ENTRAÎNÉ·E·S DANS LE « RABBIT HOLE »*

DE NOUVELLES PREUVES MONTRENT LES RISQUES DE TIKTOK POUR LA SANTÉ MENTALE DES ENFANTS

*« TERRIER DE LAPIN » : EFFET DÉSIGNANT LES SPIRALES DE CONTENUS EN LIGNE

AMNESTY INTERNATIONAL

Amnesty International est un mouvement rassemblant 10 millions de personnes qui fait appel à l'humanité en chacun et chacune de nous et milite pour que nous puissions toutes et tous jouir de nos droits humains.

Notre vision est celle d'un monde dans lequel les dirigeants et dirigeantes tiennent leurs promesses, respectent le droit international et sont tenus de rendre des comptes.

Indépendante de tout gouvernement, de toute idéologie politique, de tout intérêt économique et de toute religion, Amnesty International est essentiellement financée par ses membres et des dons de particuliers. Nous avons la conviction qu'agir avec solidarité et compassion aux côtés de personnes du monde entier peut rendre nos sociétés meilleures.

© Amnesty International 2025
Sauf exception dûment mentionnée, ce document est sous licence Creative Commons (Attribution-NonCommercial-NoDerivatives- International 4.0).
https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode
Pour plus d'informations, veuillez consulter la page relative aux autorisations sur notre site : www.amnesty.org/fr.
Lorsqu'une entité autre qu'Amnesty International est détentrice du copyright, le matériel n'est pas sous licence Creative Commons.

Édition originale publiée en 2025 par Amnesty International Ltd.
Peter Benenson House, 1 Easton Street,
London WC1X 0DW, UK

Index: POL 40/0360/2025
Original language: English

amnesty.org

Dessin de couverture : *© Marine Fustec*

AMNESTY INTERNATIONAL

AVERTISSEMENT Ce rapport traite de sujets sensibles et contient des exemples de contenus portant notamment sur l'automutilation et le suicide. Vous trouverez les coordonnées de lignes d'assistance téléphonique et d'organisations qui apportent un soutien émotionnel gratuit aux quatre coins de la planète dans le guide d'Amnesty International intitulé *Sauver le monde sans s'effondrer (volume 2). Manuel du bien-être pour les jeunes militant·e·s*[1].

En France :
3114 – Numéro National de Prévention du Suicide
Accessible 24/24 et 7/7 – gratuit et confidentiel

Remerciements

Amnesty International tient à remercier les membres du collectif Algos Victima qui ont participé à ces recherches, ainsi que ses partenaires techniques de l'Algorithmic Transparency Institute.

Le collectif Algos Victima

Une première plainte française pour provocation au suicide a été déposée contre TikTok le 12 septembre 2023 par la famille de Marie Le Tiec, représentée par Laure Boutron-Marmion. Le collectif Algos Victima a été créé en 2024 pour soutenir, à travers des actions judiciaires devant les institutions françaises, les mineur·e·s et leurs familles victimes de préjudices liés aux réseaux sociaux. En novembre 2024, sept familles ont engagé un recours collectif contre TikTok, lui reprochant d'avoir exposé leurs enfants à des vidéos qui banalisaient, voire idéalisaient, la dépression, l'automutilation ou le suicide. Aujourd'hui, ce collectif rassemble 11 familles françaises mobilisées pour faire établir la responsabilité de TikTok dans la détérioration de la santé mentale et physique de leurs enfants.

L'Algorithmic Transparency Institute

L'Algorithmic Transparency Institute (ATI) est un projet de l'organisation National Conference on Citizenship. Il vise à améliorer la transparence des plateformes numériques exerçant une influence sur le discours civique. Il développe des outils, collecte des données, produit des recherches et soutient des analyses favorisant une compréhension commune des relations entre les conversations en ligne et notre démocratie. L'ATI permet à des chercheurs et chercheuses, des journalistes et des défenseur·e·s de la société civile de suivre et de comprendre notre discours civique numérique sur les réseaux sociaux.

Levi Kaplan et Piotr Sapiezynski, de l'université du Nord-Est (États-Unis) nous ont aidé pour les tests automatisés et l'analyse des données.

1 Amnesty International, *Sauver le monde sans s'effondrer (volume 2). Manuel du bien-être pour les jeunes militant·e·s* (index : ACT 10/3231/2020), janvier 2021, amnesty.org/fr/documents/act10/3231/2020/fr/, p. 99.

SOMMAIRE

1. SYNTHÈSE 5

2. MÉTHODOLOGIE 7

3. LES JEUNES EN FRANCE MOYENNEMENT SENSIBILISÉS AUX RISQUES DES RÉSEAUX SOCIAUX 10

4. TÉMOIGNAGES DE JEUNES VICTIMES ET DE PARENTS CONCERNÉS OU ENDEUILLÉS 12

4.1 COMMENT LE FIL « POUR TOI » DE TIKTOK ENTRAÎNE LES ENFANTS ET LES JEUNES MAL DANS LEUR PEAU PLUS LOIN DANS LA DÉPRESSION, L'AUTOMUTILATION ET LES IDÉES SUICIDAIRES 12

4.2 LES CONSÉQUENCES MORTELLES DE LA SPIRALE DES IDÉES SUICIDAIRES DE TIKTOK 15

4.3 LE MODÈLE NOCIF ET LE MANQUE DE MODÉRATION DES CONTENUS DE TIKTOK – EXPÉRIENCES VÉCUES DE JEUNES ET DE PARENTS 17

5. ÉLÉMENTS DE PREUVE ISSUS DE NOS RECHERCHES TECHNIQUES 20

5.1 TESTS MANUELS 20

5.2 TESTS AUTOMATISÉS 22

6. LE CADRE DES DROITS HUMAINS 26

7. LES CADRES JURIDIQUES EUROPÉEN ET NATIONAL 29

7.1 RÉGLEMENTATION DE L'UNION EUROPÉENNE : LE RÈGLEMENT SUR LES SERVICES NUMÉRIQUES 29

7.2 LA PROTECTION DES MINEUR·E·S SUR INTERNET EN FRANCE 31

8. L'INCAPACITÉ DE TIKTOK À RENDRE SA PLATEFORME SÛRE POUR LES ENFANTS ET LES ADOLESCENT·E·S : UNE VIOLATION DU RÈGLEMENT SUR LES SERVICES NUMÉRIQUES ET DES NORMES INTERNATIONALES RELATIVES AUX DROITS HUMAINS 34

9. CONCLUSION ET RECOMMANDATIONS 38

1. SYNTHÈSE

- Deux ans après avoir lancé une première alerte, Amnesty International constate que le fil « Pour toi » de TikTok entraîne les adolescent·e·s dans une « spirale » profondément nocive : une nouvelle expérience a montré que des comptes d'adolescent·e·s manifestant un intérêt pour des contenus relatifs à la tristesse ou à un mal-être psychologique étaient orientés vers des contenus dépressifs en moins d'une heure de navigation sur ce fil.
- Au bout de trois à quatre heures, les utilisateurs et utilisatrices se sont vu proposer des vidéos donnant une vision romanesque du suicide ou montrant des jeunes faisant part de leur intention de mettre fin à leurs jours, avec notamment des informations sur les méthodes de suicide.
- Des témoignages de jeunes ayant pratiqué l'automutilation et de parents endeuillés révèlent les conséquences catastrophiques de ce fil sur la santé mentale et physique des adolescent·e·s mal dans leur peau.

En novembre 2023, Amnesty International avait publié un rapport accablant montrant comment les enfants et les jeunes manifestant un intérêt pour la santé mentale pouvaient être facilement attirés dans une « spirale » de contenus dépressifs et suicidaires sur le fil « Pour toi » de TikTok. Ses recherches avaient révélé que TikTok s'appuyait sur une conception addictive pour générer un maximum d'engagement, portant atteinte aux droits au respect de la vie privée, à la liberté de pensée et à la santé mentale et physique.

Depuis, avec des jeunes de différents pays allant de l'Argentine à l'Irlande en passant par le Kenya et le Pakistan, Amnesty International fait campagne pour que cela change. Plus de 125 000 personnes ont signé une pétition appelant TikTok à rendre sa plateforme plus sûre. Cependant, et tandis que les autorités de régulation commencent bien trop lentement à s'attaquer aux préjudices causés par le secteur des géants technologiques, TikTok ne fait toujours rien pour remédier à son modèle addictif et à ses « spirales » nocives, faisant passer le profit avant la sécurité des enfants.

C'est en France que la campagne mondiale d'Amnesty International a reçu le plus de soutien et suscité le plus de débats. Pour de trop nombreux jeunes et leurs familles, ces préjudices n'ont rien d'abstrait : ce sont des expériences qu'ils ont vécues – et continuent de vivre. Entre 2024 et 2025, 11 familles françaises se sont jointes à un recours collectif contre TikTok afin d'établir la responsabilité de cette entreprise dans la détérioration de la santé mentale et physique de leurs enfants, notamment sa contribution à la mort de deux adolescentes de 15 ans, Marie Le Tiec et Charlize Dapui-Parkiet, et de lui demander des comptes. Ce rapport complémentaire d'Amnesty International rassemble les témoignages de certaines de ces familles et jeunes victimes, et présente de nouvelles preuves qualitatives et quantitatives de la passivité persistante de TikTok face aux risques et préjudices de son modèle économique.

Les éléments de preuve qui y sont présentés montrent que TikTok ne fait rien pour remédier aux risques que son modèle fait courir aux enfants et aux jeunes, en violation des normes internationales relatives à la responsabilité des entreprises en matière de droits humains ainsi que de ses obligations

contraignantes au titre du Règlement sur les services numériques (DSA) de l'Union européenne (UE). Ce rapport lance un appel urgent à l'entreprise elle-même, mais aussi aux autorités de régulation françaises et européennes afin qu'elles agissent fermement pour l'obliger à respecter les droits fondamentaux au respect de la vie privée, à la liberté de pensée et à la santé, notamment des mineur·e·s.

Amnesty International a partagé ses principales conclusions avec TikTok et lui a offert la possibilité de commenter, mais l'entreprise n'a pas fourni de réponse avant la publication.

Pour remédier à ces risques, Amnesty International appelle la Commission européenne à :

- inclure les présentes conclusions dans son enquête en cours sur les possibles violations par TikTok de ses obligations au titre du DSA et veiller à ce que cette enquête et la mise en œuvre de ses recommandations soient promptes et efficaces ;

et TikTok à :

- passer à un modèle économique respectueux des droits qui ne soit pas fondé sur un suivi intrusif des données. Pour commencer, TikTok doit veiller à ce que ses politiques et processus de diligence nécessaire en matière de droits humains s'attaquent aux conséquences systémiques et généralisées qu'a son modèle économique sur les droits humains et en particulier sur les droits au respect de la vie privée, à la liberté d'opinion et de pensée et à la santé ;
- cesser de chercher à maximiser « l'engagement des utilisateurs et utilisatrices » aux dépens de leur santé et autres droits humains. Dans le cadre de son processus de diligence raisonnable en matière de droits humains, TikTok doit identifier, en coopération avec les utilisateurs et utilisatrices, y compris les enfants et les jeunes, ainsi qu'avec des spécialistes indépendants, les éléments de conception qui encouragent une utilisation addictive de la plateforme et des comparaisons sociales, et les remplacer par une expérience axée sur la « sécurité par la conception » et l'intérêt supérieur de l'enfant ;
- cesser de collecter des données personnelles intimes et d'opérer des déductions quant aux intérêts, à l'état émotionnel ou au bien-être des utilisateurs et utilisatrices en fonction de leur engagement et du temps de visionnage des contenus pour « personnaliser » les recommandations et cibler les publicités. Au lieu de se livrer à une surveillance généralisée pour adapter les fils aux intérêts des utilisateurs et utilisatrices, TikTok doit permettre à ceux-ci d'indiquer par eux-mêmes leurs centres d'intérêt de façon volontaire (par exemple, il pourrait leur être demandé d'entrer les sujets qui les intéressent plus particulièrement s'ils/elles souhaitent obtenir des recommandations personnalisées) et seulement sur la base d'un consentement libre, spécifique et éclairé.

2. MÉTHODOLOGIE

Ces recherches ont été engagées à la suite de la publication du rapport de 2023 d'Amnesty International intitulé *Poussé·e·s vers les ténèbres*, qui faisait état des graves risques et préjudices en matière de droits humains que le modèle économique de TikTok fondé sur l'engagement faisait subir ou contribuait à faire subir aux enfants et aux jeunes[2]. Ce rapport concluait que TikTok avait créé une plateforme addictive, en dépit des éléments de plus en plus nombreux attestant les graves risques pour la santé qui découlent d'une utilisation compulsive des réseaux sociaux par les enfants. En étudiant de plus près la personnalisation poussée à l'extrême du fil « Pour toi » de TikTok, ces recherches montraient que les enfants manifestant un intérêt pour des contenus liés à la santé mentale pouvaient facilement être entraînés dans une « spirale » de contenus nocifs, risquant d'exacerber leurs problèmes de dépression et d'automutilation. La campagne mondiale menée par la suite par Amnesty International pour rendre TikTok plus sûre a suscité un soutien et un débat public particulièrement importants en France, où TikTok est réglementée, depuis 2023, par le Règlement sur les services numériques (DSA) de l'UE, qui oblige l'entreprise à identifier et à atténuer les risques systémiques pour des droits de l'enfant. Face à la multiplication des éléments prouvant le danger de TikTok pour les jeunes et les enfants, la Commission européenne a ouvert en février 2024 une enquête sur de possibles violations du DSA par cette plateforme. Dans le contexte de cette enquête en cours, le rapport *Entraîné·e·s dans le « rabbit hole »*, nouvelle collaboration avec l'Algorithmic Transparency Institute de la National Conference on Citizenship, s'appuie sur la méthodologie des analyses relatives aux droits humains et des recherches techniques du rapport de 2023 pour présenter de nouvelles preuves de l'incapacité persistante de TikTok à rendre sa plateforme sûre pour les enfants et les jeunes.

Entretiens

En collaboration avec le collectif Algos Victima, les chercheurs et chercheuses d'Amnesty International ont interrogé trois jeunes ayant sombré dans la dépression et/ou pratiqué l'automutilation sur le rôle de TikTok dans l'exacerbation de leur mal-être psychologique, ainsi que huit mères et pères concernés (dont Stéphanie Mistre, la mère de Marie Le Tiec, et les parents de Charlize Dapui-Parkiet, deux adolescentes qui se sont suicidées). Ils/elles se sont aussi entretenus avec deux ami·e·s proches de Marie, qui ont donné leur avis sur l'utilisation de TikTok par cette jeune fille et par eux-mêmes. Amnesty International France a par ailleurs interrogé la neurologue Servane Mouton à propos des effets de TikTok sur le développement cognitif des adolescent·e·s.

Les chercheurs et chercheuses ont eu des conversations préparatoires avec les parents des jeunes victimes interrogées dans le cadre de ces recherches afin d'évaluer et d'atténuer le risque de nouveau traumatisme et de discuter du soutien psychologique à leur apporter. Une autorisation parentale a été demandée et obtenue pour tous les participant·e·s de moins de 18 ans et le consentement éclairé de chaque adolescent·e a été recueilli. Les participant·e·s de plus de 18 ans ont donné leur consentement

2 Amnesty International, *Poussé·e·s vers les ténèbres. Comment le fil « Pour toi » de TikTok encourage l'automutilation et les idées suicidaires* (index : POL 40/7350/2023), 7 novembre 2023, https://www.amnesty.org/fr/documents/POL40/7350/2023/fr.

en connaissance de cause avant les entretiens. Afin de protéger l'identité des personnes interrogées, ce rapport désigne par un pseudonyme les participant·e·s aux recherches qui ont demandé à rester anonymes (les pseudonymes sont indiqués par un astérisque).

Enquête

Afin d'étudier les comportements du grand public à l'égard des réseaux sociaux et la sensibilisation aux risques chez les jeunes en France, Amnesty International France a collaboré avec l'Institut Harris, qui a réalisé fin 2024 un sondage auprès de 1 000 jeunes âgés de 13 à 25 ans. Les résultats de cette enquête sont présentés dans le chapitre suivant.

Tests manuels

Amnesty International a créé trois faux comptes d'adolescent·e·s de 13 ans (deux filles et un garçon) sur TikTok en France afin de suivre manuellement l'effet de « spirale ». La personne responsable des recherches a fait défiler le fil « Pour toi » et a regardé deux fois les publications qui avaient un rapport avec la tristesse ou des problèmes de santé mentale, passant les autres contenus, pendant trois à quatre heures au total. Le processus a été enregistré puis analysé manuellement pour mieux comprendre quelles formes et quelle ampleur prenait l'amplification algorithmique des contenus sur la santé mentale recommandés aux mineur·e·s sur TikTok.

Tests automatisés

Les données sur l'historique de visionnage d'un des comptes manuels ont ensuite été utilisées pour créer des comptes TikTok automatiques, en plus de comptes automatiques de référence (ne regardant que des contenus bénins), tous enregistrés au nom d'adolescent·e·s de 13 ans, afin d'étudier la personnalisation ou l'effet de « spirale » à plus grande échelle. Les chercheurs et chercheuses ont partitionné cet historique de visionnage en segments contenant au total jusqu'à 387 vidéos pour tester la réponse du système algorithmique de recommandation à différents pourcentages de vidéos liées à la tristesse/la dépression ou l'automutilation/le suicide déjà visionnées.

Avant de commencer les tests automatisés, l'équipe de recherche a utilisé la fonctionnalité d'exportation des données de TikTok pour récupérer l'historique complet de visionnage d'un des comptes manuels. Le fichier ainsi téléchargé contenait les identifiants de toutes les vidéos recommandées à ce compte. Il a été considéré qu'un intervalle de temps de moins de cinq minutes entre deux recommandations indiquait que l'utilisateur·rice avait regardé une vidéo pendant ce temps. Un intervalle de plus de cinq minutes a été considéré comme indiquant un changement de session (le compte ayant été consulté en plusieurs sessions), et la dernière vidéo de chaque session a été ignorée car il était impossible de déduire le temps qui avait été passé à la regarder.

L'équipe de recherche a ensuite reproduit le compte manuel de la façon suivante : pour commencer, elle a créé un nouveau compte TikTok situé en France (en utilisant un VPN). Elle s'est connectée à ce compte depuis un appareil Android, puis a fait en sorte, au moyen d'une programmation, qu'il regarde le nombre de vidéos prédéfini issues de l'historique du compte source, chacune pendant la même durée que le compte source. Ce processus est appelé « clonage ». L'équipe a ensuite attendu 30 minutes pour laisser le temps à l'algorithme de recommandation de se mettre à jour. Puis elle a recueilli, à plusieurs reprises, les huit recommandations suivantes jusqu'à arriver à un total de 201 recommandations de vidéos, sans transmettre à TikTok aucune nouvelle donnée sur le choix de regarder ou de passer les vidéos en question, afin de ne pas influencer de nouveau l'algorithme (cette étape est appelée « récupération »).

Une fois les recommandations récupérées sur un compte, les chercheurs et chercheuses ont effacé les données de l'application TikTok, désinstallé l'application, modifié l'identifiant publicitaire Android (un identifiant semi-permanent de l'appareil permettant aux développeurs d'applications d'identifier les utilisateurs·rices d'une application installée à l'autre) et répété l'opération avec un autre nouveau compte.

Communication avec TikTok

Amnesty International a écrit à TikTok en mars 2025 pour lui demander quelles mesures elle avait mises en place depuis la publication du rapport *Poussé·e·s vers les ténèbres* fin 2023. En réponse, TikTok s'est contentée d'énumérer des mesures de « bien-être » habituelles, dont la plupart étaient déjà en place lors de nos recherches, et n'a pas reconnu le problème de l'effet de « spirale » posé par l'application. L'entreprise n'a pas non plus fourni d'éléments prouvant qu'elle avait adopté de nouvelles mesures ciblées pour résoudre ce problème[3].

3 Amnesty International, *TikTok fails to address risks to children and young people's mental health despite past warnings*, 12 mai 2025, https://www.amnesty.org/fr/documents/act10/9354/2025/en/.

3. LES JEUNES EN FRANCE MOYENNEMENT SENSIBILISÉS AUX RISQUES DES RÉSEAUX SOCIAUX

Une enquête commandée par Amnesty International France et réalisée à la fin de l'année 2024 par l'Institut Harris auprès de jeunes âgés de 13 à 25 ans montre que beaucoup d'entre eux ont conscience de l'effet addictif des réseaux sociaux et de certains des dangers qu'ils représentent. Plus de huit jeunes sur 10 reconnaissent passer trop ou beaucoup trop de temps sur les réseaux sociaux, ce taux étant encore plus élevé chez les jeunes filles (88 %). Dans ses recherches de 2023, Amnesty International avait déjà mis en avant la dimension particulièrement addictive de TikTok, réseau social plébiscité par les jeunes. Les entretiens menés par Amnesty International auprès d'enfants et de jeunes en 2023 et en 2025 montrent cependant que, souvent, ce n'est pas parce que les jeunes ont connaissance des effets addictifs des réseaux sociaux qu'ils sont capables d'arrêter d'en faire un usage excessif.

Par ailleurs, plus de la moitié des jeunes interrogés, en particulier les femmes âgées de 16 à 25 ans, déclarent être souvent exposés à des contenus dérangeants sur les réseaux sociaux, notamment des contenus montrant des corps idéalisés ou promouvant des troubles du comportement alimentaire. L'enquête révèle que 58 % des jeunes interrogés se disent affectés par ces contenus. Plus précisément, ils et elles déclarent se sentir mal à l'aise, tristes, démoralisés ou touchés dans leur estime de soi. Chez les femmes de 16 à 21 ans, ce chiffre monte à 75 %.

Si une grande majorité des jeunes interrogés (69 %) ont déjà cherché à ne plus être exposés à des contenus toxiques, seulement un jeune sur cinq déclare y être réellement parvenu. Pour des personnes vulnérables, ces « spirales » de contenus dérangeants peuvent avoir de graves conséquences dans la vie réelle, comme le montrent les entretiens réalisés par Amnesty International avec des jeunes confrontés à des problèmes de santé mentale en 2023 et 2025.

Enfin, l'enquête révèle que près de la moitié des jeunes interrogés ignorent que les contenus qui leur sont proposés sur les réseaux sociaux sont personnalisés en fonction de leurs données personnelles et de leurs centres d'intérêt. Cette méconnaissance est particulièrement marquée chez les jeunes de 13 à 15 ans (64 %). Les recherches menées par Amnesty International en 2023 avaient montré que TikTok s'immisçait dans la vie privée de ses utilisateurs·rices, en pistant tout ce qu'ils et elles faisaient sur la plateforme, afin de deviner leurs intérêts, leur état émotionnel et leur niveau de bien-être, puis

de leur proposer des contenus « personnalisés » dans le fil « Pour toi[4] ». TikTok, comme toutes les grandes entreprises de réseaux sociaux, cherche à maximiser l'engagement et le temps passé sur la plateforme afin de collecter toujours plus de données personnelles. Ces données lui permettent d'afficher des publicités ciblées et de gagner de l'argent[5]. Amnesty International a constaté à maintes reprises les graves atteintes aux droits humains associées à ce modèle économique fondé sur la surveillance, qui amène les géants de la haute technologie comme Meta, X et TikTok à privilégier l'engagement des internautes et le profit aux dépens de la sécurité de leur plateforme[6].

4 Amnesty International, *« Je me sens vulnérable ». Pris·e·s au piège de la surveillance intrinsèque à TikTok* (index : POL 40/7349/2023), 7 novembre 2023, https://www.amnesty.org/fr/documents/POL40/7349/2023/fr.

5 Le DSA de l'UE interdit les publicités ciblées à destination des mineur·e·s. Cependant, la grande majorité des jeunes utilisateurs et utilisatrices de TikTok dans le monde peuvent, eux, toujours être exposés à des publicités ciblées. Cette interdiction a donc peu de répercussions sur le modèle économique de TikTok.

6 Amnesty International, *Les géants de la surveillance. Le modèle économique de Facebook et Google menace les droits humains* (Extraits) [index : POL 30/1404/2019], 21 novembre 2019, https://www.amnesty.org/fr/documents/pol30/1404/2019/fr ; *L'atrocité des réseaux sociaux. Meta face au droit à réparations des Rohingyas (Synthèse)* [index : ASA 16/5933/2022], 29 septembre 2022, https://www.amnesty.org/fr/documents/asa16/5933/2022/fr ; *« Je me sens vulnérable ». Pris·e·s au piège de la surveillance intrinsèque à TikTok* (index : POL 40/7349/2023), 7 novembre 2023 (op. cit.) ; *"A thousand cuts": technology-facilitated gender-based violence against Poland's LGBTI community on X* (index : EUR 37/0098/2025), 1er septembre 2025, https://www.amnesty.org/fr/documents/eur37/0098/2025/en.

4. TÉMOIGNAGES DE JEUNES VICTIMES ET DE PARENTS CONCERNÉS OU ENDEUILLÉS

4.1 COMMENT LE FIL « POUR TOI » DE TIKTOK ENTRAÎNE LES ENFANTS ET LES JEUNES MAL DANS LEUR PEAU PLUS LOIN DANS LA DÉPRESSION, L'AUTOMUTILATION ET LES IDÉES SUICIDAIRES

Pour mieux comprendre les risques que la personnalisation à l'extrême du fil « Pour toi » de TikTok fait courir aux adolescent·e·s et aux jeunes adultes qui sont confrontés à des problèmes de santé mentale courants, souvent liés à des phénomènes de harcèlement et d'exclusion sociale, Amnesty International s'est entretenue avec des jeunes concernés et leurs parents en France. Les chercheurs et chercheuses ont aussi rencontré les parents et des ami·e·s de Marie, qui s'est suicidée à l'âge de 15 ans en 2021, et de Charlize, morte au même âge en novembre 2023. Ces personnes ont témoigné pour sensibiliser aux risques qui, selon eux, sont encore mal connus des parents, des écoles, des responsables politiques et même des professionnel·le·s de la santé mentale, ainsi que pour réclamer des comptes à TikTok et des mesures efficaces aux autorités de régulation.

Maëlle, 18 ans, a raconté à Amnesty International comment elle avait été entraînée dans une « spirale » de contenus liés à la dépression et à l'automutilation dans le fil « Pour toi » de TikTok peu de temps après s'être inscrite sur la plateforme, début 2021[7]. Une amie l'avait encouragée à installer cette application pour qu'elles puissent partager des vidéos amusantes et divertissantes à une période où Maëlle était mal dans sa peau à cause du harcèlement qu'elle subissait à l'école.

> *« J'ai d'abord eu ces vidéos qui étaient vraiment l'image qu'on se faisait de l'application […]. Plutôt divertissantes, pas très instructives, mais divertissantes. Et il y a eu une musique qui est revenue beaucoup de fois parce que je pense qu'elle venait de sortir. […] Et au fur et à mesure, je me suis intéressée aux paroles et je ressens quelque chose en moi, je trouve qu'elle me touche. Elle parle du mal-être du chanteur et de moments sombres. Et j'avais déjà commencé à avoir des comportements dommageables auparavant. Et donc j'ai vu de plus en plus de vidéos avec cette musique. […] Et de fil en aiguille, c'était quelque chose de plus en plus sombre, du style la mort ce n'est peut-être pas une si mauvaise idée ou des choses comme ça. »*

7 Entretien réalisé à Paris le 8 juillet 2025.

Quelques semaines après son inscription sur la plateforme, Maëlle a vu son fil TikTok devenir un flux de vidéos normalisant, voire encourageant, l'automutilation et le suicide. Elle estime que jusqu'à trois quarts des vidéos recommandées traitaient de problèmes de santé mentale d'une manière qui l'entraînait encore plus dans des comportements dommageables. En septembre 2022, Maëlle a été hospitalisée une première fois. Cependant, ni ses parents, ni ses psychologues ne savaient ce qu'elle regardait en ligne.

> *« Il y a une partie des contenus qui relève de la glamourisation. Des jeunes qui flirtent avec l'interdit. Et il y a des choses vraiment trash, des gens à hôpital ou qui ne vont vraiment pas bien. On sent qu'ils sont sur la corde raide. Ce n'est pas le même objectif mais au final sur l'utilisateur ça a quand même le même effet. […] Ça m'a attiré encore plus, ça me maintenait la tête sous l'eau. Et ça m'a fait culpabiliser de pas aller aussi mal que les gens sur la plateforme. »*

> *« D'un côté, je ressentais un sentiment d'appartenance. Je me sentais moins seule parce que je me disais que d'autres personnes vivaient ces choses-là. Mais pour autant, je savais que c'était malsain comme situation et très vicieux. »*

> *« Il n'y avait jamais quelque chose d'encourageant. Comme "allez consulter si vous vous sentez mal", des choses comme ça. C'était toujours quelque chose de très sombre. »*

La santé mentale de Maëlle s'est encore détériorée durant l'hiver 2023-2024.

> *« Il y a eu une période où je n'avais plus la force de rien faire. Je ne faisais plus de sport. Je n'allais quasiment plus en cours. Pour m'occuper, j'allais sur TikTok et ça me maintenait la tête sous l'eau. »*

> *« Voir des gens qui se scarifient en direct, des gens qui disent quoi prendre comme médicaments pour en finir, ça influence et ça t'incite à avoir des comportements délétères. Il y a des vidéos […] qui sont encore imprimées dans ma rétine. Il y a des choses qui restent, des comportements. »*

La mère de Maëlle, Morgane, n'a découvert ce type de contenus qu'en 2024, lorsqu'une émission de radio sur les débuts d'Algos Victima a déclenché une conversation avec sa fille sur sa propre expérience[8]. Elle a déclaré à Amnesty International :

> *« Quand j'ai découvert cette histoire de contenu sur TikTok, ça a été un peu la douche froide. Mais j'ai très vite compris dans quoi elle s'était enfermée. La honte et la culpabilité qu'elle pouvait ressentir de ne pas pouvoir nous en parler. […] Mais je ne pouvais même pas imaginer que ça puisse exister. Elle est devenue addict au contenu mortifère. L'addiction impacte votre vie sociale, votre santé, votre travail ou votre scolarité. Clairement, ça avait impacté tous ces milieux-là. Donc pour moi elle est devenue addict à ces contenus-là. […] Quoi que nous fassions, quoi que nous disions, c'est comme si elle était happée dans ce monde parallèle et que le monde réel n'avait plus d'impact sur ses émotions ou sur qui elle était. »*

Édouard, 17 ans, a décrit le même effet captivant des contenus nocifs sur TikTok. Il a été hospitalisé pour une grave dépression. Il regardait régulièrement des vidéos TikTok republiées sur Instagram[9].

> *« J'étais en grosse dépression. Petit à petit, je me sentais encore plus mal et je voulais aller plus loin, plus fort, plus profond. Et petit à petit, je descendais dans les abysses. […] Je tombais sur des vidéos de gens qui se scarifiaient. Pas de l'acte, mais des cicatrices. Je me disais "il faut que j'atteigne ce niveau de cicatrices sinon je ne suis pas content". Je voulais vraiment être marqué, de plus en plus profond. […] Je voulais me voir détruit, je ne m'aimais plus. Alors autant me détruire vraiment mentalement et sur le corps, physiquement. »*

8 Entretien réalisé à Paris le 8 juillet 2025.

9 Entretien réalisé près de Toulon le 29 juillet 2025.

La mère d'Édouard, Géraldine, pense que ces vidéos ont joué un rôle crucial dans la détérioration de son état de santé :

> *« Édouard ne se sentait pas très bien mais les vidéos en ligne ont été un catalyseur. Je ne dis pas que c'est l'origine de son mal-être, mais ça l'a décuplé, comme une machine qui l'a aspiré[10]. »*

Même s'il regardait principalement des contenus TikTok republiés sur Instagram, Édouard a déclaré à Amnesty International que, après avoir récemment comparé la disponibilité des contenus liés à l'automutilation et au suicide sur les deux plateformes, il avait constaté que sur TikTok « c'est vraiment beaucoup plus facile à trouver. On tape directement le drapeau de la Suisse ou le zèbre et on tombe sur ces vidéos ». (Le drapeau suisse est un trompe-algorithme courant pour évoquer le suicide et le zèbre pour faire référence à l'automutilation.)

De même, **Maëllys**, 17 ans, a décrit comment elle avait été aspirée dans une communauté TikTok complètement coupée de sa vie réelle, qui avait renforcé ses pensées suicidaires :

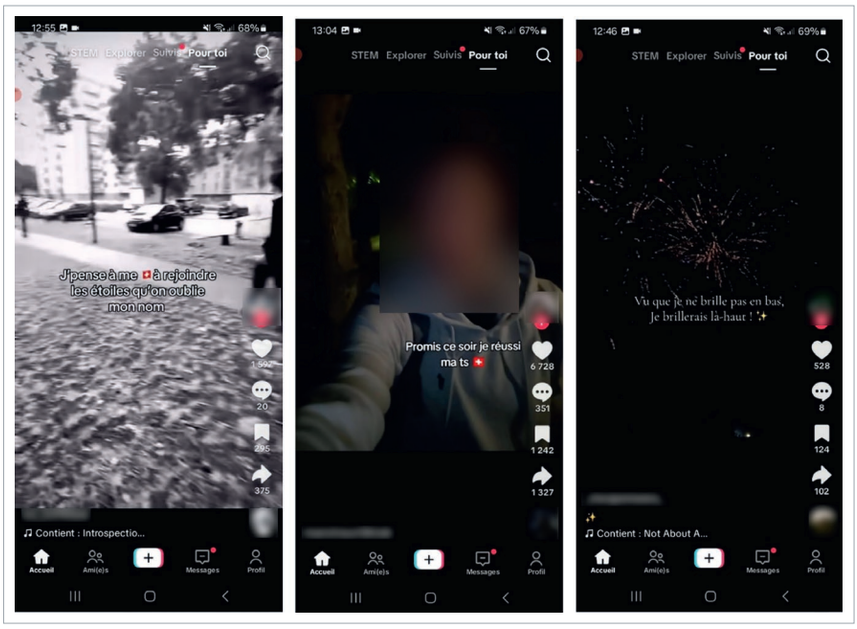
> *« Au départ, je regardais des danses, des playbacks de Disney, de téléréalité, de scènes de films. […] En 5ème, j'ai commencé à avoir des trucs tristes, j'ai très vite liké ces vidéos parce que je me sentais concernée. Puis TikTok me les a reproposées et tous les contenus de mon fil d'actualité étaient tristes, parlaient de dépression […]. C'était vraiment catastrophique, mais à ce moment-là, on ne s'en rend pas compte parce qu'on se dit qu'on n'est pas seule, que ça arrive à d'autres gens. On se sent écoutée. Mais le problème c'est qu'on se renferme sur nous-même parce qu'on se dit que seules les personnes sur TikTok peuvent nous comprendre, nous écouter et vivent les mêmes choses que nous. »*
>
> *« Plus je regardais, plus je trouvais des "solutions", par exemple des cachettes pour des lames de rasoir. […] Plus tu regardes, plus tu t'enfonces, plus tu te complais dans ton mal-être. Tu n'as plus de souvenirs d'avant, quand tu étais bien, quand tu te sentais heureuse. Tu te dis que tu as toujours été malheureuse. […] On reste là-dedans, on continue et c'est un cercle vicieux. »*
>
> *« Après ma sortie de l'hôpital, j'ai continué à aller sur TikTok. Toujours les mêmes choses, mais en pire […]. Il y avait des jeux comme celui du Labello : tu te mets du Labello tous les jours et le jour où le Labello est vide, tu te suicides. »*

Le « Labello challenge » ou « jeu du baume à lèvres » a fait l'objet de nombreux reportages dans les médias en 2022 ; à l'époque, TikTok a nié l'existence de vidéos sur ce thème sur sa plateforme[11]. Pourtant, les chercheurs et chercheuses d'Amnesty International sont tombés sur deux vidéos de ce type dans le fil d'un des comptes tests gérés manuellement en juillet 2025 (voir chapitre 5.1).

Maëllys a finalement réussi à prendre de la distance et à rompre le cercle vicieux qu'elle a décrit après des années de suivi psychologique et d'hospitalisations pour des actes d'automutilation et des tentatives de suicide. Elle utilise toujours TikTok mais elle dit que, à 17 ans, elle se sent maintenant davantage capable de ne pas se laisser aspirer dans une spirale de contenus nocifs lorsque des vidéos sur l'automutilation ou le suicide apparaissent dans son fil.

10 Entretien réalisé près de Toulon le 29 juillet 2025.

11 Euronews, "A new self-harm Tik Tok 'challenge' is worrying France. But is the panic missing the point?", 20 mai 2022, https://uk.news.yahoo.com/self-harm-tik-tok-challenge-135716162.html.



Exemples de vidéos recommandées sur des comptes tests gérés manuellement dans les quatre heures suivant l'inscription sur la plateforme (faux comptes de personnes enregistrées comme ayant 13 ans et manifestant un intérêt pour les contenus liés à la tristesse et aux problèmes de santé mentale). [drapeau suisse = suicide]

4.2 LES CONSÉQUENCES MORTELLES DE LA SPIRALE DES IDÉES SUICIDAIRES DE TIKTOK

Marie et Charlize n'ont pas échappé à la spirale d'idées suicidaires dans laquelle leur fil TikTok les a entraînées, ont raconté leurs parents à Amnesty International. Les deux familles n'ont découvert la vie cachée de leurs filles respectives sur Internet qu'après leur mort par suicide en 2021 et 2023. Stéphanie Mistre, la mère de Marie, et deux des ami·e·s les plus proches de la jeune fille ont raconté que **Marie** donnait l'impression d'aller bien à toutes les personnes qui la connaissaient dans la vraie vie. Elle a toujours été décrite comme une jeune fille rayonnante[12]. Elle n'avait aucun antécédent d'automutilation. Pourtant, à l'insu de sa mère, Marie avait fait part de son mal-être psychologique dans des vidéos TikTok, semble-t-il à la recherche de soutien et d'approbation de la part de la communauté en ligne sur son fil.

> *« Je ne savais pas du tout ce qu'était TikTok. Jusqu'à ce que Marie s'en aille et que, un mois après, j'aille voir son téléphone. Là, je suis tombée face à l'horreur : les chansons qui prônent le suicide, les contenus glauques, noirs, toujours sombres, avec une voix qui te dit "tu ne vas pas bien, tu te sens seul". Des choses qui vous mettent dans l'ambiance, sans parler des tutos sur comment se scarifier ou comment faire un nœud pour se pendre. »*

Stéphanie est convaincue que Marie a été entraînée vers le fond par ce flux d'idées suicidaires et la confortation toxique auxquels elle était exposée sur la plateforme. Une de ses vidéos dans laquelle elle se filmait en train d'exprimer son mal-être en pleurant a reçu des centaines de likes. Stéphanie a été profondément choquée par les commentaires sur le fil de sa fille, qui idéalisaient le suicide et l'encourageaient ouvertement à mettre fin à ses jours.

> *« Pour ces plateformes, nos enfants deviennent des produits et non plus des êtres humains. On utilise nos enfants comme des produits avec un algorithme et une bulle de filtres, en se servant de leurs émotions pour les capter. L'algorithme va capter vos centres d'intérêt, ce qui n'est pas*

12 Entretiens réalisés à Paris le 16 juillet 2025.

normal. On rentre dans la vie privée de l'enfant. Mais les enfants ont des droits. Là, on abuse de leur fragilité et de leur faiblesse, d'une adolescence qui est reconnue par tous les psychologues comme une période de fragilité où on se cherche, et encore plus pour une fille. Ce qui était le cas de Marie. »

Après la mort de Marie, ses ami·e·s **Manon** et **Hugo** ont eux-mêmes été exposés aux effets nocifs de la personnalisation à l'extrême de TikTok.

« Quand on n'est pas bien, on n'a que des vidéos pas bien. Parce qu'en fait, ça nous maintient dans le mood. Après le départ de Marie, pendant deux mois, je n'avais que des vidéos sur la mort, le suicide[13]. » (Manon)

« À un moment, je n'étais pas très bien. Et on trouve comme un réconfort dans les vidéos, on like et elles viennent comme un tsunami[14]. » (Hugo)

Charlize était une jeune fille sensible, elle voulait devenir psychologue pour aider les autres[15]. Les confinements liés au COVID19 ont été des moments difficiles pour elle et sa famille. Son grand-père paternel est décédé et sa mère, travaillant à l'hôpital local, était en première ligne de la lutte contre la pandémie. Quand Charlize a été victime de harcèlement à l'école, elle a commencé à passer plus de temps en ligne, mais son père a raconté qu'il la voyait rire devant des vidéos sur son téléphone et qu'il était content qu'elle ait trouvé quelque chose qui la fasse rire plutôt que de ressasser ses idées noires. Elle était suivie par une psychologue car elle avait commencé à se scarifier et avait même tenté de se suicider. Cependant, comme dans le cas de Marie, les parents de Charlize, Delphine et Jérémy, n'ont découvert la vie de leur fille sur Internet et les sentiments qu'elle exprimait en ligne que quand il a été trop tard. L'adolescente a mis fin à ses jours en novembre 2023, à l'âge de 15 ans.

Le lendemain de sa mort, Delphine et Jérémy ont appris que Charlize avait republié sur TikTok, peu avant son décès, une vidéo idéalisant des pensées suicidaires explicites. Ils ont alors découvert qu'elle republiait des vidéos exprimant des idées suicidaires depuis des mois. « TikTok a maintenu notre fille dans le mal-être », a déclaré Jérémy à Amnesty International[16]. « Comment peuvent-ils nier le fait que des contenus qui affirment : "tu as des problèmes ? Suicide-toi, ça ira beaucoup mieux" incitent à passer l'acte ? […] Signaler les vidéos, ça ne marche pas. En plus, le problème, c'est la fonctionnalité du scroll infini. Il faudrait passer ses journées à signaler. Donc ça ne peut pas relever de la responsabilité des utilisateurs. »

Ces remarques font écho à celles de Sophie*, mère d'**Emma***, qui a été exposée à des contenus nocifs sur l'automutilation et les troubles du comportement alimentaire sur TikTok[17]. Sophie* est convaincue que ces vidéos ont aggravé les problèmes d'Emma* dans ces domaines. Bien qu'elle ait limité le temps d'écran de sa fille via le contrôle parental, elle a observé que celle-ci était entraînée dans une « spirale » de contenus normalisant les problèmes de santé mentale au point que des adolescent·e·s comme sa fille « se fabriquent une identité » autour du fait d'aller mal[18]. Selon elle, ni les contrôles parentaux mis en place par sa famille, ni la modération des contenus par TikTok n'ont été suffisants pour remédier aux risques systémiques auxquels sa fille a été exposée.

« Aujourd'hui, ma fille signale tous les contenus qu'elle estime dangereux. Mais elle dit qu'ils ne sont jamais supprimés, qu'il ne se passe jamais rien. »

13 Entretien réalisé à Paris le 16 juillet 2025.

14 Entretien réalisé à Paris le 16 juillet 2025.

15 Entretien réalisé à Nice le 29 juillet 2025.

16 Entretien réalisé à Nice le 29 juillet 2025.

17 Sophie* et Emma* sont des prénoms d'emprunt utilisés à la demande des participantes.

18 Entretien réalisé à Paris le 15 juillet 2025.

Her daughter continued to use TikTok under a strict time limit of 15 minutes per day until early 2025 when she decided to quit using the app. Up until she stopped using the app, she continued to be exposed to and report what she perceived to be harmful content.

4.3 LE MODÈLE NOCIF ET LE MANQUE DE MODÉRATION DES CONTENUS DE TIKTOK – EXPÉRIENCES VÉCUES DE JEUNES ET DE PARENTS

Les témoignages des jeunes concernés et des familles et ami·e·s de Marie et de Charlize viennent corroborer ceux des enfants, des jeunes et des professionnel·le·s de la santé mentale interrogés par Amnesty International en 2023 pour son rapport *Poussé·e·s vers les ténèbres*. Comme l'organisation l'avait alors constaté, les risques liés à la conception de TikTok sont systémiques et intrinsèquement liés à son modèle économique. TikTok, comme ses concurrents sur le marché des réseaux sociaux, cherche à inciter ses utilisateurs et utilisatrices, nouveaux ou anciens, à passer le plus de temps possible sur sa plateforme afin de générer des revenus publicitaires. Pour cela, elle intègre des fonctionnalités addictives, en premier lieu desquelles le système de recommandation extrêmement invasif pour la vie privée que constitue le fil « Pour toi ». Ce système transforme les données recueillies sur les vidéos visionnées et l'engagement de l'utilisateur·rice en signaux à partir desquels le système déduit les contenus qui vont amener la personne à rester le plus longtemps possible sur son écran, avec l'aide d'autres fonctionnalités comme la possibilité de faire défiler les contenus à l'infini et le lancement automatique des vidéos. Ces fonctionnalités addictives intégrées au modèle de TikTok encouragent des usages néfastes pour la santé qui, comme l'ont montré des membres de professions médicales, entraînent de graves risques sanitaires pour les enfants et les adolescent·e·s[19].

La neurologue Servane Mouton en a expliqué les mécanismes et les effets dans un entretien accordé à Amnesty International en 2025 :

> *« Les comportements sur les réseaux sociaux se rapprochent de l'addiction, avec cette difficulté à décrocher, l'empiétement sur d'autres activités, l'obsession des pensées en boucle autour d'y avoir accès. Le manque quand on n'y est pas. Le modèle économique des réseaux sociaux va se baser sur la stimulation du système de récompense à court terme pour qu'on y vienne le plus souvent possible, le plus longtemps possible, etc. C'est un peu comme courir après des shots de dopamine qui vont produire tout de suite la sensation de plaisir et nous donner envie d'y retourner. »*

Comme les recherches de 2023 d'Amnesty International et ces nouveaux témoignages le montrent, ces choix de conception de la plateforme créent des risques supplémentaires graves pour les enfants et les jeunes dont l'intérêt pour la santé mentale les prédispose à être entraînés dans des « spirales » nocives du type de celles décrites par Maëlle, Maëllys et d'autres participant·e·s aux recherches.

Les psychologues interrogés en 2023, ainsi que de précédentes recherches indépendantes, ont aussi souligné les conséquences négatives des fonctionnalités encourageant la comparaison sociale entre les enfants ou jeunes et leurs pairs, telles que la mention « J'aime » (cœur), sur la santé mentale des adolescent·e·s[20].

---

19 Administrateur de la santé publique des États-Unis, *Social Media and Youth Mental Health*, mai 2023, hhs.gov/surgeongeneral/priorities/youth-mental-health/social-media/index.html ; Qinghua He, Ofir Turel et autres, "Brain anatomy alterations associated with Social Networking Site (SNS) addiction", *Scientific reports*, vol. 7, article 45064, mars 2017, https://www.nature.com/articles/srep45064 ; Yvonne Kelly, Afshin Zilanawala et autres, "Social Media Use and Adolescent Mental Health: Findings From the UK Millennium Cohort Study", *EClinicalMedicine*, vol. 6, p. 59-68, décembre 2018, https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370 (18) 30060-9/fulltext.

20 Amnesty International, *Poussé·e·s vers les ténèbres. Comment le fil « Pour toi » de TikTok encourage l'automutilation et les idées suicidaires*, 7 novembre 2023 (op. cit.)

Par exemple, Maëlle et Édouard ont fait part de leurs réflexions sur le bouton « J'aime » (permettant d'ajouter un cœur) en ce qui concerne les vidéos sur l'automutilation et les idées suicidaires :

> *« Quand j'ai liké la vidéo, c'était pour encourager, donner de la force, soutenir la personne dans le fait qu'elle puisse aller bien… Mais j'ai réalisé plus tard qu'en fait ça voulait peut-être dire pour la personne "Ah, on m'encourage à avoir des comportements auto-dommageables ou des choses comme ça". (Maëlle)*

> *« Le côté soutien en ligne, ça peut aussi être interprété comme "vas-y suicide toi". » (Édouard)*

Une autre constatation commune ressortant des témoignages des jeunes est le fait que, après avoir été entraînés dans la « spirale », ils et elles ont été confortés dans leurs idées négatives par la communauté en ligne dans laquelle le fil les a plongés. Comme Édouard, qui a raconté son envie d'atteindre le niveau d'automutilation qu'il voyait dans des vidéos TikTok publiées par d'autres, Maëlle a expliqué comment les commentaires incitaient les internautes à aller plus loin dans leurs comportements d'automutilation, voire les encourageaient à mettre fin à leurs jours.

> *« Il y avait des commentaires comme : si tu fais une tentative de suicide mais que tu ne te retrouves pas en réanimation ou à l'hôpital, ça n'a pas compté. Si tu n'es pas en sous-poids tu ne peux pas être anorexique. Si tu n'as pas de points de suture, c'est que ça ne compte pas, des choses comme ça qui étaient très culpabilisantes. Une scarification, profonde ou pas, ça reste quand même une scarification, et le geste est un appel à l'aide dans les deux cas. »*

Elle a ajouté :

> *« J'allais rarement dans les commentaires. Mais quand j'y allais, je trouvais ça encore plus sombre que la vidéo en elle-même. Parce que là, c'était : "Courage ma belle, tu vas y arriver demain, c'est la bonne journée. Tu vas réussir à mettre fin à tes jours. […]" Et ça faisait boule de neige. »*

Ces observations donnent un éclairage psychologique important sur les communautés en ligne liées aux idées suicidaires. Sur le plan technologique, elles soulignent également les manquements évidents du processus de modération des contenus de TikTok, qui devrait faire en sorte que les contenus promouvant ou encourageant l'automutilation et le suicide soient rapidement supprimés. Elles posent aussi d'autres questions sur les risques potentiels du modèle liés au classement des commentaires en fonction de l'engagement et à la gestion algorithmique des audiences, qui permet de cibler automatiquement des groupes spécifiques d'utilisateurs et utilisatrices en leur proposant des types de contenus particuliers, ce qui peut encourager le développement de « bulles de filtres » ou de « caisses de résonance[21] ».

Si les conclusions d'Amnesty International soulignent l'importance de s'attaquer non seulement aux insuffisances de la modération de contenu mais aussi à l'effet de « spirale » (résultant de l'amplification algorithmique des contenus sur la base du profil des utilisateurs·rices) pour prévenir les atteintes à la vie privée qui facilitent de telles concentrations dangereuses de contenus et leurs incidences néfastes, les exemples de la vie réelle issus de ces témoignages montrent aussi que les systèmes de modération des contenus de TikTok ne parviennent pas à empêcher la diffusion de contenus de toute évidence contraires à la loi promouvant l'automutilation et le suicide. Lætitia, la mère de Maëllys, a raconté que, quand elle avait découvert la vie en ligne de sa fille, elle avait passé des mois à signaler des contenus nocifs pour la santé mentale sur TikTok, mais qu'elle avait fini par laisser tomber après avoir constaté que, souvent, après contrôle, ces vidéos étaient jugées acceptables. Elle a déclaré à Amnesty International :

---

21 Owen Doyle, "Algorithms and the 'Anti-Preference': A Quantitative Investigation of 'Reaching the Wrong Audience' on TikTok", *Berkeley Undergraduate Journal*, vol. 38 (1), 2024, https://escholarship.org/uc/item/2vq7291w ; Kai Riemer et Sandra Peter, "Algorithmic Audiencing: Why We Need to Rethink Free Speech on Social Media", *Journal of Information Technology*, vol. 36, n° 4, p. 409-426, décembre 2021, https://journals.sagepub.com/doi/10.1177/02683962211013358.

« Pendant six mois, je signalais tout le temps. Dès que j'avais 5 minutes au travail, mes pauses déjeuner, mes weekends, je passais des heures et des heures à signaler des vidéos. Pour rien. Dans 90 % des cas, on nous répond que ça n'enfreint pas les règles de la communauté. Au bout d'un moment, j'ai arrêté parce que j'étais trop mal. Je ne vivais plus que pour signaler les vidéos TikTok. »

Le chapitre 8 se penche de manière plus approfondie sur l'incapacité de TikTok à identifier, atténuer et prévenir les risques et les préjudices liés à ces choix de conception de la plateforme et à ces décisions économiques.

5. ÉLÉMENTS DE PREUVE ISSUS DE NOS RECHERCHES TECHNIQUES

5.1 TESTS MANUELS

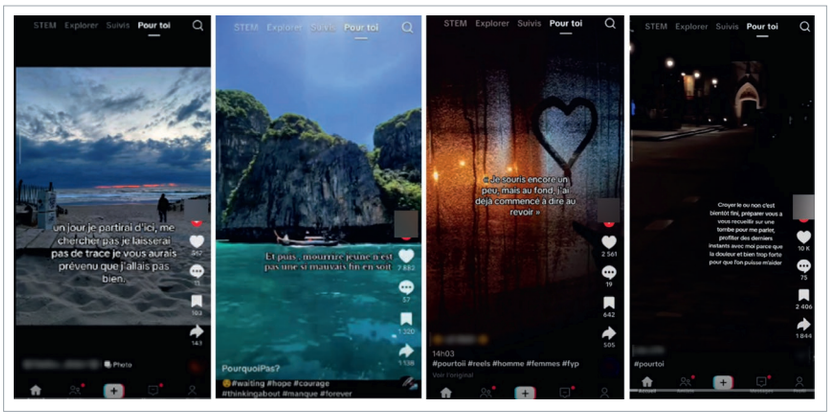
En juillet et en août 2025, Amnesty International a créé trois faux comptes d'adolescent·e·s de 13 ans (deux filles et un garçon) sur TikTok en France afin d'étudier manuellement l'amplification algorithmique des contenus sur le fil « Pour toi » de TikTok. Les chercheurs et chercheuses ont regardé plusieurs fois les contenus recommandés quand ils étaient liés à des sentiments de tristesse ou à des questions de santé mentale, et ont passé les vidéos sur d'autres sujets. Amnesty International ne prétend pas que ce comportement est totalement représentatif de celui des véritables utilisateurs et utilisatrices. Cependant, comme le système de recommandation de TikTok est connu pour utiliser le temps de visionnage comme un signal clé pour faire des déductions sur les préférences des utilisateurs·rices, ces expériences apportent un éclairage crucial sur les systèmes de recommandation de cette plateforme. Elles permettent en particulier de déterminer l'impact du temps de visionnage sur les contenus recommandés relatifs à la tristesse ou à la santé mentale, ainsi que l'efficacité des mesures prises pour atténuer les risques pour les droits de l'enfant en vertu du DSA de l'UE – mesures que l'entreprise a mises en place dans le but déclaré d'« empêcher l'affichage d'une concentration de contenus sur des sujets tels que les régimes, la perte de poids, la tristesse, la dépression, la solitude et le désespoir[22] ».

Sur les trois comptes, les chercheurs et chercheuses sont tombés sur des vidéos portant sur le sentiment de tristesse ou de désenchantement dans les cinq premières minutes de défilement du fil « Pour toi », avant même d'avoir commencé à manifester de quelconques préférences. Le fait de regarder ces vidéos a ensuite entraîné une multiplication rapide des contenus sur la tristesse et la santé mentale. Dans les 15 à 20 minutes après le début de l'expérience, les trois fils contenaient presque exclusivement des vidéos sur la santé mentale, dont jusqu'à la moitié sur des contenus tristes et dépressifs. Sur deux des trois comptes, des vidéos exprimant des pensées suicidaires sont apparues dans les 45 minutes. Après moins de trois heures sur la plateforme (sur la base de sessions d'une heure par jour), tous les comptes proposaient des contenus exprimant de telles pensées, par exemple des légendes de photos telles que « Je suis actuellement dans une période où je me demande si je dois réellement continuer à exister ou pas ». Dans certains cas, les vidéos citaient des chansons célèbres dont les paroles contenaient des allusions romantiques au désespoir, à l'idée de s'isoler ou de

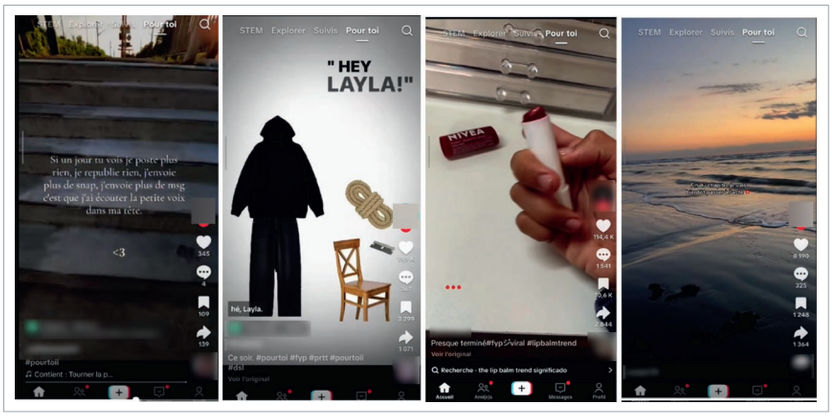
22 TikTok, *DSA Risk Assessment Report 2024*, 28 août 2024, https://www.tiktok.com/transparency/en/dsa-transparency/.

partir loin, réinterprétées par les utilisateurs et utilisatrices comme une évocation du suicide (« Un jour je partirai d'ici, me cherchez pas je laisserai pas de trace, je vous aurais prévenu que je n'allais pas bien. »).

Dans plusieurs vidéos, des internautes faisaient part de leur intention de se suicider le jour même. Les deux comptes de filles ont reçu des recommandations de contenus présentant ou faisant allusion à des méthodes spécifiques de suicide dans la troisième ou la quatrième heure sur la plateforme. L'un des deux s'est vu proposer deux vidéos sur un défi connu incitant au suicide, évoqué dans les médias français dès 2022 (voir chapitre 4.1). Les contenus suicidaires s'appuient souvent sur un langage romantique et des « trompe-algorithmes » (l'utilisation d'émoticônes ou le remplacement de lettres par des chiffres pour éviter la détection par les outils automatiques de modération de contenu). En comparant les résultats de ces tests manuels récents et les témoignages des jeunes victimes, on constate qu'au moins une émoticône couramment utilisée pour symboliser le mot « suicide » semble avoir cours depuis de nombreuses années sans pour autant déclencher de mesures appropriées pour éviter la diffusion de contenus nocifs.



Vidéo 1 *« un jour je partirai d'ici, me chercher pas je laisserai pas de trace je vous aurais prévenu que j'allais pas bien » (sic)* ***vidéo 2*** *« Et puis, mourrire jeune n'est pas une si mauvais fin en soit » (sic)* ***vidéo 3*** *« Je souris encore un peu, mais au fond, j'ai déjà commencé à dire au revoir »* ***vidéo 4*** *« Croyez le ou non c'est bientôt fini, préparer vous a vous recueillir sur une tombe pour me parler, profiter des derniers instants avec moi parce que la douleur et bien trop forte pour que l'on puisse m'aider »*



vidéo 5 *« Si un jour tu vois je poste plus rien, je republie rien, j'envoie plus de snap, j'envoie plus de msg [messages] c'est que j'ai écouter la petite voix dans ma tête »* ***vidéo 8*** *« fin du chapitre je vais bientôt passer à l'acte [drapeau suisse, un trompe-algorithme courant pour évoquer le suicide] »*

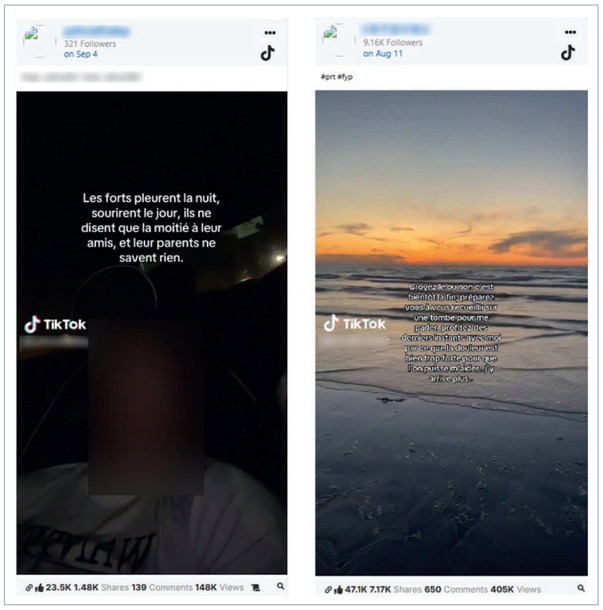
Vidéos recommandées sur le fil d'un faux compte d'adolescente de 13 ans dans le cadre d'un test manuel trois à quatre heures après la manifestation d'un intérêt pour les contenus sur la santé mentale sur le fil « Pour toi » de TikTok. Le compte a été entraîné dans une spirale de contenus mélancoliques et dépressifs en moins d'une heure, puis les contenus sur des idées suicidaires se sont multipliés dans les troisième et quatrième heures. Contrairement aux affirmations de TikTok selon lesquelles la plateforme désagrège les ensembles de contenus potentiellement nocifs, Amnesty International n'a trouvé aucune preuve d'un quelconque effort pour interrompre le flux continu de contenus néfastes.

Point particulièrement important, ces expériences ont révélé que les mesures d'atténuation des risques que TikTok a déclaré prendre au titre du DSA pour empêcher la « concentration des contenus » et mettre en œuvre de façon préventive la charte de sa communauté étaient inefficaces (voir aussi le chapitre 8). En seulement quelques heures, ces comptes ont été entraînés dans une spirale de contenus liés à la santé mentale de nature potentiellement préjudiciable, avec au fil du temps un nombre croissant de vidéos idéalisant et normalisant le suicide. Aucune de ces vidéos ne semblait avoir été réalisée par des professionnel·le·s de la santé mentale ou des organisations d'aide.

5.2 TESTS AUTOMATISÉS

Après avoir réalisé les expériences manuelles, les chercheurs et chercheuses ont mené une série de tests automatisés afin d'étudier le système de recommandation algorithmique qui se cache derrière le fil « Pour toi » de TikTok. L'expérience automatisée s'appuie sur les tests manuels : elle utilise le véritable historique de visionnage d'un des comptes (résultant de ce que le chercheur ou la chercheuse a regardé dans le fil « Pour toi » dans le cadre de l'expérience manuelle) pour construire des comptes automatiques et évaluer à plus grande échelle l'amplification algorithmique des contenus nocifs.

Les tests automatisés ont confirmé l'effet potentiellement dangereux de la personnalisation en ce qui concerne les contenus classés comme tristes ou dépressifs. Comme les comptes manuels, chaque compte automatique a été créé au nom d'un faux utilisateur ou d'une fausse utilisatrice de 13 ans. La méthodologie de collecte des données est décrite au chapitre 2. L'analyse ci-dessous détaille les conclusions tirées de la classification manuelle des recommandations de vidéos récupérées sur les comptes automatiques, dont les préférences ont été exprimées par des historiques de visionnage issus de différentes partitions de l'historique d'un des comptes manuels.



*La vidéo de **gauche** a été classée dans la catégorie des contenus **tristes/dépressifs.***

*Celle de droite a été mise dans la catégorie des contenus sur **l'automutilation/les idées suicidaires**. Cette dernière avait déjà été vue par plus de 400 000 internautes quand elle a été recommandée à notre compte.*

Exemples de classification des contenus :

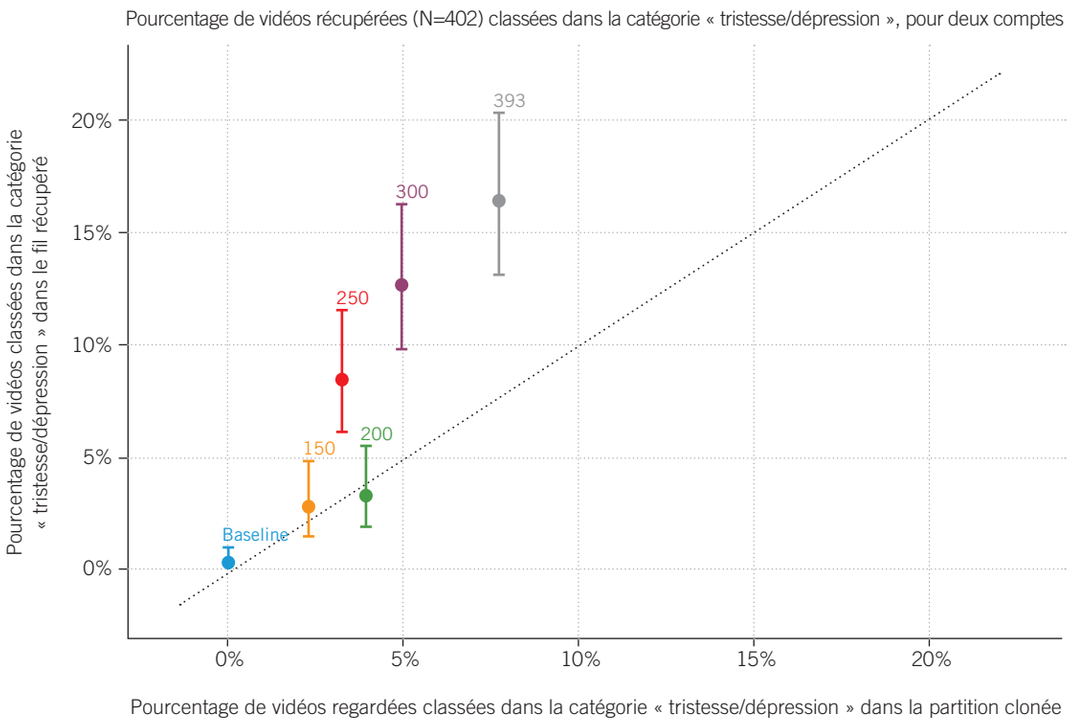
Les diagrammes de dispersion ci-dessous (graphiques 1 et 2) montrent les taux de personnalisation en réponse aux différentes proportions de contenus tristes/dépressifs ou concernant l'automutilation/les idées suicidaires figurant dans cinq partitions du fil cloné. Après avoir attribué chaque partition à deux comptes, les chercheurs et chercheuses ont récupéré 201 recommandations de vidéos sur chacun

de ces comptes[23]. À des fins d'analyse, les vidéos recommandées pour les deux comptes associés à la même partition ont été regroupées, ce qui donne un échantillon total de 402 vidéos.

L'axe des abscisses représente le pourcentage de vidéos potentiellement nocives (tristes/dépressives dans le graphique 1, automutilation/idées suicidaires dans le graphique 2) figurant dans les données importées (le fil cloné). Les vidéos ont été comptées uniquement si elles portaient sur les sujets visés (tristesse/dépression ou automutilation/suicide) et si elles avaient été regardées par le compte manuel source. Les vidéos qui soit étaient pertinentes mais avaient été ignorées par erreur dans le cadre des tests manuels, soit avaient été regardées par erreur mais ensuite considérées comme hors sujet ont donc été exclues. Une vidéo est considérée comme « vue » si le temps de visionnage excède 1,75 fois sa durée.

L'axe des ordonnées représente le pourcentage de vidéos sur le sujet (tristes/dépressives dans le graphique 1, automutilation/idées suicidaires dans le graphique 2) figurant dans les données recueillies (les 402 recommandations récupérées en réponse au fil cloné).

Graphique 1



Dans les graphiques, lorsqu'un point figure au-dessus de la ligne entre les deux axes (la diagonale en pointillés), cela signifie qu'il y a eu une *amplification* des contenus (le pourcentage de vidéos nocives étant plus élevé dans les données recueillies que dans les données importées). Le graphique 1 montre donc une nette amplification des contenus tristes ou dépressifs dans les recommandations apparues sur les comptes dans le cadre du processus de personnalisation du fil « Pour toi ». Lorsque

23 Pour confirmer que TikTok avait bien enregistré les comportements de visionnage des comptes automatiques, l'équipe de recherche a exporté l'historique de visionnage de chaque compte. Elle a découvert que 10 à 15 % des vidéos qui auraient dû être considérées comme regardées par les comptes automatiques puisque figurant dans les historiques clonés n'étaient pas présentes dans les historiques exportés. Il est difficile de savoir d'où vient cette différence, mais l'équipe en a tenu compte en ne considérant une vidéo comme visionnée que si elle apparaissait dans les données téléchargées depuis le compte. Le même nombre de vidéos sur les sujets visés a été trouvé dans les données téléchargées pour chacun des deux comptes par partition. Au total, les chercheurs et chercheuses ont constaté que, au lieu des 150, 200, 250, 300 et 393 vidéos censées figurer dans les différentes partitions, les données importées contenaient 132, 181, 217, 266 et 355 vidéos.

le point figure en dessous de la diagonale, cela signifie qu'il y a eu une *minimisation* des contenus (le pourcentage de vidéos nocives étant moins élevé dans les données recueillies que dans les données importées). Il convient toutefois de noter que toute tendance à la hausse, même si les points tombent sous la diagonale, indique que plus l'utilisateur·rice regarde des contenus problématiques, plus TikTok lui recommande ce type de contenus.

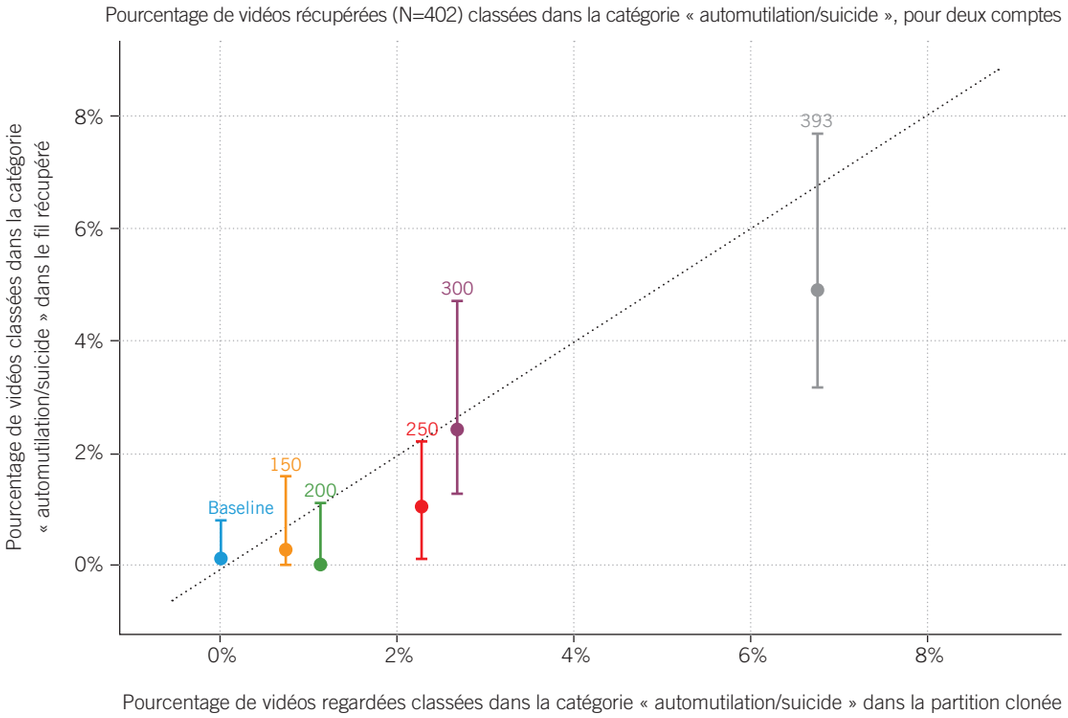
Les traits dans le prolongement des points représentent l'intervalle de confiance[24], qui est une représentation statistique du degré d'incertitude et de fiabilité de nos résultats.

Les chercheurs et chercheuses ont constaté que, plus la proportion de vidéos sur des sujets tristes ou dépressifs figurant dans le fil cloné était élevée, plus TikTok recommandait de vidéos portant sur la tristesse ou le sentiment de désespoir et donnait une réponse inappropriée. Les résultats indiquent que, avec un historique de visionnage contenant un taux de contenus tristes ou dépressifs aussi bas que 4,9 %, le pourcentage de vidéos de ce type recommandées à partir de ces signaux atteint presque 13 %, soit plus du double du pourcentage original de contenus potentiellement néfastes. À 7,6 % de vidéos tristes ou dépressives dans l'historique d'origine, le taux de tels contenus dans les recommandations passe à 16 %.

Cela signifie que, lorsqu'un·e mineur·e exprime un intérêt pour les contenus tristes ou dépressifs, le système de recommandation de TikTok continue de lui en proposer et de les multiplier dans le fil « Pour toi ».

Les pourcentages sont calculés selon une approche prudente des désaccords au sein de l'équipe sur le classement des vidéos recommandées dans les différentes catégories. En cas de désaccord sur le caractère triste ou dépressif d'une vidéo, celle-ci n'a pas été incluse dans la catégorie en question.

Graphique 2



24 Estimé au moyen d'un calcul statistique appelé méthode Agresti-Coull.

Il convient aussi de souligner que le comportement de ces comptes automatiques est nettement plus sobre que celui d'un·e véritable utilisateur·rice, qui manifesterait probablement aussi ses préférences en sauvegardant, republiant et mettant des mentions J'aime et des commentaires sur certaines vidéos, ainsi qu'en suivant des comptes. Pour des raisons éthiques et pratiques, la méthodologie choisie a été de manifester les préférences de contenus uniquement par le biais de l'historique de visionnage. Il faut aussi noter que ces conclusions sont tirées d'un échantillon de 12 comptes automatiques, ce qui, tout en permettant de généraliser davantage les résultats, reste un échantillon relativement petit, qui mériterait d'être complété par d'autres recherches comparables.

Les tests automatisés montrent que la personnalisation des contenus découlant du fait de regarder des vidéos sur l'automutilation et les idées suicidaires est moins forte qu'en ce qui concerne les vidéos sur la tristesse. Le fait de regarder ce type de contenus a néanmoins donné lieu à de nouvelles recommandations de vidéos sur l'automutilation et les idées suicidaires dans les fils.

Comparées aux contenus bénins (non problématiques), les vidéos sur l'automutilation ou le suicide ont statistiquement moins de mentions J'aime, de vues et de commentaires, et les vidéos sur la tristesse ont statistiquement moins de vues et de commentaires, les mentions J'aime n'étant pas statistiquement significatives. Cela pourrait indiquer que, du fait de la personnalisation, les contenus potentiellement nocifs n'avaient pas besoin d'être aussi populaires que les contenus bénins pour être recommandés à nos utilisateurs et utilisatrices tests.

Dans les expériences automatisées, le pourcentage le plus élevé observé de contenus potentiellement nocifs (21,4 % de contenus dépressifs ou portant sur l'automutilation/le suicide) est moins important que dans le cadre de l'effet de « spirale » constaté dans les comptes gérés manuellement. Cela s'explique en partie par la proportion plus faible de contenus potentiellement nocifs figurant dans les données importées, certaines vidéos du fil manuel d'origine n'étant plus disponibles (peut-être parce qu'elles avaient été supprimées par l'utilisateur·rice ou par TikTok). Il est possible également que le téléphone d'une véritable personne envoie plus d'informations à TikTok que la configuration automatique, ce qui donne des recommandations différentes. ***Néanmoins, le principe de base demeure : plus un·e utilisateur·rice manifeste un intérêt pour des contenus potentiellement nocifs, plus il/elle reçoit de recommandations en ce sens.***

Ces conclusions sont profondément préoccupantes car il a été démontré que même des taux plus faibles d'exposition persistante à des contenus extrêmes de ce type peuvent être source de préjudices dans le monde réel. Les individus sont plus ou moins sensibles aux contenus dépressifs ou portant sur l'automutilation ou le suicide, mais les témoignages de jeunes en chair et en os figurant dans ce rapport, ainsi que le cas très médiatisé de Molly Russell, une adolescente britannique qui s'est suicidée, donnent une idée de l'effet manipulateur d'une exposition continue à des contenus dépressifs ou liés à l'automutilation/au suicide sur des adolescent·e·s déjà confrontés à des problèmes psychologiques. Pour préciser le contexte, un *coroner* britannique a conclu que les réseaux sociaux avaient joué un rôle dans le suicide de Molly Russel, 14 ans. L'enquête a établi qu'« environ 13 % » des 16 000 publications que la jeune fille avait regardées sur Instagram dans les six mois précédant sa mort « étaient liées au suicide, à la dépression ou à l'automutilation[25] ».

25 "Digital bridge: Platforms on the hook — Transatlantic AI rulebook – Let's talk data transfers", Politico, 6 octobre 2022, https://www.politico.eu/newsletter/digital-bridge/platforms-on-the-hook-transatlantic-ai-rulebook-lets-talk-data-transfers/ ; une porte-parole de Meta qui s'est exprimée lors de l'enquête a déclaré que les contenus que Molly avait visionnés étaient « nuancés et complexes » et qu'il était important de permettre aux utilisateurs et utilisatrices ayant des pensées suicidaires de s'exprimer en ligne. "Molly Russell: Instagram posts seen by teen were safe, Meta says", BBC, 26 septembre 2022, https://www.bbc.co.uk/news/uk-england-london-63034300 ; pour consulter la réponse détaillée de Meta et les principales interrogations soulevées, voir Meta, "Response from Meta", 6 décembre 2022, https://www.judiciary.uk/wp-content/uploads/2022/10/2022-0315-Response-from-META.pdf.

6. LE CADRE DES DROITS HUMAINS

Le rapport *Poussé·e·s* vers les ténèbres publié en 2023 par Amnesty International a montré comment le modèle économique de TikTok, sa conception addictive et le manque de mesures prises par l'entreprise pour limiter le risque pour ses utilisateurs et utilisatrices de tomber dans une spirale nocive portaient atteinte aux droits humains et notamment aux droits au respect de la vie privée, à la liberté de pensée et à la santé. Nous vous invitons à consulter ce rapport pour une analyse plus détaillée des normes et instruments internationaux relatifs aux droits humains concernés. Voici cependant un résumé des responsabilités des entreprises en matière de droits fondamentaux et des principales obligations des États concernant le respect, la promotion et la mise en œuvre des droits de l'enfant et des droits humains :

Le **droit au respect de la vie privée** consacre le droit de chacun·e de ne pas faire l'objet « d'immixtions arbitraires ou illégales » dans sa vie privée, sa famille, son domicile ou sa correspondance, ce droit devant être protégé par la loi[26]. Il englobe trois notions qui sont étroitement liées : le droit de toute personne de ne pas subir d'immixtion dans sa vie privée, celui de contrôler les informations la concernant, et celui d'avoir un espace où exprimer librement son identité. Amnesty International estime que le profilage des utilisateurs et utilisatrices propre au modèle économique de TikTok entrave chacun de ces trois éléments. En effet, les plateformes de réseaux sociaux, dont TikTok, instaurent des modes de surveillance de masse pour personnaliser les recommandations de contenus et les publicités. Or, cette surveillance est par nature injustifiée et disproportionnée et ne peut en aucun cas constituer une immixtion autorisée dans la vie privée[27]. Les enfants et adolescent·e·s sont exposés à une pression sociale énorme pour participer à ces espaces. Ils partagent volontairement des informations, mais sont également incités à publier des renseignements personnels dans des publications et des commentaires, ainsi qu'à révéler leurs préférences et intérêts en attribuant des mentions « J'aime » ou en relayant les contenus publiés par d'autres personnes. Les entreprises de réseaux sociaux en apprennent encore bien plus grâce aux données sur les contenus que les enfants et les jeunes regardent, la fréquence et la durée de ces visionnages, la localisation des internautes, les heures auxquelles ils se connectent et les appareils qu'ils utilisent, pour ne citer que quelques exemples. Les enfants et adolescent·e·s accèdent donc à ces espaces au prix d'un profilage extrêmement invasif, auquel on ne peut pas réellement consentir à un si jeune âge, et qui ne peut pas être considéré comme étant pratiqué dans leur intérêt supérieur. Pourtant, les cadres juridiques de

26 Déclaration universelle des droits de l'homme (DUDH), article 12, et Pacte international relatif aux droits civils et politiques (PIDCP), article 17.

27 Amnesty International, *« Je me sens vulnérable ». Pris·e·s au piège de la surveillance intrinsèque à TikTok* (index : POL 40/7349/2023), 7 novembre 2023 (op. cit.) ; le DSA interdit la publicité ciblée à destination des personnes mineures. Cependant, la grande majorité des jeunes utilisateurs et utilisatrices de TikTok dans le monde peuvent, eux, toujours être exposés à des publicités ciblées. Cette interdiction a donc peu de répercussions sur le modèle économique de TikTok.

l'Union européenne, dont le Règlement sur les services numériques (DSA) et le Règlement général sur la protection des données (RGPD), n'ont jusqu'à présent pas permis d'enrayer ces atteintes aux droits humains, car les garanties contenues dans la législation sont insuffisantes et mal appliquées par l'UE et les États[28].

Le **droit à la liberté de pensée**, protégé par l'article 18 (1) du Pacte international relatif aux droits civils et politiques (PIDCP), est un droit absolu, ce qui signifie que personne, en aucun cas, ne peut s'immiscer dans nos pensées et croyances personnelles. Le Conseil de l'Europe a signalé que « les niveaux très subtils, subconscients et personnalisés de la persuasion algorithmique peuvent avoir des effets significatifs sur l'autonomie cognitive des citoyens et leur droit à se forger une opinion et à prendre des décisions indépendantes[29] ». Comme le montrent nos recherches et les témoignages de jeunes concernés, le système de recommandation de TikTok, basé sur la surveillance, et les caractéristiques de conception addictives de la plateforme risquent d'entraver la capacité des enfants et des jeunes à penser librement et exposent des adolescent·e·s déjà vulnérables à une « spirale » pouvant avoir de graves répercussions sur leur droit à la santé.

Les États ont l'obligation de prendre des mesures en vue de progressivement permettre la pleine réalisation du droit de toute personne au meilleur état de **santé physique et mentale** susceptible d'être atteint[30]. Les technologies numériques jouent un rôle indéniable dans la réduction des obstacles aux informations et services en matière de santé. Toutefois, il est de plus en plus admis que, en l'absence de réglementation étatique efficace, les réseaux sociaux exposent aussi les enfants et les jeunes à des risques systémiques relatifs à leurs droits fondamentaux[31]. L'administrateur de la santé publique des États-Unis a déclaré en mai 2023, au sujet des risques pour la santé des enfants et adolescent·e·s :

> *« Nous devons reconnaître le corpus croissant de recherches sur les potentiels préjudices, renforcer notre compréhension collective des risques liés à l'utilisation des réseaux sociaux et prendre de toute urgence des mesures pour limiter les préjudices et protéger la santé mentale et le bien-être des enfants et des adolescent·e·s pendant ces étapes de développement cruciales[32]. »*

Après avoir présenté les réseaux sociaux comme un déterminant de la santé, le précédent rapporteur spécial des Nations unies sur le droit à la santé physique et mentale avait formulé la recommandation suivante :

> *« Les États devraient s'acquitter de leurs obligations en adoptant des stratégies globales visant à renforcer la capacité des adolescents à se protéger des dangers en ligne et en améliorant la législation et les mécanismes d'application des lois afin de lutter contre la violence en ligne, y compris lorsqu'elle est transfrontalière, de combattre l'impunité et de former les parents et les professionnels qui travaillent avec les enfants[33]. »*

---

28 European Digital Rights (EDRi) et Amnesty International, *Civil society call and recommendations for concrete solutions to GDPR enforcement shortcomings*, 16 mars 2022, https://www.amnesty.eu/wp-content/uploads/2022/03/EDRi-recommendations-for-better-GDPR-enforcement.pdf.

29 Conseil de l'Europe, Déclaration du Comité des ministres sur les capacités de manipulation des processus algorithmiques, 13 février 2019, Decl (13/02/2019) 1, https://search.coe.int/cm/pages/result_details.aspx?ObjectId=090000168092dd4c.

30 DUDH, article 25 ; Pacte international relatif aux droits économiques, sociaux et culturels (PIDESC), article 12 ; Convention relative aux droits de l'enfant, article 24. Voir également les traités régionaux de protection des droits humains, notamment la Charte africaine des droits de l'homme et des peuples et la Charte sociale européenne.

31 Voir par exemple : Haut-Commissariat des Nations unies aux droits de l'homme (HCDH), *Innovation et technologies numériques et droit à la santé*, Rapport de la rapporteuse spéciale sur le droit qu'a toute personne de jouir du meilleur état de santé physique et mentale possible, avril 2023, https://documents-dds-ny.un.org/doc/UNDOC/GEN/G23/074/90/PDF/G2307490.pdf?OpenElement ; Comité des droits de l'enfant, Observation générale no 25 (2021) sur les droits de l'enfant en relation avec l'environnement numérique, 2 mars 2021, doc. ONU CRC/C/GC/25, https://docs.un.org/fr/CRC/C/GC/25.

32 Administrateur de la santé publique des États-Unis, *Social Media and Youth Mental Health*, mai 2023, hhs.gov/surgeongeneral/priorities/youth-mental-health/social-media/index.html.

33 HCDH, Rapport du rapporteur spécial sur le droit qu'a toute personne de jouir du meilleur état de santé physique et mentale susceptible d'être atteint, 4 avril 2016, https://docs.un.org/fr/A/HRC/32/32, § 45.

Il importe de noter que, à l'époque, devançant les débats actuels sur l'interdiction de l'utilisation des réseaux sociaux par les adolescent·e·s, le rapporteur spécial faisait déjà remarquer qu'« il n'est ni pertinent ni possible de chercher à limiter l'accès des adolescents au numérique[34] ».

Les entreprises ont l'obligation de respecter les droits humains, où qu'elles opèrent dans le monde et dans tous leurs domaines d'activité. Il s'agit d'une norme de conduite générale largement reconnue et inscrite dans les normes internationales relatives à la responsabilité des entreprises en matière de droits humains, notamment dans les Principes directeurs relatifs aux entreprises et aux droits de l'homme des Nations unies (Principes directeurs de l'ONU) et dans les Principes directeurs à l'intention des entreprises multinationales de l'Organisation de coopération et de développement économiques (Principes directeurs de l'OCDE)[35]. Les Principes directeurs de l'ONU prévoient que, pour y parvenir, les entreprises doivent mettre en œuvre de façon continue et volontariste une procédure de diligence nécessaire en matière de droits humains pour identifier, prévenir et atténuer les risques que posent leurs activités et rendre compte de la manière dont elles remédient à leur incidence potentielle et réelle. Si, au cours de la procédure de diligence requise, l'entreprise se rend compte qu'elle commet des atteintes aux droits humains ou y contribue, elle doit alors faire cesser ou prévenir ces atteintes et offrir réparation aux personnes touchées. En ce qui concerne les mineur·e·s, le Comité des droits de l'enfant a déclaré que les États devaient obliger les entreprises à prendre les précautions qui s'imposent pour garantir le respect des droits de l'enfant. Il a défini les mesures que les États doivent prendre pour empêcher les entreprises de causer des atteintes aux droits des mineur·e·s ou d'y contribuer et pour « enquêter et se prononcer sur les violations des droits de l'enfant qui ont été causées par une entreprise ou auxquelles une entreprise a contribué et […] offrir réparation[36] ».

34 Ibid.

35 Principes directeurs relatifs aux entreprises et aux droits de l'homme. Mise en œuvre du cadre de référence « protéger, respecter et réparer » des Nations unies, 2011, approuvés par la résolution 17/4 du Conseil des droits de l'homme des Nations unies : Les droits de l'homme et les sociétés transnationales et autres entreprises, adoptée le 16 juin 2011, doc. ONU A/HRC/RES/17/4, https://www.ohchr.org/sites/default/files/Documents/Publications/GuidingPrinciplesBusinessHR_FR.pdf ; OCDE, Principes directeurs pour les entreprises multinationales, 2011, https://www.oecd.org/fr/gouvernementdentreprise/mne/. Conformément aux Principes directeurs de l'ONU, la responsabilité de respecter les droits humains qui incombe aux entreprises est indépendante des obligations propres aux États en la matière et prévaut sur le respect des lois et règlements nationaux qui protègent les droits fondamentaux. Voir les Principes directeurs de l'ONU, principe 11 et commentaire associé.

36 Comité des droits de l'enfant, Observation générale no 16, 17 avril 2013, Doc. ONU CRC/C/GC/16, § 62 et 28, https://docs.un.org/fr/CRC/C/GC/16.

7. LES CADRES JURIDIQUES EUROPÉEN ET NATIONAL

7.1 RÉGLEMENTATION DE L'UNION EUROPÉENNE : LE RÈGLEMENT SUR LES SERVICES NUMÉRIQUES

En juillet 2022, l'Union européenne a adopté un texte historique : le Règlement sur les services numériques (DSA), qui définit les responsabilités et obligations des entreprises de réseaux sociaux et autres plateformes en ligne présentes dans la région[37]. Le DSA a été créé en vue d'harmoniser les règles propres à chacun des 27 États membres de l'UE afin de garantir un environnement en ligne sûr où les droits fondamentaux, dont les droits de l'enfant, soient réellement protégés. Les États membres appliquent et font appliquer les dispositions du DSA dans le cadre de leur propre système juridique, par l'intermédiaire d'autorités nationales désignées à cet effet (les coordinateurs pour les services numériques), tandis que la Commission européenne est chargée de faire respecter les règles les plus strictes, qui s'appliquent aux très grandes plateformes en ligne. Le DSA lutte contre la diffusion de contenus illégaux et la désinformation et contre les risques systémiques découlant « de la conception ou du fonctionnement de leurs services et de leurs systèmes connexes, y compris des systèmes algorithmiques, ou de l'utilisation faite de leurs services[38] ». En tant que cadre juridique contraignant, le DSA prévoit des sanctions en cas de non-respect de ses dispositions, notamment des amendes d'un montant maximum équivalant à 6 % du chiffre d'affaires mondial annuel de l'entreprise concernée.

Le DSA impose aux très grandes plateformes en ligne (plateformes définies comme comptant plus de 45 millions d'utilisateurs et utilisatrices dans l'UE), dont TikTok fait partie, de mettre en place « des mesures d'atténuation raisonnables, proportionnées et efficaces », adaptées aux risques systémiques spécifiques identifiés au moyen d'une évaluation des risques annuelle, en tenant compte de l'incidence de ces mesures sur les droits fondamentaux[39]. Ces mesures peuvent inclure l'adaptation de la conception, des caractéristiques ou du fonctionnement de leurs services, l'adaptation des processus de modération des contenus et le test et l'adaptation de leurs systèmes algorithmiques de recommandation. Le DSA interdit par ailleurs les publicités ciblées à destination des mineur·e·s et exige des grandes plateformes de réseaux sociaux qu'elles offrent en option à l'ensemble des

37 Union européenne, Règlement (UE) 2022/2065 du Parlement européen et du Conseil du 19 octobre 2022 relatif à un marché unique des services numériques et modifiant la directive 2000/31/CE (Règlement sur les services numériques), https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:32022R2065 ; Amnesty International, *Adoption de la loi de l'Union européenne sur les services numériques : quelles implications pour les droits humains et pour les modèles économiques aux effets nocifs des géants du numérique* (index : POL 30/5830/2022), 7 juillet 2022, https://www.amnesty.org/fr/documents/pol30/5830/2022/fr.

38 Règlement sur les services numériques, op. cit., art. 34

39 Ibid., art. 35.

utilisateurs et utilisatrices un fil d'actualités qui ne soit pas basé sur le profilage[40]. Les très grandes plateformes en ligne sont soumises à un audit indépendant annuel visant à vérifier si elles respectent leur devoir de vigilance.

Les obligations découlant du DSA s'appliquent depuis août 2023 aux très grandes plateformes en ligne et depuis février 2024 à toutes les plateformes numériques[41].

Le 19 février 2024, la Commission européenne a annoncé qu'elle avait ouvert une procédure formelle « afin de déterminer si TikTok a enfreint le règlement sur les services numériques dans des domaines liés à la protection des mineurs, à la transparence de la publicité, à l'accès des chercheurs aux données, ainsi qu'à la gestion des risques liés à la conception addictive et aux contenus préjudiciables[42]. » Faisant écho aux conclusions de recherche publiées par Amnesty International en novembre 2023, la Commission a annoncé qu'elle s'intéressait aux éventuelles infractions aux obligations qu'impose le DSA « en ce qui concerne les **effets négatifs** réels ou prévisibles découlant de la conception du système de TikTok, y compris ses systèmes algorithmiques, **susceptibles de stimuler les dépendances comportementales et/ou de créer des effets de "spirales infernales"** » (en gras dans le texte). Au moment de la rédaction du présent rapport, cette enquête était toujours en cours.

Une deuxième enquête, ouverte en avril 2024, a poussé TikTok à s'engager à retirer définitivement de l'UE le programme TikTok Lite Rewards[43]. TikTok Lite est une version de TikTok économe en données mobiles déjà couramment utilisée au Brésil, en Inde et en Indonésie et qui a été lancée discrètement en France et en Espagne en avril 2024. Cette version offre aux utilisateurs et utilisatrices des récompenses financières pour les inciter à utiliser davantage l'application. La Commission européenne avait mis en doute le respect par TikTok de ses obligations en matière d'évaluation et d'atténuation des risques concernant des effets potentiellement nuisibles à la santé mentale (des mineur·e·s) découlant de « la nouvelle fonctionnalité qui stimule un comportement de dépendance[44] ».

En juillet 2025, la Commission européenne a publié des Lignes directrices sur la protection des mineurs adressant des recommandations aux plateformes en ligne sur la façon de limiter les risques pour les droits de l'enfant conformément à l'article 28 du DSA, qui impose aux fournisseurs de plateformes en ligne de mettre en place « des mesures appropriées et proportionnées pour garantir un niveau élevé de protection de la vie privée, de sûreté et de sécurité des mineurs sur leur service[45]. » En ce qui concerne plus particulièrement la question du modèle addictif, ces Lignes directrices mettent l'accent sur la responsabilité qu'ont les plateformes de veiller à ce que :

> *« les mineurs ne soient pas exposés à des caractéristiques de conception persuasives qui visent principalement à les faire interagir et qui sont susceptibles de conduire à une utilisation intensive ou excessive de la plateforme ou à des schémas comportementaux problématiques ou compulsifs ». Ces caractéristiques peuvent être notamment « la possibilité de faire défiler l'écran indéfiniment, l'obligation superflue d'effectuer une action spécifique afin de recevoir des*

40 Ibid., art. 28.

41 Commission européenne, « La législation sur les services numériques », https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_fr (consulté le 8 août 2025).

42 Commission européenne, « La Commission ouvre une procédure formelle à l'encontre de TikTok au titre du règlement sur les services numériques », 19 février 2024, https://ec.europa.eu/commission/presscorner/detail/fr/ip_24_926.

43 Commission européenne, « TikTok s'engage à retirer définitivement son programme TikTok Lite Rewards dans l'UE afin de se conformer au règlement sur les services numériques », 5 août 2024, https://ec.europa.eu/commission/presscorner/detail/fr/ip_24_4161.

44 Commission européenne, Lignes directrices concernant des mesures visant à garantir un niveau élevé de protection de la vie privée, de sûreté et de sécurité des mineurs en ligne, conformément à l'article 28, paragraphe 4, du règlement (UE) 2022/2065, 10 octobre 2025, https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=OJ:C_202505519 ; DSA, op. cit., art. 28.

45 Commission européenne, Lignes directrices concernant des mesures visant à garantir un niveau élevé de protection de la vie privée, de sûreté et de sécurité des mineurs en ligne, 14 juillet 2025, https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-protection-minors; DSA, art. 28

informations actualisées sur une application, le déclenchement automatique de contenus vidéo, les notifications programmées artificiellement en vue de regagner l'attention des mineurs[46] ».

Cette recommandation est bien évidemment incompatible avec certaines des fonctionnalités principales du fil « Pour toi » de TikTok.

Concernant les systèmes de recommandation, les Lignes directrices préconisent aux plateformes :

d'« accorder la priorité aux "signaux explicites de l'utilisateur" afin de déterminer le contenu qui s'affiche pour les mineurs et qui leur est recommandé » et de « mettre en œuvre des mesures visant à prévenir l'exposition des mineurs aux recommandations de contenu susceptibles de présenter un risque pour leur sûreté et leur sécurité, en particulier lorsqu'elles leur sont présentées de manière répétée, telles que des contenus promouvant des normes de beauté ou des régimes irréalistes, des contenus qui glorifient ou banalisent des problèmes de santé mentale, tels que l'anxiété ou la dépression, les contenus discriminatoires, les contenus liés à la radicalisation et les contenus choquants contenant de la violence ou encourageant les mineurs à se livrer à des activités dangereuses[47]. »

Les Lignes directrices de la Commission européenne reconnaissent donc les risques associés aux systèmes de recommandation basés sur l'engagement et cherchent à atténuer leurs effets nocifs, mais elles n'interdisent pas totalement ces systèmes intrusifs pour la vie privée. Bien que ces recommandations n'aient pas force obligatoire, elles indiquent comment la Commission définit la notion de « sécurité par la conception » en ce qui concerne les droits et les besoins des personnes mineures et montrent ce vers quoi tend l'application du DSA à cet égard.

Dans un contexte où les pays membres de l'UE, y compris le président français, se montrent de plus en plus favorables à la limitation de l'accès des adolescent·e·s aux réseaux sociaux, les Lignes de conduite de la Commission européenne semblent ouvrir la porte à la possibilité de relever l'âge minimum de cet accès dans les États membres, même si le texte précise qu'une telle décision doit se faire « conformément au droit de l'Union » (argument qui a été invoqué par le passé pour empêcher ce genre d'initiatives par des États membres, voir chapitre 7.2). Cette dernière disposition semble contredire le DSA, qui vise officiellement à harmoniser les règles des pays de l'UE, et elle risque de restreindre indûment le droit des adolescent·e·s à la liberté d'expression et de réunion pacifique en ligne[48].

7.2 LA PROTECTION DES MINEUR·E·S SUR INTERNET EN FRANCE

7.2.1 LA MISE EN ŒUVRE DU RÈGLEMENT EUROPÉEN SUR LES SERVICES NUMÉRIQUES EN FRANCE

La loi visant à sécuriser et réguler l'espace numérique (SREN)[49], promulguée le 21 mai 2024, adapte en droit français le Règlement européen sur les services numériques (DSA). Elle confie à l'Autorité

46 Ibid., article 6.4.

47 Idib., article 6.5.1.

48 Pour une analyse plus approfondie des risques que pose pour les droits humains l'interdiction des réseaux sociaux aux adolescent·e·s, voir : Amnesty International, « Australie. Les autorités doivent réglementer efficacement les réseaux sociaux plutôt que de les interdire aux enfants et aux jeunes », 21 novembre 2024, https://www.amnesty.org/fr/latest/news/2024/11/https-www-amnesty-org-en-latest-news-2024-11-australia-must-effectively-reglate-social-media-than-ban-children/.

49 Loi no 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique, https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000049563368/.

de régulation de la communication audiovisuelle et numérique (Arcom) le rôle de coordinateur pour les services numériques (CSN). L'Arcom exerce ces missions en lien avec la Commission nationale de l'informatique et des libertés (CNIL) en ce qui concerne les dispositions protégeant la vie privée, et avec la Direction générale de la concurrence, de la consommation et de la répression des fraudes (DGCCRF), en ce qui concerne les dispositions spécifiques aux places de marchés[50]. L'Arcom participe également, avec ses homologues européens, à la supervision des obligations des acteurs régulés par la Commission européenne. Celle-ci dispose de pouvoirs exclusifs pour surveiller et faire respecter les dispositions spécifiques aux très grandes plateformes en ligne.

Le DSA prévoit également l'existence, dans chaque État membre de l'UE, de « signaleurs de confiance », chargés de signaler aux plateformes en ligne les contenus présumés illégaux[51]. Leurs signalements sont traités de manière prioritaire par toutes les plateformes en ligne soumises au DSA. Les signaleurs de confiance, désignés par l'Arcom, sont « des organisations reconnues pour leur expertise dans la détection, l'identification et la notification de contenus illicites, dans un ou plusieurs domaines déterminés[52] ». Parmi eux figure notamment l'association e-Enfance, qui lutte contre le harcèlement et les violences numériques que subissent les jeunes[53].

Près d'un an avant l'entrée en vigueur de la loi SREN, la Loi du 7 juillet 2023 visant à instaurer une majorité numérique et à lutter contre la haine en ligne[54] a établi une majorité numérique à 15 ans pour s'inscrire sur les réseaux sociaux. Cependant, la Commission européenne ayant jugé que les dispositions de cette loi n'étaient pas conformes au droit européen, le décret d'application de la loi n'a pas été publié et la mesure n'a pas été appliquée.

7.2.2 LES INITIATIVES DES AUTORITÉS FRANÇAISES

En janvier 2024, le président Emmanuel Macron a lancé « une Commission constituée d'experts issus de la "société civile" pour évaluer les enjeux attachés à l'exposition des enfants aux écrans et formuler des recommandations[55] ». Cette Commission s'est notamment penchée sur les problématiques d'addiction aux écrans et en particulier aux réseaux sociaux, leurs effets sur la santé mentale des enfants et des adolescent·e·s, ainsi que le manque de régulation des contenus auxquels les mineur·e·s peuvent être exposés. Dans son rapport rendu public en avril 2024, la Commission recommande en particulier de « s'attaquer, pour les interdire, aux conceptions addictogènes et enfermantes de certains services numériques afin de redonner du choix aux jeunes[56] ».

En mars 2025, une commission d'enquête parlementaire sur les effets psychologiques de TikTok sur les mineur·e·s a été créée[57]. Elle a notamment examiné les dispositifs de captation de l'attention utilisés par TikTok ainsi que les risques liés à l'exposition des jeunes utilisateurs·rices aux contenus dangereux. Plusieurs représentant·e·s de TikTok ont été auditionnés et questionnés sur la responsabilité de

50 Convention de coopération pour la mise en œuvre du règlement sur les services numériques entre l'Arcom, la DGCCRF et la CNIL : https://www.arcom.fr/sites/default/files/2024-06/Arcom%20-%20Convention%20coop%C3%A9ration_ARCOM_CNIL_DGCCRF.pdf

51 DSA, op. cit., art. 22.

52 Arcom, « Règlement sur les services numériques (DSA) : liste des signaleurs de confiance désignés par l'Arcom », https://www.arcom.fr/signaleurs-de-confiance.

53 https://e-enfance.org/. Voir la liste des signaleurs de confiance régulièrement mise à jour : https://www.arcom.fr/signaleurs-de-confiance.

54 Loi no 2023-566 du 7 juillet 2023 visant à instaurer une majorité numérique et à lutter contre la haine en ligne, https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000047799533.

55 Servane Mouton et. al., *Enfants et écrans. À la recherche du temps perdu*, avril 2024, https://www.elysee.fr/admin/upload/default/0001/16/fbec6abe9d9cc1bff3043d87b9f7951e62779b09.pdf.

56 Ibid.

57 Commission d'enquête sur les effets psychologiques de TikTok sur les mineurs, https://www.assemblee-nationale.fr/dyn/17/organes/autres-commissions/commissions-enquete/tiktok.

la plateforme s'agissant en particulier des « spirales » de contenus potentiellement dangereux, notamment les vidéos qui idéalisent et encouragent les pensées dépressives, l'automutilation et le suicide. Pour les représentant·e·s de TikTok, l'entreprise met tout en œuvre pour protéger ses jeunes utilisateurs et utilisatrices et leur garantir un espace « sûr » en ligne : « Nous nous engageons à agir en responsabilité […] en soutenant les contenus utiles et positifs et en veillant à la sécurité et au bien-être de notre communauté[58] ». La commission a présenté ses conclusions en septembre 2025. Arthur Delaporte, président de la commission d'enquête, a déclaré dans son avant-propos : « le verdict est sans appel : cette plateforme expose en toute connaissance de cause nos enfants, nos jeunes, à des contenus toxiques, dangereux, addictifs[59] ».

En parallèle des travaux de la commission d'enquête parlementaire, la ministre déléguée chargée de l'intelligence artificielle et du numérique s'est fortement mobilisée au niveau européen pour faire interdire l'accès aux réseaux sociaux pour les mineur·e·s de moins de quinze ans. Le président de la République s'est également exprimé en ce sens, réaffirmant la volonté de la France de relever la majorité numérique à 15 ans pour les réseaux sociaux comme prévu par la loi de juillet 2023.

Enfin, la santé mentale a été désignée Grande Cause nationale en 2025[60]. À ce titre, quatre objectifs prioritaires ont été fixés : la déstigmatisation, le développement de la prévention et du repérage précoce, l'amélioration de l'accès aux soins et l'accompagnement des personnes concernées.

58 Marlène Masure, responsable du contenu États d'Europe, du Moyen-Orient et de l'Afrique, directrice exécutive du développement commercial et marketing. Audition du 12 juin 2025 devant la Commission d'enquête parlementaire sur les effets psychologiques de TikTok sur les mineurs, https://www.assemblee-nationale.fr/dyn/actualites-accueil-hub/effets-psychologiques-de-tiktok-sur-les-mineurs-auditions-de-representants-de-tiktok-et-tiktok-france.

59 « TikTok est "un des pires réseaux sociaux à l'assaut de notre jeunesse" : les mots très durs de la commission d'enquête parlementaire », *Le Monde*, 11 septembre 2025, https://www.lemonde.fr/pixels/article/2025/09/11/commission-d-enquete-tiktok-le-rapport-etrille-une-plateforme-hors-de-controle-a-l-assaut-de-la-jeunesse-et-formule-43-recommandations_6640383_4408996.html.

60 Ministère français de la Santé, « La santé mentale, Grande Cause nationale 2025 », 24 mars 2025, https://solidarites.gouv.fr/la-sante-mentale-grande-cause-nationale-2025.

8. L'INCAPACITÉ DE TIKTOK À RENDRE SA PLATEFORME SÛRE POUR LES ENFANTS ET LES ADOLESCENT·E·S : UNE VIOLATION DU RÈGLEMENT SUR LES SERVICES NUMÉRIQUES ET DES NORMES INTERNATIONALES RELATIVES AUX DROITS HUMAINS

Le rapport *Poussé·e·s vers les ténèbres*, publié en 2023 par Amnesty International, a montré que, en ne remplissant pas correctement son devoir de vigilance en matière de droits fondamentaux et en ne limitant pas les risques graves que présente sa plateforme pour la santé physique et mentale des jeunes utilisateurs et utilisatrices souffrant déjà de problèmes psychologiques, TikTok ne s'était pas acquittée de sa responsabilité de respecter les droits humains conformément aux Principes directeurs des Nations unies relatifs aux entreprises et aux droits de l'homme.

Les conclusions de ce nouveau rapport, qui confirme que TikTok ne respecte pas les droits des enfants et des adolescent·e·s et ne s'acquitte pas des responsabilités qui lui incombent au titre des normes internationales relatives aux droits humains et du DSA, doivent être examinées dès maintenant dans le cadre de l'enquête de la Commission européenne sur les infractions que pourrait avoir commises l'entreprise au regard du DSA en tant que très grande plateforme en ligne concernant l'évaluation et l'atténuation des risques systémiques pour les droits des mineur·e·s [articles 28 (1), 34 (1), 34 (2) et 35 (1)[61]].

61 Règlement sur les services numériques, op. cit.

À l'approche de la Semaine mondiale d'information sur la santé mentale, en mai 2025, Amnesty International a demandé à TikTok quels changements elle avait adoptés depuis la publication du rapport *Poussé·e·s vers les ténèbres*, près de 18 mois plus tôt. En réponse, TikTok s'est contentée d'énumérer des mesures de « bien-être » classiques, dont la plupart avaient déjà été mises en œuvre avant 2023, et n'a pas reconnu le problème de « spirale » que posait l'application[62]. L'entreprise n'a pas non plus fourni d'éléments permettant de penser qu'elle avait adopté de nouvelles mesures ciblées pour résoudre ce problème.

Dans l'évaluation des risques effectuée par TikTok en 2024 au titre du DSA[63], l'entreprise a reconnu que « certains types de contenus concentrés peuvent, même s'ils n'enfreignent pas la charte de TikTok, causer des préjudices en renforçant involontairement chez certains internautes une expérience personnelle négative. Par exemple, des concentrations de contenus relatifs à l'image corporelle ou à des régimes particulièrement drastiques peuvent avoir des conséquences sur la santé mentale des utilisateurs et utilisatrices, notamment des plus jeunes. » Dans le même chapitre de son évaluation des risques, TikTok a énuméré des mesures d'atténuation, comme l'application préventive de sa charte, le maintien de critères à remplir pour qu'un contenu soit recommandé et l'application au fil « Pour toi » de techniques de diversification, ainsi que des outils à destination des utilisateurs et utilisatrices, comme les filtres et la fonction « rafraîchir » pour renouveler le fil. Toutes ces mesures existaient déjà en 2023, lorsqu'Amnesty International avait mené son travail de recherche et démontré que, malgré ces mesures, les jeunes utilisateurs et utilisatrices de la plateforme étaient exposés à des risques systémiques en matière de droits fondamentaux. L'évaluation des risques réalisée par TikTok ne présentait aucun indicateur pour mesurer l'ampleur de ces risques ni aucun point de référence pour définir ce que l'entreprise considère comme un niveau acceptable d'exposition à des « concentrations de contenus » (terme employé par TikTok).

En mai 2025, TikTok a intégré à son application des exercices de méditation « pour encourager les jeunes à se déconnecter la nuit[64] ». En juillet 2025, l'entreprise a lancé des « missions bien-être », promouvant de façon ludique les outils de bien-être de la plateforme[65]. Ces nouveautés montrent que TikTok a conscience de certains des risques auxquels elle expose les enfants et adolescent·e·s, mais ces mesures d'atténuation ne sont pas adaptées ni suffisantes face aux fonctionnalités addictives de l'application[66], auxquelles l'entreprise n'a pas touché. Si les lignes directrices de l'Union européenne sur la protection des mineurs n'ont pas force obligatoire, elles offrent aux plateformes des indications claires sur la manière d'interpréter les obligations qui leur incombent au titre de l'article 28 du DSA. Les fonctionnalités de base de TikTok, appliquées par défaut pour les mineur·e·s (déclenchement automatique des vidéos, possibilité de faire défiler des contenus à l'infini, recommandations de contenus fondées sur l'engagement sur le fil « Pour toi ») vont de toute évidence à l'encontre de ces indications.

Le travail de recherche technique mené pour ce rapport montre que les mesures que TikTok déclare avoir prises au titre du DSA pour limiter le risque d'exposer les jeunes utilisateurs et utilisatrices à des effets de « spirale » ou à des « concentrations de contenus » pouvant nuire à leur santé mentale et physique ne sont ni efficaces ni adaptées. Amnesty International n'a trouvé aucun élément qui permettrait de croire que les ensembles nocifs de « contenus concentrés » aient été désagrégés.

---

62 Amnesty International, *TikTok fails to address risks to children and young people's mental health despite past warnings*, op. cit.

63 TikTok, *DSA Risk Assessment Report 2024*, 28 août 2024, https://www.tiktok.com/transparency/en/dsa-transparency/.

64 TikTok, "Expanding TikTok's Mental Health Education Fund and well-being features", 15 mai 2025, https://newsroom.tiktok.com/en-us/tiktok-expanding-mental-health-education-fund-and-wellbeing-features.

65 TikTok, "New ways we're helping our community build balanced digital habits", 30 juillet 2025, https://newsroom.tiktok.com/en-us/new-ways-to-help-our-community-build-balanced-digital-habits.

66 Amnesty International, *Poussé·e·s vers les ténèbres. Comment le fil « Pour toi » de TikTok encourage l'automutilation et les idées suicidaires* (index : POL 40/7350/2023), op. cit.

Au contraire, comme en 2023, l'équipe de recherche a vérifié que l'effet de « spirale » commençait à peine quelques minutes après le début des expériences manuelles et se maintenait pendant les heures suivantes avec une multiplication des vidéos idéalisant ou normalisant le suicide.

Les témoignages de jeunes ayant pratiqué l'automutilation et le travail de recherche technique ont également confirmé que l'application de la charte que TikTok assure faire respecter de manière préventive ne protège pas les mineur·e·s d'une exposition à des contenus nocifs ou contraires à la loi. Les contenus visionnés contenaient parfois des trompe-algorithmes (stratégies employées pour contourner les systèmes de modération des contenus grâce à des émoticônes, des abréviations ou le remplacement de certaines lettres par des chiffres) faisant référence au suicide ou à l'automutilation, dont le sens était évident pour l'équipe de recherche. D'autres présentaient des défis liés au suicide, bien connus du public (voir le chapitre 4.1). Si TikTok avait exercé correctement son devoir de vigilance et s'était attaquée aux risques posés par sa plateforme en matière de droits humains, elle aurait eu connaissance de ces contenus, aurait pu les repérer et aurait identifié les risques qu'ils posaient. Les témoignages montrent également que TikTok n'a pas modéré correctement les contenus que les participant·e·s à l'étude lui ont signalés parce qu'ils/elles avaient des raisons de croire qu'ils étaient contraires à la charte de l'entreprise.

Il ne faut pas oublier que TikTok s'en remet de plus en plus à l'intelligence artificielle pour réaliser la tâche cruciale de décider quels contenus autoriser. L'entreprise a licencié des centaines de modérateurs et modératrices de contenus dans le monde en 2024 et 2025 et aurait décidé de supprimer progressivement la totalité de son service Confiance et Sécurité situé en Allemagne[67]. Contacté par Associated Press en janvier 2025 au sujet de Marie, la fille de Stéphanie Mistre qui s'est suicidée, TikTok a invoqué ses « 40 000 employés travaillant partout dans le monde à la confiance et la sécurité » pour faire valoir qu'elle veillait à supprimer rapidement les contenus incitant au suicide. L'entreprise avait pourtant déjà commencé à réduire son équipe de confiance et sécurité[68]. Une étude récente comparant les performances des principaux grands modèles de langage multimodaux et ceux des modérateurs·rices de contenu humains a conclu que « les contrôleurs humains restent plus fiables, en particulier en ce qui concerne les classifications complexes ou nuancées ». La différence est encore plus marquée pour les contenus dans d'autres langues que l'anglais[69]. Étant donnée la complexité de la modération des contenus relatifs aux idées suicidaires, qui utilisent souvent de l'argot et des trompe-algorithmes dans des langues autres que l'anglais, la différence de qualité entre la modération faite par des humains et celle faite par l'intelligence artificielle n'est sans doute pas négligeable. Cependant, l'étude a également découvert que les modérateurs et modératrices de contenu coûtent 40 fois plus cher que l'intelligence artificielle. Cette différence incite donc clairement les entreprises à privilégier cette technologie plus économique, au détriment de la sécurité de leur plateforme. Les équipes de confiance et de sécurité considèrent l'intelligence artificielle comme une aide qui leur permet d'éviter d'avoir à visionner tous les contenus explicitement nocifs[70], mais remplacer toutes ces équipes par des systèmes d'intelligence artificielle constitue un pas inquiétant dans la mauvaise direction. En effet, les défauts de la modération de contenus automatisée ont été largement démontrés.

Par ailleurs, bien que cela ne fasse pas partie de l'objet de cette étude, les témoignages concernant la grande quantité de commentaires nocifs et la formation de communautés en ligne se livrant

67 "ByteDance's TikTok cuts hundreds of jobs in shift towards AI content moderation", Reuters, 11 octobre 2024, https://ca.finance.yahoo.com/news/bytedance-cuts-over-700-jobs-022649989.html ; "TikTok content moderators in Germany strike over AI taking their jobs", Euronews, 23 juillet 2025, https://uk.news.yahoo.com/tiktok-content-moderators-germany-strike-111824712.html.

68 "Families sue TikTok in France over teen suicides they say are linked to harmful content", AP, 25 janvier 2025, https://apnews.com/article/tiktok-france-trial-suicide-lawsuit-fa8f979c3121a3c5712d52a300c9005f.

69 "Humans make better content cops than AI, but cost 40x more", The Register, 8 août 2025, https://www.msn.com/en-us/health/medical/humans-make-better-content-cops-than-ai-but-cost-40x-more/ar-AA1KbhjQ.

70 "Where are tomorrow's T&S experts coming from?", Everything in Moderation, 18 août 2025, https://www.everythinginmoderation.co/future-trust-safety-expert/.

collectivement à la normalisation de l'automutilation ou du suicide indiquent qu'il existe également de possibles risques de conception liés à la multiplication fondée sur l'engagement de *commentaires* nocifs (par opposition à la multiplication des vidéos, qui faisait l'objet de cette étude). Cette piste mériterait d'être creusée.

À la lumière de cette étude, Amnesty International demande à la Commission européenne et aux autorités de régulation nationales d'analyser de près les risques systémiques et les atteintes aux droits des mineur·e·s constatés et de poursuivre sans délai l'enquête de la Commission sur les éventuelles infractions au DSA.

9. CONCLUSION ET RECOMMANDATIONS

Près de deux ans après la publication par Amnesty International de *Poussé·e·s vers les ténèbres*, un rapport accablant sur les « spirales » nocives créées par TikTok et leurs risques pour la santé mentale des enfants et des jeunes, ce nouveau rapport montre que TikTok n'a toujours pas réglé les risques systémiques associés à son modèle économique, qui cherche à accroître l'engagement sur la plateforme. En 2024, obligée à plus de transparence par le Règlement européen sur les services numériques, l'entreprise a enfin reconnu l'existence de risques systémiques liés à la « concentration de contenus » (euphémisme utilisé par TikTok pour désigner ces « spirales ») pour les enfants et adolescent·e·s. Cette nouvelle étude montre cependant que les mesures d'atténuation que TikTok affirme avoir mises en place n'empêchent pas ces « spirales » d'attirer les mineur·e·s dans un univers en ligne où l'automutilation et le désespoir sont la norme.

Notre enquête révèle que de nombreux enfants et adolescent·e·s ont conscience de certains risques liés au modèle économique des principaux réseaux sociaux, basé sur le profilage et l'engagement, mais qu'ils n'arrivent pas à s'en protéger, même quand ils essaient. Les entretiens menés par Amnesty International auprès de jeunes ayant pratiqué l'automutilation et de parents concernés ou endeuillés ont mis en lumière le risque que représente le fait d'exposer des adolescent·e·s déjà mal dans leur peau à une plateforme conçue pour maximiser l'engagement, quel qu'en soit le prix. Leurs témoignages révèlent comment, en les attirant dans un tunnel artificiel de mal-être psychologique, TikTok a normalisé et exacerbé leur penchant pour l'automutilation et leurs idées suicidaires au point de leur recommander des méthodes de suicide et des défis en ce sens. Les parents de Marie et ceux de Charlize, deux adolescentes qui se sont suicidées à l'âge de 15 ans, pensent que le fil « Pour toi » de TikTok a poussé leurs filles à bout, aggravant leurs problèmes psychologiques jusqu'à ce qu'elles commettent l'irréparable. Ils sentent désormais qu'il est de leur devoir d'avertir les autres parents et les jeunes des risques que représentent ces réseaux sociaux et de demander que les choses changent rapidement.

Les nouvelles expériences manuelles et automatisées menées par Amnesty International ont confirmé que le fil « Pour toi » de TikTok pouvait attirer les enfants et adolescent·e·s dans une « spirale » de contenus liés à la dépression et au suicide avec la même facilité aujourd'hui qu'en 2023. Quelques heures sur la plateforme suffisent pour qu'un enfant qui regarde les contenus liés à la santé mentale qui lui sont recommandés soit exposé à un flux presque ininterrompu de vidéos potentiellement néfastes. Ces vidéos peuvent contenir des références à des représentations connues de l'automutilation et du suicide, ce qui montre que l'atténuation des risques que TikTok affirme mettre en œuvre est clairement un échec. TikTok manque donc à sa responsabilité de respecter les droits humains énoncée dans les principes directeurs des Nations unies et de l'OCDE, et elle n'honore pas ses obligations au titre du DSA.

Face à la tendance du secteur à réduire les coûts des services de confiance et de sécurité en ayant davantage recours à l'intelligence artificielle, les autorités de régulation doivent intensifier de toute urgence les mesures d'application des lois pour éviter d'exposer encore les enfants et les jeunes à ce genre de risques et de préjudices, en France et ailleurs. Les éléments exposés dans ce rapport montrent clairement que TikTok a enfreint le Règlement européen sur les services numériques. La Commission européenne doit les intégrer sans délai à son enquête à ce sujet afin d'obliger rapidement TikTok à rendre son application sûre pour les enfants et les jeunes en Europe et partout dans le monde.

RECOMMANDATIONS

RECOMMANDATIONS À LA COMMISSION EUROPÉENNE

- Veiller à ce que le Règlement sur les services numériques soit rigoureusement appliqué. La Commission européenne doit inclure sans délai les présentes conclusions dans son enquête en cours sur les possibles violations par TikTok de ses obligations au titre du DSA et veiller à ce que cette enquête et la mise en œuvre de ses recommandations soient promptes et efficaces.
- Instaurer des lignes de conduite pour les très grandes plateformes en ligne qui évaluent les risques systémiques de leurs produits et services, afin que les informations figurant dans ces évaluations soient appropriées et précises et puissent être utilisées par la société civile et les autorités de régulation pour surveiller les plateformes.
- Veiller à ce que les sanctions appliquées en cas de non-respect délibéré de la législation soient suffisantes pour inciter à ne pas récidiver.

RECOMMANDATIONS AU GOUVERNEMENT FRANÇAIS ET AUX ÉTATS MEMBRES DE L'UE

- Faire en sorte que les droits des mineur·e·s soient respectés et protégés en travaillant avec les institutions et les États membres de l'UE pour créer et appliquer des cadres réglementaires communs, notamment le DSA, le Règlement sur l'intelligence artificielle, le RGPD et les propositions à venir pour lutter contre les caractéristiques de conception addictives dans le cadre de la législation sur l'équité numérique. Veiller notamment à ce que la sécurité des internautes soit prise en compte dans la conception des applications plutôt que de restreindre l'accès des adolescent·e·s aux réseaux sociaux, ce qui limiterait injustement leur droit à la liberté d'expression et de réunion pacifique en ligne.
- Faire le nécessaire pour que l'accès aux infrastructures et services numériques incontournables (comme TikTok et d'autres plateformes de réseaux sociaux) et leur utilisation ne soient pas subordonnés à une surveillance généralisée des utilisateurs et utilisatrices, qu'il s'agisse d'enfants, de jeunes ou d'adultes. Empêcher les entreprises, en premier lieu, de conditionner l'accès à leurs services au fait d'« accepter » la collecte, le traitement ou le partage de données personnelles à des fins de ciblage de contenu et de marketing ou de publicité.
- Réglementer les entreprises de réseaux sociaux de manière que les algorithmes utilisés par les plateformes en ligne pour définir les contenus ne soient pas fondés par défaut sur le profilage, et que l'utilisation de ces algorithmes ne soit possible que si l'internaute donne son accord spécifique librement, en connaissance de cause et sans ambiguïté. La collecte et l'utilisation de données déduites à des fins de personnalisation des publicités et de recommandation de contenus doivent

être interdites (par exemple, les recommandations fondées sur le temps de visionnage ou les mentions « J'aime »). Les utilisateurs et utilisatrices doivent au contraire avoir la possibilité de contrôler les signaux envoyés ou intérêts déclarés qu'ils souhaitent que la plateforme prenne en considération pour établir leur fil d'actualité. Celles et ceux qui préfèrent un fil basé sur des recommandations personnalisées doivent avoir la possibilité de communiquer leurs intérêts personnels à la plateforme en donnant spécifiquement leur consentement, en toute liberté et en connaissance de cause et sur la base d'invites formulées en des termes compréhensibles pour les enfants.

- Promulguer/mettre à jour et faire appliquer des cadres réglementaires garantissant aux personnes l'exercice, dans la pratique, du droit de choisir des solutions respectueuses de la vie privée à la place de modèles économiques fondés sur la surveillance. Il s'agit notamment de limiter l'effet de réseau et de garantir l'interopérabilité (possibilité de communiquer avec des contacts existants en utilisant une autre plateforme compatible) plutôt que la simple portabilité des données, pour que les utilisateurs et utilisatrices puissent passer d'un service à un autre sans que leur vie sociale en pâtisse.

RECOMMANDATIONS À TIKTOK

- Passer à un modèle économique respectueux des droits qui ne soit pas fondé sur un suivi intrusif des données. Pour commencer, TikTok doit veiller à ce que ses politiques et processus de diligence nécessaire en matière de droits humains s'attaquent aux conséquences systémiques et généralisées qu'a son modèle économique sur les droits humains et en particulier sur les droits au respect de la vie privée, à la liberté d'opinion et de pensée et à la santé. TikTok doit se montrer transparente au sujet des risques et violations qu'elle décèle, notamment en ce qui concerne les droits humains, et des mesures qu'elle prend à leur égard.
- Cesser de chercher à maximiser « l'engagement des utilisateurs et utilisatrices » aux dépens de leur santé et autres droits humains, compte tenu des éléments attestant les impacts négatifs de l'utilisation compulsive de la plateforme, en particulier sur la santé et le bien-être des jeunes utilisateurs et utilisatrices. Dans le cadre de son processus de diligence raisonnable en matière de droits humains, TikTok doit identifier, en coopération avec les utilisateurs et utilisatrices, y compris les enfants et les jeunes, ainsi qu'avec des spécialistes indépendants, les éléments de conception qui encouragent une utilisation addictive de la plateforme et des comparaisons sociales, et les remplacer par une expérience axée sur la « sécurité par la conception » et l'intérêt supérieur de l'enfant.
- Faire preuve de la diligence nécessaire en matière de droits humains de façon volontariste et continue, de manière anticipée et tout au long du cycle de vie des technologies algorithmiques, avant et après le déploiement et la mise en œuvre de nouveaux systèmes et de nouvelles fonctionnalités, afin que les risques puissent être détectés dès la phase de développement et que les atteintes aux droits humains et les autres dangers puissent être gérés dès le lancement de ces technologies.
- Faire participer des enfants et des jeunes, des universitaires, des expert·e·s issus de la société civile et d'autres parties prenantes pertinentes aux processus de diligence requise en matière de droits humains menés par l'entreprise. Les enfants et les jeunes doivent aussi jouer un rôle central dans la mise en œuvre de la « sécurité par la conception », en étant associés au processus d'élaboration des outils et des fonctionnalités des plateformes de réseaux sociaux.
- Publier régulièrement des évaluations de l'impact sur les droits humains comprenant des informations détaillées sur les risques et les mesures d'atténuation prises en fonction des pays (en particulier lorsque la diligence requise est plus importante, là où les réseaux sociaux peuvent avoir un impact accru du fait de conflits politiques ou de situations d'urgence humanitaire), des

catégories d'utilisateurs et utilisatrices (les enfants et les jeunes, notamment) et des changements apportés au produit.

- Pour respecter la vie privée des utilisateurs et utilisatrices et leur donner une réelle capacité de choix et de contrôle, un écosystème de réseaux sociaux sans profilage devrait être la norme et non pas une option parmi d'autres. Les algorithmes utilisés par TikTok et les autres plateformes en ligne pour définir les contenus ne devraient donc pas être basés par défaut sur le profilage (par exemple, en fonction du temps de visionnage ou de l'engagement), et doivent reposer sur une démarche volontaire d'inscription plutôt que de désinscription, le consentement pour l'inscription devant être spécifique et donné en connaissance de cause (notamment dans un langage accessible aux enfants), librement et sans ambiguïté.
- Cesser de collecter des données personnelles intimes et d'opérer des déductions quant aux intérêts, à l'état émotionnel ou au bien-être des utilisateurs et utilisatrices en fonction de leur engagement et du temps de visionnage des contenus pour « personnaliser » les recommandations et cibler les publicités. Au lieu de se livrer à une surveillance généralisée pour adapter les fils aux intérêts des utilisateurs et utilisatrices, TikTok doit permettre à ceux-ci d'indiquer par eux-mêmes leurs centres d'intérêt de façon volontaire (par exemple, il pourrait leur être demandé d'entrer les sujets qui les intéressent plus particulièrement s'ils/elles souhaitent obtenir des recommandations personnalisées) et seulement sur la base d'un consentement libre, spécifique et éclairé.
- Prendre des mesures supplémentaires pour empêcher que les utilisateurs et utilisatrices à risque ne tombent dans des comportements compulsifs et des « spirales » de contenus potentiellement nocifs. Il pourrait s'agir notamment d'instaurer un nombre maximal quotidien de recommandations personnalisées proposées aux enfants et une liste régulièrement mise à jour de termes liés à du contenu potentiellement problématique en matière de santé mentale, termes considérés comme acceptables en tant que mots-clés de recherche mais pas à des fins d'amplification dans le fil « Pour toi ».
- Introduire des mesures de « friction » comme stratégie d'atténuation. Dans le cadre de ses processus de diligence raisonnable, TikTok doit investir dans la recherche afin de définir et d'intégrer des mesures destinées à limiter l'amplification algorithmique rapide et souvent disproportionnée des contenus potentiellement problématiques.
- Améliorer dans un premier temps l'efficacité des mesures visant à diversifier les contenus recommandés dans le fil « Pour toi » (notamment en introduisant des outils efficaces de contrôle pilotés par les utilisateurs et utilisatrices) de manière que le fil ne se remplisse pas de contenus qui risquent d'inciter les internautes à se faire du mal. Ces outils doivent être faciles à trouver et à comprendre et fournir aux utilisateurs et utilisatrices un moyen efficace d'empêcher les futures recommandations de contenu liées à un sujet, un hashtag ou un utilisateur/une utilisatrice en particulier.
- Améliorer radicalement la transparence en ce qui concerne l'utilisation des algorithmes de définition et de modération des contenus, en veillant à ce que le public puisse en connaître les mécanismes et que ceux-ci soient expliqués dans le cadre de l'utilisation de l'application, dans des termes clairs, compréhensibles et adaptés aux enfants, dans toutes les langues pertinentes.
- Veiller à la cohérence des décisions prises en matière de modération de contenu, assurer une supervision humaine de la modération automatisée et prévoir des investissements suffisants dans les ressources humaines et automatisées allouées à la modération, à la confiance et à la sécurité, et ce, dans toutes les langues.
- Modifier le système de signalement de manière à le rendre plus accessible et efficace pour que TikTok puisse traiter rapidement et adéquatement les plaintes des utilisateurs·rices ou d'autres parties concernées.

AMNESTY INTERNATIONAL EST UN MOUVEMENT MONDIAL DE DÉFENSE DES DROITS HUMAINS. LORSQU'UNE INJUSTICE TOUCHE UNE PERSONNE, NOUS SOMMES TOUS ET TOUTES CONCERNÉ·E·S.

NOUS CONTACTER

info@amnesty.org

+44 (0)20 7413 5500

PARTICIPER À LA CONVERSATION

www.facebook.com/AmnestyGlobal

@amnesty

ENTRAÎNÉ·E·S DANS LE « RABBIT HOLE »

DE NOUVELLES PREUVES MONTRENT LES RISQUES DE TIKTOK POUR LA SANTÉ MENTALE DES ENFANTS

En novembre 2023, Amnesty International avait publié un rapport accablant, intitulé *Poussé·e·s vers les ténèbres*, qui montrait comment les enfants et les jeunes manifestant un intérêt pour la santé mentale pouvaient être facilement attirés dans une « spirale » de contenus dépressifs et suicidaires sur le fil « Pour toi » de TikTok. Ces recherches avaient révélé que TikTok s'appuyait sur une conception addictive pour générer un maximum d'engagement, portant atteinte aux droits au respect de la vie privée, à la liberté de pensée et à la santé mentale et physique.

Ce nouveau rapport, établi à partir de recherches complémentaires, présente de nouveaux éléments prouvant que, près de deux ans plus tard, le fil « Pour toi » de TikTok continue de pousser les enfants et les jeunes qui manifestent un intérêt pour la santé mentale dans un cercle vicieux de contenus dépressifs ou valorisant l'automutilation et le suicide. Rassemblant des témoignages de jeunes et de parents concernés en France, ainsi que de nouvelles preuves issues de recherches techniques, ce rapport montre que TikTok ne fait rien pour remédier aux risques systémiques que son modèle fait courir aux enfants et aux jeunes, en violation des normes internationales relatives à la responsabilité des entreprises en matière de droits humains ainsi que de ses obligations contraignantes au titre du Règlement sur les services numériques de l'Union européenne. Le rapport lance un appel urgent à l'entreprise elle-même, mais aussi aux autorités de régulation françaises et européennes afin qu'elles agissent fermement pour l'obliger à respecter les droits de l'enfant et les droits humains.

INDEX: POL 40/0360/2025
OCTOBRE 2025
LANGUE : FRANCAIS

amnesty.org

AMNESTY INTERNATIONAL