

# A S S E M B L É E   N A T I O N A L E

1 7 <sup>e</sup>   L É G I S L A T U R E

## Compte rendu

### Commission d'enquête sur les effets psychologiques de TikTok sur les mineurs

- Audition, ouverte à la presse, du Lieutenant-Colonel Cyril Colliou, chef par intérim de l'Office mineurs (Ofmin), Mme Agathe Boudin, commandant de police, cheffe par intérim du pôle stratégie, et Mme Typhaine Desbordes, cheffe du bureau des partenariats et de la communication ..... 2
- Audition, à huis clos, de Mme Sarah Bouchahoua, responsable des affaires publiques France de Snapchat ..... 10
- Table ronde, ouverte à la presse, sur les radicalités : ..... 22
  - M. Tristian Boursier, Docteur associé, Centre de recherches politiques (Cevipof)
  - M. Hugo Micheron, enseignant-chercheur en sciences politiques rattaché au Centre de recherches internationales (CERI), maître de conférence à Sciences Po, spécialiste du Moyen Orient
  - Conseil représentatif des institutions juives de France (CRIF) – Mme Sophie Taïeb, responsable du pôle cybersécurité
  - M. Valentin Petit, journaliste, agence CAPA
- Présences en réunion ..... 43

Mardi

24 juin 2025

Séance de 14 heures 30

Compte rendu n° 30

SESSION ORDINAIRE DE 2024-2025

**Présidence de  
M. Arthur Delaporte,  
Président de la commission,**

**Puis présidence de  
Mme Josiane Corneloup,  
Vice-présidente,**



*La séance est ouverte à quatorze heures trente-cinq.*

*La commission auditionne le Lieutenant-Colonel Cyril Colliou, chef par intérim de l'Office mineurs (Ofmin), Mme Agathe Boudin, commandant de police, cheffe par intérim du pôle stratégie, et Mme Typhaine Desbordes, cheffe du bureau des partenariats et de la communication.*

**M. le président Arthur Delaporte.** Nous recevons le lieutenant-colonel Cyril Colliou, chef par intérim de l'Office mineurs (Ofmin), accompagné de Mme Agathe Boudin, commandante de police, cheffe par intérim du pôle stratégie, et de Mme Typhaine Desbordes, cheffe du bureau des partenariats et de la communication.

Je vous remercie de nous déclarer tout intérêt public ou privé de nature à influencer vos déclarations.

L'article 6 de l'ordonnance du 17 novembre 1958 relative au fonctionnement des assemblées parlementaires impose aux personnes auditionnées par une commission d'enquête de prêter le serment de dire la vérité, toute la vérité, rien que la vérité.

*(M. le lieutenant-colonel Cyril Colliou, Mme Agathe Boudin et Mme Typhaine Desbordes prêtent successivement serment.)*

**M. le lieutenant-colonel Cyril Colliou, chef par intérim de l'Office mineurs.** Je souhaite, à titre liminaire, revenir brièvement sur les raisons qui ont présidé à la création de l'Office mineurs. Ces dernières années, les forces de l'ordre ont constaté une hausse d'environ 30 % par an des violences commises sur les mineurs. La multiplication de ces phénomènes et des faits pédocriminels observés est notamment à mettre en lien avec le développement des technologies numériques et leur accès généralisé, ainsi qu'avec la libération progressive de la parole et une amélioration de l'écoute des victimes.

Il existe un chiffre noir en matière d'atteintes aux mineurs et plus particulièrement de violences sexuelles : les enquêtes de victimation montrent en effet que, dans la population majeure de notre pays, 5,5 millions de femmes et d'hommes ont été victimes dans leur enfance de violences sexuelles, soit un adulte sur dix. Ces violences ont un impact sur le développement affectif et social des enfants victimes et peuvent avoir des conséquences profondes tout au long de leur vie.

La minorité de la victime est, pour les phénomènes dont nous nous occupons, un critère déterminant lors du passage à l'acte de l'auteur. Elle suppose par ailleurs un traitement juridique et judiciaire spécifique des faits commis, avec un accompagnement adapté des victimes.

Ce constat a conduit à la création, au sein de la direction nationale de la police judiciaire (DNPJ), de l'Office mineurs, chargé de la lutte contre les violences faites aux mineurs. Il s'agissait de répondre à un triple défi. Le premier était de contribuer à positionner le ministère de l'intérieur comme un acteur incontournable du dispositif interministériel de protection de l'enfance, pour renforcer la coordination avec les partenaires institutionnels et associatifs et améliorer la détection et l'identification des mineurs victimes. Le deuxième défi était d'animer et de coordonner la lutte contre les violences faites aux mineurs sur l'ensemble du territoire pour développer un traitement spécifique de ces faits. Il s'agissait enfin d'améliorer le traitement judiciaire de ces violences en développant une expertise judiciaire en matière de lutte contre la pédocriminalité.

Le champ infractionnel de l'Office mineurs est extrêmement large, puisqu'il inclut toutes les infractions dont peuvent être victimes les mineurs : les infractions sexuelles, les faits d'exploitation, toutes les violences physiques et psychologiques, dont les meurtres et notamment le syndrome des bébés secoués, et l'ensemble des faits de harcèlement, y compris en ligne.

L'Office mineurs développe plusieurs méthodes et moyens d'investigation dans l'environnement numérique afin de détecter les pédocriminels, que ces derniers agissent sur le *darknet* ou utilisent des messageries cryptées.

Créé par un décret du 29 août 2023 et placé au sein de la DNPJ, il est rattaché organiquement à la sous-direction de la lutte contre la criminalité organisée et la délinquance spécialisée. Il est armé par des policiers et des gendarmes.

**Mme Laure Miller, rapporteure.** Pouvez-vous nous indiquer la part des dossiers traités par l'Ofmin présentant un lien avec l'usage des réseaux sociaux par les mineurs ?

**M. le lieutenant-colonel Cyril Colliou.** L'Office mineurs ne dispose pas à proprement parler de statistiques sur ce point. Je puis toutefois vous indiquer, de manière empirique, que cela représente un volume assez important, surtout sur le contentieux lié à la pédocriminalité en ligne. Les principaux pourvoyeurs de signalements de contenus pédocriminels sont Google, Meta avec Instagram et Facebook, Snapchat, TikTok et X.

**M. le président Arthur Delaporte.** Ne disposez-vous d'aucune ventilation plus précise ?

**M. le lieutenant-colonel Cyril Colliou.** Je n'ai pas de données chiffrées, mais je peux préciser les moyens dont dispose l'Office mineurs et la manière dont il travaille, ce qui vous permettra d'appréhender la part des dossiers ayant trait à ces outils.

L'Office mineurs est organisé autour de deux entités : un pôle stratégie, chargé de l'animation, de la coordination, de la création de partenariats et de la mise en place d'actions de sensibilisation à destination du grand public, et un pôle opérationnel disposant d'une unité de renseignement criminel, point d'entrée au niveau national des signalements de contenus pédocriminels. En 2024, cette unité a reçu 170 000 signalements, soit 465 par jour en moyenne, dont 164 516 provenant de la détection effectuée par les plateformes du numérique elles-mêmes, qui les transmettent à une fondation américaine, le NCMEC – National center for missing and exploited children –, laquelle nous les adresse lorsque la victime ou l'auteur sont localisés en France, grâce notamment à leur adresse IP. Les autres signalements proviennent de services de police étrangers, d'Interpol, d'Europol, de partenaires dont la cellule Signal-sports du ministère des sports et d'associations comme L'Enfant bleu, e-Enfance ou Rebond. Ils viennent s'ajouter à ceux de la plateforme Pharos – plateforme d'harmonisation, d'analyse, de recoupement et d'orientation des signalements – de l'Office anti-cybercriminalité (Ofac), avec lequel nous travaillons.

L'Office mineurs reçoit ces signalements, les traite, les exploite et établit des priorités en fonction du degré d'urgence et de la rapidité d'intervention requise. En 2024, 270 enquêtes ont été ouvertes sur cette base. Quelque 300 signalements ont par ailleurs été transmis à des services territoriaux de police judiciaire de la police nationale ou de la gendarmerie nationale. Cela montre que la part des dossiers relative au contentieux de la pédocriminalité en ligne en lien avec l'utilisation des réseaux sociaux est extrêmement importante dans nos activités.

**Mme Laure Miller, rapporteure.** Vous avez indiqué que les contenus à caractère pédocriminel publiés sur les principales plateformes vous étaient signalés, directement ou

indirectement, mais aussi que ces plateformes étaient utilisées par les pédocriminels pour « recruter » leurs victimes.

**M. le lieutenant-colonel Cyril Colliou.** Les deux aspects coexistent en effet. Les signalements que nous recevons de la part des plateformes numériques sont extrêmement nombreux. Ils peuvent concerner de l'échange de contenus entre pédocriminels, mais aussi des interactions de pédocriminels avec des mineurs utilisateurs de ces réseaux sociaux, qui constituent pour ces personnes un terrain d'action privilégié pour entrer en contact avec des mineurs.

**Mme Laure Miller, rapporteure.** Il semblerait que ces plateformes servent parfois simplement d'intermédiaires pour renvoyer les usagers, notamment les jeunes, vers d'autres plateformes dont le contenu est moins modéré, moins contrôlé, comme OnlyFans ou Mym. Avez-vous observé cela ?

**M. le lieutenant-colonel Cyril Colliou.** On constate en effet une utilisation des réseaux sociaux par rebond, comme une porte d'entrée, une manière pour les pédocriminels d'entrer en relation avec les mineurs ou entre eux.

L'Office mineurs n'observe pas les plateformes OnlyFans ou Mym, puisqu'elles concernent une problématique spécifique de l'exploitation sexuelle des mineurs, potentiellement prostitutionnelle. L'Ofmin cède sur cette thématique sa compétence à l'Office central pour la répression de la traite des êtres humains (OCRTEH), qui est compétent sur le sujet pour les mineurs comme pour les majeurs.

Je peux toutefois vous livrer le constat opérationnel effectué par l'OCRTEH, qui indique dans son état de la menace que « le phénomène du *caming* se développe de plus en plus ». Le *caming* a été défini par la chambre criminelle de la Cour de cassation comme une « pratique consistant à proposer, moyennant rémunération, une diffusion d'images ou de vidéos à contenu sexuel, le client pouvant donner à distance des instructions spécifiques sur la nature du comportement ou de l'acte sexuel à accomplir ». Selon l'OCRTEH, « il existe une grande quantité de plateformes de diffusion utilisées pour le *caming*, mais deux d'entre elles se partagent la majorité de l'activité : OnlyFans, site anglais, et Mym, site français. Ces sites proposent à leurs utilisateurs de s'abonner à des profils afin d'avoir accès à des contenus exclusifs sous forme de vidéos, photos ou messages privés. Certains profils offrent du contenu gratuit. La majorité des producteurs de contenu proposent des abonnements qui oscillent entre 4 et 45 euros par mois, avec un tarif dégressif selon la durée de l'abonnement. Les fans peuvent aussi verser des pourboires en échange d'une dédicace ou d'un contenu personnalisé ». L'OCRTEH détecte de plus en plus d'annonces de prostitution mentionnant un lien vers des comptes OnlyFans ou Mym. Une étude réalisée par cet office en juin 2023 mettait par exemple en lumière 496 annonces du site Sexemodel contenant des liens vers des profils OnlyFans ou Mym, et 395 fin 2023. Ces liens sont utilisés par les personnes prostituées pour promouvoir leur profil de *camgirl* ou de *camboy* et inversement pour faire de la publicité pour des rencontres physiques et des prestations sexuelles.

L'OCRTEH indique que l'état du droit ne permet pas aux services d'enquête de lutter contre le phénomène du *caming*, sauf pour les mineurs. En effet, « cette pratique ne peut satisfaire aux exigences posées par la jurisprudence de la Cour de cassation en 1996 et 2022, qui impose la présence d'un contact physique entre le client et la victime pour que la prostitution soit caractérisée ». Dans le cas des *camgirls*, les victimes sont en général seules face à la caméra ; bien que rémunérées par le client, il n'y a pas de contact physique. Une infraction spécifique a toutefois été créée par la loi du 21 avril 2021, qui vise à protéger les mineurs des crimes et délits sexuels et de l'inceste. Elle réprime notamment le fait, pour un adulte, d'inciter un mineur à se livrer à des

pratiques sexuelles sur internet. Le phénomène des *camgirls* est donc répréhensible, uniquement lorsque la victime est mineure.

**M. le président Arthur Delaporte.** Pensez-vous qu'une évolution du droit serait souhaitable afin de couvrir également le cas des majeurs ?

**M. le lieutenant-colonel Cyril Colliou.** En tant que chef par intérim de l'Office mineurs, je ne traite que de la thématique des mineurs et l'exploitation prostitutionnelle n'entre pas dans mon champ d'expertise et de compétence. Chacune et chacun d'entre vous a la réponse politique à cette question.

**Mme Laure Miller, rapporteure.** Parvenez-vous, au travers des dossiers que vous traitez, à percevoir les principaux impacts de l'usage des réseaux sociaux sur la santé mentale des mineurs ? S'agit-il par exemple essentiellement de cyberharcèlement, d'exposition à des contenus violents ou sexuels ?

**M. le lieutenant-colonel Cyril Colliou.** L'Office mineurs compte dans ses rangs une psychologue, qui est particulièrement attentive à ces questions et dont la mission principale est la prise en charge des victimes. L'analyse des dossiers que nous traitons permet d'émettre l'hypothèse que l'exposition précoce et répétée à des contenus violents génère une croissance des représentations biaisées de la réalité, avec notamment une banalisation de la violence et une représentation de la sexualité reposant sur des scénarios pornographiques. Il convient également de rappeler que l'exposition à des images et contenus violents conduit à des traumatismes, notamment chez les jeunes, et peut laisser des traces dans l'esprit de l'enfant, éventuellement sujet par la suite au développement de troubles. Dans notre champ de compétence principal, à savoir l'exploitation sexuelle, on observe par exemple des troubles d'hypersexualisation susceptibles de conduire l'enfant à s'exposer à de nouveaux dangers. Les risques sont extrêmement importants et les atteintes psychologiques peuvent être lourdes.

**Mme Laure Miller, rapporteure.** Bien que votre office soit de création récente, avez-vous observé une évolution en matière de contrôle et de modération des contenus ? Avez-vous le sentiment que les plateformes jouent le jeu ? Les dossiers liés aux réseaux sociaux sont-ils de moins en moins nombreux ou l'évolution vous semble-t-elle au contraire plutôt défavorable ?

**M. le lieutenant-colonel Cyril Colliou.** Nous observons une croissance quasi exponentielle des signalements, dont près de 165 000 nous ont été transmis par le NCMEC en 2024.

Nous notons chez la plupart des grandes plateformes, dont Meta, des progrès considérables dans les signalements qui nous sont adressés, tant d'un point de vue quantitatif que qualitatif.

Il n'en va pas de même pour la plateforme TikTok. Certes, cette dernière collabore, développe une action en matière de protection des mineurs et transmet ses signalements au NCMEC qui nous les adresse. Mais si l'on observe une explosion du nombre des signalements, il reste à notre sens de nombreux progrès qualitatifs à accomplir pour atteindre les standards des autres plateformes. Beaucoup des signalements qui nous sont transmis sont des faux positifs, qui ne correspondent pas à des contenus pédocriminels : nous passons donc du temps à traiter des signalements qui ne sont pas pertinents. Les signalements adressés sont en outre souvent assez pauvres, en matière notamment d'éléments d'identification et de localisation des auteurs et des victimes. La plupart des plateformes nous communiquent généralement plusieurs adresses IP, ainsi que des pseudonymes et des adresses électroniques vérifiées ; ce n'est pas le cas de TikTok, dont

les éléments fournis ne sont pas consolidés, ce qui nécessite de notre part des investigations complémentaires.

**Mme Laure Miller, rapporteure.** Il existe donc une marge d'amélioration en matière de qualité de la modération et des informations transmises par les plateformes.

**M. le lieutenant-colonel Cyril Colliou.** TikTok nous a adressé 4 432 signalements en 2023, 6 039 en 2024 et 17 589 au 16 juin 2025 : on observe donc une explosion du nombre de signalements. D'un point de vue opérationnel, le traitement que nous en effectuons est toutefois assez résiduel et ne concerne qu'une minorité d'entre eux, en raison du manque de qualité et de pertinence des éléments qui nous sont fournis.

De nombreux phénomènes ayant cours notamment sur TikTok sont toutefois en lien avec les thématiques que nous traitons. On observe ainsi beaucoup de vidéos et de publications virales contenant par exemple des photos assez suggestives de jeunes filles, avec des renvois vers des messageries cryptées comme Telegram. Cela fait écho à l'évolution des tendances et des pratiques évoquée précédemment. Ces applications sont souvent des portes d'entrée renvoyant vers des outils de communication beaucoup plus confidentiels. Nous avons pu observer cela sur la plateforme TikTok et avons reçu des signalements à ce sujet.

Des particuliers ont par ailleurs signalé à la plateforme Pharos, qui nous a transmis cette information opérationnelle, un phénomène que nous avons qualifié d'« emoji pizza » : des personnes considérées comme pédocriminelles s'identifient sur la plateforme TikTok en insérant dans leurs publications une émoticône en forme de pizza. En anglais, les initiales CP correspondent en effet à la fois à *cheese pizza* – pizza au fromage – et à *child pornography* – pédopornographie. Ces publications renvoient les pédocriminels vers des messageries cryptées. Or cela a été très peu signalé par TikTok, alors que le phénomène a donné lieu à de nombreux signalements de la part d'internautes auprès de la plateforme Pharos.

Nous avons également observé le phénomène dit « zizi challenge », qui n'est pas spécifique à TikTok et a été observé également sur YouTube, à la différence près que cette plateforme propose une véritable modération, qui conduit au blocage des vidéos et à de nombreux signalements de la part de la plateforme elle-même. Cette situation a un impact considérable sur notre activité, puisqu'elle nous conduit à transmettre quasiment chaque semaine des signalements opérationnels au parquet pour enquête. Il ne s'agit généralement pas de pédocriminalité, mais de jeunes, voire de très jeunes enfants, âgés parfois de 7 ou 8 ans, qui se mettent en scène, se filment le plus souvent avec le téléphone de leurs parents, dont ils utilisent le compte YouTube, et diffusent les vidéos sur internet, avec un vrai risque. Exerçant son rôle de modération, YouTube bloque ces contenus et nous les signale. Nous envoyons ensuite ces éléments aux juridictions locales pour enquête, afin notamment qu'un contact puisse être établi avec les parents concernés, dans une dimension plus sociale que répressive.

**M. le président Arthur Delaporte.** Observez-vous ce phénomène sur d'autres plateformes que YouTube ?

**M. le lieutenant-colonel Cyril Colliou.** Nous savons que le phénomène existe aussi sur TikTok, mais cela nous est très peu signalé. Cette plateforme se prêterait pourtant à ce type de pratique. Le fait que l'on ne puisse pas observer le phénomène faute de signalements nous interpelle donc. Je ne peux pas affirmer que cela traduit un problème de modération, mais la situation permet légitimement de le penser. Cela concerne le « zizi challenge », mais aussi le « challenge maillot ».

Il existe aussi des phénomènes de sextorsion, c'est-à-dire de corruption de mineurs ou *grooming*. Nous recevons ainsi par l'intermédiaire de Pharos des signalements émanant de particuliers qui relatent des prises de contact et des faits de sextorsion dont sont victimes des mineurs sur le réseau social TikTok. Toutefois, lorsque nous effectuons l'analyse des signalements qui nous parviennent de ce réseau, nous constatons qu'aucun ne concerne de tels faits. Nous pouvons donc affirmer qu'il existe un véritable problème de modération et de détection des phénomènes de sextorsion par la plateforme TikTok.

Concernant enfin les risques suicidaires, nous recevons assez régulièrement, en provenance notamment de Pharos, des signalements de messages de mineurs exprimant des idées noires à tendance suicidaire. Or nous n'avons encore reçu aucune information opérationnelle de ce type de la part de TikTok. Là encore, cela interroge, car on peut légitimement penser que de tels messages peuvent exister sur cette plateforme.

**Mme Laure Miller, rapporteure.** Pourriez-vous nous apporter des précisions sur les relations de l'Office mineurs avec Pharos ?

Entretenez-vous par ailleurs des liens similaires avec l'ensemble des plateformes ou certaines se montrent-elles plus coopératives et communiquent-elles davantage avec vous ?

**M. le lieutenant-colonel Cyril Colliou.** Pharos et l'Office mineurs sont tous les deux rattachés à la direction nationale de la police judiciaire. Les contacts sont donc étroits et permanents, quasi quotidiens. Pharos reçoit des signalements transmis par les utilisateurs de plateformes et l'Office mineurs des signalements qui lui sont adressés par les plateformes du numérique. Nous échangeons et croisons les informations reçues. Lorsque nous détectons des contenus illicites dans les signalements qui nous parviennent, nous les transmettons à Pharos pour qu'ils soient supprimés. Lorsque la plateforme Pharos reçoit des signalements correspondant à notre champ de compétence, c'est-à-dire à des infractions à caractère sexuel notamment dont peuvent être victimes des mineurs, elle nous les envoie, ce qui nous permet pour certains faits de transmettre du renseignement opérationnel à nos relais territoriaux et le cas échéant, pour les faits les plus graves et les enquêtes les plus complexes, d'ouvrir des investigations.

L'Ofac dispose d'un réseau d'antennes et de détachements dans les territoires. Quasiment toutes les semaines, l'Ofmin est amené à réaliser des investigations, des enquêtes judiciaires avec interpellation de mis en cause dans des affaires pédocriminelles. Pour procéder aux interpellations, nous nous reposons notamment sur le réseau de l'Ofac et sur ses investigateurs en cybercriminalité (ICC), spécialistes notamment de la saisie et de l'exploitation des supports numériques. Or, dans les contentieux que nous traitons, notamment les infractions pédocriminelles, les perquisitions conduisent systématiquement à la saisie et à l'exploitation des téléphones et des ordinateurs des mis en cause.

Nous avons des contacts constants avec la plateforme Pharos, partenaire privilégié de l'Ofmin. Nous entretenons des liens assez variables avec d'autres plateformes – il en existe tellement ! Nous avons des relations très privilégiées avec Meta, Instagram et Facebook, compte tenu de la qualité de la modération et du nombre de signalements qu'elles nous adressent. Des liens existent avec X et Snapchat. S'agissant de TikTok, je peux dire que la plateforme répond assez rapidement aux réquisitions et aux demandes des enquêteurs, mais qu'il existe une substantielle marge de progression en matière de qualité des retours et de la modération.

**M. le président Arthur Delaporte.** Avez-vous identifié des flous juridiques ou des qualifications juridiques trop faibles qui nécessiteraient une évolution du droit pénal, afin d'assurer une meilleure appréhension de certains comportements émergents ?

**M. le lieutenant-colonel Cyril Colliou.** Nous assistons à l'émergence de contenus pédocriminels créés par l'Intelligence artificielle (IA) – nous avons observé le phénomène dans plusieurs enquêtes. Ces contenus peuvent être créés *ab initio*, mais aussi à partir de contenus pédocriminels existants et modifiés ou bien d'images récupérées, notamment sur les réseaux sociaux, avant d'être détournées de leur usage et sexualisées. Pour nous, c'est un véritable enjeu opérationnel pour les années à venir : l'IA peut rapidement faire exploser le volume des contenus pédocriminels qui circulent en ligne, tant cet outil facilite leur création.

Or nous recevons déjà 170 000 signalements à l'année, que nous nous employons à traiter. Parmi eux, certains dossiers sont considérés comme complexes et prioritaires : les enquêtes dans lesquelles une victime encore non identifiée est potentiellement en cours d'exploitation sexuelle. Nous utilisons la base de données internationale d'images et de vidéos sur l'exploitation sexuelle des enfants (ICSE), gérée par Interpol, et le fichier français Caliope – comparaison et analyse logicielle des images d'origine pédopornographique. Quand les contenus qui circulent sur internet ne sont pas connus de ces bases de données, nous nous efforçons d'identifier les victimes et de les localiser. L'année dernière, nous avons ainsi contribué à l'identification de soixante victimes, dont vingt étaient encore en situation d'exploitation sexuelle – en moyenne âgées de moins de 8 ans, la plus jeune n'ayant que quelques semaines. L'émergence de l'IA dans ce domaine risque d'accroître notre charge de travail et de nous faire travailler sur l'identification de victimes qui n'existent pas. Dans les années à venir, il faudra que nous trouvions un moyen de faire la différence entre les images et contenus pédocriminels produits par l'IA et ceux qui sont réels.

Autre écueil : l'avenir incertain de la législation européenne en matière de détection et de signalements volontaires par les plateformes des abus sexuels commis en ligne contre des enfants. La directive 2002/58/CE du Parlement européen et du Conseil du 12 juillet 2002 concernant le traitement des données à caractère personnel et la protection de la vie privée dans le secteur des communications électroniques (directive e privacy) prévoit des dispositions dérogatoires temporaires, qui prendront fin en avril 2026. Si aucun règlement définitif ne vient la remplacer pour enjoindre aux plateformes de détecter ces contenus, nous risquons de devenir aveugles : une large partie des 165 000 signalements du NCMEC effectués en 2024 risquent de disparaître. Pour nous, il est essentiel de consolider le dispositif actuel de détection par les plateformes. Il faudrait aussi renforcer la responsabilité de ces plateformes concernant l'hébergement et la diffusion de sites et contenus à caractère pédocriminel.

**M. le président Arthur Delaporte.** Y a-t-il d'autres éléments que vous souhaiteriez aborder avec nous, qu'il s'agisse de préconisations, d'analyses ou de témoignages ? Avez-vous en tête un cas précis, lié à TikTok, que vous pourriez nous relater ?

**M. le lieutenant-colonel Cyril Colliou.** Nous avons traité peu de dossiers relatifs à TikTok, les signalements ayant été assez rares.

Mais je pense à un dossier qui illustre l'importance de la détection de contenus pédocriminels en ligne et la nécessité d'élaborer et de diffuser de bonnes pratiques au sein des forces de l'ordre, afin d'améliorer le traitement des dossiers de violences faites aux mineurs. L'affaire, qui se déroule dans la moitié nord de la France, commence il y a cinq ou six ans par un signalement à la justice de violences incestueuses commises sur une petite fille. Le service d'enquête saisi auditionne la jeune victime, le mis en cause et quelques personnes de l'entourage. Une expertise

pédopsychiatrique est diligentée, qui conclut aux tendances à l'affabulation de la fillette. Le dossier est alors classé sans suite. Le père reste en famille avec ses deux enfants, la fillette et son frère. On peut déjà dire que l'enquête n'a pas été réalisée dans les règles de l'art et que les actes accomplis apparaissent largement insuffisants.

En 2023, à la création de l'Ofmin, une plateforme d'internet nous adresse un signalement avec des vidéos à caractère pédocriminel sur lesquelles nous identifions rapidement cette fillette qui avait dénoncé des faits, rapportés par un enseignant à la justice, quelques années plus tôt. Nous reprenons alors l'enquête depuis le début et nos investigations nous permettent de démontrer que la jeune fille ainsi que son frère ont été violés à de nombreuses reprises pendant des années, le violeur utilisant notamment des sédatifs. Entre le moment où le dossier a été classé sans suite et celui où les vidéos sont arrivées à notre connaissance par le biais le biais du NCMEC, six ou sept enfants se sont ajoutés à la liste des victimes de ce même violeur.

S'agissant de la révélation des faits, je peux vous dire que nous recevons seulement 65 000 plaintes alors que la Ciivise recense 160 000 victimes de violences sexuelles par an, ce qui tend à montrer que les deux tiers des faits ne sont pas signalés aux forces de l'ordre. Il est donc essentiel que nous continuions à disposer d'outils de détection de faits pédocriminels sur les supports numériques, compte tenu de la place que ceux-ci ont pris dans le quotidien de la population.

**M. le président Arthur Delaporte.** Qu'en est-il des moyens ? Combien d'enquêteurs sont mobilisables pour traiter ce volume vertigineux de contenus pornographiques ou autres ?

**M. le lieutenant-colonel Cyril Colliou.** Toutes les forces de l'ordre s'occupent du contentieux pédocriminel. À l'Ofmin, nous serons cinquante-huit à la rentrée de septembre 2026, alors que nous devions être quatre-vingt-cinq selon les préfigurations.

Nous sommes en train de développer un réseau d'antennes et de détachements, calqué sur celui de l'Ofac et principalement localisé dans les chefs-lieux de zones de défense. À terme, nous disposerons de dix-sept antennes et détachements, soit 188 personnels spécifiquement dédiés à ce contentieux au sein de la police nationale. On peut ajouter les 1 600 personnes affectées aux brigades de protection des familles et chargées notamment du traitement des violences intrafamiliales, y compris sur des mineurs lorsqu'il n'existe pas d'antenne et de détachement. On pourrait multiplier par deux, trois, quatre ou cinq les effectifs de l'Ofmin, nous serions tous occupés. Même si nous sommes convaincus de ce que nous faisons, nous avons parfois le sentiment de vider la mer avec une petite cuillère.

**M. le président Arthur Delaporte.** Merci pour ce que vous faites, dans des conditions pas simples, en étant confrontés à un monde très dur.

**Mme Laure Miller, rapporteure.** Pouvez-vous nous dire un mot sur vos actions de prévention ?

**M. le lieutenant-colonel Cyril Colliou.** C'est le pôle stratégique, représenté par les personnes qui m'accompagnent, qui est à la manœuvre en ce qui concerne la prévention visant les mineurs. Il nous arrive de conduire quelques actions de sensibilisation, mais force est de constater qu'elles restent marginales car ce n'est pas notre mission première et que nous n'avons que peu de ressources à y consacrer. Nous voudrions conclure une convention de partenariat en matière de prévention avec l'éducation nationale, sachant que nous avons déjà avec elle des partenariats opérationnels concernant l'identification des victimes. Nous avons noué des partenariats et signé

des conventions avec des associations de protection de l'enfance telles que e-Enfance et L'Enfant bleu, avec lesquelles nous menons des actions croisées de formation et d'information.

Sur le volet préventif, il y aurait à dire sur la responsabilité cruciale des parents dans l'usage que font leurs enfants des réseaux sociaux et des médias. On pourrait penser, par exemple, à la constitution d'un réseau de parents vigilants. Dans nos interventions à destination des parents, nous avons souvent face à nous des personnes informées et sensibles à ces questions, pas forcément celles qu'il faudrait toucher en priorité. Un réseau pourrait aider les parents en fracture numérique à réagir par rapport à leurs enfants.

Il y a aussi à faire en matière de parentalité, au regard du nombre de photos d'enfants diffusées sur internet. On considère qu'à l'âge de 13 ans, un enfant a en moyenne 1 300 photos de lui en circulation sur les réseaux. Nombre de photos qui apparaissent dans les contenus pédocriminels très explicites et dans les bibliothèques des pédocriminels ont été volées sur les réseaux sociaux. Des photos d'enfants nus sur la plage, anodines pour le commun des mortels, peuvent être détournées par des pédocriminels avec l'aide de l'IA. Il faudrait sensibiliser les parents à ce risque et les conduire à développer une approche pédagogique des réseaux sociaux avec leurs enfants. Normalement, les enfants ne devraient pas avoir accès aux réseaux sociaux avant l'âge de 13 ans, et ils ne devraient y accéder qu'avec le consentement des parents s'ils sont âgés de 13 à 15 ans. Or nombre d'enfants de moins de 15 ans y accèdent en totale autonomie.

**M. le président Arthur Delaporte.** Cette audition touche à sa fin, mais nous restons à votre disposition et nous sommes preneurs de tout élément complémentaire que vous voudriez porter à notre connaissance.

**M. le lieutenant-colonel Cyril Colliou.** Nous vous adresserons les réponses que nous avons préparées sur le volet partenarial et stratégique, qui nous semble très important.

**M. le président Arthur Delaporte.** Tracez-vous des perspectives d'évolution de ces partenariats ?

**M. le lieutenant-colonel Cyril Colliou.** Nous sommes en relation constante avec nos réseaux et nous voudrions développer certains partenariats, notamment avec l'éducation nationale. La direction nationale de la police judiciaire a posé sa candidature pour assumer le pilotage du projet Impact, portant sur l'animation et la coordination en matière de prévention et de répression des violences faites aux mineurs à l'échelle des forces de l'ordre européennes. L'Ofmin et la DNPJ ont prévu de nombreuses actions.

**M. le président Arthur Delaporte.** Il me reste à vous remercier pour votre participation à nos travaux.

*Puis la commission auditionne, à huis clos, Mme Sarah Bouchahoua, responsable des affaires publiques France de Snapchat.*

**M. le président Arthur Delaporte.** Nous accueillons Mme Sarah Bouchahoua, responsable des affaires publiques France de Snapchat, qui, pour des raisons personnelles, n'a pas pu participer à la table ronde organisée la semaine dernière avec l'ensemble des plateformes. Invoquant des raisons de sécurité, vous avez demandé, madame, que cette audition se tienne à huis clos : Snapchat n'étant pas au cœur des travaux de notre commission, nous avons accepté de faire une exception.

Je vous remercie de déclarer tout intérêt public ou privé de nature à influencer vos déclarations. Par ailleurs, l'article 6 de l'ordonnance du 17 novembre 1958 relative au fonctionnement des assemblées parlementaires impose aux personnes auditionnées par une commission d'enquête de prêter le serment de dire la vérité, toute la vérité, rien que la vérité.

*(Mme Sarah Bouchahoua prête serment.)*

**Mme Sarah Bouchahoua, responsable des affaires publiques France de Snapchat.** Je vous remercie d'avoir accepté de reporter cette audition : Snapchat tenait à participer à cette discussion essentielle car la protection de notre communauté en ligne, en particulier celle des plus jeunes, est notre priorité absolue.

Snapchat a été fondé en 2011 avec une mission claire : reproduire les échanges quotidiens avec nos amis ou notre famille dans un espace numérique que nous voulons sûr, agréable et centré sur une communication authentique. La conception de la plateforme repose donc sur deux principes fondamentaux : la sécurité et la protection de la vie privée par défaut – *safety and privacy by design*. Nous sommes fermement convaincus que l'espace numérique ne devrait pas se résumer à un terrain de compétition entre utilisateurs, où la quête du sensationnel et la recherche de la viralité l'emporteraient sur l'authenticité des échanges.

Particulièrement attaché à la protection de la santé mentale des jeunes, Snapchat ne propose pas de métrique de comparaison sociale, comme les *likes*. L'application s'ouvre sur le mode caméra, et non sur un fil d'actualité infini qui inciterait à scroller, et nous n'encourageons pas la désinformation ni les contenus toxiques. Cette philosophie est partagée par notre communauté.

Selon notre dernier rapport de transparence pour l'Union européenne, plus de 850 millions d'utilisateurs uniques au niveau mondial se connectent chaque mois sur la plateforme, dont plus de 26,7 millions en France. Ces snapchateurs disent que l'application les rend heureux – ni anxieux, ni tristes : nous en sommes fiers. Ce sentiment a été confirmé par une étude indépendante, que nous n'avons pas commandée, menée récemment par l'université d'Amsterdam et la professeure Patti Valkenburg, experte ès réseaux sociaux. Pendant 100 jours, elle a évalué le bien-être, l'estime de soi et la qualité des relations de 479 adolescents de 14 à 17 ans suite à leur utilisation de plusieurs plateformes. Selon cette étude, Snapchat est la seule plateforme à avoir eu un impact positif sur le bien-être des jeunes, qui estiment que l'application renforce fortement la proximité entre amis, et n'a entraîné aucun impact négatif sur l'estime de soi.

Conscients des enjeux actuels en matière de protection des mineurs, nous travaillons chaque jour à renforcer nos mécanismes de protection et de sécurité, ainsi qu'à approfondir nos partenariats avec les acteurs publics comme privés. Notre comité consultatif de sécurité – *safety advisory board* – réunit plus de vingt experts internationaux, dont Mme Justine Atlan, directrice générale de l'association e-Enfance, que vous avez eu l'occasion d'entendre dans le cadre de vos travaux et avec laquelle nous travaillons étroitement pour développer des outils adaptés aux défis du numérique.

Développé avec ce comité, le contrôle parental permet de remettre le parent au cœur de l'activité numérique de son adolescent, tout en respectant la vie privée et l'autonomie de ce dernier. Nous proposons également un outil de signalement, développé avec l'Arcom (Autorité de régulation de la communication audiovisuelle et numérique) suite aux recommandations contenues dans son bilan sur la lutte contre la manipulation de l'information. Facilement accessible sur l'application et le site support, et organisé en catégories claires pour garantir une utilisation simple,

cet outil rappelle notamment que les signalements sont confidentiels et essentiels pour encourager un usage responsable de la plateforme.

Bien avant le 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 dit *Digital services act* (DSA), nous avons lancé un programme de signaleurs de confiance auquel participent plusieurs associations en France comme e-Enfance, Point de Contact, Stop Fisha et la Licra (Ligue internationale contre le racisme et l'antisémitisme). Après un lancement réussi aux États-Unis, nous avons récemment étendu à l'Europe notre conseil consultatif de jeunes, qui réunit des adolescents de différents pays membres de l'Union européenne et leurs parents ou tuteurs légaux. Ils rencontrent régulièrement nos équipes pour partager leur expérience et nous aider à améliorer nos outils et notre politique. Cette initiative sera officiellement lancée cette semaine : à cette occasion, les deux représentants français, âgés de 14 et 17 ans, se rendront avec leurs parents dans nos bureaux à Amsterdam.

Comme nous sommes contraints par le temps, je ne détaillerai pas davantage l'ensemble des initiatives prises par Snapchat. Reste que les approches uniformes ont montré leurs limites : il est temps de repenser la conception même des plateformes et le phénomène d'amplification virale au cœur de certains modèles économiques.

En s'ouvrant sur le mode caméra, Snapchat invite à s'exprimer plutôt qu'à faire défiler passivement un flux de contenus. Les surfaces de contenu y sont modérées : Discover propose des contenus publiés par des créateurs certifiés et des médias partenaires vérifiés – notamment *Le Monde*, *Le Nouvel Obs*, *Le Figaro* ou *Brut* –, qui sont tous soumis à des règles relatives aux contenus ; sur Spotlight, nous mettons en avant des contenus positifs et modérons les vidéos avant qu'elles ne soient visibles d'un large public. Snapchat est conçu pour favoriser un usage actif plutôt que passif. Le cœur de notre application reste la messagerie privée, fondée sur la réciprocité : pour pouvoir échanger, les utilisateurs doivent s'ajouter mutuellement. Contrairement à d'autres plateformes, on ne peut pas recevoir des messages par simple invitation. Toute capture d'écran ou enregistrement des messages est notifié, afin que l'utilisateur soit acteur de sa conversation.

Les adolescents font l'objet de mesures de protection spécifiques : ils n'apparaissent pas dans les résultats de recherche des autres utilisateurs, sauf si un lien mutuel est probable. Si un jeune accepte un inconnu dans sa liste d'amis, un bandeau s'affiche sur la partie conversationnelle pour lui signaler qu'il partage peu d'amis en commun avec cette personne. Une alerte similaire s'affiche si l'interlocuteur se trouve dans une autre zone géographique de celle où réside le mineur, ou si un écart d'âge important est avéré.

Tous ces engagements ne datent pas d'hier : contrairement à ceux de plusieurs plateformes, qui ont souligné avoir renforcé leurs mesures de protection, les comptes Snapchat ont toujours été privés. Comme le confirment notre audit sur les risques systémiques, nos rapports de transparence et nos déclarations auprès de l'Arcom – et auparavant auprès du CSA (Conseil supérieur de l'audiovisuel) – dans le cadre de la lutte contre la manipulation de l'information, Snapchat n'est pas un vecteur de diffusion massive de contenus illicites, un outil de manipulation de l'information ou d'enfermement algorithmique. Nous nous imposons des standards élevés et renforçons en permanence nos dispositifs de sécurité. Conscients du chemin qu'il reste à parcourir, nous continuons de renforcer la coopération avec le public et le privé : il est essentiel d'appeler collectivement à un environnement numérique beaucoup plus éthique, dans lequel chaque acteur assume pleinement ses responsabilités sans se défaire sur les autres.

Les défis actuels appellent une action coordonnée de l'ensemble du secteur du numérique. À cet égard, le règlement (UE) 2022/1925 du Parlement européen et du Conseil du 14 septembre 2022 relatif aux marchés contestables et équitables dans le secteur numérique et modifiant les directives (UE) 2019/1937 et (UE) 2020/1828 dit *Digital markets act* (DMA) et le DSA offrent des fondations solides : ils doivent désormais être appliqués de manière cohérente et harmonisée dans l'ensemble des États membres. Il est temps d'éviter les initiatives nationales fragmentées, qui nuisent au bon fonctionnement du marché unique numérique. Je note que nous sommes une des seules plateformes à avoir salué l'entrée en vigueur du DSA ; nous avons d'ailleurs travaillé à sa rédaction avec plusieurs eurodéputés.

Il est temps, aussi, de redonner aux parents les moyens de soutenir et accompagner leurs enfants dans leur parcours numérique, sans leur faire de reproches. L'exemple récent de l'interdiction des sites pornographiques en France illustre bien les dangers d'une régulation précipitée ou non concertée : les téléchargements de VPN – réseaux privés virtuels – ont explosé, exposant davantage les mineurs aux risques du numérique. De plus, l'arrêté ministériel interdisant ces sites a été suspendu car il ne respectait pas le droit européen. Il faut trouver des solutions réalistes, applicables, qui protègent les utilisateurs sans les inciter à adopter des comportements encore plus risqués.

**Mme Laure Miller, rapporteure.** Toutes les plateformes nous ont affirmé, comme vous, que la protection des mineurs était leur priorité absolue. Mais nous nous intéressons moins à tout ce que vous déployez déjà pour protéger les enfants qu'à ce qu'il reste à faire pour mettre fin à ce qui nous paraît problématique.

Pouvez-vous nous en dire plus sur ce qui contribue à capter l'attention des enfants sur Snapchat, notamment l'algorithme et les fonctionnalités qui les incitent à revenir sur la plateforme ? Je pense en particulier aux émojis, qui peuvent être accolés au nom de l'utilisateur – un bébé pour qui vient de rejoindre la plateforme, un cœur jaune pour les meilleurs amis, etc. – et disparaissent au bout de vingt-quatre heures, ou les snapflames. Avez-vous analysé ces dispositifs ? Avez-vous pleinement conscience qu'ils peuvent contribuer à la dépendance aux écrans, en particulier à votre réseau social ?

**Mme Sarah Bouchahoua.** Snapchat est conçu principalement pour faciliter la communication entre amis réels et encourager l'authenticité et les échanges entre utilisateurs. À cette fin, l'application s'ouvre directement sur le mode caméra, et non sur un fil d'actualité gouverné par des algorithmes de recommandation totalement opaques.

L'application offre trois espaces publics. Tout d'abord la carte, qui présente l'actualité mondiale et permet de visualiser les *stories* publiques diffusées par les autres utilisateurs. Plus elle est rouge, plus l'activité est intense – généralement, c'est en raison d'événements sportifs ou culturels, comme des festivals musicaux. Ensuite, la partie Discover fonctionne comme un kiosque à journaux ; elle propose des *stories* publiées par des médias certifiés et reconnus pour leur éthique journalistique avec lesquels nous avons noué plus de cent partenariats en France, comme *Le Monde*, *Le Nouvel Obs* ou *Konbini*. Enfin, la partie Spotlight affiche des vidéos courtes, généralement humoristiques – un format plébiscité par nos utilisateurs. Ces trois espaces font l'objet d'une modération poussée : avant d'être diffusés, les contenus sont d'abord scannés par un algorithme pour vérifier qu'ils ne contiennent rien d'explicitement illicite – terrorisme ou pédocriminalité –, puis soumis au regard d'un modérateur humain, qui s'assure qu'ils ne contreviennent pas aux règles communautaires de Snapchat, comme un placement de produit non déclaré.

Afin de lutter contre l'enfermement algorithmique et les bulles de filtre, et de proposer aux utilisateurs un contenu diversifié présentant différents points de vue, nous avons beaucoup travaillé avec l'Arcom sur l'algorithme de recommandation. Nous l'avons présenté à plusieurs reprises à la Commission européenne, qui n'a formulé aucune observation négative. Cet algorithme s'appuie à la fois sur des critères subjectifs – temps passé sur le contenu, partage – et objectifs – localisation de l'utilisateur, langue maternelle, âge – pour lui proposer des contenus d'actualité pertinents.

Comme je l'ai expliqué devant la commission d'enquête concernant l'organisation des élections en France, en juillet 2024, nous avons proposé aux utilisateurs des contenus politiques liés à la tenue des élections législatives anticipées et aux Jeux olympiques. Nous nous efforçons de proposer des contenus en lien avec les centres d'intérêt déclarés par l'utilisateur dans les paramètres de l'application mais, pour éviter de le conforter dans ses opinions, Snapchat propose également des contenus diversifiés, sans lien avec ces centres d'intérêt. Par exemple, une personne qui adore le sport se verra certes proposer des contenus sportifs provenant de médias partenaires spécialisés, comme *L'Équipe*, mais aussi des contenus liés à l'actualité politique ou simplement divertissants. Je n'aime pas trop le football mais, comme vous pouvez le voir dans la partie Discover, Snapchat me propose des contenus sur le joueur Erling Haaland, d'autres au sujet du concert de Beyoncé, que j'ai effectivement apprécié, des actualités proposées par M6 ou des informations sur les nouvelles lunettes Ray-Ban.

Nous proposons également d'autres fonctionnalités, qui peuvent être considérées comme addictives, comme les Snapstreak – les fameuses flammes. Leur utilisation est facultative : nous n'imposons rien et nous n'envoyons aucune notification relative à cette fonctionnalité – pour alerter l'utilisateur qu'une flamme est en train de s'éteindre, par exemple. D'après plusieurs enquêtes menées avec des chercheurs et notre *Teen Council* – Conseil consultatif de jeunes –, notre audience apprécie ce type d'outils car ils permettent de renforcer l'amitié à travers un petit jeu. Conscients que ces fonctionnalités entraînent des effets de bord, nous nous efforçons de les corriger : une fois par an, les utilisateurs peuvent demander une restauration gratuite de leur flamme directement sur l'application – davantage s'ils disposent de Snapchat + –, ou sur le site support de Snapchat. Je précise que l'utilisateur peut modifier les paramètres de l'application pour désactiver les notifications ou l'algorithme de recommandations spécialisé, afin que les contenus proposés ne reposent que sur l'algorithme objectif.

**Mme Laure Miller, rapporteure.** Savez-vous quelle part des jeunes a effectivement désactivé l'algorithme spécialisé ?

**Mme Sarah Bouchahoua.** Je n'ai pas de chiffres précis. La possibilité de personnaliser les paramètres est encore largement méconnue des collégiens et lycéens. Nous avons travaillé avec les médias, comme le magazine *Parents*, et la Cnil (Commission nationale de l'informatique et des libertés) pour mieux informer les jeunes sur l'existence de ces outils, mais aussi déconstruire les préjugés autour des dispositifs de signalement, trop souvent perçus comme des outils de délation.

**Mme Laure Miller, rapporteure.** De nombreux influenceurs utilisent Snapchat car le format très court des vidéos les rend plus percutantes, ce qui leur garantit un succès fulgurant. Nous avons auditionné Naldas, un cas emblématique qui ne compte pas moins de 9 millions d'abonnés. Ne pensez-vous pas que cette mise en valeur est une prime aux contenus choquants ?

**Mme Sarah Bouchahoua.** Pour être visible sur les espaces Stories ou Spotlight, les contenus doivent respecter les règles d'éligibilité à la recommandation : les contenus violents ou à caractère sexuel, sexiste, raciste ou antisémite sont interdits.

Comme je l'avais expliqué au président Delaporte lors des travaux préparatoires à l'examen de la proposition de loi visant à encadrer l'influence commerciale et à lutter contre les dérives des influenceurs sur les réseaux sociaux, Snapchat applique le principe de la réponse graduée.

**M. le président Arthur Delaporte.** Selon Nasdas, votre plateforme constitue sa principale source de revenus, notamment grâce à ses contenus viraux. Confirmez-vous cette déclaration ? Le cas échéant, combien gagne-t-il ?

**Mme Sarah Bouchahoua.** Je ne connais pas le montant exact.

**M. le président Arthur Delaporte.** Le modèle économique nous intéresse. Comment Snapchat rémunère-t-il les créateurs de contenus ? Nasdas nous a dit qu'il gagnait plusieurs centaines de milliers d'euros par mois : quelles vidéos lui rapportent le plus d'argent ? Plus largement, combien y a-t-il d'influenceurs millionnaires, en nombre de vues mais aussi en argent sonnante et trébuchante ? Plusieurs centaines de milliers d'euros par mois, ça fait plusieurs millions par an !

Par ailleurs, n'est-il pas possible de prévenir la production de contenus problématiques ? Pourquoi Nasdas, dont plusieurs contenus ont été modérés, n'a-t-il pas été banni ? À partir de combien de contenus modérés bannissez-vous un utilisateur ? C'est une des solutions à la main des plateformes.

**Mme Laure Miller, rapporteure.** Le principe même de la réponse graduée est discutable, car il signifie qu'il est possible d'insulter ou de s'attaquer à quelqu'un plusieurs fois. Le créateur d'un contenu inapproprié ou ne respectant pas vos règles devrait être banni dès la première infraction.

**Mme Sarah Bouchahoua.** Pour prévenir la production de contenus problématiques, nous avons créé en 2021 des Snap schools, où les créateurs de contenus suivent une formation aux règles communautaires et aux conditions d'utilisation de Snapchat avant leur certification. Des salariés dédiés à la relation avec les créateurs de contenus leur rappellent les règles à respecter, leur fournissent un guide et restent en relation avec eux tout au long de leur expérience.

Comme j'ai déjà eu l'occasion de l'expliquer au président Delaporte, la gradation des sanctions que nous appliquons depuis 2021 dépend de différents critères, notamment de la gravité de l'infraction aux règles de modération. Des audits trimestriels sur les créateurs de contenus qui ont déjà enfreint les règles, qu'ils soient ou non certifiés, ont conduit à bannir plus d'une quinzaine de créateurs.

J'en viens à la politique de monétisation, lancée récemment en France. Elle n'est accordée qu'à certains créateurs de contenus et obéit à des règles strictes : si un utilisateur viole les règles de modération, la monétisation est suspendue – il peut même en être définitivement privé.

**M. Thierry Sothier (SOC).** Vous avez tendance à présenter votre plateforme comme une simple messagerie mais, compte tenu du nombre de fonctionnalités que vous avez décrites, Snapchat s'apparente plutôt à un réseau social.

Vous avez indiqué qu'avant d'être certifiés sur votre plateforme, les influenceurs suivent une formation, ou au moins un temps d'accompagnement, au siège de Snapchat. Combien de personnes ont-elles été formées à ce jour ?

S'agissant de la gradation des sanctions, vous semblez vous contenter de la promesse d'un influenceur de ne plus recommencer. J'ai une fille de six ans qui me fait régulièrement ce genre de promesse et bien souvent, ça ne suffit pas à l'empêcher de recommencer. Jusqu'où vont les sanctions ? Sont-elles limitées à des avertissements à l'infini ? Combien d'influenceurs ou d'utilisateurs certifiés bannissez-vous chaque année ?

**Mme Sarah Bouchahou.** Nous distinguons les influenceurs – des personnalités publiques qui peuvent créer un compte sur Snapchat sans être soutenues par la plateforme – des créateurs de contenus certifiés, les Snap stars. Seuls ces derniers, identifiables à leur badge jaune – qui, contrairement à ce qui se fait sur d'autres plateformes, ne peut pas être acheté ni donné à n'importe qui –, suivent la formation dispensée dans les locaux de Snapchat dans le cadre du programme Snap schools.

Pour être certifiées, les personnalités publiques qui en ont fait la demande ou approchées à l'initiative de la plateforme doivent remplir plusieurs critères objectifs. Par exemple, elles doivent être reconnues pour leur talent, qu'elles soient athlètes ou artistes. Une fois certifiées, ces utilisateurs peuvent publier des contenus dans la partie Discover. Les influenceurs, eux, ne sont pas nécessairement certifiés ni formés par la plateforme aux règles de modération et de recommandation.

Le rôle de l'équipe chargée de la relation avec les Snap stars est d'alerter les utilisateurs certifiés d'éventuels problèmes avec leur compte. Il nous arrive également de retirer leur certification à des personnalités publiques qui ne respectent pas les règles de modération.

**M. le président Arthur Delaporte.** Peut-on parler d'Alex Hitchens ?

**Mme Sarah Bouchahou.** À ma connaissance, il n'a pas de compte sur Snapchat.

**M. le président Arthur Delaporte.** J'ai sous les yeux un compte qui publie sur Spotlight des vignettes comme « Alex Hitchens fait tourner sa copine pour vendre plus », « Des hommes qui se font entretenir par leur femme, leur copine et leur culot » « Alex Hitchens détruit cette femme », etc. Ce ne sont que des vignettes, mais elles méritent une analyse plus contextuelle.

**Mme Laure Miller, rapporteure.** De combien de modérateurs disposez-vous ? Parmi eux, combien sont des sous-traitants et combien des salariés de Snapchat ? Comment leur nombre a-t-il évolué ces dernières années ?

Vous avez dit tout à l'heure que vous n'encouragez pas la désinformation et les contenus toxiques ; permettez-vous néanmoins que ces contenus existent et soient diffusés sur la plateforme ?

**Mme Sarah Bouchahou.** C'est assez difficile, car nous ne pouvons malheureusement pas contrôler les utilisateurs lorsqu'ils produisent eux-mêmes leurs contenus. En revanche, nous mettons en œuvre une modération combinant l'intelligence artificielle, pour détecter les contenus explicitement violents, et l'intervention humaine : une équipe travaille sur ce qui a été repéré par l'algorithme, pour éviter notamment les faux positifs.

Nous avons 1 108 modérateurs dédiés aux pays membres de l'Union européenne, dont 163 francophones. La plupart sont basés à Londres. Ce chiffre peut sembler faible par rapport à d'autres plateformes, mais notre modèle est fondamentalement différent des autres. Pour diffuser un contenu à une large audience sur Snapchat, il faut être une personnalité publique certifiée – une Snap star – ou avoir déjà un large public. Snapchat étant essentiellement centré sur la

communication privée, entre amis et en famille, on y trouve moins de contenus susceptibles d'être considérés comme viraux ou comme des contenus illicites en propagation rapide.

Nos systèmes automatisés permettent de détecter rapidement tout contenu très violent ou explicitement illicite. Nos équipes de modérateurs prennent ensuite le relais pour contrôler. Elles sont assez efficaces et le délai médian entre la détection et l'action finale est très court : 2 minutes pour les contenus à caractère sexuel, 23 minutes pour ceux relevant de l'exploitation sexuelle des mineurs, 7 minutes pour le harcèlement et l'intimidation, 8 minutes pour les menaces et les violences, 10 minutes pour l'automutilation et le suicide, 1 minute pour la désinformation, 6 minutes pour la drogue, 1 minute pour les armes, 27 minutes pour les discours haineux, 5 minutes pour le terrorisme et l'extrémisme violent.

Les modérateurs sont formés pour répondre rapidement et le taux de précision est élevé : d'après les dernières données que nous avons transmises à la Commission européenne, il atteint 95 % s'agissant des contenus entrant en violation avec nos règles communautaires et 85 % pour les contenus non éligibles à la recommandation.

Durant plusieurs semaines, ils suivent régulièrement des modules sur les politiques internes de Snapchat, les outils utilisés, les procédures d'escalade pour remonter le plus rapidement possible les risques, les normes culturelles ou encore les événements d'actualité. Nous les avons prévenus, par exemple, des risques qui pouvaient survenir lors de la fête de la musique ; ils se sont donc tenus prêts à traiter tout signalement à cette occasion.

À la fin de leur apprentissage, les modérateurs doivent obligatoirement obtenir une certification confirmant qu'ils ont les compétences nécessaires pour devenir agent pour Snapchat. Une fois qu'ils l'ont obtenue, ils suivent, comme tous les salariés, une formation continue constituée de sessions hebdomadaires de remise à niveau portant notamment sur les dernières évolutions en matière de régulation et de législation. Pour savoir comment traiter les contenus qui sont dans la zone grise et prendre en compte le contexte, des cas limites sont présentés. Des quiz et des certifications régulières sont proposés, et des procédures d'alerte assez rapides sont mises en œuvre : je fais par exemple des points réguliers avec notre équipe de Londres pour l'alerter sur les tendances que nous percevons.

Au-delà de ces mécanismes internes, nous nous appuyons aussi beaucoup sur l'externe : pour mieux appréhender les tendances observables sur le terrain.

Si nous ne pouvons pas prévenir la publication de certains contenus par les utilisateurs, nous pouvons toutefois essayer de limiter les risques. À l'occasion des Jeux olympiques de Paris 2024, nous avons travaillé trois mois à l'avance. J'ai pris l'initiative de contacter différentes associations – SOS Racisme, la Licra, le Conseil représentatif des institutions juives de France (Crif), Stop Homophobie, SOS Homophobie, etc. Nous avons travaillé avec elles, lors de réunions bilatérales, sur les différents phénomènes qui pourraient se produire. J'ai ensuite transmis l'ensemble des informations obtenues aux équipes de modération, afin qu'elles les prennent en considération.

**Mme Laure Miller, rapporteure.** Comment est-il possible de laisser passer autant de contenus problématiques avec un tel processus de modération ? L'intelligence artificielle doit être capable d'alerter lorsque sont publiées des vidéos d'un homme piquant des gens dans la rue ! Si certains influenceurs ont été bannis récemment, c'est grâce à l'intervention de personnes dont ce n'est pas le métier – notamment du président Arthur Delaporte. Comment expliquez-vous cet énorme trou dans la raquette ?

**Mme Sarah Bouchahoua.** Je ne connais pas le détail du dossier concernant Amine Mojito.

Quant aux autres créateurs de contenus évoqués, ils ont été bannis de la plateforme en 2021. Nous n'avons pas attendu la ministre Aurore Bergé pour supprimer le compte d'AD Laurent, par exemple, mais celui-ci continue de revenir et d'être suspendu par les équipes de modération de Snapchat.

**Mme Josiane Corneloup (DR).** L'absence de vérification des comptes Snapchat permet aux prédateurs en ligne de créer de fausses identités pour tromper les jeunes. Selon les chiffres d'une organisation caritative britannique spécialisée dans la protection des enfants, Snapchat est la plateforme la plus utilisée pour le pédopiégeage en ligne : c'est là que se produisent près de la moitié des délits de ce type.

L'anonymat permet aux prédateurs de cibler, piéger et manipuler leurs victimes d'autant plus facilement que la disparition du contenu public rend les preuves de leurs interventions éphémères et donc difficiles à collecter par la suite. Alors que le nombre de faux comptes augmente sur Snapchat, ne serait-il pas temps de distinguer ceux derrière lesquels il y a de vraies personnes des autres, sans scrupules, qui sont à l'origine d'escroqueries de toute nature ?

Snapchat récompense ses utilisateurs actifs : lorsque deux d'entre eux se sont envoyé des photos ou des vidéos au cours des dernières vingt-quatre heures, ils obtiennent une snapflamme qui continue de brûler à condition qu'ils s'échangent des messages au quotidien. En l'absence d'interaction, leur série disparaît et ils doivent repartir de zéro. Ne pensez-vous pas que de telles pratiques sont de nature à provoquer très vite des addictions ?

**Mme Sarah Bouchahoua.** D'abord, l'anonymat n'existe pas sur internet. Cela fait plus de huit ans que je travaille sur le sujet, ayant été collaboratrice parlementaire de la députée Laetitia Avia, auteur de la proposition de loi contre les contenus haineux sur internet : à cette occasion, nous avons compris que c'est le régime du pseudonymat qui régnait aujourd'hui dans l'espace numérique. La question est celle de la collaboration entre les plateformes et les autorités judiciaires. Lorsque vous vous connectez, vous livrez une adresse e-mail – fût-elle totalement fictive – et un numéro de téléphone, mais surtout une adresse IP et le type de téléphone utilisé : vous êtes ainsi facilement localisable et identifiable.

Nous luttons depuis plusieurs années contre le phénomène particulièrement affreux du *grooming*, ou piégeage. Nous avons mis en place plusieurs garde-fous sur Snapchat. D'abord, le profil des mineurs qui s'inscrivent est privé, si bien qu'ils ne peuvent pas apparaître parmi les suggestions d'ajout des inconnus – lesquels peuvent mentir sur leur âge. Si un mineur ajoute une personne qu'il ne connaît pas, qui a déjà été signalée auparavant ou qui est localisée dans un autre territoire, un bandeau s'affiche pour le mettre en garde et lui demander s'il souhaite réellement poursuivre l'échange. Le bouton « bloquer » ou « signaler » apparaît en rouge, alors que le bouton « poursuivre la conversation » apparaît en gris.

À l'avenir, ce mécanisme de signalement sera disponible sur l'ensemble de la plateforme. Si l'utilisateur réalise que l'inconnu lui demande des contenus intimes, il pourra le signaler le plus rapidement possible. Nous avons rédigé un petit laïus avec e-Enfance pour réexpliquer que le signalement est totalement confidentiel – l'interlocuteur n'en est pas alerté – et qu'en cas d'urgence, il faut s'adresser directement à un adulte ou une personne de confiance, et appeler les forces de l'ordre.

Nous sommes conscients que des efforts restent à faire. Nous avons développé un centre parental visant à remettre le parent au cœur de l'expérience numérique de son enfant, qui s'efforce de reproduire leurs échanges quotidiens : ayant accès à la liste d'amis ainsi qu'au degré d'interaction avec chacun d'eux, le parent pourra signaler un ami qu'il ne connaîtrait pas, par exemple.

Nous avons travaillé avec plusieurs associations britanniques – notamment celle que vous avez évoquée, mais aussi StopNCII (*Stop Non-Consensual Intimate Image*) – spécialisées dans le *grooming* et le *hashing*. Lorsque la victime de l'un de ces méfaits leur transmet les photos ou les vidéos qu'elle a partagées avec un inconnu, ces associations nous les signalent. En scannant ensuite l'ensemble des parties publiques de la plateforme, nous pouvons mesurer si elles ont été diffusées à une large échelle et déterminer par qui elles l'ont été, afin de sanctionner directement le prédateur sexuel.

Nous estimons aujourd'hui que la solution réside dans les systèmes d'exploitation (OS) des téléphones portables, qui disposent de l'ensemble des informations relatives aux utilisateurs. Je ne dis pas cela pour me dédouaner, contrairement à certains : les plateformes sont les principales responsables de la vérification de l'âge et nous continuerons d'ailleurs d'investir sur ce sujet. Néanmoins, les OS devraient nous transférer des signaux. Nous nous réjouissons à cet égard que la société Apple ait annoncé la semaine dernière qu'elle allait transmettre aux développeurs d'application la date de naissance de ses utilisateurs, par le biais d'une API – *Application Programming Interface*.

Aujourd'hui, votre téléphone portable peut vous reconnaître grâce à la biométrie ; il contient vos cartes de fidélité ; il sait exactement ce que vous faites et qui vous êtes. Ce que nous demandons, c'est que dès l'achat du téléphone – que l'appareil soit neuf ou d'occasion – à l'étape du paramétrage, le parent puisse renseigner l'âge de son enfant, donner son consentement pour le téléchargement d'applications et limiter le temps d'écran. Aujourd'hui ces outils existent, ils sont gratuits, mais peu de parents les utilisent. Pour nous, il est essentiel qu'il y ait une seule porte d'entrée et que les OS nous fournissent un signal pour que nous puissions automatiquement activer nos contrôles parentaux. Encore une fois, je ne cherche pas à me dédouaner mais à trouver une solution efficace et effective, en lien avec le DSA et le DMA.

**M. le président Arthur Delaporte.** Vous voyez bien qu'il y a un problème. Tout ce que vous proposez pour les enfants n'est pas opérant, parce que la plupart mentent au sujet de leur âge. Quel est le pourcentage de mineurs identifiés comme tels sur Snapchat ?

**Mme Sarah Bouchahoua.** Nous comptons aujourd'hui, en France, 2,6 millions d'utilisateurs mineurs âgés de 13 à 17 ans.

**M. le président Arthur Delaporte.** Vous n'êtes donc pas capable d'identifier les enfants entre 8 et 13 ans, alors que l'on sait que nombre d'élèves de CM2 ou de sixième utilisent Snapchat. Combien de millions d'utilisateurs avez-vous en France ?

**Mme Sarah Bouchahoua.** Nous avons 26,9 millions d'utilisateurs.

**M. le président Arthur Delaporte.** Le pourcentage de mineurs est donc censé s'établir à 10 %, alors que l'on peut estimer qu'il atteint plutôt 20 à 25 %. Cela signifie qu'au moins la moitié de vos utilisateurs mineurs apparaissent comme majeurs et ne bénéficient donc pas des modalités de protection renforcée que vous avez évoquées.

**Mme Sarah Bouchahoua.** Je ne peux pas me prononcer sur des chiffres non vérifiés mais d'après Médiamétrie, dont vous avez auditionné le président, l'âge moyen déclaré sur Snapchat est de 36 ans. C'est le fameux effet Harry Potter : Snapchat a été créé en 2011, et les adolescents d'alors sont devenus des adultes.

**M. le président Arthur Delaporte.** Et selon vous, quel est l'âge moyen sur Snapchat ?

**Mme Sarah Bouchahoua.** Il est de 35 ans.

**M. le président Arthur Delaporte.** Il est donc à peu près cohérent.

**Mme Sarah Bouchahoua.** Médiamétrie a une méthode de calcul différente, fondée sur un panel et sur une API directement branchée sur Snapchat, au travers de laquelle ils répondent aux questions des utilisateurs.

**Mme Josiane Corneloup (DR).** Ce que l'on constate, c'est que les enfants sont seuls devant l'application et que la plupart des parents n'en connaissent pas le fonctionnement. Personne ne vérifie rien et tout se fait dans l'anarchie totale, sans contrôle. Il faudrait que les parents soient au moins formés et avisés des risques que prend leur enfant en se connectant.

**Mme Sarah Bouchahoua.** C'est un très bon point. Nous organisons des campagnes promotionnelles au sujet des outils que je vous ai présentés, afin d'informer les parents de leur existence.

Nous avons fait réaliser des études indépendantes : d'après le Family Online Safety Institute (FOSI), plus de 75 % des parents ne sont pas au courant des différents outils existant sur les espaces numériques. La plupart d'entre eux disent qu'ils n'ont pas le temps de se connecter et de créer des comptes. C'est la raison pour laquelle nous plaçons pour une solution reposant sur les systèmes d'exploitation : ils offrent une porte unique, qui permettrait aux parents d'autoriser ou non le téléchargement d'applications et de limiter le temps d'écran facilement.

Nous avons développé un site internet, parents.snapchat, qui explique ce qu'est Snapchat, comment l'utiliser, comment en parler avec son enfant et comment activer l'outil de contrôle parental. Nous faisons des efforts, mais le sujet doit aussi être traité au niveau industriel et sociétal : l'éducation nationale – avec laquelle nous essayons de travailler – nous a indiqué, par exemple, que de moins en moins de parents assistaient aux réunions avec les professeurs.

**Mme Laure Miller, rapporteure.** Qu'est-ce qui vous empêche d'instaurer un dispositif de vérification de l'âge au moment de l'inscription ?

**Mme Sarah Bouchahoua.** C'est également une bonne question. Nous poursuivons nos discussions avec les différents vérificateurs d'âge. Malheureusement, aucune solution ne permet de distinguer les jeunes de 13 ans de ceux qui ont 15 ans. Pour l'accès aux sites pornographiques, la limite est fixée à 18 ans...

**M. le président Arthur Delaporte.** Je ne crois pas que cela réponde à la question de Mme la rapporteure. Celle-ci voudrait savoir pourquoi l'âge de l'utilisateur n'est pas vérifié dès son inscription sur la plateforme, ce qui permettrait de certifier – de façon certes imparfaite – l'âge de l'ensemble des utilisateurs. Nous avons posé la même question à TikTok.

**Mme Sarah Bouchahoua.** Aucun outil, justement, ne permet cette vérification d'âge. Nous avons discuté avec Yoti – je ne sais pas si vous les avez auditionnés...

**M. le président Arthur Delaporte.** C'est le système utilisé par TikTok, notamment.

**Mme Sarah Bouchahoua.** Yoti ne permet pas de faire la distinction entre un utilisateur âgé de 14 ans et 7 mois et un autre de 15 ans et 2 mois. En outre, il y a quelques temps, la marge d'erreur était de 30 %.

**M. le président Arthur Delaporte.** Vous avez donc connaissance de cette marge d'erreur, que Yoti vous a communiquée. Les représentants de TikTok, à qui nous avons posé la même question mot pour mot, ne nous en ont pas parlé.

**Mme Sarah Bouchahoua.** C'est connu...

**M. le président Arthur Delaporte.** C'est une question importante. Vous n'utilisez donc pas Yoti ?

**Mme Sarah Bouchahoua.** Non, nous ne l'utilisons pas en raison de la marge d'erreur, qui n'est pas améliorable aujourd'hui.

**M. le président Arthur Delaporte.** Ne serait-ce pas un moindre mal ? On pourrait au moins vérifier qu'un utilisateur a entre 14 et 16 ans, et pas 8 ans.

**Mme Sarah Bouchahoua.** La marge d'erreur concerne le seuil de 18 ans, pas la différence entre 14 ans et 7 mois et 15 ans et 2 mois.

Nous estimons qu'aucun outil ne permet aujourd'hui, dans le respect des droits fondamentaux des adolescents, de mettre en œuvre une vérification de l'âge lors de l'inscription. Mais nous continuons d'échanger avec des vérificateurs d'âge.

Nous travaillons aussi sur la *biodata* et, avec la Commission européenne, sur le *digital wallet*. Nous sommes ouverts à tout et nous nous efforçons simplement de trouver la solution la plus efficace pour les adolescents.

**Mme Josiane Corneloup (DR).** Ne pourraient-ils pas fournir une pièce d'identité, tout simplement ?

**Mme Sarah Bouchahoua.** On en parle depuis des années. La Cnil y a apporté une réponse assez claire. D'abord, les utilisateurs n'ont pas forcément de pièce d'identité ; ensuite, une pièce d'identité française vaut-elle une pièce espagnole ? Le DSA s'applique aux vingt-sept pays membres de l'Union, c'est donc au niveau européen que nous devons trouver une solution. Enfin, le risque en matière de cybersécurité est élevé : moi-même, je ne souhaite pas fournir ma pièce d'identité à certains acteurs discutables de l'espace numérique.

**Mme Laure Miller, rapporteure.** Même si elles sont imparfaites, il existe des solutions. Vous ne pouvez pas commencer votre propos en expliquant que la priorité absolue de Snapchat est la protection des enfants et nous dire ensuite qu'en l'absence de solution tout à fait satisfaisante, vous préférez ne rien faire. C'est totalement contradictoire.

**Mme Sarah Bouchahoua.** Ce n'est pas que nous préférons ne rien faire. Nous développons des outils de protection, nous lançons des campagnes au sujet des usages et des mesures de protection à disposition des adolescents et des parents. Mais nous devons trouver une solution industrielle à l'échelle européenne. À cet égard, le texte solide adopté à l'initiative de la France est une bonne base.

**M. le président Arthur Delaporte.** En tant que responsable des affaires publiques de Snapchat, recevez-vous des alertes Google ? Faites-vous de la veille d'actualité pour identifier ce qui pourrait nuire à l'image de l'entreprise ?

**Mme Sarah Bouchahoua.** Non, je reçois des alertes sur des articles de presse, mais pas d'alertes Google. Pour renforcer nos actions, nous nous appuyons essentiellement sur les alertes transmises par des associations au sujet de phénomènes perçus sur Snapchat ou sur d'autres espaces numériques.

**M. le président Arthur Delaporte.** Avez-vous lu l'article évoquant l'arrestation par des gendarmes de Corrèze, le 24 avril 2025, de dealers qui avaient ouvert une boutique en ligne sur Snapchat pour vendre de la drogue ?

Pourquoi le trafic de drogue se fait-il sur Snapchat ?

**Mme Sarah Bouchahoua.** Il ne se fait pas uniquement sur Snapchat.

**M. le président Arthur Delaporte.** Non, un peu sur Telegram aussi...

**Mme Sarah Bouchahoua.** Je suis très au fait de cette question. Le trafic commence à se diriger vers d'autres réseaux sociaux. Snapchat a renforcé ses différents outils, notamment en matière de modération.

Enfin, nous travaillons à la sensibilisation. Aujourd'hui, si un utilisateur tape le mot « cannabis », les contenus qui s'affichent, travaillés avec la mission interministérielle de lutte contre les drogues et les conduites addictives (Mildeca), ne promeuvent pas cette substance mais renvoient vers Drogues Info Service.

**M. le président Arthur Delaporte.** Nous arrivons au terme de l'audition. Nous sommes preneurs de vos réponses écrites au formulaire ainsi que de précisions sur les profils des utilisateurs et sur leurs revenus. Nous sommes très intéressés, notamment, par les éléments qui nous aideraient à comprendre le modèle économique de Nasdas.

*Enfin la commission auditionne M. Tristan Boursier, Docteur associé, Centre de recherches politiques (Cevipof), M. Hugo Micheron, enseignant-chercheur en sciences politiques rattachées au Centre de recherches internationales (CERI), maître de conférence à Sciences Po, spécialiste du Moyen Orient, Mme Sophie Taïeb, responsable du pôle cybersécurité au Conseil représentatif des institutions juives de France (CRIF), M. Valentin Petit, journaliste, agence CAPA.*

**M. le président Arthur Delaporte.** Nous reprenons nos auditions sous la forme d'une table ronde sur les radicalités avec M. Tristan Boursier, docteur associé au Centre de recherches politiques (Cevipof) ; M. Hugo Micheron, enseignant-chercheur en sciences politiques, rattaché au Centre de recherches internationales (Ceri), professeur associé à Sciences Po, spécialiste du Moyen-Orient ; Mme Sophie Taïeb, responsable du pôle cybersécurité du Conseil représentatif des institutions juives de France (Crif) ; et M. Valentin Petit, journaliste à l'agence CAPA. Je vous remercie de nous déclarer tout intérêt public ou privé qui serait de nature à influencer vos déclarations, en particulier toute rémunération que vous pourriez percevoir de la part d'un acteur du numérique.

L'article 6 de l'ordonnance du 17 novembre 1958 relative au fonctionnement des assemblées parlementaires impose aux personnes auditionnées par une commission d'enquête de prêter le serment de dire la vérité, toute la vérité, rien que la vérité.

*(M. Valentin Petit, Mme Sophie Taïeb, M. Hugo Micheron et M. Tristan Boursier prêtent successivement serment.)*

**M. Tristan Boursier, docteur associé au Centre de recherches politiques (Cevipof).** Je suis chercheur postdoctoral, associé au Cevipof et à l'UQO (Université du Québec en Outaouais). Mes recherches ne sont pas financées par des entreprises privées ou des groupes du numérique.

Dans le cadre de mes recherches, je m'intéresse aux discours et aux idéologies d'extrême droite et antiféministes, plus spécifiquement aux convergences entre ces deux discours sur les médias sociaux. Mes recherches se focalisent sur YouTube, TikTok et, dans une moindre mesure, Instagram. J'associe des méthodes qualitatives – analyse de discours et immersion en ligne, notamment en menant des enquêtes ethnographiques en ligne au sein de communautés d'extrême droite en ligne – à des méthodes quantitatives appliquées sur un nombre important de vidéos YouTube, dont j'analyse les discours.

Je travaille sur le sujet depuis cinq ans et j'ai observé trois changements principaux, notamment chez les influenceurs d'extrême droite qui sont au centre de mes recherches : d'abord, une professionnalisation de ces acteurs, avec une meilleure maîtrise des codes du numérique et du capitalisme de l'attention ; deuxièmement, un phénomène de prolifération – multiplication des microcélébrités, des microacteurs dont les publications sont fortement visibles, malgré le faible nombre d'abonnés ou de contenus ; troisièmement, une grande adaptation de ces acteurs aux différentes modes et tendances des médias sociaux et aux évolutions des réglementations des plateformes.

**M. Hugo Micheron, enseignant-chercheur en sciences politiques rattaché au Centre de recherches internationales (Ceri), professeur associé à Sciences Po, spécialiste du Moyen-Orient.** Je suis professeur associé à Sciences Po, spécialiste du Moyen-Orient. Je travaille abondamment sur le djihadisme et, depuis 2020, je m'intéresse énormément à la question informationnelle, notamment à la diffusion des différentes idéologies en ligne et au pouvoir des algorithmes. Nous avons déployé un outil, Arlequin AI, qui permet de documenter les diffusions d'un certain nombre de ce qu'on appelle désormais les « narratifs » en ligne.

Dans un premier temps, je m'en tiendrai à présenter ce qui me semble être la réalité du monde politique dans lequel nous évoluons tous. L'espace informationnel est en proie à des mutations extrêmement rapides, violentes et profondes, qui redéfinissent en grande partie les attributs de la puissance, notamment publique et politique. Ainsi, des acteurs non étatiques, parfois très minoritaires, se retrouvent à avoir un poids considérable, complètement disproportionné par rapport à leur capacité à représenter les groupes dont ils se réclament.

Parmi eux, on trouve des acteurs de comités dits à risque. Je pense à mon sujet de recherche de ces quinze dernières années : le djihadisme. Nous avons affaire, par définition, à des groupes très minoritaires qui parviennent néanmoins à diffuser en ligne des éléments de propagande, caractéristique très importante de ces groupes. En effet, dès les années 1980, au tout début du djihad contemporain, ces acteurs étaient déjà très investis dans les supports de propagande – cassettes vidéo et audio, pamphlets qui s'échangeaient sous le manteau ou sur les places de marché quand ils y parvenaient. À partir des années 2000, les supports se sont en

partie numérisés, ils ont été diffusés sur des forums et stockés sur des formats dématérialisés, comme des fichiers PDF. À partir des années 2010, au moment du pic de Daech, ils sont diffusés sur les réseaux sociaux sous une forme souvent qualifiée d'hollywoodienne, qui reprend tous les codes de la communication culturelle occidentale. Depuis quelques années, on voit apparaître les premières vidéos et les premiers produits propagandiques liés à l'intelligence artificielle. La première vidéo francophone de Daech produite à 100 % par l'intelligence artificielle (IA) a été mise en ligne sur les réseaux sociaux en 2024.

Ce sujet n'est pas à la marge mais au cœur du débat politique. Dans le monde dans lequel nous vivons, les algorithmes forment un pouvoir à part entière. Jusque-là, nous nous sommes projetés dans un certain modèle de séparation des pouvoirs – l'Assemblée nationale fait partie des lieux où ces principes ont été pensés et mis en pratique : le politique s'incarne dans l'exécutif et le législatif, qui sont séparés l'un de l'autre et également distincts du judiciaire. Un autre pouvoir, qui est un contre-pouvoir, a émergé au fil du temps : les médias et la société civile – associations et corps intermédiaires. Au gré du temps, des affrontements et des ajustements, tout cela a constitué un socle démocratique européen et, plus largement, une base démocratique qui est un modèle mondial.

Aujourd'hui, il existe un nouveau pouvoir, le pouvoir algorithmique, qui, pour l'instant, ne fait face à aucun contre-pouvoir. Il redéfinit les attributs de la puissance entre les différents pouvoirs. Dans le cadre de l'étude de l'évolution de la propagande djihadiste, en m'intéressant aux algorithmes, j'ai abordé de nombreux sujets, parmi lesquels la question des grandes plateformes du numérique. M. Elon Musk, le patron de X, a trafiqué l'algorithme de X dont il est le propriétaire afin que, pendant la campagne électorale, les tweets politiques de M. Trump soient vus 17 milliards de fois par les citoyens américains – je ne parviens même pas à concevoir ce que ce chiffre représente. J'ignore quel en a été l'impact. Néanmoins, cette diffusion est énorme, et bien plus importante que la capacité de M. Elon Musk à toucher ces personnes. Dans cette démocratie, comme en France et en Europe, les campagnes électorales sont très encadrées, tant du point de vue du financement que du calendrier – on ne peut pas dire tout et n'importe quoi et à quarante-huit heures du scrutin, la campagne s'arrête. Or ce n'est pas le cas sur les réseaux sociaux.

En Roumanie, pays plus proche de nous, un candidat, considéré à l'extrême droite de l'extrême droite, a fait 100 % campagne sur TikTok et a déclaré zéro euro de frais de campagne. Il n'était connu de personne – il était crédité de moins de 1 % des votes – mais il a fini par obtenir 43 % des votes au premier tour de l'élection présidentielle. Ce qui a sauvé la situation, en quelque sorte, c'est qu'il n'a pas déclaré de frais de campagne alors qu'il a été prouvé qu'il avait été financé à hauteur de 1 million d'euros, notamment par des agents extérieurs. Cet individu, inconnu en Roumanie, a été hissé à la huitième place des tendances mondiales sur TikTok durant cette période. Il existe donc une déformation algorithmique car rien ne laissait présager que sur un réseau social chinois en Roumanie, il s'élèverait au-dessus des contenus relatifs au conflit à Gaza ou aux commentaires portant sur M. Joe Biden ou M. Donald Trump. Ce qui est certain, c'est qu'il a surperformé.

En étudiant des événements antérieurs, on s'est rendu compte que les mêmes méthodes avaient été utilisées neuf mois plus tôt pour pousser la vente d'aspirateurs en ligne, qui ont été en rupture de stock en Roumanie deux semaines plus tard. Des systèmes de marketing ont été mis en place puis appliqués au marketing politique. À partir de cas très concrets, liés aux élections, avec un impact politique direct aux États-Unis ou en Europe, on constate que les algorithmes sont arsenalisés. Ces questions touchent tous les sujets politiques, et les tendances à risque que je couvre en font partie.

L'audition est suspendue de dix-sept heures trente à dix-sept heures trente-cinq.

**Mme Sophie Taïeb, responsable du pôle cybersécurité du Conseil représentatif des institutions juives de France (Crif).** Je suis chargée de la cybersécurité et des signalements pour le compte du Crif. Le Crif, créé en 1943, est une association universaliste qui fédère soixante-cinq associations. L'une de nos missions statutaires est de lutter contre l'antisémitisme et toutes les formes de haine.

Depuis dix ans, nous sommes partenaires des principales plateformes – X, Meta, Google et TikTok. Nous avons été certifiés signaleurs de confiance par l'Autorité de régulation de la communication audiovisuelle et numérique (Arcom) au printemps 2025. Nous sommes également en relation avec certaines institutions – la plateforme Pharos (d'harmonisation, d'analyse, de recoupement et d'orientation des signalements) et la délégation interministérielle à la lutte contre le racisme, l'antisémitisme et la haine anti-LGBT (Dilcrah). Notre expertise en matière de détection de signalements de contenus haineux est reconnue.

Chaque année, plusieurs milliers de signalements sont reçus par le Crif par mail, via un formulaire sur notre site internet ou la messagerie. Nous les traitons tous manuellement : je décide des contenus qui seront signalés aux plateformes et/ou à Pharos. Nous ne signalons que les contenus illicites qui enfreignent soit la loi française, soit les règles d'utilisation des plateformes. C'est pourquoi notre taux de retrait est d'environ 90 %, toutes plateformes confondues.

Parmi les contenus les plus signalés qui entrent dans notre domaine d'expertise figurent les contenus négationnistes, la discrimination en raison de la race ou de la religion, les discours de haine, l'apologie du terrorisme et le *doxxing*, à savoir la divulgation d'informations et de données personnelles.

Nous rencontrons des difficultés spécifiques à TikTok. Si vous tapez « Hitler » sur TikTok, les vidéos dans lesquelles ce mot est prononcé n'apparaîtront pas. Selon les études, on estime qu'il y a entre 300 et 400 vidéos téléchargées chaque seconde sur TikTok. Or il est impossible de récupérer les contenus dans lesquels sont prononcés les mots recherchés, contrairement à X où il suffit de saisir le mot-clé Hitler ou Hamas pour qu'apparaissent les tweets qui utilisent ces mots. La recherche de contenus illicites ne s'effectue donc pas à partir du contenu de la vidéo, mais grâce aux signalements que nous recevons et à l'algorithme ultrapuissant. En effet, à force de consulter des contenus antisémites, TikTok me propose automatiquement des vidéos antisémites.

De plus, les utilisateurs savent tromper la modération. Si vous tapez « *Hitler was right* » – « Hitler avait raison » – sur Tiktok, vous recevrez un avertissement vous invitant à vérifier les faits, notamment dans des livres d'histoire. Mais si vous faites partie de la génération qui sait contourner la modération, vous remplacez le i de Hitler par un point d'exclamation et le e par un trois, et vous accédez alors à des vidéos dont le contenu est potentiellement nazi.

L'antisémitisme sur Tiktok est diffus, il est partout : dans les noms d'utilisateurs, les vidéos, les commentaires. Je vous en donne quelques exemples : des comptes s'intitulent *gazeur2juif*, des *lives* sont accompagnés de commentaires tels que « on va attaquer les synagogues jusqu'à la mort », des vidéos très vulgaires sont publiées, que je ne lirai pas ici.

La dissuasion est faible. Voici une vidéo dans laquelle une jeune femme dit : « Vous allez tous finir en enfer, bande de Juifs, vous allez tous cramer [...] », avant de prononcer des

propos très vulgaires. Il a fallu plusieurs jours pour faire retirer ce contenu, puis plusieurs autres pour faire suspendre le compte. Cette jeune femme de Chambéry est déjà de retour sur TikTok car elle a pu recréer un compte immédiatement.

L'antisémitisme est global. Selon des études, 69 % des Juifs disent avoir été victimes d'antisémitisme sur la plateforme. Les commentaires ne sont pas forcément liés au contexte politique. Sous cette vidéo où l'on voit une famille en train de danser pendant une bar-mitsvah, 3 500 commentaires ont été publiés parmi lesquels « Le grand-père, il a du gaz qui sort de sa bouche, c'est normal ? », « Papy, va à la douche », « Free Palestine », « Ça fête la mort du H », « bande de juifs », etc.

La modération est aléatoire. À mon grand étonnement, on m'a répondu que les deux comptes intitulés gazeur2juif n'enfreignaient aucune règle. Dans ce genre de cas, je fais appel par mail à notre contact chez TikTok France, mais je ne peux le faire à chaque fois que je ne comprends pas une décision de modération.

**M. le président Arthur Delaporte.** Malgré le fait que vous ayez été désigné signaleur de confiance, si vous n'appelez pas votre contact, le contenu n'est pas retiré.

**Mme Sophie Taïeb.** Absolument. Il est ici question de la plateforme *Trusted flaggers*, qui est partenaire de TikTok, la plateforme Signaleurs de confiance sera, quant à elle, mise en place prochainement.

Le puissant algorithme TikTok est utilisé comme source d'information par les jeunes. Il suggérera des vidéos allant dans le même sens que celles que l'utilisateur a déjà regardées. S'il commence à regarder des contenus antisémites ou avec des *fake news*, il lui en proposera davantage. C'est cela qui entraîne une spirale de radicalisation.

Pour finir, j'ai quelques recommandations à vous suggérer. Nous souhaiterions qu'une transparence algorithmique soit instaurée et qu'une interface de programmation d'application (API) en temps réel soit mise à la disposition des chercheurs agréés afin d'extraire les données relatives aux contenus haineux.

Nous imaginons un monde où l'algorithme pourrait être limité et où on pourrait pousser du contenu pédagogique aligné avec les valeurs de la République. Les utilisateurs extrêmes pourraient et devraient être bannis et, dans un monde idéal, ne devraient pas pouvoir recréer un compte aussi facilement qu'ils le peuvent aujourd'hui. Le Crif est également engagé sur la question de la création de l'amende civile qui pourrait dissuader les utilisateurs de poster du contenu haineux.

Enfin, s'agissant de l'éducation – car tout passe aussi par là –, il conviendrait de sensibiliser à la lutte contre l'antisémitisme les créateurs dont les comptes dépassent les 50 000 abonnés ainsi que les modérateurs, par la diffusion des modules.

**M. Valentin Petit, journaliste, agence CAPA.** Je travaille pour une émission d'investigation numérique, qui est produite par l'agence CAPA et diffusée sur YouTube et arte.tv.

Pour mener une enquête numérique, nous utilisons divers outils, allant de l'analyse d'images satellites et de gigantesques bases de données à celle de réseaux sociaux, notamment de TikTok.

Au mois de janvier 2025, j'ai découvert, complètement par hasard, d'étranges vidéos diffusées sur TikTok : il s'agissait d'extraits de jeux vidéo. Jusque-là, il n'y avait rien d'anormal. Mais en y regardant de plus près, ces jeux étaient des reconstitutions, au détail près, d'attentats djihadistes ou d'extrême droite ayant été commis. Comme on peut le voir sur ces images, il s'agit, d'un côté, de versions vidéoludiques et, de l'autre, d'images filmées par les tueurs eux-mêmes – les tueries dans la synagogue de Halle et dans une mosquée de Christchurch. J'en ai trouvé d'autres, notamment relatives à l'attentat homophobe du Pulse en Floride et à l'attentat raciste de Buffalo, etc.

Pour vous donner la mesure du contexte, ces comptes ne se contentent pas de diffuser ou de republier des extraits de jeux vidéo, ils publient aussi d'autres types de contenus qui sont tout aussi problématiques, voire davantage. Notamment, ils utilisent les codes de la plateforme TikTok pour pousser des contenus particulièrement mortifères et promouvoir des idéologies particulièrement morbides. Ainsi, ils utilisent des stickers, des filtres, et un ensemble d'éléments disponibles sur la plateforme.

Les *edits* traduisent ce décalage morbide entre l'utilisation des codes de TikTok et les idéologies promues. Il s'agit de petits clips avec des musiques à la mode et des séquences très surcotées, qui visent d'habitude à mettre à l'honneur des personnages ou des stars. Or ils peuvent aussi être utilisés pour glorifier des terroristes et des tueurs de masse. Voici la capture d'écran d'une vidéo mettant en avant le tueur de la synagogue de Halle dans laquelle est incrusté le personnage de Hello Kitty.

Dans les forums d'extrême droite, sont également publiés des mèmes – outil très classique – à savoir des images humoristiques qui reprennent l'iconographie de TikTok, d'internet et de la culture web. Grâce à l'IA, les créateurs créent des contenus dans lesquels ils font danser les terroristes devant les lieux où ils ont commis leurs attentats.

Plus inquiétant encore, on trouve aussi de vraies images déguisées sur ces réseaux. Je pense à un exemple particulièrement révélateur, un contenu violent qui est resté six mois sur la plateforme. Bien que la vidéo ait été supprimée, le compte est encore actif – j'ai vérifié il y a quarante-huit heures. Il s'agit de la vidéo de la tuerie de Buffalo, ce sont les vraies images tournées par le tueur où on voit les personnes s'effondrer. Ce sont de vraies personnes qui ont été tuées. Dans cette vidéo, on voit le tueur tirer une autre balle dans la tête d'une personne. Pour contourner la modération, les personnes qui ont diffusé la vidéo ont ajouté des éléments en lien avec le jeu vidéo Fortnite, par exemple « *Let's play Fortnite* », des stickers, etc.

Nous avons bien entendu analysé les comptes qui diffusaient ce genre de contenus. Ils ne se cachent pas de la modération dont ils n'ont pas peur car ils savent qu'ils peuvent recréer très facilement leur compte. Ils arborent des photos de profil sans équivoque, avec des croix gammées, des dignitaires nazis, des terroristes djihadistes – j'ai vu de nombreux comptes avec la photo de Jihadi John ou d'Omar Mateen, l'auteur de la tuerie du Pulse. Dans leur bio, ils parlent de terroristes comme s'il s'agissait de saints – Saint Brenton pour Brenton Tarrant, le tueur de Christchurch, et Saint Breivik pour Anders Behring Breivik – ou de martyrs, par exemple Jihadi John, le tueur de l'État islamique.

Ce qui est intéressant, c'est que ce type de contenus se rencontre moins fréquemment sur d'autres plateformes. Sur Instagram et Facebook, nous avons trouvé très peu de contenus similaires et la modération intervient assez rapidement. Sur TikTok, c'est effarant : la modération n'est pas du tout à la hauteur du problème. Pire encore, la plateforme nous proposait tous les jours de suivre jusqu'à une centaine de comptes problématiques – @jihadist\_pakistan

par exemple, et un autre où le i de Hitler est remplacé par un l pour échapper à une modération qu'on pourrait qualifier de premier degré.

Sous ce type de contenus, on trouve des commentaires très inquiétants rappelant les attaques au couteau qui ont eu lieu dans un lycée à Nantes ou un collège à Nogent. Par exemple, sous une vidéo représentant, sous forme d'un jeu vidéo, une tuerie, a été publié le commentaire suivant, qui a fait l'objet de 94 likes : « *How my school will look like one day* » – « c'est ce à quoi ressemblera mon école un jour ». Et le créateur du contenu de répondre : « *I hope bro* » – je l'espère, mon frère. Sous une vidéo qui rendait hommage à Brenton Tarrant, on peut lire : « *Need more people like him in this world. I hope the universe gives me the opportunity to do something terrible but great* » – « Nous avons besoin de plus de gens comme lui dans le monde. J'espère que l'univers me donnera l'opportunité de faire quelque chose de terrible mais de grand ». Ce sont deux exemples parmi des centaines et des centaines.

C'est tout un écosystème de radicalisation au sein duquel Tiktok agit comme une vitrine. On pourrait penser que ces contenus sont de simples trolls, des bravades adolescentes un peu subversives avec un humour douteux. Mais non, les commentaires sous les contenus ou la bio de ces comptes TikTok renvoient vers des dizaines et des dizaines de chaînes Telegram dont les contenus sont encore plus violents. On peut trouver notamment les manifestes des tueurs et les vidéos originales qui permettent de créer des contenus qu'on pourra ensuite diffuser sur Tiktok.

Nous nous sommes rendu compte que Tiktok servait à orienter des jeunes vers ces discours et ces contenus. Nous avons pu infiltrer ces différentes chaînes Telegram où des jeunes consommaient ces contenus à outrance, *ad nauseam*. Au cours de cette infiltration – j'ai infiltré plus de vingt chaînes Telegram –, j'ai assisté au passage à l'acte d'un jeune Afro-Américain de 17 ans, qui a attaqué son lycée et tué une personne avec une arme de poing après avoir consommé ces types de contenus. Il connaissait les administrateurs des chaînes Telegram auxquelles j'étais abonné.

J'ai aussi pu interviewer un Hollandais de 17 ans, qui créait ce type de contenus, à la fois des jeux vidéo et des edits qu'il était possible de republier sur Tiktok. Du haut de ses 17 ans, il m'a dit que les *edits*, les memes qu'il publiait *ad nauseam* sur Tiktok, c'était une forme de propagande politique car : « Il n'y a aucune solution politique. La seule solution, c'est la violence ».

En conclusion, voici un meme qui montre l'ampleur de l'absence de modération. Le premier contenu s'intitule « *beauty inside the mosque* » – beauté à l'intérieur d'une mosquée en 2024. Le second, qui a pour titre « *beauty inside the mosque on a good day* » – la beauté à l'intérieur d'une mosquée dans un bon jour –, est une des vidéos du tueur Brenton Tarrant tirant sur des corps sans vie dans la mosquée de Christchurch.

Ces comptes et ces contenus diffusés en masse sont, dans de nombreux cas, créés par des mineurs pour des mineurs. Ils participent à une glorification de la violence ainsi qu'à une forme de radicalisation puisqu'ils rendent complètement normales des formes de violence extrême.

Pour conclure, sur TikTok, j'ai vu des gens mourir, se prendre des balles dans la tête, le tout agrémenté de musiques à la mode, de stickers rose flashy et de filtres roses. L'utilisation de ces codes sur la plateforme, le manque de modération et le type de contenus qui y est poussé témoignent d'un décalage total.

**M. Tristan Boursier.** Mon terrain d'étude se distingue de celui de mes collègues car je m'intéresse à des contenus qui sont moins spectaculaires mais plus populaires. Je me focalise sur des acteurs spécifiques, qui produisent et diffusent des contenus en France, et ce, pour les plus anciens d'entre eux, depuis au moins une dizaine d'années. Vous les connaissez sans doute, ils sont devenus populaires : Papacito, qui, en 2021 a publié la fameuse vidéo où il tirait sur un mannequin grimpé en électeur de La France insoumise ; Matthieu Valet ; Le Raptor ; Le Lapin du Futur, etc.

Je m'intéresse à la quinzaine d'influenceurs les plus populaires sur le site français de YouTube, que j'ai identifiés comme étant d'extrême droite selon une définition académique. Étant donné que ce sont des acteurs complets ayant adopté une stratégie cross-media, mon analyse porte aussi sur les contenus qu'ils diffusent et produisent sur d'autres plateformes, notamment Tiktok et Instagram.

Je rebondis sur ce que disait Valentin Petit sur le rôle de Tiktok dans cet écosystème social médiatique : Tiktok joue vraiment un rôle de vitrine par rapport aux autres plateformes. Les contenus sont beaucoup plus courts et potentiellement plus viraux que ceux publiés sur d'autres médias sociaux, ce qui permet aux influenceurs d'attirer un public qui ne les connaît pas nécessairement.

TikTok soumet directement des contenus aux utilisateurs, sans qu'ils aient forcément le temps d'identifier leur créateur, ni même leur titre. Vous êtes donc directement exposé à des images et à des discours, ce qui laisse plus de chance à des propos plus radicaux, différents de votre idéologie ou de votre vision du monde, de vous convaincre.

**Mme Laure Miller, rapporteure.** Estimez-vous, à la suite des recherches que vous avez menées, qu'il existe une spécificité de TikTok, qui serait notamment liée à son algorithme ou à son design ?

Je vais me faire l'avocate du diable : les responsables de TikTok pourraient répondre qu'il faut, pour trouver les contenus radicaux que vous avez évoqués, les chercher – on ne tomberait pas dessus de façon spontanée. Avez-vous des éléments qui démontreraient qu'un enfant de 11 ou 14 ans peut tomber fréquemment, sans les chercher, sur ce type de contenus ?

**M. Hugo Micheron.** Les deux premiers éléments que je mettrai en avant sont la jeunesse, déjà évoquée, du public de TikTok, plateforme privilégiée des 12-25 ans, et l'abondance du temps qu'ils y passent. La moyenne est, je crois, d'environ quatre heures par jour. J'ai fait un peu de terrain en région parisienne, notamment en Seine-Saint-Denis : quand vous demandez à des enfants, dans une classe, combien de temps ils passent en moyenne sur TikTok, c'est plutôt six ou sept heures par jour. Une enseignante qui fait un travail de conscientisation en classe m'a signalé que certaines réponses étaient même « autant que possible ». Dans les faits, TikTok est une plateforme de socialisation de premier ordre.

Lors d'une étude faite à Sciences Po, dans le 7<sup>e</sup> arrondissement de Paris, qui n'était pas une recherche au sens propre, devant déboucher sur une publication, mais un test destiné à mettre des étudiants en situation, une cinquantaine de comptes vierges, paramétrés avec le moins d'informations possible, ont été créés sur TikTok. La première recherche de contenu a consisté à taper le mot islam, on ne peut plus générique, dans la barre des tâches. Il s'agissait de voir au bout de combien de vidéos on tomberait sur des contenus indiscutablement – j'insiste sur ce point – salafistes, qui font la promotion, par exemple, des auteurs de référence de cette

doctrine radicale de l'islam ou de ses éléments constitutifs, comme le port du voile intégral et de gants par les femmes.

Le résultat n'a jamais été supérieur à cinq vidéos – c'était entre trois et cinq. Quand on prend en considération ce que pèse cette doctrine dans l'islam mondial et de France, il est en soi assez hallucinant de découvrir qu'elle est à ce point présente sur la plateforme privilégiée des jeunes. La visibilité de ces vidéos était très importante, à en juger par le nombre de vues : 300 000 pour une vidéo de jeunes femmes faisant la promotion du corpus salafiste dans une librairie islamique. On ne trouve sur aucune autre plateforme de tels scores – s'ils sont vrais, une autre caractéristique de TikTok étant son opacité. La situation est d'autant plus troublante que la plateforme dispose de moyens, qu'elle utilise, pour déprioriser des contenus en fonction de certains mots-clefs, dont l'usage peut être pénalisant quand il s'agit d'autres causes.

Il y a visiblement un phénomène de viralité sur TikTok, au moins dans ce domaine – mais je crois que les autres présentations étaient tout à fait convergentes. Cette société dit publiquement ne pas avoir la capacité d'agir sur les algorithmes ; on ne pourrait donc rien y faire, et pourtant leur comportement montre clairement l'existence d'une éditorialisation. La logique suivie n'est pas toujours la même que pour les comptes neutres.

**Mme Sophie Taïeb.** J'ai plusieurs comptes TikTok, organisés par thèmes – selon les types de menaces. Je vois bien la puissance de l'algorithme, très supérieure à ce qu'on trouve sur Instagram, Facebook, X ou YouTube. Vous recevez, sous forme de suggestions, du contenu qui vous entraîne dans une spirale et empêche toute pensée critique, puisqu'il va toujours dans le même sens, qu'il s'agisse de comptes en français ou en anglais. Je ne peux donc que rejoindre M. Micheron.

**M. le président Arthur Delaporte.** Je reviens sur la question de la rapporteure, qui portait sur la spécificité de TikTok. En voyez-vous une ?

**Mme Sophie Taïeb.** La spécificité de TikTok, d'après ce que nous voyons, est de diffuser beaucoup plus que toute autre plateforme auprès de la jeunesse, qu'il s'agisse du volume ou du type de contenus. Si on veut comprendre à quoi la jeunesse est exposée, c'est TikTok qu'il faut regarder. Si elle est exposée à trop de désinformation, trop de haine ne faisant pas l'objet d'une modération, des conséquences assez terribles sont malheureusement à craindre.

**M. Valentin Petit.** Un mot sur la méthodologie. La première fois que ce type de contenus m'a été soumis par l'algorithme, c'était sur un compte qui n'y était pas dédié. J'ai ensuite créé un compte vierge pour essayer de creuser la question. Ce qui est assez symptomatique et problématique sur TikTok, même si on le retrouve aussi sur d'autres plateformes, ce sont les comptes qu'on vous propose de suivre toutes les dix vidéos – on peut alors *swiper* pour accéder jusqu'à une vingtaine de comptes. Avec le compte plutôt orienté sur le djihadisme que j'avais créé, je tombais sur des influenceurs de cette tendance et des anachid, c'est-à-dire des chants djihadistes. Pour ce qui est de la sphère d'extrême droite, on me conseillait par vingtaines des comptes problématiques avec des photos de profil de phalangistes, de néonazis, d'accélérationnistes, etc.

**M. Tristan Boursier.** J'irai dans le même sens. Selon des études faites par des équipes de recherche états-uniennes, il faut environ une quinzaine de minutes pour tomber sur des contenus masculinistes radicaux sur TikTok. Mes propres recherches montrent aussi que tout va extrêmement vite s'agissant de l'extrême droite. Pour un des groupes que j'ai infiltrés, tout

a commencé sur TikTok : c'est là que j'ai accédé, par un profil dédié à la recherche, à cette communauté.

**Mme Laure Miller, rapporteure.** Nous avons compris lors de l'audition de responsables de TikTok qu'ils acceptent une espèce de taux d'échec en matière de modération. Nous avons perçu chez eux une absence de volonté de regarder les logiques de contournement utilisées par ceux qui tentent de diffuser des contenus radicaux, comme le recours à des émojis et à des mots-clefs. Ce que nous avons compris en peu de temps, sans être des chercheurs ou des spécialistes mais en travaillant un peu, par exemple, sur les challenges – je sais que ce n'est pas l'objet de vos travaux –, les responsables de TikTok semblaient le découvrir en notre présence. Diriez-vous qu'il n'y a pas forcément grand-chose à attendre de la modération, en particulier de la part de TikTok ? Quels sont les intérêts sous-jacents ? Nos interlocuteurs ont dit qu'il n'y avait pas de question économique derrière, car c'était la publicité qui comptait, mais comment analysez-vous la situation ? Malgré les injonctions de nombreux États, rien n'est fait pour empêcher les contenus radicaux.

**Mme Sophie Taïeb.** Nous espérons que la mise en œuvre du 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 dit *Digital services act* (DSA), notamment le fait que des associations, dont nous faisons partie, ont été nommées signaleurs de confiance, permettra de faire retirer des contenus haineux, même si nous sommes bien conscients qu'agir dans ce sens sur une plateforme où plusieurs millions de vidéos sont téléchargées tous les jours s'apparente à essayer de vider la mer avec une petite cuillère trouée. C'est aussi pour cette raison que nous avons parlé de dissuasion dans nos recommandations, par des amendes civiles et en empêchant la recréation de comptes très radicaux. Il est certain que cela ne va pas dans le sens du modèle économique de la plateforme, mais on s'expose, en n'agissant pas contre la diffusion des discours de haine, à une situation dans laquelle la jeunesse est inondée de contenus non pas problématiques mais radicaux, avec les conséquences auxquelles cela peut conduire.

**M. Valentin Petit.** Je vais vous donner un exemple de manque de sérieux en matière de modération. Le sujet que j'ai réalisé a été retiré de la chaîne d'Arte sur YouTube en France, parce qu'il comportait, notamment, des images de tueries – nous avons pourtant pris toutes les précautions journalistiques d'usage. Certaines plateformes arrivent à modérer les contenus, parfois de manière très sévère, mais TikTok dit ne rien pouvoir faire, parce que des comptes utilisent des codes – ce qu'on appelle le *leetcode* – comme le remplacement de lettres par des chiffres et d'autres formes de contournement telles que l'emploi de stickers. J'ai la prétention d'être né à l'époque du numérique, je suis sur Reddit, Instagram et toutes les autres plateformes : je peux vous dire que ces techniques, qui sont connues depuis vingt ans, sont vraiment le niveau zéro du contournement de la modération.

**M. Hugo Micheron.** Il y a non seulement la question du modèle économique de cette plateforme, mais aussi celle du modèle politique. La modération se fait *a posteriori*. Les algorithmes, quant à eux – cela commence à être très bien documenté –, suivent une logique très spécifique dont on pourrait presque dire qu'elle est éditorialisée. Les mêmes algorithmes font la promotion de contenus très radicaux liés au salafisme et au djihadisme, à l'extrême droite et à d'autres mouvances dites à risque, mais censurent en Chine ce qui concerne la répression des Ouïghours et font la promotion des sciences mathématiques auprès des jeunes. Cela signifie que des gens paramètrent les algorithmes dans un sens ou dans un autre, en fonction de leurs intérêts.

La question du modèle politique doit être posée. Il faut, pour cela, évaluer systématiquement les biais de la plateforme, ce qui ne pourra être fait en multipliant des petites études, comme autant de pièces d'un puzzle. Si les démocraties, comme la France, veulent tenir un discours structurant au sujet de la jeunesse, notamment sur la nécessité d'encadrer les contenus auxquels elle est exposée, pour éviter des dérives violentes et polarisantes, il faut produire des algorithmes de contrôle, de supervision, c'est-à-dire un contre-pouvoir algorithmique qui permette de contrôler systématiquement les biais éditoriaux, *a priori*, de plateformes telles que TikTok. Cela pourrait également s'appliquer, dans le cadre de l'article 40 du DSA, à d'autres algorithmes.

Ces algorithmes, à bien des égards, deviennent des concurrents, des rivaux des enseignants et des éducateurs. D'un côté, on constate une montée des comportements violents chez les adolescents, notamment sous l'influence d'une consommation excessive de contenus sur les réseaux sociaux, comme TikTok, et d'un autre côté on surresponsabilise l'école, alors qu'elle est la première à faire les frais de cette évolution. On pourrait ajouter vingt ou vingt-cinq heures d'éducation civique et citoyenne, mais si les gamins passent six ou sept heures par jour sur TikTok, où ce qui est mis en avant est disproportionnellement du contenu haineux, qu'on cherche précisément à bannir des cours de récréation ou à transcender par des valeurs collectives et universelles, alors on n'arrivera à rien. Il faut donc se poser la question non pas de la modération en aval, mais de la politique qui peut être construite pour traiter ces questions. Pour être tout à fait clair, je suis absolument convaincu qu'on ne réglera pas le problème du djihadisme tant que, dans l'état actuel des choses, on n'aura pas empêché TikTok de fonctionner comme il le fait actuellement.

Quand on regarde la question de la diffusion des idéologies violentes, on voit que des groupes militants, très minoritaires et très radicaux, ont cherché à répandre leurs idées dans leur entourage en passant par des méthodes de prédication classiques. J'ai montré, dans mes travaux, comment le djihadisme s'était construit, en dehors des attentats, par la diffusion d'une idéologie dans des environnements qui n'étaient pas les plus pauvres, ni ceux où il y avait le plus de musulmans en France ou en Europe. Désormais, les machines de prédication sont les algorithmes. Celui de TikTok a, au moins, un biais en la matière, et le djihadisme n'est qu'un exemple de sujets très problématiques pour les sociétés européennes. Une vraie question de fond se pose : ce n'est pas un détail du débat public, mais le cœur des enjeux pour les dix prochaines années en France.

**M. Kévin Mauvieux (RN).** Mme la rapporteure a évoqué les challenges – ces défis entre utilisateurs qui présentent des risques importants pour les jeunes, selon de précédentes auditions –, en soulignant d'emblée que ce n'était pas la question du jour. C'est vrai, mais M. Boursier a parlé, dans son propos liminaire, de l'adaptation des radicaux aux modes. Comment analysez-vous leur adaptation à la mode des challenges ? Ces derniers sont-ils utilisés pour diffuser auprès des jeunes, d'une manière plus subtile ou plus ludique, en quelque sorte, des idées radicales, qu'elles soient antisémites, islamistes ou d'extrême droite ? Est-ce un des dangers que vous avez identifiés ?

Mme Taïeb a dit avoir détecté que l'algorithme de TikTok était beaucoup plus puissant que ceux d'autres réseaux sociaux, ce qui est un fait indéniable. Au-delà de cette question, constatez-vous, vous qui faites et recevez des signalements, que ceux-ci sont beaucoup plus nombreux proportionnellement en ce qui concerne TikTok que s'agissant d'autres réseaux sociaux et sont-ils plus graves ou, en tout cas, portent-ils sur des contenus d'une plus grande virulence sur TikTok que sur d'autres plateformes ?

Nous n'avons pas du tout parlé d'un élément qui doit pourtant jouer une place prépondérante dans la diffusion des idées radicales – mais vous me contredirez peut-être –, à savoir les lives, qui me semblent beaucoup moins contrôlables que les contenus classiques. Quelle est votre analyse de la diffusion de contenus extrêmes, radicaux au moyen de lives ?

Avez-vous constaté lors de vos recherches ou enquêtes une diffusion de contenus radicaux, en live ou non, par des mineurs ou mettant en scène des enfants ? Il est apparu lors de précédentes auditions que certains influenceurs ou certaines personnes qui souhaitent le devenir utilisaient leurs enfants sur TikTok. Avez-vous vu des contenus mettant en scène un ou des enfants et ayant une visée radicale ?

J'en viens à une dernière question, peut-être provocatrice ou radicale, mais qui fait écho à des propos qui ont été tenus devant nous – je crois que c'était lors de l'audition de parents d'enfants victimes de scarifications ou ayant fait une tentative de suicide. Un père nous a dit qu'il se demandait si la Chine, par l'utilisation de TikTok, n'essayait pas d'affaiblir nos populations pour obtenir une supériorité sur les pays occidentaux en abrutissant, en quelque sorte, leurs nouvelles générations. J'y pense car l'un de vous a déclaré qu'en Chine, et c'est tout à fait vrai, TikTok sait très bien contrôler les contenus diffusés. Quand nous auditionnons les responsables de cette plateforme en France, ils nous regardent un peu en chiens de faïence et disent qu'ils ne comprennent pas et ne savent pas faire. Serait-il possible que ce père ait raison ? Une puissance étrangère propriétaire de TikTok serait-elle désireuse de laisser passer certains contenus pour affaiblir les futures générations dans nos pays et développer une supériorité, technique, technologique ou intellectuelle, par rapport à nous dans les années à venir ?

**M. Tristan Boursier.** S'agissant de l'adaptation aux modes, je vais vous décevoir car sur mon terrain, les influenceurs d'extrême droite et masculinistes, les challenges n'occupent vraiment pas une place importante – surtout à l'extrême droite.

Quand je parlais d'adaptation aux modes, je pensais plus largement à la stratégie métapolitique qui est employée pour créer des contenus : on se situe en amont de la politique institutionnelle, à un niveau extraparlémentaire, pour mobiliser des éléments culturels déjà présents, déjà populaires afin de diffuser des idées. On trouve ainsi beaucoup de références à des mangas chez les influenceurs d'une trentaine d'années : les personnes nées dans les années 1990 ont été des consommateurs importants de mangas, et la France est un des pays européens où les jeunes continuent à en lire beaucoup. Ces influenceurs utilisent, par exemple, une référence aux dragons célestes, dont il a beaucoup été question il y a quelques mois, pour faire passer des messages antisémites. On trouve aussi des références au nuage de *One Piece*. D'autres références sont liées à la culture populaire : des extraits du film *Les Visiteurs* évoquent ainsi l'homme viril, le guerrier ou le chevalier d'un Moyen-Âge fantasmé.

Second volet de l'adaptation aux modes, j'ai constaté une évolution des formats. L'usage du podcast semble vraiment exploser, suivant une tendance d'origine états-unienne que j'explique notamment par son modèle économique. Le recours au podcast permet de produire un contenu qui peut être monétisé, notamment sur Spotify, et scindé en séquences un peu plus choc qui sont diffusées sur TikTok et d'autres plateformes. On peut ainsi faire une seule prise, d'un format assez long, qui est diffusée sur des médias sociaux différents et monétisée de diverses manières.

Même si ce n'était pas le centre de mes recherches, j'ai assisté à différents lives : ils offrent beaucoup plus de liberté et donnent lieu à beaucoup plus de dérives, car le contrôle est bien moindre – la modération, déjà faible en général, doit être quasi nulle s'agissant des lives.

Je n'ai pas vu de mineurs mis en scène dans ce cadre, mais le public peut interpellé les créateurs de contenu sur des plateformes telles que YouTube, qui est moins utilisée pour les *lives*, ou Twitch, qui l'est davantage, et je suppose, compte tenu des interactions que j'ai observées, qu'il y avait dans le public beaucoup de mineurs – ils disaient être encore à l'école ou au collège. Des influenceurs que j'identifie comme étant d'extrême droite mais qui se positionnent de plus en plus sur des thématiques antiféministes et masculinistes utilisent leur stature et leur crédibilité pour donner ce qu'ils appellent des conseils de drague – ce sont en fait des conseils de harcèlement, notamment envers les jeunes filles. C'est là que j'ai vu une interaction vraiment concrète avec des mineurs.

Pour le reste, un influenceur qui est récemment devenu père, M. Julien Rochedy, met en scène son enfant sur les médias sociaux, notamment Instagram. Il reprend, ce qui est assez intéressant, l'esthétique des *tradwives*, les épouses traditionnelles. Au moyen d'une esthétique qui a l'air complètement dépolitisée, ce courant états-unien promeut une façon de vivre conforme à des valeurs conservatrices ou réactionnaires. L'influenceur en question est par ailleurs connu pour des contenus qui, eux, sont complètement politiques. Il se présente même sur Twitter comme un penseur d'extrême droite – il n'y a aucune ambiguïté sur ce point.

En ce qui concerne le rôle de la Chine, je n'ai pas grand-chose à dire, au-delà de la remarque suivante : si TikTok est effectivement un réseau social chinois, la plupart des médias sociaux, qui forment un écosystème avec TikTok, sont états-uniens – la Chine n'est donc pas le seul pays qu'il faut mettre dans la balance.

**Mme Sophie Taïeb.** S'agissant des signalements que nous faisons, je soulignerai surtout la viralité des contenus. Un tweet qui a 120 ou 200 vues n'est rien par rapport à une vidéo qui en a des millions sur TikTok. L'ampleur est complètement différente et les délais de retrait font que les contenus concernés sont bien trop vus. Il y a du contenu violent partout, sur X, sur Facebook ou sur Telegram ; la différence, c'est le nombre de vues sur TikTok. On ne connaît pas les règles internes des plateformes, mais elles sont censées avoir une sorte de règle des *strikes* : si vos contenus sont retirés plusieurs fois, à la fin votre compte est suspendu. Comme je ne connais pas la règle, je ne peux vous donner que mon ressenti : j'ai l'impression qu'Instagram suspend plus facilement que TikTok des comptes qui franchissent à répétition la ligne.

Je vais vous donner un exemple. Lorsqu'un couple qui se rendait à une soirée dans un musée de Washington D.C. a été assassiné, il y a quelques semaines, Guy Christensen, un américain qui a des centaines de milliers de *followers*, a d'abord publié une vidéo dans laquelle il disait condamner toute attaque contre des civils, puis il a publié une autre vidéo dans laquelle il expliquait, en gros, qu'il s'agissait d'agents sionistes, du Mossad, et que c'était bien fait pour eux car ils l'avaient bien cherché. Il a fallu plusieurs jours pour que cette vidéo soit retirée, mais elle est partout tant elle a été partagée, y compris sur d'autres plateformes, et son auteur a toujours son compte. Il apparaît même tout le temps dans mon *feed*, à cause de l'algorithme, ce qui m'inquiète.

En ce qui concerne les *lives*, il est quasiment impossible de signaler du contenu parce qu'il n'est pas pérenne. Grâce à notre accès professionnel, nous pouvons envoyer à Pharos des captures d'écran – comme celle que je vous ai montrée tout à l'heure, où il était dit qu'il fallait « attaquer les synagogues jusqu'à la mort » – mais eux ne parviennent pas à retrouver le contenu problématique ni *a fortiori* à obtenir sa suppression. Pourtant ledit contenu a bien été vu. Nous avons le même problème avec la fonctionnalité *Spaces* sur X. Il est évidemment hors de question d'enregistrer des heures de *live*. Faute de pouvoir être signalés comme un post, un

commentaire ou un compte, les *lives* sont dans une zone grise. C'est une difficulté et un autre chantier à ouvrir.

S'agissant des différences entre le TikTok chinois et le TikTok que nous connaissons, je ne fais pas de politique mais il est vrai que la première plateforme pousse des contenus éducatifs et limite le temps d'utilisation quand la seconde laisse diffuser des contenus complètement en roue libre, avec les conséquences que nous évoquons depuis tout à l'heure.

**M. Valentin Petit.** Tout TikTok, et l'algorithme au premier chef, est basé sur la mode. Les vidéos sont accompagnées de musiques – anachid chez les djihadistes, hardtek ou pop pour les trolls d'extrême droite – et de *hashtags*, qui ont vocation à pousser des messages. Le montage, effectué grâce aux applications Capcut ou Viggie AI notamment, cette dernière permettant de donner vie à des images, joue aussi un rôle pour créer des phénomènes viraux, qui seront ensuite poussés par la plateforme. Les acteurs radicaux, que ce soient les djihadistes ou l'extrême droite accélérationniste et néonazie, utilisent les codes de la mode pour pousser leurs contenus sur la plateforme. Des contenus auparavant dans les marges, qui apparaissaient dans des forums obscurs réunissant 300 personnes, tel que 4chan, sont désormais vus par 300 000 personnes.

Si l'algorithme est capable de pousser ces contenus viraux, qui associent *hashtags*, musiques, challenges, etc., il est tout aussi capable de les détecter. La modération pourrait donc facilement être plus active.

**M. Hugo Micheron.** Les *lives* sont beaucoup moins faciles à contrôler. L'une des spécificités de TikTok tient au volume des flux vidéo – c'est le principe même de cette plateforme. Il est donc très dur de retrouver une vidéo vieille d'un mois. Cela limite considérablement les possibilités d'analyse *a posteriori*. C'est aussi ce qui rend la plateforme attrayante aux yeux des jeunes et qui rend très difficile de comprendre ce qui s'y joue.

En ce qui concerne la dimension géopolitique, cela mériterait évidemment une étude approfondie. Ce qui est sûr, c'est que la France, et plus généralement l'Europe, n'est pas du tout souveraine pour exercer un contrôle sur l'espace premier de socialisation de sa jeunesse, qui vient y communier plusieurs heures par jour et ce faisant, est exposée à des contenus dont on ignore tout des conditions de production. Il est question rien de moins que de l'avenir du pays.

Ce que nous savons avec certitude, c'est qu'il existe un alignement, un énorme biais : l'algorithme promeut certains contenus et en disqualifie d'autres, en se fondant sur leur forme – musique, longueur, couleur, etc. – mais aussi sur le fond. C'est ce dernier aspect qu'il faut documenter, et ce n'est pas simple.

Il faut cesser de croire que TikTok est un espace public au motif que tout le monde peut y avoir accès gratuitement et y passer des heures sans avoir l'impression de consommer quoi que ce soit. On ignore tout des ombres dans le décor. C'est là tout le problème. Il y a d'un côté le fonctionnement démocratique, très transparent, qui repose sur la confiance, et de l'autre, une boîte noire qui s'appelle TikTok, qui amuse la galerie, notamment les plus jeunes, et à laquelle on accorde une confiance aveugle.

Il est important de le comprendre si l'on veut agir sur le fond : le fait de demander un contrôle des contenus promus sur TikTok n'est pas attentatoire à la liberté d'expression. Tracer un chemin dans une jungle informationnelle pour pouvoir s'y repérer est bon pour la

démocratie. Les démocraties se sont construites contre d'autres jungles informationnelles : au XIX<sup>e</sup> siècle, il était possible de dire tout et n'importe quoi dans des torchons qui s'appelaient des journaux et qui étaient financés par certaines personnes pour défendre leurs intérêts, lancer des modes et calomnier leurs opposants. À un moment, des lois sont venues encadrer la presse, en imposant notamment une vérification des sources ; une industrie s'est développée et tout d'un coup, la démocratie a été mieux informée. S'agissant des réseaux sociaux, nous sommes encore à l'étape antérieure. La démocratie ne pourra pas s'en sortir, du moins sans violence, si elle ne s'attaque pas frontalement à ces questions et TikTok est un exemple hors pair pour cela.

**M. Thierry Sother (SOC).** Vous nous livrez des éclairages très inquiétants.

Notre assemblée a adopté il y a quinze jours une résolution appelant la Commission européenne à faire pleinement respecter le DSA face aux ingérences étrangères. Monsieur Micheron, vous avez rappelé l'influence de X et de son propriétaire ainsi que les récents événements en Roumanie. Lorsque nous l'avons interrogée, la représentante de la plateforme X nous a répondu qu'à sa connaissance, il était naturellement impossible que les contenus de son propriétaire soient mis en avant, en dépit des études qui démontrent le contraire. C'est évidemment pour nous une source d'inquiétude importante.

Je reviens aux contenus toxiques et radicaux dont M. David Chavalarias, dans ses études sur X, estime la part très importante – ce doit être analogue sur TikTok. Pourquoi l'algorithme – ce que vous appelez les ombres et moi les amis invisibles – choisit-il de les mettre en avant ? Est-ce une application de l'économie de l'attention, en vertu de laquelle plus les utilisateurs restent captifs longtemps, plus les plateformes peuvent monétiser les espaces publicitaires ? Ou y a-t-il encore d'autres intérêts cachés ?

**M. Hugo Micheron.** De nombreuses enquêtes montrent – et M. Elon Musk s'en est même vanté publiquement – que l'algorithme de X est programmé pour favoriser les *tweets* de son propriétaire. Si l'algorithme de Twitter est entre les mains d'une personne, qui peut décider de faire prévaloir sa parole sur toutes les autres, on imagine ce que cela peut donner sur d'autres sujets.

La parole de la représentante de X est importante mais l'avenir de la démocratie ne devrait pas en dépendre. La situation est évidemment anxiogène, mais les possibilités de reprendre le contrôle sont nombreuses. Je ne pense pas que le renforcement de l'éducation aux médias ou le triplement des heures d'éducation civique soit la solution – cela reviendrait à écoper la mer avec une petite cuillère. En revanche, il faut travailler sur les algorithmes, et le DSA nous offre le cadre juridique pour cela. L'article 40, que l'on doit à une plume française, est bien pensé : par la référence qu'il comporte aux risques systémiques, les plateformes sont comptables du contenu toxique qui vient polluer les places algorithmiques auxquelles tout le monde vient s'abreuver. Nous devons prendre à bras-le-corps la question de la souveraineté algorithmique, être capables de contrôler les algorithmes et d'en démontrer systématiquement les biais. Il serait étonnant que dans un pays comme la France, on ne s'accorde pas au moins sur le fait qu'un contenu similaire à 90 % à la propagande de Goebbels ou de l'État islamique ne soit pas propulsé ou mis en avant. La quasi-totalité de la population y serait favorable.

Avant même de s'interroger sur une éventuelle intention maligne ou de quelconques intérêts, sur lesquels de nombreux travaux de recherche restent à mener, l'urgence commande de comprendre car c'est l'avenir de la politique qui est en jeu. Les représentants du peuple ne peuvent pas être à la merci d'une saillie sur les réseaux sociaux, qui vient ruiner tout leur travail

ou casser leur dynamique de campagne. Si l'on ne contrôle rien et si l'on se contente de subir, la question de la politique et de sa représentation va se poser à tous les niveaux.

Nous sommes au début d'une transformation. La France a toutes les cartes en main pour agir et apporter une réponse démocratique à ces logiques plutôt autoritaires, je vous l'accorde.

**Mme Sophie Taïeb.** Nos recherches montrent que les acteurs hostiles qui postent des contenus toxiques maîtrisent parfaitement les codes de la plateforme. Ils les ont étudiés et savent exactement quoi dire et comment pour pouvoir être mis en avant. Ils sont parfaitement au fait de ce qui est viral. Ce sont vraiment des professionnels.

**M. Tristan Boursier.** Tant que la haine sera rentable sur les plateformes, nous serons confrontés à ce genre de problème. Le concept d'économie de l'attention nous aide à éclairer une bonne partie de la question ainsi que les enjeux de souveraineté.

Puisque je suis les mêmes acteurs depuis cinq ans environ, j'ai pu observer des évolutions stratégiques visant à monétiser du contenu. En ce moment, au sein de l'extrême droite, la mode est à l'antiféminisme parce que les propos de cette nature semblent beaucoup moins sujets à la régulation que d'autres. On constate une percée des propos antiféministes chez des acteurs traditionnellement d'extrême droite. Je pense à Thaïs d'Escufon, une militante de Génération identitaire, qui s'est lancée en ligne avec ce mouvement puis, il y a quelques années, a fait un virage masculiniste pour prodiguer des conseils à des jeunes hommes sur la manière de draguer les femmes dans la rue ou de trouver une épouse. Récemment, elle a pris un nouveau virage après avoir constaté son échec – elle ne le dit pas en ces termes-là, c'est mon interprétation – afin de se lancer dans de nouvelles activités. Sa réorientation aboutit à une association avec un autre influenceur du nom de Valek, qui lance le Valek Studio, une sorte de service-communauté dans lequel il vend des prestations. Pour l'instant, ce n'est pas très clair ; on comprend que le service est monétisé et vise à capter la petite frange de leur communauté qui est prête à les suivre en s'engageant financièrement davantage.

**M. le président Arthur Delaporte.** Pouvez-vous nous en dire un peu plus sur ce modèle économique ?

**M. Tristan Boursier.** Les contenus les plus problématiques, en tout cas ceux qui font le plus réagir, qui misent sur l'outrance, créent de l'engagement, ce qui favorise leur visibilité.

Certains ont néanmoins compris qu'ils ne pouvaient pas fonder leur modèle économique sur les seuls médias sociaux. Des recherches, menées notamment en Allemagne, montrent très bien la diversité des sources économiques des influenceurs. Cette diversité concerne bien sûr les réseaux sociaux sur lesquels ils sont présents, mais aussi d'autres supports tels que les caisses de soutien sur les plateformes de *crowdfunding* ou la création de communautés – les dix plus gros influenceurs français ont tenté ou tentent encore de le faire –, par le biais desquelles ils vont pouvoir capter de l'argent de la partie la plus crédule et manipulable de leurs fans. Ils peuvent ainsi simplement vendre un accès à un espace en ligne fermé. Par exemple, sur Discord, en payant un abonnement régulier ou une somme mirobolante, on peut entrer dans un espace où l'on trouve des services plus ou moins bien structurés, comme des formations aux sports de combat, aux cryptomonnaies ou à l'investissement immobilier. Ces services sortent du cadre politique mais ils sont utilisés dans une stratégie métapolitique. On a affaire à des businessmen, mais des businessmen qui surfent sur une vision du monde

assez spécifique. Il faut bien avoir en tête ces deux facettes pour comprendre leur activité et leur succès.

**M. le président Arthur Delaporte.** Si je comprends bien, la militante dont vous parliez s'associe pour vendre ce type de formations.

**M. Tristan Boursier.** Elle est ambassadrice du Valek Studio – je suppose que cela passe par une association économique contractuelle – à l'intérieur duquel on trouve apparemment aussi des formations au montage vidéo. Je ne sais pas quel est l'objectif recherché mais on peut supposer qu'il s'agit de former d'autres influenceurs.

Valek a également diffusé dans les médias sociaux des contenus qui reprennent les discours masculinistes. On peut y voir un syncrétisme de différentes tendances réactionnaires. Christine Bard parle d'intersectionnalité des haines dans d'autres contextes. Je reprends volontiers à mon compte l'expression pour le monde numérique.

**M. le président Arthur Delaporte.** J'ai observé la reprise par certains influenceurs ou responsables politiques d'extrême droite de sujets très *mainstream* dont on sait qu'ils marchent, comme l'alimentation. Parmi les rares contenus que j'ai diffusés sur TikTok, ceux qui ont eu le plus de succès sont ceux liés à la bouffe. La vidéo où l'on voit le président de la République boire des verres a cartonné ; M. Bardella passe son temps à se mettre en scène avec du pain. Est-ce une tendance que vous avez également constatée ?

**M. Tristan Boursier.** Tout à fait, et c'est vraiment intéressant parce qu'il est question d'éléments culturels de notre société – un nombre important de Français consomment du vin, par exemple, qui, jusqu'à présent, n'était pas un objet politique. Or le fait de boire du vin ou de manger de la viande est devenu un sujet dans le débat public. Il a d'ailleurs été lié à différents camps politiques. Selon que vous êtes pour ou contre le carnisme, vous pouvez être associé à une tendance politique ou à une idéologie.

M. Baptiste Marchais, un influenceur que j'étudie, a beaucoup joué sur cette tendance en organisant ce qu'il appelle les « repas de seigneur ». Il s'agit d'un format vidéo assez long – environ une heure voire plusieurs heures –, parfois découpé en plusieurs publications, dans lequel il invite des amis, issus de l'extrême droitosphère – Jean Lassalle a aussi été convié dans l'un des épisodes –, à des repas gargantuesques où la viande est vraiment mise en avant. Les convives font beaucoup de plaisanteries sur la viande ; ils disent que les députés écologistes feraient des syncopes en voyant les tablées qui s'empiffrent.

Les propos ne sont pas forcément haineux ou radicaux en tant que tels, mais des choses aussi banales que la consommation de viande sont mises au service d'un appel au drapeau de leur communauté ou servent à interpeller des gens, qui, probablement, ne se reconnaissent pas dans les idéaux d'extrême droite mais qui, par le biais de la viande, peuvent éprouver une sympathie à l'égard des influenceurs.

**M. Hugo Micheron.** C'est un point clé, commun aux autres réseaux sociaux. Les groupes extrémistes sont très roués à la banalisation. On parle de groupes radicaux extrémistes idéologisés, mais ils se présentent autour d'une blague, d'une situation de la vie quotidienne – les djihadistes en Syrie se mettaient en scène avec des petits chatons qu'ils nourrissaient – dont le sous-titre est : « ils disent qu'on est ceci ou cela alors qu'en fait, on est comme vous, on fait les mêmes choses que vous, on aime les mêmes choses que vous, on rigole aux mêmes

choses que vous ». C'est le principe élémentaire de la propagande : banaliser une idéologie et la faire passer pour totalement acceptable et ordinaire.

La réunion, suspendue à dix-huit heures trente, est reprise aussitôt.

Présidence de Mme Josiane Corneloup, vice-présidente de la commission

**Mme Josiane Corneloup, présidente.** L'un d'entre vous a évoqué une question de fond pour les sociétés européennes, et j'ajoute, pour l'avenir de nos démocraties. Du fait des codes utilisés, il est très difficile de repérer les messages de haine, qui polluent complètement notre jeunesse. Dans vos fonctions respectives, vous en débusquez un certain nombre mais il en reste probablement d'autres. Plutôt que d'agir sur les conséquences et de chercher à faire retirer les contenus, il serait plus opportun d'essayer d'éviter qu'ils soient publiés. Que préconisez-vous pour y parvenir ?

**Mme Laure Miller, rapporteure.** Lors de son audition, Laurent Marcangeli a parlé du DSA comme d'une ode à l'impuissance publique. Il faisait allusion au fait qu'aucun décret d'application de la loi visant à établir une majorité numérique à 15 ans, dont il était à l'initiative, n'a été pris pour ne pas être en contradiction avec le droit de l'Union européenne.

Monsieur Micheron, vous avez dit que la France pouvait déjà faire beaucoup. Pouvez-vous préciser votre pensée ?

**M. Hugo Micheron.** J'ai une appréciation différente sur le DSA.

Je ne pense pas qu'il soit difficile d'intervenir sur le contenu. Actuellement, la modération s'opère par mot-clé. C'est la raison pour laquelle vous pouvez y échapper en remplaçant le « e » dans le mot Hitler par un « 3 » : le robot est dupé mais n'importe quel humain va lire Hitler car l'effet visuel est identique. Il est tout à fait possible, grâce à l'intelligence artificielle, de mettre en place une modération basée sur la sémantique dès lors qu'on a répondu par la négative à ces questions : tolère-t-on un propos qui emprunte à 90 % sa sémantique à tel groupe extrémiste et tolère-t-on la visibilité disproportionnée dont bénéficie sur les réseaux sociaux comme nulle part ailleurs ?

On peut tout à fait nourrir l'IA – l'*embedder* – avec la sémantique de groupes donnés pour pouvoir ensuite la détecter automatiquement – nous le faisons à Sciences Po. Des vidéos qui lancent des appels à la violence atroce peuvent être repérées. Les grandes compagnies de technologies, qu'elles soient chinoises, américaines ou russes, ont les moyens de détecter la sémantique et de la signaler. Si elles ne l'ont pas encore fait, elles ont les moyens de développer les procédés nécessaires.

Le DSA est probablement incomplet mais il est intéressant car, pour la première fois, un cadre juridique dit clairement qu'un contenu illégal n'a rien à faire sur les réseaux sociaux – encore heureux ! Par ailleurs, il existe des groupes très habiles, qui se placent en dessous de la ligne de détection du radar légal, de telle sorte que le contenu, bien que très problématique, n'est pas considéré comme illégal. Le contenu antisémite est un très bon exemple : il y a plein de manières d'être juste en dessous du seuil de détection légale et pourtant de produire un

discours totalement toxique. On peut faire en sorte que la machine soit capable de détecter ces stratagèmes.

Si je devais faire une seule recommandation, ce serait celle-là. Nous avons en France tous les ingénieurs tech qu'il faut pour nous outiller et penser une démocratie numérique costaude, capable de résister à du contenu haineux en ligne.

**Mme Sophie Taïeb.** Il y a en effet une zone grise pour ces contenus qui sont sous la ligne légale mais évidemment problématiques. S'il était possible de retirer ce qui est évident, on aurait bien avancé.

Je souhaiterais revenir sur l'amende civile que j'ai mentionnée parmi nos recommandations. De la même manière que l'installation de radars et le port obligatoire de la ceinture de sécurité, malgré les protestations initiales, ont permis de faire diminuer le nombre d'accidents de la route, nous pensons qu'il serait utile de sanctionner les contenus indiscutablement hors-la-loi ou dangereux par une amende. On ne peut pas laisser dire tout et n'importe quoi sur les réseaux sociaux en toute impunité. Peut-être les gens continueraient-ils à le penser mais peut-être arrêteraient-ils de les diffuser et de les publier sur les plateformes.

**M. Valentin Petit.** Il n'est pas difficile de repérer les codes, au contraire : c'est très simple. En une semaine, on a compris de quoi il retourne. Quand par exemple l'emoji « jus » est utilisé – parce qu'en anglais le mot *juice* est proche du mot qui signifie « Juifs » –, accompagné d'un emoji « gaz », on comprend très vite de quoi il est question.

Les codes qui sont utilisés, souvent à destination des jeunes, voire très jeunes pour créer un phénomène d'*insider* – en leur faisant des clins d'œil, en utilisant des memes et l'IA pour faire danser Anders Breivik devant l'île d'Utøya –, sont loin d'être impossibles à déchiffrer. C'est assez simple et frontal : ces codes ne sont pas impossibles ni même difficiles à déchiffrer.

Il faut évidemment utiliser l'analyse sémantique et l'intelligence artificielle mais je plaide aussi pour un investissement dans le travail humain. Les plateformes sont très peu désireuses de faire appel à des êtres humains pour opérer la modération des contenus. Elles font essentiellement confiance à l'automatisation par mots-clés, notamment TikTok. Mais il y a très peu de personnes chargées d'analyser la sémantique et d'alimenter une IA pour améliorer l'automatisation. Le manque d'outils d'analyse humaine et la prépondérance des outils d'analyse automatiques sur ces plateformes est très problématique.

**Mme Sophie Taïeb.** Nous fournissons tous les ans à l'ensemble des plateformes avec lesquelles nous travaillons ainsi qu'à Pharos, à la Dilcrah et à d'autres une liste de termes et d'emojis antisémites – le jus, le nez, les savonnettes, etc. – pour qu'ils puissent informer leurs modérateurs. On ne sait pas si tous le font mais tous reçoivent la liste.

**Mme Josiane Corneloup, présidente.** Avez-vous l'impression que c'est suivi d'effet ?

**Mme Sophie Taïeb.** Pas forcément. Le plus emblématique d'entre ces termes, « dragon céleste », est désormais connu. Un travail de réflexion a été mené avec Meta sur l'utilisation du terme « sioniste » : quand ce mot a-t-il une connotation politique, et quand devient-il une insulte antisémite ? Si je traite un rabbin de « sale sioniste », suis-je en train de l'insulter en le traitant de sale juif, ou bien est-ce que je m'exprime sur ses prises de position

politiques ? Nous avons travaillé avec eux pour différencier le contenu insultant du contenu politique.

L'antisémitisme est une forme de haine compliquée. La pédopornographie, c'en est ou c'en n'est pas ; le terrorisme, c'en est ou c'en n'est pas. En revanche, l'antisémitisme comporte une zone grise assez importante ; c'est une menace qui évolue, parce que de nouveaux termes, de nouvelles tendances émergent, parce qu'il se passe des choses dans le monde. C'est pour cela que nous travaillons avec les plateformes : nous essayons de les informer autant que possible sur les nouvelles tendances. Nos efforts sont parfois suivis d'effet, et parfois un peu moins.

**M. Kévin Mauvieux (RN).** Madame Taïeb, vous avez indiqué à plusieurs reprises qu'il fallait créer des amendes civiles pour sanctionner l'antisémitisme sur les réseaux sociaux. Je ne suis même pas sûr que ce soit la solution, car l'antisémitisme est déjà interdit et répréhensible. Nous devons nous interroger sur l'existence d'un droit parallèle pour les réseaux sociaux par rapport à la vie de tous les jours. Une personne qui tient des propos ouvertement antisémites dans la rue sera arrêtée, placée en garde à vue, auditionnée et condamnée. Sur les réseaux sociaux, il faut se battre pendant quinze jours pour retirer une vidéo, et cela s'arrête là, alors que le délit est le même que celui commis dans la rue. Il faudrait trouver la solution pour que les peines existantes s'appliquent aussi aux réseaux sociaux.

La monétisation de leurs contenus par les influenceurs radicaux est problématique. Il est en effet possible de payer des publicités sur TikTok pour sponsoriser ses contenus, les mettre en avant et les faire apparaître de manière encore plus extravagante. Qu'Adidas ou McDonald's se paient une publicité, c'est leur droit et cela ne dérange personne. Mais qu'en est-il pour les influenceurs radicaux ? Avez-vous déjà effectué des recherches sur la modération et sur le contrôle des publicités ?

Je prendrai un exemple simple : les responsables politiques n'ont pas le droit de monétiser le contenu de leurs réseaux sociaux. Je dispose d'un compte sur TikTok qui, très régulièrement, m'envoie des notifications pour m'inciter à le faire. Étant membre de cette commission d'enquête, j'ai fait le test – sans avoir l'intention d'aller jusqu'au bout puisque cela m'aurait mis hors la loi : TikTok m'a bloqué en me rappelant qu'un responsable politique n'avait pas le droit de monétiser ses contenus.

Un système équivalent existe-t-il pour les influenceurs radicaux ? Comment sont contrôlées les publicités d'un utilisateur de TikTok ? Je ne parle pas des entreprises mais bien des utilisateurs lambda, qui n'ont pas forcément de très gros comptes mais créent beaucoup de petits comptes. Ont-ils accès à du *sponsoring* et, si oui, sont-ils contrôlés ?

**Mme Sophie Taïeb.** Concernant les amendes civiles, c'est l'absence de dissuasion qui nous dérange. Si, avant de poster, on se dit « soit je vais payer une amende, soit je vais être poursuivi en justice et cela sera suivi d'effets », alors les gens arrêteront de poster. Nous sommes donc d'accord.

Je n'ai pas étudié en revanche la question des publicités.

**M. Tristan Boursier.** Votre question est très pertinente car elle permet de souligner la difficulté que rencontrent les chercheurs à comprendre ces mécanismes, notamment les relations que ces acteurs ont avec les régies publicitaires. Ils se plaignent souvent de la démonétisation de leurs contenus et du fait que les plateformes les empêchent de faire de

l'argent. Or cette affirmation ne repose que sur leur parole ; les plateformes ne collaborent pas pour nous transmettre des informations.

Néanmoins, quand on regarde une vidéo, par exemple sur YouTube, on peut voir si de la publicité apparaît ou non. De plus, même si les plateformes démonétisent une partie des contenus, ce risque est intégré par les plus gros influenceurs, qui ont diversifié leurs sources de revenus. Un écosystème de *sponsoring* s'est mis en place, avec notamment le fonds Périclès, financé par Pierre-Édouard Stérin. Ce fonds soutient la marque Terre de France, qui sponsorise nombre de vidéos d'influenceurs, notamment Valek et Le Raptor. D'autres entreprises le font aussi, comme Kalos, dont les liens avec Périclès ne sont pas clairs pour moi. Cet écosystème permet de soutenir l'activité économique des influenceurs en dehors des canaux de monétisation classiques des plateformes.

**M. Valentin Petit.** Les petits comptes n'ont pas besoin de diffuser des publicités : le fonctionnement opaque de la plateforme et de l'algorithme se charge de pousser les contenus problématiques. Ainsi, un compte de moins de 300 abonnés pourra percer et atteindre 300 000 vues grâce à une vidéo glorifiant par exemple un tueur djihadiste ayant fait cinquante morts. C'est un phénomène que l'on voit rarement sur d'autres plateformes, mais qui est très fréquent sur TikTok.

**Mme Josiane Corneloup, présidente.** Ces vues par milliers sont-elles réelles ou artificielles ? Si elles sont réelles, cela signifie qu'il y a un vrai public.

**M. Valentin Petit.** Les contenus sur TikTok étant très courts, les vues commencent très rapidement. Cela ne signifie pas que 300 000 personnes ont regardé la vidéo en entier, mais que ce contenu a été conseillé à 300 000 personnes : la nuance est importante.

Ainsi, lors de l'élection présidentielle roumaine, la plateforme a favorisé le candidat Călin Georgescu de façon absolument astronomique, en utilisant notamment des comptes dormants, qui publiaient habituellement du contenu footballistique sur Neymar ou sur Messi et qui se sont mis d'un seul coup à publier des vidéos pro-Georgescu qui atteignaient 400 000 ou 500 000 vues – l'algorithme avait en fait conseillé ces vidéos à toutes ces personnes. L'opacité de l'algorithme repose aussi sur l'opacité des comptes qui poussent ce genre de contenus.

**Mme Josiane Corneloup, présidente.** Je vous remercie pour votre présence et pour les précisions fort utiles que vous avez apportées à nos travaux.

*La séance s'achève à dix-neuf heures.*

**Membres présents ou excusés**

*Présents.* – Mme Josiane Corneloup, M. Arthur Delaporte, M. Kévin Mauvieux, Mme Laure Miller, Mme Constance de Pélichy, M. Thierry Sother, M. Stéphane Vojetta