SÉCURITÉ ET AUTONOMISATION EN LIGNE DES UTILISATEURS ET DES CRÉATEURS DE CONTENU

Recommandation CM/Rec(2026)4 et exposé des motifs

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

SÉCURITÉ ET AUTONOMISATION EN LIGNE DES UTILISATEURS ET DES CRÉATEURS DE CONTENU

Recommandation CM/Rec(2026)4
adoptée par le Comité des Ministres
du Conseil de l'Europe
le 8 avril 2026

Conseil de l'Europe

Édition anglaise :
Online safety and empowerment of users and content creators

Conception de la couverture et mise en page :
Division publications et identité visuelle (DPIV), Conseil de l'Europe
Photos : © Shutterstock

Cette publication n'a pas fait l'objet d'une relecture typographique et grammaticale de l'Unité éditoriale de la SPDP.

Imprimé dans les ateliers du Conseil de l'Europe

Table des matières

Recommandation CM/Rec(2026)4

du Comité des Ministres aux États membres sur la sécurité et l'autonomisation en ligne des utilisateurs et des créateurs de contenu

(adoptée par le Comité des Ministres le 8 avril 2026, lors de la 1556e réunion des Délégués des Ministres)

Préambule

Le Comité des Ministres, en vertu de l'article 15.*b* du Statut du Conseil de l'Europe (STE n° 1),

Considérant que le but du Conseil de l'Europe est de réaliser une union plus étroite entre ses membres afin de sauvegarder et de promouvoir les idéaux et les principes qui sont leur patrimoine commun, notamment en encourageant des politiques et des normes communes ;

Compte tenu des obligations des États membres au titre de la Convention de sauvegarde des droits de l'homme et des libertés fondamentales (STE n° 5, « la Convention »), telle qu'interprétée par la Cour européenne des droits de l'homme (« la Cour ») dans sa jurisprudence, de reconnaître à toute personne relevant de leur juridiction les droits et libertés définis dans la Convention ;

Réaffirmant leur engagement en faveur de la promotion et de la protection des droits humains dans l'environnement en ligne ;

Soulignant que les États membres ont l'obligation positive de garantir l'exercice du droit à la liberté d'expression (article 10 de la Convention) et des autres droits conventionnels, comme le droit à la vie privée et familiale (article 8 de la Convention), aussi bien hors ligne qu'en ligne, mais aussi l'obligation négative de ne pas imposer de restrictions aux droits, à l'exception de celles prévues par la loi et nécessaires, dans une société démocratique, à la poursuite d'un but légitime ;

Conscient de la nécessité de créer les conditions d'un internet libre, ouvert et accessible à toutes et à tous, en mettant en place, dans le même temps, un environnement favorable à l'exercice en ligne du droit à la liberté d'expression et des autres droits ;

Ayant à l'esprit que l'environnement en ligne est devenu l'un des principaux lieux d'exercice du droit à la liberté d'expression y compris du droit à recevoir des informations, que l'activité expressive générée par les utilisateurs constitue un exercice riche et diversifié du droit à la liberté d'expression, et que toute approche réglementaire devrait valoriser et préserver les avantages uniques associés à cette possibilité ;

Soulignant que les technologies numériques ont renforcé la capacité des individus et des groupes d'individus à recevoir et à communiquer des informations, et qu'elles ont augmenté l'étendue et la diversité des informations auxquelles les individus et les groupes peuvent avoir accès ;

Rappelant que les prestataires de services de la société de l'information contribuent fortement à faciliter l'accès à l'information et au débat sur un large éventail de sujets politiques, sociaux et culturels ;

Rappelant que le droit à la liberté d'expression protège non seulement les informations et les idées accueillies avec faveur ou considérées comme inoffensives ou indifférentes, mais aussi celles qui peuvent heurter, choquer ou inquiéter les pouvoirs publics ou une partie de la population, et reconnaissant que cette protection est essentielle dans une société démocratique et qu'elle s'étend aux propos en ligne ainsi qu'aux travaux des créateurs de contenu dans l'environnement numérique ;

Ayant à l'esprit l'urgente nécessité de faire en sorte que les femmes et les filles, les enfants et les personnes en situation de vulnérabilité ou exposées à la discrimination, notamment sur la base de leur handicap ou de leur appartenance à une minorité nationale, ethnique, linguistique ou religieuse, ou aux communautés de personnes lesbiennes, gays, bisexuelles, transgenres et intersexes (LGBTI), ainsi que les migrants et les personnes issues de l'immigration, puissent effectivement, individuellement ou collectivement, accéder à l'environnement en ligne et y exercer en toute autonomie leur capacité d'action ;

Rappelant le rôle des médias et des autres « chiens de garde publics » dans une société démocratique, ainsi que la nécessité de garantir le pluralisme des médias, la protection du journalisme et la sécurité des journalistes et des autres acteurs des médias à la fois en ligne et hors ligne ;

Reconnaissant que l'environnement en ligne comporte des risques et que les préjudices potentiels qui en découlent sont susceptibles de porter atteinte à l'exercice des droits humains et au fonctionnement de la démocratie ;

Constatant que les femmes et les filles, les enfants et les personnes en situation de vulnérabilité ou exposées à la discrimination rencontrent dans l'environnement en ligne des risques spécifiques et accrus, dont le ciblage fondé sur l'identité et des obstacles intersectionnels, au plein exercice de leurs droits humains, et reconnaissant que ces risques peuvent s'étendre à l'environnement physique, ce qui renforce les inégalités et les préjudices déjà existants ;

Soulignant la nécessité de mesures destinées à protéger aussi bien le droit à la liberté d'expression des utilisateurs qui partagent des contenus que la sécurité des utilisateurs qui risquent d'être réduits au silence par des contenus potentiellement préjudiciables, afin de permettre la participation pleine et entière de toutes et de tous ;

Reconnaissant le besoin de cadres juridiques et d'autres initiatives transparents et fondés sur des informations factuelles pour veiller à ce que les risques en ligne, ainsi que les préjudices qui en résultent, soient évalués, traités et atténués dans le respect des droits humains et en excluant toute ingérence disproportionnée dans le droit à la liberté d'expression et les autres droits humains ;

Soulignant la nécessité que ces mesures d'évaluation et d'atténuation des risques soient entreprises en consultation avec les utilisateurs, y compris les créateurs de contenu, les groupes et les communautés concernés, et les autres acteurs pertinents de la société civile ;

Reconnaissant que l'autonomisation des utilisateurs est fondée sur leur dignité humaine et leur indépendance, et qu'elle contribue à l'équité dans l'accès aux technologies numériques, rend possible le plein exercice des droits humains dans l'environnement en ligne et favorise la participation inclusive de toutes et de tous aux espaces numériques ;

Soulignant que l'autonomisation des utilisateurs représente un important moyen d'assurer la réalisation de l'ensemble des droits humains dans l'environnement en ligne, et insistant en particulier sur le fait qu'un espace en ligne plus sûr peut créer un environnement favorable à l'exercice de la liberté d'expression ;

Soulignant que dans tous les cas où il est avéré, ou flagrant, que l'autonomisation des utilisateurs ne suffit pas à atténuer les conséquences néfastes des risques en ligne, les États devraient étudier des moyens alternatifs et proportionnés de traiter les préjudices qui découlent de ces risques, y compris l'imposition d'obligations de vigilance aux plateformes et l'application proportionnée de restrictions aux contenus ou à leur accessibilité ;

Soulignant que toutes les lois, réglementations ou mesures visant à prévenir ou à atténuer les préjudices découlant des risques en ligne doivent être fondées sur des informations factuelles, nécessaires et proportionnées au but poursuivi, rédigées en termes précis et prévisibles quant à leurs effets ;

Reconnaissant que l'adoption de mesures restrictives disproportionnées pour prévenir ou atténuer les préjudices découlant des risques en ligne nuit à l'exercice du droit à la liberté d'expression et d'information, aux débats sur des sujets d'intérêt public, à la jouissance des autres droits humains, à la capacité des utilisateurs à gérer les risques auxquels ils peuvent être exposés et à leur confiance envers les médias et les autres contenus en ligne, et sape à terme le fonctionnement de la démocratie ;

Reconnaissant que les mesures prises par les plateformes, notamment par le biais de la sélection, de l'organisation et de la modération des contenus, peuvent également constituer une ingérence dans l'exercice de la liberté d'expression et d'information et d'autres droits, et en affecter l'exercice de manière disproportionnée ;

Notant en outre que les choix de conception des plateformes, notamment lorsqu'ils visent à engendrer la viralité des contenus et l'engagement des utilisateurs, peuvent renforcer la visibilité et la pertinence de

contenus susceptibles de nuire à la sécurité des usagers, aux droits humains des utilisateurs, dont leur droit à la liberté d'expression, ainsi qu'à la cohésion sociale et, à terme, à la démocratie ;

Constatant en outre les fortes différences qui existent entre les plateformes en termes de taille, de portée et d'impact, ainsi que la nécessité d'une approche graduelle et proportionnée pour veiller à ce que tous les prestataires observent leurs obligations envers les droits humains des créateurs de contenu et des utilisateurs, tout en évitant d'imposer une charge excessive aux petits et très petits prestataires, et tenant compte de la responsabilité et du devoir de rendre des comptes accrus de ceux qui sont particulièrement influents ;

Rappelant sa vive inquiétude devant la concentration des pouvoirs entre les mains de quelques plateformes en ligne, l'asymétrie du rapport de force entre ces plateformes et leurs utilisateurs, et les conséquences de ces dynamiques sur la sécurité des utilisateurs, leurs droits humains et les processus et institutions démocratiques ;

Jugeant impératif de réaffirmer et de clarifier encore, notamment eu égard à la sécurité et à l'autonomisation des utilisateurs, le rôle et les effets sur les droits humains des plateformes en ligne qui exercent une grande influence sur l'environnement des communications publiques, ainsi que leurs devoirs et responsabilités à cet égard ;

Réaffirmant que les plateformes et autres intermédiaires en ligne privés, par le biais de leurs activités, ne doivent ni porter atteinte aux droits humains, ni contribuer à de telles atteintes, et qu'ils doivent prendre des mesures effectives pour prévenir et atténuer les effets néfastes de leurs opérations, produits ou services, notamment en mettant en place des mécanismes pour leur responsabilisation, des voies de recours et des mesures d'autonomisation des utilisateurs ;

Prenant en considération les principes applicables dans ce domaine et sans préjudice des obligations énoncées dans les conventions du Conseil de l'Europe pertinentes, dont :

- la Convention pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel (STE n° 108) et son Protocole d'amendement (STCE n° 223) ;
- la Convention sur la cybercriminalité (STE n° 185), son Protocole additionnel relatif à l'incrimination d'actes de nature raciste et xénophobe commis par le biais de systèmes informatiques (STE n° 189) et son Deuxième Protocole additionnel relatif au renforcement de la coopération et de la divulgation de preuves électroniques (STCE n° 224) ;
- la Convention sur la protection des enfants contre l'exploitation et les abus sexuels (STCE n° 201) ;
- la Convention du Conseil de l'Europe sur l'accès aux documents publics (STCE n° 205) ;
- la Convention du Conseil de l'Europe sur la prévention et la lutte contre la violence à l'égard des femmes et la violence domestique (STCE n° 210) ; et
- la Convention-cadre du Conseil de l'Europe sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit (STCE n° 225) ;

Considérant et appelant à mettre en œuvre les importantes recommandations et déclarations déjà adoptées par le Comité des Ministres concernant l'exercice et la protection du droit à la liberté d'expression en ligne, dont :

- la Recommandation CM/Rec(2016)5 sur la liberté d'internet ;
- la Recommandation CM/Rec(2018)2 sur les rôles et les responsabilités des intermédiaires d'internet ;
- la Recommandation CM/Rec(2018)7 sur les Lignes directrices relatives au respect, à la protection et à la réalisation des droits de l'enfant dans l'environnement numérique ;
- la Recommandation CM/Rec(2022)4 sur la promotion d'un environnement favorable à un journalisme de qualité à l'ère du numérique ;
- la Recommandation CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication ;
- la Recommandation CM/Rec(2022)13 sur les effets des technologies numériques sur la liberté d'expression ;
- la Recommandation CM/Rec(2022)16 sur la lutte contre le discours de haine ;
- la Recommandation CM/Rec(2026)2 sur l'obligation de rendre des comptes en matière de violence à l'égard des femmes et des filles facilitée par la technologie ; et

– la Déclaration relative à la protection du droit au respect de la vie privée des enfants dans l'environnement numérique ; Prenant également en compte les documents d'orientation pertinents adoptés par le Comité directeur sur les médias et la société de l'information, notamment :

– la Note d'orientation sur la lutte contre la propagation de la mésinformation et de la désinformation en ligne par le biais de la vérification des faits et de la conception de plateformes dans le respect des droits de l'homme ;

– la Note d'orientation sur la modération de contenu ; et la Note d'orientation sur la hiérarchisation des contenus d'intérêt public en ligne ;

Soulignant la nécessité d'une mise en œuvre rapide et approfondie de la jurisprudence pertinente de la Cour européenne des droits de l'homme,

Recommande aux gouvernements des États membres :

– de revoir leurs cadres législatifs, leurs politiques et leurs propres pratiques à la lumière des principes figurant dans l'annexe à la présente recommandation, et de promouvoir leur application dans tous les domaines concernés ;

– de prendre en compte, dans la mise en œuvre de ces principes, les normes énoncées dans la Convention, la jurisprudence pertinente de la Cour européenne des droits de l'homme et les recommandations antérieures du Comité des Ministres aux États membres, ainsi que les déclarations du Comité des Ministres concernant la mise en œuvre des droits humains dans l'environnement en ligne, dont notamment le droit à la liberté d'expression, le droit au respect de la vie privée, le droit à la liberté de réunion et d'association, et le droit à la protection des groupes qui risquent d'être pris pour cibles ou dont la sécurité et le bien-être encourent des risques accrus ;

– de promouvoir les objectifs de la présente recommandation aux niveaux national et international en faisant en sorte qu'ils soient traduits et diffusés le plus largement possible, ainsi qu'en nouant un dialogue et une coopération avec tous les acteurs pertinents et intéressés par la réalisation de ces objectifs, dont les instances de régulation, les organisations de la société civile, les entreprises et les autres parties prenantes ;

– d'évaluer périodiquement, afin d'en améliorer l'efficacité, les mesures prises pour mettre en œuvre la présente recommandation et d'informer le Comité des Ministres, dans un délai de quatre ans à compter de son adoption, des mesures prises par les États membres et par les autres parties prenantes, des progrès réalisés et des lacunes qui subsistent ;

– de veiller à ce que, en plus des organes législatifs et exécutifs, toutes les parties prenantes concernées (notamment les plateformes en ligne et les autres acteurs du secteur privé, les journalistes et les autres acteurs des médias, les organismes d'autorégulation et de corégulation, les organisations de la société civile et les milieux universitaires) soient associées aux processus d'examen, de mise en œuvre et d'évaluation susmentionnés, et de faire en sorte que toutes soient conscientes de leurs rôles, droits et responsabilités respectifs dans le cadre des efforts visant à assurer la sécurité en ligne et l'autonomisation des utilisateurs et des créateurs de contenu.

ANNEXE À LA RECOMMANDATION CM/REC(2026)4 SUR LA SÉCURITÉ ET L'AUTONOMISATION EN LIGNE DES UTILISATEURS ET DES CRÉATEURS DE CONTENU

Principes pour la sécurité et l'autonomisation en ligne des utilisateurs et des créateurs de contenu

I. Raison d'être, champ d'application et définitions

Raison d'être

1. Les prestataires de services de la société de l'information contribuent fortement à faciliter l'accès à l'information et aux débats sur un large éventail de sujets politiques, sociaux et culturels. L'environnement en ligne, et particulièrement un petit nombre de plateformes influentes, est devenu le principal moyen d'exercice du

droit à la liberté d'expression, ainsi que d'autres droits. Cette expansion des possibilités d'expression s'accompagne d'inquiétudes croissantes ; en effet, la sécurité en ligne des utilisateurs et des créateurs de contenu est exposée à des risques qui sont susceptibles de causer de graves préjudices aux individus, aux groupes, à la société ou à des intérêts sociétaux tels que la démocratie, l'État de droit et la libre circulation des informations. Ces risques sont à la fois une spécificité de l'environnement en ligne ainsi qu'une prolongation et une amplification des risques déjà présents dans la société.

2. La sécurité en ligne devrait être considérée comme l'un des éléments d'un environnement en ligne favorable. Un tel environnement devrait être accessible sans discrimination, il devrait être sûr, inclusif, pluriel et fiable. Il devrait permettre aux utilisateurs de jouir de leurs droits humains et de les exercer sans ingérence injustifiée, et leur offrir le plus d'autonomie et de possibilités de participation et d'engagement possible.

3. La sécurité en ligne suppose que les États et les plateformes adoptent, dans leurs domaines de compétence respectifs :

- des mesures proportionnées, respectueuses des droits humains et nécessaires dans une société démocratique pour lutter contre le risque que des personnes ou des groupes de personnes ne soient exposées à la violence, l'exploitation, la discrimination, y inclus les crimes de haine, et à d'autres ingérences illégales dans leurs droits humains dans l'environnement en ligne ou à la suite d'activités se déroulant dans cet environnement ;
- des mesures qui donnent aux utilisateurs les moyens de comprendre ces risques et d'ajuster leur expérience en ligne à leurs choix et à leurs préférences.

4. Les règles relatives aux contenus et l'application de la responsabilité des utilisateurs et des plateformes ne peuvent, à elles seules, résoudre le problème à grande échelle, et la mise en œuvre de ces règles peut s'avérer arbitraire et sélective. Elles risquent aussi d'imposer ou d'inciter à imposer aux contenus des restrictions et des pratiques de modération excessives ou autrement disproportionnées, avec des effets négatifs sur les droits humains, en particulier la liberté d'expression et le droit au respect de la vie privée. Par conséquent, en n'abordant la sécurité en ligne qu'à travers ce type de mesures, on passe à côté des défis à relever pour promouvoir un environnement en ligne favorable aux utilisateurs et aux créateurs de contenu.

5. La promotion et la protection effectives des droits humains en ligne exigent l'élaboration de cadres de régulation et de corégulation proportionnés et fondés sur des informations factuelles, destinés à la fois à améliorer la transparence et la responsabilisation des plateformes quant à leurs choix de conception et à leur fonctionnement, et à autonomiser en ligne les utilisateurs et les créateurs de contenu. Ces cadres, complétés par des politiques adéquates de promotion de l'autonomisation au sein de la société, contribuent à construire des environnements en ligne plus sûrs dès leur conception et à sensibiliser davantage les utilisateurs aux risques en ligne, à améliorer leurs connaissances en la matière et leur capacité à y faire face, rendant ainsi les espaces en ligne plus favorables à la liberté d'expression.

Champ d'application

6. Les présents principes se concentrent sur les moyens de répondre, au travers de mesures conformes aux droits humains, aux risques en ligne qui résultent de l'exercice du droit à la liberté d'expression ou qui ont une incidence sur cet exercice.

7. Ils visent à orienter les États dans l'adoption, la mise en œuvre et l'exécution de politiques et de cadres juridiques, ainsi que d'autres mesures, destinés à répondre aux risques qui pèsent sur la sécurité en ligne, à atténuer les risques de préjudices et à promouvoir un environnement en ligne favorable à l'exercice des droits humains.

8. Ils abordent également les mesures que les plateformes devraient prendre, ou que les États devraient exiger d'elles, afin de s'acquitter de leurs propres responsabilités dans la création d'un tel environnement.

9. La présente recommandation vise deux objectifs distincts, mais interdépendants : protéger la sécurité des utilisateurs et leur donner des moyens d'agir (« autonomisation ») en ligne.

10. L'autonomisation et la sécurité emportent, pour l'État, à la fois l'obligation positive de prendre des mesures pour assurer la jouissance des droits humains et l'obligation négative de ne pas s'ingérer dans ces droits au-delà de ce qui est nécessaire, dans une société démocratique, pour atteindre un but légitime. Ensemble, ces deux obligations de l'État visent à promouvoir un accès équitable aux technologies de la communication, à rendre possible l'exercice plein et entier des droits humains, et à favoriser la participation inclusive de toutes et tous dans le cadre des espaces en ligne.

Définitions

11. Aux fins de la présente recommandation, on entend par :

- « utilisateurs », toute personne physique ou morale, ou groupe de personnes, qui utilise des services en ligne ;
- « créateurs de contenu », les utilisateurs qui produisent et diffusent, via une plateforme, de manière régulière ou à titre professionnel, des informations et des idées, sous forme écrite, sonore, visuelle, audiovisuelle ou autre, dans l'intention de toucher un public qui dépasse leur cercle privé ;
- « intermédiaires d'internet », les acteurs définis par la Recommandation CM/Rec(2018)2 sur les rôles et les responsabilités des intermédiaires d'internet ; gardant à l'esprit que les intermédiaires d'internet offrent des fonctions et des services variés, et qu'ils peuvent exercer plusieurs fonctions en parallèle ; le cas échéant, ce document fait référence à certaines fonctions spécifiques qu'ils exercent ;
- « plateformes », des prestataires de services numériques en ligne qui ont comme but, fonction ou utilisation principale de mettre en lien leurs utilisateurs et de faciliter l'échange d'informations et d'idées entre eux, sur des forums en accès public, et qui fixent les règles applicables à leurs interactions, y inclus l'usage fréquent de systèmes algorithmiques pour collecter et analyser des données ou personnaliser leurs services. Dans le domaine des communications, de telles plateformes englobent les réseaux sociaux, mais aussi les moteurs de recherche, les agrégateurs d'actualités et les services de partage de vidéos, dans la mesure où ces plateformes offrent une fonctionnalité d'interaction entre utilisateurs ;
- « plateformes exerçant une influence significative », des plateformes qui, en raison de leur taille, de leur portée ou de leur impact, contribuent d'une manière significative à façonner le paysage informationnel au niveau mondial ou sur un territoire spécifique et ont donc des effets concrets sur la jouissance et l'exercice de la liberté d'expression et d'information, et des autres droits humains, ainsi que sur le fonctionnement de la démocratie. Les critères sur la base desquels la taille, la portée et l'impact des différentes plateformes sont évalués devraient être spécifiés clairement et revus périodiquement, dans le cadre du droit national ;
- « conception de plateforme », toutes les grandes décisions qui structurent le fonctionnement d'une plateforme en ligne et déterminent la manière dont les utilisateurs l'expérimentent ; la définition englobe également les moyens techniques permettant aux plateformes de mettre en œuvre, de maintenir et de mettre à jour leurs architectures et interfaces, dont les fonctions de fiabilité et de sécurité destinées aux utilisateurs, mais elle ne s'étend pas aux choix, fondés sur les contenus, concernant la sélection et la modération des contenus licites ;
- « autonomisation des utilisateurs », les moyens qui permettent aux utilisateurs de mieux comprendre et contrôler leur expérience en ligne, et d'opérer à son sujet des choix éclairés, afin de bénéficier pleinement des possibilités offertes par cette expérience et d'en gérer les risques sans que cela ne représente pour eux une charge excessive. Ces moyens englobent à la fois des mesures à adopter dans le monde physique, comme l'éducation à la citoyenneté numérique, les stratégies et initiatives pour l'éducation aux médias et à l'information, l'association des utilisateurs aux processus décisionnels qui les concernent, ainsi que des mesures à adopter dans les environnements en ligne, telles que de mettre à la disposition des utilisateurs les outils leur permettant de personnaliser efficacement leur expérience sur les plateformes en fonction de leurs préférences, les possibilités d'exercer et de protéger les droits des utilisateurs, et les pistes d'actions collectives ;
- « autorégulation », le processus par lequel un acteur privé ou le secteur lui-même élabore et applique ses propres règles afin d'atteindre un objectif de politique publique ou sectorielle ; cela inclut les politiques et règle contractuelles des plateformes qui affecte les utilisateurs de leurs services ;
- « corégulation », l'autorégulation du secteur exercée dans le cadre d'un mandat de l'État et/ou faisant l'objet d'une forme de supervision par l'État ;
- « contenu restreint par la loi », une expression, ou une manifestation d'un comportement, par des utilisateurs, qui n'est pas conforme au droit applicable, ce qui englobe les contenus illégaux et les contenus légaux, mais réglementés ;
- « contenu illégal », une expression, ou une manifestation d'un comportement, par des utilisateurs, qui est interdite par le droit pénal, civil ou administratif ;

– « contenu légal mais réglementé », une expression, ou une manifestation d'un comportement, par des utilisateurs, qui ne constitue pas un « contenu illégal », mais dont la publication, la diffusion ou la visibilité est restreinte, eu égard au contenu spécifique et dans un cadre précis, notamment pour en réduire la visibilité par un groupe protégé, tels que les enfants, ou pour réduire l'ampleur de sa diffusion auprès du grand public, comme lorsqu'un moteur de recherche risque de révéler des données à caractère personnel ou d'exposer des résultats de sondages juste avant un scrutin ;

– « contenu légal », une expression, ou une manifestation d'un comportement, par des utilisateurs, qui ne constitue pas un contenu restreint par la loi ;

– « signalement », une alerte émanant d'un utilisateur, au moyen d'une fonctionnalité intégrée à la conception de la plateforme, attirant l'attention sur un contenu ou comportement susceptible d'enfreindre les politiques et règles contractuelles de la plateforme ou les normes juridiques émanant de l'État ;

– « notification », une demande formelle, adressée à un intermédiaire et emportant des effets juridiques, par laquelle un utilisateur ou un tiers sollicite le retrait ou la restriction d'un contenu. Les notifications peuvent émaner entre autres d'utilisateurs, de détenteurs de droits, d'autorités de régulation ou d'autres autorités publiques ;

– « injonction », une directive juridiquement contraignante par laquelle une autorité publique exige d'une plateforme qu'elle accomplisse une action, telle que retirer un contenu ou lui appliquer des restrictions, suspendre ou supprimer un compte, ou donner la priorité à l'évaluation de la légalité de contenus spécifiques.

II. Risques en ligne relatifs à la liberté d'expression

12. L'exercice de la liberté d'expression en soi est susceptible de heurter, de choquer ou d'inquiéter les pouvoirs publics ou certains pans de la population. Cela ne suffit pas, dans une société démocratique, à justifier l'adoption de mesures restreignant cette liberté. Les opinions qui remettent en question les points de vue dominants sont indispensables à un débat public permettant aux sociétés démocratiques de corriger leurs erreurs, de renforcer leur gouvernance et d'encourager une amélioration constante.

13. Cependant, certains risques en ligne peuvent affecter les utilisateurs et les créateurs de contenu en portant atteinte à leurs droits ou en les amenant par ailleurs à craindre pour leur bien-être et leur sécurité, avec pour effet d'entraver leur volonté, leur capacité et leur détermination à s'exprimer librement. Les risques en ligne peuvent aussi limiter l'aptitude du public à accéder à des informations fiables, à découvrir des points de vue variés et à se forger une opinion éclairée sur les questions d'intérêt général. Ils peuvent aussi entraîner des conséquences sociétales plus vastes, tels que l'érosion de la cohésion sociale et de la confiance envers les institutions, des menaces pour la santé publique et l'affaiblissement des processus démocratiques.

14. Les risques en ligne relatifs à la liberté d'expression comprennent :

a. les risques pour la sécurité et le bien-être des personnes ou des communautés de personnes du fait de contenus ou de communications auxquels les utilisateurs peuvent se trouver exposés, avec lesquels ils peuvent interagir ou dont ils peuvent être la cible ;

b. les risques pour les processus démocratiques, pour l'intégrité de l'information et pour le caractère éclairé du discours public ;

c. les risques associés aux systèmes déployés par des prestataires et susceptibles de constituer une ingérence dans la liberté d'expression, le respect de la vie privée, la protection des données personnelles et d'autres droits des utilisateurs.

15. Ces risques peuvent découler directement de l'activité en ligne d'autres utilisateurs et d'autres créateurs de contenu, mais aussi être suscités ou exacerbés par la conception et le fonctionnement de la plateforme.

16. La disponibilité à grande échelle de systèmes d'intelligence artificielle qui permettent de produire, de publier et d'accroître ou, au contraire, de réduire la visibilité de certains contenus crée de nouveaux risques ; elle est également susceptible d'amplifier les risques existants.

17. Certaines catégories d'utilisateurs et de créateurs de contenu rencontrent, dans l'environnement en ligne, des risques spécifiques et accrus, dont un ciblage fondé sur l'identité et des obstacles intersectionnels au plein exercice de leur droit à la liberté d'expression. Ces risques peuvent gagner l'environnement physique, ce qui renforce les inégalités et les préjudices existants. Les groupes à risque accru comprennent :

a. les enfants, qui se trouvent dans une situation accrue de vulnérabilité dans l'environnement en ligne et qui ont le droit d'être tenus à l'abri des contenus auxquels la loi interdit de les exposer, conformément à la Convention du Conseil de l'Europe sur la protection des enfants contre l'exploitation et les abus sexuels et à la Recommandation CM/Rec(2018)7 sur les Lignes directrices relatives au respect, à la protection et à la réalisation des droits de l'enfant dans l'environnement numérique ;

b. les femmes et les filles, en particulier lorsqu'elles sont créatrices de contenu, qui rencontrent un risque accru de violences et d'agressions en ligne, qui souvent revêtent un caractère genré et visent à les réduire au silence. En outre, les femmes et les filles peuvent être affectées négativement, et de manière disproportionnée, par diverses formes de contenus en ligne susceptibles de refléter et d'exacerber des dynamiques préjudiciables déjà présentes dans la société ;

c. des individus et des groupes perçus comme se trouvant en situation vulnérable ou exposés à la discrimination, notamment sur la base de leur handicap ou de leur appartenance à une minorité nationale, ethnique, linguistique ou religieuse ou aux communautés LGBTI, ainsi que les migrants et les personnes issues de l'immigration, qui subissent un ciblage fondé sur leur identité et destiné à les réduire au silence ; et

d. les journalistes, responsables politiques, chercheurs, militants et autres acteurs contribuant fréquemment au débat sur les questions d'intérêt général et à d'éventuelles controverses sont souvent ciblés par des contenus qui visent à les empêcher de prendre part aux discussions ou d'exercer leur liberté d'expression à l'avenir. Cela peut passer par des menaces ou d'autres types d'agressions, avec pour cibles leur propre personne, leur famille, leurs collaborateurs ou leurs communautés.

18. La présence de risques en ligne ne demande pas toujours l'introduction de mesures constituant une ingérence dans l'exercice des droits humains, notamment de la liberté d'expression. Les mesures prises devraient être proportionnées au risque d'un préjudice.

19. Les mesures adoptées dans l'intérêt de la sécurité en ligne, qu'elles soient mises en œuvre par les États ou par les plateformes, peuvent elles-mêmes présenter un risque pour la jouissance des droits humains, notamment la liberté d'expression. Les États et les acteurs privés, lorsqu'ils répondent aux inquiétudes que soulèvent les effets potentiellement néfastes des différents types de risques exposés ci-dessus, devraient par conséquent s'assurer que ces mesures ne constituent pas une ingérence disproportionnée dans la liberté d'expression et dans les autres droits.

III. Principes généraux d'un environnement en ligne favorable

Principes à l'attention des États

20. La gouvernance de la sécurité en ligne et l'autonomisation des utilisateurs devrait avoir pour but, pour les États comme pour les intermédiaires d'internet, y compris les plateformes, de créer un environnement en ligne favorable, tel que décrit au paragraphe 2.

21. Les mesures spécifiques à l'espace en ligne, telles que celles recommandées à la partie IV, paragraphes 57 à 65 (Règles relatives à la responsabilisation des plateformes et à l'autonomisation des utilisateurs), et à la partie V (Mesures d'autonomisation des utilisateurs en ligne), ne suffisent pas à garantir un environnement en ligne favorable. Elles devraient venir compléter et prolonger des actions plus larges dans le monde hors ligne. Les politiques et interventions des États à cet égard devraient s'inscrire dans une stratégie, globale et coordonnée, axée sur les réalités sociétales et les inégalités qui sont à l'origine de comportements abusifs en ligne et qui exposent des utilisateurs à ces comportements. Elles devraient promouvoir l'égalité, la cohésion sociale et les valeurs démocratiques, renforcer l'État de droit, veiller à la sûreté publique et donner aux utilisateurs les moyens de prendre des décisions éclairées sur leur expérience en ligne. Parmi ces mesures devraient figurer des projets éducatifs en faveur de la citoyenneté numérique, des stratégies et politiques renforçant l'éducation aux médias et à l'information, des initiatives pour l'autonomisation au sein des communautés, des mesures destinées à promouvoir des médias libres, indépendants, responsables et pluralistes et un journalisme de qualité, et des mécanismes effectifs visant à préserver la sécurité et le bien-être physiques et mentaux des utilisateurs, à enquêter sur les infractions pénales facilitées par les technologies et à s'assurer que les auteurs de ces infractions rendent des comptes, conformément à la loi.

22. Les États devraient s'abstenir d'actions susceptibles de compromettre la sécurité en ligne, que ce soit par l'aggravation des risques de préjudices ou par la réduction des possibilités de protection et d'autonomisation. En particulier, les États devraient éviter toute mesure susceptible de créer de nouvelles failles ou

vulnérabilités dans les composantes techniques des services en ligne qui constituent des garanties cruciales pour l'exercice du droit à la vie privée et des autres droits humains en ligne.

23. Les États devraient étudier, évaluer et prendre en compte les effets de leurs interventions – ou de leur absence d'intervention – sur le caractère accessible et inclusif des plateformes pour l'ensemble des utilisateurs, indépendamment de leur situation socio-économique, de leurs handicaps ou de toute autre situation constituant un désavantage. Ils devraient, par ailleurs, prêter une attention particulière aux risques accrus auxquels peuvent être exposées en ligne les catégories d'utilisateurs énoncées au paragraphe 17.

24. Les mesures destinées à évaluer et à traiter les risques à l'égard des enfants, à réduire les préjudices, à autonomiser les enfants et à les protéger devraient avoir l'intérêt supérieur de l'enfant pour considération primordiale et prendre en compte l'âge des enfants, leurs vulnérabilités et l'état d'évolution de leurs capacités. Toutes ces mesures devraient être prises dans le respect de leurs droits, y compris les droits à la liberté d'expression et à la vie privée.

25. Les pouvoirs publics devraient faire preuve de transparence quant à leurs interactions avec les plateformes. Notamment, ils devraient garantir l'accès à l'information et veiller à ce que les exceptions relatives à des intérêts commerciaux et autres intérêts économiques ne soient pas rédigées ou interprétées de manière à exclure indûment l'intérêt du public à obtenir des informations sur ces interactions, conformément aux normes d'accès à l'information, telles que promues par la Convention du Conseil de l'Europe sur l'accès aux documents publics. De plus, les États devraient promouvoir la transparence sur le fonctionnement des plateformes, y compris en veillant à ce que les lanceurs d'alerte soient pleinement protégés, conformément à la Recommandation CM/Rec(2014)7 sur la protection des lanceurs d'alerte.

Principes à l'attention des plateformes

26. Toutes les plateformes, d'autant plus lorsqu'elles exercent une influence significative, jouent un rôle central dans la possibilité d'exercer le droit à la liberté d'expression ; par conséquent, il est de leur responsabilité d'intégrer les considérations de sécurité et d'autonomisation des utilisateurs à toutes leurs décisions clés sur la conception de leurs services, y compris concernant l'intelligence artificielle et autres systèmes algorithmiques, de façon à promouvoir un environnement en ligne favorable. La nature et l'étendue de cette responsabilité devraient varier en fonction de la taille, de la portée et de l'impact de chaque plateforme.

27. Protéger la sécurité des créateurs de contenu et des utilisateurs devrait être une considération clé dans le développement, la conception, la gouvernance et le fonctionnement des plateformes. L'intégration des considérations de sécurité dans la conception et le fonctionnement des plateformes, notamment des plateformes exerçant une influence considérable, contribue à favoriser l'exercice effectif du droit à la liberté d'expression en créant des environnements où les créateurs de contenu et les utilisateurs peuvent évoluer sans craindre de subir des violences, du harcèlement ou des ingérences injustifiées. Dans le même temps, ces mesures ne devraient pas nuire au pluralisme des médias, à la diversité des voix qui se font entendre et à la nature ouverte et inclusive du discours public. Toutes les interventions devraient être transparentes, proportionnées et ancrées dans le droit international des droits humains, afin que les efforts de promotion de la sécurité ne reviennent pas à marginaliser ou à étouffer les points de vue minoritaires ou dissidents.

28. Dans l'exercice de leur responsabilité, les plateformes devraient prêter une attention particulière aux risques accrus auxquels peuvent être exposées, en ligne, les catégories d'utilisateurs énoncées au paragraphe 17.

29. Dès lors qu'une plateforme compte un nombre significatif d'utilisateurs dans un pays, une région ou un territoire donné, elle devrait s'attacher à comprendre les enjeux de sécurité en ligne spécifiques au contexte local, y compris les risques liés au genre. Elle devrait également désigner des points de contact pour la communication et la conformité, et recruter un nombre suffisant de salariés ou de sous-traitants connaissant le contexte politique, culturel et social local, et parlant couramment les langues officielles concernées, afin de pouvoir évaluer les risques et mettre en œuvre des réponses appropriées.

Principes à l'attention des créateurs de contenu

30. Il est de la responsabilité des créateurs de contenu de contribuer à un discours public sain, éclairé, démocratique et respectueux des droits d'autrui. Leur degré de responsabilité peut varier selon des facteurs tels que la nature et le format des contenus, leur apport au débat sur des questions d'intérêt général ou encore l'âge du public visé. Les créateurs de contenu qui ont une audience importante ou se présentent comme disposant d'une expertise professionnelle sont d'autant plus tenus d'agir de bonne foi, d'observer les principes d'exactitude, d'équité et d'intégrité, et de respecter les droits d'autrui.

31. Le rôle particulier, le statut professionnel ou la position sociale de certains créateurs de contenu peuvent emporter des responsabilités supplémentaires spécifiques, assorties de mécanismes leur permettant de rendre compte de leurs actions. En particulier, les créateurs de contenu qui exercent une responsabilité éditoriale sur les contenus qu'ils produisent et diffusent et qui ont pour principal objectif d'informer le public ou d'influencer son opinion sur l'actualité et sur des questions d'intérêt général devraient respecter les principes du journalisme, indépendamment de leur audience, de leur forme juridique et du fait qu'ils soient reconnus ou non comme journalistes en vertu du droit interne. Ils devraient faire preuve de transparence à l'égard de leurs sources de revenus et étiqueter clairement les contenus rémunérés et sponsorisés. Les créateurs de contenu ayant le statut de personne morale devraient faire preuve de transparence quant à l'identité de leurs propriétaires.

32. En plus des responsabilités énoncées aux paragraphes ci-dessus, les parents ou représentants légaux des enfants agissant en tant que créateurs de contenu devraient veiller, par rapport à la publication de tout contenu, à ce que l'intérêt supérieur de l'enfant soit la considération primordiale et à ce que la dignité et la sécurité de l'enfant soient préservées.

IV. Principes applicables aux cadres juridiques sur la sécurité en ligne et l'autonomisation des utilisateurs, à leur déploiement et à leur mise en œuvre

Principes communs

33. Les États ont l'obligation positive de traiter efficacement les risques de préjudice en ligne. Les États peuvent assumer cette responsabilité en adoptant et en faisant appliquer des cadres juridiques :

a. qui spécifient les contenus restreints par la loi et précisent quand les restrictions s'appliquent (règles relatives aux contenus) ;

b. qui précisent les cas exceptionnels et les conditions dans lesquels les intermédiaires d'internet peuvent être tenus responsables pour la violation par leurs utilisateurs de règles relatives au contenu (règles relatives à la responsabilité des intermédiaires) ;

c. qui imposent aux intermédiaires, tels que les plateformes, des devoirs et responsabilités de nature systémique visant à améliorer leur capacité à rendre des comptes et à renforcer la sécurité en ligne et l'autonomisation des utilisateurs au moyen d'améliorations de leurs systèmes et de leurs processus (règles relatives à la responsabilisation des plateformes et à l'autonomisation des utilisateurs).

34. Bien que les plateformes ne produisent pas de contenus, elles peuvent jouer un rôle actif dans les communications publiques dans la mesure où elles gèrent et/ou organisent des contenus, y compris par la conception et le déploiement de systèmes algorithmiques. Les devoirs imposés aux plateformes devraient donc différer de ceux qui peuvent être imposés aux médias, lesquels assument une responsabilité éditoriale sur le contenu. Cela n'empêche pas certaines règles d'être applicables à l'ensemble des médias et des plateformes, selon leurs cadres juridiques respectifs.

35. Les États devraient établir une distinction claire entre les réponses aux risques posés par la diffusion de contenus restreints par la loi, d'une part, et de contenus légaux, d'autre part. Les contenus légaux ne devraient faire l'objet, de la part des États, que de mesures d'atténuation des risques conformes aux principes concernant la responsabilisation des plateformes et l'autonomisation des utilisateurs, tels qu'indiqués plus loin, y compris les mesures d'autonomisation en ligne des utilisateurs abordées à la partie V. En plus de mesures d'atténuation des risques, les contenus restreints par la loi peuvent faire l'objet de mesures restrictives proportionnées et conformes aux principes énoncés plus loin concernant les règles relatives aux contenus et à la responsabilité des intermédiaires.

36. Les devoirs et responsabilités systémiques imposés aux intermédiaires à l'égard des contenus et comportements légaux ne devraient pas servir de prétextes pour contourner le processus législatif en appliquant aux contenus des restrictions dépourvues de base juridique claire. Cela ne devrait pas empêcher les États de demander des comptes aux plateformes quant aux procédures qu'elles appliquent pour mettre en œuvre et faire respecter les règles et restrictions relatives aux contenus tels qu'inscrites dans leurs propres politiques et règles contractuelles.

37. Le blocage ou l'interdiction d'un service en ligne, d'un domaine ou d'un site web entier constitue une restriction préalable, et donc une ingérence d'une gravité exceptionnelle dans le droit à la liberté d'expression. Toute action de ce type devrait être ordonnée par une autorité judiciaire ou une autre autorité publique indépendante dont les décisions sont soumises à un contrôle juridictionnel et répondre à une exigence de

justification très élevée. Les opérateurs, ainsi que les utilisateurs directement touchés, devraient recevoir des explications pertinentes quant à la mesure prise et pouvoir accéder à un recours effectif.

38. Les États ne devraient pas exercer de pression sur les intermédiaires d'internet et les créateurs de contenu pour qu'ils instaurent des mesures affectant la disponibilité de contenus en ligne par des moyens autres que ceux prévus par la loi. Les intermédiaires d'internet et les utilisateurs devraient disposer d'un recours effectif pour se défendre contre de telles formes de pression.

39. Tout cadre juridique régissant les contenus et la responsabilité des intermédiaires, ou la responsabilisation des plateformes et l'autonomisation des utilisateurs, qui est susceptible d'entraîner des ingérences dans la liberté d'expression, devrait respecter l'article 10 de la Convention et suivre les lignes directrices énoncées en annexe à la Recommandation CM/Rec(2022)13 sur les effets des technologies numériques sur la liberté d'expression, ainsi que les principes procéduraux énoncés dans la Recommandation CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication.

40. Tout cadre juridique régissant les contenus et la responsabilité des intermédiaires, ou la responsabilisation des plateformes et l'autonomisation des utilisateurs, applicable aux intermédiaires d'internet et à leurs relations avec les États et les utilisateurs, ainsi que toute action de mise en œuvre ou d'exécution menée conformément à ce cadre devraient suivre les principes énoncés à la partie I de l'annexe à la Recommandation CM/Rec(2018)2 sur les rôles et les responsabilités des intermédiaires d'internet.

41. Compte tenu du rôle joué par les intermédiaires en tant que facilitateurs de la liberté d'expression, les exigences liées à la responsabilisation des plateformes et, le cas échéant, la responsabilité des intermédiaires conformément à la présente recommandation, devraient suivre une approche graduée et différenciée selon les capacités technologiques et les moyens économiques de chaque intermédiaire, afin que soient associés à chacun à la fois le type de protection spécifique et le niveau de responsabilité qui lui correspond. Les actions de régulation et corégulation des plateformes devraient être proportionnées et suivre une approche graduée en fonction de leur taille, leur portée ou leur impact, d'une part pour ne pas surcharger les petits ou très petits acteurs, d'autre part pour tenir compte des responsabilités accrues qui reviennent aux plateformes exerçant une influence significative.

42. Les autorités nationales ne devraient pas imposer aux intermédiaires, directement ou indirectement, une obligation générale de surveiller, les contenus auxquels ils donnent simplement accès ou qu'ils transmettent ou stockent, que ce soit par des moyens automatisés ou non automatisés.

43. Pour être efficaces, les cadres juridiques mis en place afin de répondre aux risques en ligne devraient conserver, autant que possible, leur cohérence au-delà des frontières. Les États devraient donc coopérer pour éviter la fragmentation et pour soutenir l'application cohérente de règles conformes aux droits humains.

Règles relatives aux contenus

44. Par principe, les contenus qui sont légaux hors ligne devraient aussi l'être en ligne.

45. Toute règle restreignant la publication, la diffusion ou l'accessibilité en ligne de certains types d'expression ou de manifestation d'un comportement (règles relatives aux contenus) devrait être prévue par la loi, poursuivre l'un des buts légitimes énumérés à l'article 10.2 de la Convention, recourir à des moyens proportionnés et satisfaire aux exigences de sécurité juridique, de nécessité et de prévisibilité. Le cas échéant, elles devraient également préciser les devoirs et les responsabilités qui en découlent pour les intermédiaires d'internet. Toute règle de ce type devrait reposer sur des informations factuelles réunies de manière transparente. Le champ matériel et géographique de l'application des restrictions et des mesures correctrices associées devrait être proportionné, afin d'éviter le retrait collatéral de contenus légaux.

46. La loi devrait identifier les contenus restreints avec suffisamment de clarté pour que son application soit prévisible, anticipable et non arbitraire. Ces critères peuvent varier selon la sévérité de la restriction imposée à la liberté d'expression.

47. Les contenus susceptibles ou devant faire l'objet de restrictions légales présentent des différences quant à leur degré de gravité et à l'imminence des risques qu'ils comportent. Les règles relatives au contenu ne devraient prévoir que des restrictions nécessaires et proportionnées à la gravité des contenus proscrits et à leur potentiel de nuisance. Elles peuvent aller de l'interdiction pure et simple pour les contenus tombant sous le coup du droit pénal, administratif ou civil, qui devraient être retirés ou bloqués (contenus illégaux), à des mesures réduisant l'accessibilité, la diffusion ou la visibilité de contenus dans des circonstances spécifiques (contenu légal mais réglementé), tels que les limites d'accès fondées sur l'âge, les règles spécifiques aux médias audiovisuels ou aux annonces à caractère commercial, les mesures destinées à concrétiser le

droit à l'oubli ou les restrictions uniquement en vigueur pendant les périodes électorales. Les restrictions imposées aux contenus légaux mais réglementés doivent toujours être appréciées au cas par cas et ne pas être présumées moins sévères que celles imposées aux contenus illégaux. Comme l'environnement en ligne ne cesse d'évoluer, entraînant des situations et des difficultés nouvelles, les États membres devraient examiner régulièrement leurs règles juridiques relatives aux contenus pour s'assurer qu'elles sont assez claires et à jour pour répondre aux défis émergents dans des domaines spécifiques.

48. Les États ne sauraient appliquer de restrictions que par rapport aux contenus entrant dans la catégorie des contenus restreints par la loi et dans le respect de l'article 10, paragraphe 2, de la Convention. La mise en œuvre par les pouvoirs publics de restrictions juridiques concernant des contenus publiés sur des plateformes doit être prévue par la loi, s'exercer dans les limites définies par la loi et s'accompagner de garanties contre toute application sélective, discriminatoire ou arbitraire. Toute décision de mise en œuvre devrait rester dans les limites de ce qui est nécessaire et proportionné, et fournir des motifs pertinents et suffisants.

49. Les mesures visées au paragraphe 48 devraient, par principe, n'être adoptées que sur la base d'injonctions formelles, émanant d'une autorité judiciaire ou d'une autre autorité publique indépendante dont les décisions sont soumises à un contrôle juridictionnel. Il convient de ne pas donner à de simples notifications, pour la seule raison qu'elles émanent de pouvoirs publics, les mêmes effets juridiques qu'à des injonctions. Tout utilisateur ou intermédiaire dont la liberté d'expression ou d'autres droits sont restreints du fait de telles mesures devrait avoir accès à un recours effectif devant une juridiction compétente.

50. Les plateformes peuvent aussi restreindre les contenus et les comportements légaux produits par des utilisateurs par le biais de leurs politiques et règles contractuelles, telles que les accords sur les conditions de service ou les standards de la communauté. Ce faisant, elles devraient pleinement évaluer et prendre en compte les effets de ces restrictions sur les droits humains des utilisateurs. Les politiques et règles contractuelles en question devraient être transparentes, communiquées en termes clairs et élaborées en consultation avec les utilisateurs et les communautés d'utilisateurs, qui doivent y apporter une véritable contribution. Elles devraient être appliquées de manière cohérente et sans discrimination ni arbitraire.

51. Tout en prenant en considération les garanties énoncées aux paragraphes précédents, les plateformes devraient adopter des mesures proportionnées et fondées sur des informations factuelles contre les créateurs de contenu ou les utilisateurs qui diffusent de manière répétée des contenus dont il est manifeste qu'ils sont restreints par la loi, et contre les signaleurs qui envoient de manière répétée des signalements dont il est manifeste qu'ils sont inexacts ou infondés. Ces mesures peuvent comprendre la réduction de la visibilité de leurs publications ou de leurs comptes, ou de leur capacité à les monétiser, des limites temporaires à leur capacité d'envoyer des signalements, la suspension ou la clôture de leur compte, et l'exclusion de l'utilisateur. Il convient de n'appliquer de telles mesures que lorsque les raisons d'intervenir sont claires, obéissent à un haut niveau de justification et s'accompagnent de garanties procédurales, comme énoncé plus loin, aux paragraphes 90 à 93.

52. Les États peuvent obliger les plateformes à divulguer des informations d'intérêt public spécifiques sur leurs services. Cependant, étant donné que cette obligation constitue une ingérence dans les droits des plateformes en vertu de l'article 10 de la Convention, elle doit être prévue par la loi. À chaque application concrète d'une telle obligation, les États devraient démontrer que cette dernière obéit à un besoin social impérieux fondé sur des motifs pertinents et suffisants, et qu'elle est nécessaire et proportionnée pour atteindre un but légitime.

53. Les créateurs de contenu peuvent être soumis à des obligations en vertu des cadres juridiques existants, tels que ceux qui s'appliquent à la diffusion de services de médias audiovisuels ou à la protection des consommateurs. Les créateurs de contenu exerçant certaines professions, tels que les journalistes, les avocats ou les médecins, peuvent en outre être soumis à des obligations professionnelles et aux mécanismes d'autorégulation correspondants. Les États devraient exiger des créateurs de contenu professionnels la transparence sur la manière dont leurs contenus sont monétisés. Sans préjudice de ces obligations, les États devraient encourager le développement et la promotion de cadres d'autorégulation transparents, inclusifs et fondés sur les droits humains, destinés aux créateurs de contenu qui n'en disposent pas encore. Ces mécanismes devraient soutenir l'adhésion des créateurs de contenu à des normes éthiques et professionnelles, et améliorer la qualité et la fiabilité de leurs contenus, renforçant par là leur autonomisation et leur capacité à rendre des comptes.

Règles relatives à la responsabilité des intermédiaires

54. Imposer aux intermédiaires une obligation excessive de modération des contenus les contraint, en pratique, à exercer une censure pour le compte des autorités étatiques, approche incompatible avec l'article 10 de la Convention. Les États devraient s'abstenir d'imposer aux intermédiaires d'internet une responsabilité disproportionnée à l'égard des contenus restreints par la loi émanant des utilisateurs de leurs services, qui pourrait inciter ou engendrer un blocage excessif de contenus légaux.

55. Les États devraient veiller, en droit et en pratique, à ce que les intermédiaires d'internet, y compris les plateformes, ne puissent être tenus responsables des contenus de tiers auxquels ils donnent simplement accès, qu'ils transmettent ou qu'ils stockent. Les autorités étatiques peuvent tenir les intermédiaires pour coresponsables des contenus qu'ils stockent lorsque ces derniers n'agissent pas avec la diligence voulue pour restreindre l'accès aux contenus ou aux services dès lors qu'ils ont connaissance de leur caractère restreint par la loi, y compris par le biais de procédures transparentes, accessibles et efficaces reposant sur la notification. Les conditions de retrait de contenus illégaux ou l'exécution d'autres restrictions sur les contenus légaux mais réglementés devraient être prévues par la loi. Ces conditions, y inclus le cas échéant des délais de retrait, devraient être différenciées en fonction de la nature du contenu et du caractère plus ou moins grave et imminent des risques provoqués par sa diffusion.

56. Lorsque la loi, y inclus lorsqu'une autorité publique agit sous l'égide de la loi et dans le contexte de cadres de corégulation, exige des plateformes qu'elles retirent, bloquent ou rendent moins accessibles ou visibles des contenus restreints par la loi, des garanties devraient être en place pour éviter toute restriction disproportionnée. En dernier ressort, la protection des droits incombe toujours à l'État, qui ne saurait déléguer cette obligation à des entités privées. Toutes les mesures adoptées par des acteurs privés sur instruction de l'État devraient respecter la liberté d'expression et les autres droits humains. Les plateformes, lorsqu'elles donnent suite à des injonctions légales ou appliquent des restrictions à l'initiative d'une autorité publique ou sous sa direction, devraient fournir à leurs utilisateurs suffisamment d'informations pour que ces derniers puissent contester les décisions en question.

Règles relatives à la responsabilisation des plateformes et à l'autonomisation des utilisateurs

57. L'existence de cadres législatifs applicables à la responsabilisation des plateformes et à l'autonomisation des utilisateurs est essentielle pour promouvoir un environnement en ligne favorable, propice à l'exercice de la liberté d'expression. Obliger les plateformes à favoriser l'autonomie de leurs utilisateurs constitue un élément clé des lois relatives à la sécurité en ligne et à la responsabilisation des plateformes. La réglementation de la sécurité en ligne devrait poursuivre comme objectif clé de permettre aux utilisateurs d'évoluer en toute confiance dans l'environnement en ligne, d'en négocier efficacement les risques et d'avoir la maîtrise de leurs expériences en ligne.

58. La législation relative à la responsabilisation des plateformes et à l'autonomisation des utilisateurs devrait se concentrer sur les devoirs et responsabilités systémiques que les plateformes devraient assumer pour créer un environnement en ligne favorable. Ces devoirs et responsabilités devraient porter sur les systèmes et les processus des plateformes, notamment la conception globale, les procédures, la gouvernance et le fonctionnement des services offerts sur les plateformes, en suivant une approche graduée en fonction de la taille, de la portée et de l'impact de la plateforme. Ils devraient couvrir les questions de transparence, les systèmes et processus de sélection, d'organisation et de modération des contenus, la conception pratique des plateformes, la gestion des risques et la gouvernance de l'entreprise.

59. Les États devraient exiger des plateformes qu'elles élaborent leurs services de manière à y intégrer, par défaut et dès la conception, les considérations de sécurité des utilisateurs, en respectant le droit à la liberté d'expression et la nécessité d'un environnement informationnel pluriel. Les interfaces des plateformes et leurs algorithmes de sélection, d'organisation et de modération de contenus influent fortement sur l'expérience et le comportement des utilisateurs et des créateurs de contenu. Par conséquent, leur conception devrait s'accompagner de mesures d'atténuation des risques visant à éviter que des contenus et des comportements qui risquent clairement d'entraîner des préjudices ne se trouvent amplifiés.

60. Les États devraient exiger des plateformes exerçant une influence significative qu'elles réalisent une évaluation des risques liés aux décisions qu'elles prennent concernant la conception, le fonctionnement et l'utilisation de leurs services, y compris leurs politiques et règles contractuelles, afin d'étudier de près les effets de leurs services sur les droits humains et la démocratie. Si l'évaluation conclut que les interventions proposées entraînent des risques, les interventions devraient intégrer des mesures concrètes destinées à

atténuer ces risques et les plateformes devraient être tenues de déployer ces mesures avant de procéder à la mise en œuvre des changements.

61. Lorsqu'elles réalisent des évaluations de risques, les plateformes exerçant une influence significative devraient prendre l'initiative de consulter les parties prenantes concernées et offrir des possibilités de participation au grand public en temps utile et en continu. Les résultats de ces consultations devraient être dûment pris en compte.

62. Les États devraient exiger des plateformes exerçant une influence significative qu'elles produisent et publient des documents sur leurs évaluations des risques et des effets sur les droits humains et la démocratie, afin que le public puisse vérifier que ces plateformes prennent en compte les risques de préjudice associés à leurs services et agissent en conséquence. Certaines informations peuvent être tues ou modifiées si cela s'avère nécessaire à la poursuite d'un but légitime, en particulier lorsque leur publication nuirait à la sécurité des utilisateurs, en tenant compte de l'intérêt que leur divulgation présente pour le public.

63. Les États devraient soutenir le public, les experts et les chercheurs dans leurs efforts pour identifier les risques à traiter en priorité sur leurs différents services en ligne et pour trouver des stratégies efficaces d'atténuation des risques. Les États devraient œuvrer activement à donner au public les moyens d'examiner les évaluations des risques et les mesures d'atténuation, d'y contribuer et de commenter la manière dont elles sont mises en œuvre dans les règles et réglementations concernées.

64. Les États devraient assurer, en droit comme en pratique, l'indépendance des autorités de régulation chargées de superviser, de mettre en œuvre ou de faire respecter le cadre législatif applicable à la responsabilisation des plateformes et à l'autonomisation des utilisateurs, et veiller à ce que ces autorités disposent de ressources et compétences adéquates, travaillent sur la base d'informations factuelles et pèsent toujours avec soin les types de risques qu'elles supervisent. Les plateformes devraient être tenues par la loi de fournir en temps utile les informations nécessaires à ces activités de contrôle, sous peine de sanctions proportionnées en cas de manquement à ces obligations. Les décisions de ces autorités de régulation devraient être susceptibles d'un recours juridictionnel.

65. Les États devraient s'attacher, dans l'intérêt du public, à partager les responsabilités en matière de sécurité en ligne en y associant des acteurs non étatiques, tels que des chercheurs, des groupes d'utilisateurs, des professionnels du signalement et de la notification, des organes extrajudiciaires de règlement des différends et des tiers œuvrant à l'étiquetage des contenus. Ces acteurs non étatiques devraient être soumis à des mécanismes de transparence et de responsabilisation. Les États devraient étudier les moyens de les inciter à participer et de récompenser leurs efforts par le biais de mesures ne dépendant pas de considérations politiques ou commerciales.

V. Mesures d'autonomisation en ligne des utilisateurs

Dispositions générales

66. Les États devraient adopter des cadres législatifs fondés sur des informations factuelles imposant aux plateformes les obligations énoncées ci-dessous en matière d'autonomisation. L'étendue et les conditions de ces obligations devraient être affinées en fonction des informations factuelles qui devraient être analysées au moyen d'un processus consultatif inclusif. Les principales obligations des plateformes visant à renforcer l'autonomisation sont les suivantes :

a. obligations relatives à la conception des services, comme le fait d'assurer aux utilisateurs le droit de personnaliser leur expérience en ligne, de désactiver certains types de recommandations, de masquer certains types de contenus, de bloquer d'autres utilisateurs ou d'activer l'étiquetage des contenus par des tiers ;

b. obligations relatives à la transparence, comme celles garantissant aux utilisateurs le droit de comprendre le fonctionnement général des systèmes de sélection, d'organisation et de modération des contenus, ainsi que des mécanismes de leur monétisation, et d'inspecter la provenance des publicités, ainsi que le droit des chercheurs à étudier la plateforme et à s'appuyer sur elle pour étudier des phénomènes sociaux ;

c. obligations de procédure équitable, comme celles qui garantissent aux utilisateurs le droit d'être avertis des décisions concernant la modération des contenus, d'en comprendre les bases légales ou contractuelles et de contester les décisions de modération de contenus prises par la plateforme ;

d. obligations relatives aux actions collectives, permettant aux utilisateurs de signaler individuellement ou collectivement des violations des politiques et règles contractuelles, d'envoyer des notifications concernant les contenus restreints par la loi et de bénéficier d'une représentation professionnelle.

67. La gestion des risques en ligne devrait toujours englober l'étude des possibilités d'y répondre (entièrement ou comme un moyen parmi d'autres) en donnant aux utilisateurs davantage de pouvoir sur leur expérience en ligne. Il convient d'encourager et, lorsque nécessaire, de requérir l'exercice par les utilisateurs de différents types de contrôle sur l'organisation, la sélection et la modération des contenus et des comportements, comme l'ajustement du champ des contenus légaux pouvant être priorisés ou restreints par des utilisateurs spécifiques, y compris avec le soutien de tiers de leur choix. Cette approche de l'organisation, de la sélection et de la modération de contenus centrée sur les utilisateurs, qui leur délègue des pouvoirs sur la gestion des risques en ligne, ne devrait cependant pas leur en conférer la responsabilité.

68. Les obligations d'autonomisation imposées aux plateformes devraient être proportionnées et suivre une approche graduée selon leur taille, leur portée ou leur impact, d'une part pour ne pas surcharger les petits ou très petits acteurs, d'autre part pour tenir compte des responsabilités accrues qui reviennent aux plateformes exerçant une influence significative.

69. Sauf mention contraire, les obligations d'autonomisation formulées dans cette partie s'appliquent par principe à toutes les plateformes. Toutefois, leurs modalités de mise en œuvre devraient être ajustées en fonction de la nature de chaque plateforme et de sa taille, de sa portée ou de son impact. Pour les obligations d'autonomisation indiquées aux paragraphes 71, 72, 86, 91 et 92, il convient de prévoir des exemptions pour les petites et très petites plateformes. Conformément à leurs responsabilités telles qu'énoncées au paragraphe 26, les plateformes sont encouragées à mettre en œuvre des mesures d'autonomisation même si les cadres juridiques en vigueur ne les y obligent pas.

Autonomisation dès la conception

70. Les plateformes devraient concevoir leurs services de manière à conférer aux utilisateurs le plus de moyens d'action possible.

71. La conception des systèmes automatisés d'organisation, de sélection et de modération des contenus légaux devrait permettre aux utilisateurs de personnaliser leur expérience en ligne en fonction de leurs préférences. Pour cela, des outils conviviaux et faciles d'accès devraient leur permettre de désactiver certains types de recommandations, de masquer certains types de contenus et de bloquer d'autres utilisateurs.

72. La conception de chaque plateforme devrait renforcer la capacité des utilisateurs à opérer des choix éclairés quant aux contenus qu'ils consultent. Cela pourrait impliquer le fait de faciliter l'étiquetage des contenus légaux par des tiers, qu'il s'agisse d'experts, de vérificateurs de faits ou de communautés. Les utilisateurs devraient alors avoir la possibilité d'utiliser cet étiquetage, dans leurs paramètres, pour affiner la personnalisation de leur expérience en ligne en masquant ou en priorisant les contenus correspondant à certaines étiquettes.

73. Les États devraient explorer des options visant à ouvrir les plateformes exerçant une influence significative aux outils développés par des tiers, que les utilisateurs devraient pouvoir choisir afin de personnaliser leur expérience en ligne dans la gestion de contenus légaux, y compris leur organisation, leur sélection et leur modération, sous réserve d'un niveau approprié de responsabilisation et d'autonomisation des utilisateurs, tel que prévu dans les présents principes. Une telle possibilité devrait être prise en compte sans préjudice des mesures prises par les États conformément au paragraphe 52. Afin de faciliter les processus de responsabilisation, la fourniture de tels outils peut être soumise à une condition de certification.

74. La conception des plateformes devrait activement promouvoir l'autonomisation des utilisateurs et la sécurité en ligne des personnes présentant des incapacités, y compris à l'aide d'outils fournis par des tiers. Cela suppose de veiller à ce que ces personnes puissent utiliser et mettre en œuvre ces outils pour surmonter les barrières à l'accessibilité qui les empêchent de bénéficier des mesures de sécurité et d'autonomisation.

75. Outre les autres mesures appropriées d'atténuation des risques qui peuvent être prises par les plateformes et conformément à la Recommandation CM/Rec(2018)7 sur les Lignes directrices relatives au respect, à la protection et à la réalisation des droits de l'enfant dans l'environnement numérique, les États devraient exiger le recours à des systèmes effectifs de contrôle de l'âge, afin d'assurer la protection des enfants contre les produits, services et contenus de l'environnement numérique qui sont restreints par la loi sur des critères d'âge. Il convient, en particulier, d'imposer de tels systèmes aux plateformes qui proposent principalement des services ou des contenus restreints par la loi dans l'intérêt des enfants. Ces systèmes devraient respecter

les droits humains et utiliser des méthodes qui respectent la liberté d'expression et la protection des données à caractère personnel et de la vie privée, conformément à l'intérêt supérieur de l'enfant. Les États, lorsqu'ils exigent la mise en œuvre de tels systèmes, devraient prévoir des garanties pour s'assurer que ces systèmes n'aboutissent pas à, de manière disproportionnée, exclure les enfants des espaces en ligne et à restreindre leur droit de participer aux débats sur des questions d'intérêt général. Des garanties devraient également être prévues pour s'assurer que ces systèmes n'excluent pas des espaces en ligne, ou n'excluent pas encore davantage, les personnes se trouvant en situation vulnérable et exposées à la discrimination.

76. Les États devraient exiger des plateformes qu'elles développent, produisent et mettent régulièrement à jour d'autres outils efficaces et adaptés à l'âge afin d'atténuer les risques pour les enfants dans l'environnement en ligne. Ces outils, à l'intention des enfants ou des parents selon qu'il convient, devraient accorder une considération primordiale à l'intérêt supérieur de l'enfant et être élaborés et déployés en tenant compte du développement des capacités de l'enfant, conformément à son âge et à son degré de maturité. Ils ne devraient avoir pour effet ni de renforcer des attitudes discriminatoires, ni de porter atteinte au droit des enfants à la vie privée ou à leur intérêt supérieur, ni de les priver de leur droit à la liberté d'expression.

77. Les plateformes exerçant une influence significative devraient veiller à ce que les créateurs de contenu disposent de l'option d'étiqueter leurs contenus selon l'âge du public visé, que ce soit par choix ou parce qu'ils y sont juridiquement tenus.

78. Le déploiement de systèmes et d'outils visant à protéger et à autonomiser les enfants en ligne ne devrait pas être conçu comme un transfert des responsabilités en matière de sécurité en ligne des plateformes vers les parents, les enfants ou les créateurs de contenu, ni aboutir à un tel transfert.

79. Les plateformes ne devraient pas entraver la capacité de leurs utilisateurs, y compris les créateurs de contenu, de transférer leurs profils en ligne vers d'autres plateformes complémentaires ou concurrentes.

Transparence

80. Les plateformes devraient fournir des explications claires et utiles sur la conception de leurs systèmes de sélection et d'organisation des contenus, afin que leurs utilisateurs puissent comprendre quelles informations sont mises en avant ou, au contraire, reléguées.

81. Les plateformes devraient dévoiler au public des informations détaillées sur leurs pratiques de modération des contenus et publier périodiquement des statistiques globales sur le nombre et les types de décisions prises en matière de modération de contenu sur leurs services.

82. Les plateformes devraient décrire les systèmes automatisés sur lesquels se fonde leur modération des contenus, préciser dans quels domaines ces systèmes s'appliquent et publier des rapports qualitatifs sur la précision de ces outils et les garanties qui les entourent.

83. Les plateformes devraient faire preuve de transparence sur l'identité de leurs annonceurs, leur recours aux techniques de ciblage publicitaire et le coût de chaque annonce.

84. Les plateformes devraient faire preuve de transparence sur la monétisation des contenus publiés par des utilisateurs sur leurs services : quels contenus sont monétisés, par qui et sur quels grands principes repose l'attribution de ressources aux créateurs de contenu.

85. Les plateformes exerçant une influence significative devraient fournir aux créateurs de contenu des outils leur permettant d'assurer la transparence sur la manière dont leurs contenus sont monétisés.

86. Les chercheurs indépendants devraient avoir un accès effectif aux données détenues par les plateformes, sans discrimination et de manière sécurisée, légale et conforme à la protection de la vie privée, afin de mener, de manière éthique et responsable, des recherches poursuivant un intérêt public. Cela inclut notamment les recherches portant sur l'étude des phénomènes sociaux ainsi que sur la nature et l'impact des risques, et sur l'efficacité des différentes stratégies visant à les atténuer. Lorsque l'accès à des données personnelles, individuelles ou confidentielles est nécessaire à la réalisation de la recherche, cet accès devrait être accordé aux chercheurs qui ont préalablement fait l'objet d'une vérification par un organisme indépendant externe. Des garanties solides pour la protection de la vie privée et des données à caractère personnel devraient être mises en place, conformément à la partie 6 de la Recommandation CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression.

87. Les chercheurs indépendants devraient avoir la possibilité technique et juridique, sans discrimination, d'utiliser les plateformes pour mener leurs travaux de recherche, dès lors que ces derniers obéissent aux

principes établis en matière de déontologie et d'intégrité de la recherche, et dans le respect des garanties requises en matière de protection des données.

Droits procéduraux

88. Les plateformes devraient énoncer clairement et sans ambiguïté les politiques et les règles contractuelles en vertu desquelles elles proposent leurs prestations au public. Les utilisateurs concernés devraient être informés à l'avance de toute modification importante apportée à ces règles. Les politiques et les règles contractuelles et toutes leurs modifications devraient être dûment expliquées aux utilisateurs concernés, en des termes accessibles leur permettant de comprendre à la fois les règles elles-mêmes et les répercussions que les changements vont avoir sur leur activité future. Les règles applicables aux contenus et aux comportements des utilisateurs et les modifications de ces règles devraient être assez prévisibles pour éviter que les éventuelles sanctions ne revêtent un caractère arbitraire. Les politiques et règles contractuelles applicables aux enfants devraient être expliquées d'une manière compréhensible pour les enfants et adaptées à leur âge, à leur degré de maturité et à leur situation.

89. Les utilisateurs et les créateurs de contenu devraient être capables de contester effectivement les décisions de modération de contenus prises par les plateformes, qui affectent leur droit à la liberté d'expression, y compris le droit de recevoir des informations, ou d'autres droits.

90. Les décisions prises par les plateformes en matière de modération de contenus qui ciblent ou affectent des utilisateurs en particulier, dont celles imposées à la visibilité des contenus des utilisateurs, à leur monétisation et aux privilèges associés à leur compte, devraient être promptement notifiées et expliquées aux utilisateurs ayant créé les contenus ou les comptes touchés par la décision. La notification doit préciser les motifs de la décision, expliquer par quel processus elle a été prise et indiquer toutes les possibilités de recours en termes clairs, non techniques et adaptés à l'âge des personnes concernées.

91. Les plateformes devraient également adresser une notification à tout autre utilisateur identifiable directement concerné ou touché par ces contenus et ayant choisi de recevoir ce type de notification.

92. Les décisions des plateformes en matière de modération des contenus ayant des effets importants sur les utilisateurs devraient s'accompagner de voies de recours extérieures et indépendantes, telles que des organes extrajudiciaires de règlement des différends ou toute autre forme de contrôle indépendant. Ces dispositifs devraient être accessibles, transparents, permettre d'examiner les plaintes rapidement et selon un calendrier transparent, et déboucher sur un recours effectif. Ils ne devraient empêcher ni les utilisateurs ni les plateformes d'opter pour d'autres voies de recours, judiciaires ou autres, prévues par le droit interne.

Actions collectives d'utilisateurs

93. Les plateformes devraient veiller à ce que les utilisateurs puissent signaler facilement les atteintes par d'autres utilisateurs aux politiques et aux règles contractuelles et offrir aux auteurs de signalement un retour d'informations suffisant sur les suites données à leur démarche.

94. Les plateformes devraient veiller à ce que les utilisateurs puissent soumettre une notification concernant des contenus susceptibles d'être restreints par la loi, et offrir aux auteurs de notifications un retour d'informations suffisant sur les suites données à leur démarche. Si la décision de la plateforme en matière de modération de contenus n'est pas satisfaisante, les auteurs de notifications devraient avoir la possibilité de la contester, y inclus conformément au paragraphe 92.

95. Les États devraient encourager l'identification de professionnels susceptibles, en qualité d'experts indépendants, d'envoyer des notifications concernant les contenus restreints par la loi ou de signaler les violations des politiques et des règles contractuelles sur les plateformes. Les États devraient favoriser la reconnaissance de ces professionnels en leur accordant certains privilèges, tels qu'un traitement prioritaire de leurs signalements, notifications et recours, un soutien financier ou un meilleur accès aux interfaces techniques.

96. Les États devraient encourager la création de groupes d'utilisateurs professionnels capables d'agir en tant qu'experts indépendants pour défendre les intérêts des utilisateurs et des créateurs de contenu auprès des plateformes ou des autorités publiques. Les États peuvent octroyer à ces groupes d'utilisateurs certains privilèges, comme un traitement prioritaire de leurs recours, des financements, un meilleur accès aux interfaces techniques ou le droit d'action collective contre les atteintes aux droits des utilisateurs.

Exposé des motifs

PRÉAMBULE

1. Cette recommandation vise à fournir aux États membres un cadre cohérent et applicable dans la pratique pour promouvoir la sécurité en ligne et l'autonomisation des utilisateurs et des créateurs de contenu.

2. L'environnement en ligne constitue aujourd'hui un espace public essentiel, qui permet l'échange d'informations et d'idées, la communication, la participation civique, l'expression culturelle et l'activité économique, pour ne citer que quelques activités. Cependant, s'il ne fait aucun doute que l'environnement en ligne favorise considérablement l'exercice du droit à la liberté d'expression ainsi que la jouissance d'autres droits, il comporte également des risques importants, tels que le harcèlement, le discours de haine, l'exposition à la désinformation et à la mésinformation, ainsi que les biais algorithmiques, qui peuvent porter atteinte aux droits humains, affecter l'intégrité de l'information et dissuader la participation. À bien des égards, l'environnement en ligne reflète la vie quotidienne : il offre des espaces d'apprentissage, de créativité et de connexion, mais expose également les individus à des risques et à des préjudices qui doivent être gérés si l'on veut que la participation demeure sûre et significative.

3. La recommandation réaffirme l'engagement du Conseil de l'Europe à lutter contre ces risques sans sacrifier l'ouverture et l'accessibilité de l'internet. Elle souligne qu'une approche fondée sur les droits humains, ayant pour principes fondamentaux la légalité, la nécessité et la proportionnalité, et pour axe central la sécurité par l'autonomisation, est essentielle pour garantir que les mesures prises pour la sécurité des utilisateurs et des créateurs de contenu ne restreignent pas involontairement de manière ou disproportionnée la liberté d'expression et d'autres droits.

4. Fondés sur la Convention de sauvegarde des droits de l'homme et des libertés fondamentales (Convention européenne des droits de l'homme, la Convention), les principes énoncés dans l'annexe à la recommandation s'appuient sur la jurisprudence de la Cour européenne des droits de l'homme (la Cour) relative à la jouissance et à l'exercice des droits humains dans l'environnement en ligne. Ils soulignent la nécessité pour les États de créer un internet libre, ouvert et accessible à tous, ainsi qu'un environnement en ligne propice où les utilisateurs peuvent jouir de leurs droits sans discrimination. La Cour a de plus en plus abordé les responsabilités des États à cet égard, soulignant leur obligation de protéger les utilisateurs contre les ingérences préjudiciables des autorités publiques et des entités privées, ainsi que leur obligation de s'abstenir d'imposer des restrictions aux droits autres que celles qui sont prescrites par la loi et nécessaires dans une société démocratique pour atteindre un but légitime. La Cour a également précisé que les entités privées peuvent, dans certaines circonstances, être tenues responsables des contenus illégaux publiés par les utilisateurs et a souligné les devoirs et responsabilités qui accompagnent l'exercice du droit à la liberté d'expression (voir, par exemple, *Delfi AS c. Estonie* [GC], n° 64569/09, 16 juin 2015).

5. Le préambule rappelle un principe clé de la jurisprudence de la Cour : le droit à la liberté d'expression protège non seulement les propos inoffensifs ou neutres, mais aussi ceux qui peuvent « choquer, heurter ou déranger » (*Handyside c. Royaume-Uni*, n° 5493/72, 7 décembre 1976, par. 49). Ce principe est particulièrement important dans le contexte de l'expression en ligne, où il existe une grande diversité de voix et où les discours polarisés ou provocateurs sont fréquents. La Cour a longtemps soutenu qu'une protection solide des opinions dissidentes ou minoritaires est essentielle dans une société démocratique, et que cela s'applique tant dans l'environnement physique qu'en ligne, y compris pour les contenus générés par les utilisateurs, le travail journalistique et l'expression politique (*Delfi AS c. Estonie* [GC], précité, paragraphes 131 à 139). Si la Cour a

observé qu'un certain degré d'insultes vulgaires est courant dans de nombreux forums en ligne, elle a souligné que cela doit être compris dans son contexte et qu'un certain niveau de tolérance est attendu, en particulier de la part des personnalités publiques telles que les politiciens (voir par exemple *Tamiz c. Royaume-Uni* (déc.), n° 3877/14, 19 septembre 2017, par. 81, disponible seulement en anglais, dans laquelle le requérant, un homme politique, avait lui-même initié l'utilisation d'un langage vulgaire). La recommandation indique donc qu'une protection solide de la liberté d'expression est un élément essentiel de toute réglementation visant à promouvoir la sécurité en ligne.

6. En même temps, le préambule reconnaît l'existence de risques dans l'environnement en ligne et les préjudices potentiels qui peuvent en résulter, parfois graves, pour la sécurité des utilisateurs, la jouissance des droits, le fonctionnement de la démocratie et d'autres intérêts sociétaux. Ces risques existent pour tous. Cependant, certaines personnes et certains groupes peuvent être plus exposés que d'autres en raison de leur identité, de leur rôle ou de leur propre contribution au débat public. Cela inclut, sans s'y limiter, les femmes et les filles, les enfants, les personnes en situation de vulnérabilité et exposés à la discrimination, notamment les personnes handicapées, les minorités ethniques, linguistiques et religieuses nationales, les communautés LGBTI[1], ainsi que les migrants et les personnes issues de l'immigration. Toute personne perçue comme appartenant à ces groupes peut être confrontés à des abus ciblés, y compris de nature intersectionnelle, à une discrimination structurelle ou à une exclusion algorithmique, ce qui limite leur capacité à exercer leurs droits en ligne. En particulier, la violence à l'égard des femmes et des filles facilitée par la technologie est un problème croissant, avec une prévalence mondiale estimée à environ 85 %, les femmes étant 27 fois plus susceptibles d'en être victimes que les hommes. Ces risques sont plus importants pour les créateurs de contenu (y compris les médias, les activistes et les ONG), qui peuvent être victimes d'abus visant à les empêcher de s'exprimer. Ces risques menacent non seulement les individus, mais aussi les groupes, la jouissance plus large des droits dans la société et, en fin de compte, la démocratie.

7. La recommandation préconise la mise en place de cadres juridiques transparents et fondés sur des données probantes, ainsi que d'autres initiatives visant à garantir que les risques en ligne, ainsi que les préjudices potentiels qui en découlent, soient évalués, traités et atténués de manière non discriminatoire et conforme aux droits humains, afin de prévenir toute ingérence disproportionnée dans la liberté d'expression et d'autres droits humains. Elle souligne la nécessité que ces mesures d'évaluation et d'atténuation des risques soient prises en consultation avec divers utilisateurs, notamment les créateurs de contenu, les groupes et communautés concernés, les plateformes et autres entreprises, ainsi que toutes les autres parties prenantes concernées. Les évaluations et les mesures d'atténuation doivent être adaptées et intersectionnelles, afin de garantir que tous les utilisateurs puissent jouir de leurs droits en ligne en toute sécurité.

8. Il est essentiel de veiller à ce que les utilisateurs soient effectivement équipés pour comprendre, naviguer et réagir aux risques en ligne. L'autonomisation est essentielle à cet égard. L'autonomisation repose sur la dignité humaine et l'autonomie des utilisateurs et contribue à garantir un accès équitable aux technologies en ligne, permettant la pleine jouissance des droits humains dans l'environnement en ligne et favorisant la participation inclusive de tous aux espaces en ligne. L'autonomisation ne concerne pas seulement la protection, mais aussi la possibilité pour les utilisateurs de s'engager de manière significative en ligne. Cela inclut l'accès aux outils en ligne, l'alphabétisation numérique, la représentation et la capacité à participer aux processus de gouvernance. La réalisation de l'autonomisation nécessite un cadre réglementaire fondé sur les droits, dans lequel la sécurité des utilisateurs, l'autonomisation et la responsabilité systémique sont légalement obligatoires et soumises à un contrôle indépendant.

9. Tous les risques ne peuvent pas être traités par l'autonomisation, et la charge de traiter les risques et les dommages potentiels ne devrait pas peser principalement sur ceux qui y sont les plus exposés. Chaque fois qu'il est établi que l'autonomisation échoue ou est susceptible d'échouer, afin d'atténuer les effets néfastes des risques en ligne, les États devraient envisager d'autres moyens proportionnés pour traiter les préjudices découlant des risques en ligne, y compris l'imposition de restrictions proportionnées aux contenus ou à leur accessibilité sur les plateformes. Conformément à l'article 10, paragraphe 2, de la Convention, toute limitation du droit à la liberté d'expression doit être « prévue par la loi » et « nécessaire dans une société démocratique » pour la protection d'un but légitime. Des mesures trop générales ou trop vagues risquent d'étouffer le discours légitime. L'article 17 de la Convention (Interdiction de l'abus de droit) autorise également des

1. L'acronyme LGBTI désigne les personnes lesbiennes, gays, bisexuelles, transgenres et intersexes, selon le sens ordinaire donné à ces termes dans les standards juridiques et les documents pertinents du Conseil de l'Europe. Voir, en particulier, la Recommandation CM/Rec(2010)5 relative aux mesures de lutte contre la discrimination fondée sur l'orientation sexuelle ou l'identité de genre, et la Recommandation CM/Rec(2025)7 relative à l'égalité des droits pour les personnes intersexes, ainsi que la Commission européenne contre le racisme et l'intolérance (ECRI), Recommandation de politique générale n° 17 sur la prévention et la lutte contre l'intolérance et la discrimination envers les personnes LGBTI, disponible à l'adresse https://go.coe.int/8gURb.

restrictions à certaines expressions en privant de la protection offerte par l'article 10 les activités expressives qui détournent le droit à la liberté d'expression de son objectif réel en l'invoquant pour justifier, promouvoir ou accomplir des actes contraires au texte et à l'esprit de la Convention ou incompatibles avec la démocratie ou d'autres valeurs fondamentales de la Convention. Selon les affaires examinées jusqu'à présent par la Cour, celles-ci peuvent comprendre, selon les circonstances spécifiques, l'incitation à la violence et à la haine, la promotion et la justification du terrorisme et des crimes de guerre, la promotion d'idéologies totalitaires et la négation de l'Holocauste. La Cour a plus généralement souligné que « certaines catégories de discours, tels que les propos obscènes et indécents, ne jouent aucun rôle essentiel dans l'expression des idées » (*Rujak c. Croatie* (déc.), n° 57942/10, 2 octobre 2012, par. 29) et ne relèvent donc pas du champ d'application de l'article 10 de la Convention. Un raisonnement similaire pourrait s'appliquer à la diffusion de contenus qui n'ont aucun rapport raisonnable avec l'expression d'idées, tels que les matériels d'abus sexuels sur enfants ou les contenus sexuels privés non consensuels partagés dans le but de causer une détresse à une personne. Ces restrictions doivent toutefois être strictement encadrées et ne s'appliquer qu'aux propos clairement incompatibles avec le système de la Convention lui-même. La présente recommandation conseille donc aux États d'éviter les restrictions générales et de veiller à ce que les normes juridiques soient clairement définies, proportionnées et soumises à un contrôle judiciaire.

10. Les principes énoncés dans l'annexe à la recommandation reconnaissent et traitent du rôle central des plateformes en ligne, en particulier celles exerçant une influence significative et qui hébergent et régulent une grande partie de la vie publique en ligne (voir l'exposé des motifs concernant le paragraphe 11, sur la définition de « plateforme exerçant une influence significative »). Ils soulignent que ces acteurs ont la responsabilité de respecter les droits humains et de créer un environnement en ligne propice, garantissant la sécurité de leurs utilisateurs, et qu'ils ne doivent pas opérer dans un vide réglementaire. Ce principe est inscrit dans les Recommandations CM/Rec(2016)3 sur les entreprises et les droits de l'homme et CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, ainsi que dans les Principes directeurs des Nations Unies relatifs aux entreprises et aux droits de l'homme[2]. Dans le contexte de la sécurité en ligne, les plateformes ont la responsabilité de ne pas contribuer aux violations des droits humains, notamment en amplifiant les contenus présentant un risque de préjudice, d'exercer leurs activités avec la diligence requise et de garantir l'accès à des recours efficaces. Elles doivent également réaliser des évaluations d'impact sur les droits humains et mettre en place des mécanismes de responsabilisation, de redressement et d'autonomisation des utilisateurs.

11. Le préambule souligne que les mesures prises par le biais de l'organisation, de la sélection et de la modération des contenus constituent des ingérences dans la jouissance de la liberté d'expression et d'information et d'autres droits et peuvent affecter de manière disproportionnée l'exercice de ces droits. Conformément à la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, les Principes énoncés dans l'annexe soulignent la nécessité de mettre en place des mécanismes de responsabilisation, de transparence et de recours conformes aux normes relatives aux droits humains. Dans le même temps, ils reconnaissent la nécessité de prendre des mesures pour protéger les personnes qui risquent d'être réduites au silence par des contenus susceptibles de causer un préjudice.

12. Le préambule souligne la vive préoccupation suscitée par la concentration du pouvoir entre les mains de quelques plateformes en ligne, l'asymétrie de pouvoir entre ces plateformes et leurs utilisateurs, et les implications de cette dynamique pour la sécurité des utilisateurs, les droits humains des utilisateurs et les processus et institutions démocratiques. Conformément à la Recommandation CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication, le préambule souligne la nécessité d'une approche graduée de la réglementation, garantissant que les plateformes exerçant une influence significative soient soumises à des obligations renforcées en matière de sécurité et d'autonomisation des utilisateurs, tout en mettant en garde contre des exigences disproportionnées pour les fournisseurs qui n'ont pas un tel impact.

2. Nations Unies, *Principes directeurs relatifs aux entreprises et aux droits de l'homme : mise en œuvre du cadre « Protéger, respecter et réparer » des Nations Unies,* doc. ONU HR/PUB/11/04, Nations Unies, New York et Genève, 2011, ci-après « Principes directeurs des Nations Unies relatifs aux entreprises et aux droits de l'homme », disponibles à l'adresse https://digitallibrary.un.org/record/720245?ln=fr&v=pdf, approuvés par le Conseil des droits de l'homme des Nations Unies, résolution 17/4, 16 juin 2011, doc. ONU A/HRC/RES/17/4, disponible à l'adresse https://ap.ohchr.org/documents/dpage_f.aspx?si=A/HRC/RES/17/4.

ANNEXE À LA RECOMMANDATION CM/REC(2026)4 – PRINCIPES POUR LA SÉCURITÉ ET L'AUTONOMISATION EN LIGNE DES UTILISATEURS ET DES CRÉATEURS DE CONTENU

I. Raison d'être, champ d'application et définitions

Raison d'être

Concernant les paragraphes 1 à 5

13. Le paragraphe 1 de l'annexe reconnaît que l'internet offre à la fois des possibilités en matière de liberté d'expression, y compris l'accès à l'information, et des risques pour la sécurité des utilisateurs, car des contenus susceptibles de nuire aux individus et à la société peuvent être diffusés à une échelle et à une vitesse sans précédent. Il reconnaît que les plateformes en ligne, en particulier certaines d'entre elles qui ont une portée considérable, fonctionnant grâce à l'organisation algorithmique de contenu, ont acquis un rôle central tant dans l'offre de ces possibilités que dans la contribution à ces risques.

14. Les paragraphes 2 et 3 précisent que la sécurité en ligne ne peut être comprise comme étant séparée ou isolée de la société et de la protection et de la promotion des droits humains. La sécurité en ligne fait partie d'un concept plus large d'environnement en ligne favorable aux droits humains. Elle ne peut être assimilée à l'absence de risques en ligne ou liés aux activités en ligne, car cela, même si cela était possible, pourrait se faire au détriment d'autres droits humains et des libertés fondamentales. La recommandation part du principe que la sécurité n'est pas un concept statique et immuable. Elle varie considérablement dans le temps et dans l'espace, car elle peut être influencée à la fois par le développement technologique et les valeurs sociétales. En outre, la sécurité est inextricablement liée aux besoins spécifiques de certains secteurs de la société, tels que les enfants, ainsi qu'aux perceptions et préférences des individus concernant leur propre niveau acceptable de risques et la nécessité connexe de ne pas y être exposés. À ce titre, la sécurité en ligne ne diffère pas de la sécurité hors ligne et doit être comprise dans une double dimension. Les mesures prises par les États pour protéger la sécurité en ligne devraient établir quel est le niveau acceptable de risques en ligne dans une société démocratique et quelles sont les mesures de protection appropriées pour y faire face, soit pour le grand public, soit pour des secteurs spécifiques de la population, sans interférer de manière disproportionnée avec l'exercice des droits humains. En outre, ces interventions devraient viser à donner aux individus la capacité de reconnaître et de comprendre les risques, ainsi que les outils nécessaires pour contrôler leur expérience en ligne d'une manière qui s'adapte à leurs préférences et à leurs choix, et leur permettre d'atteindre le niveau de bien-être qu'ils souhaitent. Garantir la sécurité dans l'environnement en ligne nécessite également de s'attaquer aux inégalités et à la discrimination structurelles et intersectionnelles, notamment celles fondées sur le genre, car celles-ci peuvent influencer à la fois l'exposition aux risques en ligne et la capacité à rechercher une protection et un recours.

15. Des mesures visant à réduire la disponibilité, l'accessibilité et la visibilité de certains contenus en ligne peuvent être nécessaires dans une société démocratique qui assure une protection adéquate des droits humains, à condition qu'elles soient clairement définies par la loi, qu'elles répondent à un besoin social urgent et qu'elles soient proportionnées au but légitime poursuivi. Les paragraphes 4 et 5 reconnaissent qu'une approche de la sécurité en ligne fondée exclusivement ou principalement sur ce type de mesures ne peut à elle seule favoriser un environnement en ligne favorable. Un recours excessif à ces mesures est insuffisant et potentiellement préjudiciable, car il peut conduire à des restrictions arbitraires ou disproportionnées qui portent atteinte aux droits humains, en particulier à la liberté d'expression et à la vie privée. Conformément à l'approche générale adoptée dans certaines juridictions en Europe et au-delà[3], la recommandation préconise donc l'élaboration d'une nouvelle génération d'instruments juridiques proportionnés et fondés sur des données probantes, complémentaires aux approches existantes, qui renforcent la responsabilisation et le contrôle public des plateformes en ce qui concerne leurs choix de conception et leurs pratiques générales de gestion des risques, tout en renforçant l'autonomisation des utilisateurs et des créateurs de contenu.

3. Voir, Union européenne, *Règlement (UE) 2022/2065 du Parlement européen et du Conseil du 19 octobre 2022 relatif à un marché unique des services numériques et modifiant la directive 2000/31/CE (règlement sur les services numériques)*, disponible à l'adresse http://data.europa.eu/eli/reg/2022/2065/oj, ci-après « Règlement de l'Union européenne sur les services numériques », et Royaume-Uni, *Online Safety Act 2023*, c. 50, disponible à l'adresse www.legislation.gov.uk/ukpga/2023/50 , ci-après « loi sur la sécurité en ligne du Royaume-Uni ».

Champ d'application

Concernant le paragraphe 6

16. Ce paragraphe, en conjonction avec la définition des plateformes, délimite le champ d'application de la recommandation. Les principes énoncés dans son annexe n'ont pour objectif ni de traiter à tous les types de risques en ligne en général, ni d'y apporter des solutions. Ils se concentrent plutôt sur la manière de gérer, dans le respect des droits humains, les risques étroitement liés à l'exercice de la liberté d'expression. Ces risques peuvent être compris de deux manières qui se recoupent : les risques qui découlent de l'exercice de la liberté d'expression et les risques qui ont pour effet d'entraver cet exercice.

17. Les risques résultant de l'exercice de la liberté d'expression surviennent en premier lieu lorsque des personnes font l'objet de représailles pour ce qu'elles disent ou publient. Par exemple, une journaliste peut publier un article d'investigation sur la corruption et recevoir par la suite des menaces de mort, ou voir ses informations personnelles divulguées en ligne. Un autre exemple est celui où des discours de haine ciblés sont diffusés en réponse aux opinions d'une personne, ce qui peut mettre en danger sa sécurité. En outre, des risques découlent également des activités d'expression en ligne qui peuvent porter atteinte aux droits d'autrui, telles que l'incitation à la violence, le discours de haine ou la diffamation, ainsi qu'aux intérêts de la société, tels que l'intégrité de l'information et le processus démocratique.

18. Les risques qui entravent l'exercice de la liberté d'expression sont ceux qui dissuadent les personnes de s'exprimer. Cela inclut ce que la Cour européenne des droits de l'homme a qualifié de « effet dissuasif », ou « *chilling effect* », où la crainte des conséquences réduit au silence un débat légitime. Par exemple, des lois formulées de façon trop générale sur la diffamation ou des sanctions disproportionnées peuvent dissuader les reportages critiques sur des questions d'intérêt public. En outre, le risque d'être victime de violence en ligne peut entraver l'expression publique, par exemple en décourageant les femmes et les filles de participer à des débats publics ou de partager leurs opinions par crainte de harcèlement.

19. De nombreux risques relèvent des deux perspectives. Des systèmes de modération de contenu mal conçus, par exemple, peuvent signaler ou supprimer à tort des contenus légitimes, en particulier ceux provenant de voix marginalisées ou de créateurs abordant des sujets sensibles. Un militant des droits humains qui met en ligne une vidéo sur les violations des droits humains peut voir celle-ci supprimée par un système automatisé qui la qualifie de « contenu violent ». Ces risques résultent à la fois de l'exercice de la liberté d'expression et entravent son exercice futur. La partie II du présent exposé des motifs examine plus en détail les différents types de risques.

Concernant les paragraphes 7-8

20. Les principes énoncés dans l'annexe s'adressent principalement aux États, mais ils traitent également des responsabilités des plateformes. En vertu du droit international des droits humains, les États ont l'obligation et la responsabilité primordiale de garantir et de protéger les droits, mais les acteurs privés, y compris les plateformes, ont également des responsabilités (voir CM/Rec(2016)3 sur les droits de l'homme et les entreprises, CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet et CM/Rec(2022)13 sur les impacts des technologies en ligne sur la liberté d'expression, ainsi que les Principes directeurs des Nations Unies relatifs aux entreprises et aux droits de l'homme). Ces responsabilités peuvent devenir des obligations légales en vertu du droit interne. La Cour européenne des droits de l'homme a reconnu que, si les obligations et les responsabilités des plateformes peuvent différer de celles d'un éditeur traditionnel en ce qui concerne les contenus tiers, « lorsque les intermédiaires d'internet gèrent les contenus disponibles sur leurs plateformes ou jouent un rôle de curateur ou d'éditeur, y compris par le biais d'algorithmes, leur fonction importante dans la facilitation et l'orientation du débat public engendre des obligations de vigilance et de diligence raisonnable, qui peuvent également augmenter proportionnellement en fonction de la portée de l'activité expressive concernée » (*Google LLC et autres c. Russie*, n° 37027/22, 8 juillet 2025, disponible seulement en anglais, par. 79).

21. Conformément à la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, les plateformes sont donc tenues de respecter les droits humains de leurs utilisateurs. Afin de s'acquitter de ces responsabilités, les plateformes sont tenues de prendre des mesures pour protéger la sécurité des utilisateurs et des créateurs de contenu. Ce faisant, elles doivent agir avec transparence, responsabilité et diligence raisonnable ; les mesures qu'elles prennent doivent être conformes aux droits humains, en veillant à ce que leurs systèmes et leurs décisions n'entraînent pas de restrictions injustifiées à la liberté d'expression ou à d'autres droits.

Concernant les paragraphes 9 et 10

22. La recommandation poursuit deux objectifs distincts mais interdépendants : (1) la protection de la sécurité des utilisateurs dans les environnements en ligne, et (2) la promotion de l'autonomisation des utilisateurs, c'est-à-dire qu'ils soient informés et maîtrisent leur environnement en ligne, et qu'ils puissent participer pleinement, librement et de manière égale à la sphère en ligne. La sécurité et l'autonomisation sont deux dimensions qui se renforcent mutuellement dans un environnement en ligne fondé sur les droits. Ces objectifs reflètent un double engagement : permettre aux individus d'exercer leurs droits sans ingérence indue et veiller à ce que les espaces en ligne soient structurés et régis de manière à respecter la dignité humaine, la sécurité et l'inclusion.

23. Ces objectifs sont fermement ancrés dans la Convention, en particulier dans l'article 10, qui protège le droit à la liberté d'expression, l'article 8, qui garantit le droit au respect de la vie privée et familiale, une notion large qui englobe la protection des données à caractère personnel ainsi que l'intégrité physique et psychologique d'une personne, y compris des aspects tels que l'orientation sexuelle et la protection contre les atteintes graves à la réputation, et l'article 14, qui protège le droit de ne pas être victime de discrimination dans la jouissance des droits (voir, par exemple, *Denisov c. Ukraine* [GC], n° 76639/11, 25 septembre 2018, par. 95 ; *Beizaras et Levickas c. Lituanie*, n° 41288/15, 14 janvier 2020, par. 109 ; *Minasyan et autres c. Arménie*, n° 59180/15, 7 janvier 2025, disponible seulement en anglais, par. 53). Dans les cas les plus graves, la violence en ligne peut même entraîner des obligations positives de prévention et de protection en vertu des articles 2 (droit à la vie) et 3 (interdiction de la torture).

24. La Cour européenne des droits de l'homme a toujours considéré que les droits consacrés par la Convention imposent des obligations négatives et positives à l'État. L'obligation négative exige que les États s'abstiennent de toute ingérence injustifiée dans l'exercice des droits. Toute restriction à la liberté d'expression doit remplir les conditions énoncées à l'article 10, paragraphe 2 : elle doit être prévue par la loi, poursuivre l'un des buts légitimes énumérés et être nécessaire dans une société démocratique, ce qui signifie qu'elle doit être proportionnée et répondre à un besoin social impérieux.

25. Les obligations positives impliquent que les États doivent prendre des mesures raisonnables et appropriées pour garantir la jouissance effective des droits consacrés par la Convention, y compris dans les sphères privées et en ligne, ce qui peut les obliger à imposer et à faire respecter des obligations aux acteurs privés. La Cour européenne des droits de l'homme a affirmé que les États doivent garantir un environnement dans lequel les individus peuvent exercer leur droit à la liberté d'expression sans être confrontés à des menaces, des harcèlements ou des violences, qu'ils proviennent des autorités publiques ou d'acteurs privés : « les obligations positives découlant de l'article 10 de la Convention impliquent que les États sont tenus de créer … un environnement favorable à la participation aux débats publics de toutes les personnes concernées, leur permettant d'exprimer sans crainte leurs opinions et idées, même si celles-ci vont à l'encontre de celles défendues par les autorités officielles ou par une partie importante de l'opinion publique, voire même sont irritantes ou choquantes pour ces dernières » (*Khadija Ismayilova c. Azerbaïdjan*, nos 65286/13 et 57270/14, 10 janvier 2019, disponible seulement en anglais, par. 158 (traduction non officielle) ; Voir également *Özgür Gündem c. Turquie*, n° 23144/93, 16 mars 2000, par. 43 et 44 ; *Dink c. Turquie*, n° 2668/07 et autres, 14 septembre 2010, par. 137).

26. L'article 8 impose une obligation similaire. La Cour a souligné que « si l'article 8 tend pour l'essentiel à prémunir l'individu contre des ingérences arbitraires des pouvoirs publics, il ne se contente pas de commander à l'Etat de s'abstenir de pareilles ingérences : à cet engagement négatif peuvent s'ajouter des obligations positives inhérentes à un respect effectif de la vie privée, qui peuvent impliquer l'adoption de mesures visant au respect de la vie privée jusque dans les relations des individus entre eux » (voir *Aksu c. Turquie* [GC], nos. 4149/04 et 41029/04, 15 mars 2012, par. 59 ; et *Minasyan et autres c. Arménie,* précité, par. 58).

27. Dans le contexte de la sécurité en ligne, cette double obligation implique que les États doivent prendre des mesures actives pour répondre de manière appropriée aux contenus présentant un risque de préjudice, notamment les abus, les intimidations et les discriminations susceptibles d'entraver la pleine participation des individus, en particulier ceux appartenant à des catégories à haut risque, telles qu'identifiées au paragraphe 16. Ce faisant, les États doivent toutefois s'abstenir d'imposer des restrictions disproportionnées. Par exemple, les autorités publiques ne devraient pas ordonner la suppression de l'ensemble d'un site web ou d'un domaine internet lorsque seules certaines pages de ce domaine contiennent des contenus soumis à des restrictions légales.

28. Les obligations positives des États englobent en outre le devoir de promouvoir un accès équitable aux infrastructures de communication en ligne, de promouvoir l'éducation aux médias et à l'information, et de

prendre des mesures réglementaires qui favorisent le développement de plateformes en ligne inclusives et respectueuses des droits (voir CM/Rec(2016)1 sur la protection et la promotion du droit à la liberté d'expression et du droit à la vie privée en lien avec la neutralité du réseau et CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet).

Définitions

Concernant le paragraphe 11

29. Les définitions figurant au paragraphe 11 sont essentielles pour comprendre la portée et les objectifs de la recommandation.

30. La définition du terme « utilisateur » est volontairement large, englobant toute personne physique, tout groupe de personnes physiques, ainsi que toute personne morale (y compris les entreprises de médias et les organisations de la société civile).

31. Le « créateur de contenu » est défini comme un sous-ensemble des utilisateurs. Les éléments saillants de la définition, qui s'inspire de la définition de « journaliste » dans la Recommandation Rec(2000)7 sur le droit des journalistes de ne pas révéler leurs sources d'information, sont les suivants. Un créateur de contenu doit :

- viser à toucher un public au-delà de son cercle privé : cela signifie que toute personne disposant d'un compte sur les réseaux sociaux et publiant régulièrement des messages à l'intention de ses amis ou de sa famille n'est pas nécessairement un créateur de contenu ;
- être engagé dans la diffusion ou la production de contenus de manière régulière ou professionnelle, y compris toute personne physique ou morale dont l'activité professionnelle, ou un aspect significatif de celle-ci, consiste à créer du contenu en ligne, sur lequel elle exerce une responsabilité éditoriale. Ces personnes peuvent opérer en tant qu'entités commerciales en vue de réaliser un profit, mais peuvent également être constituées en organisations non gouvernementales, en entités à but non lucratif, en associations de citoyens ou sous toute autre forme juridique autorisée par le droit interne. Les utilisateurs qui, sans être enregistrés en tant qu'entreprise ou autre et sans nécessairement avoir un but commercial, publient fréquemment et régulièrement des contenus sur des plateformes, peuvent également y être compris[4] ;
- produire et diffuser des informations et des idées, sous forme écrite, sonore, visuelle, audiovisuelle ou autre. Cette définition est volontairement large ; les termes « informations et idées » reflètent le texte de l'article 10 de la Convention, et les termes « sous forme textuelle, audio, visuelle, audiovisuelle ou autre » sont délibérément formulés de manière ouverte et reflètent le texte de la Recommandation CM/Rec(2011)7 sur une nouvelle conception des médias. Elle est toutefois limitée aux informations et aux idées produites ou diffusées via une plateforme.

32. La définition du terme « plateforme » s'appuie sur la définition similaire donnée dans la Recommandation CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication, mais elle la restreint aux fins de la présente recommandation. Elle met en avant, en tant qu'éléments qui déterminent sa portée, soit la fonction de « mettre en lien leurs utilisateurs » et de « faciliter l'échange … entre eux », soit la circonstance que cet échange porte sur des « informations » et des « idées ». En conséquence, les plateformes, telles que les places de marché et les plateformes d'économie collaborative, qui connectent bien les utilisateurs, mais principalement à des fins différentes, sont en principe totalement exclues. Néanmoins, si l'échange d'idées et d'informations devient prépondérant sur des plateformes initialement conçues à d'autres fins, telles que le jeu, ces plateformes entreraient dans le champ d'application de la définition[5]. En revanche, pour les intermédiaires qui diffusent des contenus médiatiques sur lesquels ils exercent une surveillance ou un contrôle éditorial, tels que les portails d'information ou les plateformes de vidéo à la demande, seules les fonctionnalités permettant l'interaction entre utilisateurs, notamment les sections de commentaires, seraient couvertes. La définition inclut explicitement uniquement les « forums accessibles au public ». Cela signifie également que les plateformes offrant exclusivement des services de communication privés, y compris les discussions de groupe fermées, seraient exclues, tandis que des fonctionnalités spécifiques telles que

4. Pour un aperçu des définitions émergentes dans le droit interne, voir Observatoire européen de l'audiovisuel, *National rules applicable to influencers,* Strasbourg, 2024, p. 22 et suivantes, disponible en anglais seulement à l'adresse https://go.coe.int/WaGoW. L'étude couvre les États membres de l'Union européenne, la Norvège et la Suisse.
5. Voir, par exemple, Parlement européen, Direction générale des services de recherche parlementaire, *Preventing radicalisation in the European Union : how EU policy has evolved : in-depth analysis*, Parlement européen, 2025, disponible en anglais seulement à l'adresse https://data.europa.eu/doi/10.2861/3681328, section 4.2 et références qui y sont mentionnées.

les canaux publics et les applications de messagerie de groupe sont incluses, car il est prouvé que des acteurs extrémistes utilisent des groupes de messagerie publics pour s'organiser et se mobiliser[6].

33. La recommandation établit une nouvelle définition des « plateformes exerçant une influence significative », qu'elle définit comme des plateformes qui, en raison de leur taille, de leur portée ou de leur impact, jouent un rôle important dans la configuration de l'environnement informationnel à l'échelle mondiale ou dans des territoires particuliers. Les critères sont de nature générale et alternatifs, plutôt que cumulatifs, dans leur portée. Cela signifie que les États disposent d'une marge d'appréciation pour choisir ceux qu'ils souhaitent appliquer, et que leur application est soumise à la spécification dans le droit interne de critères plus précis et mesurables, fournis *ex ante* afin d'éviter une classification arbitraire des plateformes dans cette catégorie. Ils peuvent par exemple se baser sur le nombre d'utilisateurs actifs, comme dans le Règlement de l'Union européenne sur les services numériques. D'ailleurs, ils peuvent également décider d'inclure des plateformes qui, malgré une portée moindre en termes de chiffres, exercent une influence significative sur le discours public en raison de leur rôle dans la création, l'amplification ou la coordination des discours, pouvant ainsi amplifier les risques (voir également l'exposé des motifs concernant le paragraphe 41).

34. Les définitions des termes « conception de plateforme » et « autonomisation des utilisateurs » s'appuient sur les définitions fournies dans la Note d'orientation sur la lutte contre la mésinformation et la désinformation en ligne[7]. La définition de « conception de plateforme » souligne que les fonctionnalités de confiance et de sécurité destinées aux utilisateurs y sont comprises. Il s'agit par exemple des systèmes de recommandation, de l'utilisation d'étiquettes d'avertissement et des systèmes de modération de contenu. En revanche, les choix des plateformes en matière de modération des contenus licites en tant que fondés sur le point de vue ou l'opinion exprimés dans des contenus individuels n'y sont pas comprises. La définition de « autonomisation des utilisateurs » apporte des précisions supplémentaires à la définition fournie dans ladite note d'orientation, en précisant qu'elle peut également demander des mesures en ligne et hors ligne. Il peut s'agir, par exemple, de campagnes d'éducation à l'information et aux médias visant à aider les utilisateurs à reconnaître les informations erronées, à comprendre les biais algorithmiques et à évaluer les sources ; de tableaux de bord faciles à utiliser et efficaces en matière de confidentialité et de consentement ; de la personnalisation de la conception des systèmes de recommandation ou d'autres fonctionnalités des services ; et de la transparence et du contrôle par les utilisateurs des algorithmes qui déterminent la visibilité des contenus.

35. La définition de l'« autorégulation » s'appuie sur celle fournie dans la Note d'orientation sur la lutte contre la mésinformation et la désinformation en ligne. Elle ajoute une mention explicite selon laquelle « cela inclut les politiques et règles contractuelles des plateformes qui affectent les utilisateurs de leurs services ». La définition de la « corégulation » est tirée de la Recommandation CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication.

36. Les définitions des termes « contenu restreint par la loi », « contenu illégal », « contenu légal mais réglementé » et « contenu légal » aux fins de la recommandation et ses Principes doivent être comprises comme un ensemble de définitions connexes :

- Le terme « contenu illégal » désigne tout contenu interdit par le droit pénal, administratif ou civil. Il s'agit par exemple des contenus suivants :
  - les actes d'exploitation et d'abus sexuels sur des enfants, y compris la mise à disposition de matériels d'abus sexuels sur enfants tel que défini à l'article 20 de la Convention du Conseil de l'Europe sur la protection des enfants contre l'exploitation et les abus sexuels (STE n° 201) (Convention de Lanzarote) et à l'article 9 de la Convention sur la cybercriminalité (STE n° 185) ;
  - les actes à caractère raciste et xénophobe, tels que définis dans le Protocole additionnel à la Convention sur la cybercriminalité, relatif à l'incrimination d'actes de nature raciste et xénophobe commis par le biais de systèmes informatiques (STE n° 189) ;
  - les actes de violence à l'égard des femmes, tels que définis dans la Convention du Conseil de l'Europe sur la prévention et la lutte contre la violence à l'égard des femmes et la violence domestique (STCE n° 210) (Convention d'Istanbul) [et conformément à la Recommandation CM/Rec(2026)2 sur l'obligation de rendre des comptes en matière de violence à l'égard des femmes et des filles facilitée par la technologie] ;

6. Conseil de l'Europe, Comité de la lutte contre le terrorisme (CDCT), *Rapport sur les modèles émergents de détournement des technologies par les acteurs terroristes*, disponible à l'adresse https://go.coe.int/JsJB9.
7. Conseil de l'Europe, *Note d'orientation sur la lutte contre la propagation de la mésinformation et de la désinformation en ligne par le biais de la vérification des faits et de la conception de plateformes*, 2023, adoptée par le CDMSI lors de sa 24e réunion, 29 novembre-1e décembre 2023, CM(2024)9-add1.

- le discours de haine illégal, conformément à la Recommandation CM/Rec(2022)16 sur la lutte contre le discours de haine ;
- l'incitation à la violence et les menaces graves de dommages physiques ;
- les contenus terroristes, tel que les matériels incitant ou donnant des instructions pour commettre des actes de terrorisme.

– Le terme « contenu légal mais réglementé » désigne un contenu qui n'est pas illégal en soi, mais dont la publication, la diffusion ou la visibilité peuvent être restreintes par la loi dans certains contextes particuliers. Cela peut inclure :
  - les règles de silence électoral ou l'interdiction des sondages d'opinion en période électorale, afin de protéger l'intégrité du processus démocratique (voir CM/Rec(2022)12 sur la communication électorale et la couverture médiatique des campagnes électorales, par. 4.6) ;
  - la pornographie ou d'autres contenus tels que les représentations graphiques de violence ou les documents incitant à l'automutilation, qui peuvent être légaux mais dont l'accès est soumis à une restriction d'âge afin de protéger les enfants ;
  - l'affichage de données ou d'informations personnelles (telles que les adresses personnelles ou les numéros de téléphone), qui peuvent être désindexées par les moteurs de recherche afin de protéger la vie privée ;
  - les contenus faisant la promotion de produits tels que l'alcool, le tabac ou les jeux de hasard, dont la publicité ou le placement peuvent être soumis à des restrictions.

– Le terme « contenu restreint par la loi » couvre à la fois les catégories de contenu illégal et de contenu légal mais réglementé. Dans la pratique, cela signifie qu'il couvre les expressions clairement illégales, telles que certaines formes de discours de haine et d'incitation à la violence, ainsi que les expressions légales soumises à des restrictions contextuelles, telles que la diffusion soumise à des restrictions d'âge ou de durée.

– Le terme « contenu légal » désigne tout contenu, y compris les expressions ainsi que toute manifestation du comportement des utilisateurs, qui n'est pas considéré comme étant restreint par la loi.

37. Les définitions des termes « signalement », « notification » et « injonction » renvoient aux différentes manières dont les utilisateurs ou les autorités publiques peuvent porter à l'attention d'une plateforme un contenu nécessitant une action. Un exemple de « signalement » serait l'action d'un utilisateur qui porte à l'attention d'une plateforme un message sur un réseau social pour discours de haine ou bien une vidéo contenant des informations erronées sur la santé publique, en utilisant le bouton de signalement intégré à la plateforme, lorsque cela est jugé contraire soit à la loi, soit aux règles contractuelles de la plateforme. Un exemple de « notification » serait celui d'un détenteur de droits d'auteur soumettant une demande de suppression d'une plateforme de partage de vidéos de contenu piraté, ou une personne concernée demandant la suppression d'informations personnelles en vertu des lois sur la protection des données. Une « injonction » serait, par exemple, l'ordre d'un tribunal qu'une plateforme supprime un contenu diffamatoire visant un tiers, ou d'une autorité de régulation imposant la suspension d'un compte diffusant des propos de haine illégaux.

II. Risques en ligne relatifs à la liberté d'expression

Concernant les paragraphes 12 à 16

38. La Cour européenne des droits de l'homme a toujours souligné que la liberté d'expression va au-delà des discours neutres ou agréables et s'applique aussi aux opinions qui peuvent heurter, choquer ou inquiéter. Ce principe est particulièrement pertinent en ligne, où les débats sont souvent vifs et diversifiés, et où les voix minoritaires ou dissidentes doivent pouvoir participer pleinement. Si la Cour a constaté la prévalence des insultes vulgaires dans certains espaces en ligne, elle a également souligné l'importance du contexte et le seuil de tolérance plus élevé attendu, en particulier dans les débats sur des questions d'intérêt public (voir, par exemple, *Handyside c. Royaume-Uni,* précité, par. 49 ; *Delfi AS c. Estonie* [GC], précité, par. 131-139 ; *Tamiz c. Royaume-Uni* (déc.), précité, par. 81).

39. Tout en réaffirmant une protection solide de la liberté d'expression, les principes énoncés dans l'annexe reconnaissent que l'environnement en ligne génère également des risques susceptibles de compromettre ou de dissuader l'exercice même de ce droit, ainsi que d'autres droits. Ils affirment également la nécessité de

trouver un juste équilibre, préservant l'essence de tous les droits humains en jeu. Ces risques touchent les individus et les groupes, ainsi que le grand public, et peuvent avoir de graves conséquences pour la société dans son ensemble. Lorsque des personnes sont confrontées à des comportements hostiles ou manipulateurs en ligne, tels que le harcèlement coordonné de journalistes, la divulgation d'informations personnelles d'activistes ou la publication d'images intimes sans consentement, elles peuvent être intimidées et renoncer à s'exprimer à l'avenir sur des questions d'intérêt public. La crainte d'atteintes à leur réputation ou à leur intégrité physique, de poursuites judiciaires ou d'un épuisement émotionnel pur et simple peut étouffer leur volonté et leur capacité à s'exprimer de façon ouverte, en particulier sur des sujets controversés.

40. Les menaces en ligne plus générales peuvent fausser la manière dont les gens trouvent et comprennent l'information. Par exemple, lorsque des contenus partagés via des comptes non authentiques ou des vidéos manipulées en cachette sont propulsés en tête des résultats de recherche ou des flux des réseaux sociaux, cela peut noyer les informations plus fidèles et rendre plus difficile l'accès à un éventail d'opinions diverses. Cela ne se limite à affecter pas les individus, mais engendre également des conséquences graves pour la société. Par exemple, les fausses allégations sur les vaccins peuvent entraîner une baisse des taux de vaccination et accroître les risques pour la santé (voir *Bielau c. Autriche,* n° 20007/22, 27 août 2024, disponible seulement en anglais, par. 44-45). Pendant les élections, les efforts coordonnés visant à influencer les algorithmes des réseaux sociaux et la perception du public sur des sujets d'intérêt public, ainsi que la désinformation sur le processus électoral lui-même, pourraient ébranler la confiance du public dans ces processus ou dans les résultats électoraux. De plus, l'exposition répétée des personnes à des messages polarisants peut aggraver la méfiance entre les groupes et rendre plus difficile la recherche de solutions négociées aux problèmes publics.

41. Le paragraphe 13 classe les risques en trois catégories :

- risques pour la sécurité et le bien-être des personnes et des communautés ;
- risques pour le processus démocratique, l'intégrité de l'information et le débat public éclairé ; et
- risques liés aux systèmes déployés par les fournisseurs qui peuvent porter atteinte aux droits à la liberté d'expression, à la vie privée et à la protection des données à caractère personnel, ainsi qu'à d'autres droits.

42. Voici quelques exemples de risques pour la sécurité et le bien-être des personnes et des communautés :

- le risque d'être confronté à, ou victime de, discours de haine, y compris les discours sexistes, les crimes haineux et la discrimination, les menaces et la coercition ;
- le risque d'être confronté à, ou victime de, harcèlement, traque, abus et cyberintimidation, tant généraux que fondés sur l'identité ;
- le risque d'être exposé à des contenus perturbateurs pour le bien-être mental et émotionnel, notamment des contenus encourageant le suicide et l'automutilation, ou des contenus susceptibles de contribuer à des troubles alimentaires ou à une image négative du corps, y compris les risques liés au trouble dysmorphique corporel ;
- le risque d'être victime de harcèlement sexuel, d'exploitation sexuelle ou d'abus d'images intimes, y compris la falsification numérique à caractère sexuel ;
- les risques pour la vie privée, allant de la surveillance à l'usurpation d'identité, la fraude, le chantage et les escroqueries financières ;
- le risque d'être exposé au recrutement et à la radicalisation par des groupes terroristes et extrémistes.

43. Voici quelques exemples de risques pour le processus démocratique, l'intégrité de l'information et le débat public éclairé :

- le risque que des comportements inauthentiques coordonnés, y compris des campagnes de désinformation pendant les élections, influencent les électeurs, réduisent la participation électorale ou sèment le doute sur le résultat final ;
- le risque que des poursuites judiciaires, des menaces réglementaires ou des campagnes coordonnées de signalement massif poussent les médias ou les plateformes à s'autocensurer ou à supprimer des contenus légitimes, affaiblissant ainsi le rôle de « chien de garde » des médias et l'accès du public à des informations essentielles ;

- le risque que des contenus produits uniquement à des fins lucratives, y compris des contenus générés par l'intelligence artificielle (communément appelés « *AI slop* »), évincent les informations d'intérêt public ;
- le risque que des vidéos *deepfake* ou d'autres formes convaincantes d'usurpation d'identité circulent largement avant de pouvoir être démenties, sapant ainsi la confiance dans les preuves vérifiables.
- le risque que les femmes et les filles, ainsi que les personnes en situation de vulnérabilité ou exposées à la discrimination, se retirent du débat public, réduisant ainsi la diversité des voix et des points de vue.

44. Voici quelques exemples de risques associés aux systèmes déployés par les fournisseurs :

- le risque d'exclusion, de refus d'accès et d'autres obstacles à l'utilisation des systèmes publics en ligne ;
- le risque que les algorithmes de recommandation privilégient les informations sensationnelles ou fausses au détriment des reportages précis, ce qui réduit l'éventail des points de vue auxquels les gens sont exposés et accentue la polarisation ;
- le risque que les choix de conception facilitent et amplifient la diffusion de contenus présentant un risque plus élevé d'être soumis à des restrictions légales ;
- le risque de rétrogradation des contenus licites, affectant leur visibilité et réduisant leur trafic (ce que l'on appelle le « *shadow banning* »).

45. Tous ces risques ne découlent pas forcément d'un comportement criminel ou illégal. Si des actes tels que certains discours de haine, l'incitation à la violence, la cyberviolence ou le harcèlement et l'intimidation en ligne sont illégaux dans la plupart des États membres du Conseil de l'Europe, d'autres formes de contenus potentiellement préjudiciables, tels que les contenus susceptibles de nuire au bien-être mental ou émotionnel, peuvent ne pas relever du champ d'application des interdictions et réglementations légales. Pourtant, ces contenus peuvent néanmoins nuire de manière significative aux individus et à la société dans son ensemble. La question cruciale à trancher n'est donc pas de savoir si un comportement est illégal ou pas, mais d'identifier quel type de réponse est approprié, proportionné et efficace pour minimiser les risques de préjudice. Ces réponses devraient être conçues de manière à établir un juste équilibre entre les différents droits humains, notamment les droits à la liberté d'expression, à la vie privée, y compris la protection des données à caractère personnel.

46. Comme le souligne le paragraphe 16, la disponibilité et l'utilisation généralisées de l'intelligence artificielle dans la production et la diffusion de contenus peuvent amplifier considérablement les risques existants. Les Lignes directrices sur les implications de l'intelligence artificielle générative pour la liberté d'expression[8], explorent par exemple les risques, ainsi que les opportunités, associés à l'utilisation des outils d'intelligence artificielle générative.

Concernant le paragraphe 17

47. Le paragraphe 17 souligne que certaines catégories d'utilisateurs, en particulier parmi les créateurs de contenu, sont plus exposés aux risques que d'autres, soit en raison de leur identité, soit en raison de leur position. Il souligne également que les risques en ligne peuvent entraîner des répercussions dans l'environnement physique. Le harcèlement et les menaces en ligne sont souvent le préalable à d'autres formes de violence : on recense d'innombrables cas de journalistes, d'activistes et de personnalités politiques qui ont été harcelés, agressés ou pris pour cible dans le monde physique à la suite de campagnes de cyberharcèlement[9]. Cela est particulièrement vrai pour les femmes occupant des fonctions publiques, qui sont victimes de niveaux disproportionnés de harcèlement en ligne et de ciblage fondé sur le genre[10].

48. Les risques accrus auxquels sont exposés les enfants sont bien documentés et ont déjà fait l'objet d'instruments du Conseil de l'Europe, tels que la Convention de Lanzarote et la Recommandation CM/Rec(2018)7 sur les Lignes directrices relatives au respect, à la protection et à la réalisation des droits de l'enfant dans l'environnement numérique. Ces lignes directrices reconnaissent que les enfants peuvent être exposés à

8. Conseil de l'Europe, *Note d'orientation sur les incidences de l'intelligence artificielle générative sur la liberté d'expression*, adoptée par le CDMSI lors de sa 28e réunion, 3-5 décembre 2025, CDMSI(2025)15-rev disponible à l'adresse https://go.coe.int/uvwOU.
9. Zamfir I. et Murphy C., *Cyberviolence against women in the EU,* Service de recherche du Parlement européen, Note d'informations politique, PE 767.146, Bruxelles, 2024, disponible en anglais seulement à l'adresse www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2024)767146.
10. Posetti J. et al., *Violence en ligne à l'égard des femmes journalistes : un aperçu mondial des incidences et impacts,* UNESCO, Paris, 2020, disponible à l'adresse https://unesdoc.unesco.org/permalink/P-bd67e208-065d-41f5-ba4e-ca4aa6098871.

toute une série de préjudices graves, y compris l'exploitation sexuelle et le harcèlement, ainsi que d'autres menaces pour leur bien-être. Elles précisent également que l'exposition au risque n'est pas uniforme : les besoins des enfants évoluent avec l'âge et la maturité. Une protection efficace doit donc être graduée, axées sur l'autonomisation et centrée sur l'enfant, en trouvant un équilibre entre la sécurité et le droit de grandir, d'explorer et de participer de manière significative à la vie en ligne. Le Comité de Lanzarote, organe de suivi de la Convention de Lanzarote, a publié, en 2022, un Rapport de mise en œuvre sur la protection des enfants contre l'exploitation et les abus sexuels facilités par les technologies de l'information et de la communication, en 2019, un Avis sur les images et/ou vidéos d'enfants sexuellement suggestives ou explicites produites, partagées ou reçues par des enfants et, en 2024, une Déclaration sur la protection des enfants contre l'exploitation et les abus sexuels facilités par les technologies émergentes.

49. Les risques en ligne auxquels sont exposées les femmes et les filles, en particulier les créatrices de contenu, sont tout aussi bien documentés et font l'objet d'un travail de longue date du Conseil de l'Europe. Le Commissaire aux droits de l'homme du Conseil de l'Europe a lancé un signal d'alerte sur la violence à l'égard des femmes et des filles dans le monde en ligne, décrivant l'internet comme « un terrain fertile pour la violence sexiste à l'égard des femmes et des filles dans une mesure alarmante, et avec peu de responsabilité » et appelant à l'action[11]. La Cour a rendu plusieurs arrêts dans lesquels elle a constaté une violation des droits humains dans des affaires où les États n'avaient pas suffisamment protégé le droit des femmes à être protégées contre les abus en ligne (*M.Ş.D. c. Roumanie*, n° 28935/21, 3 décembre 2024 ; *Volodina c. Russie (n° 2)*, n° 40419/19, 14 septembre 2021 ; *Buturugă c. Roumanie*, n° 56867/15, 11 février 2020 ; disponibles seulement en anglais). Le Groupe d'experts sur l'action contre la violence à l'égard des femmes et la violence domestique (GREVIO), qui supervise la mise en œuvre de la Convention d'Istanbul, a publié en 2021 sa Recommandation générale n° 1 sur la dimension en ligne de la violence à l'égard des femmes, définissant diverses formes de cyberviolence (telles que l'abus d'images intimes sans consentement, le harcèlement en ligne et les menaces), et exhortant les États à prendre des mesures pour protéger les femmes et les filles en ligne, ainsi que pour prévenir et poursuivre la violence en ligne. En outre, le Comité des Ministres du Conseil de l'Europe a adopté la Recommandation CM/Rec(2026)2 sur l'obligation de rendre des comptes en matière de violence à l'égard des femmes et des filles facilitée par la technologie, qui fournit des orientations aux États membres sur le renforcement des réponses juridiques, institutionnelles et réglementaires à cette violence.

50. Les femmes et les filles, les enfants, les personnes en situation de vulnérabilité et les individus et groupes exposés à la discrimination, notamment les personnes handicapées, les minorités ethniques, linguistiques et religieuses, les communautés LGBTI, ainsi que les migrants et les personnes issues de l'immigration, sont exposés au risque d'être ciblés en ligne en raison de leur identité. Dans son rapport annuel 2024, l'Agence des droits fondamentaux de l'Union européenne (FRA) a souligné la nécessité de réglementer les espaces en ligne, les qualifiant de « domaine à haut risque pour les droits fondamentaux, en particulier pour les personnes vulnérables et marginalisées », notamment en raison du risque « d'altérisation de certains groupes par la désinformation ou la propagation de la haine en ligne et par la création d'obstacles pour les groupes démographiques vulnérables », concluant que « la prise en compte de ces risques dans le cadre de la création de l'environnement en ligne de demain est essentielle pour construire une Europe plus inclusive »[12].

51. La Recommandation CM/Rec(2016)4 sur la protection du journalisme et la sécurité des journalistes et autres acteurs des médias souligne les risques auxquels sont confrontés les journalistes et note que « ces violations sont de plus en plus souvent commises en ligne ». Les menaces qui pèsent sur la sécurité en ligne ont de graves conséquences sur la capacité des journalistes à rendre compte de questions d'intérêt public et sur le droit du public à en être informé. Une étude du Conseil de l'Europe réalisée en 2017 a révélé que 37 % des journalistes interrogés avaient autocensuré des reportages potentiellement critiques en raison d'un risque de préjudice. Une étude de suivi réalisée en 2020 a mis en évidence les expériences de vingt journalistes[13]. L'une des journalistes interrogées dans le cadre de cette étude était Daphne Caruana Galizia. Elle a été assassinée quelques jours après avoir été interviewée. Les femmes journalistes sont beaucoup exposées que leurs homologues masculins aux abus en ligne. Un rapport de l'UNESCO publié en 2021 montre l'ampleur des

11. Commissaire aux droits de l'homme, *Carnet des droits de l'homme : Pas de place pour la violence à l'égard des femmes et des filles dans le monde numérique*, 2022, disponible à l'adresse https://go.coe.int/qNAxw.
12. Agence des droits fondamentaux de l'Union européenne (FRA), *Rapport sur les droits fondamentaux 2024,* 2024, disponible en anglais seulement à l'adresse https://fra.europa.eu/fr/publication/2024/fundamental-rights-report-2024, p. 12, traduction non-officielle. Voir également Recommandation CM/Rec(2026)2 sur l'obligation de rendre des comptes en matière de violence à l'égard des femmes et des filles facilitée par la technologie.
13. Clark M. et Grech A., *Journalists under pressure : Unwarranted interference, Fear and self-censorship among Journalists in Europe,* Conseil de l'Europe, Strasbourg, 2017, disponible en anglais seulement à l'adresse https://go.coe.int/RoHyP. Voir également Clark M. et Horsley W., *A mission to inform : Journalists at risk speak out,* Conseil de l'Europe, Strasbourg, 2020, disponible en anglais seulement à l'adresse https://go.coe.int/s1kfx.

attaques contre les femmes journalistes et leur impact sur leur bien-être, leur travail et la liberté de la presse en général[14].

52. Dans l'une des affaires les plus médiatisées d'abus en ligne visant des femmes journalistes, *Khadija Ismayilova c. Azerbaïdjan* (précité), la Cour européenne des droits de l'homme a examiné la surveillance secrète d'une journaliste d'investigation dont la chambre à coucher et d'autres espaces privés avaient été filmés à son insu et dont les images avaient été diffusées en ligne. La Cour a qualifié ces abus de « graves et contraires à la dignité humaine » et d'« atteinte grave, flagrante et extrêmement intense à sa vie privée ». Elle a en outre souligné que « la requérante est une journaliste connue et qu'il existait un lien plausible entre son activité professionnelle et les intrusions susmentionnées, dont le but était de la réduire au silence » (par. 116, traduction non officielle).

53. Également, les personnalités politiques, les chercheurs, les éducateurs, les scientifiques, les activistes et d'autres personnes qui contribuent fréquemment au débat sur des questions d'intérêt public sont souvent la cible d'abus visant à les empêcher de participer à ce débat[15].

54. Les personnes qui appartiennent simultanément à plusieurs de ces groupes sont encore plus susceptibles d'être exposées à des risques de nature intersectionnelle. Quand les divers motifs ou facteurs de risque se cumulent, les individus se trouvent exposés à des formes plus complexes de discrimination, d'exclusion et de violence, donnant lieu à des vécus et des vulnérabilités spécifiques[16]. Le Commissaire aux droits de l'homme du Conseil de l'Europe a fourni plusieurs exemples de ces formes intersectionnelles d'abus dans une déclaration de 2022 sur la violence en ligne à l'égard des femmes et des filles[17].

Concernant le paragraphe 18

55. Le paragraphe 18 précise que l'existence de risques en ligne, en tant que tel, ne justifie pas toujours l'introduction de mesures qui portent atteinte à l'exercice de la liberté d'expression et d'autres droits humains, notamment en restreignant ou en régulant les contenus. Les mesures adoptées par les États doivent répondre à un besoin social impérieux et être proportionnées, ce qu'implique l'indisponibilité d'autres mesures appropriées, qui ne soient pas restrictives des droits ou qui le soient moins (*Glor c. Suisse*, n° 13444/04, 30 avril 2009, par. 94), telles que, par exemple, l'éducation aux médias et d'autres initiatives d'autonomisation. Les plateformes devraient adopter la même approche, par exemple dans le contexte de la modération de contenu (voir l'Exposé des motifs relatif à la Recommandation CM/Rec(2022)11, par. 12.2).

56. Afin de garantir le respect des exigences de nécessité et de proportionnalité, de telles mesures ne devraient être introduites que lorsque les informations disponibles indiquent l'existence d'un risque réel ou probable de préjudice. Cette approche fondée sur des informations factuelles est renforcée tout au long de la recommandation et ses principes. Le Préambule souligne la nécessité de cadres juridiques transparents et fondés sur des preuves pour évaluer et traiter les risques en ligne d'une manière conforme aux droits humains. Le paragraphe 5 appelle à la mise en place de cadres réglementaires et coréglementaires fondés sur des informations factuelles ; le paragraphe 45 exige que les règles relatives aux contenus soient établies sur la base d'informations recueillies de manière transparente ; et les paragraphes 51, 64 et 66 obligent également les autorités réglementaires et les législateurs à s'appuyer sur des informations factuelles lors de la conception de leurs interventions.

Concernant le paragraphe 19

57. Le paragraphe 19 souligne que les mesures visant à améliorer la sécurité en ligne peuvent elles-mêmes menacer les droits humains. Une préoccupation majeure réside dans l'imposition, par les États ou les régulateurs, de restrictions disproportionnées à la liberté d'expression, en particulier lorsqu'elles amènent les plateformes à agir de manière préventive, ce qui conduit à des retraits excessifs de contenus ou à une forme de « censure par ricochet ». La *loi Avia* (loi n° 2020-766 du 24 juin 2020 relative à la lutte contre les contenus

14. Posetti J. *et al.*, *The chilling : Global trends in online violence against women journalists,* UNESCO, Paris, 2021, disponible en anglais seulement à l'adresse https://unesdoc.unesco.org/ark :/48223/pf0000377223.
15. Voir : Congrès des pouvoirs locaux et régionaux, Résolution 459 (2020) et Recommandation 449 (2020) Lutter contre la violence sexiste à l'égard des femmes en politique aux niveaux local et régional ; Commissaire aux droits de l'homme, *Rapport : Les défenseurs des droits de l'homme dans l'espace du Conseil de l'Europe en temps de crise,* 2023, disponible en anglais seulement à l'adresse https://go.coe.int/yuHwm ; GREVIO, *6e rapport général sur les activités du GREVIO, couvrant la période de janvier à décembre 2024,* Conseil de l'Europe, 2025, p. 33-38, disponible à l'adresse https://go.coe.int/GjJ3X.
16. Voir Conseil de l'Europe, *Stratégie pour l'égalité de genre (2024-2029)*, CM(2024)17-final, adoptée par le Comité des Ministres le 6 mars 2024, CM/Del/Dec(2024)1491/4.3.
17. Commissaire aux droits de l'homme, *Carnet des droits de l'homme : Pas de place pour la violence à l'égard des femmes et des filles dans le monde numérique*, 2022, disponible à l'adresse https://go.coe.int/qNAxw.

haineux sur Internet), en France, illustre ce risque : bien qu'elle vise à lutter contre le discours de haine, ses dispositions fondamentales ont été annulées en 2020 par le Conseil constitutionnel au motif qu'elles constituaient une menace disproportionnée pour la liberté d'expression, à défaut d'un contrôle judiciaire suffisant et de garanties suffisantes contre les retraits excessifs (décision n° 2020-801 DC, 18 juin 2020). Dans le contexte de la réglementation des services audiovisuels, la Cour européenne des droits de l'homme a souligné qu'une publicité télévisée en faveur des droits des animaux ne pouvait être interdite simplement parce que les téléspectateurs pouvaient la trouver « désagréable » (*Verein gegen Tierfabriken Schweiz (VgT) c. Suisse (n° 2)* [GC], n° 32772/02, 30 juin 2009, par. 96).

58. Des risques similaires apparaissent lorsque les systèmes de recommandation ou les outils de hiérarchisation des contenus sont, pour des raisons de sécurité des utilisateurs, conçus et utilisés de manière à réduire de façon invisible la portée de certains points de vue, affectant de manière disproportionnée les voix minoritaires ou les perspectives controversées, bien que légitimes. Par exemple, une étude publiée par l'Agence des droits fondamentaux de l'Union européenne en 2022 a révélé que les algorithmes de détection des propos offensants et haineux produisent des résultats biaisés et peuvent avoir un effet discriminatoire[18].

59. Les États et les acteurs privés doivent donc veiller à ce que les mesures de sécurité soient nécessaires, proportionnées, transparentes et fondées sur des informations factuelles, et qu'elles fassent l'objet d'un contrôle indépendant et de voies de recours effectives. L'objectif n'est pas d'opposer la sécurité aux droits, mais de garantir que les mesures de sécurité renforcent la protection des droits dans leur ensemble.

III. Principes généraux d'un environnement en ligne favorable

Principes à l'attention des États

Concernant les paragraphes 20 à 23

60. L'accent mis sur la création d'un environnement en ligne favorable reflète un principe de longue date dans la jurisprudence de la Cour européenne des droits de l'homme et dans les documents du Comité des Ministres : en vertu de l'article 10 de la Convention, les États ont une obligation positive de favoriser des conditions dans lesquelles la liberté d'expression peut être exercée efficacement et par tous. Dans sa jurisprudence relative aux agressions contre des journalistes, la Cour a souligné à plusieurs reprises que ces conditions exigent une protection efficace des personnes les plus exposées. Dans l'affaire *Dink c. Turquie* (précité) et dans des affaires similaires, la Cour a estimé que des États avaient manqué à leurs obligations de protéger les journalistes contre les menaces, le harcèlement et la violence, notant que « les obligations positives en la matière impliquent, entre autres, que les États sont tenus de créer, tout en établissant un système efficace de protection des auteurs ou journalistes, un environnement favorable à la participation aux débats publics de toutes les personnes concernées, leur permettant d'exprimer sans crainte leurs opinions et idées » (par. 137).

61. La Recommandation CM/Rec(2016)4 sur la protection du journalisme et la sécurité des journalistes et autres acteurs des médias souligne qu'un environnement favorable suppose que « au minimum la sûreté, la sécurité et la protection de tout un chacun, en particulier des journalistes et des autres acteurs des médias, soient réellement garanties dans la pratique, et que chacun puisse s'attendre à pouvoir contribuer au débat public sans crainte et sans avoir à modifier sa conduite sous l'effet de la peur » (principes, paragraphe 18).

62. Le même principe s'applique à la sécurité en ligne : les États ont le devoir de promouvoir des conditions permettant à chacun de contribuer au débat public sans crainte d'intimidation ou de représailles. L'existence d'une obligation positive de « créer un environnement sûr et favorable à la participation de tous au débat public » est reconnue par le Comité des Ministres dans les préambules des Recommandations CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, CM/Rec(2022)16 sur la lutte contre le discours de haine et CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression. La Recommandation CM/Rec(2016)5 sur la liberté de internet invite les États à « créer un environnement favorable à la liberté de l'Internet » et la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet affirme que la législation applicable aux intermédiaires d'internet « devrait créer un environnement en ligne sûr, propice aux communications privées et au débat public ». En outre, la Recommandation CM/Rec(2020)1 sur les impacts des systèmes algorithmiques sur les droits de l'homme invite les États à « veiller à ce que les processus de conception, développement et de déploiement

18. Agence des droits fondamentaux de l'Union européenne (FRA), *Bias in algorithms - Artificial intelligence and discrimination*, 2022, disponible en anglais seulement à l'adresse https://data.europa.eu/doi/10.2811/25847, p. 12.

permanent des algorithmes intègrent des mécanismes permettant d'assurer par essence la sûreté, le respect de la vie privée, la protection des données et la sécurité » (section B, par. 3.2).

63. La présente recommandation et ses principes s'appuient sur cet acquis et soulignent que la sécurité, l'inclusivité, le pluralisme et l'autonomisation des utilisateurs sont essentiels à un environnement en ligne respectueux des droits. Si la recommandation se concentre sur les mesures relatives à l'environnement en ligne, ses principes reconnaissent que la réalisation de cet objectif nécessite une combinaison de mesures réglementaires et éducatives. Les politiques dans ce domaine devraient par conséquent s'inscrire dans une stratégie globale et coordonnée qui s'attaque aux conditions sociales et aux inégalités sous-jacentes qui donnent lieu à des abus en ligne et déterminent l'exposition des utilisateurs à ces abus. Une telle stratégie devrait englober des mesures visant à promouvoir l'égalité, la cohésion sociale et les valeurs démocratiques ; des mesures visant à renforcer l'État de droit et la sécurité publique ; ainsi que des mesures visant à donner aux utilisateurs les moyens de faire des choix éclairés concernant leur expérience en ligne. Dans la pratique, cela peut englober :

- l'éducation et la sensibilisation : intégrer la citoyenneté en ligne et la sécurité en ligne dans les programmes scolaires ; mener des campagnes visant à renforcer la résilience face à la désinformation ; mettre en place des programmes encourageant un comportement respectueux en ligne[19] ;
- l'éducation aux médias et à l'information : stratégies nationales visant à renforcer l'engagement critique vis-à-vis des contenus en ligne ; soutien public aux initiatives de vérification des faits ; partenariats avec la société civile pour améliorer l'accès à des informations fiables[20] ;
- l'autonomisation des communautés : financement d'initiatives visant à soutenir les groupes disproportionnellement ciblés en ligne (par exemple, les femmes, les minorités, les personnes LGBTI, les journalistes) ; financement de lignes d'assistance téléphonique ou de services de soutien aux victimes d'abus en ligne ; financement de formations sur la sécurité en ligne pour les ONG et les leaders communautaires[21] ;
- le soutien à un journalisme de qualité et à des médias pluralistes : garantir l'indépendance des médias de service public ; offrir des systèmes transparents et équitables de soutien financier au journalisme d'investigation ; protéger les journalistes contre le harcèlement judiciaire et les poursuites-bâillons (SLAPP)[22] ;
- la mise en œuvre de la loi et des responsabilités : enquêtes criminelles efficaces sur les violences et les abus facilités par la technologie ; unités de police spécialisées formées pour traiter les cybercrimes avec sensibilité et efficacité ; garanties, y compris la protection contre la responsabilité pénale des intermédiaires pour les actes de leurs utilisateurs[23], visant à assurer que les poursuites et les procédures judiciaires respectent les droits humains ; mécanismes de plainte et de recours accessibles aux utilisateurs[24].

64. Le paragraphe 22 affirme qu'un environnement en ligne sûr et favorable repose sur la préservation des garanties techniques qui protègent les droits, y compris le chiffrement de bout en bout des communications privées. Dans l'affaire *Podchasov c. Russie* (n° 33696/19, 13 février 2024, par. 76-79), la Cour a observé que les mesures qui affaiblissent ou contournent le chiffrement (par exemple, les obligations de déchiffrement, les dispositifs d'accès réservés de type « backdoors » ou toute mesure équivalente sur le plan fonctionnel)

19. Convention du Conseil de l'Europe sur la protection des enfants contre l'exploitation et les abus sexuels (STCE n° 201) (Convention de Lanzarote) ; Convention du Conseil de l'Europe sur la prévention et la lutte contre la violence à l'égard des femmes et la violence domestique (STCE n° 210) (Convention d'Istanbul) ; Recommandation CM/Rec(2019)10 visant à développer et à promouvoir l'éducation à la citoyenneté numérique ; Recommandation CM/Rec(2019)1 sur la prévention et la lutte contre le sexisme. Le Comité directeur pour les droits de l'enfant (CDENF) prépare un projet de recommandation sur la protection des enfants contre la violence par le biais d'une éducation complète à la sexualité adaptée à leur âge, qui sera soumis au Comité des Ministres pour adoption en 2026, pour plus d'informations voir https://go.coe.int/qz0FM.
20. Recommandation CM/Rec(2018)1 sur le pluralisme des médias et la transparence de leur propriété. Voir également Conseil de l'Europe, *Stratégies nationales d'éducation aux médias et à l'information (EMI) – Indicateurs et aspects pratiques*, adopté par le CDMSI lors de sa 28e réunion, 3-5 décembre 2025, CDMSI(2025)09, disponible à l'adresse https://go.coe.int/2vdHr.
21. Conformément aux ressources suggérées sur le portail Cyberviolence du Conseil de l'Europe : www.coe.int/fr/web/cyberviolence/home. Voir également la Recommandation CM/Rec(2018)7 sur les Lignes directrices relatives au respect, à la protection et à la réalisation des droits de l'enfant dans l'environnement numérique.
22. Recommandation CM/Rec(2022)4 sur la promotion d'un environnement favorable à un journalisme de qualité à l'ère du numérique ; Recommandation CM/Rec(2024)2 sur la lutte contre l'utilisation des poursuites stratégiques contre la participation publique (poursuites-bâillons).
23. Rapport explicatif de la Convention sur la cybercriminalité (ETS n° 185), par. 125.
24. Recommandation CM/Rec(2016)4 sur la protection du journalisme et la sécurité des journalistes et autres acteurs des médias.

risquent de s'apparenter à une surveillance générale et indifférenciée, considérée comme disproportionnée et contraire à l'article 8 de la Convention.

65. Le paragraphe 23 souligne la nécessité de veiller à ce que les risques d'exclusion et de marginalisation de certaines catégories de personnes des espaces en ligne, résultant soit de mesures prises par les États, soit de l'absence de telles mesures, soient dûment examinés et pris en compte. Tant l'intervention réglementaire de l'État concernant la manière dont les plateformes protègent et autonomisent les utilisateurs que l'absence d'intervention lorsqu'une protection est nécessaire peuvent, involontairement, créer des obstacles à l'accessibilité et à l'inclusion. Par exemple, les outils de sécurité tels que les systèmes de vérification de l'âge ou de l'identité des utilisateurs mis en place pour protéger les enfants, s'ils sont mal conçus, peuvent poser des obstacles à l'inclusion de certains adultes n'ayant pas de documents d'identité officiels ou ne possédant pas de compétences numériques adéquates. De même, les systèmes automatisés de modération de contenu visant à lutter contre le discours de haine ou la désinformation peuvent réduire de manière disproportionnée la voix des minorités, des femmes ou des personnes handicapées qui dépendent de modèles linguistiques particuliers ou de technologies d'assistance. À l'inverse, l'absence d'intervention, par exemple en n'exigeant pas des plateformes qu'elles disposent de canaux accessibles pour signaler les abus, peut laisser sans protection des groupes déjà particulièrement exposés. Les États ont donc la responsabilité de veiller à ce que les mesures destinées à promouvoir la sécurité n'ancrent pas involontairement des désavantages et que l'inaction ne perpétue pas une exposition inégale aux risques en ligne.

66. Le paragraphe 25 s'appuie sur le principe selon lequel la transparence est essentielle à la responsabilité démocratique, tant dans les activités des États que dans les opérations d'acteurs privés influents tels que les plateformes en ligne. La Convention du Conseil de l'Europe sur l'accès aux documents publics (CETS n° 205) établit que le public a le droit de savoir comment les autorités publiques exercent leurs pouvoirs, sous réserve d'exceptions qui doivent être interprétées de manière restrictive et mises en balance avec l'intérêt public supérieur à la divulgation. La protection des lanceurs d'alerte joue un rôle important dans ce cadre. Comme l'indique clairement la Recommandation CM/Rec(2014)7 sur la protection des lanceurs d'alerte, les personnes qui divulguent des informations dans l'intérêt public – par exemple, sur des demandes de surveillance illégales, des pratiques de modération dangereuses ou des risques systémiques pour les utilisateurs – devraient être protégées contre les représailles.

Principes à l'attention des plateformes

Concernant les paragraphes 26 à 29

67. En raison de leur rôle central dans la facilitation et le façonnement de l'expression en ligne, toutes les plateformes devraient intégrer des considérations de sécurité et d'autonomisation dans la conception de leurs services de base et dans leurs choix de gouvernance. Il ne suffit pas que les plateformes se limitent à prendre des mesures de manière réactive : la sécurité et l'autonomisation doivent être intégrées dès la conception, y compris dans le développement et le déploiement de systèmes d'intelligence artificielle, d'outils de recommandation et de mécanismes de modération de contenu. Elles devraient accorder une attention particulière aux risques accrus auxquels peuvent être confrontées les femmes et les filles, les enfants, les personnes en situation de vulnérabilité et les individus et groupes exposés à la discrimination, notamment les personnes handicapées, les minorités ethniques, linguistiques et religieuses nationales, les communautés LGBTI et les personnes issues de l'immigration. Ces dispositions s'appuient sur des principes similaires déjà énoncés dans toute une série d'instruments. Par exemple, la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet souligne que les intermédiaires devraient intégrer le respect des droits humains dans leurs processus de conception et de prise de décision, et devraient procéder à des évaluations régulières de l'impact de leurs services sur les droits humains. La Recommandation CM/Rec(2020)1 sur les impacts des systèmes algorithmiques sur les droits de l'homme exige que les droits humains, la démocratie et l'État de droit soient préservés tout au long de la conception et du cycle de vie des systèmes d'IA et algorithmiques. L'évaluation et l'atténuation des risques et des effets négatifs constituent également un élément central de la Convention-cadre sur l'intelligence artificielle et les droits de l'homme, la démocratie et l'État de droit (CETS n° 225).

68. Conformément au principe de proportionnalité et à l'approche graduée de la réglementation des plateformes, ces responsabilités et les moyens de les assumer prennent des formes différentes en fonction de la taille, de la portée ou de l'impact des plateformes. Les plateformes exerçant une influence significative ont des responsabilités accrues et peuvent donc être soumises à des exigences légales spécifiques, telles que celles détaillées dans les parties IV et V de l'annexe. Si les responsabilités en matière de sécurité pour les

plateformes de plus petite taille doivent être plus limitées afin de ne pas les surcharger et d'éviter étouffer l'innovation et la concurrence, elles n'en restent pas moins importantes. On peut par exemple attendre de ces plateformes qu'elles adoptent des normes contractuelles claires et accessibles, qu'elles veillent à ce que les utilisateurs disposent d'outils simples pour signaler les abus et qu'elles répondent rapidement aux plaintes. En général, lorsqu'elles sont soumises à un contrôle réglementaire, toutes les plateformes devraient être en mesure de démontrer qu'elles ont agi avec la diligence requise, en tenant compte des risques et en intégrant dans leur conception des mesures raisonnables et adaptées à leur niveau de risque et à leurs capacités.

69. Les mesures visant à la sécurité des utilisateurs ne devraient pas être mises en œuvre au détriment du pluralisme des médias, de la diversité des opinions ou du caractère ouvert et inclusif du débat public. Par exemple, les règles qui amènent les plateformes à supprimer automatiquement de grandes quantités de contenus signalés peuvent affecter de manière disproportionnée la voix des minorités ou les points de vue controversés mais légitimes, en particulier lorsque les outils de modération automatisés interprètent de manière erronée les expressions culturelles ou linguistiques. De même, les obligations imposant aux plateformes de donner la priorité aux sources fiables d'information peuvent involontairement marginaliser les médias locaux ou indépendants, réduisant ainsi la diversité des points de vue accessibles aux utilisateurs. Les efforts visant à lutter contre la publication de contenus restreints par la loi devraient éviter d'imposer des restrictions excessivement larges, qui pourraient être exploitées pour réprimer les critiques légitimes à l'égard des gouvernements ou des acteurs puissants.

70. Pour prévenir ces risques, les interventions doivent être transparentes, proportionnées et fondées sur le droit international des droits humains. La transparence exige l'établissement de règles claires relatives à la manière dont les contenus sont modérés et hiérarchisés, ainsi que la mise en place de recours accessibles pour les utilisateurs dont la liberté d'expression est injustement restreinte. La proportionnalité exige que les restrictions soient précisément ciblées afin de remédier à des préjudices clairement définis, plutôt que d'être imposées par des mesures générales qui risquent de supprimer tout débat légitime. Le respect du droit international des droits humains garantit que les mesures de protection sont compatibles, entre autres, avec les droits à la liberté d'expression, à la vie privée et à la non-discrimination.

71. Le paragraphe 29 souligne que les plateformes qui opèrent à grande échelle dans un pays ou une région devraient comprendre les contextes locaux et y répondre. Cela signifie, par exemple, reconnaître et prendre en considération que dans un certain pays ou une région donnée, des groupes minoritaires peuvent être disproportionnellement pris pour cible par des campagnes de harcèlement coordonnées ; que la désinformation peut être diffusée dans des langues locales que les systèmes de modération automatisés ne détectent pas ; et que le discours de haine peut prendre la forme de termes codés, inconnus du personnel opérant depuis l'étranger. Elles doivent également tenir compte des risques spécifiques liés au genre, notamment de la manière dont les normes sociales et les comportements en ligne propres à chaque contexte peuvent exposer les femmes et les filles à des risques particuliers. Pour lutter efficacement contre ces risques, les plateformes devraient employer ou recruter du personnel connaissant les dynamiques politiques, culturelles et sociales locales ; désigner des points de contact accessibles aux utilisateurs et aux régulateurs ; et s'assurer que les équipes chargées de la sécurité maîtrisent les langues officielles de la juridiction concernée. Il s'agit d'une question d'efficacité : sans expertise locale, les systèmes de modération passent à côté des abus, suppriment par erreur des contenus légitimes ou ne parviennent pas à réagir à des risques urgents tels que l'incitation à la violence. Parmi les exemples de bonnes pratiques, on peut citer les plateformes qui mettent en place des équipes dédiées à l'intégrité des élections dans les pays en période politique sensible ou qui font appel à des ONG locales pour fournir une expertise sur les menaces pesant la sécurité et les droits humains.

Principes à l'attention des créateurs de contenu

Concernant les paragraphes 30 à 32

72. Conformément au texte de l'article 10, paragraphe 2, de la Convention, les créateurs de contenu, comme toute personne qui exerce son droit à la liberté d'expression, ont des devoirs et des responsabilités. Le niveau de ces devoirs et responsabilités dépend de facteurs tels que le format du contenu et sa pertinence pour le débat public. Par exemple, un journaliste qui réalise des reportages d'investigation sur la corruption politique a une responsabilité plus grande en matière d'exactitude et d'impartialité qu'un influenceur de mode ou qu'un créateur de mèmes. Cependant, tous les créateurs de contenu sont tenus de respecter les droits et la dignité humains.

73. La Cour européenne des droits de l'homme a constamment jugé que, s'agissant de la protection offerte par l'article 10, le rôle des blogueurs et des utilisateurs influents des réseaux sociaux qui diffusent

des contenus sur des questions d'intérêt public peut être assimilé à celui de « chien de garde », traditionnellement joué par la presse (*Magyar Helsinki Bizottság c. Hongrie* [GC], 2016, para. 168). Cette protection est soumise à la condition qu'ils respectent les devoirs et responsabilités traditionnellement liés à la fonction de journaliste. Les créateurs de contenu qui publient sur des questions d'intérêt public, revendiquent une expertise professionnelle ou touchent un public important ont un devoir accru d'agir de bonne foi et de respecter les principes d'exactitude, d'équité et d'intégrité. Par exemple, les chercheurs peuvent être soumis à une obligation renforcée de garantir l'exactitude lorsqu'ils communiquent avec le grand public[25]. Cela reflète la jurisprudence de la Cour européenne des droits de l'homme sur les « devoirs et responsabilités » de ceux qui contribuent au débat public (*Stoll c. Suisse,* n° 69698/01, 10 décembre 2007, par. 104 ; *Jersild c. Danemark*, n° 15890/89, 23 septembre 1994, par. 31 ; *Savva Terentyev c. Russie*, n° 10692/09, 28 août 2018, disponible seulement en anglais, par. 79). Dans l'affaire *Bielau c. Autriche* (précité), un médecin avait été condamné à une amende par le Conseil disciplinaire de l'Association médicale autrichienne pour avoir publié des affirmations scientifiquement indéfendables sur les vaccins, décision qui avait été confirmée par les tribunaux nationaux. La Cour a souligné que « [l]a restriction de la liberté d'expression des médecins peut être nécessaire en cas d'informations publiques catégoriques et mensongères sur des questions médicales, en particulier si ces informations sont publiées sur un site web, afin de protéger la santé et le bien-être d'autrui » (par. 44, traduction non officielle). Puisque le médecin faisait la promotion de « l'auto-guérison et de l'homéopathie » et utilisait ces déclarations pour faire la publicité de ses services, la Cour l'a considéré comme un créateur de contenu ayant une qualification professionnelle dont l'expression pouvait être restreinte lorsqu'il ne respectait pas les devoirs et responsabilités y associés.

74. Les responsabilités énoncées aux paragraphes 30 et 31 s'alignent sur les nouvelles réglementations relatives à la catégorie des créateurs de contenu souvent appelés « influenceurs »[26]. L'attention est attirée sur la nécessité pour les créateurs de contenu dont l'activité porte sur l'actualité et les questions d'intérêt public, parfois appelés « influenceurs-journalistes » ou « newsfluenceurs », de s'aligner sur les normes journalistiques.

75. Plusieurs États européens ont adopté, ou envisagent d'élaborer, des règles applicables aux personnalités des réseaux sociaux qui rassemblent un large public et influencent l'opinion publique. La France, par exemple, a adopté une loi qui, *entre autres*, interdit la promotion commerciale par les influenceurs de certains produits à haut risque (par exemple, la chirurgie, les cryptomonnaies, la nicotine) et exige la transparence en matière de publicité et de retouche photographique[27]. En Espagne, les « utilisateurs particulièrement influents » des plateformes de partage de vidéos sont tenus de s'enregistrer et sont soumis aux règles des médias audiovisuels en matière de protection des mineurs et de communications commerciales audiovisuelles. L'élaboration de codes de conduite d'autorégulation et de corégulation est également encouragée[28]. Dans les pays dépourvus de législation spécifique, les influenceurs sont généralement régis par le droit de la protection des consommateurs et, s'ils répondent à la définition et remplissent les conditions pour être considérés comme des prestataires de tels services, les règles relatives aux services de médias audiovisuels s'appliquent alors. Souvent, ces dernières transposent et mettent en œuvre la Directive sur les services de médias audiovisuels de l'Union européenne, qui interdit l'incitation à la haine, protège les enfants et encadre la publicité, le parrainage et le placement de produits[29].

76. Dans ce contexte, les régulateurs, ainsi que les organismes d'autorégulation de la publicité, à travers l'Europe ont produit des réglementations et des documents d'orientation afin de clarifier les obligations légales et éthiques des influenceurs, en particulier en ce qui concerne les contenus commerciaux, la transparence et la protection des enfants. Par exemple, le régulateur belge des médias audiovisuels a adopté un protocole pour les créateurs de contenu couvrant les contenus commerciaux, la protection des enfants et le discours de haine ; le régulateur estonien a publié des lignes directrices en matière d'étiquetage et élabore actuellement des règles pour les influenceurs en tant que fournisseurs de services de médias audiovisuels à la demande ; le régulateur norvégien a fixé des règles d'étiquetage obligatoires ; en Italie, les lignes directrices

25. Voir, par exemple, All European Academy (ALLEA), *The European Code of Conduct for Research Integrity – Revised Edition 2023,* Berlin, 2023, disponible en anglais seulement à l'adresse www.doi.org/10.26356/ECOC, par. 2.7.
26. Voir, par exemple, Union européenne, *Conclusions du Conseil sur le soutien aux influenceurs en tant que créateurs de contenus en ligne*, 23 juillet 2024, JO C/2024/3807, disponible à l'adresse https://op.europa.eu/s/AaUz.
27. France, *Loi n° 2023-451 du 9 juin 2023 visant à encadrer l'influence commerciale et à lutter contre les dérives des influenceurs sur les réseaux sociaux*, disponible à l'adresse www.legifrance.gouv.fr/eli/loi/2023/6/9/ECOX2308125L/jo/texte.
28. Espagne, *Loi n° 13/2022 du 7 juillet 2022, Loi générale sur la communication audiovisuelle*, disponible à l'adresse www.boe.es/eli/es/l/2022/07/07/13/con, et *Décret royal 444/2024 du 30 avril 2024*, disponible à l'adresse www.boe.es/eli/es/rd/2024/04/30/444/con.
29. Union européenne, *Directive 2010/13/UE du Parlement européen et du Conseil du 10 mars 2010 visant à la coordination de certaines dispositions législatives, réglementaires et administratives des États membres relatives à la fourniture de services de médias audiovisuels (directive « Services de médias audiovisuels »)*, telle que modifiée par la *Directive (UE) 2018/1808 du Parlement européen et du Conseil du 14 novembre 2018*. Le texte consolidé est disponible à l'adresse suivante : http://data.europa.eu/eli/dir/2010/13/2025-02-08. Pour un aperçu, voir Observatoire européen de l'audiovisuel, *National rules applicable to influencers*, précité.

en matière de régulation audiovisuelle s'appliquent aux influenceurs considérés comme des fournisseurs de services de médias audiovisuels. En Suède et en France, les organismes d'autorégulation de la publicité appliquent officiellement le Code de la publicité et de la communication commerciale de la Chambre de commerce internationale de 2024, qui couvre également les influenceurs[30]. La tendance générale est au renforcement de la transparence et de la protection des utilisateurs, souvent dans le cadre de la réglementation des services de médias audiovisuels, avec des attentes croissantes en matière de bonnes pratiques et de conformité.

77. Le paragraphe 32 décrit les responsabilités des parents et des représentants légaux des enfants qui, dans les limites d'âge et autres limites réglementaires pouvant être fixées par le droit interne, agissent en tant que créateurs de contenu. En revanche, il ne couvre pas le cas des parents ou tuteurs légaux qui, en tant que créateurs de contenu, utilisent ou exploitent l'image de leurs enfants (ce que l'on appelle le « *sharenting* »). La nécessité pour les parents et les représentants légaux d'agir dans l'intérêt supérieur de leurs enfants est un principe bien établi dans les normes internationales et celles du Conseil de l'Europe en matière de droits humains. La Convention des Nations unies relative aux droits de l'enfant, ainsi que la Recommandation CM/Rec(2018)7 sur les Lignes directrices relatives au respect, à la protection et à la réalisation des droits de l'enfant dans l'environnement numérique, soulignent que la dignité, la sécurité et le développement de l'enfant doivent guider toutes les actions le concernant. Lorsque les enfants agissent en tant que créateurs de contenu, les parents ou les représentants légaux ont un devoir de diligence particulier, non seulement pour se conformer aux règles applicables en matière de travail, de publicité et de protection des données, mais aussi pour préserver le bien-être de l'enfant. Dans la pratique, cela signifie veiller à ce que les enfants ne soient pas contraints de produire du contenu, que leur éducation, leur repos et leurs loisirs ne soient pas compromis par leurs activités en ligne et qu'ils soient protégés contre des risques tels que la surexposition, le harcèlement ou l'exploitation commerciale. Par exemple, les parents doivent être attentifs à la quantité d'informations personnelles qu'un enfant révèle en ligne et éviter que des contenus publiés puissent nuire au bien-être et à la réputation de l'enfant plus tard dans sa vie. L'objectif est de trouver un équilibre entre les possibilités offertes aux enfants de participer, de créer et de s'exprimer en ligne et la mise en place de garanties solides qui protègent leurs droits et leurs intérêts à long terme. Les enfants, en particulier lorsqu'ils agissent en tant que créateurs de contenu, devraient bénéficier d'une éducation numérique, ainsi que d'une éducation aux media et a l'information, afin de leur permettre de faire des choix éclairés, d'accéder à des voies de recours appropriées et à un soutien lorsque leurs contenus sont utilisés à mauvais escient ou entraînent une victimisation quelconque.

78. Des pratiques nationales commencent à voir le jour dans ce domaine. En France, par exemple, une législation spécifique réglemente l'exploitation commerciale de l'image des enfants de moins de seize ans sur les plateformes en ligne[31]. Cette loi, qui couvre également le cas où les enfants sont les sujets, plutôt que les auteurs, de contenus, introduit une procédure de contrôle par une autorité publique sur la production et la diffusion non occasionnelles de contenus, destinés à être distribués sur des plateformes, dans lesquels un enfant mineur de seize ans est le sujet principal. À la suite d'une déclaration des titulaires de l'autorité parentale, l'autorité compétente peut leur adresser des directives visant à préserver la santé et le bien-être de l'enfant, ainsi que le respect de son droit à l'éducation. En outre, elle protège les revenus qu'ils génèrent et consacre le droit à l'oubli, ce qui signifie que les plateformes seront tenues de supprimer les contenus à la demande de l'enfant.

IV. Principes applicables aux cadres juridiques sur la sécurité en ligne et l'autonomisation des utilisateurs, à leur déploiement et à leur mise en œuvre

Principes communs

Concernant le paragraphe 33

79. L'obligation positive qui incombe aux États d'aborder les risques de préjudice en ligne découle de leur devoir plus large de protéger les droits humains, y compris les droits à la vie privée et à la liberté d'expression, qui peuvent tous être affectés par les activités en ligne. De plus en plus, ces risques sont visés par des cadres

30. Pour des références et d'autres exemples, voir Observatoire européen de l'audiovisuel, *National rules applicable to influencers*, 2024, précité.
31. France, *Loi nº 2020-1266 du 19 octobre 2020 visant à encadrer l'exploitation commerciale de l'image d'enfants de moins de seize ans sur les plateformes en ligne*, texte consolidé disponible à l'adresse www.legifrance.gouv.fr/loda/id/JORFTEXT000042439054/.

législatifs nationaux ou supranationaux, tels que le Règlement de l'Union européenne sur les services numériques et la loi sur la sécurité en ligne du Royaume-Uni.

80. Dans le contexte de la sécurité en ligne, les cadres juridiques et réglementaires peuvent jouer un rôle essentiel pour responsabiliser les intermédiaires d'internet, car leurs infrastructures et leurs services peuvent être détournés pour diffuser rapidement de grandes quantités de contenus susceptibles de causer un préjudice, ainsi que pour permettre des comportements inauthentiques tels que l'utilisation de faux comptes ou l'augmentation artificielle de la visibilité des contenus afin d'élargir leur portée. Ces cadres constituent également une condition nécessaire pour que les mesures adoptées pour répondre aux préoccupations liées à la sécurité en ligne soient compatibles avec les droits humains des utilisateurs et, dans la mesure pertinente, des intermédiaires d'internet. Par conséquent, ces cadres juridiques devraient traiter à la fois les risques posés par certains types de contenus disponibles en ligne et le rôle des intermédiaires dans l'habilitation, la facilitation, l'amplification et la prévention de ces risques. Cependant, les États doivent également veiller à ce que ces cadres ne conduisent pas à des pratiques de conformité excessive ou à une mise en œuvre discriminatoire.

81. Le paragraphe 33 énumère les trois types complémentaires d'approches réglementaires que les États peuvent adopter en matière de règles de sécurité en ligne. La première catégorie, visée à la lettre (a), regroupe les règles qui limitent certains types d'expressions ou de manifestations de comportement, ainsi que leur diffusion, en raison de leur contenu, et prévoient les conséquences juridiques découlant de leur production et de leur diffusion. En général, ces règles liées au contenu ne sont pas propres à l'environnement en ligne, mais constituent l'extension à l'expression en ligne de limitations qui s'appliqueraient également dans d'autres contextes ou médias. Elles incluent notamment des dispositions de nature très variée, telles que celles qui interdisent les images d'abus sexuels sur enfants, l'incitation à la violence, le discours de haine, les propos diffamatoires ou la publicité trompeuse. Les règles relatives au contenu peuvent toutefois être adaptées à l'environnement en ligne, notamment lorsqu'il est nécessaire de prévoir des garanties spécifiques dans l'espace en ligne ou des conséquences liées à la nature du média utilisé. La deuxième catégorie, visée au point b), concerne les règles qui définissent les cas et les conditions dans lesquels les plateformes peuvent être exceptionnellement tenues juridiquement responsables du contenu généré par les utilisateurs qu'elles stockent.

82. Les principes applicables à ces deux types d'interventions sont rappelés et détaillés dans les sous-parties correspondantes de l'annexe. Toutefois, le paragraphe 33 souligne déjà que seules les formes d'expression et les types de contenu clairement définis comme étant restreints par la loi peuvent faire l'objet de restrictions en vertu des règles relatives au contenu. Cela implique également que les conditions dans lesquelles les intermédiaires peuvent être tenus responsables de cas spécifiques d'expression restreints par la loi générée par les utilisateurs doivent être définies par la loi.

83. Le paragraphe 33, lettre (c), concerne des règles qui suivent une approche systémique plutôt qu'une approche fondée sur le contenu en ce qui concerne les responsabilités des intermédiaires dans la promotion d'un environnement en ligne favorable. Elles mettent l'accent sur les processus par lesquels les intermédiaires d'internet classent, modèrent et suppriment le contenu, plutôt que sur le contenu spécifique ciblé, comme déjà établi dans de nombreuses normes du Conseil de l'Europe (CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication, principe 12, et son exposé des motifs (CM(2022)44-addfinal), par. 12.4 ; CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression, par. 1.6 ; Note d'orientation sur la lutte contre la propagation de la désinformation et des fausses informations en ligne, par. 23). Ce type de législation devrait traiter, notamment en favorisant la responsabilisation des plateformes et en renforçant l'autonomisation des utilisateurs, des obligations et responsabilités systémiques que les plateformes devraient avoir à l'égard de leurs propres systèmes et processus, y compris la manière dont ceux-ci peuvent affecter négativement les risques de préjudices en ligne.

Concernant le paragraphe 34

84. Les intermédiaires d'internet remplissent diverses fonctions et fournissent divers services (CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, préambule, paragraphe 4). Les plateformes, en particulier, font désormais partie intégrante des pratiques d'information et de communication des citoyens, exerçant une influence considérable sur la manière dont les utilisateurs produisent, accèdent à et interagissent avec les informations et les contenus médiatiques. Lorsque les plateformes modèrent et classent les contenus, notamment par le traitement automatisé des données à caractère personnel, elles exercent des formes de contrôle qui influencent l'accès des utilisateurs à l'information en ligne d'une manière comparable à celle des médias, ou elles peuvent remplir d'autres fonctions qui s'apparentent à celles des éditeurs (CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, préambule, paragraphe 5). Néanmoins, malgré

leur pouvoir et leur rôle, les plateformes continuent de fonctionner comme des intermédiaires. Ce rôle, même lorsqu'il implique la gestion des contenus générés par leurs utilisateurs et leur organisation et sélection par le biais d'un système algorithmique, nécessite une approche différenciée dans leur gouvernance par rapport à celle des acteurs des médias (CM/Rec(2022)11 sur les principes de gouvernance de la communication et des médias). La distinction essentielle réside dans la responsabilité : les plateformes ne sont généralement pas responsables pour les contenus spécifiques générés par leurs utilisateurs, contrairement aux médias qui exercent une responsabilité éditoriale et peuvent donc être tenus responsables des contenus qu'ils publient. L'absence de responsabilité éditoriale ne dégage toutefois pas les plateformes de toutes leurs obligations et responsabilités. Comme l'a fait remarquer la Cour dans l'affaire *Google LLC et autres c. Russie,* « lorsque les intermédiaires d'internet gèrent les contenus disponibles sur leurs plateformes ou jouent un rôle de curateur ou d'éditeur, y compris par le biais d'algorithmes, leur fonction importante dans la facilitation et l'orientation du débat public engendre des obligations de diligence raisonnable et de prudence, qui peuvent également augmenter proportionnellement à la portée de l'activité expressive concernée » (précité, par. 79 ; voir également l'exposé des motifs concernant la partie sur les règles relatives à la responsabilité des intermédiaires ci-dessous). Toutefois, la nécessité de différencier les responsabilités des plateformes de celles des médias traditionnels ne signifie pas que certaines règles ayant le même objectif ne puissent pas s'appliquer à la fois aux plateformes et aux médias. Par exemple, les plateformes et les médias devraient avoir l'obligation de ne pas diffuser de contenu illégal. On peut citer à titre d'exemple l'ensemble d'obligations imposées aux plateformes de partage de vidéos par l'article 28.b de la Directive de l'Union européenne sur les services de médias audiovisuels, qui vise à protéger les mineurs et le grand public contre certains types de contenus qui sont également soumis à des restrictions pour les fournisseurs de services de médias audiovisuels. Même si les responsabilités des plateformes et des médias à cet égard sont différentes, les règles applicables aux deux types d'acteurs partagent le même objectif, qui est de prévenir les conséquences préjudiciables pour les utilisateurs.

Concernant les paragraphes 35 et 36

85. Les paragraphes 35 et 36 réaffirment le principe fondamental d'un cadre de gouvernance pour les intermédiaires d'internet fondé sur les droits humains, tel qu'énoncé dans la Recommandation CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression (par. 1.1). Les réponses apportées par les États aux risques en ligne devraient identifier clairement, dans les cadres juridiques pertinents, les contenus qui doivent faire l'objet de mesures constituant une ingérence dans l'exercice du droit à la liberté d'expression des utilisateurs, tels que le retrait ou le déclassement de contenus, ou la suspension ou résiliation des comptes d'utilisateurs spécifiques. Cela découle directement des exigences de l'article 10, paragraphe 2, de la Convention, en particulier du fait que toute limitation à l'exercice de la liberté d'expression doit être prévue par la loi. Les contenus qui ne font pas l'objet de restrictions par la loi, en particulier lorsque ces restrictions auraient un effet disproportionné sur les droits consacrés à l'article 10 de la Convention, peuvent néanmoins susciter des inquiétudes quant à leur impact potentiel sur d'autres droits humains et, par conséquent, sur la sécurité et le bien-être des utilisateurs. Les États devraient traiter ces contenus exclusivement par le biais d'autres mesures qui n'interfèrent pas avec le droit à la liberté d'expression des utilisateurs et qui visent plutôt à atténuer les risques de préjudice. Il s'agit notamment des mesures relatives à la responsabilisation des plateformes et à l'autonomisation des utilisateurs énoncées dans les parties pertinentes de l'annexe.

86. Toute politique ou action des autorités publiques qui interfère avec le droit à la liberté d'expression devrait être prescrite par la loi, poursuivre un but légitime et répondre aux exigences de sécurité juridique, de nécessité et de prévisibilité (CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression, paragraphe 1.2). Les obligations et responsabilités systémiques générales ne doivent être ni interprétées comme exigeant la suppression ou la limitation de contenus légaux, ni mises en œuvre de façon à imposer aux plateformes des restrictions fondées sur le contenu en dehors des cas prévus par la loi. Par exemple, les plateformes peuvent être tenues de mettre en place des mesures systémiques pour atténuer le risque que des enfants accèdent à des contenus qui leur sont interdits par la loi, telles que des systèmes de vérification de l'âge ou la mise en place de systèmes de recommandation visant à empêcher les enfants d'être exposés à des recommandations de contenus qui pourraient présenter un risque pour leur sécurité, en particulier lorsqu'ils y sont exposés de manière répétée. Toutefois, elles ne devraient pas être tenues de supprimer ces contenus pour l'ensemble de leurs utilisateurs afin de protéger les enfants. De même, les mesures d'atténuation qui peuvent être exigées par les autorités réglementaires pour lutter contre les risques posés au débat public éclairé par des contenus délibérément trompeurs qui ne sont pas en soi restreints par la loi devraient se concentrer sur des critères neutres quant au contenu, tels que l'existence de schémas de diffusion susceptibles de révéler un comportement inauthentique coordonné.

87. Les plateformes ont souvent des règles et des politiques de contenu qui interdisent certains types de contenu, même si celui-ci n'est pas restreint par le droit interne. Ces politiques et règles sont généralement décrites dans les conditions d'utilisation ou les règles de communauté des plateformes, que les utilisateurs acceptent lors de l'inscription, et elles constituent des engagements contractuels entre la plateforme et ses utilisateurs. Si les États ne doivent pas exploiter ces règles internes pour imposer *de facto* des obligations de suppression d'expressions à caractère légal, ils doivent toutefois tenir les plateformes responsables de la manière dont ces règles sont appliquées et mises en œuvre. Dans ce contexte, le contrôle exercé par l'État devrait se concentrer sur les garanties procédurales, notamment la transparence, la protection contre les décisions arbitraires et la mise en place de mécanismes de recours, afin de garantir la protection des droits et libertés, en veillant, par exemple, à ce que le contenu ne soit pas supprimé ou restreint d'une manière qui porte atteinte de manière injustifiée à la liberté d'expression.

Concernant le paragraphe 37

88. Le blocage ou l'interdiction d'accès à l'ensemble d'un site web, d'un domaine ou d'une plateforme en ligne est considéré comme l'une des formes les plus graves d'ingérence dans le droit à la liberté d'expression. De telles mesures restreignent non seulement l'accès à certains contenus illégaux, mais suppriment souvent une grande quantité d'expressions à caractère légal. La Recommandation CM/Rec(2016)5 sur la liberté de internet stipule que toute mesure de ce type prise par les autorités publiques ou toute demande émanant de celles-ci visant à mettre en œuvre de telles mesures doit respecter les conditions énoncées à l'article 10 de la Convention concernant la légalité, la présence d'un but légitime et la proportionnalité des restrictions.

89. Le blocage total d'une plateforme en ligne dans son ensemble constitue une ingérence qui ne peut être ordonnée que comme mesure de dernier ressort, dans des cas très exceptionnels. Toute préoccupation liée au contenu doit être traitée par des mesures proportionnées et fondées sur des éléments factuels, telles que la suppression ciblée de contenus illégaux, ou par l'imposition d'obligations et de responsabilités systémiques pour répondre à des préoccupations légitimes, telles que la protection et le bien-être des enfants. La Cour européenne des droits de l'homme a traité cette question dans plusieurs affaires. Dans l'affaire *Ahmet Yıldırım c. Turquie* (n° 3111/10, 18 décembre 2012), la Cour a estimé que la décision des autorités turques de bloquer l'ensemble de la plateforme Google Sites en raison d'une page prétendument illégale constituait une restriction disproportionnée et non nécessaire, car elle rendait inaccessibles de grandes quantités d'informations, ce qui affectait directement les droits des utilisateurs d'internet et avait un effet collatéral important. De même, dans l'affaire *Cengiz et autres c. Turquie* (nos 48226/10 et 14027/11, 1er décembre 2015), la Cour a jugé que le blocage total de YouTube – sur la base de quelques vidéos jugées illégales – constituait une violation du droit des requérants de recevoir et de communiquer des informations.

90. Les restrictions d'accès à un service, à un domaine ou à un site web devraient être considérées comme une mesure de dernier recours. Elles devraient être ordonnée par une autorité judiciaire ou une autre autorité publique indépendante, telle qu'une autorité de régulation indépendante, dont les décisions sont soumises à un contrôle juridictionnel, et n'être appliquées que dans les cas les plus graves. Cela peut être le cas, par exemple, lorsqu'un service héberge exclusivement des contenus illégaux ou est utilisé pour commettre des infractions pénales mettant en danger la vie ou la sécurité des personnes. Ces restrictions doivent être spécifiques, temporaires et imposées uniquement lorsque d'autres mesures n'ont pas produit l'effet voulu et que l'infraction en question continue de causer un préjudice grave. Des mesures préliminaires pourraient être prévues dans la législation, telles que le pouvoir conféré aux autorités publiques compétentes d'ordonner la cessation de l'infraction ou d'adopter des mesures provisoires visant à prévenir le risque de préjudice grave. Une autre circonstance dans laquelle l'accès à un service pourrait faire l'objet d'une restriction générale se présente lorsqu'une plateforme manque ou refuse de manière persistante de se conformer à des obligations légales ou à des exigences réglementaires conformes aux droits humains, dans le cadre de la législation sur la responsabilité des plateformes, à condition que ce non-respect entraîne un préjudice réel. Parmi ces dommages, on peut mentionner, entre autres, la diffusion à grande échelle de contenus illégaux, tels que des matériels d'abus sexuels sur enfants ou des images à caractère sexuel partagées sans consentement, ou le fait de permettre aux enfants d'accéder à du matériel pornographique.

91. Les États devraient prévoir des garanties procédurales efficaces, notamment un contrôle judiciaire indépendant, des possibilités de recours ou de révision tant pour les opérateurs que pour les utilisateurs concernés, et des limites claires au pouvoir discrétionnaire des autorités publiques. La Recommandation CM/Rec(2016)5 sur la liberté de internet définit également des obligations en matière de transparence (par. 2.2.5). Elle recommande aux États de publier des informations sur les sites web qui ont été bloqués ou sur lesquels des informations ont été supprimées, y compris des détails sur la base juridique, la nécessité et la

justification de ces restrictions, la décision judiciaire qui les autorise et le droit de recours. Des indications supplémentaires sur le contenu de ces rapports que les États doivent publier figurent dans l'exposé des motifs relatif à la Recommandation CM/Rec(2022)16 sur la lutte contre le discours de haine (CM(2022)43-addfinal, par. 107). Des explications pertinentes pourraient également être fournies par les boutiques d'applications ou les fournisseurs d'accès à internet lorsque les utilisateurs recherchent une application qui a été supprimée ou tentent d'accéder à un site web bloqué. Ces explications devraient clairement indiquer les raisons de la restriction et offrir aux utilisateurs concernés des indications pratiques sur la manière de s'y opposer.

Concernant le paragraphe 38

92. Conformément à l'article 10 de la Convention, et comme le rappelle la Recommandation CM/Rec(2018)2 sur le rôle et les responsabilités des intermédiaires d'internet, toute demande, exigence ou autre mesure adressée par les autorités publiques aux intermédiaires d'internet qui porte atteinte aux droits humains et aux libertés fondamentales doit être prévue par la loi. Les États ne devraient pas exercer de pression sur les intermédiaires d'internet par des moyens qui ne sont pas prévus par la loi. Le même principe devrait s'appliquer aux mesures visant les créateurs de contenu, en particulier lorsque ces mesures pourraient porter atteinte à leur liberté d'expression. Tout pouvoir de ce type accordé aux autorités publiques, y compris aux autorités chargées de l'application de la loi, devrait être clairement défini afin d'éviter toute application arbitraire.

93. Imposer des mesures aux intermédiaires d'internet et aux créateurs de contenu qui affectent la disponibilité du contenu sans base juridique claire est particulièrement préoccupant en période de crise et dans des circonstances exceptionnelles pouvant menacer la sécurité ou la santé publiques, telles que les conflits, les troubles civils, les actes de terrorisme, les catastrophes naturelles ou les urgences de santé publique. La Recommandation CM/Rec(2024)7 sur la protection effective des droits humains en situation de crise fournit aux États un cadre général sur la manière d'exercer leurs pouvoirs dans de telles situations et rappelle que « [l]es États membres devraient préserver la liberté d'expression et l'accès du public à des informations précises et fiables en situation de crise » (par. 16). Les demandes et les mesures prises en réponse à de telles crises ne devraient être adoptées que dans la mesure où elles sont strictement nécessaires et urgentes, en tenant compte du fait que ces mesures doivent être proportionnées à la gravité de la situation et à leurs implications potentielles sur les droits et les intérêts, y compris la liberté d'expression. En outre, ces mesures ne devraient s'appliquer que pendant une période limitée, telle que prévue par la loi. Le cas échéant, les mesures liées à la crise devraient être fondées sur des obligations de diligence renforcées pour les plateformes, comme le prévoit le mécanisme de réponse aux crises de l'article 36 du Règlement de l'Union européenne sur les services numériques. Il peut s'agir, par exemple, d'adapter les processus de modération de contenu, notamment en augmentant les ressources qu'y sont consacrées, en intensifiant la coopération avec les signaleurs de confiance, en mettant en œuvre des mesures de sensibilisation et en promouvant les informations fiables (voir, par exemple, Règlement de l'Union européenne sur les services numériques, préambule, considérant 91).

94. Les États devraient veiller à ce que la législation comporte des garanties contre tout abus de pouvoir, tel que la délivrance d'injonctions par les autorités administratives ou l'exercice d'autres formes de pression réglementaire sur les intermédiaires internet sans base juridique. La législation devrait également accorder aux intermédiaires et aux utilisateurs le droit à un recours effectif dans les cas où une telle pression est exercée.

Concernant le paragraphe 41

95. Ce paragraphe renforce l'approche graduée et proportionnée de la réglementation des plateformes, telle que soulignée par la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, selon laquelle la portée et l'approche adoptées pour remplir les responsabilités des plateformes peuvent varier en fonction du risque de préjudice et de la gravité potentielle de leur impact sur les droits humains (par. 2.1.2). Les micro-plateformes et les petites plateformes fonctionnent généralement avec des ressources limitées et exercent une influence minimale sur le débat public. Soumettre ces acteurs aux mêmes exigences réglementaires que les plateformes opérant à l'échelle mondiale, dotées de capacités technologiques avancées et de ressources économiques plus importantes, risque de freiner la croissance, de décourager les nouveaux entrants et de compromettre les objectifs plus larges d'un écosystème numérique diversifié et concurrentiel.

96. Les plateformes exerçant une influence significative se définissent par leur capacité à façonner l'environnement informationnel et à influer sur la jouissance des droits humains. À ce titre, elles devraient être

soumises à des obligations de diligence plus strictes et à un contrôle réglementaire plus rigoureux, telles que l'obligation de procéder à des évaluations des risques et le respect d'exigences accrues en matière de transparence.

97. En proposant des critères alternatifs pour définir les différentes catégories de plateformes, le paragraphe 41 laisse aux États une grande marge de manœuvre pour décider des critères appropriés afin de graduer les responsabilités des différentes plateformes. Il permet l'utilisation de critères quantitatifs, tels que la base d'utilisateurs, comme c'est le cas dans le Règlement de l'Union européenne sur les services numériques pour la désignation des très grandes plateformes en ligne et des moteurs de recherche (article 33). Le cadre réglementaire peut également, à cette fin, recourir à des critères qualitatifs, tels que le niveau d'impact sociétal ou le niveau de risque des plateformes. Les critères sur la base desquels le niveau de risque des différents intermédiaires d'internet est évalué doivent être clairement précisés, révisés périodiquement, mesurés avec précision et communiqués de manière transparente par l'autorité publique compétente (CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression, par. 1.4).

98. L'exemption générale de responsabilité des intermédiaires pour les contenus générés par les utilisateurs et ses exceptions, telle que détaillée aux paragraphes 54 à 56 de l'annexe, devrait toujours s'appliquer à tous les intermédiaires d'internet, y compris les plateformes, indépendamment de leur taille, de leur portée ou de leur impact. Toutefois, lorsque, exceptionnellement, les exemptions de responsabilité ne s'appliquent pas dans les circonstances spécifiques examinées par la Cour dans l'affaire *Delfi AS c. Estonie* [GC] (précité, par. 115-116, voir également l'exposé des motifs concernant le paragraphe 55), le type et la taille du fournisseur peuvent constituer des facteurs déterminants pour limiter cette responsabilité et éviter des exigences disproportionnées à leur égard (*Pihl c. Suède* (déc.), n° 74742/14, 7 février 2017, paragraphe 31 ; *Magyar Tartalomszolgáltatók Egyesülete et Index.hu Zrt c. Hongrie,* n° 22947/13, 2 février 2016, disponible seulement en anglais, par. 82).

Concernant le paragraphe 42

99. Le paragraphe 42 réitère un principe inscrit dans la Recommandation CM/Rec(2018)2 sur le rôle et les responsabilités des intermédiaires d'internet. Les États ne devraient imposer aucune mesure susceptible d'obliger les intermédiaires à surveiller systématiquement l'activité des utilisateurs ou à rechercher activement des faits ou des circonstances qui indiqueraient une activité illégale, par exemple en exigeant l'utilisation de systèmes automatisés pour analyser de manière proactive les données ou le contenu des utilisateurs à la recherche d'infractions à la loi. Toutefois, les intermédiaires d'internet devraient pouvoir mener volontairement leurs propres enquêtes sur les contenus illégaux. Cette attitude proactive ne devrait pas, en soi, engager leur responsabilité.

Concernant le paragraphe 43

100. À l'ère numérique, la réglementation de la sécurité en ligne ne peut pas être efficacement mise en œuvre par des mesures nationales isolées. Compte tenu de la nature transfrontalière de l'internet, les États devraient coopérer afin de garantir que leurs cadres juridiques et réglementaires reposent solidement sur les normes et principes internationaux en matière de droits humains et qu'ils soient aussi cohérents que possible. Cela s'avère particulièrement crucial dans les régions avec des contextes linguistiques, culturels ou politiques partagés, car elles sont souvent confrontées à des risques en ligne et à des impacts sociétaux similaires.

101. Une approche commune en matière de réglementation et de coopération dans l'application des règles est essentielle tant du point de vue du marché que de la sécurité juridique. Des règles harmonisées peuvent favoriser un meilleur respect par les plateformes numériques opérant dans plusieurs juridictions. En revanche, des approches réglementaires fragmentées peuvent entraîner une application incohérente et une protection inégale des utilisateurs. Par exemple, des législations divergentes peuvent conduire à un « *forum shopping* », les plateformes choisissant d'opérer dans les juridictions les plus indulgentes, ce qui compromet les normes de protection.

102. Il convient également d'encourager une participation active au sein des organismes intergouvernementaux, des réseaux transnationaux et des autres organisations internationales concernées afin d'assurer l'harmonisation avec les instruments européens et internationaux pertinents (voir également CM/Rec(2022)16 sur la lutte contre le discours de haine, par. 63).

103. Les cadres conventionnels existants relatifs aux comportements criminels en ligne et à la coopération dans le cadre des enquêtes et des poursuites relatives à ces crimes, tels que la Convention sur la

cybercriminalité et ses deux protocoles additionnels (STE n° 189 et STCE n° 224), fournissent des exemples de coopération intergouvernementale renforcée. Ces instruments visent à garantir des approches communes en matière de contenus illégaux et une coopération efficace dans leur mise en œuvre transfrontalière, tout en respectant les droits humains.

104. Les États peuvent également mettre en place des mécanismes de coordination réglementaire transfrontalière, comprenant notamment des protocoles d'enquêtes conjointes, des normes techniques communes et, le cas échéant, des mesures d'application coordonnées.

Règles relatives aux contenus

Concernant le paragraphe 44

105. Conformément à la Recommandation CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication, la promotion des droits humains et des libertés fondamentales dans la communication implique « également d'harmoniser les règles dans les environnements en ligne et hors ligne » (principe 6). L'exposé des motifs concernant cette disposition précise que cela suppose de prendre en compte de certaines différences, sans pourtant « introduire une réglementation plus stricte des contenus diffusés par le biais des plateformes » (par. 6.2). Comme l'a souligné, par exemple, la Commission de Venise, « la mise en place d'un régime spécifique uniquement applicable aux éditions électroniques des médias écrits engendre une différence de traitement juridique de contenus identiques. ... Toute distinction entre les réglementations juridiques applicables à la presse écrite, à la presse en ligne et aux médias audiovisuels doit être justifiée » (CDL-AD(2020)013, Albanie - Avis sur le projet de modification de la loi n° 97/2013 sur les Services des médias audiovisuels, par. 26). L'environnement en ligne ne devrait donc pas être soumis à des restrictions arbitraires ou disproportionnées qui ne seraient pas admissibles hors ligne. Toutefois, les plateformes en ligne et la communication numérique présentent des défis spécifiques et introduisent de nouveaux risques en raison de l'ampleur et de la vitesse de la diffusion, du potentiel d'anonymat ou de l'amplification algorithmique. Ces facteurs peuvent nécessiter des mesures de gouvernance adaptées pour atténuer les préjudices, comme indiqué ci-dessous.

Concernant le paragraphe 45

106. Les règles relatives au contenu visent à restreindre la publication, la diffusion ou l'accessibilité de certains types d'expressions ou de manifestations de comportements. En tant que telles, elles affectent directement la capacité des individus et des groupes à recevoir et à diffuser des informations, y compris par le biais des technologies numériques. Conformément aux principes énoncés dans la Recommandation CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication, le paragraphe 45 souligne la nécessité d'un processus normatif transparent et fondé sur des informations factuelles, condition indispensable pour satisfaire aux critères de nécessité et de proportionnalité prévus à l'article 10, paragraphe 2, de la Convention. Ces informations factuelles doivent confirmer l'existence d'atteintes claires aux droits qui exigent d'être traitées par des mesures restrictives des contenus. Comme indiqué au paragraphe 1.1.4 de la Recommandation CM/Rec(2018)2 sur le rôle et les responsabilités des intermédiaires de l'internet, l'adoption de lois ou de règlements devrait être précédée d'évaluations d'impact sur les droits humains.

107. Une collecte transparente des éléments factuels implique que les États, lorsqu'ils définissent des politiques et des cadres juridiques visant des règles relatives aux contenus, devraient solliciter la participation d'un large éventail d'acteurs non étatiques, notamment la société civile et la communauté scientifique, et soutenir la recherche indépendante visant à comprendre et à traiter de manière adéquate et spécifique les risques.

108. Une approche proportionnée à la restriction des contenus exige que les États définissent de manière limitée et juridiquement claire les types de contenus à limiter, choisissent des mesures correctives ciblées et aussi peu intrusives que possible, et qu'ils n'appliquent les restrictions que dans le périmètre géographique strictement nécessaire, afin d'éviter des effets extraterritoriaux[32].

Concernant le paragraphe 46

109. L'article 10, paragraphe 2, de la Convention exige que toute ingérence dans l'exercice de la liberté d'expression soit « prévue par la loi ». Ce paragraphe de l'annexe traite du principe de légalité, du point de vue de

32. Voir, par exemple, Cour de justice de l'Union européenne (grande chambre), affaire C-507/17, *Google LLC c. CNIL*, 24 septembre 2019, sur l'absence d'effets extraterritoriaux du droit à l'oubli.

la « qualité » que doit présenter une telle loi. Selon la Cour, on ne peut considérer comme une « loi » qu'une norme énoncée avec assez de précision pour permettre aux justiciables de régler leur conduite et de prévoir, en s'entourant au besoin de conseils éclairés, de prévoir les conséquences à dériver d'un acte déterminé, à un degré raisonnable dans les circonstances concrètes. En même temps, ces conséquences n'ont pas à être prévisibles avec une certitude absolue, la loi pouvant laisser une marge d'interprétation (*Perinçek c. Suisse* [GC], nº 27510/08, 15 octobre 2015, par. 131 ; *Sanchez c. France* [GC], nº 45581/15, 15 mai 2023, par. 125-126). Seuls les contenus définis par la loi de manière suffisamment claire et précise devraient faire l'objet de restrictions portant atteinte à la liberté d'expression.

110. La loi doit être accessible et prévisible, ce qui signifie que les individus doivent être en mesure de comprendre ce qui est interdit et quelles en sont les conséquences. Une telle clarté et prévisibilité juridiques profitent à toutes les parties concernées. Les autorités publiques chargées d'ordonner la suppression ou le blocage de contenus soumis à des restrictions par la loi devraient être en mesure de déterminer clairement et sans ambiguïté si un contenu spécifique relève de leur mandat, en permettant une application rapide et efficace de la loi. Les intermédiaires sont mieux à même de se conformer à leurs obligations légales lorsqu'ils comprennent précisément ce qui leur est demandé. Cela réduit le risque de retraits excessifs pris par précaution. Enfin, les utilisateurs bénéficient d'une plus grande transparence et d'une meilleure protection juridique, leurs droits et libertés étant moins susceptibles d'être restreints de manière injustifiée. En revanche, des définitions vagues ou excessivement larges ouvrent la voie à des interprétations subjectives, qui peuvent entraîner une application arbitraire de la loi et des entraves disproportionnées et injustifiées à la liberté d'expression (CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression, par. 1.3). À titre d'exemple, la Commission de Venise a reconnu que « les interdictions de diffuser des informations fondées sur des notions vagues et ambiguës, telles que les "fausses nouvelles" ou les "informations non objectives", sont incompatibles avec les normes internationales en matière de restrictions à la liberté d'expression ... et devraient être abolies » (CDL-AD(2019)016, Rapport conjoint sur les technologies numériques et les élections, par. 90).

111. La Cour a souligné que la portée des notions de prévisibilité et d'accessibilité dépend dans une large mesure du contenu du texte concerné, du domaine qu'il couvre ainsi que du nombre et de la qualité de ses destinataires. Les règles relatives aux contenus régissant la conduite générale des utilisateurs devraient donc être formulées avec un degré de clarté et de précision plus élevé que celles qui s'adressent aux créateurs de contenu professionnels ou aux plateformes : les professionnels sont censés être mieux conscients de leurs responsabilités légales et apporter un soin particulier à l'évaluation des risques que leur activité comporte (*Chauvy et autres c. France*, nº 64915/01, 29 juin 2004, par. 43-45). En particulier pour les dispositions pénales (telles que celles relatives au discours de haine), la Cour exige que la portée des infractions concernées soit définie « de manière claire et précise », afin d'éviter que le pouvoir discrétionnaire de l'État de poursuivre ces infractions ne devienne trop large et ne donne lieu à des abus par le biais d'une application sélective (*Savva Terentyev c. Russie*, précité, par. 85 ; voir également *Altuğ Taner Akçam c. Turquie*, nº 27520/07, 25 octobre 2011, disponible seulement en anglais, par. 93-94). De même, les contenus soumis aux restrictions les plus sévères, telles que le retrait ou le blocage, doivent être définis avec le plus haut degré de clarté et de précision juridiques.

Concernant le paragraphe 47

112. Le principe de proportionnalité exige que les restrictions soient adaptées à la gravité du préjudice causé, ou susceptible d'être causé, par différents types de contenus. Le droit interne, y compris lorsque cela est requis par les obligations découlant de la Convention ou d'autres instruments internationaux relatifs aux droits humains, devrait exiger des plateformes qu'elles suppriment ou bloquent les formes les plus graves de contenus préjudiciables, tels que les le matériel d'abus sexuels sur enfants, la propagande terroriste ou extrémiste violente, ou les discours incitant directement à la violence ou à la haine, qui devraient être interdits par le droit pénal. Ces mesures sont justifiées par le préjudice grave, ou le risque imminent de préjudice, causé par le contenu concerné. Le processus de retrait devrait être conforme aux principes de légalité, de nécessité et de proportionnalité, et respecter les normes internationales en matière de droits humains, ainsi que les principes énoncés dans la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet. En outre, conformément au paragraphe 2.3.6 de cette dernière, lorsque les plateformes suppriment ou bloquent des contenus illégaux, elles doivent veiller à la conservation d'éléments probatoires en vue d'enquêtes pénales efficaces. Les restrictions, y compris le blocage ou la suppression, peuvent également être justifiées dans les cas où le contenu viole des dispositions de droit administratif ou civil qui ne se limitent pas à l'environnement en ligne, suivant le principe selon lequel ce qui est illégal hors ligne doit également être illégal en ligne. Tel serait le cas, par exemple, des violations du droit d'auteur ou des contenus constituant un discours de haine illégal qui n'atteignent pas le seuil de responsabilité pénale

(voir CM/Rec(2022)16 sur la lutte contre le discours de haine, par. 3). Enfin, des restrictions et des règles spécifiques liées au contenu peuvent être imposées à certains services réglementés, tels que les services de médias audiovisuels, comme le prévoit la législation sectorielle spécifique (voir également l'exposé des motifs concernant le paragraphe 53). Toutefois, les États doivent veiller à ce que les mesures de modération de contenu soient adaptées au type et à la nature des différents contenus restreints par la loi, ainsi qu'au problème spécifique qu'ils cherchent à résoudre, en tenant compte du fait que les contenus restreints par la loi varient considérablement quant à leur caractéristiques et à la gravité de leurs conséquences[33], ainsi qu'à assurer qu'il existe des garanties contre les restrictions injustifiées de la liberté d'expression.

113. Certains types de contenus, bien qu'ils ne soient pas illégaux en toutes circonstances, peuvent néanmoins présenter un risque selon le contexte. Dans ces cas, les risques liés à certains types de contenus soumis par la loi à des restrictions pourraient être efficacement atténués par des mesures moins intrusives. Ces contenus, dont l'identification par le droit interne reste soumise au principe de légalité et de sécurité juridique, sont considérés comme des contenus légaux mais réglementés. En voici quelques exemples : l'accès restreint aux contenus violents ou pornographiques en fonction de l'âge, qui permet d'éviter que les enfants y soient exposés sans pour autant supprimer ces contenus destinés à un public adulte ; les règles de transparence et de ciblage pour la publicité politique ou la promotion commerciale de certains produits et services (par exemple, les jeux de hasard ou l'alcool) ; et la réduction de la visibilité dans les résultats de recherche ou les flux de recommandations, ce qui réduit la diffusion et la visibilité, ou la démonétisation, qui supprime les incitations financières, de contenus par ailleurs licites mais présentant des risques de préjudice, tels que certaines formes de désinformation. Ces mesures ne comportent pas le retrait du contenu concerné, mais limitent les effets néfastes de sa diffusion non réglementée. À cet égard, il convient également d'accorder une attention particulière à l'existence d'une législation en matière de responsabilisation, qui établit des devoirs et des responsabilités systémiques pour les plateformes dans le cadre d'un modèle de gouvernance favorisant l'autonomisation, comme décrit aux paragraphes 57 et 58 de l'annexe à la recommandation et aux paragraphes 149 à 152 du présent Exposé des motifs.

114. Les restrictions d'accès à des contenus légaux mais réglementés doivent toujours être évaluées au cas par cas. Elles ne doivent pas être présumées comme moins intrusives ou moins impactantes que les mesures appliquées aux contenus illégaux. Par exemple, le déclassement de contenus légaux peut les rendre pratiquement invisibles. De telles mesures peuvent avoir un effet dissuasif sur la liberté d'expression comparable à celui d'un retrait pur et simple. Par conséquent, avant d'imposer toute restriction sur un contenu, les autorités publiques devraient procéder à une évaluation approfondie et fondée sur des preuves des avantages escomptés et des conséquences imprévues possibles. Cela implique notamment l'évaluation de l'impact que la mesure proposée pourrait avoir sur la liberté d'expression, le pluralisme des médias, l'accès à l'information et d'autres droits humains. Les autorités devraient également tenir compte des effets plus larges sur les communautés et l'environnement en ligne. Une fois qu'une mesure est mise en œuvre, son efficacité et sa proportionnalité devraient être régulièrement suivies et réévaluées afin d'éviter tout excès et abus, et une suppression excessive « par ricochet » de contenus licites. Conformément au principe de proportionnalité, la mesure la moins restrictive et la plus ciblée capable d'atteindre le but légitime devrait toujours être privilégiée.

115. Les États devraient réexaminer périodiquement leurs cadres législatifs et réglementaires relatifs au contenu afin de garantir leur clarté, leur efficacité et leur adéquation aux évolutions sociétales, technologiques et juridiques. Cet examen devrait considérer à la fois si leur portée est trop large ou, au contraire, trop limitée. Ces révisions devraient s'appuyer sur la collecte d'informations factuelles transparentes et indépendantes, et sur des consultations inclusives impliquant la société civile, les chercheurs, les intermédiaires et les communautés concernées. En outre, les États devraient soutenir, y compris financièrement, la recherche indépendante sur les risques de préjudice en ligne, les risques liés au contenu et l'évolution du paysage numérique, afin d'adapter les réponses politiques aux nouveaux défis et de mieux cibler les mesures existantes.

Concernant les paragraphes 48 et 49

116. Conformément au principe de légalité et de sécurité juridique, les autorités publiques, lorsqu'elles appliquent des restrictions spécifiques au contenu, notamment en exigeant des intermédiaires qu'ils restreignent l'accès ou suppriment des contenus, ne devraient le faire que pour les contenus qui entrent dans les catégories de contenus restreints par la loi en vertu du droit interne, et dans les limites que celui-ci leur

33. Conseil de l'Europe, *Note d'orientation sur la modération de contenu. Meilleures pratiques en vue de la mise en place de cadres juridiques et procéduraux efficaces pour les mécanismes d'autorégulation et de corégulation de la modération de contenu*, 2021, adoptée par le CDMSI lors de sa 19e réunion, 19-21 mai 2021, pp. 36-39, disponible à l'adresse https://go.coe.int/D6PMH, ci-après « Note d'orientation sur la modération de contenu ».

confère. La loi devrait également prévoir des garanties contre l'exercice sélectif, discriminatoire ou arbitraire de tout pouvoir d'imposer la restriction de contenus en ligne. Avant et après avoir ordonné l'application d'une restriction, les autorités publiques devraient évaluer soigneusement l'impact que cette mesure peut avoir, y compris sur la liberté d'expression, et devraient appliquer la mesure la moins intrusive possible (CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, par. 1.3.1).

117. Le paragraphe 49 réaffirme le principe énoncé au paragraphe 1.3.2 de la Recommandation CM/Rec(2018)2, selon lequel les autorités publiques devraient obtenir une décision formelle d'une autorité judiciaire ou d'une autre instance administrative indépendante, telle qu'un régulateur indépendant, dont les décisions sont soumises à un contrôle juridictionnel, lorsqu'elles exigent d'un intermédiaire la restriction de l'accès à des contenus. Les notifications reçues d'autres autorités publiques ne devraient pas être automatiquement considérées comme des injonctions juridiquement contraignantes obligeant les intermédiaires à s'y conformer. Dans ces cas, toutefois, l'absence de mesures peut aussi avoir des conséquences sur le régime de responsabilité des intermédiaires (voir l'exposé des motifs concernant les paragraphes 54 à 56). Lorsque le droit interne confère à une autorité publique le pouvoir d'émettre des « notifications » qui entraînent des conséquences juridiques autres que l'application d'une restriction, telles que l'obligation de donner la priorité à l'examen de la légalité d'un contenu donné, celles-ci devraient être traitées comme des « injonctions » aux fins de la présente recommandation. Le mandat et l'étendue des pouvoirs conférés aux autorités publiques habilitées à émettre de telles injonctions devraient être clairement définis dans la loi.

118. L'utilisation de l'expression « en principe » indique qu'il peut y avoir des situations dans lesquelles la gravité ou l'imminence d'un préjudice peut justifier des interventions rapides de la part d'autorités autres que celles mentionnées au paragraphe 49. Le paragraphe 1.3.2. de la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, par exemple, indique que l'exigence ci-dessus ne s'applique pas dans les cas impliquant des contenus manifestement illégaux, tels que matériel d'abus sexuels sur les enfant, ou dans les cas où des mesures accélérées sont requises et conforment aux conditions prescrites à l'article 10 de la Convention. Les cas dans lesquels des autorités publiques autres que judiciaires ou indépendantes peuvent ordonner des restrictions à des contenus devraient être définis par la loi.

119. Toute injonction imposant des restrictions devrait clairement indiquer sa base juridique et fournir des motifs pertinents et suffisants pour évaluer sa conformité avec la loi et, le cas échéant, expliquer également comme l'impact de la mesure sur la liberté d'expression du producteur ou de l'émetteur du contenu a été pris en considération conformément aux exigences de l'article 10 de la Convention. Un recours judiciaire accessible, suffisamment rapide et susceptible d'apporter un redressement devrait être mis à la disposition de l'intermédiaire destinataire de l'injonction, ainsi que de tout utilisateur dont la liberté d'expression, y compris le droit de recevoir des informations, est affectée par la mesure.

Concernant le paragraphe 50

120. Les plateformes peuvent imposer des restrictions sur les contenus légaux générés par les utilisateurs par le biais de leurs propres politiques et règles contractuelles. Dans certains cas, ces restrictions liées au contenu peuvent être introduites à titre de mesures visant à prévenir ou à atténuer les risques, sur la base de l'obligation d'effectuer des évaluations des risques, comme indiqué au paragraphe 60 de l'annexe. Toutefois, il doit être clair que de telles restrictions sur les contenus légaux découlent exclusivement des politiques et règles propres aux plateformes et ne sont pas imposées par la loi.

121. Les règles et politiques contractuelles des plateformes devraient être élaborées en dialogue avec les utilisateurs et les communautés d'utilisateurs, et avec leur contribution significative. Ceci est essentiel pour garantir qu'elles offrent une protection efficace contre les préjudices auxquels les utilisateurs sont confrontés et qu'elles ne vont pas au-delà de ce qui est nécessaire pour restreindre leurs droits et libertés. Il s'agit notamment de prendre en compte la manière dont des risques spécifiques se manifestent, ainsi que la manière dont la conception et les fonctionnalités propres à la plateforme, y compris l'utilisation de systèmes automatisés, peuvent influencer l'exposition des utilisateurs aux risques de préjudice. Les lignes directrices de l'UNESCO pour la gouvernance des plateformes numériques identifient la diligence raisonnable en matière de droits humains comme l'un de leurs principes fondamentaux, imposant « un engagement significatif avec diverses parties prenantes afin d'identifier les risques spécifiques pour les groupes en situation de vulnérabilité et de marginalisation »[34]. Les lignes directrices soulignent que les plateformes numériques devraient

34. UNESCO, *Principes pour la gouvernance des plateformes numériques : préserver la liberté d'expression et l'accès à l'information - une approche multipartite*, 2023, disponible à l'adresse https://unesdoc.unesco.org/ark:/48223/pf0000387359, par. 90.

également être ouvertes aux contributions d'experts et d'indépendants sur la manière dont ces évaluations sont structurées.

122. Les États devraient veiller à ce que les plateformes examinent et évaluent régulièrement et systématiquement l'impact de leurs politiques, systèmes et pratiques, y compris la manière dont ces dernières traitent les risques nouveaux et émergents. Ces évaluations devraient également être menées avant l'introduction de modifications de conception ou de nouvelles politiques. Elles devraient garantir la participation des parties prenantes concernées, en particulier de celles qui sont les plus touchées par les risques. En outre, elles devraient être transparentes et facilement accessibles, conformément au paragraphe 88 de l'annexe.

Concernant le paragraphe 51

123. Ce paragraphe traite directement des risques découlant de l'inaction des plateformes à l'égard des utilisateurs et des créateurs de contenu qui abusent clairement de leurs droits de publier des informations et de soumettre des notifications. Il couvre deux cas distincts, concernant les utilisateurs et les créateurs de contenu qui : premièrement, agissent en violation flagrante de la loi, en diffusant systématiquement des contenus illégaux ou soumis à des restrictions légales ; et, deuxièmement, abusent systématiquement de leur droit de soumettre des notifications, en particulier lorsqu'il est possible d'observer qu'ils le font dans le but de harceler ou de réduire au silence d'autres utilisateurs.

124. Les actions recommandées comprennent des mesures restrictives telles que la rétrogradation et la démonétisation de tous leurs contenus et, dans les cas les plus graves, la suspension et la résiliation de leurs comptes actuels et futurs. De telles mesures ont toutefois des conséquences très graves sur le droit à la liberté d'expression des utilisateurs visés et sur le droit des autres utilisateurs de recevoir les informations, et comportent un risque de censure par ricochet. Il convient donc de faire preuve de la plus grande prudence dans leur application. Elles devraient être réservées aux cas d'action systématique où l'intention de diffuser des contenus légalement restreints est évidente, lorsque d'autres mesures moins restrictives n'ont pas réussi à encadrer le comportement de l'utilisateur.

125. Les garanties procédurales contre une application arbitraire, telles que détaillées aux paragraphes 90 à 93 de l'annexe, s'appliquent : elles comprennent le droit des utilisateurs concernés d'être dûment informés et d'avoir accès à des recours extrajudiciaires équitables et efficaces.

Concernant le paragraphe 52

126. Comme l'a réaffirmé la Cour dans l'affaire *Google LLC et autres c. Russie* (précité, par. 90), l'article 10 de la Convention protège également le droit de ne pas être contraint de s'exprimer (*Gillberg c. Suède* [GC], n° 41723/06, 3 avril 2012, par. 85-86) et le droit de garder le silence (*Kobaliya et autres c. Russie*, nos 39446/16 et 106 autres, 22 octobre 2024, disponible seulement en anglais, par. 84). Par conséquent, toute disposition légale ou mesure administrative ayant pour effet de contraindre les plateformes à héberger des contenus spécifiques constitue une ingérence dans leurs droits au titre de l'article 10 de la Convention et doit être justifiée conformément à son paragraphe 2 : la mesure doit être prévue par la loi, s'appuyer sur un cadre juridique clair, accessible et prévisible ; elle doit poursuivre un but légitime ; elle doit être nécessaire dans une société démocratique, ce qui nécessite de démontrer un besoin social impérieux, étayé par des raisons pertinentes et suffisantes, et une réponse proportionnée à l'objectif poursuivi.

127. La notion de « besoin social impérieux » exige que la mesure réponde à un intérêt public véritable et important. Un tel besoin doit être étayé par des raisons pertinentes et suffisantes, y compris des preuves claires et fondées qui démontrent que la mesure est nécessaire pour servir cet intérêt public. On peut citer comme exemples la fourniture d'informations lors d'une urgence publique ou l'imposition de contenus dont l'indisponibilité réduirait considérablement le pluralisme des médias et compromettrait la diversité des points de vue, essentielle au discours démocratique. Toutefois, ces questions devraient être traitées en premier lieu par des mesures moins intrusives, en particulier des obligations et des responsabilités systémiques en matière d'atténuation des risques et d'autonomisation des utilisateurs imposées aux plateformes, plutôt que par l'imposition d'obligations *ad hoc* de publier des contenus ou des informations spécifiques.

Concernant le paragraphe 53

128. La Recommandation CM/Rec(2011)7 sur une nouvelle conception des médias recommande aux États d'adopter une notion large des médias, englobant « tous ceux qui participent à la production et à la diffusion, à un public potentiellement vaste, de contenus … tout en conservant … la surveillance ou le contrôle éditorial de ces contenus ». La Recommandation CM/Rec(2022)11 sur les principes de gouvernance des médias

et de la communication réaffirme que « la gouvernance des médias et de la communication devrait viser à garantir que les médias, les journalistes et autres acteurs respectent les obligations relatives aux contenus, en accord avec l'article 10 de la Convention et les normes professionnelles » (principe 10).

129. Les nouveaux acteurs des médias, tels que les blogueurs, les vlogueurs, les influenceurs, les podcasteurs, les journalistes citoyens et autres voix numériques indépendantes, ont un impact croissant sur les contenus en ligne et le flux d'informations. Ils opèrent généralement de manière indépendante, conservant le contrôle éditorial sur leur contenu, mais en dehors de la supervision éditoriale formelle propre aux médias traditionnels. Compte tenu du fait que le public se tourne de plus en plus vers les sources d'information et de divertissement en ligne, les créateurs de contenu opérant en dehors du cadre des médias traditionnels ont une influence croissante sur la manière dont les personnes reçoivent et accèdent à l'information, y compris aux actualités, et exercent un impact plus large sur la formation des valeurs et des opinions. Leur portée, leur visibilité et leur relation avec leur public, fondée sur l'authenticité, leur confèrent une capacité considérable à influencer les attitudes démocratiques, les valeurs et les opinions politiques, ainsi que la santé, les attitudes personnelles et les décisions de carrière. Si ces acteurs peuvent avoir des effets positifs, ils peuvent également produire des effets préjudiciables en diffusant de la mésinformation et de la désinformation, du discours de haine ou des contenus discriminatoires. En particulier, lorsqu'ils opèrent sur des plateformes particulièrement populaires auprès des enfants, ils peuvent exercer une influence significative sur eux et leur bien-être, ce qui soulève des inquiétudes quant à des messages par ailleurs licites, tels que des représentations irréalistes, la promotion d'habitudes malsaines et les risques pour la santé mentale. Par conséquent, leur comportement et leurs pratiques ont un impact direct sur la qualité et la sécurité de l'environnement en ligne.

130. Comme indiqué précédemment, certaines catégories de créateurs de contenu peuvent être soumises à des obligations légales relatives aux contenus, notamment dans le cadre de la législation sur les médias audiovisuels et la protection des consommateurs. D'autres professionnels, lorsqu'ils agissent en tant que créateurs de contenu, peuvent également être soumis à des cadres éthiques et professionnels d'autorégulation existants. Cela peut être le cas des avocats, des médecins ou des journalistes professionnels, mais aussi d'autres catégories telles que les chercheurs, les enseignants et, plus généralement, les fonctionnaires (voir l'exposé des motifs concernant les paragraphes 30 à 32). Dans ce cadre, le paragraphe 53 adresse deux recommandations aux États.

131. Premièrement, les États devraient veiller à ce que tous les créateurs de contenu professionnels, tels que définis ci-dessus (voir l'exposé des motifs concernant le paragraphe 11, deuxième tiret), soient tenus de divulguer des informations sur la manière dont ils monétisent leur contenu. Cette notion doit être comprise au sens large et ne pas se limiter à la monétisation par le biais des vues, des adhésions et des abonnements, des parrainages et des partenariats ou de la publicité et du placement de produits, mais englober toutes les stratégies et pratiques commerciales qui visent à tirer un profit économique du contenu, telles que la redirection vers d'autres plateformes ou sites web pour y vendre des produits et des services.

132. Deuxièmement, les États sont appelés à promouvoir l'élaboration de cadres d'autorégulation, en particulier pour les catégories de créateurs de contenu qui ne relèvent pas du champ d'application de la réglementation et de l'autorégulation existantes. On peut citer à titre d'exemple les créateurs qui, bien qu'ils ne soient pas enregistrés en tant que fournisseurs de services de médias ou journalistes professionnels conformément à la législation nationale, agissent effectivement en tant que tels. Des cadres d'autorégulation bien conçus peuvent profiter à la fois aux créateurs de contenu et à leur public, contribuant ainsi à la création d'un paysage médiatique en ligne plus sûr et plus fiable.

133. Les cadres d'autorégulation peuvent jouer un rôle clé en aidant les créateurs de contenu à respecter les normes éthiques et professionnelles, notamment celles visant à protéger les enfants, à améliorer la fiabilité des informations et à renforcer la confiance du public. Ces cadres doivent être transparents, avec des règles et des procédures accessibles au public, inclure des voix et des types de contenu diversifiés, et être fondés sur les droits humains. Plutôt que de se concentrer sur la restriction de l'expression licite, ces mécanismes doivent mettre l'accent sur la responsabilité, l'équité et la prévention des préjudices. À cette fin, ils pourraient prévoir des principes clairs sur des questions telles que la transparence des parrainages, l'amplification responsable des contenus, l'engagement respectueux et la correction des erreurs. Les créateurs de contenu devraient garantir que le contenu qu'ils fournissent respecte les obligations pertinentes en matière de contenu afin de protéger les groupes vulnérables, en particulier les enfants, contre tout préjudice. Les initiatives d'autorégulation ou les procédures de conformité internes devraient prévoir la mise en place d'un système de classification par âge, la classification indépendante du contenu avant sa diffusion et le traitement des plaintes

(voir l'Exposé des motifs relatif à la CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication, par. 10.5).

134. Des orientations utiles à cet égard figurent dans les Conclusions du Conseil de l'Union européenne sur le soutien aux influenceurs en tant que créateurs de contenu en ligne. Ce document propose diverses mesures, notamment le renforcement des compétences en matière d'éducation aux médias chez les créateurs de contenu, la facilitation du dialogue politique avec leurs organisations représentatives et les agences d'influence, le soutien au développement d'organismes ou de mécanismes d'autorégulation tels que des codes éthiques, et la participation des créateurs de contenu à l'élaboration des mesures de politique des média susceptibles de les concerner.

135. Les cadres d'autorégulation devraient être encouragés, sans préjudice de la possibilité pour les États de viser les créateurs de contenu dans la réglementation légale des médias : si des mécanismes d'autorégulation indépendants font défaut ou sont inefficaces, ou si l'intérêt public exige une implication plus forte de l'État en tant que garant de certains intérêts, les États ne devraient pas être empêchés d'adopter des cadres de corégulation appropriés et proportionnés (Exposé des motifs relatif à la CM/Rec(2022)11, par. 10.5).

Règles relatives à la responsabilité des intermédiaires

Concernant le paragraphe 54

136. Dans l'affaire *Magyar Tartalomszolgáltatók Egyesülete et Index.hu Zrt c. Hongrie*, précité, par. 86), la Cour européenne des droits de l'homme a mis en garde contre une extension excessive de la responsabilité des intermédiaires, qui pourrait avoir un effet dissuasif sur la liberté d'expression. Plus récemment, dans l'affaire *Google LLC et autres c. Russie*, elle a observé que les lourdes sanctions imposées à une plateforme pour non-respect d'injonctions de retrait formulées en termes généraux « imposaient une charge excessive aux intermédiaires ..., les contraignant de fait à agir en tant que censeurs du discours politique pour le compte des autorités publiques, une approche incompatible avec l'approche de la Cour en matière de liberté d'expression » (précité, par. 79). S'inspirant de ces préoccupations et de cette formulation, le paragraphe 54 met en garde contre le risque que les intermédiaires, s'ils sont tenus pour responsables de manière disproportionnée du contenu généré par les utilisateurs, anticipent leur responsabilité éventuelle en supprimant tout contenu dès que le moindre doute sur sa légalité surgit. Cette pratique, souvent qualifiée de « blocage excessif » ou de « censure par ricochet », ne résulte pas du fait que le contenu soit restreint par la loi, mais plutôt de la volonté des intermédiaires de minimiser les risques juridiques, d'éviter les amendes ou de préserver leur réputation. Afin d'éviter le blocage excessif, les États devraient veiller à ce que les cadres de responsabilité soient proportionnés, clairement définis et ciblés, et à ce que les intermédiaires d'internet puissent opérer dans un régime offrant une sécurité juridique suffisante.

Concernant le paragraphe 55

137. Le paragraphe 55 réitère le principe énoncé au paragraphe 1.3.7 de la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet[35]. Ces acteurs facilitent la transmission, l'accès ou le stockage de contenus pour le compte des utilisateurs. En reconnaissance de ce rôle, ils ne sauraient, en règle générale, être tenus responsables des contenus tiers qu'ils ne créent pas et se contentent d'héberger, de transmettre ou de rendre accessibles. Les autorités publiques ne peuvent par conséquent tenir les intermédiaires pour responsables à l'égard d'éléments individuels de contenu généré par les utilisateurs, qu'ils stockent, que lorsque deux conditions sont remplies : (1) l'intermédiaire a une connaissance du caractère restreint par la loi d'un contenu spécifique ; (2) l'intermédiaire omet d'agir rapidement pour restreindre l'accès au contenu incriminé. Les États devraient veiller à ce que les conditions, y compris les éventuels délais lorsque cela s'avère approprié, pour le retrait des contenus illicites ou l'application d'autres restrictions soient fixées par la loi.

138. Les intermédiaires peuvent prendre connaissance du fait qu'un contenu est restreint par la loi soit au moyen de notifications suffisamment étayées émanant d'utilisateurs, de groupes d'utilisateurs professionnels ou d'autorités publiques, soit à la suite de leurs propres enquêtes ou dans le cadre de l'exécution de leurs obligations légales en matière de sécurité et de responsabilité en ligne. Les procédures fondées sur la notification doivent être transparentes, accessibles et efficaces. Les autorités nationales doivent veiller à ce que ces

35. Voir également : Règlement de l'Union européenne sur les services numériques, articles 4-7 ; Royaume-Uni, *Règlement de 2002 sur le commerce électronique (directive CE), n° 2013*, disponible à l'adresse www.legislation.gov.uk/uksi/2002/2013/contents, articles 17-19.

procédures ne soient pas conçues de manière à encourager le retrait de contenus licites (CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, par. 1.3.7). La législation devrait également préciser que les intermédiaires d'internet ne peuvent être tenus responsables sur la seule base d'une connaissance générale du fait que leur service est utilisé pour stocker ou diffuser des contenus soumis à des restrictions légales ou parce qu'ils entreprennent volontairement, de bonne foi et avec diligence, des actions visant à détecter et à lutter contre les contenus soumis à des restrictions légales qu'ils peuvent héberger.

139. L'action requise par l'intermédiaire après avoir pris connaissance de la situation devrait consister à évaluer rapidement et de bonne foi si le contenu est restreint par la loi et, dans ce cas, à en restreindre l'accès ou à le supprimer, selon les cas. Les cadres juridiques pertinents devraient également préciser que les intermédiaires d'internet ne seront pas tenus responsables s'ils choisissent de ne pas supprimer un contenu sur la base d'une évaluation de bonne foi, fondée sur des informations de fait et juridiquement solides, même si ce contenu est ultérieurement qualifié par les autorités compétentes comme étant contraire au droit pénal, civil ou administratif (voir également l'Exposé des motifs relatif à la Recommandation CM/Rec(2022)16 sur la lutte contre le discours de haine). Lorsque la loi fixe des délais spécifiques pour agir, ceux-ci devraient toujours se rapporter à une catégorie spécifique de contenu et tenir compte de sa nature et de sa gravité. Les contenus dont la diffusion présente le risque de préjudice le plus imminent, tels que le matériel d'abus sexuel sur enfants ou l'incitation directe à la violence, devraient être soumis à des conditions plus strictes, notamment à des délais de retrait plus courts. Des délais d'action trop courts pour les notifications concernant des contenus relevant de catégories définies de manière large qui ne présentent pas de risque de préjudice imminent peuvent inciter à un retrait excessif, y compris de contenus licites.

140. Dans le cadre de l'évaluation, au titre de l'article 10 de la Convention, de la responsabilité d'un opérateur de portail internet pour ne pas avoir supprimé des commentaires publiés par un tiers, la Cour européenne des droits de l'homme a identifié quatre critères visant à établir un juste équilibre entre le droit à la liberté d'expression et les droits de la personne ou de l'entité lésée, à savoir : le contexte et le contenu des commentaires ; la possibilité de tenir les auteurs des commentaires pour responsables ; les mesures prises par l'intermédiaire pour prévenir les contenus illicites et le comportement de la partie lésée ; les conséquences pour la partie lésée et pour l'intermédiaire (*Magyar Tartalomszolgáltatók Egyesülete et Index.hu Zrt c. Hongrie,* précité, par. 60 et suivants ; *Delfi AS c. Estonie* [GC], précité, par. 142 et suivants).

141. Sur la base de ces critères, la Cour n'a pas exclu la possibilité, dans des cas exceptionnels et spécifiques, que certains intermédiaires d'internet puissent être tenus responsables s'ils « ne prennent pas de mesures pour retirer les commentaires clairement illicites sans délai après leur publication, et ce même en l'absence de notification par la victime alléguée ou par des tiers » (*Magyar Tartalomszolgáltatók Egyesülete et Index.hu Zrt c. Hongrie*, précité, par. 91). Dans l'affaire *Delfi AS c. Estonie* [GC] (précité), la Cour a estimé que l'article 10 de la Convention ne s'opposait pas à ce qu'un grand portail d'information internet, géré de manière professionnelle et exploité à des fins commerciales, et qui publie ses propres articles d'actualité et invite ses lecteurs à les commenter, soit condamné à verser des dommages-intérêts pour des commentaires anonymes extrêmes, qualifiés de discours de haine illégal ou d'incitation à la violence, publiés sous un article sur son site web. La Cour a également souligné que l'affaire ne concernait pas « d'autres types de forums sur Internet susceptibles de publier des commentaires provenant d'internautes, par exemple les forums de discussion ou les sites de diffusion électronique, où les internautes peuvent exposer librement leurs idées sur n'importe quel sujet sans que la discussion ne soit canalisée par des interventions du responsable du forum, ou encore les plateformes de médias sociaux où le fournisseur de la plateforme ne produit aucun contenu et où le fournisseur de contenu peut être un particulier administrant un site ou un blog dans le cadre de ses loisirs » (par. 116). Dans des affaires ultérieures, après avoir pris en considération, entre autres, l'absence de contenu illégal d'une gravité extrême, tel que des propos haineux illégaux ou des menaces directes à l'intégrité physique, dans les commentaires des utilisateurs en question, la Cour a estimé que la responsabilité des portails internet pour les commentaires de tiers sans connaissance effective de leur illégalité n'était pas compatible avec l'article 10 de la Convention (*Magyar Tartalomszolgáltatók Egyesülete et Index.hu Zrt c. Hongrie*, précité, par. 91 ; *Pihl c. Suède* (déc.), précité, par. 32 ; *Tamiz c. Royaume-Uni* (déc.), précité, par. 84 ; *Høiness c. Norvège*, n° 43624/14, 19 mars 2019, disponible seulement en anglais, par. 73-74).

142. La jurisprudence de la Cour n'implique ni que l'imposition d'une responsabilité sur cette base soit nécessaire pour trouver un juste équilibre entre les différents droits en jeu, ni que son raisonnement permettant de ne pas accorder l'exemption de responsabilité dans ces affaires s'applique à d'intermédiaires de nature différente. Elle indique simplement que, dans certains cas exceptionnels, le fait d'imposer une responsabilité à un type spécifique d'intermédiaire qui n'a pas connaissance effective du caractère illicite de certains contenus générés par les utilisateurs peut ne pas violer les droits de cet intermédiaire au titre de l'article 10 de la Convention. Le manque de diligence de la part de l'intermédiaire dans l'adoption de mesures raisonnables

et appropriées de modération de contenu est l'une des conditions nécessaires à une telle conclusion. Par conséquent, lorsque des cadres juridiques sont adoptés afin d'assurer un contrôle public de la diligence des plateformes en ce domaine, conformément à la présente recommandation, l'évaluation quant à la nature adéquate des mesures préventives adoptées devrait avoir lieu dans le cadre de ces mécanismes de responsabilisation des plateformes.

Concernant le paragraphe 56

143. Des obligations en matière de garanties procédurales doivent être mises en place pour protéger contre les demandes ou mesures arbitraires ou disproportionnées des autorités publiques. Comme l'indique la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, la législation devrait définir clairement les pouvoirs accordés aux autorités publiques, en particulier lorsqu'ils sont exercés par les autorités chargées de l'application de la loi (par. 1.2.2). Les obligations en matière de transparence constituent une autre garantie essentielle contre les retraits excessifs. Cela peut inclure la nécessité pour les autorités publiques d'indiquer clairement la base juridique de leur action et fournir des motifs étayés, lorsqu'elles adressent des demandes relatives au contenu à des intermédiaires. En outre, les États devraient publier des informations sur le nombre, la nature et la base légale des demandes de restriction de contenu soumises aux intermédiaires (ibidem, par. 1.2.3).

144. Dans ce contexte, une garantie essentielle consiste à veiller à ce que les utilisateurs disposent d'informations suffisantes pour contester les mesures prises par les intermédiaires en réponse à des injonctions de suppression, de blocage ou autres injonctions relatives aux contenus. Premièrement, les intermédiaires devraient être tenus de fournir des informations claires, facilement accessibles et suffisantes chaque fois qu'une décision de restreindre un contenu est prise, qu'elle soit fondée sur une décision judiciaire ou administrative, une notification soumise par un utilisateur individuel ou un groupe d'utilisateurs, ou bien l'application des conditions et politiques propres à la plateforme. Les explications devraient être communiquées directement aux utilisateurs concernés, dans un langage concis et simple qu'ils peuvent comprendre, et devraient indiquer la base juridique sur laquelle la décision est fondée. Deuxièmement, les utilisateurs concernés devraient être clairement informés de toute les voies de recours disponibles, notamment du système interne de traitement des plaintes de la plateforme, des voies de recours extérieures, des procédures devant les autorités de réglementation indépendantes et de la possibilité de saisir directement un tribunal compétent (voir également l'exposé des motifs concernant les paragraphes 89 à 92). La communication de ces informations peut toutefois être retardée, comme le prévoit la loi, en particulier lorsque cela est nécessaire dans le cadre d'une procédure judiciaire en cours.

145. Les États devraient s'abstenir de créer des conditions qui peuvent encourager les intermédiaires d'internet à supprimer ou à bloquer davantage de contenus au-delà de ce qui est requis par la loi. Comme indiqué ci-dessus, les ambiguïtés dans les définitions des contenus soumis à des restrictions, en particulier lorsqu'elles sont associées à des limitations supplémentaires, telles que des délais de suppression injustement courts, ne permettent souvent pas une évaluation approfondie dans les cas complexes. Des dispositions devraient donc être prévues pour décourager les suppressions préventives et précipitées de contenus potentiellement licites.

146. Les intermédiaires d'internet devraient être autorisés à appliquer des mesures provisoires dans certains cas incertains et complexes nécessitant une enquête factuelle ou une analyse juridique détaillée, par exemple en les renvoyant pour une évaluation plus approfondie à des organismes indépendants, notamment des autorités de réglementation, ou des organismes de corégulation et d'autorégulation. Les mesures provisoires peuvent également inclure la dépriorisation ou la contextualisation du contenu jusqu'à ce qu'une décision finale soit prise par l'organisme de la plateforme approprié ou un autre mécanisme. La dépriorisation consiste à accorder une priorité moindre au contenu et à en limiter ainsi la diffusion, tandis que la contextualisation consiste à publier le contenu accompagné d'une note indiquant qu'il pourrait s'agir d'un contenu non conforme à la loi.

147. L'annexe réaffirme un principe important énoncé dans la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, à savoir que les États restent en dernier ressort responsables de la protection des droits humains. Ils ne peuvent transférer ou déléguer cette obligation à des entités privées, qu'il s'agisse d'intermédiaires d'internet ou de tiers auxquels ils ont délégué certaines responsabilités en matière de signalement de contenus ou de résolution de litiges. Tous les cadres réglementaires devraient inclure des mécanismes de contrôle efficaces et indépendants afin de garantir que toutes les tâches déléguées sont conformes aux normes en matière de droits humains.

148. Certaines questions peuvent être gérées efficacement grâce à des cadres de corégulation, en particulier lorsque le secteur adopte des codes dont le respect peut être utilisé pour démontrer la conformité aux obligations légales. Par exemple, les codes de conduite volontaires dans le cadre du Règlement de l'Union européenne sur les services numériques, constituent un ensemble d'engagements qui peuvent servir de mesure d'atténuation adaptée à des risques systémiques spécifiques, tels que ceux touchant au discours civique et à l'intégrité des élections, pour lesquels l'intégrité de l'information est essentielle, ou à des défis spécifiques liés à la lutte contre les contenus restreints par la loi, tels que le discours de haine illégal. Toutefois, tout mécanisme de corégulation doit s'appuyer sur un cadre juridique établi par l'État, qui en définit la portée, garantit la responsabilité et comprend des garanties contre les décisions arbitraires prises par des acteurs non étatiques. La Note d'orientation sur la modération de contenu énonce les principes essentiels à cet égard, que les États devraient utiliser comme guide pour concevoir et mettre en œuvre leurs approches de corégulation. Elle souligne également les écueils des approches d'autorégulation et de corégulation motivées par la pression gouvernementale, concluant que les interventions politiques ayant pour but de minimiser les risques « devraient également être assorties d'objectifs clairs, de mécanismes d'ajustement et de supervision, d'une protection significative de la liberté d'expression, ainsi que d'outils permettant d'identifier les effets contre-productifs » (p. 31).

Règles relatives à la responsabilisation des plateformes et à l'autonomisation des utilisateurs

Concernant les paragraphes 57 et 58

149. Un environnement en ligne favorable est un environnement qui favorise une participation significative des utilisateurs, protège les individus contre les préjudices et leur donne les moyens d'agir et de contrôler leurs expériences en ligne. Un tel environnement est essentiel à la réalisation des droits numériques, notamment la liberté d'expression, l'accès à l'information et la participation au débat public. Cependant, les problèmes structurels propres aux plateformes en ligne, tels que l'opacité des algorithmes, les caractéristiques de conception peu sûres et la modération incohérente des contenus, peuvent créer des conditions qui compromettent, plutôt que de les favoriser, l'exercice et la jouissance de ces droits, notamment pour les groupes marginalisés et vulnérables.

150. Compte tenu de l'influence croissante des plateformes en ligne sur la manière dont les personnes accèdent à l'information et exercent leurs droits, il est essentiel de veiller à ce que les utilisateurs soient en mesure de comprendre, d'atténuer et de réagir aux risques en ligne. Comme le souligne la Note d'orientation sur la lutte contre la désinformation et la mésinformation en ligne, l'intégrité de l'information en ligne nécessite une stratégie holistique. L'impact des contenus « de mauvaise qualité », tels que la désinformation, dépend en fin de compte du fait que les utilisateurs (a) soient régulièrement exposés à des contenus de haute qualité et en reconnaissent l'importance ; (b) soient capables de faire la distinction entre les contenus de haute qualité et ceux de mauvaise qualité ; (c) agissent de manière responsable envers les autres lorsqu'ils partagent et discutent des différents types d'informations qu'ils peuvent rencontrer ; et (d) bénéficient de protections solides pour leurs droits humains, sachent comment les exercer et sont convaincus qu'ils peuvent avoir un impact positif, tout en se protégeant eux-mêmes, en les exerçant. La législation devrait donc placer l'autonomisation des utilisateurs au cœur des cadres juridiques pour la sécurité en ligne et la responsabilisation des plateformes. Les mesures d'autonomisation devraient fournir aux utilisateurs des outils, des processus et des systèmes leur permettant de gérer leur expérience en ligne, de comprendre comment le contenu est présenté, organisé et recommandé par les plateformes, de prendre des décisions éclairées et de contester les résultats de la modération du contenu, ainsi que des mécanismes pour protéger leurs droits.

151. Les utilisateurs devraient être en mesure de participer activement à l'environnement en ligne. Toutefois, leur autonomisation ne devrait pas se faire au prix d'une charge excessive pesant sur eux pour la protection de leurs propres droits (CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression, par. 1.7). En outre, les obligations en matière d'autonomisation ne sauraient être considérées comme un substitut aux responsabilités plus larges qui incombent aux plateformes quant à la gestion des risques de préjudice en ligne (voir également CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication, principe 15). Certaines obligations clés en matière d'autonomisation devraient donc être imposées par la loi et soumises à un contrôle, sans être laissées à la discrétion des plateformes, y compris par le biais de mesures appliquées dans le cadre d'accords d'autorégulation.

152. Il existe un consensus mondial croissant sur le fait que la protection des droits en ligne nécessite des mesures proactives et structurelles, pas seulement une modération réactive des contenus. Comme déjà

exposé précédemment (voir l'exposé des motifs concernant le paragraphe 33), les cadres juridiques devraient porter sur les devoirs et les responsabilités systémiques qui incombent aux plateformes à l'égard de leurs propres systèmes et processus, plutôt que d'imposer une responsabilité au titre de l'hébergement de contenus spécifiques contraires à la loi générés par les utilisateurs, hors des cas visés au paragraphe 55 de l'annexe. Cela implique que la législation établisse des obligations générales concernant la manière dont les caractéristiques structurelles et opérationnelles des plateformes, telles que les modèles économiques, les conditions générales de service et leur application, les paramètres de conception, les systèmes algorithmiques et publicitaires, les processus d'organisation et de modération des contenus ou les pratiques en matière de transparence, contribuent à divers risques de préjudice en ligne. L'intégration de devoirs et de responsabilités systémiques dans le cadre de gouvernance des plateformes est conforme au principe de promotion d'un environnement en ligne favorable, car cette approche leur confère la responsabilité d'identifier, d'évaluer et d'atténuer de manière proactive ces risques, actuels et émergents, y compris par des mesures qui responsabilisent les utilisateurs et soutiennent leur capacité d'action.

Concernant le paragraphe 59

153. Les États devraient exiger que les plateformes en ligne intègrent la sécurité des utilisateurs au cœur même de l'architecture de leurs services dès la conception, plutôt que d'aborder les risques uniquement après la survenue de préjudices. Cela signifie qu'il faut veiller à ce que la prévention des préjudices soit intégrée dans la conception, le développement et le déploiement des fonctionnalités des plateformes, en alignant l'architecture du système sur la sécurité et les droits des utilisateurs. En outre, la conception devrait faire l'objet de mesures régulières d'atténuation des risques afin de prendre en compte ceux découlant des caractéristiques de conception qui facilitent l'amplification du contenu ou des comportements. L'amplification algorithmique, en particulier par le biais des systèmes de recommandation et de publicité, peut considérablement étendre la portée des contenus présentant un risque de préjudice, y compris les contenus soumis à des restrictions par la loi. De même, les caractéristiques de conception de l'interface utilisateur telles que la lecture automatique, les incitations comportementales ou les interfaces trompeuses peuvent manipuler les utilisateurs ou créer des environnements à risque. Les éléments fondamentaux d'une approche axée sur la sécurité des utilisateurs par défaut et dès la conception peuvent inclure des choix techniques visant à limiter l'amplification des contenus soumis à des restrictions légales par le biais des systèmes de recommandation, à définir par défaut des paramètres de confidentialité renforcés pour les mineurs, ou à désactiver le ciblage de contenus ou de publicités fondé sur le profilage, sauf consentement exprès de l'utilisateur. Par exemple, la Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet qui préconise, en ce qui concerne l'utilisation des données à caractère personnel, l'application des principes de « confidentialité par défaut » et de « confidentialité dès la conception » à tous les stades afin de prévenir ou de minimiser le risque d'atteinte aux droits et libertés fondamentaux des utilisateurs (par. 2.4.3). Par exemple, les sections 9 à 13 de la loi sur la sécurité en ligne du Royaume-Uni imposent des obligations de sécurité dès la conception aux plateformes réglementées, notamment des paramètres par défaut pour les enfants, des évaluations de la conception de fonctionnalités telles que les flux algorithmiques et des obligations en matière d'autonomisation des utilisateurs par le biais de réglages et de contrôles.

154. La conception des plateformes en ligne, y compris les interfaces utilisateur, l'architecture du contenu et les interactions entre les systèmes, joue un rôle central dans la manière dont les utilisateurs accèdent à l'information en ligne, interagissent avec celle-ci et y contribuent. La Recommandation CM/Rec(2022)13 sur les effets des technologies numériques sur la liberté d'expression préconise que l'infrastructure numérique de communication soit conçue de manière à promouvoir les droits humains. Conformément aux Principes de l'UNESCO pour la gouvernance des plateformes numériques, « les plateformes numériques doivent garantir la non-discrimination et l'égalité de traitement dans leurs processus de conception, ainsi que dans leurs politiques, pratiques et systèmes de modération et de conservation de contenus. Cela comprend la lutte contre les préjugés, les stéréotypes et les algorithmes discriminatoires ou les pratiques de modération de contenus qui affectent les femmes et les filles, ainsi que les groupes en situation de vulnérabilité et de marginalisation, y compris les communautés autochtones » (par. 93).

155. Les obligations en matière de conception devraient tenir compte de la nécessité de préserver la liberté d'expression et un environnement informationnel pluraliste dans lequel des voix et des perspectives diverses sont présentes et accessibles en ligne. Les services en ligne devraient être conçus de manière à minimiser les restrictions à l'expression légitime et la marginalisation des opinions minoritaires ou dissidentes. Ces risques peuvent provenir de la conception des systèmes algorithmiques utilisés pour recommander, classer, hiérarchiser ou sélectionner des contenus ; des éléments de l'interface utilisateur optimisés pour maximiser l'engagement des utilisateurs ; ou des paramètres par défaut qui limitent la visibilité ou la possibilité de découvrir

certains types de contenus ou de comptes d'utilisateurs. Pour atténuer ces risques, les plateformes devraient fournir des contrôles clairs, accessibles et granulaires qui permettent aux utilisateurs d'influencer la manière dont le contenu leur est proposé et recommandé, tels que des paramètres de recommandation ajustables, la transparence sur les raisons pour lesquelles un contenu spécifique est affiché, ou la possibilité de choisir de se désinscrire des recommandations basées sur des algorithmes qui déterminent soit la visibilité du contenu dans leurs flux ou résultats de recherche, soit la sélection de contenu basée sur le profilage des utilisateurs[36].

Concernant les paragraphes 60 à 63

156. En raison de leur taille, de leur portée et de leur rôle structurel dans la formation du discours public, les plateformes exerçant une influence significative devraient être soumises à des responsabilités accrues en ce qui concerne l'impact de leur conception et de leurs décisions opérationnelles. Elles devraient donc être tenues d'examiner l'impact de leurs services sur la liberté d'expression et d'autres droits humains, tels que le droit à la vie privée, le droit à la non-discrimination et les droits de l'enfant. En outre, les évaluations devraient tenir compte de tout impact sur les processus démocratiques et électoraux, ainsi que des risques liés à la diffusion de contenus soumis à des restrictions légales.

157. Les risques peuvent découler de la conception même des services et de leur fonctionnement. Ils peuvent également résulter des processus de modération de contenu des plateformes. Par exemple, certaines caractéristiques de conception, telles que les mécanismes d'enregistrement des utilisateurs, peuvent offrir une protection insuffisante contre la création de faux comptes utilisés à des fins de harcèlement ou de fraude. De même, le fonctionnement des plateformes peut présenter un risque de diffusion et d'amplification des contenus, à la fois par la priorisation algorithmique des contenus les plus susceptibles d'attirer l'attention du public et par le détournement des politiques publicitaires des plateformes qui permettent la monétisation de ces contenus.

158. Une prise en compte effective de la sécurité des utilisateurs dès la conception suppose que les plateformes concernées procèdent à des évaluations de risques rigoureuses tout au long de la mise en œuvre, de la maintenance et de la mise à jour des décisions de conception. Il peut s'agir d'évaluations de risques avant le lancement (par exemple, des tests d'abus) afin de minimiser les risques de préjudice avant que des décisions de conception soient mises en œuvre, ainsi que de mettre en place des mécanismes de supervision continue afin de garantir que tout risque éventuel puisse être efficacement atténué. À cette fin, les États peuvent introduire des intervalles réguliers auxquels les plateformes en ligne sont tenues de procéder à des évaluations de risques.

159. Les risques systémiques se manifestent différemment selon les groupes d'utilisateurs. Une gestion et une atténuation efficaces des risques nécessitent donc la prise en compte des contributions des personnes concernées, ainsi que des éléments contextuels reflétant leur situation particulière. Les États devraient donc veiller à ce que, lorsqu'elles procèdent à des évaluations des risques, les plateformes exerçant une influence significative mènent des consultations ouvertes, transparentes et efficaces (CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression, par. 3.5) avec les parties prenantes concernées, en particulier celles dont les droits et les expériences sont directement affectés par la conception et les pratiques de gouvernance des plateformes. Une attention particulière devrait être accordée à la participation des femmes, compte tenu de la prévalence des préjudices en ligne liés au genre et des risques spécifiques auxquels elles sont confrontées dans ces environnements. Les consultations avec les parties prenantes sont également cruciales pour évaluer, par exemple : si les politiques et les systèmes de modération de contenu des plateformes offrent une protection efficace contre les risques de préjudices, y compris les plus subtils et évolutifs ; s'ils prennent en compte de manière spécifique et adéquate les particularités liées au contexte, telles que les conséquences des particularités linguistiques, culturelles ou régionales ; s'ils ont des conséquences préjudiciables dans la vie réelle. Une consultation significative garantit que les plateformes intègrent les expériences vécues et les vulnérabilités des différents groupes d'utilisateurs et peuvent mieux identifier et atténuer les risques involontaires de préjudices découlant de leur conception et de leurs pratiques de gouvernance des contenus. Par exemple, les Lignes directrices sur la protection des mineurs adoptées en 2025 par la Commission européenne conformément à l'article 28.4 du Règlement de l'Union européenne sur les services numériques prévoient que la conception et le fonctionnement des systèmes de recommandation destinés aux mineurs, ainsi que tout élément de la plateforme qui concerne leur vie privée, leur sûreté et leur sécurité, doivent être testés avec des mineurs et que leurs commentaires soient pris en compte, notamment

36. UNESCO, *Towards user empowerment : a multistakeholder action plan for integrating media and information literacy on digital platforms*, 2025, disponible en anglais seulement à l'adresse https://unesdoc.unesco.org/ark :/48223/pf0000394855.locale=en, action 27.

en consultant des mineurs d'âges différents, issus de milieux culturels et linguistiques divers et présentant des handicaps[37].

160. Le processus de consultation est non seulement essentiel pour garantir une protection efficace contre les préjudices, mais aussi pour éviter la mise en œuvre de mesures d'atténuation qui vont au-delà de ce qui est nécessaire et proportionné dans une société démocratique, risquant de restreindre l'expression légitime ou le pluralisme. Les plateformes concernées devraient être tenues de documenter la manière dont les contributions des parties prenantes ont été prises en compte et de démontrer que les résultats de ces consultations ont été dûment pris en considération dans l'élaboration de leurs mesures d'atténuation des risques, conformément à la Recommandation CM/Rec(2022)13 sur les effets des technologies numériques sur la liberté d'expression, paragraphe 3.5, qui stipule que les intermédiaires d'internet doivent fournir des informations complètes sur le processus, le contenu et les résultats des consultations, en divulguant tous les commentaires qu'ils reçoivent et en expliquant si et de quelle façon ils en tiennent compte.

161. La publication par les plateformes de leurs évaluations des risques et des incidences sur les droits humains est de plus en plus reconnue comme un élément nécessaire à une gouvernance transparente et responsable. Le règlement de l'Union européenne sur les services numériques, par exemple, exige que les très grandes plateformes en ligne et moteurs de recherche publient les évaluations des risques systémiques qu'ils réalisent, ainsi que les mesures prises pour atténuer ces risques (article 42.5). La documentation publique des évaluations des risques et des incidences sur les droits humains est un outil important pour l'autonomisation des utilisateurs, car elle leur permet de mieux comprendre les risques associés aux services des plateformes. Elle permet également à la société civile et aux chercheurs d'examiner et d'évaluer les pratiques des plateformes, aux régulateurs d'évaluer la conformité aux exigences légales et au grand public de participer à un débat éclairé sur le rôle sociétal et la gouvernance des plateformes exerçant une influence significative.

162. Le paragraphe 63 traite de l'obligation positive des États de favoriser un environnement favorable à une participation et à un contrôle efficace. Outre la garantie juridique de la transparence, des consultations et de l'accès à l'information sur les évaluations des risques et les mesures d'atténuation, les États pourraient également créer, faciliter et soutenir des forums consultatifs ou de supervision multipartite. Ce soutien pourrait se concrétiser entre autres par la mise à disposition de ressources et le renforcement des capacités afin de garantir la participation de la société civile, des chercheurs et des experts indépendants à l'évaluation de la conformité et à l'identification des domaines à risque prioritaires sur l'ensemble des plateformes. Ces organismes devraient être en mesure d'exercer une véritable influence, par exemple en émettant des recommandations, en lançant des enquêtes ou en demandant des données supplémentaires aux plateformes.

163. Les plateformes peuvent expurger ou ne pas divulguer certaines informations spécifiques concernant leurs évaluations des risques et de l'impact sur les droits humains lorsque la divulgation publique complète de certaines données internes ou procédures opérationnelles de la plateforme peut créer des failles de sécurité ou être exploitée par des acteurs malveillants, entraînant un risque accru de préjudice pour les utilisateurs, par exemple lorsqu'il est raisonnablement prévisible que ces informations pourraient être exploitées par des acteurs malveillants pour coordonner des actes de harcèlement ou d'abus. Toutefois, ces expurgations devraient être strictement limitées à ce qui est nécessaire pour prévenir ces effets négatifs et devraient être soumises à un contrôle réglementaire afin d'éviter tout abus. Les organismes de réglementation devraient avoir accès à l'intégralité des informations non expurgées afin d'évaluer la légitimité des expurgations et de s'assurer que les exceptions ne sont pas appliquées de manière arbitraire ou excessive.

Concernant le paragraphe 64

164. Une supervision et une application efficaces des cadres législatifs relatifs à la responsabilité des plateformes et à l'autonomisation des utilisateurs exigent que les autorités réglementaires fonctionnent avec un haut niveau d'indépendance, d'intégrité et de capacité institutionnelle. Les États devraient veiller, tant par la législation que dans la pratique, à ce que ces autorités puissent exercer leurs fonctions sans subir d'influence politique, commerciale ou de la part des plateformes, qu'elle soit directe ou indirecte. Par exemple, les autorités devraient utiliser des infrastructures techniques et des capacités d'analyse de données indépendantes, évitant une dépendance excessive vis-à-vis des plateformes pour le contrôle de conformité et la collecte d'éléments de preuve. Cela peut impliquer de développer une expertise technique interne ou de faire appel à des partenaires externes spécialisés, tels que des auditeurs et des chercheurs indépendants

37. Commission européenne, *Lignes directrices concernant des mesures visant à garantir un niveau élevé de protection de la vie privée, de sûreté et de sécurité des mineurs en ligne, conformément à l'article 28, paragraphe 4, du règlement (UE) 2022/2065*, communication C/2025/5519, 7 octobre 2025, disponible à l'adresse http://data.europa.eu/eli/C/2025/5519/oj, ci-après « Lignes directrices sur la protection des mineurs », parr. 65 et 89.

(voir les paragraphes 86-87 de l'annexe). Toutefois, afin d'exercer efficacement leurs pouvoirs de contrôle et de mise en œuvre, les autorités de régulation devraient avoir le pouvoir de demander aux plateformes toutes les informations nécessaires à l'exercice de ces fonctions. La législation sur la responsabilité des plateformes devrait donc prévoir l'obligation légale pour les plateformes de fournir ces informations à des fins de contrôle réglementaire, ainsi que des sanctions proportionnées en cas de manquement ou de refus de donner suite à une telle demande des autorités de régulation.

165. Il est primordial, pour l'indépendance des autorités de régulation, qu'il leur soit accordé une indépendance et une autonomie au niveau budgétaire, afin de garantir qu'elles disposent des ressources financières, techniques et humaines adéquates pour s'acquitter efficacement de leurs fonctions. La Recommandation Rec(2000)23 concernant l'indépendance et les fonctions des autorités de régulation du secteur de la radiodiffusion offre des orientations précieuses à cet égard, qui devraient être adaptées et appliquées de manière appropriée au contexte de la réglementation des plateformes numériques. En outre, la Recommandation CM/Rec(2022)11 sur les principes de gouvernance des médias et de la communication définit des normes de gouvernance clés qui sont directement pertinentes, en particulier le principe 3, qui souligne l'importance de l'indépendance et de l'impartialité des organismes de régulation et de gouvernance, et le principe 4, qui met en évidence la nécessité de choix de gouvernance fondés sur des éléments factuels et axés sur les résultats. Ces normes renforcent l'obligation des États de mettre en place et de soutenir des autorités de régulation fonctionnellement autonomes, dotées de ressources suffisantes et habilitées à agir sur la base de leur expertise et d'éléments factuels. La réglementation des plateformes numériques, en particulier, nécessite un large éventail de compétences spécialisées, ainsi qu'une compréhension approfondie des principes de la réglementation des plateformes. Cela inclut la capacité d'évaluer comment les risques se manifestent différemment selon les plateformes, les services et les groupes d'utilisateurs. Pour relever les défis complexes et en constante évolution de l'environnement numérique, une expertise multidisciplinaire est essentielle.

Concernant le paragraphe 65

166. La promotion d'un environnement favorable à la sécurité en ligne dans une société démocratique est un effort collectif, dans lequel la société civile et les autres acteurs non étatiques agissant dans l'intérêt public jouent un rôle fondamental. Si l'annexe met l'accent sur les responsabilités respectives des États et des plateformes, elle confie également des rôles et des tâches spécifiques à ces acteurs, comme le précisent les sections de la partie V consacrées à la transparence, aux droits procéduraux et à l'action collective des utilisateurs. Le paragraphe 65 reconnaît la nécessité de veiller à ce que ces acteurs exercent leurs fonctions dans l'intérêt public, restent indépendants tant des autorités étatiques que des plateformes, et opèrent dans le cadre d'un régime de transparence et de responsabilité, avec des garanties appropriées contre tout abus potentiel de leur position. À cette fin, ils peuvent être soumis à des processus de certification et à des obligations d'établir des rapports. Afin de garantir que ces entités servent véritablement l'intérêt public et s'acquittent de leurs tâches déléguées avec précision et objectivité, les États devraient créer un environnement durable pour leur fonctionnement indépendant, notamment grâce à des sources de financement fiables leur permettant d'être rémunérées équitablement, ce qui garantirait leur efficacité et leur indépendance. Parmi les formes possibles de rémunération qui ne dépendent pas du budget de l'État, les États peuvent envisager l'introduction de taxes ou de redevances imposées aux plateformes.

V. Mesures d'autonomisation en ligne des utilisateurs

Dispositions générales

Concernant les paragraphes 66 et 67

167. Les paragraphes 66 et 67 reflètent une approche des risques en ligne axée sur l'utilisateur et son autodétermination, dans laquelle les utilisateurs se voient accorder une plus grande influence sur leur expérience en ligne. Cela garantit le respect de leur autonomie individuelle et renforce leur capacité à gérer leur propre exposition aux risques. Cette approche reconnaît que de nombreux préjudices rencontrés dans les environnements numériques dépendent du contexte et peuvent souvent être atténués, en tout ou en partie, en donnant aux utilisateurs les moyens de faire des choix éclairés concernant les contenus et les interactions auxquels ils sont exposés. Elle contribue également à renforcer la résilience des utilisateurs et leur permet de jouer un rôle plus actif et participatif dans le développement futur de la gouvernance numérique, faisant d'eux des participants actifs dans l'environnement numérique. À cette fin, les cadres juridiques devraient soutenir activement - et, le cas échéant, exiger - des mécanismes permettant une organisation et une modération

des contenus et des comportements en ligne centrées sur l'utilisateur. Reconnaissant que la plupart des plateformes en ligne fonctionnent déjà selon des modèles mixtes de modération de contenu, avec un certain niveau d'autonomie accordé aux utilisateurs, les outils d'autonomisation des utilisateurs devraient aller au-delà des approches binaires traditionnelles pour offrir également des options graduelles. Ces mécanismes peuvent porter sur la possibilité d'ajuster la visibilité ou la priorité des contenus en fonction des préférences individuelles, ou de limiter l'exposition à certains types de contenus grâce à des outils personnalisables. Par exemple, la Loi sur la sécurité en ligne du Royaume-Uni introduit des « obligations d'autonomisation des utilisateurs » qui obligent certains services à proposer aux adultes des outils permettant de filtrer ou de restreindre des contenus qui ne sont légalement restreints qu'aux enfants, tels que les abus ou la glorification de l'automutilation, sans pour autant supprimer complètement ces contenus. Les utilisateurs devraient également avoir la possibilité de faire appel à des tiers de confiance pour les aider à gérer ces paramètres ou les gérer à leur place.

168. L'autonomisation des utilisateurs ne peut se substituer à une approche systémique et fondée sur les droits humains en matière de gouvernance des contenus et d'atténuation des risques. Elle doit plutôt être conçue et mise en œuvre dans un cadre plus large où les plateformes demeurent pleinement responsabilisées en matière de sécurité de l'environnement en ligne. La redistribution des pouvoirs d'action doit donc être considérée comme une responsabilité partagée, qui nécessite un engagement soutenu de la part de tous les acteurs afin d'éviter de renforcer les déséquilibres de pouvoir ou de faire peser la charge de la sécurité uniquement sur les utilisateurs (voir l'exposé des motifs concernant les paragraphes 57 et 58).

Concernant les paragraphes 68-69

169. Les paragraphes 68 et 69 réaffirment le principe de responsabilité proportionnée et graduée des plateformes (voir l'exposé des motifs concernant le paragraphe 41), en mettant particulièrement l'accent sur les obligations en matière d'autonomisation des utilisateurs.

170. Les obligations imposées aux plateformes en vertu des cadres juridiques applicables devraient représenter des exigences minimales, leur laissant une marge de manœuvre importante pour la mise en œuvre volontaire de mesures supplémentaires. L'annexe soutient le principe selon lequel les plateformes devraient être encouragées à adopter des mesures d'autonomisation même lorsqu'elles n'y sont pas strictement tenues par la loi, en reconnaissant que les actions volontaires peuvent jouer un rôle complémentaire aux cadres réglementaires et coréglementaires dans la promotion d'un environnement en ligne favorable.

171. Dans ce contexte, le rôle des États pourrait consister à engager un dialogue régulier avec les plateformes afin de faciliter la mise en place de forums collaboratifs multipartites entre les différentes parties prenantes, notamment les plateformes, les groupes d'utilisateurs, les acteurs de la société civile, la communauté universitaire, etc. En outre, les États peuvent jouer un rôle important en soutenant des initiatives d'alphabétisation numérique et d'éducation aux médias et a l'information impliquant les plateformes, ce qui peut encourager des mesures proactives visant à améliorer les compétences numériques des utilisateurs, y compris leur compréhension de la manière dont les plateformes présentent et recommandent les contenus (voir CM/Rec(2022)4 sur la promotion d'un environnement favorable à un journalisme de qualité à l'ère du numérique, par. 1.4.5 ; Principes de l'UNESCO pour la gouvernance des plateformes numériques, par. 80-84).

Autonomisation par la conception

Concernant le paragraphe 70

172. Le paragraphe 70 énonce un principe très général : les choix de conception des plateformes devraient donner aux utilisateurs le contrôle maximal possible sur leur expérience en ligne. Cela implique que les outils soient conçus et mis en œuvre de manière à faciliter leur utilisation. La maximisation de l'autonomisation des utilisateurs ne dispense toutefois pas les plateformes de leurs obligations, le cas échéant, d'évaluer et d'atténuer les risques que leurs systèmes font peser sur les droits humains.

Concernant le paragraphe 71

173. Les systèmes de recommandation, c'est-à-dire les outils algorithmiques utilisés pour hiérarchiser, filtrer et suggérer des contenus individuels aux utilisateurs, font partie des principaux domaines où les choix de conception des plateformes peuvent donner lieu à des risques systémiques.

174. En façonnant la manière dont les individus accèdent à l'information et interagissent avec le contenu, les systèmes de recommandation constituent un élément central de la conception des plateformes en ligne. Les

mécanismes de classement et de recommandation optimisés pour l'engagement donnent souvent la priorité aux contenus dont le potentiel d'interaction est élevé et les affichent aux utilisateurs les plus susceptibles d'interagir avec eux. Cela peut contribuer à l'amplification et à la diffusion rapide de contenus à la limite de l'illicéité, notamment du discours de haine et de la désinformation. Les systèmes de recommandation optimisés principalement pour l'engagement peuvent nuire à la découvrabilité des contenus en limitant l'exposition des utilisateurs à une diversité de sources et de points de vue, compromettant ainsi le pluralisme des médias et l'exercice effectif du droit de recevoir des informations.

175. Les plateformes devraient être tenues de mettre à disposition une gamme d'outils qui donnent aux utilisateurs une possibilité véritable de choisir les types de contenus qu'ils souhaitent voir et de personnaliser les algorithmes de la plateforme en fonction de leurs préférences, de leurs valeurs et de leurs sensibilités. Entre autres, les utilisateurs devraient avoir la possibilité de définir activement leurs propres choix, de désactiver les paramètres par défaut ou de refuser certaines fonctionnalités ou certains éléments de conception imposés par la plateforme. Par exemple, la Loi sur la sécurité en ligne du Royaume-Uni prescrit explicitement que des outils destinés aux utilisateurs soit prévues en tant qu'exigence obligatoire en matière de conception du service. Elle impose aux grands services utilisateur à utilisateur (notamment ceux désignés comme « services de catégorie 1 ») une obligation de responsabilisation des utilisateurs afin de permettre aux adultes de personnaliser leur expérience en choisissant de bloquer, de filtrer ou de recevoir des avertissements concernant les contenus qui encouragent le suicide ou l'automutilation, encouragent les troubles alimentaires, impliquent des abus ou de la haine fondés sur des caractéristiques protégées (section 16). L'article 38 du règlement de l'Union européenne sur les services numériques impose aux très grandes plateformes en ligne et aux moteurs de recherche qui utilisent des systèmes de recommandation l'obligation d'introduire au moins une option alternative qui ne soit pas basée sur le profilage. Il peut s'agir de fils d'actualité chronologiques, de recommandations des contenus les plus populaires ou de contenus tendance à un moment donné. Le système de gestion des risques ou la protection des mineurs dans le cadre du règlement pourrait imposer des mesures supplémentaires d'autonomisation concernant les systèmes de recommandation[38].

176. Les plateformes devraient fournir aux utilisateurs des outils leur permettant de protéger efficacement leur vie privée, notamment la possibilité de choisir parmi un ensemble d'options de paramètres de confidentialité. Comme indiqué dans la Recommandation CM/Rec(2020)1 sur les impacts des systèmes algorithmiques sur les droits de l'homme, les options par défaut ne devraient conduire qu'à la collecte des données nécessaires et proportionnées à la finalité légitime spécifique du traitement des données, tandis que les paramètres de suivi devraient être réglés par défaut en mode « *opt-out* » (par. 2.2).

Concernant le paragraphe 72

177. La Note d'orientation sur la lutte contre la mésinformation et la désinformation en ligne offre des recommandations précieuses sur la conception des plateformes et les solutions visant à autonomiser les utilisateurs, qui facilitent leur autodétermination et leur prise de décision éclairée concernant les contenus auxquels ils s'intéressent. Il peut s'agir, entre autres, de fournir des informations supplémentaires aux utilisateurs, des alertes liées à l'âge, des avertissements déclencheurs et des contenus supplémentaires provenant de sources officielles et indépendantes faisant autorité, telles que les organes de presse professionnels et les médias de service public[39].

178. Les étiquettes de contenu sont une caractéristique importante de la conception qui fournit des indications sur l'exactitude des informations et permet aux utilisateurs d'évaluer leur fiabilité et leur intégrité contextuelles. L'utilisation d'étiquettes ou de drapeaux placés par des vérificateurs de faits indépendants en est un excellent exemple. Les étiquettes de vérification des faits jouent un rôle important en permettant aux utilisateurs de prendre des décisions éclairées sur le contenu qu'ils consultent, avec lequel ils interagissent et qu'ils partagent. En fournissant un contexte précieux, la vérification des faits réduit la crédibilité perçue de la désinformation, ce qui rend les utilisateurs moins enclins à partager ce type de contenu. La Note d'orientation sur la lutte contre la mésinformation et la désinformation en ligne souligne que « les États et les plateformes devraient reconnaître la centralité de la vérification des faits en tant que pratique essentielle pour la santé du débat public », et que « Les plateformes devraient collaborer avec les organismes de vérification des faits de sorte que leurs utilisateurs soient exposés à des informations de qualité sur des sujets de débat public et encouragés à les partager, ainsi qu'à remettre en question et à démystifier la mésinformation et la désinformation auxquelles ils sont confrontés » (par. 9).

38. Commission européenne, *Lignes directrices sur la protection des mineurs*, précitées, section 6.5.
39. Voir également la Recommandation CM/Rec(2022)13 sur les effets des technologies numériques sur la liberté d'expression, paragraphe 1.5.

179. Le rôle important des dispositifs de labellisation par des tiers et de modération de contenu a également été reconnu dans les efforts réglementaires visant à lutter contre la désinformation, notamment à travers le Code de conduite sur la désinformation[40], un instrument de corégulation relevant du Règlement de l'Union européenne sur les services numériques. Les engagements pris dans le cadre du code comprennent l'utilisation et l'intégration de la vérification des faits dans les services des signataires, grâce à des mesures telles que la collaboration avec des vérificateurs de faits indépendants et l'utilisation de mécanismes tels que l'étiquetage, les panneaux d'information ou l'application de politiques visant à accroître l'impact des vérifications des faits sur le public.

Concernant le paragraphe 73

180. Les plateformes exerçant une influence significative, grâce à leurs systèmes de recommandation et de modération de contenu, orientent fondamentalement le discours public, influençant à la fois la visibilité et l'accessibilité de l'information. Les systèmes qui sont conçus, exploités et contrôlés exclusivement par les fournisseurs de plateformes eux-mêmes peuvent offrir aux utilisateurs un choix limité et peu significatif quant aux mécanismes qui façonnent leur expérience en ligne, ce qui pourrait soulever d'importantes préoccupations en matière de liberté d'expression, de pluralisme et de participation démocratique. Le paragraphe 73 recommande aux États d'explorer, par le biais de processus multipartites et fondés sur des éléments factuels, conformément aux principes généraux, la possibilité d'introduire dans leurs cadres réglementaires l'obligation pour ces plateformes d'autoriser le déploiement d'outils développés par des tiers pour remplir ces fonctions. Ces outils, parfois appelés « *middleware* », peuvent offrir une plus grande diversité d'expression et avoir un impact positif sur la circulation de l'information[41]. Ils pourraient notamment permettre de mieux desservir différents marchés régionaux et linguistiques, ainsi que l'émergence de nouveaux services exclusivement axés sur l'intérêt public. Par exemple, les utilisateurs pourraient choisir un système de recommandation donnant la priorité aux sources journalistiques vérifiées ou un système adapté à la communauté locale.

181. Le pouvoir d'action des utilisateurs à cet égard pourrait être renforcé en leur offrant la possibilité de confier les décisions relatives à leur expérience utilisateur à des services tiers de leur choix, ce qui leur permettrait de sélectionner parmi une variété d'algorithmes et de fournisseurs tiers ceux qui correspondent le mieux à leurs préférences et à leurs besoins, élargissant ainsi potentiellement les options des utilisateurs et améliorant le choix individuel quant au contenu consulté. Permettre la personnalisation de l'organisation et de la modération des contenus pourrait aider les utilisateurs à adapter leur environnement en ligne à leurs propres préférences et valeurs, qu'il s'agisse d'un filtrage plus strict des contenus, d'une priorité accordée aux sources d'information vérifiées, d'une plus grande importance accordée aux contenus locaux ou pertinents pour la communauté, ou de pratiques de modération sensibles aux contextes culturels et linguistiques. Toutefois, il est également reconnu que le recours à des fournisseurs tiers d'outils et de services de contenu peut susciter des inquiétudes quant au pluralisme, car il peut renforcer les chambres d'écho, les bulles d'information et l'effet « *rabbit hole* », entraîner des questions délicates concernant le partage des responsabilités entre le fournisseur d'outils tiers et les plateformes d'hébergement, et soulever des problèmes techniques.

182. Par conséquent, tout en encourageant les États à envisager cette approche, le paragraphe 73 prévoit également certaines limites préliminaires. L'intégration d'outils et de services tiers ne devrait être autorisée que dans des conditions non discriminatoires et équitables, conformément aux principes établis par la Recommandation CM/Rec(2022)13 sur les effets des technologies numériques sur la liberté d'expression, en particulier son paragraphe 2.3. En outre, les fournisseurs de ces outils devraient être soumis à des normes de transparence, de responsabilisation et de respect des obligations en matière de droits humains, y compris le droit à la vie privée et à la liberté d'expression. Des garanties devraient être mises en place pour protéger les données des utilisateurs, pour assurer l'interopérabilité sans compromettre la sécurité et pour empêcher l'accès à des systèmes tiers abusifs ou discriminatoires. À cet égard, les exigences de transparence devraient s'étendre à la divulgation de la nature et de l'étendue de la participation de tiers au processus de modération de contenu afin de garantir un contrôle efficace.

40. Commission européenne, Direction générale des réseaux de communication, du contenu et des technologies, *Code of conduct on disinformation – as amended in October 2024*, Office des publications de l'Union européenne, 2025, disponible seulement en anglais à l'adresse https://data.europa.eu/doi/10.2759/5029213.
41. France, Conseil national du numérique, *Cultiver la richesse des réseaux*, 2024, disponible à l'adresse https://www.conseil-ia-numerique.fr/nos-travaux/cultiver-la-richesse-des-reseaux.

Concernant le paragraphe 74

183. La nécessité de concevoir des plateformes qui favorisent de manière proactive l'autonomisation et la sécurité des utilisateurs handicapés repose sur un cadre normatif du Conseil de l'Europe (CM/Rec(2016)5 sur la liberté de internet, par. 2.1.1) ainsi que sur le droit international plus large relatif aux droits humains (Convention relative aux droits des personnes handicapées, articles 2 et 21)[42]. Le droit à la liberté d'expression inscrit dans la Convention comprend l'égalité d'accès à l'internet pour tous, sans discrimination. Il est en outre mis en évidence dans la Recommandation CM/Rec(2019)6 sur le développement et la promotion de l'éducation à la citoyenneté numérique, qui stipule que les personnes handicapées, en tant que citoyens numériques, ont le droit d'accéder à des environnements en ligne sûrs et inclusifs et à une éducation à la culture numérique adaptée à leurs besoins. Les personnes handicapées sont souvent confrontées à des obstacles à l'accessibilité qui les empêchent de bénéficier pleinement des mesures de sécurité et d'autonomisation sur les plateformes numériques. Ces obstacles peuvent résulter d'interfaces inaccessibles ou d'un manque de compatibilité avec les technologies d'assistance. Pour remédier à ces risques, la conception des plateformes devrait garantir que les personnes handicapées puissent utiliser et déployer des outils d'accessibilité tiers, tels que des lecteurs d'écran, des systèmes de sous-titrage ou des applications de synthèse vocale, afin que les fonctionnalités d'accessibilité intégrées puissent être efficacement complétées par des outils indépendants adaptés aux besoins individuels. Cela est particulièrement important pour les fonctions de sécurité et d'autonomisation en ligne, telles que les mécanismes de plainte et de signalement, qui ne sont efficaces que si les utilisateurs peuvent y accéder et les utiliser.

Concernant le paragraphe 75

184. Les enfants sont particulièrement vulnérables dans l'environnement numérique, car ils n'ont peut-être pas encore la capacité de reconnaître pleinement les risques en ligne ou de prendre les mesures appropriées pour se protéger. Il est donc essentiel de mettre en place des mesures de protection efficaces qui garantissent leurs droits, assurent leur sécurité et favorisent leur développement sain en ligne. La mise en place de systèmes efficaces de vérification de l'âge est essentielle pour protéger les enfants contre l'exposition à des produits, services et contenus illégaux (tels que le matériel d'abus sexuels sur enfants) ou interdits par la loi à leur tranche d'âge : sites pornographiques et sites de rencontre, ainsi que tout autre contenu spécifique soumis à une restriction d'âge, tel que les jeux de hasard en ligne, la vente en ligne de tabac et d'alcool, les communications commerciales pour des produits et services non destinés aux enfants, ainsi que d'autres types de contenus susceptibles de nuire à leur développement physique, mental ou moral.

185. Les récentes initiatives législatives reflètent une prise de conscience croissante de la nécessité de mettre en place des mécanismes solides de vérification de l'âge dans l'environnement en ligne. La loi sur la sécurité en ligne du Royaume-Uni impose aux fournisseurs réglementés (de services utilisateur à utilisateur) de mettre en œuvre une « vérification très efficace de l'âge » afin d'empêcher les enfants d'accéder à certains types de contenus restreints par la loi, notamment la pornographie et les contenus qui encouragent, promeuvent ou fournissent des instructions pour l'automutilation, les troubles alimentaires ou le suicide. Les lignes directrices connexes sur la vérification très efficace de l'âge, adoptées par l'Ofcom, l'autorité de régulation indépendante chargée des fonctions réglementaires en vertu de la loi, définissent les critères d'efficacité des systèmes (précision technique, robustesse, fiabilité et équité) et exigent que ces systèmes soient proportionnés et préservent la vie privée[43]. L'article 28.b de la Directive de l'Union européenne sur les services de médias audiovisuels exige que les plateformes de partage de vidéos mettent en place et exploitent des systèmes de vérification de l'âge afin de protéger les mineurs contre les contenus audiovisuels susceptibles de nuire à leur développement physique, mental ou moral. L'article 28 du règlement de l'Union européenne sur les services numériques, quant à lui, oblige les plateformes en ligne accessibles aux mineurs à prévoir « des mesures appropriées et proportionnées pour garantir un niveau élevé de protection de la vie privée, de sûreté et de sécurité des mineurs sur leur service ». Dans les Lignes directrices sur la protection des mineurs, adoptées en vertu de cette disposition, la Commission européenne a précisé que la vérification de l'âge peut constituer une telle mesure, mais a souligné que les plateformes doivent procéder à une évaluation préalable afin de déterminer si la mesure est à la fois appropriée et proportionnée, et si le même niveau de protection pourrait être atteint en recourant à d'autres « mesures moins contraignantes » (par. 31). Les lignes directrices adoptent une approche fondée sur les risques en matière de vérification de l'âge. Elles considèrent par exemple que la vérification de l'âge est nécessaire pour protéger les enfants contre les risques élevés, tels que

42. Nations Unies, Convention relative aux droits des personnes handicapées, UNTS 2515, p. 3, disponible à l'adresse https://treaties.un.org/pages/viewdetails.aspx?src=treaty&mtdsg_no=iv-15&chapter=4&clang=_en
43. Royaume-Uni, Ofcom, Partie 3 Guidance on Highly Effective Age Assurance, 2025, disponible à l'adresse www.ofcom.org.uk/online-safety/online-safety-regulatory-documents, pp. 10-12.

la vente d'alcool, de tabac, de produits liés à la nicotine ou l'accès à tout contenu pornographique. À l'inverse, lorsque les risques pour la sécurité des enfants sont moyens, l'estimation de l'âge serait appropriée et proportionnée. En vertu de l'article 35, les très grandes plateformes en ligne et les moteurs de recherche ont l'obligation supplémentaire de prendre des mesures ciblées pour protéger les droits de l'enfant, y compris la vérification de l'âge et des outils de contrôle parental, ou des outils permettant d'aider les mineurs à signaler les abus ou à obtenir un soutien, s'il y a lieu.

186. Le paragraphe 75 recommande une approche plus stricte en matière de vérification de l'âge lorsque les plateformes proposent principalement des contenus ou des services, tels que la pornographie ou les jeux de hasard, dont l'accès est restreint aux mineurs par la loi. Cette expression vise à couvrir les plateformes sur lesquelles une grande partie des interactions des utilisateurs concerne, par exemple, des contenus pornographiques, même si elles hébergent d'autres contenus non soumis à des restrictions légales. Des exigences plus strictes en matière de vérification de l'âge des utilisateurs pour ce type de plateformes, et en particulier pour celles sur lesquelles des contenus pornographiques sont partagés, sont déjà prévues par la législation de plusieurs États membres du Conseil de l'Europe[44].

187. Les outils de vérification de l'âge doivent être conçus non seulement pour restreindre l'accès des enfants, mais aussi pour défendre leurs droits en trouvant un juste équilibre : les protéger des contenus et services qui leur sont interdits par la loi tout en évitant le risque de barrières excessives qui les excluraient des contenus et services légitimes. Par exemple, la simple déclaration d'âge est largement reconnue comme insuffisante, car elle peut être facilement détournée et ne peut donc être considérée comme une forme adéquate de vérification de l'âge. Certaines méthodes qui peuvent sembler techniquement efficaces pour vérifier l'âge peuvent, dans la pratique, porter atteinte aux droits des enfants d'une manière qui n'est pas immédiatement apparente. En particulier, les méthodes d'estimation de l'âge qui utilisent des données biométriques telles que la voix ou les traits du visage (analyse faciale par IA ou reconnaissance vocale) ou le profilage des utilisateurs sur la base de leur comportement en ligne suscitent des préoccupations considérables en matière de confidentialité et de sécurité des données. C'est pourquoi les exigences relatives au déploiement de systèmes de vérification de l'âge devraient s'accompagner de lignes directrices claires et pratiques sur leur sécurité, leur transparence et leur caractère inclusif, en accordant une attention particulière aux méthodes utilisant l'intelligence artificielle. Des garanties supplémentaires devraient inclure un contrôle indépendant et l'imposition d'obligations de transparence et de responsabilité aux plateformes. Les plateformes devraient être tenues de publier des explications claires et accessibles sur le fonctionnement de leurs systèmes de vérification de l'âge, en précisant quelles données sont collectées, comment les décisions sont prises et comment les utilisateurs peuvent demander une correction en cas d'erreur. En outre, les plateformes devraient être tenues de fournir régulièrement des rapports de transparence sur les performances de ces systèmes, y compris des informations sur leur précision, leurs taux d'erreur et les mesures prises pour atténuer les risques pour les droits des enfants.

188. Les systèmes de vérification de l'âge peuvent également porter atteinte aux droits des adultes, par exemple en exigeant le partage excessif d'informations personnelles ou en restreignant l'accès à certains contenus et services. De plus, le recours excessif à des systèmes qui exigent des compétences numériques avancées ou la fourniture de documents d'identité officiels peut, sans le vouloir, aggraver les risques d'exclusion des personnes déjà menacées de marginalisation dans la société ou dans les espaces en ligne. Par conséquent, les garanties exposées ci-dessus devraient tenir compte de la nécessité d'éviter toute ingérence non nécessaire ou disproportionnée dans les droits des enfants comme des adultes, et accorder une attention particulière à leurs effets sur les personnes à risque de marginalisation et de discrimination.

Concernant le paragraphe 76

189. Le déploiement d'outils parentaux doit être considéré à la fois comme une mesure d'atténuation des risques et comme une mesure d'autonomisation. Ces outils permettent aux parents et aux tuteurs non seulement de gérer l'exposition des enfants à des contenus et services potentiellement soumis à des restrictions légales, mais aussi de guider activement leur engagement numérique et leurs habitudes de consommation. Toutefois, ces outils, ainsi que toute mesure visant à atténuer les risques de préjudice pour les enfants, doivent placer l'intérêt supérieur de l'enfant au premier plan, conformément à l'article 3 de la Convention des Nations Unies relative aux droits de l'enfant. En outre, le niveau de contrôle parental devrait être proportionné à l'âge et à la maturité de l'enfant, afin de garantir que les enfants bénéficient d'une autonomie croissante à mesure qu'ils grandissent. Cette approche est conforme au principe 2.2 sur l'évolution des capacités de l'enfant figurant dans la Recommandation CM/Rec(2018)7 sur les Lignes directrices relatives au respect, à

44. Royaume-Uni, Online Safety Act, partie 5 ; Commission européenne, *Lignes directrices sur la protection des mineurs*, précitées.

la protection et à la réalisation des droits de l'enfant dans l'environnement numérique, qui reconnaît que les politiques et les pratiques doivent répondre de manière appropriée aux besoins différents des jeunes enfants et des adolescents.

Concernant le paragraphe 77

190. Le paragraphe 77 introduit une mesure visant à permettre aux créateurs de contenu de signaler, à toute plateforme hébergeant leur contenu ainsi qu'à leurs utilisateurs, qu'un contenu déterminé peut ne pas convenir à un public en dessous d'un certain âge. Une telle indication pourrait être exigée en vertu des cadres juridiques nationaux, des normes professionnelles ou éthiques, d'autres régimes d'autorégulation. Elle peut être aussi prise sur volontairement par le créateur de contenu afin d'assumer la responsabilité de la protection des enfants. Cette option doit être comprise comme une mesure qui donne aux créateurs de contenu les moyens de contribuer à un environnement en ligne plus sûr pour les enfants. Par exemple, un diffuseur qui est tenu, en vertu de la législation applicable aux médias, d'utiliser des labels d'âge pour les contenus qui ne conviennent pas aux enfants d'un certain âge devrait pouvoir appliquer les mêmes labels à des contenus identiques diffusés via ses réseaux sociaux. De même, les créateurs de contenu qui ne sont pas des acteurs professionnels des médias mais qui souhaitent exercer une responsabilité professionnelle devraient avoir accès à des fonctionnalités qui leur permettent de le faire.

Concernant le paragraphe 78

191. Le déploiement de systèmes de vérification de l'âge, de contrôles parentaux, d'étiquetage des contenus et d'autres outils ne doit pas être considéré comme suffisant en soi ou comme une solution aux risques auxquels les enfants sont exposés en ligne. La mise à disposition d'outils parentaux par les plateformes ne doit pas être interprétée comme un transfert de responsabilité des plateformes vers les parents, et la mise à disposition d'outils d'étiquetage ne doit pas décharger les plateformes de leurs propres responsabilités en ce qui concerne les contenus qui sont restreints par la loi ou contraires aux règles et politiques contractuelles destinées à protéger les enfants. Si ces outils peuvent aider les parents et les représentants légaux à guider et à protéger leurs enfants et les créateurs de contenu à agir de manière responsable et à se conformer à leurs propres devoirs et responsabilités, les plateformes restent en fin de compte responsables de la réduction des risques. Cette responsabilité comprend notamment la prévention de l'hébergement et de la diffusion de contenus illégaux, tels que les matériels d'abus sexuels sur enfants.

Concernant le paragraphe 79

192. La possibilité pour les utilisateurs de transférer leur identité, leurs données et leurs contenus en ligne d'un service à l'autre est un aspect essentiel de leur autonomie numérique et, par conséquent, de leur autonomisation. Les restrictions à la portabilité des profils en ligne enferment les utilisateurs dans un service déterminé, sous peine de perdre pour leur visibilité et leur audience, ce qui réduit également les possibilités de pluralisme dans l'environnement numérique. Pour les créateurs de contenu en particulier, la possibilité de transférer leurs profils et leur audience d'une plateforme à l'autre contribue à préserver la liberté d'expression et la créativité artistique. Elle évite une dépendance excessive à l'égard des politiques ou des algorithmes d'une seule plateforme et contribue à une sphère publique numérique plus diversifiée et plus résiliente. Les États devraient donc exiger des plateformes qu'elles adoptent des choix de conception et des normes techniques qui favorisent la portabilité des profils.

193. Faciliter la portabilité est également conforme aux normes européennes plus larges, notamment aux droits des personnes en vertu de la législation sur la protection des données sur la protection des données qui leur permettent d'accéder à leurs données à caractère personnel et de les transférer, ainsi qu'aux objectifs d'interopérabilité et d'ouverture dans la gouvernance des services numériques. Le Règlement de l'Union européenne sur les marchés numériques[45], par exemple, accorde des droits de portabilité des données à tous les utilisateurs dans le but de promouvoir l'équité et la contestabilité sur les marchés numériques en limitant le pouvoir de contrôle des plateformes dominantes, notamment en améliorant l'accès aux données et la portabilité pour les utilisateurs.

45. Union européenne, *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 (règlement sur les marchés numériques)*, disponible à l'adresse http://data.europa.eu/eli/reg/2022/1925/oj.

Transparence

Concernant le paragraphe 80

194. Les exigences en matière de transparence sont essentielles pour comprendre comment les politiques et les pratiques des plateformes influencent la liberté d'expression, car elles fournissent aux utilisateurs les connaissances nécessaires pour interpréter la manière dont leur expérience en ligne est façonnée et pour faire des choix éclairés quant à leur engagement vis-à-vis des contenus. Conformément à la Recommandation CM/Rec(2022)13 sur les effets des technologies numériques sur la liberté d'expression, les plateformes devraient assurer une transparence adéquate dans la conception et la mise en œuvre de leurs conditions d'utilisation et de leurs principales politiques, telles que les informations concernant la suppression, la recommandation, l'amplification, la promotion, le déclassement, la monétisation et la distribution de contenus, en particulier en ce qui concerne leurs conséquences sur la liberté d'expression. L'obligation d'expliquer les systèmes algorithmiques d'organisation et de sélection des contenus exige des plateformes qu'elles informent clairement les utilisateurs sur la manière dont les contenus sont classés, hiérarchisés et personnalisés, par exemple sur la manière dont leur comportement et leurs interactions avec les contenus ou d'autres utilisateurs influencent leur expérience utilisateur par le biais de recommandations personnalisées et du classement des contenus. Sur la base de cette compréhension, les utilisateurs devraient être en mesure de modifier les paramètres utilisés pour l'organisation et la sélection des contenus et ainsi prendre des décisions éclairées sur la visibilité des contenus dans leurs flux ou leurs résultats de recherche. L'objectif de cette exigence de transparence est donc double : premièrement, favoriser l'autonomie éclairée des utilisateurs et leur capacité à influencer les informations qui leur sont présentées ; deuxièmement, réduire les risques de préjudice en renforçant le contrôle des utilisateurs sur l'organisation et la sélection algorithmiques des contenus.

195. Ces explications devraient figurer dans les conditions d'utilisation de la plateforme. Pour être pertinentes, elles devraient être compréhensibles pour le public ciblé, ce qui peut impliquer leur disponibilité dans les langues locales et minoritaires et la fourniture d'exemples contextuels[46]. En outre, elles devraient être aussi précises que possible dans l'explication des critères qui déterminent l'ordre d'apparition des contenus, ainsi que dans la justification de la pondération ou de l'importance accordée à ces critères (CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, par. 2.2.3). L'autonomie et le contrôle des utilisateurs peuvent être renforcés en expliquant comment les paramètres d'organisation et de sélection des contenus peuvent être modifiés.

Concernant les paragraphes 81 et 82

196. La transparence des pratiques et des décisions en matière de modération de contenu est un mécanisme essentiel de la responsabilisation des plateformes, compte tenu notamment de leurs implications importantes pour la liberté d'expression, l'accès à l'information et la diversité du débat public. La publication de rapports de transparence est largement reflétée dans les normes internationales et les cadres juridiques. La Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, au paragraphe 2.2.4, préconise la publication « régulière » de rapports de transparence fournissant des informations simples, facilement accessibles et significatives sur toutes les restrictions de contenu, y compris sur les motifs de ces décisions, par exemple les injonctions judiciaires, les demandes des utilisateurs ou l'application des politiques propres aux plateformes. Les plateformes devraient donc être tenues de fournir des rapports de transparence à intervalles réguliers. Le Règlement de l'Union européenne sur les services numériques exige que toutes les plateformes publient des rapports annuels sur la transparence détaillant les mesures de modération de contenu, tandis que les très grandes plateformes en ligne et les moteurs de recherche sont tenus de les soumettre tous les six mois (articles 15 et 42).

197. L'obligation de déclaration devrait inclure des informations quantitatives et qualitatives. Les statistiques devraient indiquer le nombre de signalements reçus, les mesures prises et les résultats de ces mesures, ventilés selon des catégories qualitatives. Le principe opérationnel n° 1 (« Chiffres ») des Principes de Santa Clara[47] décrit en détail les attentes en matière de transparence à cet égard, telles que la fourniture d'informations sur la source des signalements (acteurs étatiques, signaleurs de confiance, utilisateurs, automatisation), les

46. Voir, par exemple, UNESCO, *Towards user empowerment : a multistakeholder action plan for integrating media and information literacy on digital platforms*, précité.

47. *Principes de Santa Clara 2.0 sur la transparence et la responsabilité dans la modération des contenus*, 2021, disponibles à l'adresse https://santaclaraprinciples.org/fr/. Il convient de noter que les principes de Santa Clara « été développée pour soutenir les entreprises à respecter leurs responsabilités envers les droits humains et améliorer leur redevabilité, et pour soutenir les défenseurs des droits humains dans son travail. Elles ne sont pas proposées comme un modèle de régulation ».

motifs de l'action (violation de la loi ou des règles et politiques propres aux plateformes) et le fait que les décisions de modération de contenu ont été prises sous contrôle humain ou par des processus automatisés.

198. Les systèmes automatisés de modération de contenu ne sont pas en mesure de comprendre le contexte et les nuances nécessaires à une évaluation précise, ce qui peut entraîner un excès ou une insuffisance de suppression de contenus. Ces risques sont particulièrement élevés lorsque ces systèmes fonctionnent sans supervision humaine adéquate ou lorsque les garanties visant à prévenir les résultats discriminatoires ou biaisés sont insuffisantes ou inexistantes. Outre les statistiques sur les décisions de modération de contenu prises dans le cadre de processus automatisés, les rapports de transparence devraient donc inclure des explications qualitatives relatives à l'utilisation de la modération algorithmique. Ces rapports devraient indiquer les systèmes algorithmiques utilisés pour la modération de contenu, les catégories ou types de contenu auxquels ils s'appliquent (par exemple, discours de haine, application des droits d'auteur ou contenu terroriste) et fournir des descriptions claires du processus décisionnel, y compris les critères clés et la logique utilisés[48]. Les rapports devraient également inclure des évaluations significatives des performances du système, telles que les taux de précision et de réussite, les marges d'erreur et les garanties mises en place pour protéger les droits des utilisateurs. Des mesures de performance significatives peuvent nécessiter la communication séparée des taux de faux positifs et de faux négatifs (plutôt que la seule précision globale), la ventilation des performances par type de contenu et par langue, ainsi que des mesures de précision contextuelle, en reconnaissant que les systèmes automatisés ne peuvent pas distinguer le contexte. L'obligation de rendre compte des indicateurs de précision, des taux d'erreur et des garanties associés aux systèmes de modération automatisés est reflétée dans le Règlement de l'Union européenne sur les services numériques (article 15). Au Royaume-Uni, les directives finales sur la transparence en matière de rapports sur la sécurité en ligne, élaborées par l'Ofcom, indiquent que le régulateur peut exiger des fournisseurs de services qu'ils rendent compte de l'utilisation de systèmes automatisés dans la modération de contenu, y compris les mesures de précision ou les mécanismes d'assurance qualité qui sont appliqués[49].

199. Une garantie essentielle réside dans la possibilité pour les utilisateurs d'obtenir un examen humain de toute décision de modération automatisée du contenu. Un examen humain significatif nécessite plus qu'une simple implication symbolique et doit être effectué par des personnes disposant d'une autorité réelle et de ressources adéquates. Ainsi, les examinateurs doivent disposer de l'autorité et des capacités appropriées pour modifier une décision prise par des moyens automatisés et avoir accès à toutes les données pertinentes, telles que le contexte d'un contenu signalé et le raisonnement algorithmique qui a déclenché la décision automatisée. Des allocations de temps réalistes, permettant un examen approfondi et le déploiement adéquat du personnel, sont également essentielles à une supervision efficace.

Concernant le paragraphe 83

200. La publicité est un élément fondamental du financement des services des plateformes en ligne. Cependant, la publicité numérique prend souvent des formes diverses, dont certaines sont conçues pour se fondre parfaitement dans le contenu habituel, ce qui rend difficile pour les utilisateurs de faire la distinction entre les informations éditoriales et les contenus sponsorisés. Ce flou nuit à la capacité des utilisateurs à évaluer de manière critique ce qu'ils voient en ligne et augmente le risque de manipulation. Le manque de transparence en matière publicitaires facilite les pratiques commerciales cachées qui exploitent les vulnérabilités des consommateurs. Ces risques sont accrus dans le cas de la publicité ciblée, qui repose sur le profilage des utilisateurs grâce à la collecte de données à caractère personnel et au suivi de leur comportement en ligne. De telles pratiques peuvent avoir de graves conséquences pour les droits des utilisateurs en matière de protection des données, tout en causant des préjudices plus larges à la société. Elles peuvent également amplifier les préjudices causés aux processus démocratiques et électoraux, en particulier compte tenu des préoccupations liées à la propagation de la désinformation et au microciblage politique.

201. La transparence quant à la source de la publicité est un outil essentiel pour l'autonomisation des utilisateurs. Elle permet à ces derniers de reconnaître clairement quand un contenu est payant, par qui il l'est et quels sont les montants dépensés en publicité. Ces informations devraient être facilement et clairement accessibles afin que les utilisateurs puissent identifier sans ambiguïté le contenu comme étant une publicité, le distinguer d'autres formes de contenu et comprendre la personne ou l'entité pour le compte de

48. Ces informations devraient être mises à la disposition des utilisateurs dans les conditions d'utilisation, conformément au principe énoncé au paragraphe 3.4 de la Recommandation CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expression voir également l'exposé des motifs concernant le paragraphe 80.
49. Royaume-Uni, Ofcom, Online Safety Transparency Reporting. Final Transparency Guidance, 2025, disponible à l'adresse https://www.ofcom.org.uk/online-safety/online-safety-regulatory-documents, en particulier l'annexe A.

laquelle la publicité est placée. Une pratique courante consiste à apposer des étiquettes directement sur les publicités. Par exemple, le code de conduite de l'UE sur la désinformation prévoit de telles mesures pour la publicité politique. Il exige en outre des signataires qu'ils tiennent à jour un registre accessible au public des publicités politiques et thématiques. L'article 39 de la loi sur les services numériques impose également aux très grandes plateformes en ligne et aux moteurs de recherche de tenir à jour des registres accessibles au public contenant des données sur les publicités affichées, l'annonceur et le bénéficiaire, le public cible et les dépenses publicitaires globales.

202. La transparence publicitaire exige également que les utilisateurs puissent comprendre pourquoi ils voient une publicité particulière. Cela inclut des explications claires et accessibles sur l'utilisation éventuelle de techniques de ciblage et les paramètres appliqués, tels que les critères démographiques, géographiques, contextuels, basés sur les centres d'intérêts ou comportementaux (code de conduite de l'UE sur la désinformation, mesure 9.2). En fournissant ces informations dans un format compréhensible, les plateformes favorisent le choix éclairé des utilisateurs et atténuent les risques d'influence et de manipulation dissimulées.

Concernant le paragraphe 84

203. Les plateformes offrent de plus en plus aux créateurs de contenu la possibilité de monétiser directement leur contenu, par exemple en participant à des programmes de redistribution des revenus. À mesure que l'économie des créateurs gagne en ampleur et en influence, la nécessité d'une transparence dans la manière dont les plateformes monétisent le contenu généré par les utilisateurs est de plus en plus reconnue, en particulier en ce qui concerne les bénéficiaires, les conditions et les modèles d'allocation. Des pratiques de monétisation opaques peuvent créer des risques tant pour les utilisateurs que pour les créateurs de contenu. Pour les utilisateurs, le manque de clarté quant au contenu monétisé, par qui il l'est et dans quelles conditions, limite leur capacité à évaluer les incitations commerciales potentielles qui se cachent derrière le contenu qu'ils consomment. Dans le même temps, les créateurs de contenu dépendent des systèmes de monétisation des plateformes pour atteindre leur public et générer des revenus. Si les critères qui déterminent la répartition des ressources sont opaques ou imprévisibles, les créateurs peuvent être confrontés à une dépendance excessive ou à une pression pour adapter leur contenu aux préférences algorithmiques ou politiques. Une autre préoccupation concerne la démonétisation excessive, souvent effectuée par des systèmes automatisés, qui peut avoir un impact significatif sur les moyens de subsistance des créateurs.

204. La transparence en matière de monétisation est également étroitement liée à l'intégrité des processus démocratiques. L'un des risques liés aux pratiques de monétisation des plateformes est qu'elles créent des incitations financières à la publication ou à l'amplification de contenus et d'activités en ligne soumis à des restrictions légales. En particulier, la monétisation de la désinformation pose de sérieux risques pour la confiance du public et le discours démocratique, car les systèmes publicitaires opaques canalisent souvent les revenus vers des sites web qui diffusent délibérément de la désinformation. La divulgation transparente des pratiques de monétisation, y compris les critères utilisés pour allouer les ressources, est donc essentielle pour garantir que les contenus restreints par la loi ne soient pas récompensés financièrement, tout en préservant des opportunités équitables pour les créateurs légitimes et en soutenant un écosystème d'information numérique plus sain.

205. L'orientation fournie par la Commission européenne concernant l'interprétation et l'application de la Directive sur les pratiques commerciales déloyales précise que le marketing d'influence (y compris les publications rémunérées, le contenu affilié, les retweets ou le taggage du commerçant/de la marque) doit clairement indiquer lorsque la promotion est rémunérée, et que cette mention doit être bien visible. Cette obligation s'applique également lorsque des influenceurs font la promotion de leurs propres produits ou activités[50]. Au niveau international, les travaux de l'OCDE sur les travailleurs des plateformes et l'économie numérique soulignent la nécessité de mettre en place des systèmes de rémunération équitables et transparents, y compris dans le contexte de la création de contenu, et exhortent les gouvernements à veiller à ce que les créateurs ne soient pas soumis à des changements opaques ou arbitraires dans les politiques de monétisation[51].

50. Commission européenne, *Orientations concernant l'interprétation et l'application de la directive 2005/29/CE du Parlement européen et du Conseil relative aux pratiques commerciales déloyales des entreprises vis-à-vis des consommateurs dans le marché intérieur*, Notice 2021/C 526/01, 2021, disponible à l'adresse https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021XC1229(05), section 4.2.6.
51. Lane M., « Regulating platform work in the digital age », *OECD Going Digital Toolkit Notes*, n° 1, Éditions OCDE, Paris, 2020, disponible à l'adresse https://doi.org/10.1787/181f8a7f-en.

Concernant le paragraphe 85

206. Les plateformes exerçant une influence significative devraient fournir des outils permettant aux créateurs de contenu de divulguer la manière dont leur contenu est monétisé, c'est-à-dire comment il génère des revenus. Cela peut inclure la divulgation de l'identité de leurs partenaires commerciaux, la garantie de la visibilité de leurs accords publicitaires tels que les parrainages ou les partenariats d'affiliation, et la participation à des programmes de monétisation spécifiques à la plateforme. Pour les utilisateurs, la transparence de ces relations financières est cruciale, car elle leur permet de comprendre les potentiels biais commerciaux, d'évaluer la crédibilité du contenu et de prendre des décisions éclairées sur ce qu'ils consomment. Ces outils pourraient prendre la forme d'étiquettes visibles ou de balises de métadonnées sur les publications ou les vidéos, ou encore de tableaux de bord accessibles aux utilisateurs, indiquant les sources de revenus ou les relations commerciales liées à un contenu spécifique.

207. L'article 28, point (b), de la Directive de l'Union européenne sur les services de médias audiovisuels, par exemple, exige que les plateformes de partage de vidéos disposent d'une fonctionnalité permettant aux utilisateurs qui téléchargent des vidéos générées par les utilisateurs de déclarer si ces vidéos contiennent des communications commerciales audiovisuelles, reconnaissant que les plateformes ne peuvent pas toujours avoir connaissance de ces arrangements. Les nouveaux accords d'autorégulation qui voient le jour à travers l'Europe fournissent de bons exemples de pratiques en matière de divulgation de l'identité des sponsors. En Irlande, par exemple, la Commission de la concurrence et de la protection des consommateurs et l'Autorité de régulation de la publicité ont publié des lignes directrices communes sur la publicité et le marketing d'influence[52], dont les mesures s'appliquent aux créateurs de contenu, qu'ils soient ou non enregistrés en tant que fournisseurs de services de médias audiovisuels auprès de l'autorité de régulation des médias.

Concernant les paragraphes 86 et 87

208. Un examen indépendant est essentiel pour comprendre les risques en ligne, y compris ceux qui découlent de la conception, du fonctionnement et des politiques des plateformes, ainsi que pour évaluer l'efficacité des mesures adoptées par les plateformes pour atténuer ces risques. Sans accès garanti par la loi aux données des plateformes, la recherche repose sur la coopération volontaire, incohérente et sélective des plateformes ou sur l'accès par le biais d'outils tiers. À l'échelle internationale, l'accès des chercheurs aux données demeure *ponctuel* et difficile. L'obligation pour les plateformes d'accorder aux chercheurs l'accès aux données est de plus en plus soutenue au niveau international, notamment dans les Principes de l'OCDE sur l'intelligence artificielle[53] et les Principes de l'UNESCO pour la gouvernance des plateformes numériques (principe 5). Ces deux instruments soulignent que cet accès doit être sécurisé, éthique et conforme aux lois sur la protection des données.

209. Étant donné qu'un examen approfondi des plateformes nécessite l'accès à des données granulaires et souvent sensibles pour évaluer les risques en ligne, les chercheurs ne devraient pas être empêchés d'accéder à des données à caractère personnel et confidentiel et de les traiter lorsque cela est nécessaire à des fins de recherche, sous réserve que toutes les garanties nécessaires en matière de protection des données soient en place. Ces garanties devraient porter à la fois sur la protection de la vie privée et des données à caractère personnel des utilisateurs (par exemple, grâce à l'anonymisation des ensembles de données) et sur la protection des informations propriétaires et des secrets commerciaux des entreprises.

210. Comme indiqué au paragraphe 6.6 de la Recommandation CM/Rec(2022)13 sur les effets des technologies numériques sur la liberté d'expression, les plateformes devraient être tenues d'accorder l'accès aux données individuelles aux chercheurs qui ont été sélectionnés par un établissement scientifique indépendant, remplissant des critères en matière de qualifications, d'expertise et d'indépendance vis-à-vis des intérêts commerciaux et politiques, et approuvés par le comité d'éthique de leur université affiliée. Les États peuvent également mettre en place des systèmes dans lesquels les plateformes sont tenues de donner accès aux chercheurs agréés par une autorité indépendante sur la base des mêmes critères, y compris ceux qui ne sont pas affiliés à des universités.

211. L'accès aux données publiques devrait être soumis à des exigences moins strictes et pourrait également être accordé à d'autres parties prenantes menant des recherches, telles que la société civile ou les médias, à

52. Irlande, Commission de la concurrence et de la protection des consommateurs (CCPC) et Autorité de régulation de la publicité (ASAI), *Guidance on influencer advertising and marketing*, 2023, disponible à l'adresse www.ccpc.ie/business/help-for-business/guidelines-for-business/influencer-advertising-and-marketing/.

53. OCDE, *Recommandation du Conseil sur l'intelligence artificielle*, 2019, disponible à l'adresse https://legalinstruments.oecd.org/fr/instruments/oecd-legal-0449, principes 116-118 (accès aux données à des fins de recherche).

condition qu'elles démontrent leur respect des normes en matière de protection et de sécurité des données et confirment que leurs travaux sont menés de manière indépendante et dans l'intérêt public. En outre, les plateformes devraient permettre aux chercheurs indépendants d'accéder aux données accessibles au public sans restrictions techniques. Cet accès ne devrait pas se limiter aux interfaces gérées par les plateformes, telles que les interfaces de programmation d'applications (API), mais devrait également permettre la collecte de données par « *scraping* », c'est-à-dire l'extraction automatisée de données à partir des interfaces accessibles aux utilisateurs des sites web ou des applications concernés.

212. Le paragraphe 87 renvoie notamment à l'obligation pour les plateformes de permettre aux chercheurs indépendants de mener des recherches sur les plateformes et d'interagir directement avec les systèmes des plateformes afin d'examiner l'impact que les services des plateformes peuvent avoir sur les droits humains et d'autres risques découlant de leur conception, de leur fonctionnement, de leurs politiques et de leurs règles, ainsi que d'évaluer l'efficacité des mesures visant à autonomiser les utilisateurs. Par exemple, les chercheurs pourraient, à titre expérimental, publier certains types de contenus soumis à des restrictions légales afin d'évaluer soit leur visibilité, soit la manière dont les systèmes de modération de la plateforme réagissent à de tels contenus.

Droits procéduraux

Concernant le paragraphe 88

213. Les politiques et règles contractuelles des plateformes, telles que les conditions de service et les standards communautaires, définissent les limites du comportement, du contenu et des interactions autorisées aux utilisateurs. Elles devraient donc être rendues publiques dans un langage clair et simple, qui soit compréhensible et facilement accessible à l'utilisateur moyen du service. En outre, elles doivent être disponibles dans les langues couramment parlées par les utilisateurs de la plateforme et les communautés concernées, afin de garantir leur accessibilité et leur compréhension. Elles doivent également expliquer explicitement les conséquences potentielles de leur application. Les motifs des décisions d'application les plus sévères, telles que la suspension d'un compte ou la suppression de contenu, doivent être indiqués avec une clarté particulière, en évitant toute formulation vague ou ambiguë.

214. Une attention particulière devrait être accordée aux besoins des enfants, qui représentent une part importante des utilisateurs des plateformes et qui sont particulièrement vulnérables aux dispositions contractuelles opaques ou trop complexes. Conformément à la Recommandation CM/Rec(2018)7 sur les Lignes directrices relatives au respect, à la protection et à la réalisation des droits de l'enfant dans l'environnement numérique, les règles contractuelles applicables aux enfants devraient être formulées dans un langage simple et adapté à leur âge, leur permettant de comprendre à la fois leurs droits et leurs responsabilités lorsqu'ils utilisent des services en ligne. Un service peut être considéré comme pertinent pour les enfants s'il leur est destiné ou s'il est principalement utilisé par eux.

215. Des modifications importantes des conditions d'utilisation ou des normes communautaires peuvent avoir des implications considérables sur les droits et les activités des utilisateurs, notamment leur capacité à créer, partager ou monétiser du contenu. Les plateformes devraient donc être tenues d'informer les utilisateurs à l'avance de toute modification substantielle, en fournissant des explications claires qui soulignent non seulement la nature des changements, mais aussi leurs conséquences pratiques. Les explications relatives aux mises à jour devraient être clairement signalées, par exemple par le biais de notifications directes dans les fils d'actualité des utilisateurs ou en fournissant des versions « annotées » des conditions contractuelles mises à jour mettant en évidence les changements spécifiques de manière à permettre aux utilisateurs de comparer les versions précédentes et mises à jour.

Concernant les paragraphes 89 à 92

216. Lorsqu'elles appliquent leurs politiques et règles, les plateformes peuvent prendre des mesures susceptibles de restreindre la liberté d'expression des utilisateurs et des créateurs de contenu. Ces mesures peuvent inclure la suppression de contenu tel que des publications ou des vidéos, la suspension ou la suppression permanente de comptes d'utilisateurs, la démonétisation de contenu, son déclassement dans les systèmes de recommandation, ou encore la rétrogradation et la suppression de la visibilité de contenu, d'utilisateurs ou de groupes d'utilisateurs. Chacune de ces mesures peut affecter de manière significative sur la capacité des utilisateurs à communiquer, à accéder à l'information ou à maintenir leur présence et leur audience en ligne. Il est donc essentiel que des garanties soient en place pour veiller à ce que les contenus ne soient pas supprimés ou censurés de manière injustifiée. Le bannissement fantôme (« *shadow banning* ») est

particulièrement problématique en raison de son opacité : il restreint la visibilité d'un utilisateur à son insu, le privant ainsi de la possibilité de comprendre ou de contester la décision de la plateforme. Les instruments du Conseil de l'Europe (notamment, CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet et CM/Rec(2022)13 sur les effets des technologies numériques sur la liberté d'expression, partie 4), ainsi que les normes internationales et les cadres réglementaires[54] appellent à la transparence dans la modération de contenu, exigeant des plateformes qu'elles informent les utilisateurs chaque fois que des restrictions sont imposées à leur liberté d'expression, et qu'elles fournissent des voies de recours claires et accessibles. L'absence de retour d'information significatif sur les décisions de modération de contenu peut saper la confiance des utilisateurs dans le système de modération, les décourager de participer à la gouvernance des plateformes et nuire à leur autonomisation.

217. Les notifications envoyées aux utilisateurs dont le contenu ou le compte a été affecté par une décision de modération du contenu doivent expliquer les motifs de cette décision, c'est-à-dire indiquer si elle repose sur une violation de dispositions légales ou des règles propres aux plateformes, et faire référence spécifique à la disposition légale ou à la politique ou règle interne présumée violée. Les utilisateurs doivent également être informés de ce qui a déclenché la procédure, qu'il s'agisse d'une détection automatisée, de signalements ou notifications émanant d'autres utilisateurs ou de signaleurs de confiance, ou d'une demande ou d'une injonction des autorités publiques. L'explication du processus décisionnel consiste à indiquer si la décision a été prise par des moyens automatisés ou après examen humain. En outre, les notifications doivent fournir une explication claire de la procédure disponible pour faire appel de la décision. L'article 17 du Règlement de l'Union européenne sur les services numériques impose aux plateformes de fournir aux utilisateurs une déclaration claire, détaillée et rapide des motifs de leur action, chaque fois qu'elles suppriment, restreignent ou modèrent de toute autre manière des contenus, des comptes ou des services. Il oblige également les plateformes à veiller à ce que ces déclarations comprennent la base légale ou contractuelle de la mesure, le type de restriction appliquée (par exemple, suppression de contenu, suspension de compte, déclassement, démonétisation ou désactivation de l'accès) et des informations sur les voies de recours disponibles. En vertu de l'article 20, les utilisateurs et les créateurs de contenu peuvent contester ces décisions par le biais d'un système interne de traitement des plaintes convivial et gratuit, que les plateformes sont tenues de traiter en temps utile et auquel elles doivent répondre par un retour d'information clair et motivé.

218. Lorsque des plateformes exerçant une influence significative prennent des mesures concernant un contenu, leur obligation de notification devrait s'étendre au-delà du créateur original de ce contenu. La notification devrait également être adressée à tout utilisateur identifiable directement concerné ou affecté, pour autant que ces utilisateurs aient choisi de recevoir ces notifications. Il peut s'agir, par exemple, d'une personne mentionnée dans le contenu, telle qu'une personne faisant l'objet d'une publication diffamatoire, qui peut avoir un intérêt légitime à savoir si le contenu reste disponible ou a été supprimé. Il pourrait également s'agir des personnes ou des organisations qui ont signalé le contenu, car elles ont initié une demande de modération et ont donc un intérêt évident dans l'issue de la décision de la plateforme.

219. En plus de la disponibilité des systèmes de recours internes, les utilisateurs devraient également pouvoir exercer leur droit de recours par le biais de voies extérieures. Les États devraient créer les conditions nécessaires au fonctionnement de ces mécanismes de recours ou de contrôle indépendants, en garantissant leur légitimité et leur responsabilité, ainsi que des garanties d'indépendance, de transparence et d'accès équitable[55]. Ces mécanismes devraient également être en mesure d'offrir des recours efficaces, notamment la restauration de contenus ou de comptes lorsque les décisions sont jugées injustifiées. Ces mécanismes peuvent prendre la forme d'organismes de règlement alternatifs des litiges, tels que les organismes extrajudiciaires de règlement des litiges (Règlement de l'Union européenne sur les services numériques, article 21), des médiateurs indépendants ou d'autres mécanismes de contrôle indépendants par des tiers. La disponibilité de mécanismes externes ne devrait pas dispenser les plateformes de leur obligation de maintenir un système de recours interne solide, ni limiter le droit des utilisateurs de saisir la justice (CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires d'internet, par. 1.5.2 et 2.5.5).

54. Principes de Santa Clara ; Principes de l'UNESCO pour la gouvernance des plateformes numériques, principe 5 ; Règlement de l'Union européenne sur les services numériques.
55. Par exemple, les Principes directeurs des Nations Unies relatifs aux entreprises et aux droits de l'homme soulignent clairement le rôle des États dans la garantie que les restrictions aux droits humains liées aux entreprises soient corrigées. Ils soulignent également les principes pour des mécanismes de plainte efficaces à cette fin.

Action collective des utilisateurs

Concernant les paragraphes 93 et 94

220. Les signalements par les utilisateurs complètent les efforts de surveillance des plateformes en permettant l'identification à grande échelle des contenus qui enfreignent les règles des plateformes, ce qui est particulièrement important étant donné que la surveillance systématique et proactive est rarement faisable ou proportionnée. Ces mécanismes sont particulièrement utiles pour détecter les violations fortement dépendantes du contexte, telles que le harcèlement ou le discours de haine, qui nécessitent souvent une intervention humaine pour être évaluées avec précision. Il est essentiel que des outils de signalement accessibles et efficaces permettent aux utilisateurs de jouer un rôle actif dans la configuration de leur environnement en ligne et favorisent une meilleure compréhension des politiques et des pratiques d'application des plateformes. Afin de renforcer cette fonction participative, les personnes qui signalent des contenus devraient recevoir un retour d'information significatif, par exemple la confirmation que leur signalement a été reçu, des informations sur toute suite qui y a été donnée et, lorsqu'aucune mesure n'est jugée nécessaire, une explication claire des raisons qui ont motivé cette décision.

221. Les plateformes devraient être tenues de permettre aux utilisateurs de signaler les contenus qu'ils estiment être soumis à des restrictions légales. Il s'agit là d'un mécanisme d'autonomisation important, car ce sont souvent les notifications des utilisateurs qui permettent aux plateformes de prendre conscience de la nature potentiellement contraire à la loi d'un contenu et de prendre les mesures qui s'imposent, comme le souligne le paragraphe 55 de l'annexe. Afin de faciliter la soumission de notifications suffisamment précises et étayées, les intermédiaires devraient être tenus de concevoir leurs mécanismes de notification et d'action de manière à permettre aux utilisateurs de leur fournir tous les éléments nécessaires afin qu'ils puissent agir rapidement et efficacement. À titre d'exemple, le Règlement de l'Union européenne sur les services numériques exige que les mécanismes de notification et d'action permettent la soumission d'une « explication suffisamment étayée des raisons pour lesquelles le particulier ou l'entité allègue que les informations en question sont du contenu illicite », ainsi que l'emplacement électronique exact du contenu en question (article 16). Une norme similaire s'applique au Royaume-Uni au titre du Règlement sur le commerce électronique[56]. L'article 28.b de la Directive européenne sur les services de médias audiovisuels exige que les plateformes de partage de vidéos fournissent aux utilisateurs des systèmes de classification des contenus et mettent en place et exploitent des mécanismes transparents et conviviaux permettant aux utilisateurs de signaler les contenus incitant à la haine ou à la violence, les contenus préjudiciables aux mineurs ou certaines catégories de contenus illégaux. Afin de faciliter la participation des utilisateurs, les mécanismes de notification doivent être clairement visibles, facilement accessibles et conviviaux, à l'image de la conception et de la facilité d'utilisation des outils de signalement utilisés pour signaler les violations des règles propres à une plateforme.

Concernant le paragraphe 95

222. Le paragraphe 95 encourage les États à promouvoir la mise en place d'acteurs professionnels pouvant servir d'experts de confiance pour notifier les contenus soumis à des restrictions légales ou signaler les violations des politiques et règles contractuelles sur les plateformes en ligne. Grâce à leur expertise, ces professionnels, souvent appelés « signaleurs de confiance » ou « notificateurs de confiance », peuvent fournir des signalements et des notifications de meilleure qualité et mieux étayées que les utilisateurs ordinaires, améliorant ainsi la précision et la rapidité des réponses des plateformes. Par exemple, une organisation de la société civile spécialisée dans le contrôle du discours de haine peut identifier les contenus contraires à la loi dans ce domaine de manière plus fiable que le grand public. De même, les signalements effectués par des associations de journalistes ou des conseils des médias pourraient aider les plateformes à répondre aux risques systémiques pesant sur la liberté d'expression sans la restreindre inutilement (voir également la Note d'orientation sur la lutte contre la mésinformation et la désinformation en ligne, par. 37).

223. Les États devraient inciter les plateformes à reconnaître et coopérer avec ces acteurs experts, en leur accordant des privilèges spécifiques, tels que le traitement prioritaire de leurs notifications et de leurs signalements, l'accélération des procédures d'appel et l'accès à des interfaces techniques améliorées pour le signalement, telles que des outils de signalement groupés ou des API. Les États devraient mettre en place des programmes de soutien financier ou institutionnel afin de garantir la viabilité de ces acteurs. Dans le même temps, des garanties sont nécessaires pour assurer leur indépendance et leur responsabilité. La reconnaissance des notificateurs professionnels devrait être fondée sur des critères transparents, tels que

56. Royaume-Uni, *Règlement de 2002 sur le commerce électronique (directive CE), 2002 nº 2013*, disponible à l'adresse www.legislation.gov.uk/uksi/2002/2013/contents, Règlement 19.

l'expertise démontrée, l'indépendance vis-à-vis des intérêts politiques ou commerciaux et le respect des normes éthiques et de protection des données. Des mécanismes de contrôle devraient être mis en place pour le suivi des performances des groupes reconnus et, le cas échéant, révoquer leur statut en cas d'abus.

Concernant le paragraphe 96

224. Les États devraient également favoriser un environnement favorable à la création et à la professionnalisation de groupes ou d'associations d'utilisateurs indépendants capables d'agir collectivement au nom des utilisateurs et des créateurs de contenu. Ces organisations pourraient fournir un soutien juridique, un travail de plaidoyer et une assistance dans le cadre des procédures d'appel, contribuant ainsi à rééquilibrer l'asymétrie de pouvoir entre les utilisateurs individuels et les grandes plateformes en ligne. Pour garantir leur efficacité, les États peuvent accorder à ces groupes des privilèges spécifiques, notamment la prise en compte prioritaire de leurs recours collectifs devant les plateformes, un soutien financier ou logistique pour garantir leur viabilité à long terme, l'accès à des outils de signalement ou à des données améliorées permettant d'identifier les problèmes systémiques, et la reconnaissance formelle de leur droit d'engager une action collective en cas de violation des droits. Une attention particulière devrait être accordée à la promotion de la diversité et de l'inclusion au sein de ces organisations, afin qu'elles représentent adéquatement les intérêts non seulement des créateurs de contenu professionnels, mais aussi des groupes vulnérables, des enfants et des communautés marginalisées qui peuvent être exposés à des risques disproportionnés en ligne.

La Recommandation CM/Rec(2026)4 constitue une avancée significative dans la régulation des plateformes en plaçant les droits humains au cœur de l'environnement en ligne. Elle reconnaît que des espaces numériques sûrs sont essentiels à la liberté d'expression et appelle à des cadres juridiques assurant un contrôle public et une supervision démocratique de la manière dont les plateformes remplissent leurs responsabilités en matière de droits humains. L'accent se déplace des restrictions de contenu vers des exigences portant sur la conception et les systèmes, y compris la curation algorithmique et la modération des contenus. Des mesures efficaces d'évaluation et d'atténuation des risques sont donc attendues des plateformes plus influentes. L'autonomisation des utilisateurs, grâce à la transparence, des outils conviviaux de personnalisation, des voies de recours effectives et des moyens d'action collective, est considérée comme un élément clé de la sécurité en ligne. La recommandation fixe aussi des attentes vis-à-vis des créateurs de contenus responsables et insiste sur des mesures ciblées pour protéger et autonomiser les femmes et les filles, ainsi que les groupes vulnérables, y compris les enfants. En combinant régulation, responsabilité et autonomisation, elle ouvre la voie à une approche de la sécurité en ligne fondée sur les droits humains, dans laquelle la protection de la liberté d'expression et la sécurité des utilisateurs vont de pair.

PREMS 051322 6

FRA

www.coe.int

Le Conseil de l'Europe est la principale organisation de défense des droits humains du continent. Il comprend 46 États membres, dont l'ensemble des membres de l'Union européenne. Tous les États membres du Conseil de l'Europe ont signé la Convention européenne des droits de l'homme, un traité visant à protéger les droits humains, la démocratie et l'État de droit. La Cour européenne des droits de l'homme contrôle la mise en œuvre de la Convention dans les États membres.

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE