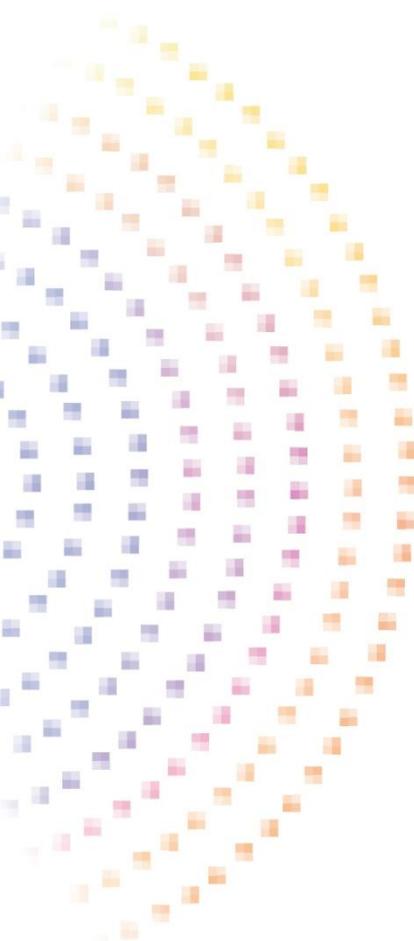




APPEL DE CHRISTCHURCH

À AGIR CONTRE LE TERRORISME
ET L'EXTREMISME VIOLENT EN LIGNE



Déclaration de la coprésidence : Sommet des dirigeants de l'Appel de Christchurch 2023

Aujourd'hui, présidents et chefs de gouvernement, dirigeants des entreprises du numérique, et représentants de la société civile et des partenaires de l'Appel de Christchurch se sont réunis au palais de l'Élysée, à Paris, afin de poursuivre leurs efforts communs visant à supprimer les contenus terroristes et extrémistes violents en ligne.

Le Sommet de cette année se déroule alors que la communauté de l'Appel se trouve confrontée à des enjeux majeurs :

- Le conflit à Gaza et en Israël a provoqué l'apparition d'un nombre considérable de contenus terroristes et extrémistes violents ainsi que de contenus incitant à la violence, l'antisémitisme et l'islamophobie ou en faisant l'apologie.
- Les progrès rapides destechologies, notamment des modèles d'intelligence artificielle (IA), offrent de puissants instruments innovants permettant de mieux lutter contre les contenus terroristes et extrémistes violents en ligne, mais ils sont également sources de nouveaux mécanismes et de nouvelles vulnérabilités pouvant être exploités par les organisations terroristes et extrémistes violents.

Aujourd'hui, les hauts responsables ont réaffirmé leur engagement en faveur d'un Appel fort, résilient et efficace et leur volonté de supprimer les contenus terroristes et extrémistes violents, de protéger et de promouvoir les droits de l'homme en ligne ainsi qu'un internet libre, ouvert et sûr, en en faisant une force positive et une plateforme d'innovation et de progrès social.

Notre communauté continue de grandir. La communauté a souhaité la bienvenue à quatre nouveaux soutiens de l'Appel : Anthropic, Discord, OpenAI et Vimeo ; ainsi qu'à sept nouvelles organisations partenaires : le Réseau de recherche sur l'extrémisme et les jeux vidéo (EGRN), CASM Technology, Memetica, Moonshot, Muflehun, Point de Contact et Tremau. Le réseau consultatif de

l'Appel de Christchurch, constitué d'organisations de la société civile, s'est également agrandi, avec onze nouveaux membres. Cette communauté élargie donne des capacités, des compétences et une énergie nouvelles à l'Appel.

Nouvelles technologies et conséquences en matière de contenus terroristes et extrémistes violents

L'intelligence artificielle transforme la manière dont nous faisons face aux contenus terroristes et extrémistes violents, en nous offrant de meilleures capacités de détection, de prévention et d'intervention ainsi que la possibilité d'adapter les outils à de nouveaux contextes, à de nouvelles langues et de les diversifier.

Anthropic et OpenAI se joignent à un certain nombre de soutiens de l'Appel qui disposent de capacités avancées en matière d'intelligence artificielle. D'importants travaux sont déjà en cours pour permettre de tirer profit tout ce potentiel. Une nouvelle collaboration entre Microsoft et *Tech Against Terrorism* a ainsi été lancée pour expérimenter l'utilisation d'intelligences artificielles de pointe afin d'améliorer les mécanismes existants et, à terme, de mettre à la disposition des entreprises moins outillées des services de détection de nouvelle génération.

En l'absence de garde-fous, les capacités des intelligences artificielles de pointe seront inexorablement exploitées par les organisations terroristes et extrémistes violentes à des fins de propagande, de recrutement et de préparation d'attaques. Les biais ou les vulnérabilités intégrées lors de l'entraînement des modèles d'intelligence artificielle pourraient représenter de nouveaux vecteurs de nuisances ou de nouveaux moyens d'exploitation néfaste. Les dispositifs de sécurité mis en place pourraient

aussi avoir des répercussions néfastes inattendues sur les droits de l'homme.

Le recours à des équipes de prévention de crise (« Red team ») et l'analyse indépendante des risques constitueront une part importante des outils pour gérer ces risques. La communauté de l'Appel a réalisé des progrès dans la définition des paramètres de sécurité en matière de gestion des risques, notamment grâce au rôle joué par les soutiens de l'Appel et à la création de nouveaux forums et de nouvelles procédures visant à aider et à relever les défis globaux posés par les intelligences artificielles en matière de sécurité.

Les dirigeants de l'Appel se sont prononcés en faveur des actions suivantes :

- travailler de concert avec la communauté multipartite de l'Appel de Christchurch pour contribuer à la mise en place de cadres, notamment dans le contexte du G7, du Partenariat mondial sur l'intelligence artificielle et du Sommet sur la sécurité de l'IA, afin d'identifier, de signaler et de limiter l'exploitation de ces outils par les organisations terroristes et extrémistes violentes ;
- collaborer avec les acteurs du secteur des technologies de la sécurité (industrie « Trust & Safety ») et les fournisseurs d'outils commerciaux et sous licence libre afin de favoriser la mise à disposition de solutions utiles dans le secteur, dans le respect des droits de l'homme et d'une façon compatible avec un internet libre, ouvert et sûr ;
- continuer d'attirer dans une communauté d'experts, de partenaires et des organisations compétents ;
- œuvrer à l'évaluation des risques et à la mise en place d'équipes de prévention des crises pour réduire les risques d'usage abusif de l'IA par les organisations terroristes et extrémistes violentes, en s'appuyant sur les points forts de la communauté multipartite de l'Appel ;

- tirer profit des travaux menés par les fournisseurs de services en ligne pour aligner les outils de gouvernance des contenus générés par des IA avec les engagements pris dans le cadre de l'Appel et pour favoriser l'établissement de rapports et d'un dialogue plus réguliers concernant les efforts déployés ;
- s'appuyer sur l'expertise de la communauté pour veiller à ce que nos efforts contribuent au maintien d'un internet mondial libre, ouvert, sûr et interopérable et ne le mettent pas en péril ; et
- collaborer avec la société civile pour veiller à ce que les droits de l'homme soient promus et respectés.

Radicalisation et résultats algorithmiques

Les engagements pris dans le cadre de l'Appel cherchent à s'attaquer non seulement aux facteurs sous-jacents de la diffusion des contenus terroristes et extrémistes violents en ligne, mais aussi à aborder les questions plus pressantes posées par la modération des contenus. Ils exigent que nous tenions compte du défi que représente la lutte contre la radicalisation violente, qui mène souvent à la création de nouveaux contenus terroristes et extrémistes violents en ligne.

La radicalisation n'est pas uniquement un phénomène en ligne, mais nos travaux ont mis en évidence la nécessité de mieux comprendre les activités en ligne pour pouvoir mettre au point des interventions, des mesures de prévention et d'autres mesures efficaces. Les internautes et les communautés impactées jouent un rôle important en matière de prévention, et ils devraient avoir la possibilité de contribuer à ce travail.

Les points de vue des plus jeunes internautes sont essentiels pour comprendre comment le parcours utilisateur en ligne de personnes de plus en plus jeunes peut créer un terrain propice à la radicalisation et contribuer à une

plus grande exposition à des contenus terroristes et extrémistes violents en ligne.

L'année dernière, les dirigeants ont demandé à la communauté de l'Appel de trouver des façons de mobiliser les jeunes efficacement, et en toute sécurité, et de faire l'écho à leurs idées dans nos travaux. Nous avons organisé ce jour notre première réunion de mobilisation des jeunes, au cours de laquelle nous avons eu l'occasion d'entendre directement leurs priorités en matière de lutte contre les contenus terroristes et extrémistes violents en ligne et la manière dont ils considèrent que nous devrions aborder ces problématiques. Les dirigeants ont accueilli favorablement l'initiative visant à créer un groupe de référence pour la jeunesse dans le cadre de l'Appel de Christchurch pour que les jeunes puissent contribuer aux efforts et aux travaux tout au long de l'année.

Les dirigeants ont salué le rapport intitulé [« Les voies misogynes de la radicalisation » \(Misogynistic Pathways to Radicalisation\)](#), élaboré par l'Institut pour le Dialogue Stratégique en collaboration avec l'Appel de Christchurch et le Partenariat mondial pour l'action contre les violences et le harcèlement sexistes en ligne. Les dirigeants ont souligné qu'il était important de prendre des mesures pour lutter contre les violences à l'encontre des personnes LGBTQIA+ et les violences sexistes, qui constituent des vecteurs de radicalisation, d'élaborer, de mettre en œuvre des recommandations multipartites, et de continuer à réunir des données factuelles pour soutenir l'adoption de mesures nouvelles.

L'adoption de nouvelles mesures réglementaires visant à favoriser l'accès des chercheurs aux données et de nouvelles mesures volontaristes de transparence, telles que la transmission, par les fournisseurs de services en ligne, d'informations détaillées concernant les éléments entrant en compte dans leurs systèmes de recommandation de contenus, ainsi que les codes informatiques soutenant ces systèmes, ont étendu les possibilités en matière de sensibilisation du grand public. Ces mesures promeuvent aussi

la promotion de processus décisionnels fondés sur des données factuelles pour les décideurs, les développeurs de technologies et les représentants de la société civile.

Les dirigeants ont salué les travaux du Forum de l'UE sur Internet (« EU Internet Forum »), notamment son [Étude relative au rôle et aux effets du recours à l'amplification algorithmique pour diffuser des contenus terroristes et extrémistes violents et des contenus à la limite de la légalité \(« borderline content »\)](#). Cette étude, publiée en 2023, présente un aperçu des différentes répercussions que peuvent avoir les systèmes de recommandation et les interactions entre utilisateurs et incite à poursuivre les efforts déployés pour faire face à ces risques.

L'initiative de l'Appel de Christchurch sur les résultats algorithmiques (« Christchurch Call Initiative on Algorithm Impact » - CCIAO), a été annoncée l'année dernière par la Nouvelle-Zélande, les États-Unis, Microsoft et X, en collaboration avec l'organisation à but non lucratif OpenMined. Cette initiative a permis de réaliser d'importants progrès dans l'élaboration d'un système d'échange de données sur les algorithmes en offrant une confidentialité renforcée. Au cours de la première phase de l'initiative, ce système a été mis en place et testé, démontrant qu'il était en mesure de répondre en toute sécurité aux questions posées par des chercheurs tiers accrédités en utilisant des données conservées de manière sécurisée par les réseaux sociaux.

Aujourd'hui, nous avons annoncé que la France, Dailymotion et le Centre pour la sécurité et les technologies émergentes de l'Université de Georgetown ont rejoint le CCIAO, annonçant un financement supplémentaire de 1,3 million de dollars. Cette contribution financière permettra de poursuivre la mise en œuvre du CCIAO, de développer ses activités ainsi que d'étudier et de tester son potentiel en tant que moyen de faciliter l'accès des personnes tierces aux données, notamment dans le cadre de nouvelles mesures réglementaires.

Les dirigeants se sont prononcés en faveur des actions suivantes :

- la mise au point d'une nouvelle gouvernance et de nouveaux cadres éthiques pour le CCIAO, afin de permettre la poursuite de la mise en œuvre du projet ;
- la poursuite des travaux visant à faire du CCIAO un réseau mondial qui exercera une influence majeure en faveur de l'Appel de Christchurch, du déploiement éthique et responsable de l'intelligence artificielle et du bien public ;
- des efforts spécifiques pour améliorer l'expérience en ligne des jeunes et pour lutter contre la misogynie et la haine anti-LGBTQIA+ en ligne, qui constituent des vecteurs de l'extrémisme violent.

Prochaines étapes de l'Appel de Christchurch

L'Appel de Christchurch, qui agit comme une communauté multipartite guidée par les gouvernements français et néo-zélandais et qui s'appuie sur des valeurs et des engagements communs, est un vecteur important de changements positifs.

Aujourd'hui, les dirigeants ont examiné diverses solutions visant à donner à l'Appel une base résiliente à long terme qui lui permettra de remplir ses objectifs dans un paysage technologique, diplomatique et réglementaire international en constante évolution et de tirer parti des avantages de son modèle novateur pour contribuer à relever les défis numériques.

En tant que cofondateurs, la Nouvelle-Zélande et la France sont déterminées à assurer le succès à long terme de l'Appel de Christchurch.

Les dirigeants se sont prononcés en faveur des actions suivantes :

- Les deux pays travailleront de concert avec la communauté de l'Appel ainsi que d'éventuels donateurs pour envisager des

solutions et mettre en place une structure pour que le secrétariat de l'Appel soit résilient, permette à l'ensemble des membres de cette communauté multipartite de contribuer à nos efforts, produise des résultats à long terme et conserve les qualités et les valeurs importantes qui ont contribué à notre succès.

- Le secrétariat de l'Appel recherchera des partenariats avec des initiatives complémentaires qui partagent les engagements de l'Appel en faveur des droits de l'homme et d'un internet libre, ouvert et sûr, afin de coordonner les efforts déployés pour supprimer les contenus terroristes et extrémistes violents en ligne, d'étudier comment les travaux relatifs à des questions connexes, telles que l'intégrité de l'information, le déploiement sûr et responsable de l'intelligence artificielle et la radicalisation des jeunes, peuvent permettre à l'Appel de progresser dans la tenue de ses engagements, et de faciliter la participation des gouvernements, des organisations et des institutions de taille réduite aux efforts multipartites visant à élaborer les normes en matière de technologies.
- Dans la perspective du cinquième anniversaire de l'Appel de Christchurch, l'année prochaine, les dirigeants se prononceront sur les structures devant être mises en place pour soutenir son action et faire de l'Appel une initiative multipartite influente à l'avenir.

Réaction face aux crises et contenus en ligne liés au conflit à Gaza et en Israël

Les dirigeants ont reconnu que la forte augmentation du nombre de contenus terroristes et extrémistes violents découlant de ce conflit qui incitent à la violence et à la haine contre les communautés juives, musulmanes, les communautés au Moyen-Orient, et toutes les communautés impactées

par ce conflit, pris individuellement ou en tant que groupes, sont sources de peur et de souffrances dans le monde entier.

Dans le contexte de ce conflit, les gouvernements et les fournisseurs de services en ligne ont réaffirmé leurs engagements pris au titre de l'Appel de supprimer les contenus terroristes et extrémistes violents en ligne, de manière transparente et tout en respectant et en protégeant les droits de l'homme ainsi que la défense d'un internet libre, ouvert et sûr.

S'agissant des capacités de réponse aux crises, les dirigeants se sont félicités des progrès réalisés l'année dernière pour permettre à plus de fournisseurs de services en ligne de rejoindre l'Appel. Ils ont salué l'élargissement de la Plateforme d'analyse des contenus terroristes (*Terrorist Content Analytics Platform*) de *Tech Against Terrorism*, soutenue par un financement du Canada, et l'annonce d'un financement de l'Australie en faveur des capacités de réponse aux crises 24 heures sur 24, 7 jours sur 7. Ces deux évolutions permettent de mieux soutenir les plus petits acteurs privés et de remédier aux lacunes identifiées dans la réaction face aux crises.

Les dirigeants se sont félicités du soutien supplémentaire dont bénéficieront les plus petits acteurs privés, notamment grâce à des initiatives comme l'outil « [Altitude](#) », gratuit et sous licence libre, conçu par Jigsaw, filiale de Google, et *Tech Against Terrorism*, en collaboration avec le Forum mondial de l'internet contre le terrorisme, pour aider les plateformes à protéger leurs utilisateurs des contenus terroristes et extrémistes violents.

Les dirigeants ont également approuvé les mesures suivantes :

- procéder à l'examen de nos systèmes de réaction face aux crises afin de garantir une réaction rapide, efficace et qui permette aux personnes concernées de faire valoir leurs droits face aux crises en ligne impliquant la prolifération de contenus terroristes ou extrémistes violents, notamment ceux issus d'incidents et de conflits prolongés hors

ligne, ou qui impliquent des contenus émanant de témoins ;

- faire évoluer le protocole commun de réaction face aux crises de l'Appel de Christchurch afin de l'adapter aux circonstances et aux menaces nouvelles, en s'appuyant sur le savoir-faire et les capacités de nos partenaires et du réseau consultatif de l'Appel de Christchurch pour surveiller et contrer les contenus terroristes et extrémistes violents et les répercussions sur les droits de l'homme ;
- envisager et inclure des approches de réduction des tensions et de prévention de la haine et des violences en ligne et hors ligne, notamment des interventions positives et des communications stratégiques, dans le cadre de l'évolution du protocole commun de réponse face aux crises ;
- tirer parti des efforts déployés au cours de l'année écoulée pour mettre en place de nouveaux outils et renforcer le soutien aux petites entreprises dans leur lutte contre les contenus terroristes et extrémistes violents sur leurs plateformes, notamment en situation de crise. Intégrer de nouvelles capacités dans les infrastructures et protocoles existants, notamment la Plateforme d'analyse des contenus terroristes de Tech Against Terrorism et l'évolution du protocole commun de réaction face aux crises ;
- continuer à travailler avec tous les secteurs pour procéder à la mise à l'essai et à l'amélioration du système ainsi qu'au renforcement de la transparence et de la protection des droits de l'homme.

Activités en cours dans le cadre de l'Appel

Les membres de l'Appel de Christchurch ont réalisé des progrès importants sur plusieurs points supplémentaires, notamment les nouvelles recommandations en matière de transparence gouvernementale, les efforts de

recherche sur l'évolution des menaces, l'accroissement de la coordination et des synergies avec la Coalition pour la liberté en ligne, *Tech for Democracy* et d'autres initiatives similaires pour améliorer les normes applicables en matière de technologie, ainsi que l'utilisation par la France de certains aspects du modèle de l'Appel dans la mise en place du Laboratoire pour la protection de l'enfance en ligne.

Les dirigeants se sont prononcés en faveur des actions suivantes :

- en collaboration avec la Coalition de l'action pour une véritable transparence (*Action Coalition on Meaningful Transparency*), surmonter les défis liés à la transparence dans le cadre des efforts pour éliminer les contenus terroristes et extrémistes violents, notamment en promouvant la mise à disposition d'informations pour la conduite d'évaluations et de recherches indépendantes, en recourant au [portail commun des initiatives sur la transparence](#) (en anglais) ;
- échanger au sujet de la cohérence réglementaire pour promouvoir des approches efficaces et complémentaires au niveau mondial pour l'élimination des contenus terroristes et extrémistes violents ;
- renforcer le rôle des représentants du secteur technique au sein de l'Appel pour promouvoir un internet libre, ouvert et sûr ;
- mettre en place un processus multipartite pour évaluer nos actions au regard des engagements pris au titre de l'Appel et améliorer la transparence et la communication sur l'importance et la portée de l'Appel ;
- encourager le Forum mondial de l'internet contre le terrorisme à harmoniser ses efforts avec l'action menée dans le cadre de l'Appel. Accueillir favorablement ses efforts pour intégrer de nouveaux membres et encourager à inclure une plus

grande variété de plateformes, notamment des entreprises de jeux vidéo et des entreprises de petite taille, pour accroître son influence sur l'industrie ;

- continuer d'œuvrer à un dialogue constructif avec la société civile et les autres parties prenantes, prendre part et contribuer à la diversité de points de vues, des cultures, des régions et des expériences dans le cadre de notre action, en lien notamment avec les victimes et les survivants.

Conclusion

En quatre ans depuis que l'Appel de Christchurch a été lancé, nous avons accompli des progrès considérables concernant nos engagements et pour prévenir l'instrumentalisation d'internet par des groupes terroristes et extrémistes violents.

L'environnement numérique joue aujourd'hui un rôle central et essentiel dans la vie des populations du monde entier, et les nouveaux outils numériques déployés grâce à internet ouvrent des perspectives enthousiasmantes et inédites pour contribuer au bien-être social et à la connectivité de nos sociétés.

La nature des contenus terroristes et extrémistes violents et l'exploitation des services en ligne évoluent, et nous devons nous y adapter. Nous sommes profondément convaincus que le modèle de l'Appel de Christchurch, reposant sur les droits de l'homme, un internet libre, ouvert et sûr et la volonté commune d'agir, peut continuer de fournir des moyens utiles pour aller de l'avant dans un environnement numérique complexe et contesté.

Nous avons montré aujourd'hui que l'Appel continue de jouer un rôle majeur dans la création d'un meilleur environnement numérique.