**.

# Prochaines étapes À partir de 2023

**En 2022, nous avons affiné notre travail avec Meta, publié notre premier avis consultatif en matière de politiques et avons continué à avoir un impact via nos décisions et nos recommandations. En 2023, nous continuerons à apporter notre pierre à cet édifice solide. En réponse aux retours des parties prenantes, nous continuons à mettre en œuvre notre approche et avons défini six objectifs dans le cadre de notre travail avec Meta.**

## Conseil de surveillance | Objectifs 2023

1. Publier nos toutes premières décisions sommaires
2. Publier nos toutes premières décisions accélérées
3. Atteindre notre objectif révisé de nombre total de membres du Conseil, pour un maximum d'efficacité
4. Intensifier notre engagement autour de nos sept priorités stratégiques
5. Élaborer des plans à long terme en vue d'élargir notre champ d'action
6. Contrôler la manière dont Meta applique nos recommandations et inciter l'entreprise à produire des preuves de la mise en œuvre de ces recommandations et de leur impact

### 1. PUBLIER NOS TOUTES PREMIÈRES DÉCISIONS SOMMAIRES.

Sur les 50 cas sélectionnés en 2022, Meta a annulé immédiatement sa décision initiale pour 32 d'entre eux, en restaurant ou en supprimant le contenu, selon le cas. Au total, Meta a retiré ou restauré plus de 80 publications depuis 2020 à la suite du signalement de ces erreurs par le Conseil, fourni des voies de recours aux utilisateurs et révélé des informations cruciales sur sa gestion de contenu. En 2023, les Membres du Conseil, dans le cadre du comité de sélection des cas, sélectionneront certains de ces cas, qui seront ensuite examinés au titre de décisions sommaires. Ces examens mettront en lumière la raison pour laquelle nous considérons le cas comme important et évoqueront les recommandations que nous avons effectuées dans des situations similaires ainsi que nos avis consultatifs en matière de politiques. Les décisions sommaires feront l'objet d'une première ébauche réalisée par le Comité de sélection des cas. Elles seront approuvées, puis un vote sera réalisé par le Comité de sélection des cas, et non par le Conseil, qui ne tiendra pas compte des commentaires publics.

### 2. PUBLIER NOS TOUTES PREMIÈRES DÉCISIONS ACCÉLÉRÉES.

En 2023, nous envisageons de publier nos toutes premières décisions accélérées. Ces cas nous seront recommandés par Meta en accéléré. L'ébauche et la publication d'une décision sous quelques jours nous permettront d'apporter notre contribution sur des problématiques et des situations d'intérêt public qui ont un fort impact sur les droits de la personne, et ce, au moment opportun. Un jury délibérera, puis rédigera et approuvera une décision écrite, qui sera ensuite publiée sur notre site Web. Les décisions accélérées quant

au retrait ou au maintien de contenu incomberont à Meta. En raison de contraintes de temps, pour ces cas, il ne sera pas tenu compte des commentaires publics et la décision y afférente sera basée sur les informations disponibles au moment de la délibération.

### **3. ATTEINDRE NOTRE OBJECTIF RÉVISÉ DE NOMBRE TOTAL DE MEMBRES DU CONSEIL, POUR UN MAXIMUM D'EFFICACITÉ.**

Si nous avons initialement escompté un nombre de 40 membres pour ce Conseil, trois années de mise en pratique nous ont montré que le nombre optimal de membres pour des délibérations et des prises de décision réalisées en temps voulu, régulièrement et de façon efficace, est de 26. Après le renouvellement de la plupart de nos Membres en avril 2023, et une fois achevés les processus de sélection de deux nouveaux Membres, encore en cours, Meta se retirera du processus de sélection. Après cela, la sélection des Membres du Conseil sera réalisée par ses Membres et par les fiduciaires sans que Meta ne soit impliquée. Le remplacement de tout Membre du Conseil relève et relèvera uniquement de la compétence de Membres du Conseil et des fiduciaires. Nous escomptons que le Conseil disposera d'un groupe complet de Membres d'ici la fin de l'année 2023.

### **4. INTENSIFIER NOTRE ENGAGEMENT AUTOUR DE NOS SEPT PRIORITÉS STRATÉGIQUES.**

En 2022, nous avons sélectionné sept priorités stratégiques à partir d'une analyse de cas soumise au Conseil, et de problématiques rencontrées par les utilisateurs du monde entier. En 2023, ces priorités constitueront le fondement des recommandations de cas par Meta à notre intention, ainsi que de notre sélection finale pour révision. Pour l'ensemble de nos priorités, nous continuerons à travailler avec les parties prenantes afin de comprendre les politiques et pratiques de mise en application qui doivent être améliorées d'urgence chez Meta, ainsi que les types de cas permettant de se pencher dessus. Nous encourageons les parties prenantes spécialisées dans ces domaines à nous contacter via notre procédure de commentaires publics, ou via des entretiens individuels.

### **5. ÉLABORER DES PLANS À LONG TERME EN VUE D'ÉLARGIR NOTRE CHAMP D'ACTION.**

En 2022, nous avons acquis la capacité à émettre des décisions contraignantes quant à la mise en place d'un écran d'avertissement lors de la restauration ou du maintien de contenu concerné. En 2023, nous continuerons à dialoguer avec Meta sur l'amplification continue de notre champ d'application afin d'y inclure les groupes et les comptes. Si nous prévoyons le lancement de cette initiative en 2023, Meta a indiqué que, pour des raisons opérationnelles et techniques, les groupes et les comptes ont peu de chances d'entrer dans le champ d'application avant 2024. En tant que Conseil, nous sommes intéressés par l'exploration du champ d'application en vue d'une expansion à d'autres domaines, y compris à ceux de l'amplification de contenu et de la rétrogradation.

### **6. CONTRÔLER LA MANIÈRE DONT META APPLIQUE NOS RECOMMANDATIONS ET INCITER L'ENTREPRISE À PRODUIRE DES PREUVES DE LA MISE EN ŒUVRE DE CES RECOMMANDATIONS ET DE LEUR IMPACT.**

Nous continuerons à surveiller l'application des recommandations par Meta, et à faire le point dans des rapports de transparence trimestriels. Nous incitons également l'entreprise à fournir des preuves de sa mise en œuvre des recommandations dans ses politiques et produits, et à partager des statistiques sur l'effet de ces recommandations sur l'expérience des personnes qui utilisent ses plateformes et peuvent en être impactées.



# Partager les avantages d'une surveillance indépendante.

En 2022, un nombre plus important d'entreprises a souhaité bénéficier d'une expertise externe quant à ses décisions de modération de contenu. Spotify et le site de microblogging Koo ont respectivement mis en place un conseil consultatif de sécurité et un conseil consultatif. L'ancien responsable de la sécurité et des relations de confiance chez Twitter a également appelé à la mise en place de conseils de modération des contenus pour les boutiques d'application de Google et d'Apple.

Dans ses fondements, la surveillance indépendante consiste pour les entreprises à montrer au grand jour leurs processus internes et à convier des personnes extérieures à examiner leurs décisions. Ce type d'examen et de mise à l'épreuve entraîne une meilleure prise de décision, qui s'avérera également plus solide, et aidera à l'établissement d'une relation de confiance avec les utilisateurs sur le long terme : une relation gagnant-gagnant pour les communautés et les entreprises.

Dès le départ, le Conseil de surveillance ne consistait pas simplement à améliorer Facebook et Instagram. Il s'agissait également d'une opportunité d'expérimenter une nouvelle approche indépendante de la modération du contenu pouvant être appliquée à d'autres entreprises technologies et plateformes. En tant que Conseil, nous avons acquis une large expérience durant les trois dernières années, qui pourrait aider d'autres entreprises à prendre des décisions plus éclairées et à mieux satisfaire leurs utilisateurs.

## ||| UNE EXPÉRIENCE UNIQUE DE CRÉATION D'UN CONSEIL DE SURVEILLANCE

La création et la gestion d'un conseil de surveillance pour la gouvernance de contenu, quelque chose d'entièrement nouveau, se sont avérées bien plus compliquées que prévu. Depuis le premier jour, nous avons travaillé dur pour surmonter un certain nombre de défis, en matière d'infrastructure technique et d'élaboration d'une culture d'équilibre entre une modération de contenu rapide et une délibération juste. Pendant des mois, nous avons travaillé avec Meta pour créer un système d'appels indépendant accessible aux milliards d'utilisateurs du monde entier. Nous avons mis en place un processus de commentaires publics pour donner à chacun la possibilité de s'exprimer dans le cadre de nos prises de décisions. Nous avons appris, en nous appuyant sur nos différentes nationalités, nos différents parcours et nos différents points de vue, à délibérer de cas qui ne donnent pas lieu à des réponses faciles.



**La liberté d'expression de milliards de personnes sur cette planète est trop importante pour être mise entre les mains d'une unique entreprise. »**

**Alan Rusbridger**  
MEMBRE DU CONSEIL DE SURVEILLANCE



En trois ans d'existence, nous avons appris bien des choses sur la façon dont un conseil de surveillance pour la gouvernance de contenu doit fonctionner. Plus spécifiquement, nous avons identifié cinq caractéristiques pouvant aider d'autres entreprises technologiques cherchant à mettre en place ce type de surveillance.

- **Indépendance** : tout organe de surveillance doit être structuré de manière à pouvoir émettre des opinions indépendantes. Pour prendre les bonnes décisions en matière d'examen de contenu, le conseil ne doit pas subir de restrictions de la part des plateformes pour des motifs politiques, commerciaux, ou encore en lien

avec l'image. En tant que Conseil, nous n'avons pas hésité à annuler les décisions de Meta. Cette distinction est cruciale pour établir la légitimité auprès des utilisateurs et des groupes de la société civile dans le monde.

- ⦿ **Transparence** : pour avoir confiance dans les décisions d'une entreprise, les personnes doivent pouvoir comprendre le processus de décision. Avant de rendre nos décisions, nous tenons compte d'un maximum d'informations auparavant non publiques concernant les politiques et leur mise en application. Pour être considérées comme authentiques et légitimes dans leur engagement à promouvoir la liberté d'expression, les entreprises doivent indiquer aux utilisateurs pourquoi leurs publications ont été supprimées, ou leur compte désactivé. Elles doivent également mieux expliquer les motifs de leur prise de décisions et faire preuve de davantage de transparence lorsque les gouvernements ou les représentants de l'État les appellent à retirer des contenus.
- ⦿ **Diversité** : la plupart des utilisateurs des plateformes de réseaux sociaux sont basés en dehors des États-Unis et d'Europe. Nombre d'inquiétudes liées aux impacts négatifs des réseaux sociaux, ainsi qu'à leurs avantages, sont plus ancrées dans des pays situés en dehors des États-Unis et d'Europe. Pour avoir confiance en la prise de décision d'une entreprise, les habitants de ces pays ont besoin de bien davantage d'une « implication » en faveur de leur région du monde. Ils doivent se sentir entendus et représentés au moment de la prise de décision. Les Membres de notre Conseil de surveillance vivent dans 27 pays et parlent 29 langues. La diversité constitue un réel avantage dans nos prises de décision.
- ⦿ **Droits de la personne** : aujourd'hui, les entreprises technologiques sont confrontées à un défi important en matière de modération de contenu : quelles règles mettre en place pour des milliards de personnes aux nationalités, aux langues et aux cultures diverses ? Nous pensons que les Normes internationales des droits de l'homme constituent la clé de voûte de la réponse à cette question. Elles s'appliquent de façon égalitaire à chacun et constituent un cadre cohérent pour l'évaluation de la liberté d'expression d'un utilisateur, en parallèle d'autres droits de la personne, tels que le droit de vivre et le droit à la vie privée.
- ⦿ **Partenariat** : enfin, le partenariat est un élément essentiel. Des organes de surveillance indépendante n'auront un impact sur le long terme que si les entreprises portent à leur connaissance leurs données et leurs processus. S'il existe parfois des situations dans lesquelles le partage d'informations peut compromettre la confidentialité des informations sur un utilisateur ou permettre à des personnes malintentionnées de se servir des politiques à mauvais escient, notre travail dépend toutefois de

données et d'informations que seule Meta peut fournir. Cette transparence permet également d'établir une relation de confiance et une image légitime auprès des utilisateurs et de la société civile. Les entreprises doivent également se montrer disposées à mettre en œuvre les recommandations, à l'instar de Meta pour nombre de nos propositions. Le partenariat entre une entreprise et un conseil de surveillance nécessite une volonté d'apprendre de la part des deux entités.

« Nous avons constaté que nos processus de prise de décision faisaient désormais l'objet d'une réflexion pour d'autres plateformes de réseaux sociaux qui sont confrontées aux mêmes problématiques. »

Kristina Arriaga  
FIDUCIAIRE DU CONSEIL  
DE SURVEILLANCE



En tant que Conseil, nous espérons que l'approche décrite précédemment fournit un cadre crédible pour les autres entreprises technologiques qui souhaitent bénéficier d'une surveillance indépendante. Dans le cadre de notre travail avec Meta, nous avons déjà surmonté bien des obstacles opérationnels et appris bien des leçons quant à la création d'un organe de cette envergure. Si toute approche doit être adaptée aux caractéristiques propres à une entreprise, notre expérience peut aider les entreprises à bénéficier d'une surveillance indépendante plus rapidement, avec des coûts réduits.

Dans le sillage de notre réflexion quant au meilleur moyen de responsabiliser le secteur des technologies, nous encourageons l'ensemble des entreprises à envisager la création d'un organe de surveillance indépendant pour leurs plateformes et services.

## Aider les entreprises à s'adapter aux nouvelles réglementations

Nous sommes intéressés à l'idée de travailler avec des entreprises qui partagent la conviction qu'une gouvernance de contenus responsable et transparente, surveillée par des organes indépendants, fait partie intégrante de la mise en place d'un environnement en ligne qui respecte la liberté d'expression et les autres droits de la personne. Les entreprises disposées à prendre un engagement fort au regard de ces normes et de ces structures au bénéfice de leurs utilisateurs et, à plus grande échelle, de la société gagneront, à l'instar de Meta, la confiance des utilisateurs et pourront témoigner de leur intention rigoureuse auprès des autorités de régulations.

Durant l'année, de nouvelles réglementations imposeront de nouvelles exigences aux entreprises technologiques. Nombre d'entre elles devront modifier leur approche. Les débats antérieurs liés à la régulation des réseaux sociaux présentaient un choix binaire pour la régulation gouvernementale et la régulation autonome dans ce secteur. Toutefois, aujourd'hui, nous voyons émerger une possibilité de co-régulation, dont les éléments du mécanisme de régulation indépendant sont fondés sur la législation. Si nombre d'intervenants joueront leur rôle dans ce paysage réglementaire émergent, nous pensons que notre approche indépendante ainsi que notre priorisation de la transparence et de la justice auprès des utilisateurs peuvent constituer un élément de solution.

« Nous ne cherchons pas à devenir le Conseil de l'ensemble du secteur. Mais nous souhaitons partager ce que nous avons appris et travailler avec les entreprises intéressées par la mise en place de divers organes pour la définition de normes et la surveillance de la gouvernance de contenu. »

**Thomas Hughes**  
DIRECTEUR DU CONSEIL  
DE SURVEILLANCE.

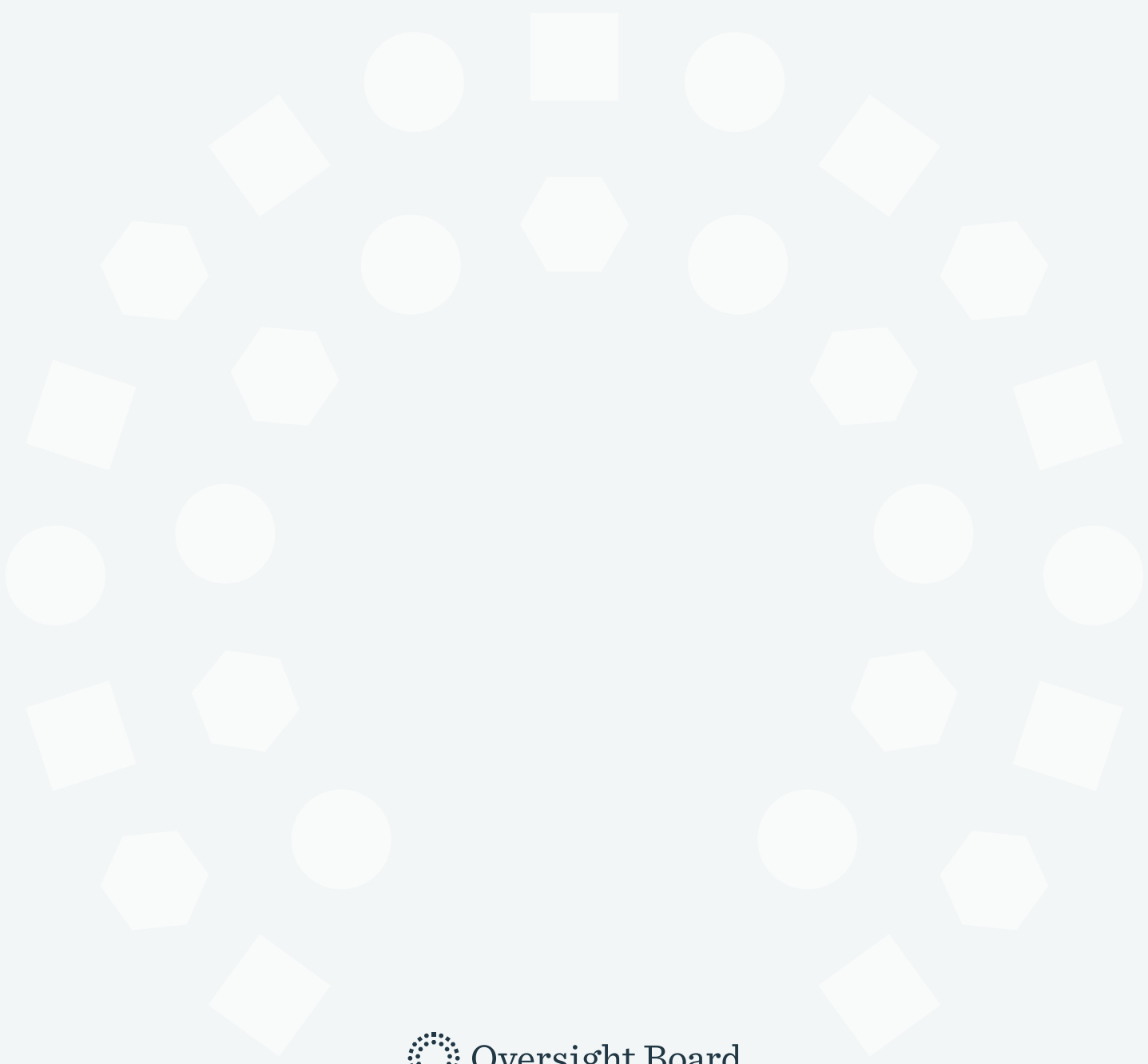


# Conclusion

Les réseaux sociaux, sous leur meilleur jour, peuvent vraiment constituer un vecteur d'exception pour les échanges et les interactions à l'échelle mondiale. Bénéficier des avantages tout en limitant les préjudices, ce n'est pas chose aisée. Ayant ouvert la voie, le Conseil de surveillance continue à mettre en application les leçons apprises pour s'améliorer. Nous pensons que les entreprises de réseaux sociaux prendront des décisions de modération du contenu de façon plus juste et raisonnée si elles basent ces décisions sur les normes internationales des droits de l'homme. Nous sommes prêts à partager ce que nous avons appris jusqu'à présent avec d'autres entreprises et organisations qui partagent nos aspirations à favoriser la transparence et à améliorer l'expérience humaine des utilisateurs en ligne.







Oversight Board

[www.oversightboard.com](http://www.oversightboard.com)

© 2023 Oversight Board LLC