

Distr. générale
28 avril 2023
Français
Original : anglais

Assemblée générale

Soixante-dix-septième session

Point 126 a) de l'ordre du jour

Renforcement du système des Nations Unies :

Renforcement du système des Nations Unies

Notre Programme commun

Note d'orientation n° 8 : intégrité de l'information sur les plateformes numériques

Résumé

Seule une coopération internationale renforcée nous permettra de relever les défis qui se posent à nous. Au Sommet de l'avenir qui doit se tenir en 2024, nous pourrons définir ensemble des solutions multilatérales en faveur d'un avenir meilleur en renforçant la gouvernance mondiale pour le bien des générations actuelles et futures (résolution 76/307 de l'Assemblée générale). J'ai été invité, en ma qualité de Secrétaire général, à contribuer aux préparatifs du Sommet en formulant des recommandations pratiques s'inscrivant dans la continuité des propositions que j'ai présentées dans mon rapport intitulé « Notre Programme commun » (A/75/982), établi pour donner suite à la déclaration faite à l'occasion de la célébration du soixante-quinzième anniversaire de l'Organisation des Nations Unies (résolution 75/1 de l'Assemblée générale). La présente note d'orientation constitue une déclinaison de cette contribution. J'y développe les idées initialement proposées dans Notre Programme commun en tenant compte des orientations données ultérieurement par les États Membres et des consultations intergouvernementales et multipartites tenues pendant plus d'un an, ces idées étant fondées sur les buts et principes consacrés dans la Charte des Nations Unies, la Déclaration universelle des droits de l'homme et autres instruments internationaux.

La présente note d'orientation porte principalement sur la manière dont les menaces pesant sur l'intégrité de l'information influencent les progrès réalisés aux niveaux mondial, national et local. Dans Notre Programme commun, j'ai préconisé un consensus construit collectivement sur l'expérience et portant sur les faits, la science et la connaissance. À cette fin, la présente note d'orientation expose les principes potentiels d'un code de conduite qui aidera à guider les États Membres, les plateformes numériques et les autres parties prenantes dans les efforts qu'ils déploient pour rendre l'espace numérique plus inclusif et plus sûr pour tous et toutes, tout en



défendant vigoureusement le droit à la liberté d'opinion et d'expression ainsi que le droit d'accès à l'information. Le code de conduite portant sur l'intégrité de l'information diffusée sur les plateformes numériques est en cours d'élaboration dans le cadre des préparatifs du Sommet de l'avenir. J'espère qu'il servira de guide dans la formulation de mesures de renforcement de l'intégrité de l'information.

Les plateformes numériques sont des outils essentiels qui ont transformé les interactions sociales, culturelles et politiques partout dans le monde, en mettant en relation des citoyens préoccupés par des questions importantes. Grâce à elles, les Nations Unies informent le public et dialoguent directement avec les personnes qui œuvrent en faveur de la paix, de la dignité et de l'égalité sur une planète saine. Ces plateformes ont donné de l'espoir aux peuples pendant les périodes de crise et de lutte, elles ont amplifié des voix qui n'étaient pas entendues auparavant et fait naître des mouvements mondiaux.

Cependant, ces mêmes plateformes ont aussi révélé un pan plus sombre de l'écosystème numérique. Elles ont favorisé la propagation rapide de mensonges et de discours de haine qui ont fait beaucoup de mal partout dans le monde. L'optimisme quant au potentiel des médias sociaux en matière de mise en relation et de dialogue entre les personnes a été tempéré par la mésinformation, la désinformation et les discours haineux qui sont sortis des marges de l'espace numérique et se sont banalisés. On ne saurait trop insister sur les dangers créés par ce phénomène. Les discours haineux et la désinformation véhiculés par les médias sociaux peuvent conduire à la violence et à la mort^a. La capacité de diffuser de la désinformation à grande échelle pour remettre en question des faits scientifiquement établis met en péril la survie de l'humanité (A/75/982, par. 26) ainsi que les institutions démocratiques et les droits humains fondamentaux. Les progrès rapides de la technologie, tels que l'intelligence artificielle générative, ne font qu'intensifier ce risque. Dans le monde entier, les Nations Unies surveillent la façon dont la désinformation et les discours de haine peuvent menacer la réalisation des objectifs de développement durable. Il semble évident qu'il est désormais impossible de continuer « comme si de rien n'était ».

^a A/HRC/42/50 ; A/77/287 ; A/HRC/51/53 ; Nations Unies. Déclaration d'Alice Wairimu Nderitu, Conseillère spéciale pour la prévention du génocide, condamnant la récente escalade des combats en Éthiopie, communiqué de presse, 19 octobre 2022 ; Haut-Commissariat des Nations Unies aux droits de l'homme (HCDH), *Myanmar: Social media companies must stand up to junta's online terror campaign say UN experts*, 13 mars 2023 ; HCDH, *Freedom of speech is not freedom to spread racial hatred on social media: UN experts*, déclaration, 6 janvier 2023 ; Rapporteuse spéciale sur la promotion et la protection du droit à la liberté d'opinion et d'expression, *#JournalistsToo – Les femmes journalistes prennent la parole*, 24 novembre 2021 ; HCDH, *Sri Lanka: Experts dismayed by regressive steps, call for renewed UN scrutiny and efforts to ensure accountability*, communiqué de presse, 5 février 2021.

I. Qu'est-ce que l'intégrité de l'information ?

1. Le concept d'intégrité de l'information fait référence à l'exactitude, la cohérence et la fiabilité des informations. La désinformation, la mésinformation et les discours haineux constituent donc des menaces. S'il n'existe pas de définition universellement acceptée de ces termes, les entités des Nations Unies ont formulé des définitions opérationnelles.
2. La Rapporteuse spéciale sur la promotion et la protection du droit à la liberté d'opinion et d'expression définit la désinformation comme « la diffusion intentionnelle d'une fausse information en vue de causer un préjudice social grave »¹. L'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO) utilise le terme « désinformation » pour décrire des contenus faux ou trompeurs qui peuvent causer un préjudice particulier, indépendamment des intentions, du degré de conscience et des comportements qui sous-tendent leur production et leur diffusion².
3. Aux fins de la présente note d'orientation, la mésinformation et la désinformation se distinguent par l'intention qui les motive.
4. La désinformation concerne des informations qui sont non seulement fausses, mais aussi destinées à tromper et diffusées dans le but de nuire. Elles peuvent être véhiculées par des acteurs étatiques ou non étatiques, dans différents contextes tels que des conflits armés, et peuvent toucher tous les domaines du développement, de la paix et la sécurité aux droits humains, en passant par la santé publique, l'aide humanitaire et l'action climatique.
5. La mésinformation correspond à la diffusion non intentionnelle d'informations inexacts mais partagées de bonne foi par des personnes qui ne sont pas conscientes de la nature de ces informations. La mésinformation peut prendre racine dans la désinformation, dans la mesure où les mensonges délibérés et les discours trompeurs sont utilisés comme des armes à mesure que le temps passe, intégrés aux discours publics et transmis de façon involontaire³. Dans la pratique, il est parfois difficile de distinguer la mésinformation de la désinformation.
6. Selon la Stratégie et le Plan d'action des Nations Unies pour la lutte contre les discours de haine, le discours de haine couvre tout type de communication par la parole, l'écrit ou le comportement qui attaque quiconque ou utilise un langage péjoratif ou discriminatoire à l'endroit d'une personne ou d'un groupe en raison de leur identité, en d'autres termes, en raison de l'appartenance religieuse, de l'origine ethnique, de la nationalité, de la race, de la couleur de peau, de l'ascendance, du genre ou de tout autre élément constitutif de l'identité⁴.
7. La mésinformation, la désinformation et les discours haineux sont liés à des phénomènes distincts : leur mise au jour, leur atténuation et leur résolution

¹ A/HRC/47/25, par. 15.

² Kalina Bontcheva et Julie Posetti, éd., *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression – Broadband Commission Research Report on Freedom of Expression and Addressing Disinformation on the Internet*, Genève (Union internationale des télécommunications, UIT), Paris (UNESCO), 2020.

³ Voir Nations Unies, *Combattre la désinformation*, disponible à l'adresse <https://www.un.org/fr/countering-disinformation>, et A/77/287.

⁴ Disponible à l'adresse https://www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_FR.pdf.

comportent des différences et des points communs. Tous trois polluent l'écosystème de l'information et menacent les progrès humains⁵.

8. Les menaces visant l'intégrité de l'information sont loin d'être nouvelles : les mensonges et la haine ont été diffusés depuis longtemps pour obtenir des avantages politiques ou financiers. Cependant, à l'ère numérique ces opérations peuvent être menées à une échelle jusqu'alors inimaginable. Des outils de communication puissants peuvent désormais répandre des informations instantanément dans le monde entier, ce qui crée un problème si important que même les plateformes en ligne sont parfois incapables d'en prendre la pleine mesure. L'absence de définitions faisant autorité au niveau gouvernemental ne devrait pas créer la paralysie. Nous devons tout mettre en œuvre pour pallier les maux causés par ces phénomènes.

II. Intégrité de l'information et plateformes numériques

9. Les plateformes numériques devraient être des acteurs à part entière des activités visant à préserver l'intégrité de l'information. Certains médias traditionnels peuvent être des sources de mésinformation et de désinformation, mais la rapidité, le volume et la viralité de ces contenus sur les canaux numériques nécessite une réponse immédiate et adaptée. Aux fins de la présente note d'orientation, l'expression « plateforme numérique » renvoie à un service numérique facilitant les échanges entre deux utilisateurs ou plus et couvrant une large gamme d'activités, des médias sociaux et moteurs de recherche aux applications de messagerie. Ces plateformes collectent habituellement des données sur leurs utilisateurs et les interactions de ceux-ci⁶.

10. La mésinformation et la désinformation sont créées par un vaste éventail d'acteurs aux motivations multiples, qui restent en grande partie anonymes. Les campagnes de désinformation coordonnées par des acteurs étatiques et non étatiques ont exploité les failles de systèmes numériques pour promouvoir des discours toxiques, ce qui a eu des répercussions graves.

11. De nombreux États ont lancé des initiatives visant à réglementer les plateformes numériques, 70 lois de ce type au moins ayant été adoptées ou envisagées au cours des quatre dernières années⁷. De par leur nature même, ces approches législatives comprennent généralement un ensemble restreint de solutions visant à définir et à supprimer les contenus toxiques. En axant les mesures sur la suppression du contenu toxique, certains États ont passé des lois inadéquates ou trop larges qui ont eu pour effet de réduire au silence des formes protégées d'expression autorisées par le droit international. D'autres mesures, comme les coupures généralisées d'Internet ou les bannissements de plateformes, peuvent être dénuées de fondement juridique ou porter atteinte aux droits humains.

12. De nombreux États et personnalités politiques ont prétendu s'inquiéter de l'intégrité de l'information et ont utilisé ces préoccupations comme prétexte pour restreindre l'accès à l'information, discréditer et restreindre l'établissements de rapports, et pour cibler les journalistes et les opposants⁸. Les acteurs étatiques ont

⁵ L'Organisation des Nations Unies mène actuellement une étude sur les liens entre la mésinformation, la désinformation et les discours haineux et les points de convergence et de divergence aux niveaux conceptuel et opérationnel entre ces phénomènes distincts mais connexes.

⁶ La Commission européenne définit les plateformes en ligne sur sa page « Façonner l'avenir numérique de l'Europe : plateformes en ligne », 7 juin 2022, disponible à l'adresse <https://digital-strategy.ec.europa.eu/fr/policies/online-platforms-0>.

⁷ Voir HCDH, *Moderating online content: fighting harm or silencing dissent?*, 23 juillet 2021.

⁸ Voir Nations Unies, *Combattre la désinformation*, et A/77/287.

aussi fait pression sur les plateformes pour qu'elles cèdent à leurs exigences, sous le couvert de la lutte contre la mésinformation et la désinformation⁹. Des spécialistes de la liberté d'expression ont souligné que les acteurs étatiques avaient un devoir particulier dans un tel contexte et ne devaient pas produire, cautionner, encourager ou disséminer de fausses informations (A/77/287, par. 45).

13. Toute volonté de réglementer l'expression, même lorsqu'elle répond à un objectif légitime d'intérêt public, présente des risques intrinsèques qui obligent à adopter une approche soigneusement adaptée, qui respecte les exigences de légalité, de nécessité et de proportionnalité fixées par le droit des droits de l'homme (ibid., par. 42).

14. La désinformation constitue aussi un marché juteux. Qu'elles aient pignon sur rue ou non, les sociétés de relations publiques auxquels font appel les États, les personnalités politiques et le secteur privé sont des sources majeures d'informations fausses et trompeuses¹⁰. L'une des tactiques utilisées consiste à publier des contenus sur des sites Web copiant des sites d'information officiels, pour faire croire que ces fausses informations sont en fait véridiques¹¹. Il est très difficile de tracer ces activités obscures et d'enquêter à leur sujet, et donc d'estimer l'ampleur exacte du problème. Les personnes répandent elles aussi des informations fausses pour vendre des produits ou des services et ciblent souvent des groupes vulnérables en période de crise ou d'insécurité.

15. La plupart des plateformes numériques actuelles basent leur modèle économique sur l'économie de l'attention. Des algorithmes sont créés pour donner la priorité à certains contenus qui vont captiver les utilisateurs, ce qui optimise leur utilisation de la plateforme et donc les recettes tirées de la publicité. Des contenus faux et haineux conçus pour polariser les utilisateurs et susciter des émotions fortes sont souvent ceux qui créent le plus d'intérêt, ce qui veut dire que les algorithmes récompensent et amplifient la mésinformation, la désinformation et les discours de haine¹².

16. Face à une baisse des recettes générées par les publicités, les plateformes numériques cherchent d'autres sources de bénéfices au-delà de l'économie de l'attention. Ainsi, les programmes d'authentification payants, grâce auxquels les utilisateurs peuvent acheter un badge jusqu'alors utilisé pour confirmer l'authenticité d'un compte, ont suscité de graves préoccupations en matière d'intégrité de l'information, compte tenu des possibilités d'abus par des acteurs de la désinformation¹³.

⁹ A/HRC/47/25.

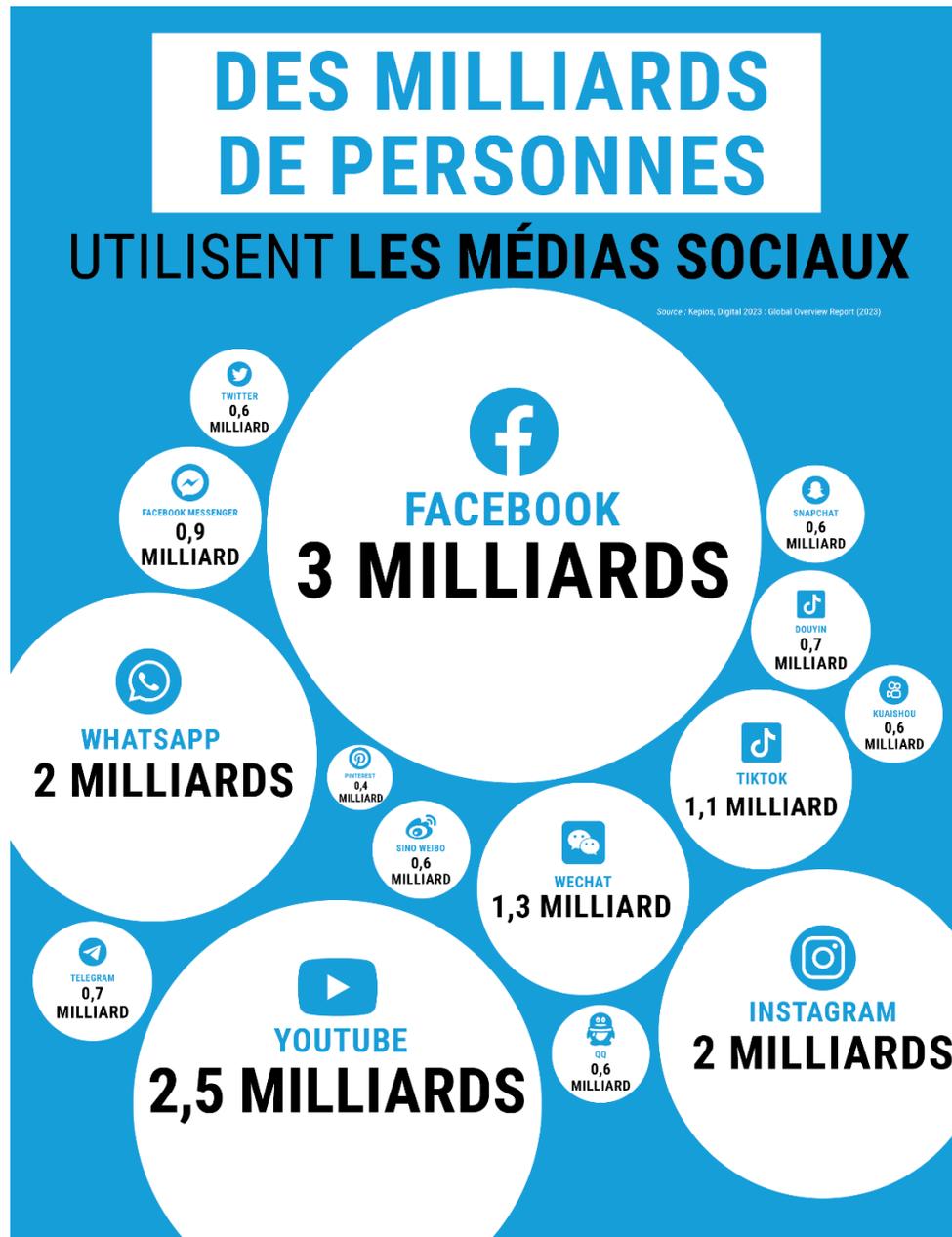
¹⁰ Stephanie Kirchgaessner et al., *Revealed: the hacking and disinformation team meddling in elections*, The Guardian, 14 février 2023.

¹¹ Alexandre Alaphilippe et al., *Doppelgänger – media clones serving Russian propaganda*, EU DisinfoLab, 27 septembre 2022.

¹² Réseau des économistes des Nations Unies, *New economics for sustainable development: attention economy*.

¹³ Twitter, *À propos de Twitter Blue ; Meta, Testing Meta Verified to help creators establish their presence*, 17 mars 2023.

Figure I
Des milliards de personnes utilisent les médias sociaux



Source : Kepios, Digital 2023 : Global Overview Report (2023).

III. Quel est le cadre juridique applicable au niveau international ?

17. La promotion de l'intégrité de l'information doit être pleinement ancrée dans les normes et règles internationales applicables, notamment le droit des droits humains et les principes de la souveraineté et de la non-intervention dans les affaires intérieures d'autres États. En août 2022, j'ai transmis à l'Assemblée générale un rapport intitulé « Combattre la désinformation pour promouvoir et protéger les droits

humains et les libertés fondamentales »¹⁴, dans lequel j'ai dressé la liste des principes de droit international des droits humains applicables à la désinformation, dont la Déclaration universelle des droits de l'homme et le Pacte international relatif aux droits civils et politiques. Selon ces instruments juridiques internationaux, toute personne a droit à la liberté d'expression¹⁵.

18. L'article 19 de la Déclaration et le paragraphe 2 de l'article 19 du Pacte protègent aussi le droit à la liberté d'expression, notamment le droit de chercher, de recevoir et de répandre, sans considérations de frontières, des informations et des idées par quelque moyen d'expression que ce soit. Le droit des personnes à la liberté d'expression ne se limite pas aux informations accueillies avec faveur (A/77/287, par. 13). Liée à la liberté d'expression, la liberté d'information est un droit en soi. L'Assemblée générale a déclaré que « la liberté de l'information [était] un droit fondamental de l'homme et la pierre de touche de toutes les libertés à la défense desquelles se consacr[aient] les Nations Unies » (ibid., par. 14). La liberté d'expression et l'accès à l'information peuvent être soumis à certaines restrictions répondant à des critères précis, décrits au paragraphe 3 de l'article 19 du Pacte¹⁶. Les États ne peuvent pas ajouter des motifs supplémentaires ou restreindre l'expression au-delà de ce qui est permis par le droit international.

19. Adopté en 2012, le Plan d'action de Rabat sur l'interdiction de l'appel à la haine nationale, raciale ou religieuse qui constitue une incitation à la discrimination, à l'hostilité ou à la violence donne des orientations générales pratiques et juridiques aux États sur la meilleure façon d'appliquer le paragraphe 2 de l'article 20 du Pacte et l'article 4 de la Convention internationale sur l'élimination de toutes les formes de discrimination raciale, qui interdit certaines formes de discours haineux. Le Plan d'action de Rabat a déjà été utilisé par des États Membres dans différents contextes¹⁷.

20. Les discours de haine ont été annonciateurs des crimes atroces qui les ont suivis, génocides compris. La Convention de 1948 pour la prévention et la répression du crime de génocide interdit « l'incitation directe et publique à commettre le génocide ».

21. Dans sa résolution 76/227 adoptée en 2021, l'Assemblée générale souligne que la désinformation, sous toutes ses formes, peut nuire à l'exercice des droits humains et des libertés fondamentales, ainsi qu'à la réalisation des objectifs de développement durable. De même, dans sa résolution 49/21 adoptée en 2022, le Conseil des droits de l'homme a affirmé que la désinformation peut avoir des effets négatifs sur l'exercice et la réalisation de tous les droits de l'homme.

¹⁴ A/77/287.

¹⁵ En février 2023, 173 États Membres étaient parties au Pacte international relatif aux droits civils et politiques.

¹⁶ Les limites imposées à la liberté d'expression doivent respecter les conditions suivantes établies de longue date : légalité (les restrictions doivent être imposées par une loi, de façon à distinguer de façon précise les formes d'expression légales et illégales), nécessité et proportionnalité (il peut être démontré que cette mesure est celle qui restreint le moins l'exercice du droit et qu'elle protège effectivement, ou est susceptible de protéger, l'intérêt légitime de l'État en jeu), et légitimité (pour être légales, ces restrictions doivent protéger uniquement les intérêts énumérés au paragraphe 3 de l'article 19 du Pacte international relatif aux droits civils et politiques).

¹⁷ Il s'agit entre autres de communications audiovisuelles, en Côte d'Ivoire, au Maroc et en Tunisie et du suivi des incitations à la violence par la Mission multidimensionnelle intégrée des Nations Unies pour la stabilisation en République centrafricaine.

IV. Quels sont les préjudices causés par la mésinformation, la désinformation et les discours de haine diffusés en ligne ?

22. La diffusion en ligne de contenus de mésinformation, de désinformation et de discours de haine crée de graves problèmes pour le grand public. Une étude basée sur les résultats d'une enquête menée dans 142 pays a établi que 58,5 % des utilisateurs d'Internet et des médias sociaux dans le monde sont préoccupés à l'idée d'être victimes de mésinformation en ligne, les jeunes et les personnes ayant le moins de ressources étant les personnes chez qui le sentiment de vulnérabilité était le plus prononcé¹⁸. Les jeunes d'aujourd'hui sont des enfants du numérique qui sont plus susceptibles de passer du temps en ligne que le reste de la population, ce qui fait d'eux la génération la plus connectée de l'histoire de l'humanité¹⁹. Partout dans le monde, un enfant se connecte à Internet pour la première fois de sa vie toutes les demi-secondes, ce qui crée un risque d'exposition à des discours haineux et à des dangers, et peut affecter sa santé mentale dans certains cas²⁰.

23. Les répercussions de la mésinformation, de la désinformation et des discours de haine en ligne se manifestent dans le monde entier, et ce même dans les domaines de la santé, de l'action climatique, de la démocratie et des élections, de l'égalité des genres, de la sécurité et de l'action humanitaire. Dans une enquête menée en 2021 dans les bureaux de pays du Programme des Nations Unies pour le développement, la pollution de l'information a été qualifiée de motif de préoccupation majeur par 75 % des participants et participantes. Elle a une influence énorme sur la confiance, la sécurité, la démocratie et le développement durable, selon une enquête récente menée à la demande de l'UNESCO dans plus de 800 documents publiés dans le monde universitaire, la société civile, les journaux et les entreprises²¹.

24. La mésinformation et la désinformation peuvent être dangereuses, voire mortelles, en particulier en période de crise, d'urgence ou de conflit. Pendant la pandémie de maladie à coronavirus (COVID-19), une avalanche de fausses informations sur le virus, les mesures de santé publique et les vaccins s'est abattue sur Internet²². Certains acteurs ont exploité la confusion à des fins personnelles : des militants antivaccins ont par exemple envoyé des utilisateurs vers des sites vendant de faux remèdes ou faisant la promotion de fausses mesures préventives²³. De nombreuses victimes de la COVID-19 ont refusé de se faire vacciner ou de prendre des précautions sanitaires de base après avoir été exposées à la mésinformation et à la désinformation en ligne²⁴.

25. La désinformation peut aussi s'avérer mortelle dans des contextes sociaux et politiques déjà instables. Dans un rapport datant de 2022, la Rapporteuse spéciale sur la promotion et la protection du droit à la liberté d'opinion et d'expression a examiné

¹⁸ Aleksi Knuutila, Lisa-Maria Neudert et Philip N. Howard, *Who is afraid of fake news? Modeling risk perceptions of misinformation in 142 countries*, Harvard Kennedy School (HKS) Misinformation Review, vol. 3, n° 3 (avril 2022).

¹⁹ UIT, *Mesurer la société de l'information* (Genève, 2013).

²⁰ Fonds des Nations Unies pour l'enfance (UNICEF), *Protecting children online*, 23 juin 2022, disponible à l'adresse www.unicef.org/protection/violence-against-children-online.

²¹ UNESCO, documents de travail sur la gouvernance numérique et les problèmes qu'elle pose en matière de confiance et de sécurité. Disponibles à l'adresse www.unesco.org/en/internet-conference/working-papers.

²² Voir Julie Posetti et Kalina Bontcheva, *Désinfodémie : déchiffrer la désinformation sur le COVID-19, note d'orientation 1* (Paris, UNESCO, 2020), et *Désinfodémie : dissection des réponses à la désinformation sur le COVID-19, note d'orientation 2* (Paris, UNESCO, 2020).

²³ Center for Countering Digital Hate, *Pandemic Profiteers: The Business of Anti-Vaxx* (2021).

²⁴ Michael A Gisondi et al., *A deadly infodemic: social media and the power of COVID-19 misinformation*, Journal of Medical Internet Research, vol. 24, n°2 (février 2022).

les incidences de l'utilisation de l'information comme une arme qui sert à semer la confusion, à nourrir la haine, à inciter à la violence et à prolonger les conflits²⁵. Dans un autre rapport publié en 2022, il a été établi que la désinformation pouvait prendre la forme d'un discours sectaire et haineux visant les minorités, les femmes et toute personne perçue comme « autre », ce qui constituait une menace non seulement pour les personnes directement visées, mais aussi pour l'inclusion et la cohésion sociale. Elle pouvait aggraver les tensions et les divisions en période d'urgence ou de crise, lors de moments politiques clés ou de conflits armés²⁶.

26. Certaines des pires incidences des préjudices en ligne se situent dans des contextes négligés par les plateformes numériques, même dans les lieux où elles bénéficient d'un taux de pénétration élevé. Les pays en proie à des conflits ou à des contextes instables, qui constituent souvent des marchés moins lucratifs, n'ont pas bénéficié de ressources suffisantes pour la modération des contenus ou l'assistance aux utilisateurs. Les médias traditionnels constituent toujours une source d'information importante pour la plupart des habitants des zones de conflit, mais la haine diffusée sur les plateformes numériques a aussi encouragé et alimenté la violence²⁷. Certaines plateformes numériques ont été critiquées pour leur rôle dans les conflits, notamment la guerre qui se déroule actuellement en Ukraine²⁸.

27. De même, la mésinformation et la désinformation sur l'urgence climatique retardent l'adoption de mesures urgentes visant à rétablir la perspective d'un avenir viable sur la planète. La mésinformation et la désinformation climatiques peuvent être considérées comme des contenus faux ou trompeurs qui remettent en cause les fondements scientifiques de l'existence des changements climatiques dus aux activités humaines, de leurs causes et de leurs conséquences. Des campagnes coordonnées cherchent à nier, minimiser ou faire oublier le consensus scientifique recueilli par le Groupe d'experts intergouvernemental sur l'évolution du climat et à torpiller les mesures urgentes devant être prises pour atteindre les objectifs de l'Accord de Paris de 2015. Une minorité, petite mais bruyante, de climatosceptiques²⁹ continue de rejeter ce consensus et d'occuper une place démesurée sur certaines plateformes numériques. En 2022 par exemple, des simulations aléatoires réalisées par des organisations de la société civile ont révélé que l'algorithme de Facebook recommandait des contenus climatosceptiques au détriment de contenus en lien avec la climatologie³⁰. Sur Twitter, l'utilisation du mot-dièse #climatescam (arnaque climatique) est passée de moins de 2 700 par mois au premier semestre 2022 à 80 000 en juillet et à 199 000 en janvier 2023. Cette expression a aussi été présentée par la

²⁵ A/77/288.

²⁶ A/77/287, par. 6.

²⁷ En 2018, une mission internationale indépendante d'établissement des faits nommée par le Conseil des droits de l'homme a déclaré que Facebook était la plateforme qui accueillait le plus grand nombre de discours haineux au Myanmar (A/HRC/42/50, par. 72).

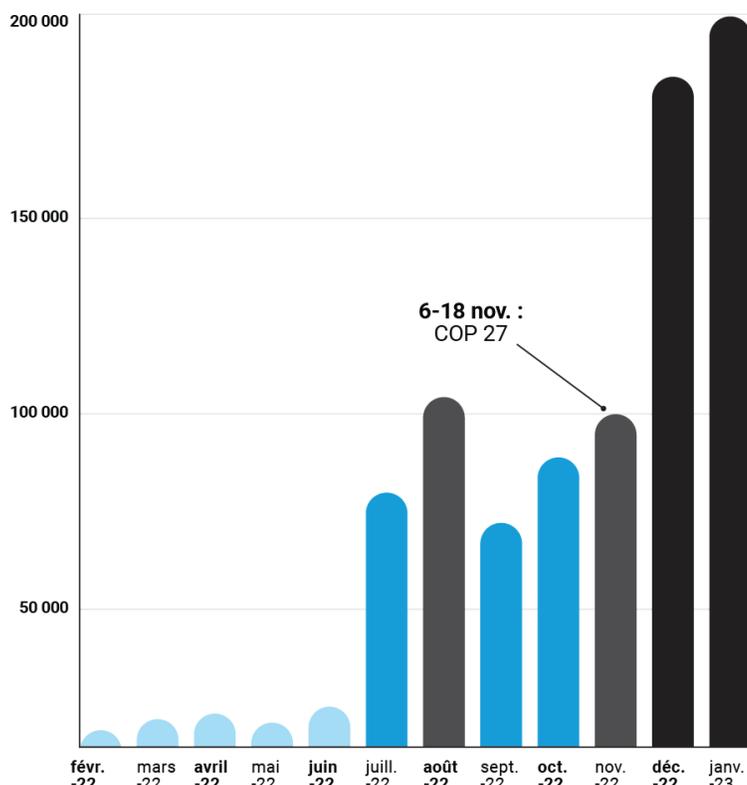
²⁸ Voir ONU Info, *Les discours de haine : une menace internationale grandissante*, 28 janvier 2023, et *Digital technology, social media fuelling hate speech like never before, warns UN expert*, 20 octobre 2022.

²⁹ Voir John Cook, *Deconstructing climate science denial*, dans *Research Handbook in Communicating Climate Change*, David C. Holmes et Lucy M. Richardson, éd., Cheltenham (Royaume-Uni), Edward Elgar, 2020. Cook a indiqué qu'Abraham et al. (2014) avaient résumé la manière dont les articles contenant des discours climatosceptiques, tels que des affirmations de refroidissement dans les mesures satellitaires ou des estimations de faible sensibilité du climat, ont été réfutés avec autorité dans la littérature scientifique. De même, Benestad et al. (2016) ont tenté de reproduire les conclusions issues d'articles contradictoires et détecté un certain nombre d'erreurs, telles que des méthodes statistiques inappropriées, de fausses dichotomies et des conclusions basées sur des éléments de physique erronés.

³⁰ Global Witness, *The climate divide: how Facebook's algorithm amplifies climate disinformation*, 28 mars 2022.

plateforme parmi les premiers résultats des recherches utilisant le mot « climat »³¹. En février 2022, le Groupe d'experts intergouvernemental sur l'évolution du climat a dénoncé pour la première fois la désinformation climatique, en déclarant qu'un travail de sape délibéré de la science contribuait à des perceptions erronées du consensus scientifique, créait l'incertitude, faisait fi des risques et de l'urgence, et attisait la dissidence³².

Figure II
Nombre d'utilisations du mot-dièse #climatescam chaque mois sur Twitter



Source : Département de la communication globale à partir de données communiquées par Talkwalker.

28. Certaines entreprises du secteur des combustibles fossiles déploient souvent des stratégies d'écoblanchiment, qui consistent à faire croire au public qu'une entreprise ou une entité en fait plus pour protéger l'environnement que pour lui nuire, alors que la réalité est tout autre. Les entreprises n'agissent pas seules. Les efforts visant à semer la confusion dans l'esprit du public et à détourner l'attention de la responsabilité du secteur des combustibles fossiles sont facilités et soutenus par les agences de publicité et de relations publiques, les entreprises de technologie

³¹ Analyse du Département de la communication globale à partir de données communiquées par Talkwalker.

³² Jeffrey A. Hicke et al., Groupe d'experts intergouvernemental sur l'évolution du climat, *Climate Change 2022: Impacts, Adaptation and Vulnerability – Contribution du Groupe de travail II au sixième rapport d'évaluation du Groupe d'experts intergouvernemental sur l'évolution du climat* (Cambridge, Royaume-Uni, Cambridge University Press, 2022).

publicitaire, les organes de presse et les plateformes numériques³³. Dans l'ensemble, les agences de publicité et de relations publiques qui créent des contenus d'écoblanchiment et les tiers qui les distribuent gagnent des milliards grâce à ces efforts visant à éviter à l'industrie des combustibles fossiles d'être surveillée de trop près et de rendre des comptes. Les sociétés de relations publiques ont mené des centaines de campagnes pour des entreprises du secteur du charbon, du pétrole et du gaz³⁴.

29. En affaiblissant la confiance dans les institutions démocratiques et les médias indépendants et en réduisant la participation aux affaires politiques et publiques, la mésinformation et la désinformation ont des effets considérables sur la démocratie. Tout au long du cycle électoral, l'exposition à des informations fausses et trompeuses peut priver les électeurs de la possibilité de faire des choix éclairés. La diffusion d'informations fausses et trompeuses peut miner la confiance du public dans les institutions et les opérations électorales (comme l'inscription sur les listes électorales, le scrutin et les résultats) et potentiellement entraîner l'apathie des électeurs ou le rejet de résultats électoraux crédibles. Les États et les dirigeants politiques se sont avérés de puissantes sources de désinformation, qui diffusent délibérément et stratégiquement des informations erronées pour conserver ou assurer leur pouvoir, ou pour saper les processus démocratiques dans d'autres pays³⁵.

30. Les groupes marginalisés et vulnérables sont également la cible fréquente de la mésinformation, de la désinformation et des discours haineux, ce qui aggrave leur exclusion sociale, économique et politique. Les candidates, les électrices, les responsables électorales, les journalistes et les représentantes de la société civile sont les cibles de la désinformation générée en ligne³⁶. Ces attaques compromettent la participation politique et affaiblissent les institutions démocratiques et les droits humains, y compris la liberté d'expression et l'accès de ces groupes à l'information³⁷. Il s'agit d'une priorité de plus en plus urgente pour la communauté internationale, notamment parce que plus de 2 milliards d'électeurs et d'électrices sont appelés à se rendre aux urnes partout dans le monde en 2024.

31. La mésinformation et la désinformation contribuent aussi à une contamination croisée entre les plateformes et les médias traditionnels, ce qui complique leur suivi

³³ Mei Li, Gregory Trencher et Jusen Asuka, *The clean energy claims of BP, Chevron, ExxonMobil and Shell: a mismatch between discourse, actions and investments*, PLOS ONE, volume 17, n°2 (février 2022).

³⁴ Robert J. Brulle et Carter Werthman, *The role of public relations firms in climate change politics*, Climatic Change, vol. 169, n°1-2 (novembre 2021). Selon Global Disinformation Index, un organisme de surveillance à but non lucratif, les annonceurs de l'industrie des technologies ont versé 36,7 millions de dollars à 98 sites Web diffusant de la désinformation climatique en anglais en 2021. Un rapport de novembre 2022 du Center for Countering Digital Hate, une association de lutte, a révélé que, rien que sur Google, près de la moitié des 23,7 millions de dollars dépensés en publicités publiées sur les pages de résultats de recherche par les sociétés des secteurs du pétrole et du gaz au cours des deux dernières années ciblaient des mots clés liés à la durabilité environnementale. Dans le cadre de recherches menées par InfluenceMap, on a trouvé 25 147 publicités trompeuses émanant de 25 organisations des secteurs du pétrole et du gaz, qui avaient été publiées sur Facebook aux États-Unis d'Amérique en 2020 et représentaient un montant total de 9 597 376 dollars. Jusqu'à présent, la réponse a été sans commune mesure avec l'ampleur du problème.

³⁵ Voir la résolution 76/227 de l'Assemblée générale, la résolution 49/21 du Conseil des droits de l'homme et Service européen d'action extérieure de l'Union européenne, *Tackling disinformation, foreign information manipulation and interference*, 27 octobre 2021.

³⁶ Lucina Di Meco, *Monetizing misogyny: gendered disinformation and the undermining of women's rights and democracy globally*, #ShePersisted, février 2023.

³⁷ Voir Andrew Puddephatt, *Social media and elections*, Cuadernos de Discusión de Comunicación e Información, n° 14 (Montevideo, UNESCO, 2019), et Julie Posetti et al., *The Chilling: Global trends in online violence against women journalists* (UNESCO, 2021).

et leur traitement si elles ne sont pas repérées à la source. La désinformation peut être une tactique délibérée des médias influencés par des idéologies et récupérés par des intérêts politiques et commerciaux³⁸. Dans le même temps, l'essor des plateformes numériques a précipité l'effondrement des médias indépendants et dignes de confiance. Les lecteurs des journaux et les recettes publicitaires ont migré en masse vers les plateformes en ligne, tendance exacerbée par la pandémie de COVID-19. Dans certaines régions ou pays³⁹, on constate une « extinction des médias » ou des « déserts de l'information », dans la mesure où les populations perdent des sources d'information locales dignes de confiance, ce qui contribue à la pollution de l'écosystème de l'information. Le « newswashing », qui consiste à déguiser un contenu sponsorisé pour qu'il ressemble à un reportage, est souvent mal signalé lorsqu'il est publié sur des plateformes numériques, ce qui lui confère un vernis de légitimité. Une fois ces informations reprises par d'autres médias, citées par des hommes politiques ou largement partagées sur des plateformes, leur origine devient de plus en plus obscure et les consommateurs et consommatrices d'informations sont incapables de les distinguer d'un fait réel.

32. La désinformation a aussi des répercussions sur les activités des Nations Unies. Les coordinateurs résidents, les envoyés, les médiateurs et les soldats de la paix se sont inquiétés des effets de la désinformation sur la sécurité opérationnelle, l'efficacité et la capacité d'action du système. Dans une enquête réalisée en 2022, 70 % des soldats de la paix ont déclaré que la mésinformation et la désinformation avaient des conséquences graves, critiques ou modérées sur leur travail, tandis que 75 % ont déclaré qu'elles avaient une incidence sur leur sûreté et leur sécurité. La mésinformation et la désinformation peuvent aussi être utilisées pour cibler les intervenants humanitaires et entraver les opérations de sauvetage dans les zones de conflit.

V. Intégrité de l'information et objectifs de développement durable

33. Comme le montre la figure III, les menaces qui pèsent sur l'intégrité de l'information peuvent nuire à la réalisation des objectifs de développement durable.

³⁸ EU DisinfoLab, *The role of "media" in producing and spreading disinformation campaigns*, 13 octobre 2021.

³⁹ Voir ONU Info, *Le modèle économique des médias d'information n'est plus viable, le droit à l'information est menacé* (UNESCO), 10 mars 2022, et Anya Schiffrin et al., *Trouver des financements pour un journalisme florissant : viabilité des médias, les options stratégiques, Tendances mondiales en matière de liberté d'expression et de développement des médias* (Paris, UNESCO, 2022).

Figure III
Intégrité de l'information et objectifs de développement durable

	<p>La désinformation continue d'avoir une incidence sur les efforts déployés pour éliminer la pauvreté et sur l'économie mondiale. Les difficultés économiques peuvent aussi alimenter la diffusion de mensonges haineux et polarisants, qui concernent entre autres les groupes marginalisés. La flambée du coût de la vie a été un terreau particulièrement fertile pour la diffusion d'informations accusant faussement le passage aux énergies renouvelables d'être à l'origine de l'envolée des coûts de l'énergie ou des pertes d'emplois, par exemple.</p>		<p>La désinformation et les discours haineux diffusés en ligne polarisent les sociétés et ciblent des populations déjà marginalisées et vulnérables, ce qui peut aggraver leur exclusion sociale, économique et politique.</p>
	<p>Les menaces qui pèsent sur l'intégrité de l'information peuvent aggraver la faim dans le monde, notamment en exacerbant les conflits, les changements climatiques, les catastrophes, la pauvreté et les inégalités. La désinformation peut détourner l'attention et faire oublier les défis que les conflits créent en matière de sécurité alimentaire mondiale.</p>		<p>Les efforts visant à rendre les villes et les communautés plus durables peuvent être sapés par la désinformation qui nie ou fait oublier les répercussions de l'activité humaine sur l'environnement.</p>
	<p>Pendant la pandémie de maladie à coronavirus (COVID-19), une infodémie de fausses informations et de désinformation a sapé les mesures de santé publique et les campagnes de vaccination. La menace qui pèse sur la santé et le bien-être des enfants du fait de l'exposition à des contenus préjudiciables persiste.</p>		<p>Les militants à l'origine d'initiatives visant à favoriser l'économie circulaire et à encourager les pratiques zéro déchet ont été la cible de discours haineux et de désinformation en ligne.</p>
	<p>La désinformation et les discours haineux peuvent avoir des répercussions négatives sur l'accès à un enseignement de qualité, en particulier pour les groupes marginalisés, notamment les jeunes femmes et les filles. Les efforts menés en faveur de l'accès à l'information et de l'initiation aux médias numériques afin d'accroître la résilience joueront un rôle clé dans la limitation de l'incidence des dangers en ligne sur la société.</p>		<p>La désinformation climatique et l'inertie qu'elle encourage sapent les efforts visant à prendre sans tarder des mesures pour faire face à la crise climatique, notamment en empêchant la transition cruciale des combustibles fossiles polluants aux énergies renouvelables et les investissements urgents dans la résilience aux changements climatiques.</p>
	<p>Les discours haineux et la désinformation fondés sur le genre visent à soumettre systématiquement les femmes en les réduisant au silence et en les excluant de la sphère publique. Ils peuvent avoir des conséquences dévastatrices, allant de l'étouffement de la voix des femmes à l'autocensure, en passant par des atteintes aux perspectives de carrière et à la réputation, et par l'incitation à la violence physique.</p>		<p>La désinformation peut nuire à l'action menée sur le plan international pour conserver et exploiter de manière durable les océans, les mers et les ressources marines.</p>
	<p>Deux milliards de personnes vivent sans accès à des services d'eau potable gérés de façon sûre. La désinformation liée à la sécurité de l'eau potable et de l'assainissement peut avoir des conséquences dangereuses pour la santé^a.</p>		<p>Les défenseurs de l'environnement qui œuvrent à la protection de la vie sur terre ont été la cible de campagnes en ligne d'incitation à la haine et de désinformation, qui ont eu des conséquences concrètes. Les fausses informations et la désinformation sur le climat sapent l'action climatique^c.</p>
	<p>La désinformation sur le climat, alimentée en grande partie par le secteur des combustibles fossiles, nuit à la transition urgente vers des formes plus propres de production d'énergie, réduisant ainsi la possibilité d'offrir un avenir durable à tous et à toutes.</p>		<p>La désinformation et les discours haineux ont été utilisés pour influencer les élections et les discours publics et semer la confusion. On y a eu recours pour affaiblir des adversaires, contrecarrer des efforts de rétablissement de la paix, inciter à la violence, prolonger des conflits et ébranler la confiance dans l'état de droit. Les efforts visant à promouvoir des sociétés pacifiques et inclusives ainsi que le rôle de l'Organisation des Nations Unies pour ce qui est de soutenir la paix et la sécurité ont été particulièrement entravés^d.</p>
	<p>Des recherches ont mis en évidence les effets néfastes de la désinformation et des discours haineux sur l'économie^b.</p>		<p>La désinformation et les discours haineux peuvent entraver la mise en place de partenariats efficaces aux fins de la réalisation des objectifs, tandis que les ressources détournées pour traiter le problème peuvent affaiblir les activités menées pour ne laisser personne de côté.</p>
	<p>La désinformation et les discours haineux, ainsi que les réponses trop larges à ces phénomènes, peuvent influencer négativement l'innovation, notamment en limitant le potentiel des groupes marginalisés et en rendant les espaces numériques moins égaux et moins inclusifs.</p>		

^a Organisation mondiale de la Santé et Fonds des Nations Unies pour l'enfance (UNICEF), *Progress on Household Drinking Water, Sanitation and Hygiene 2000–2020: Five Years into the SDGs* (Geneva, 2021).

^b Voir Roberto Cavazos et CHEQ, "The economic cost of bad actors on the Internet: fake news, 2019" ; et London Economics, "The cost of lies: assessing the human and financial impact of COVID-19 related online misinformation on the UK", décembre 2020.

^c Global Witness, *Last Line of Defence: The Industries Causing the Climate Crisis and Attacks against Land and Environmental Defenders* (2021).

^d A/77/288.

VI. Comment renforcer l'intégrité de l'information ?

34. La mésinformation, la désinformation et les discours haineux n'existent pas hors de tout contexte : ils se propagent lorsque des personnes se sentent exclues et non entendues, lorsqu'elles se heurtent aux conséquences des disparités économiques et lorsqu'elles sont déçues par la politique. Les réponses doivent tenir compte de ces problèmes concrets. Les efforts déployés pour atteindre les objectifs de développement durable sont essentiels pour construire un monde dans lequel la confiance peut être rétablie.

35. En élaborant des réponses, il est important de ne pas perdre de vue l'énorme valeur que les plateformes numériques apportent au monde. Elles ont révolutionné la communication de masse en temps réel, en permettant la diffusion d'informations vitales lors de catastrophes naturelles et de pandémies. Elles ont contribué à mobiliser un appui en faveur des objectifs que les Nations Unies s'efforcent d'atteindre, et se sont souvent avérées des forces bénéfiques pour l'inclusion et la participation à la vie publique. Elles ont relié des populations dispersées sur le plan géographique et jusqu'alors exclues, des personnes souffrant de maladies rares, et un grand nombre de militants s'efforçant de rendre le monde meilleur.

Mesures réglementaires

36. La question de savoir si les plateformes numériques peuvent et doivent être tenues légalement responsables du contenu qu'elles hébergent a fait l'objet d'un long débat. Dans certains contextes, certaines lois existantes portant sur la diffamation, la cyberintimidation et le harcèlement ont été utilisées de manière efficace pour contrer des menaces ciblant l'intégrité de l'information sans imposer pour autant de nouvelles restrictions à la liberté d'expression ([A/77/287](#), par. 44).

37. De plus, certaines mesures législatives ont été récemment prises pour lutter contre ce problème aux niveaux régional et national. On peut citer, par exemple, le cadre législatif adopté par l'Union européenne en 2022, qui comprend la législation sur les services numériques, l'initiative sur la transparence et le ciblage de la publicité à caractère politique et le Code de bonnes pratiques contre la désinformation. La législation sur les services numériques établit de nouvelles règles pour les utilisateurs et les utilisatrices, les plateformes numériques et les entreprises opérant en ligne dans l'Union européenne. Les mesures visent les contenus, biens et services en ligne illégaux et prévoient un mécanisme permettant aux utilisateurs et utilisatrices de signaler les contenus illégaux et de contester les décisions de modération qui les ciblent. Elles obligent les plateformes numériques à améliorer la transparence (notamment en ce qui concerne l'utilisation et la nature des algorithmes recommandant des contenus), et les grandes plateformes à donner aux chercheurs un accès aux données.

38. Afin de lutter contre la propagation de la désinformation en ligne dans l'Union européenne, le Code de bonnes pratiques contre la désinformation définit des principes et des engagements pour les plateformes en ligne et le secteur de la publicité, que les signataires ont accepté de mettre en œuvre⁴⁰. Il s'agit entre autres d'engagements volontaires visant à démonétiser la désinformation, en empêchant la diffusion de publicités contenant de la désinformation et en évitant de placer des publicités à côté de contenus de désinformation. Les signataires sont aussi convenus d'étiqueter plus clairement les publicités politiques, en communiquant sur les demandeurs, les dépenses publicitaires et la période d'affichage, et de créer des bases

⁴⁰ Commission européenne, *Façonner l'avenir numérique de l'Europe : le Code européen renforcé de bonnes pratiques contre la désinformation*, 4 juillet 2022.

de données consultables sur les publicités politiques. De plus, ils se sont engagés à partager des informations sur les comportements manipulateurs malveillants utilisés pour diffuser de la désinformation (tels que les faux comptes, l'amplification de contenus par des robots, l'usurpation d'identité et les hypertrucages malveillants) détectés sur leurs plateformes, ainsi qu'à mettre régulièrement à jour et à appliquer les politiques visant à les combattre. D'autres engagements visent à donner aux utilisateurs et utilisatrices les moyens de reconnaître, de comprendre et de signaler la désinformation, à renforcer la collaboration avec les vérificateurs de faits et à offrir aux chercheurs et chercheuses un meilleur accès aux données. Ces nouveaux mécanismes seront vraiment mis à l'épreuve lors de leur utilisation.

39. L'un des principaux objectifs du Code de bonnes pratiques contre la désinformation est d'améliorer la transparence des plateformes. En février 2023, les signataires du Code de bonnes pratiques ont publié leurs premiers rapports sur la manière dont ils appliquent leurs engagements. Ces rapports donnent des informations sur le montant de recettes publicitaires non empêchées par les acteurs de la désinformation et d'autres comportements manipulateurs repérés : on peut citer, par exemple, une campagne coordonnée à grande échelle dans plusieurs pays européens pour manipuler l'opinion publique au sujet de la guerre en Ukraine⁴¹.

Réponses apportées par les plateformes numériques

40. Les plateformes numériques varient en taille, fonction et structure et ont adopté un large éventail de mesures pour lutter contre les dommages. Plusieurs grandes plateformes se sont engagées publiquement à respecter les Principes directeurs relatifs aux entreprises et aux droits de l'homme, mais des lacunes persistent en matière de mesures, de transparence et de mise en œuvre⁴². Certaines plateformes n'appliquent pas leurs propres normes et, à des degrés divers, autorisent et amplifient les mensonges et la haine⁴³. Les algorithmes créés pour favoriser le modèle de rentabilité des plateformes sont conçus pour maximiser délibérément l'engagement et monopoliser l'attention, ce qui tend à pousser les utilisateurs vers des contenus polarisants ou provocateurs.

41. La plupart des plateformes numériques disposent d'un système d'autoréglementation, de modération ou de mécanismes de contrôle, mais la transparence en matière de mesures et de pratiques de suppression de contenus continue de poser problème⁴⁴. L'investissement dans ces mécanismes d'une région et d'une langue à l'autre est très disparate et largement concentré dans le monde du Nord, tout comme l'application par les plateformes de leurs propres règles. Selon une enquête récente, les outils de modération et les mécanismes de contrôle n'ont pas été complètement traduits dans les langues locales sur toutes les plateformes⁴⁵. Dans le même temps, la modération est souvent externalisée et manque cruellement de

⁴¹ Voir les remarques de la Vice-Présidente de la Commission européenne chargée des valeurs et de la transparence, Věra Jourová, dans Commission européenne, *Code de bonnes pratiques contre la désinformation : Le nouveau Centre de transparence fournit pour la première fois des informations et des données sur la désinformation en ligne*, communiqué de presse, 9 février 2023. Disponible à l'adresse <https://digital-strategy.ec.europa.eu/fr/news/code-practice-disinformation-new-transparency-centre-provides-insights-and-data-online>.

⁴² Disponible à l'adresse <https://unglobalcompact.org/library/2>.

⁴³ Center for Countering Digital Hate et Human Rights Campaign, *Digital hate: social media's role in amplifying dangerous lies about LGBTQ+ people*, 10 août 2022.

⁴⁴ Voir Andrew Puddephatt, *Laissons entrer le soleil : transparence et responsabilité à l'ère numérique, Tendances mondiales en matière de liberté d'expression et de développement des médias* (Paris, UNESCO, 2021).

⁴⁵ Whose Knowledge?, Oxford Internet Institute et The Centre for Internet and Society, *State of the Internet's Languages Report* (2022).

ressources dans les langues autres que l'anglais⁴⁶. Les témoignages de modérateurs et modératrices ont soulevé des questions troublantes quant aux mauvais traitements, aux normes de travail et aux traumatismes secondaires⁴⁷. Les modérateurs déclarent être constamment exposés à des contenus violents et dérangeants, et n'avoir que quelques secondes pour déterminer si un message signalé est contraire à la politique de l'entreprise. Les systèmes automatisés de modération de contenus peuvent jouer un rôle essentiel, mais les données et les structures utilisées pour les former risquent de créer des biais. On constate aussi que leur taux d'erreur est élevé en anglais et qu'il est encore plus mauvais dans les autres langues. Un certain nombre de plateformes numériques emploient des équipes chargées de la confiance et de la sécurité, des droits humains et de l'intégrité de l'information mais ces experts et expertes sont rarement associés aux premières étapes du développement des produits et leurs postes sont souvent les premiers supprimés lorsqu'il faut faire des économies.

Accès aux données

42. L'accès aux données pour les chercheurs et les chercheuses constitue aussi une priorité immédiate à l'échelle mondiale. La recherche et les ressources existantes restent fortement axées sur les États-Unis d'Amérique et l'Europe. À quelques exceptions près, telles que les rapports sur les opérations de maintien de la paix des Nations Unies en Afrique et la mission internationale indépendante d'établissement des faits sur le Myanmar⁴⁸, et certains rapports d'enquête et reportages réalisés par des journalistes⁴⁹, peu d'études portant sur les répercussions sur le reste du monde ont été publiées. Cela s'explique en partie par le fait que les chercheurs n'ont pas accès aux plateformes et à leurs données. Les outils nécessaires pour mener des études concrètes sur les données limitées communiquées par les plateformes sont généralement conçus à des fins de marketing et leur coût est souvent prohibitif. Le passage des plateformes d'une approche fondée sur un accès à la demande à une approche de divulgation par défaut, assortie des garanties nécessaires en matière de protection de la vie privée, permettrait aux chercheurs et chercheuses d'évaluer correctement les préjudices subis.

Moyens d'action pour les utilisateurs et les utilisatrices

43. Des groupes de la société civile et des universitaires ont mené des recherches approfondies sur la meilleure façon de lutter contre la désinformation, la désinformation et le discours haineux tout en protégeant la liberté d'expression. Un certain nombre d'entre eux ont souligné qu'il fallait trouver des solutions élaborées par la base et donnant davantage de moyen aux internautes, afin de limiter les répercussions des préjudices en ligne subis par leurs propres communautés et de décentraliser le pouvoir monopolisé par les plateformes.

44. Les utilisateurs et utilisatrices des plateformes, y compris les groupes marginalisés, devraient être encouragés à participer et pris en compte dans la prise de décisions. Les jeunes, en particulier, disposent d'une expertise aussi riche que dense. En tant qu'enfants du numérique, les jeunes, et notamment les jeunes femmes, et les enfants sont déjà souvent la cible d'informations erronées, de désinformation, de désinformation et de discours haineux, et seront directement touchés par les nouvelles plateformes. Les jeunes utilisateurs et utilisatrices peuvent s'appuyer sur leur expérience pour juger des répercussions et des défauts potentiels de différentes

⁴⁶ [A/HRC/38/35](#).

⁴⁷ Billy Perrigo, *Inside Facebook's African sweatshop*, Time, 17 février 2022.

⁴⁸ [A/HRC/42/50](#).

⁴⁹ Parmi les exemples notables, on peut citer Maria Ressa, *Résistez aux dictateurs* (Paris, Fayard, 2023) et Max Fischer, *The Chaos Machine* (New York, Little, Brown and Company, 2022).

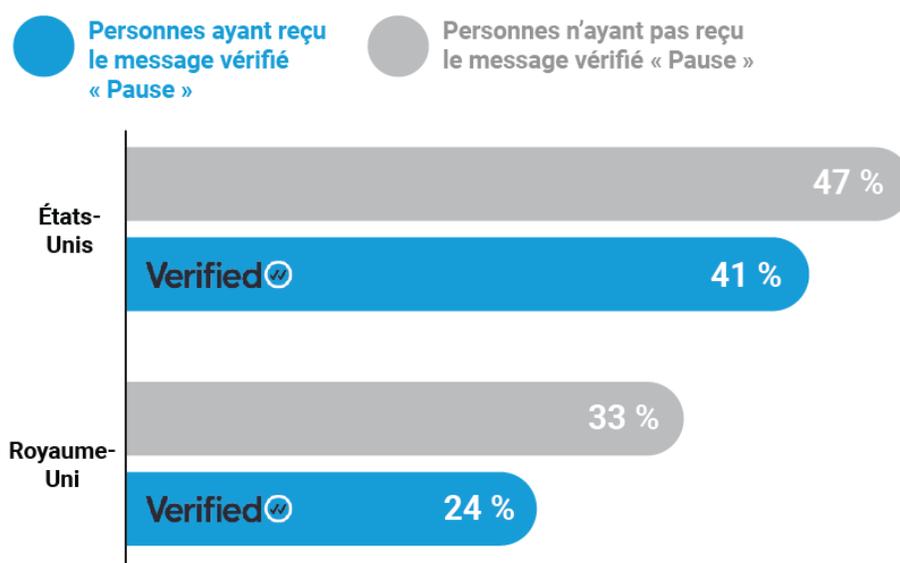
propositions. Ils ont aussi contribué activement aux activités de plaidoyer et de vérification des faits organisées en ligne⁵⁰.

45. L'amélioration de la pensée critique peut rendre les utilisateurs et utilisatrices plus résistants à la manipulation numérique. Plus précisément, l'habileté numérique peut les aider à mieux évaluer les informations trouvées en ligne et à les relayer de manière responsable. Plusieurs entités des Nations Unies ont acquis une expérience précieuse dans ce domaine. Dans le cadre de l'initiative Verified⁵¹ des Nations Unies, une série de tactiques ont été déployées, notamment le ciblage des messages destinés aux utilisateurs, la publication de messages d'information qui avertissent les utilisateurs afin qu'ils soient en mesure de repérer les informations erronées lorsqu'ils les rencontrent, et les campagnes de formation aux outils numériques.

Figure IV

Les campagnes menées par les Nations Unies sont efficaces pour contrer la désinformation et les fausses informations

Probabilité qu'une personne partage des informations fallacieuses (2021)^a



^a Sur la base de recherches menées par le Massachusetts Institute of Technology en mars 2021.

Mesures de dissuasion

46. Dans le modèle économique actuel de la plupart des plateformes numériques, l'engagement passe avant les droits humains, la vie privée et la sécurité, et les données personnelles sont monétisées à des fins lucratives, malgré les éléments de plus en plus nombreux attestant des effets dévastateurs de ce modèle sur les sociétés.

47. Certains groupes de la société civile et chercheurs ont étudié des solutions permettant de démonétiser et donc de décourager la création et la diffusion de contenus de mésinformation et de désinformation et de discours haineux en ligne, en faisant remarquer que si la liberté d'expression est un droit humain fondamental, en tirer profit ne l'est pas⁵². Les propositions visent à lutter contre la rentabilité de la

⁵⁰ Voir UNICEF, *Young reporters fact-checking COVID-19 information*.

⁵¹ Voir <https://shareverified.com/fr/>.

⁵² Global Disinformation Index, une organisation à but non lucratif, suit la publicité associée à la désinformation. Les Nations Unies ont été victimes de cette pratique, Global Disinformation Index

désinformation, à garantir une transparence totale en matière de monétisation des contenus et d'évaluation indépendante des risques, et à dissuader les acteurs de la publicité en ligne de favoriser la désinformation.

48. Les marques qui font paraître leurs publicités aux côtés de contenus de mésinformation ou de désinformation et de discours haineux risquent de réduire l'efficacité de ces campagnes et, à terme, de salir leur réputation. Les publicitaires peuvent élaborer des politiques claires visant à éviter de financer et de légitimer par inadvertance des contenus de mésinformation, de désinformation et des discours haineux et à contribuer à ce que ces contenus ne soient plus rentables. Ces mesures peuvent inclure la gestion de listes d'inclusion et d'exclusion actualisées, et l'utilisation d'outils de vérification des publicités. Les publicitaires peuvent aussi faire pression sur les plateformes numériques pour intensifier les mesures de protection de l'intégrité de l'information et s'abstenir de confier des contenus publicitaires à des organes de presse qui alimentent la haine et entretiennent la désinformation⁵³.

Médias indépendants

49. Dans des dizaines de pays, de nouvelles mesures continuent de porter atteinte à la liberté de la presse. Selon le rapport mondial publié par l'UNESCO pour 2022 dans sa série phare *Tendances mondiales en matière de liberté d'expression et de développement des médias*, 85 % de la population mondiale a connu un recul de la liberté de la presse dans son pays au cours des cinq années précédentes⁵⁴. Dans la mesure où 2,7 milliards de personnes n'ont toujours pas accès à Internet⁵⁵, une autre priorité est de renforcer les médias indépendants, d'intensifier les initiatives de vérification des faits et de soutenir la fiabilité et l'exactitude des informations dans l'intérêt du public. Un véritable débat public s'appuie sur des faits clairement exposés et rapportés de manière éthique et indépendante. Les journalistes respectueux de l'éthique, ayant reçu une formation de qualité et travaillant dans de bonnes conditions, ont les compétences nécessaires pour rétablir l'équilibre face aux informations erronées et à la désinformation. Ils sont en mesure d'offrir un service vital : des informations précises, objectives et fiables portant sur les questions importantes.

Anticiper l'évolution technologique

50. Même si nous cherchons des solutions pour protéger l'intégrité de l'information dans le contexte actuel, nous devons veiller à ce que les recommandations résistent à l'épreuve du temps, et qu'elles tiennent compte des technologies émergentes et de celles à venir. Lancée en novembre 2022, la plateforme ChatGPT-3 d'Open AI comptait 100 millions d'utilisateurs en janvier 2023, ce qui en fait l'application grand public ayant enregistré la croissance la plus rapide de l'histoire⁵⁶, de nombreuses autres entreprises s'étant empressées de développer des outils concurrents. Même s'ils offrent des possibilités quasiment inimaginables de s'attaquer aux problèmes

ayant trouvé des publicités de l'UNICEF placées à côté d'articles antivaccins, et des publicités du Haut-Commissariat des Nations Unies pour les réfugiés placées à côté de contenus anti-réfugiés.

⁵³ *Conscious Advertising Network*, manifestes. Disponibles à l'adresse <https://www.consciousadnetwork.com/the-manifestos/> (en anglais uniquement).

⁵⁴ UNESCO, *Le journalisme est un bien public : tendances mondiales en matière de liberté d'expression et de développement des médias : rapport mondial 2021/2022* (Paris, 2022).

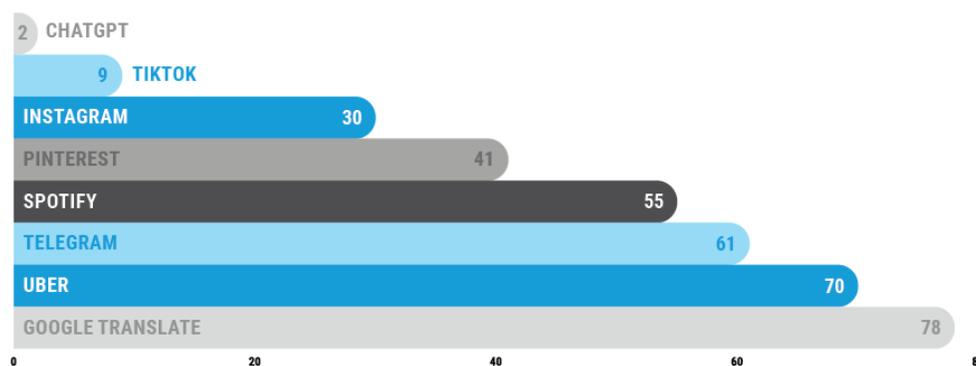
⁵⁵ UIT, *Facts and figures 2021: 2.9 billion people still offline*, 29 novembre 2021. Le Pacte numérique mondial, qui sera adopté par les États membres lors du Sommet de l'avenir qui se tiendra en 2024, définira les principes communs d'un avenir numérique ouvert, libre et sûr pour tout le monde (voir <https://www.un.org/techenvoy/fr/global-digital-compact>).

⁵⁶ Krystal Hu, *ChatGPT sets record for fastest-growing user base - analyst note*, Reuters, 2 février 2023.

mondiaux, les progrès récemment réalisés en matière d'intelligence artificielle (notamment les générateurs d'images et les hypertrucages vidéo) suscitent des inquiétudes sérieuses et immédiates quant à leur potentiel tout aussi formidable de menace pour l'intégrité de l'information. Des rapports et des recherches récemment menés ont montré que les outils d'intelligence artificielle générative créaient des informations erronées, de la mésinformation, de la désinformation et des discours haineux, qui sont ensuite présentés de manière convaincante aux utilisateurs comme s'il s'agissait de faits véridiques⁵⁷.

Figure V

Nombre de mois nécessaires à ChatGPT pour atteindre 100 millions d'utilisateurs actifs par mois, par rapport à d'autres applications populaires



Source : Similarweb, sur la base de données provenant de Sensor Tower.

51. Mon envoyé pour les technologies dirige les activités visant à évaluer les effets de l'intelligence artificielle générative et d'autres plateformes émergentes. À cette fin, nous devons aussi tirer des enseignements des erreurs du passé. Lorsque les plateformes numériques ont été lancées, les dommages potentiels pour les sociétés et les personnes n'ont pas été suffisamment perçus ou évalués. Nous avons aujourd'hui la possibilité de faire en sorte que l'histoire ne se répète pas avec les technologies qui arrivent. La Silicon Valley, dont le principe de base était jusqu'à présent « Moving fast and break things » (Foncer et bousculer les conventions à tout prix) doit adopter une nouvelle devise. Il est essentiel que le respect de la vie privée des utilisateurs et utilisatrices, la sécurité, la transparence et la prise en compte de la sécurité soient intégrés dès la conception dans toutes les nouvelles technologies et tous les nouveaux produits.

Mesures prises par les Nations Unies

52. Des mesures sont prises, notamment par les opérations de paix et les bureaux de pays des Nations Unies, pour surveiller, analyser et combattre la menace que la mésinformation et la désinformation font peser sur l'exécution des mandats. La Stratégie et le Plan d'action des Nations Unies pour la lutte contre les discours de haine définissent les orientations stratégiques à suivre pour combattre ces discours aux niveaux national et mondial. En février 2023, l'UNESCO a organisé une conférence intitulée Pour un Internet de confiance afin de discuter d'un projet de

⁵⁷ Voir Center for Countering Digital Hate, *Misinformation on Bard, Google's new AI chat*, 5 avril 2023, et Tiffany Hsu et Stuart A. Thompson, *Disinformation researchers raise alarms about A.I. chatbots*, The New York Times, 13 février 2023.

principes mondiaux de régulation des plateformes numériques, qui devrait être achevé dans le courant de l'année⁵⁸.

53. Ensemble, ces initiatives et ces approches indiquent la voie à suivre pour définir les principes de base d'un code de conduite pour les Nations Unies.

VII. Vers un code de conduite pour les Nations Unies

54. Le code de conduite des Nations Unies pour l'intégrité de l'information sur les plateformes numériques, que je vais proposer, s'appuierait sur les principes suivants :

- Engagement envers l'intégrité de l'information ;
- Respect des droits humains ;
- Appui aux médias indépendants ;
- Amélioration de la transparence ;
- Moyens d'action pour les utilisateurs ;
- Renforcement de la recherche et de l'accès aux données ;
- Montée en puissance des réponses apportées ;
- Mesures de dissuasion plus fortes ;
- Renforcement de la confiance et de la sécurité.

55. Les principes précédemment énoncés ont été définis à partir des idées centrales exposées dans la présente note d'orientation, dans la droite ligne de ma note d'orientation sur l'élaboration d'un Pacte numérique mondial. Les États Membres seront invités à appliquer le code de conduite au niveau national. Les consultations tenues avec les parties prenantes vont se poursuivre afin de compléter le code de conduite et de mettre au jour des méthodes concrètes visant à traduire ces principes en termes opérationnels.

56. Le code de conduite pourrait s'appuyer sur les recommandations suivantes :

Engagement envers l'intégrité de l'information

a) Toutes les parties prenantes doivent s'abstenir d'utiliser, de soutenir ou d'amplifier la désinformation et les discours de haine à quelque fin que ce soit, notamment dans la poursuite d'objectifs politiques, militaires ou stratégiques, dans l'incitation à la violence, la remise en question de processus démocratiques ou le ciblage de populations civiles, de groupes vulnérables, de communautés ou de personnes ;

Respect des droits humains

b) Les États Membres devraient :

i) Faire en sorte que les mesures prises pour combattre la désinformation, la désinformation et les discours de haine respectent le droit international, y compris le droit international des droits humains, et ne soient pas utilisées à mauvais escient pour bloquer toute expression légitime de vues ou d'opinions, notamment par des coupures généralisées d'Internet ou l'interdiction de plateformes ou de médias ;

⁵⁸ On peut consulter le projet de principes à l'adresse <https://www.unesco.org/fr/internet-conference>.

- ii) Instaurer des mesures réglementaires visant à protéger les droits fondamentaux des utilisateurs et utilisatrices de plateformes numériques, dont des mécanismes d'application de la loi, en faisant preuve de transparence quant aux exigences imposées aux sociétés du secteur des technologies ;
- c) Toutes les parties prenantes devraient appliquer les Principes directeurs relatifs aux entreprises et aux droits de l'homme ;

Appui aux médias indépendants

- d) Les États Membres devraient garantir un environnement médiatique libre, viable, indépendant et pluriel, doté de mesures de protection solides pour les journalistes et les médias indépendants, et soutenir la création, le financement et les activités de formation d'organisations indépendantes de vérification des faits dans les langues locales ;
- e) Les organes de presse devraient veiller à ce que les informations soient communiquées de façon précise et éthique, grâce à une formation de qualité et des conditions de travail appropriées, dans le respect des normes internationales en matière de travail et de droits humains ;

Amélioration de la transparence

- f) Les plateformes numériques devraient :
 - i) Assurer une véritable transparence pour ce qui est des algorithmes, des données, de la modération de contenus et de la publicité ;
 - ii) Publier et faire connaître des politiques accessibles sur la mésinformation, la désinformation et les discours de haine, et rendre compte de la présence d'éléments de désinformation coordonnée sur leurs plateformes et de l'efficacité des mesures visant à contrer de telles opérations ;
- g) Les organes de presse devraient faire en sorte que les sources de financement et les politiques liées à la publicité soient vraiment transparentes, et établir une distinction claire entre le contenu éditorial et les publicités payantes, y compris lorsqu'ils publient des contenus sur des plateformes numériques ;

Moyens d'action pour les utilisateurs

- h) Les États Membres devraient faire en sorte que le public ait accès à des informations officielles précises, transparentes, et fondées sur des sources fiables, particulièrement lorsqu'elles sont d'intérêt public, et notamment pour tout ce qui concerne les objectifs de développement durable ;
- i) Les plateformes numériques devraient faire en sorte que les utilisateurs et utilisatrices disposent de moyens d'action et soient protégés, en toute transparence, afin qu'ils aient davantage de choix quant aux contenus qui leur sont présentés et à l'utilisation de leurs données. Elles devraient permettre aux utilisateurs et utilisatrices d'apporter la preuve de leur identité et de l'authenticité de leur compte sans contrepartie financière ou liée à leur vie privée et mettre en place des mécanismes transparents de réclamation et de signalement par les utilisateurs, appuyés par des mécanismes d'examen des plaintes accessibles et bien connus ;
- j) Toutes les parties prenantes devraient investir dans des campagnes dynamiques d'éducation aux outils numériques visant à permettre aux utilisateurs et utilisatrices de tous âges de comprendre comment fonctionnent les plateformes numériques, comment leurs données personnelles peuvent être utilisées, et comment repérer la mésinformation, la désinformation et les discours haineux et y répondre. Il

convient de veiller tout particulièrement à ce que les jeunes, les adolescents et les enfants aient pleinement connaissance de leurs droits dans les espaces en ligne ;

Renforcement de la recherche et de l'accès aux données

k) Les États Membres devraient investir dans la recherche indépendante sur la présence et les conséquences de la mésinformation, de la désinformation et des discours de haine dans différents pays et langues et soutenir ce type de recherche, particulièrement dans les contextes sous-représentés et dans des langues autres que l'anglais, en permettant à la société civile et aux universitaires de travailler librement et en toute sécurité ;

l) Les plateformes numériques devraient :

i) Autoriser les chercheurs et chercheuses et les universitaires à accéder aux données tout en respectant la vie privée des utilisateurs et utilisatrices. Les chercheurs et chercheuses devraient avoir les moyens de recueillir des exemples et des données de qualité sur les personnes et les groupes ciblés par la mésinformation, la désinformation et les discours de haine afin de mieux comprendre l'ampleur et la nature des préjudices subis, tout en protégeant les données et en respectant les droits humains ;

ii) Garantir la pleine participation de la société civile dans les efforts déployés pour lutter contre la mésinformation, la désinformation et les discours de haine ;

Intensification des réponses apportées

m) Toutes les parties prenantes devraient :

i) Affecter des ressources à la recherche de solutions et à l'établissement de rapports sur les origines, la diffusion et l'incidence de la mésinformation, de la désinformation et des discours de haine, tout en respectant les normes en matière de droits humains, et investir davantage dans les capacités de vérification des faits dans tous les pays et contextes ;

ii) Former de vastes coalitions sur l'intégrité de l'information, qui réuniraient différentes formes d'expertises et d'approches afin de réduire l'écart entre les organisations locales et les entreprises du secteur de la technologie opérant à l'échelle mondiale ;

iii) Promouvoir la formation et le renforcement des capacités afin de mieux comprendre de quelle manière la mésinformation, la désinformation et les discours de haine apparaissent et de renforcer la prévention ainsi que les stratégies d'atténuation ;

Mesures de dissuasion plus fortes

n) Les plateformes numériques devraient s'écarter des modèles économiques qui font passer l'engagement avant les droits humains, la vie privée et la sécurité ;

o) Les annonceurs et les plateformes numériques devraient faire en sorte que les publicités ne soient pas placées à côté de contenus de mésinformation ou de désinformation, ou de discours haineux, et que les publicités contenant de tels messages ne sont pas mises en avant ;

p) Les organes de presse devraient veiller à ce que l'ensemble des publicités et contenus publicitaires soient clairement libellés comme tels et ne contiennent pas de discours de mésinformation, de désinformation ou de haine ;

Renforcement de la confiance et de la sécurité

- q) Les plateformes numériques devraient :
 - i) Veiller à ce que la sécurité et la protection de la vie privée soient prises en compte dès la conception, y compris en utilisant les ressources internes appropriées en matière d'expertise quant à la confiance et à la sécurité, et en appliquant les mesures de façon cohérente dans tous les pays et les langues ;
 - ii) Investir dans des systèmes de modération de contenus par des humains et par des intelligences artificielles, dans toutes les langues utilisées dans les pays dans lesquels elles mènent leurs activités, et veiller à ce que les mécanismes de communication de l'information soient transparents et comportent un taux de réponse accéléré, en particulier en cas de conflit ;
- r) Toutes les parties prenantes devraient prendre le plus rapidement possible des mesures pour que l'utilisation de l'intelligence artificielle soit sûre, sécurisée, responsable, éthique et respectueuse des droits humains, et s'atteler aux incidences des progrès récemment réalisés dans ce domaine en lien avec la mésinformation, la désinformation et les discours de haine.

VIII. Prochaines étapes

57. Le Secrétariat de l'Organisation des Nations Unies va entreprendre de vastes consultations avec différentes parties prenantes aux fins de l'élaboration d'un code de conduite pour les Nations Unies, qui prévoirait aussi des mécanismes de suivi et d'application. La création d'un observatoire indépendant composé d'experts reconnus et chargé d'évaluer les mesures prises par les acteurs qui s'engageront à respecter le code de conduite est aussi envisagée, de même que d'autres mécanismes de communication de l'information.

58. Pour soutenir et éclairer l'élaboration du code, le Secrétariat pourrait mener des études approfondies afin de mieux comprendre l'intégrité de l'information dans le monde, en particulier dans les régions moins prises en compte dans la recherche.

59. Le Secrétaire général va doter le Secrétariat d'une capacité spécialisée afin d'intensifier la réponse apportée à la mésinformation, à la désinformation et aux discours de haine en ligne, qui nuisent à l'exécution des mandats de l'Organisation et à ses priorités de fond. Grâce au suivi et à l'analyse réalisés par des experts, cette capacité pourrait élaborer des stratégies de communication adaptées afin d'anticiper ou d'atténuer rapidement les menaces avant qu'elles ne fassent du tort en ligne et hors ligne, et soutenir le renforcement des capacités du personnel de l'Organisation des Nations Unies et des États Membres. Elle appuierait les efforts déployés par les États Membres, les plateformes numériques et d'autres parties prenantes pour se conformer au code et appliquer ses dispositions, lorsqu'il sera achevé.

IX. Conclusion

60. Le renforcement de l'intégrité de l'information sur les plateformes numériques est une priorité absolue pour la communauté internationale. De la santé à l'égalité des genres, en passant par la paix, la justice, l'éducation et l'action climatiques, les mesures limitant les conséquences de la mésinformation, de la désinformation et des discours de haine donneront un coup de fouet aux efforts visant à créer un avenir durable et à ne laisser personne de côté. Même s'il est possible d'agir au niveau national, ces problèmes ne peuvent être complètement résolus que par une

coopération mondiale plus solide. Les idées fondamentales mises en avant dans la présente note d'orientation montrent que la voie vers un renforcement de l'intégrité de l'information doit être centrée sur les droits humains, faire intervenir plusieurs parties prenantes et tenir compte de dimensions multiples. Elles ont été condensées en différents principes qui doivent être examinés aux fins de l'élaboration d'un code de conduite des Nations Unies pour l'intégrité de l'information sur les plateformes numériques, qui pourrait servir de modèle pour renforcer l'intégrité de l'information tout en respectant fermement les droits humains. Je me réjouis à la perspective de collaborer avec les États Membres et d'autres partenaires pour faire de ces principes des engagements concrets.

Annexe

Consultations avec les États Membres et les autres parties prenantes concernées

1. Les idées qui sont exposées dans la présente note d'orientation s'inspirent des propositions présentées dans le rapport intitulé « Notre Programme commun » (A/75/982), qui a fait l'objet de consultations approfondies avec les États Membres, le système des Nations Unies, des leaders d'opinion, des jeunes et des acteurs de la société civile du monde entier. La présente note d'orientation donne suite en particulier aux réflexions riches et détaillées développées par les États Membres et d'autres parties prenantes sur Notre Programme commun au détour de 25 séances de concertation tenues dans le cadre de l'Assemblée générale.
2. Avant de publier la présente note d'orientation, des consultations ont été menées avec les États Membres, notamment lors d'un exposé au Comité de l'information, auquel tous les États non membres du Comité ont été invités à participer. Des discussions ont aussi eu lieu avec les partenaires de la société civile, le milieu universitaire, des experts et expertes et le secteur privé, notamment les entreprises du secteur des technologies.
3. De vastes consultations vont être menées sur l'élaboration du code de conduite avant le Sommet de l'avenir.