COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

CDMSI(2023)016
12 décembre 2023

Projet d'exposé des motifs de 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

1. Introduction

La désinformation est devenue une préoccupation majeure dans les démocraties du monde entier et fait partie intégrante de l'ère numérique. Les individus naviguent désormais dans un paysage médiatique fragmenté, interconnecté et complexe qui s'étend au-delà des médias de masse traditionnels pour inclure de nombreuses plateformes de médias sociaux et de canaux d'information. Comprendre le degré de diffusion et l'impact de la désinformation sur les démocraties est devenu de plus en plus crucial pour les observateurs de cet environnement d'information complexe.

L'hypothèse largement répandue au sujet de la désinformation est qu'elle existe à des niveaux alarmants, imprégnant toutes les couches de la société et de la gouvernance démocratique. Toutefois, les données empiriques complètes sur sa prévalence sont rares. Le rapporteur spécial des Nations unies sur la liberté d'expression a récemment souligné que "[m]ême si la recherche empirique suggère que seule une petite proportion de personnes est exposée à la désinformation, les incidences sur les institutions, les communautés et les individus sont réels, étendus et légitimes",[1] en s'appuyant sur des communications d'institutions académiques sur la recherche empirique.[2] Alors que des rapports spécialisés continuent de mettre l'accent sur les menaces posées par la désinformation dans des domaines spécifiques, tels que l'ingérence russe dans la diffusion d'informations concernant l'agression de l'Ukraine depuis 2022,[3] ou les niveaux de désinformation dans les rapports des médias sur la crise climatique et la pandémie de COVID-19,[4] , le paysage empirique global reste complexe. La recherche universitaire a également exploré les effets secondaires de la désinformation, mettant en évidence un décalage entre sa présence perçue et la gravité des problèmes qu'elle engendre à l'échelle mondiale. Elle révèle que la peur des individus d'être exposés à la désinformation peut être tout aussi puissante que les fausses informations elles-mêmes, en alimentant l'appréhension, la colère et la polarisation des opinions.[5]

Une autre conclusion essentielle de la recherche sur la désinformation est qu'il s'agit d'une cible mouvante. La technologie permettant la création et la diffusion de fausses informations évolue rapidement, ce qui nécessite une attention constante aux développements récents tels que le rôle de l'IA générative dans la production de textes, d'images et de vidéos. Des technologies clés telles que les deepfakes et le ChatGPT soulignent la nécessité d'une mise à jour constante et d'une vigilance accrue dans la compréhension de la désinformation. En outre, le fait que la désinformation soit utilisée comme une arme par des acteurs malveillants est un sujet d'intérêt mondial qui requiert une attention particulière en raison des risques qu'il fait peser sur la démocratie.

[1] Rapport du rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, "Désinformation et liberté d'opinion et d'expression", (A/HRC/47/25), 13 avril 2021, paragraphe 22.

[2] Voir, par exemple, la communication NYU Center for Social Media and Politics ; et la communication de l'Institut Reuters pour l'étude du journalisme, Université d'Oxford.

[3] Service d'action extérieure de l'Union européenne (2023), 1st EEAS report on foreign information manipulation and interference threats : Towards a framework for networked defence. Disponible à l'adresse : https://euvsdisinfo.eu/uploads/2023/02/EEAS-ThreatReport-February2023-02.pdf, consulté le 26 septembre 2023.

[4] First Draft research ; disponible à l'adresse : https://firstdraftnews.org/research/, consulté le 26 septembre 2023.

[5] Lecheler, S. & Egelhofer, J. L. (2022), "Disinformation, misinformation, and fake news : understanding the supply side", in Strömbäck J. et al. (Eds.), *Knowledge resistance in high-choice information environments*, Routledge, pp. 69-87.

Le travail des vérificateurs de faits joue un rôle central dans la lutte mondiale contre la désinformation, à la fois par la vérification des informations et par l'émergence d'une communauté professionnelle dédiée à cet effort. La vérification des faits est à la fois un processus et une profession, et il a fait l'objet d'études approfondies ces dernières années. En outre, au-delà de la vérification des faits, les solutions de conception des plateformes peuvent également contribuer à réduire la propagation et les effets négatifs de la désinformation en ligne. En effet, les solutions de conception de plateformes peuvent promouvoir les conditions d'un débat public sain et constructif et garantir un environnement favorable à la liberté d'expression, ce qui est essentiel pour lutter contre la désinformation. Les utilisateurs sont un troisième élément clé du renforcement de la résilience face à la propagation de la mésinformation et de la désinformation. Il est essentiel de leur donner les moyens d'agir : si la vérification des faits peut contribuer à la démystification et si les solutions de conception des plateformes peuvent aider à réduire la portée de la désinformation, les utilisateurs seront toujours exposés à un certain type de contenu erroné et doivent être équipés pour relever ce défi.

Enfin, il convient de noter que les débats publics, universitaires et réglementaires s'appuient souvent sur la distinction entre la désinformation et la mésinformation. La "mésinformation" peut se référer à des informations fausses, inexactes ou trompeuses diffusées sans intention d'induire en erreur, de causer un préjudice ou de rechercher un gain économique ou politique ; les utilisateurs qui partagent des informations erronées les croient généralement vraies.[6] La "désinformation" fait référence à des informations vérifiables, fausses, inexactes ou trompeuses, délibérément créées et diffusées dans le but de causer un préjudice ou de rechercher un gain économique ou politique en trompant le public, selon la définition du Comité des Ministres du Conseil de l'Europe.[7] Toutefois, la distinction entre la désinformation et la mésinformation est difficile à appliquer dans la pratique, car il est souvent impossible de déterminer l'intention de ceux qui ont créé et diffusé certains contenus et le même contenu peut être diffusé avec des intentions différentes par des individus et des groupes différents..[8] En fin de compte, les deux formes contribuent à la diffusion de fausses informations, ce qui affaiblit l'écosystème de l'information et met en péril les valeurs fondamentales. Il est donc essentiel de mettre au point des mécanismes permettant de lutter plus efficacement contre la diffusion intentionnelle et non intentionnelle de fausses informations. Pour ces raisons, il est suggéré qu'il serait utile que le Conseil de l'Europe poursuive sa réflexion, ses recherches et ses orientations en la matière.

[6] Voir également Wardle, C., Derakhshan, H. (2017), *Information Disorder : Toward an interdisciplinary framework for research and policymaking*, rapport du Conseil de l'Europe DGI(2017)09, p. 20 (qui définit la désinformation comme une "[i]nformation qui est fausse, mais qui n'a pas été créée dans l'intention de nuire").

[7] Voir Recommandation CM/Rec(2022)11 du Comité des Ministres aux Etats membres sur les principes de gouvernance des médias et de la communicationannexe, paragraphe 4 ; et la Recommandation CM/Rec(2022)12 du Comité des Ministres aux Etats membres sur la communication électorale et la couverture médiatique des campagnes électorales annexe, paragraphe 4.

[8] Pour des discussions critiques sur les définitions de la désinformation, voir, par exemple, Ó Fathaigh, R., Helberger, N., & Appleman N. (2021), "The perils of legally defining disinformation", 10(4) *Internet Policy Review* 1 ; Lecheler, S., & Egelhofer, J. L. (2022), "Disinformation, misinformation, and fake news : understanding the supply side", in Strömbäck J. et al. (Eds.), *Knowledge resistance in high-choice information environments*, Routledge, pp. 69-87 ; Pérez-Escolar, M., Lilleker, D., & Tapia-Frade, A. (2023), "A systematic literature review of the phenomenon of disinformation and misinformation", *Media and Communication*, 11(2), pp. 76-87 ; Vraga, E. K., & Bode, L. (2020), "Defining misinformation and understanding its bounded nature : Using expertise and evidence for describing misinformation", *Political Communication*, 37(1), pp. 136-144 ; Freelon, D., & Wells, C. (2020), "Disinformation as Political Communication", *Political Communication*, 37(2) ; Hameleers, M., Brosius, A., Marquart, F., Goldberg, A. C., van Elsas, E., & de Vreese, C. H. (2022), "Mistake or Manipulation ? Conceptualizing Perceived Mis- and Disinformation among News Consumers in 10 European Countries", *Communication Research*, 49(7), pp. 919-941.

Cet exposé des motifs donne un aperçu concis des principes qui sous-tendent la Note d'orientation, en abordant ses trois piliers centraux : le rôle des vérificateurs de faits, la conception des plateformes et l'autonomisation des utilisateurs. En s'appuyant sur la recherche scientifique, les normes en matière de droits de l'homme et d'autres sources pertinentes, il développe davantage ces piliers, offrant des standards et d'autres éléments précieux dans la lutte contre la désinformation.

2. Vérification des faits

Cette première section de l'exposé des motifs vise à faire comprendre l'émergence de la vérification des faits et son importance au sein des écosystèmes médiatiques démocratiques. Elle cherche à contextualiser l'impact que la vérification des faits peut avoir sur ces écosystèmes, en soulignant son rôle et ses limites dans la promotion de l'exactitude, de la responsabilité et d'un discours public informé. En outre, cette section vise à présenter des exemples de bonnes pratiques et de normes communes qui sont actuellement développées au sein des réseaux de vérification des faits. Ces pratiques et normes servent de lignes directrices pour garantir la crédibilité et l'efficacité des initiatives de vérification des faits, encourager la transparence et renforcer la confiance du public dans l'écosystème de l'information.

a. Situation actuelle : la pratique et la profession de la vérification des faits

À une époque caractérisée par la diffusion rapide d'informations sur diverses plateformes et canaux, l'importance d'informations précises et fiables est devenue de plus en plus cruciale. La vérification des faits a émergé en tant que (a) pratique et (b) profession de premier plan à cette époque, assumant un rôle important dans le paysage moderne de l'information. Cependant, il est important de reconnaître que ces deux aspects de la vérification des faits ne sont pas uniquement liés à l'"ère de la désinformation", mais s'appuient également sur des décennies de culture journalistique et sur l'application de méthodes scientifiques de vérification au cours du processus de production des informations. Cela signifie que tout effort visant à relever les défis actuels et futurs liés à la désinformation peut s'appuyer sur les bonnes pratiques existantes dans ce domaine, les appliquer de manière plus globale et les développer davantage à la lumière des réalités technologiques en évolution.

Tout d'abord, la vérification des faits peut être compris comme une combinaison de normes, de routines et de comportements visant à vérifier l'exactitude des informations à l'ère numérique.[9] Ce processus implique non seulement des vérificateurs de faits professionnels, mais aussi d'autres communautés engagées dans ces objectifs et d'autres moyens technologiques, tels que les outils de vérification des faits qui qui utilisent le traitement du langage naturel,[10] , tous dans le but de vérifier l'information. La vérification des faits est une boîte à outils puissante pour lutter contre la diffusion de mésinformation et de désinformation et vise à renforcer la confiance dans les nouvelles et la communication publique. D'un point de vue conceptuel, la vérification des faits englobe l'utilisation de routines et de normes journalistiques liées à la vérification, ainsi que de méthodes scientifiques pour mener à bien son analyse.[11] Allant au-delà de l'objectif abstrait de correction des fausses informations, la pratique de la vérification des faits implique aujourd'hui des "évaluations régulières de la validité des affirmations douteuses faites par un large éventail de sources, y compris les gouvernements, les politiciens, les institutions, les organismes de presse et les utilisateurs de médias sociaux".[12] Elle englobe non seulement le processus de vérification, mais aussi la publication d'informations corrigées et l'identification des acteurs responsables de la désinformation, ce qui permet d'atteindre un public plus large.

Il est important de noter que la vérification des faits englobe aujourd'hui un large éventail de techniques employées pour vérifier l'exactitude des informations dans le domaine public, ainsi que les identités, les stratégies et les pratiques des personnes responsables de la production ou de la diffusion de ces informations. Il s'agit d'une pratique et d'une profession en constante évolution qui reflète la manière dont l'information est gérée au sein des sociétés et qui met en lumière les individus ou les entités responsables de la sauvegarde de l'information tout au long du processus politique. La vérification des faits ne se limite pas au seul processus de production des informations, mais s'étend aux informations déjà présentes dans le discours public. Les vérificateurs de faits sont les gardiens de l'intégrité de l'information, avant et après sa diffusion. Cette forme de vérification des faits implique une collaboration entre les journalistes, les spécialistes de la communication, les chercheurs et les organisations de vérification des faits pour lutter contre la désinformation. En outre, la vérification professionnelle des faits est de plus en plus souvent entreprise par des institutions gouvernementales, ainsi que dans le cadre d'efforts de collaboration entre les organisations de vérification des faits et les gouvernements. Cette vérification des faits par l'État-Cette vérification des faits par l'État doit cependant être évaluée avec prudence car elle comporte des menaces potentielles pour l'indépendance des acteurs des médias dans les démocraties.

[9] Graves, L. (2016), *Deciding what's true : The rise of political fact-checking in American journalism*, Columbia University Press. Voir également Silverman, C. (2015), *Lies, Damn Lies, and Viral Content : How News Websites Spread (and Debunk) Online Rumors, unverified, Claims, and Misinformation*, Tow Center for Digital Journalism ; et Guerrini, F. (2013), "From traditional to online fact-checking", Oxford Magazine, Eighth Week, Trinity Term.

[10] Zeng, X., Abumansour, A. S., & Zubiaga, A. (2021), "Automated fact-checking : A survey", *Language and Linguistics Compass*, 15(10), disponible sur : https://compass.onlinelibrary.wiley.com/doi/10.1111/lnc3.12438, consulté le 28 septembre 2023.

[11] Mena, P. (2019), "Principles and boundaries of fact-checking : Journalists' perceptions", *Journalism Practice*, 13(6), pp. 657-672.

[12] Vinhas, O., & Bastos, M. (2022), "Fact-Checking Misinformation : Eight Notes on Consensus Reality", *Journalism Studies*, 23(4), pp. 448-468, en particulier p. 449.

Deuxièmement, la vérification des faits est une profession journalistique. Depuis de nombreuses années, des vérificateurs de faits spécialisés travaillent dans les salles de rédaction du monde entier, en particulier aux États-Unis. Ces vérificateurs de faits sont chargés de vérifier l'exactitude du travail des journalistes, de revérifier les détails importants, tout en laissant à d'autres le soin de rédiger les articles. Leur objectif principal est d'empêcher la publication d'informations inexactes dans les organes d'information. La propagation croissante de la mésinformation et de la désinformation, ainsi que les préoccupations qu'elles suscitent, ont entraîné une croissance substantielle de la profession de vérificateurs de faits et un élargissement de son champ d'action. Les grands organes de presse et les radiodiffuseurs publics mettent de plus en plus en place des bureaux spécialisés dans la désinformation et la vérification des faits, qui utilisent des outils numériques pour vérifier les informations, parfois dans le cadre d'une coopération transfrontalière. Ces bureaux procèdent à des vérifications factuelles en se concentrant sur la vérification "ex post" des informations et des reportages déjà présents dans le domaine public.

L'émergence de la vérification des faits en tant que profession est toutefois plus évidente avec la création d'organisations de vérification des faits ou de projets de vérification des faits indépendants, c'est-à-dire d'organisations dont l'objet principal est la vérification des faits, mais qui ne sont pas liées à des organisations médiatiques traditionnelles. Ce type d'organisation est devenu un écosystème professionnel à part entière. Les organisations de fact-checking et leurs sites web et/ou outils dédiés au fact-checking emploient souvent aussi des journalistes mais ont des valeurs et des normes différentes de celles des organes de presse traditionnels. Certaines organisations de vérification des faits sont basées sur des projets, issus d'universités, de groupes de réflexion, d'organisations gouvernementales et d'autres parties prenantes afin de relever des défis spécifiques en matière de désinformation.[13] Graves et Cherubini (2016) résument la mission des vérificateurs de faits professionnels en trois identités de rôle. Tout d'abord, il y a un groupe important de "reporters", ceux qui s'identifient comme journalistes et considèrent que la vérification des faits fait partie des services d'information. Les "réformateurs" conçoivent la vérification des faits comme un activisme et associent leurs activités à des changements politiques. Les "experts" sont ceux qui "mettent particulièrement l'accent sur leur propre expertise dans le domaine ou sur une méthodologie distincte, se positionnant comme un groupe de réflexion plutôt que comme des journalistes ou des militants". [14]

b. Principaux défis de la vérification des faits dans le nouvel environnement numérique de l'information

La vérification des faits peut donc être considérée à la fois comme une pratique et comme une profession, et ces deux aspects ont un impact sur l'environnement d'information dans lequel ils opèrent. Toutefois, la nature exacte de cet impact fait encore l'objet de débats dans la littérature de recherche actuelle, ce qui met en évidence un certain nombre de défis majeurs.

[13] Graves, L. (2018), "Boundaries Not Drawn", *Journalism Studies*, 19:5, pp. 613-631.
[14] Graves, L., & Cherubini, F. (2016), "The Rise of Fact-Checking Sites in Europe", in *Digital News Project Report (Reuters Institute Digital News Report)*, Reuters Institute for the Study of Journalism, p. 17.

La pratique de la vérification des faits a longtemps été principalement associée au travail effectué dans les salles de rédaction. Les techniques de vérification fondées sur le professionnalisme journalistique se concentrent principalement sur l'étude de l'origine des informations et la validation croisée des sources à travers différents canaux d'information. Traditionnellement, ces canaux impliquaient la vérification d'informations par le biais de réseaux personnels et l'investigation d'autres sources au sein de l'élite des médias. Par exemple, un journaliste politique qui reçoit de nouvelles informations sur les développements politiques en cours dans un pays spécifique tentera généralement de vérifier ces informations par le biais d'un réseau de sources au sein des institutions gouvernementales et des organisations affiliées. Une autre étape consisterait à vérifier si et comment d'autres médias ou agences de presse, tant au niveau national qu'international, ont rendu compte de ces informations. En outre, lors de la publication d'un nouvel article, celui-ci est souvent soumis à une vérification des faits par les rédacteurs en chef et les vérificateurs de faits afin de s'assurer de son exactitude.[15] Aujourd'hui, ces techniques et processus traditionnels sont confrontés à des défis en raison de plusieurs changements dans l'écosystème journalistique.

Tout d'abord, le journalisme s'oriente vers le généralisme, ce qui suppose que les journalistes eux-mêmes possèdent les compétences nécessaires pour gérer l'ensemble du processus de production, de la recherche à la publication. En conséquence, la responsabilité de la vérification des faits au sein des salles de rédaction incombe en grande partie à des journalistes individuels plutôt qu'à des services spécialisés dans la vérification des faits. Cette responsabilité et cette pression accrues limitent le temps qu'un journaliste peut consacrer à des pratiques de vérification des faits telles que la vérification de ses sources pour un article, ce qui l'expose au risque de passer à côté d'un contenu potentiellement faux. Tous ces défis surviennent à un moment où les conditions de travail des journalistes se détériorent régulièrement et considérablement, en raison de la crise des modèles économiques établis des organismes de presse,[16] , qui affecte également les vérificateurs de faits.

[15] Van Leuven et al. (2018), "Online and newsworthy : Les sources en ligne ont-elles changé le journalisme ?", *Digital Journalism*, 6(7), pp. 798-806.
[16] Recommandation CM/Rec(2022)4 sur la promotion d'un environnement favorable au journalisme de qualité à l'ère numérique.

Le nombre croissant d'activités dont les journalistes sont responsables au cours de la production d'informations est exacerbé par la numérisation croissante de la production et de la diffusion des informations. de la production d'informations et de la de la vérification des faits. À l'ère du numérique, les sources ne sont plus seulement les agences de presse traditionnelles ou les bureaux de presse institutionnels, mais contiennent une myriade de sites web, de pages personnelles sur les médias sociaux et de sources de données complexes. Cela signifie qu'aujourd'hui, la vérification des faits doit impliquer une vérification par le biais de technologies numériques complexes, y compris des moteurs de recherche, des outils d'intelligence artificielle et des applications dédiées à la vérification des faits. L'intégration de l'information numérique dans le processus de recherche journalistique a posé de nouveaux défis aux compétences traditionnelles de vérification qui sont toujours enseignées aux aspirants journalistes dans le monde entier. La complexité et la nature algorithmique de ces technologies ont posé des difficultés aux journalistes. Des études expérimentales indiquent que la vérification des faits dans les salles de rédaction continue de s'appuyer fortement sur des méthodes traditionnelles, telles que les conversations téléphoniques, et que seuls les journalistes spécialisés, comme par exemple les journalistes de données et les journalistes d'investigation qui utilisent systématiquement et efficacement des outils technologiques avancés dans le cadre des processus de vérification.[17] Par conséquent, la vérification des faits dans les salles de rédaction peut nécessiter des formes innovantes de collaboration. Les journalistes et les salles de rédaction s'associent de plus en plus à d'autres organes de presse, à des ONG spécialisées dans la vérification des faits, à des scientifiques spécialisés en matière des données et à des plateformes de médias sociaux au cours de la recherche.

Comme indiqué plus haut, la pratique de la vérification des faits évolue de plus en plus et comprend désormais une série de nouvelles organisations à but non lucratif visant à vérifier les informations disponibles dans le domaine public. Il ne fait aucun doute que ces organisations ont conduit à une professionnalisation accrue des techniques de vérification des faits au sein d'un secteur en pleine expansion. Tout comme leurs collègues des organisations journalistiques, les vérificateurs de faits professionnels travaillant dans les organisations de vérification de faits emploient des méthodes de vérification systématiques et rigoureuses tout en possédant les compétences nécessaires pour communiquer efficacement leurs résultats au public. Bien que certaines de ces organisations aient expérimenté des outils numériques innovants, il semble que la plupart d'entre elles s'appuient encore sur des techniques de vérification traditionnelles. Certaines études indiquent que les organisations de vérification des faits s'appuient sur des méthodes de vérification manuelles, en particulier dans les cas où des outils plus importants ou des approches automatisées ne sont pas disponibles en raison de limitations linguistiques ou de contraintes financières [18] . Cela signifie que les organisations de vérification des faits peuvent être vulnérables aux mêmes défis que ceux décrits ci-dessus dans le contexte du journalisme, y compris les conditions de travail.

[17] Lecheler, S., & Kruikemeier, S. (2016), "Re-evaluating journalistic routines in a digital age : A review of research on the use of online sources", *New media & society*, 18(1), pp. 156-171 ; de Haan, Y., van den Berg, E., Goutier, N., Kruikemeier, S., & Lecheler, S. (2022), "Invisible Friend or Foe ? How Journalists Use and Perceive Algorithmic-Driven Tools in Their Research Process", *Digital Journalism*, 10(10), pp. 1775-1793.

[18] Hrckova, A., Moro, R., Srba, I., Simko, J., & Bielikova, M. (2022), "Automated, not Automatic : Needs and Practices in European Fact-checking Organizations as a basis for Designing Human-centered AI Systems", *arXiv preprint arXiv:2211.12143*.

La promotion des organisations de vérification des faits est cependant cruciale : les organisations de de vérification des faits sont une alternative au journalisme traditionnel à une époque où les citoyens ont un accès direct à l'information et à la communication politique sans dépendre uniquement des filtres des médias traditionnels. C'est particulièrement le cas en Europe, alors que la vérification des faits en salle de rédaction est encore plus couramment pratiquée aux États-Unis.[19] Dans de nombreux pays, en particulier en Europe de l'Est, les organisations de vérification des faits visent à renforcer la responsabilité des hommes politiques et des entités médiatiques extérieures au système médiatique traditionnel.[20]

Comme nous l'avons mentionné précédemment, les vérificateurs de faits, qu'ils soient basés dans les salles de rédaction ou non, s'engagent de plus en plus auprès d'un troisième groupe d'acteurs, qui comprend les plateformes de médias sociaux et d'autres entreprises technologiques. Étant donné qu'une grande partie des informations et des discussions se déroulent sur les plateformes de médias sociaux dans le monde entier, la vérification des faits est devenue une pratique courante dans ces espaces. Certaines plateformes ont intensifié leurs efforts de modération du contenu ou ont mis en place des partenariats avec des organisations externes de vérification des faits. partenariats avec des organisations externes de vérification des faits. Dans le domaine des plateformes, la vérification des faits est largement discutée dans la littérature plus développée sur la modération du contenu des plateformes, en particulier en ce qui concerne l'efficacité et les considérations éthiques des interventions humaines par rapport aux interventions mécaniques sur les plateformes de médias sociaux. Si les plateformes ont eu tendance à laisser la vérification des faits aux utilisateurs individuels, elles ont surtout négligé de fournir l'infrastructure et les outils qui rendraient ces activités individuelles plus faciles et plus évolutives.[21] En outre, cela signifie que les frontières entre les utilisateurs et les vérificateurs de faits professionnels deviennent floues, ce qui pose sa propre série de défis.

[19] Vinhas, O., & Bastos, M. (2022), "Fact-Checking Misinformation : Eight Notes on Consensus Reality", *Journalism Studies*, 23(4), pp. 448-468.
[20] Voir Graves, L. & Cherubini, F. (2016), *The Rise of Fact-Checking Sites in Europe*, Reuters Institute for the Study of Journalism, p.10.
[21] Les notes communautaires de X (anciennement Twitter), qui permettent aux utilisateurs d'ajouter des informations contextuelles au contenu publié par d'autres, sont un exemple de plateforme majeure fournissant ce type d'infrastructure. Disponible à l'adresse : https://communitynotes.twitter.com/guide/it/about/introduction, consulté le 28 septembre 2023.

En effet, il existe un intérêt croissant pour la production de données permettant d'évaluer l'efficacité des méthodes professionnelles de vérification des faits dans la lutte contre la désinformation. Comme indiqué précédemment, les recherches indiquent que les techniques traditionnelles de vérification journalistique sont confrontées à des défis dans les environnements complexes de l'information numérique. En ce qui concerne la vérification des informations publiques, les méta-analyses suggèrent que les méthodes de vérification des faits peuvent être efficaces, mais que leur succès dépend fortement des différences individuelles et contextuelles.[22] Par exemple, l'acceptation des résultats de la vérification des faits varie en fonction de l'arrière-plan idéologique des individus. Les individus peuvent être réticents à accepter qu'une information particulière soit fausse si elle contredit leur propre idéologie politique.[23] Outre les facteurs individuels tels que l'idéologie et les traits de personnalité, l'acceptation de mésinformation et de désinformation corrigées dépend également des caractéristiques du message et du sujet.[24] Par exemple, l'hypothèse selon laquelle la désinformation audiovisuelle est plus difficile à corriger est répandue dans la littérature actuelle, mais n'a pas encore été pleinement explorée.[25]

c. Relever les défis : meilleures pratiques et normes de vérification des faits

La professionnalisation de la vérification des faits entraîne la nécessité d'établir des cadres de référence et des normes internationales.[26] Ces principes, aujourd'hui formalisés dans des ouvrages tels que « Code de principes » du réseau international de vérification des faits (IFCN), sont principalement ancrés dans le principe d'indépendance, qui souligne l'importance d'être libre de tout parti pris partisan. La transparence en matière de sources, de financement, de méthodologie et de communication avec le public est également essentielle. Il est essentiel de mettre l'accent sur le développement de méthodologies et de bonnes pratiques reconnues au niveau international, ainsi que sur l'établissement de normes dans le domaine de la vérification des faits.

Une grande partie de la littérature se concentre sur la relation étroite entre la vérification des faits et le travail journalistique. Les meilleures pratiques en matière de vérification des faits sont intrinsèquement liées aux normes du "bon" journalisme.[27] La pertinence, la crédibilité et l'indépendance font partie des critères d'évaluation de la vérification des faits. Cependant, la vérification des faits n'est pas simplement une extension du journalisme, mais plutôt une pratique et une profession distincte. Alors que l'identité professionnelle de la vérification des faits continue d'évoluer, des normes basées sur le travail journalistique se développent également.[28]

---

[22] Oeldorf-Hirsch, A., Schmierbach, M., Appelman, A., & Boyle, M. P. (2023), "The influence of fact-checking is disputed ! The role of party identification in processing and sharing fact-checked social media posts", *American Behavioral Scientist*, 00027642231174335 ; Nieminen, S., & Rapeli, L. (2019), "Fighting misperceptions and doubting journalists' objectivity : A review of fact-checking literature", *Political Studies Review*, 17(3), pp. 296-309.
[23] Voir Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020), "Fact-checking : A meta-analysis of what works and for whom", *Political Communication*, 37(3), pp. 350-375.
[24] Voir Chung, M., & Kim, N. (2021), "When I learn the news is false : How fact-checking information stems the spread of fake news via third-person perception", *Human Communication Research*, 47(1), pp. 1-24.
[25] Voir Thomson, T. J., Angus, D., Dootson, P., Hurcombe, E., & Smith, A. (2022), "Visual mis/disinformation in journalism and public communications : Current verification practices, challenges, and future opportunities", *Journalism Practice*, 16(5), pp ; 938-962 ; Weikmann, T., & Lecheler, S. (2022), "Visual disinformation in a digital age : A literature synthesis and research agenda", *new media & society*, disponible sur : https://journals.sagepub.com/doi/10.1177/14614448221141648, consulté le 28 septembre 2023.
[26] Humprecht, E. (2020), "How do they debunk "fake news" ? A cross-national comparison of transparency in fact checks", *Digital journalism*, 8(3), pp. 310-327.
[27] Himma-Kadakas, M., & Ojamets, I. (2022), "Debunking false information : investigating journalists' fact-checking skills", *Digital Journalism*, 10(5), pp. 866-887.
[28] Graves, L., & Amazeen, M. A. (2019), "Fact-checking as idea and practice in journalism", in *Oxford research encyclopedia of communication*.

Par exemple, jusqu'à récemment (ou peut-être encore aujourd'hui), il n'y avait guère de consensus sur la longueur optimale des vérifications de faits, les informations qu'elles devraient inclure ou les méthodes spécifiques par lesquelles la véracité devrait être recherchée (par exemple, par de nouvelles sources ou par l'analyse des sources existantes dans le texte/visuel original vérifié). En outre, les vérificateurs de faits ont mis au point des systèmes uniques de mesure de la véracité, tels que des systèmes d'évaluation et de classement.[29] Ces systèmes s'écartent des normes journalistiques traditionnelles de vérification et peuvent présenter des difficultés, car ils appliquent une logique quasi quantitative au processus de vérification de faits complexes dans un contexte en ligne.[30]

En raison de la nature globale de la profession de la vérification des faits, des initiatives ont vu le jour pour créer des normes internationales et des meilleures pratiques. L'International Fact-Checking Network (IFCN) a d'abord élaboré un code de principes accompagné d'une structure de gouvernance. Les signataires du code s'engagent à adhérer à cinq principes : impartialité et équité, transparence des sources, transparence du financement et de l'organisation, transparence de la méthodologie et politique de correction ouverte et honnête.[31] Par ailleurs, le réseau européen des normes de vérification des faits (EFCSN) a publié un code d'intégrité professionnelle spécifiquement destiné aux vérificateurs de faits européens, élaboré en collaboration avec des organisations de vérification des faits et approuvé par elles. Le code définit les projets européens comme ceux qui ont un "intérêt substantiel et démontrable pour un (ou plusieurs) des pays du Conseil de l'Europe, plus le Kosovo ou la Biélorussie".[32] Il comprend des lignes directrices en matière de conformité et de gouvernance. Le code met l'accent sur d'importantes normes de bonnes pratiques, telles que des méthodologies claires et transparentes de vérification des faits, l'établissement de normes éthiques et la transparence du financement et des pratiques organisationnelles. Dans l'ensemble, la mise en place de ces codes et principes directeurs signifie le développement d'identités professionnelles structurelles au sein de la communauté de la vérification des faits, qui se distingue davantage des pratiques traditionnelles de la vérification des faits journalistique.

Les codes et lignes directrices émergents en matière de vérification des faits visent à préserver l'indépendance des modèles d'activité de vérification des faits (à la fois en tant que projets indépendants et dans le cadre d'organisations médiatiques traditionnelles), en particulier à une époque où la vérification et l'établissement de la vérité sont de plus en plus politisés. Ces initiatives ascendantes sont cruciales car un nombre croissant d'institutions gouvernementales et internationales s'investissent également dans la lutte contre les campagnes de désinformation à l'échelle nationale et internationale. Les codes et lignes directrices reflètent une approche ascendante, axée sur le terrain, qui va à l'encontre des initiatives des plateformes visant à établir leurs propres lignes directrices en matière de modération du contenu et de vérification des faits. Les initiatives des plateformes se concentrent souvent sur la visibilité et la facilité de trouver le contenu plutôt que sur la vérification elle-même.[33] Par conséquent, l'élaboration de codes et de lignes directrices indépendants garantit que les efforts de vérification des faits restent consacrés à la vérification et conservent leur intégrité malgré les pressions politiques et les intérêts et modèles commerciaux des plateformes.

---

[29] Graves, L. (2018), "Boundaries not drawn : Mapping the institutional roots of the global fact-checking movement", *Journalism studies*, 19(5), pp. 613-631.

[30] Nieminen, S., & Sankari, V. (2021), "Checking PolitiFact's fact-checks", *Journalism Studies*, *22*(3), pp. 358-378.

[31] International Fact-Checking Network (IFCN), Code of Principles, disponible à l'adresse : https://ifcncodeofprinciples.poynter.org/know-more/the-commitments-of-the-code-of-principles, consulté le 28 septembre 2023.

[32] European Fact-Checking Standards Network (EFCSN), Code européen de normes pour les organisations indépendantes de vérification des faits, Art. 5.2, C, disponible sur : https://eufactcheckingproject.com/app/uploads/2022/10/EU-CODE-EFCSN-.pdf ; consulté le 28 septembre.

[33] Cavaliere, P. (2020), "From journalistic ethics to fact-checking practices : defining the standards of content governance in the fight against disinformation", *Journal of media law*, 12(2), pp. 133-165.

Enfin, lorsqu'il s'agit de relever les défis liés aux conditions de travail des vérificateurs de faits, il peut être utile de s'appuyer sur la Recommandation CM/Rec(2022)4 du Conseil de l'Europe sur la promotion d'un environnement favorable à un journalisme de qualité à l'ère numérique.

3. Solutions pour la conception des plateformes

Cette section explique le contexte, le raisonnement et les principes qui sous-tendent les recommandations sur les solutions de conception des plateformes contenues dans la Note d'orientation ; elle sert de point de référence pour les normes pertinentes en matière de droits de l'homme et les meilleures pratiques lors de la mise en œuvre des recommandations.

a. Principaux défis

Il est important de commencer par reconnaître les défis considérables que la Note d'orientation cherche à relever en ce qui concerne les solutions de conception des plateformes pour lutter contre la désinformation. À cet égard, le manque de transparence, de responsabilité et d'explication des systèmes algorithmiques des plateformes qui affectent la diffusion de la désinformation figure parmi les principaux défis à relever. En effet, ces défis spécifiques ont été soulignés par les organes européens et internationaux de défense des droits de l'homme, y compris le Comité des Ministres, et par divers rapports du Conseil de l'Europe. Il a été spécifiquement décrit comment le manque de transparence et d'accès aux données sont des "défaillances majeures" des plateformes dans "presque toutes les préoccupations liées à la désinformation et à la mésinformation", ce qui empêche un examen indépendant et sape la responsabilité et la confiance.[34] Le Comité des Ministres a souligné la nécessité de garantir la transparence, la responsabilité et l'explicabilité des systèmes algorithmiques des plateformes, et de renforcer la responsabilité de ceux qui "développent et mettent en œuvre" ces systèmes. [35]

Il a également été souligné que le manque de transparence de la modération des contenus et des systèmes algorithmiques des plateformes rend "impossible" l'évaluation de "l'efficacité" des mesures adoptées par les plateformes pour lutter contre la désinformation et ces mesures "ont un impact sur les droits de l'homme".[36] Cela est d'autant plus difficile que les chercheurs indépendants, la société civile, les journalistes et les régulateurs indépendants n'ont pas accès aux données nécessaires pour garantir des évaluations indépendantes et objectives. Il est absolument essentiel que les réponses des plateformes à la désinformation soient fondées sur la transparence et la responsabilité.[37] Ainsi, les recommandations contenues dans la Note d'orientation sur les solutions de conception des plateformes visent à répondre spécifiquement aux défis liés au manque de transparence et de responsabilité, et à la préoccupation centrale selon laquelle la transparence est essentielle pour évaluer l'efficacité des mesures adoptées par les plateformes pour lutter contre la mésinformation et la désinformation.

b. Situation actuelle : normes en matière de droits de l'homme et la régulation de la désinformation

[34] Rapport du rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, "Désinformation et liberté d'opinion et d'expression", (A/HRC/47/25), 13 avril 2021, paragraphe 80.
[35] Recommandation CM/Rec(2022)11 du Comité des Ministres aux Etats membres sur les principes de gouvernance des médias et de la communication, pp. 6-7.
[36] Rapport du rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, "Désinformation et liberté d'opinion et d'expression", (A/HRC/47/25), 13 avril 2021, paragraphe 81.
[37] Voir, par exemple, la Recommandation CM/Rec(2022)13 sur l'impact des technologies numériques sur la liberté d'expression, annexe (pour les normes relatives à la garantie de la responsabilité et de la transparence).

Les dispositions de la Note d'orientation relatives aux solutions de conception des plateformes pour lutter contre la désinformation reposent sur un point essentiel : toute mesure réglementaire doit être compatible avec le droit à la liberté d'expression, garanti par l'article 10 de la Convention européenne des droits de l'homme et par le droit international en matière de droits de l'homme. Des organes tels que la Commissaire aux droits de l'homme du Conseil de l'Europe ont fait part de leurs préoccupations concernant les mesures mises en œuvre pour lutter contre la désinformation, notamment la réglementation autorisant le retrait et le blocage de contenus considérés comme de "fausses informations".[38] La Commissaire a recommandé que les mesures de lutte contre la désinformation ne conduisent jamais à ce que des contenus soient "indûment bloqués" en ligne.[39] En tant que telles, les mesures visant les plateformes en particulier doivent être pleinement compatibles avec les normes de liberté d'expression relatives à la désinformation.

Il convient de noter que la Cour européenne des droits de l'homme a examiné des cas où des personnes ont été poursuivies pour "diffusion d'informations fausses", la Cour établissant le principe fondamental selon lequel "l'article 10 de la Convention en tant que tel n'interdit pas la discussion ou la diffusion d'informations reçues même si l'on soupçonne fortement que ces informations pourraient ne pas être véridiques".[40] Un principe similaire a été établi dans le cadre de la législation internationale sur les droits de l'homme. Quatre mandats spéciaux internationaux sur la liberté d'expression ont publié une Déclaration commune sur la désinformation, dans laquelle ils affirment que les interdictions de diffusion d'informations fondées sur des concepts "vagues et ambigus" tels que les "fausses nouvelles" sont "incompatibles" avec les normes internationales en matière de liberté d'expression.[41] En effet, le Rapporteur spécial des Nations unies sur la liberté d'expression a déclaré que "la pénalisation de la désinformation est disproportionnée", tandis que le concept de désinformation lui-même est un "concept extraordinairement insaisissable" à définir en droit, et "susceptible de donner aux autorités exécutives un pouvoir discrétionnaire excessif pour déterminer ce qui est de la désinformation, ce qui est une erreur, ce qui est de la vérité".[42] En tant que telles, toutes les mesures réglementaires qui ciblent la conception des plateformes sur la désinformation doivent être cohérentes avec ces principes de liberté d'expression, et les dispositions de la Note d'orientation sont également étayées par ces normes.

c. Relever les défis : des solutions de conception de plateformes fondées sur les droits de l'homme

Il est essentiel que les dispositions de la Note d'orientation s'appuient sur un nombre considérable de normes importantes du Conseil de l'Europe afin de garantir que les solutions de conception de plateformes soient pleinement ancrées dans le droit relatif aux droits de l'homme. À cet égard, un certain nombre de principes essentiels sous-tendent ces recommandations.

---

[38] Commissaire aux droits de l'homme du Conseil de l'Europe, La liberté de la presse ne doit pas être compromise par des mesures visant à lutter contre la désinformation au sujet du COVID-19, 3 avril 2020.
[39] Ibid.
[40] *Salov c. Ukraine*, Requête n° 65518/01, arrêt du 6 septembre 2005, paragraphe 113. Voir Ó Fathaigh, R., Helberger, N., & Appleman N. (2021), "The perils of legally defining disinformation", 10(4) *Internet Policy Review* 1. Voir également Nenandic, I. & Verza, S. (2022), "European Policymaking on Disinformation and the Standards of the European Court of Human Rights", in Psychogiopoulou E. et De la Sierra S. (Eds.), *Digital Media Governance and Supranational Courts, Selected Issues and Insights from the European Judiciary*, Edward Elgar Publishing, pp. 175-198.
[41] Rapporteur spécial des Nations unies sur la liberté d'opinion et d'expression, représentant de l'Organisation pour la sécurité et la coopération en Europe sur la liberté des médias, Rapporteur spécial de l'Organisation des États américains sur la liberté d'expression et rapporteur spécial de la Commission africaine des droits de l'homme et des peuples sur la liberté d'expression et l'accès à l'information, Déclaration conjointe sur la liberté d'expression et les "fausses nouvelles", la désinformation et la propagande (FOM.GAL/3/17), 3 mars 2017, paragraphe 2(a).
[42] *Les pandémies et la liberté d'opinion et d'expression*, Rapport du rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, A/HRC/44/49, paragraphe 42.

Tout d'abord, un principe fondamental qui ressort des dispositions relatives aux solutions de conception des plateformes (et de la Note d'orientation dans son ensemble) est qu'il est essentiel de garantir un environnement favorable et propice à la liberté d'expression pour lutter contre la désinformation. En effet, la Cour européenne des droits de l'homme a estimé que les États ont l'obligation positive (c'est-à-dire le devoir) de créer un tel environnement favorable à la liberté d'expression.[43] Il est important de noter qu'en vertu des normes relatives aux droits de l'homme, la création d'un environnement propice à la liberté d'expression, avec un environnement de communication diversifié, y compris la diversité des médias, est un "moyen essentiel de lutter contre la désinformation".[44] Il est essentiel que les plateformes appliquent les normes relatives aux droits de l'homme à leurs systèmes, et la Recommandation CM/Rec(2018)2 du Comité des Ministres sur les rôles et les responsabilités des intermédiaires de l'internet reste un instrument normatif de référence que les plateformes devraient suivre, et elle contient des dispositions détaillées sur la manière de rendre les droits de l'homme opérationnels dans toutes les actions des plateformes.[45]

Deuxièmement , la sécurité dès la conception est essentielle pour la conception algorithmique, ce qui a également été souligné par le Comité des Ministres dans la Recommandation CM/Rec(2020)1 sur les incidences des systèmes algorithmiques sur les droits de l'homme.[46] Ce principe est crucial pour prévenir et atténuer le risque d'effets négatifs de la désinformation sur les individus et la société, en particulier la désinformation ciblant certains groupes et les campagnes de désinformation ayant un impact sur la sécurité en ligne des journalistes, ce qui a été documenté dans des études importantes élaborées pour l'UNESCO et le Centre international des journalistes, entre autres.[47]

[43] Voir, par exemple, *OOO Memo c. Russie*, Requête n° 2840/10, arrêt du 15 mars 2022, paragraphe 9.
[44] Déclaration conjointe sur la liberté d'expression et les "fake news", la désinformation et la propagande, (FOM.GAL/3/17), 3 mars 2017, paragraphe 3(a).
[45] Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires de l'internet.
[46] Recommandation CM/Rec(2020)1 sur les impacts des systèmes algorithmiques sur les droits de l'homme, annexe, B3.2.
[47] Voir, par exemple, Posetti, J., Bontcheva, K., & Shabbir, N. (2022), *The Chilling : Assessing Big Tech's Response to Online Violence Against Women Journalists*, UNESCO, disponible sur : https://unesdoc.unesco.org/ark:/48223/pf0000383044.locale=en, consulté le 28 septembre 2023 ; International Center for Journalists, " How Disinformation Fuels Online Violence Storms Targeting Women Journalists ", 14 février 2023. disponible sur : https://www.icfj.org/news/how-disinformation-fuels-online-violence-storms-targeting-women-journalists, consulté le 28 septembre 2023.

Troisièmement, le principe de proportionnalité sous-tend également les dispositions relatives aux solutions de conception des plateformes, qui est un principe fondamental en vertu de l'article 10 de la Convention et de la jurisprudence de la Cour européenne des droits de l'homme. Il s'ensuit que lorsque les plateformes appliquent des mesures pour restreindre la désinformation, ces mesures doivent respecter les droits des utilisateurs, être mises en œuvre en utilisant les "moyens les moins restrictifs", de manière "transparente et non discriminatoire", et être limitées dans leur portée et leur durée à ce qui est strictement nécessaire pour éviter la "restriction ou le retrait collatéral injustifié d'un contenu légal".[48] À cet égard, il a été noté que les mesures anti-désinformation de certaines plateformes ont entraîné des restrictions sur les contenus d'information indépendants.[49] Lors de la mise en œuvre de solutions de conception pour la désinformation, il convient de veiller tout particulièrement à ce qu'il n'y ait pas d'interférences disproportionnées avec la liberté des médias. Les récentes mesures réglementaires européennes, telles que le règlement sur les services numériques de l'UE (DSA), visent également à garantir que les plateformes appliquent les conditions générales de manière "proportionnée" et compatible avec la liberté d'expression et la liberté des médias, ce qui est également conforme aux recommandations du Conseil de l'Europe.[50]

L'accent mis sur les garanties procédurales est particulièrement important. Cet accent est cohérent avec les normes internationales en matière de droits de l'homme applicables à la lutte contre la désinformation, dans la mesure où la focalisation sur les obligations en matière de procédure peut apporter une « contribution positive » à la protection des droits de l'homme et à une « plus grande responsabilité publique des plates-formes », à la place d'une réglementation basée sur le point de vue ou le contenu.[51] Cet accent mis sur la garantie des droits procéduraux sous-tend également les récentes mesures réglementaires de l'UE, avec des dispositions de l'ASD obligeant les plateformes à motiver les restrictions de contenu et à mettre en place des systèmes internes de traitement des plaintes pour les utilisateurs.[52] Il est important de noter que la Cour européenne des droits de l'homme a également estimé que l'article 10 comportait d'importantes garanties procédurales, notamment que les utilisateurs doivent pouvoir connaitre les motifs du blocage d'un contenu et disposer d'un forum pour contester une restriction à la liberté d'expression.[53]

[48] Recommandation CM/Rec(2018)2 sur les rôles et responsabilités des intermédiaires de l'internet, annexe, paragraphes 2.3.1 - 2.3.2.
[49] Rapport du rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression (A/HRC/38/35), 6 avril 2018, paragraphe 31.
[50] Voir le règlement (UE) 2022/2065 du Parlement européen et du Conseil du 19 octobre 2022 relatif à un marché unique des services numériques et modifiant la directive 2000/31/CE (loi sur les services numériques), article 14. Voir également Quintais, J.P, Appelman, N., & Ó Fathaigh, R. (2023), "Using Terms and Conditions to Apply Fundamental Rights to Content Moderation", *German Law Journal*, 24(5), pp. 881-911.
[51] Rapport de la Rapporteuse spéciale sur la promotion et la protection du droit à la liberté d'opinion et d'expression, "Désinformation et liberté d'opinion et d'expression", (A/HRC/47/25), 13 avril 2021, paragraphe 59.
[52] Voir le Règlement (UE) 2022/2065 du Parlement européen et du Conseil du 19 octobre 2022 sur un marché unique des services numériques et modifiant la directive 2000/31/CE (loi sur les services numériques), articles 17 et 20.
[53] Voir, par exemple, *Vladimir Kharitonov c. Russie*, Requête n° 10795/14, arrêt du 23 juin 2022, paragraphe 36.

Enfin, les recommandations de la Note d'orientation sur la transparence et la responsabilité sont essentielles. Comme indiqué plus haut, le manque de transparence et d'accès aux données a été décrit comme un manquement crucial de la part des plateformes en ce qui concerne la désinformation. Il s'agit notamment du manque de données spécifiques sur l'engagement des utilisateurs à l'égard des la mésinformation et de la désinformation, du manque de transparence des systèmes automatisés utilisés pour identifier et supprimer la mésinformation et la désinformation, et d'un manque général de transparence concernant les systèmes de modération du contenu des plateformes, ce qui rend presque impossible l'évaluation de l'efficacité ou de l'impact des mesures prises par les plateformes en matière de mésinformation et de désinformation.[54] Là encore, les récentes politiques européennes, telles que l'ASD, sont fermement ancrées dans la mise en place de cadres visant à garantir la transparence des systèmes des plateformes. En outre, l'accent mis sur la transparence est également cohérent avec les normes internationales en matière de droits de l'homme dans l'approche de la lutte contre la désinformation.[55] Les recommandations sur la transparence et la responsabilité contenues dans la Note d'orientation doivent être lues à la lumière de la Note d'orientation sur les meilleures pratiques pour des cadres juridiques et procéduraux efficaces pour les mécanismes d'autorégulation et de corégulation de la modération de contenu, adoptée par le Comité directeur sur les médias et la société de l'information (CDMSI), qui contient des dispositions utiles sur la garantie de la transparence.[56] L'objectif de la Note d'orientation est de contribuer à la mise en œuvre de l'appel du Comité des ministres à garantir "la transparence, la responsabilité et l'explicabilité" des systèmes algorithmiques des plateformes.[57]

d. Priorité aux sources d'information professionnelles et aux contenus d'intérêt public

[54] Rapport du rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, "Désinformation et liberté d'opinion et d'expression", (A/HRC/47/25), 13 avril 2021, paragraphe 81.
[55] Rapport du rapporteur spécial sur la promotion et la protection du droit à la liberté d'opinion et d'expression, (A/HRC/38/35), 6 avril 2018, paragraphe 66.
[56] Note d'orientation sur la modération de contenu : Meilleures pratiques en matière des cadres juridiques et procéduraux efficaces pour les mécanismes d'autorégulation et de corégulation de la modération de contenu, adoptée par le Comité directeur pour les médias et la société de l'information, juin 2021.
[57] Recommandation CM/Rec(2022)11 du Comité des Ministres aux Etats membres sur les principes de gouvernance des médias et de la communication, pp. 6-7.

Des recommandations notables ont été formulées sur le rôle de la priorisation des sources d'information professionnelles et du contenu d'intérêt public en tant qu'outil spécifique pour les plateformes afin de lutter contre la mésinformation et la désinformation. En effet, l'Assemblée parlementaire du Conseil de l'Europe, dans sa Résolution 2255(2019) sur les médias de service public dans le contexte de la désinformation et de la propagande, a spécifiquement recommandé que les plateformes en ligne coopèrent avec les organes de presse européens publics et privés afin d'améliorer la "visibilité des informations fiables et dignes de confiance et de faciliter l'accès des utilisateurs à ces informations".[58] De même, quatre mandats spéciaux internationaux sur la liberté d'expression et la liberté des médias ont publié une Déclaration commune en 2023, avec une recommandation spécifique pour les grandes plateformes en ligne de "privilégier les médias indépendants de qualité et les contenus d'intérêt public sur leurs services afin de faciliter le discours démocratique".[59] Le Comité des Ministres a reconnu que certaines plateformes ont fait "des efforts pour donner plus d'importance à des sources d'information généralement fiables" ; cependant, il a également souligné que "l'impact de ces mesures sur la libre circulation de l'information et des idées dans les sociétés démocratiques doit être étudié avec soin".[60]

Il est essentiel que les critères d'identification et de labellisation soient élaborés dans le cadre d'un processus transparent et multipartite. En effet, certains États membres ont mis en place des régimes de prééminence, et il existe des recherches sur ces mécanismes, y compris ceux qui s'appliquent aux médias de service public.[61] Les États et les plateformes devraient notamment s'inspirer d'exemples tels que l'Initiative pour la confiance dans le journalisme de Reporters sans frontières.[62] À cet égard, lors de la mise en œuvre de la présente Note d'orientation, toute mesure devrait être mise en œuvre conformément à la Note d'orientation détaillée et complète sur la hiérarchisation des contenus d'intérêt public en ligne, adoptée par le CDMSI en 2021.[63] En particulier, les États devraient être juridiquement empêchés d'obliger les plateformes à mettre en avant leurs propres déclarations et communications, sauf en cas d'urgence publique telle que définie par l'article 15 de la Convention et interprétée dans la jurisprudence de la Cour européenne des droits de l'homme.[64]

e. Recherche indépendante

[58] Assemblée parlementaire, Résolution 2255 (2019) sur les médias de service public dans le contexte de la désinformation et de la propagande, 23 janvier 2019, paragraphe 8.2.
[59] Rapporteur spécial des Nations Unies sur la promotion et la protection du droit à la liberté d'opinion et d'expression, Représentant de l'Organisation pour la sécurité et la coopération en Europe pour la liberté des médias, Rapporteur spécial de l'Organisation des États américains sur la liberté d'expression et Rapporteur spécial de la Commission africaine des droits de l'homme et des peuples sur la liberté d'expression et l'accès à l'information en Afrique, Déclaration conjointe sur la liberté des médias et la démocratie, 2023, p. 8.
[60] Recommandation CM/Rec(2022)4 sur la promotion d'un environnement favorable au journalisme de qualité à l'ère numériqueannexe, paragraphe A6.

[61] Voir, par exemple, Cole M. D. et al. (Eds), (2022), *Prominence of European works and of services of general interest*, IRIS Special, Observatoire européen de l'audiovisuel.
[62] Initiative du fonds pour le journalisme, https://www.journalismtrustinitiative.org.
[63] Note d'orientation sur la hiérarchisation des contenus d'intérêt public en ligneadoptée par le Comité directeur pour les médias et la société de l'information, CDMSI(2021)009, 2 décembre 2021.
[64] Ibid, paragraphe 21.

Enfin, les dispositions relatives à la recherche indépendante s'appuient sur le principe primordial selon lequel l'élaboration de politiques efficaces en matière de solutions de conception de plateformes pour lutter contre la désinformation requiert des "connaissances précises, nuancées et complètes" fondées sur des "recherches rigoureuses et indépendantes".[65] Les décideurs politiques en Europe, tels que l'ASD de l'UE, mettent en œuvre des cadres pour faciliter (et obliger) l'accès aux données de certaines plateformes en ligne pour une recherche indépendante sur les aspects des systèmes des plateformes.[66] À cet égard, en mettant en œuvre les dispositions de la Note d'orientation relatives la garantie de la recherche indépendante, les États devraient suivre les dispositions détaillées de la section 6 de la recommandation CM/Rec(2022)13 du Comité des ministres sur les incidences des technologies numériques sur la liberté d'expression. Ces dispositions portent notamment sur (i) la garantie que les chercheurs peuvent accéder aux données détenues par les plateformes de manière sécurisée, légale et conforme à la vie privée ; (ii) le rôle des autorités compétentes dans la création d'environnements sécurisés facilitant la recherche ; (iii) l'accès aux données individuelles disponibles pour la recherche indépendante ; (iv) le filtrage ; (v) la responsabilité ; et les accords de partage de données entre les plateformes et les chercheurs. Le rapport du Groupe de travail sur l'accès aux données entre plateformes et chercheurs de l'Observatoire européen des médias numériques est une excellente source d'informations sur la manière dont les plateformes peuvent partager leurs données avec les chercheurs.[67]

4. Autonomisation des utilisateurs

Cette section explique le contexte, le raisonnement et les principes qui sous-tendent les recommandations sur l'autonomisation des utilisateurs contenues dans la Note d'orientation.

a. Situation actuelle

Le renforcement de l'intégrité de l'information en ligne nécessite une stratégie globale qui va au-delà de la limitation de l'exposition à la mésinformation et à la désinformation. Dans les sociétés ouvertes et démocratiques, il est difficile d'empêcher complètement la production et la diffusion d'informations de mauvaise qualité sans mettre en péril la liberté d'expression. Si les autres piliers du présent exposé des motifs mettent en évidence les moyens par lesquels l'écosystème médiatique contemporain peut favoriser la circulation d'informations de qualité et limiter la diffusion de contenus de qualité médiocre, l'impact de ces informations dépend en fin de compte de la question de savoir si les utilisateurs (a) sont régulièrement exposés à des contenus de qualité et en reconnaissent l'importance ; (b) sont capables de faire la distinction entre un contenu de haute qualité et un contenu de faible qualité ; (c) agissent de manière responsable envers les autres en partageant et en discutant des différents types d'informations qu'ils peuvent rencontrer ; et (d) bénéficient de solides protections pour leurs droits de l'homme, savent comment les exercer et sont convaincus qu'ils peuvent faire une différence positive, tout en se protégeant eux-mêmes, en les exerçant.

[65] Recommandation CM/Rec(2022)13 sur les impacts des technologies numériques sur la liberté d'expressionpréambule.
[66] Voir, par exemple, la loi sur les services numériques, article 40 (accès aux données et contrôle).
[67] EDMO, (2022), *Rapport du groupe de travail de l'Observatoire européen des médias numériques sur l'accès aux données de plateforme à chercheur*.

En ce sens, le rôle des médias de service public en Europe est plus pertinent que jamais. En remplissant leur mission de fournir des informations exactes et objectives librement accessibles à tous les secteurs de la société, les médias de service public peuvent faire en sorte que la plupart des individus aient régulièrement accès à des informations fiables. Les modèles économiques des médias commerciaux ayant été remis en question par les changements technologiques et sociaux, le rôle central des médias de service public en tant que « norme d'excellence » dans l'information et l'autonomisation du public devient encore plus crucial.[68]

L'autonomisation des utilisateurs est particulièrement pertinente lorsque la communication en ligne s'effectue sur des plateformes privées ou semi-privées, telles que WhatsApp, Facebook Messenger, Snapchat et Telegram, et via des services de messagerie tels qu'iMessage et Android Message. Ces applications sont très populaires et constituent un élément important de la vie sociale quotidienne, car elles sont utilisées pour maintenir des liens avec des amis proches et des membres de la famille, des connaissances plus éloignées et, au moins dans certains cas, des groupes plus importants composés principalement d'étrangers qui s'intéressent à un certain sujet ou vivent dans une certaine région. Étant donné que la communication sur ces applications est privée et souvent cryptée de bout en bout, les possibilités d'appliquer des solutions de conception et des interventions fondées sur la technologie pour lutter contre la propagation de la mésinformation de la désinformation sont beaucoup plus limitées que sur les plateformes publiques de médias sociaux. Par ailleurs, il est prouvé que les utilisateurs de messageries personnelles sont réticents à contester et à corriger les personnes qui diffusent de la désinformation sur ces applications, car ils craignent de provoquer un conflit avec des personnes avec lesquelles ils entretiennent des relations étroites de parenté ou d'amitié, ou de les mettre dans l'embarras vis-à-vis d'autres personnes lorsque l'interaction a lieu en groupe.[69] Par conséquent, il est encore plus important sur les messageries privées que sur les médias sociaux publics de donner aux utilisateurs les moyens de discerner entre les vraies et les fausses informations, d'être responsables de ce qu'ils partagent et d'être capables d'avoir des discussions respectueuses et constructives avec ceux qui peuvent partager des contenus de qualité médiocre.

Si des informations de mauvaise qualité circulant en ligne atteignent des utilisateurs conscients et habilités, les dommages potentiels résultant de cette exposition seront probablement minimes. Ces utilisateurs seront en mesure d'identifier les marqueurs qui caractérisent généralement les messages faux ou inauthentiques, seront capables de les vérifier, seront moins susceptibles de les partager, seront mieux préparés à corriger les faussetés qu'ils contiennent en s'adressant à la personne qui les a partagés d'une manière appropriée au contexte social de l'interaction, et seront au courant des outils à leur disposition pour signaler ce contenu à la plateforme où il circule, contribuant ainsi à la capacité de la plateforme à reconnaître et à réduire la propagation des faussetés. Ainsi, l'autonomisation des utilisateurs pourrait faciliter un cercle vertueux dans lequel les utilisateurs non seulement éviteraient de tomber dans le piège des contenus de mauvaise qualité, mais protégeraient et alerteraient également les autres utilisateurs susceptibles de le faire, tout en améliorant la responsabilité et la capacité des plateformes numériques à lutter contre la propagation de la mésinformation et de la désinformation.

[68] Voir les recommandations de l'Office britannique des communications sur l'avenir des médias de service public, disponibles à l'adresse suivante : https://www.smallscreenbigdebate.co.uk/data/assets/pdf_file/0023/221954/statement-future-of-public-service-media.pdf, consulté le 26 septembre 2023.

[69] Chadwick, A., Vaccari, C., & Hall, N. A. (2023), "What Explains the Spread of Misinformation in Online Personal Messaging Networks ? Exploring the Role of Conflict Avoidance", *Digital Journalism*, pp. 1-20.

Lorsque ces conditions sont réunies, les individus et les sociétés développent une résilience face à la désinformation. Dans les environnements informationnels turbulents, et en particulier pendant les crises, la résilience est une ressource clé contre les dommages individuels et sociétaux qui peuvent résulter de la diffusion de fausses informations.

Au niveau individuel, on peut s'attendre à ce que la résilience à la mésinformation et à la désinformation soit plus élevée chez les citoyens qui ont un niveau d'instruction plus élevé, qui s'intéressent à l'actualité, qui ont une bonne maîtrise du numérique et de l'information, qui sont familiarisés avec la technologie et qui ont accès à des informations de qualité, entre autres. Les interventions politiques visant à renforcer la résilience au niveau individuel devraient donc se concentrer principalement sur l'amélioration de ces caractéristiques chez les individus et les groupes qui en sont dépourvus.

Nous pouvons également considérer la résilience face à la mésinformation et à la désinformation comme une ressource sociétale. De ce point de vue, la résilience peut être considérée comme le résultat de conditions politiques, médiatiques et économiques. Dans le système politique, des niveaux plus élevés de polarisation politique et des partis et dirigeants populistes de premier plan peuvent diminuer la résilience. Les acteurs étrangers et nationaux peuvent également diminuer la résilience en ciblant activement un pays par la désinformation et des campagnes visant à cultiver la méfiance à l'égard des experts et des journalistes. Les systèmes médiatiques dans lesquels les citoyens ont peu confiance en l'information, les médias de service public sont faibles et les audiences sont très fragmentées, peuvent également favoriser l'émergence de sociétés moins résilientes. En ce qui concerne les conditions économiques, les conditions commerciales dans lesquelles les organismes de presse opèrent, la taille du marché publicitaire d'un pays et les niveaux d'utilisation des médias sociaux peuvent réduire la résilience.[70] Les politiques visant à renforcer la résilience face à la désinformation au niveau sociétal devraient s'attaquer à certains des facteurs systémiques qui facilitent son développement.

b. Principaux défis

À tout moment, dans une société donnée, certains utilisateurs bénéficieront déjà de niveaux élevés d'autonomisation, tandis que d'autres n'en bénéficieront pas. Les processus par lesquels les êtres humains prennent conscience des problèmes, acquièrent de nouvelles compétences et développent de nouvelles habitudes sont complexes et nécessitent de la motivation, du temps et des ressources. En outre, des groupes différents bénéficient de niveaux d'autonomisation différents et nécessitent des approches spécifiques pour les renforcer. Une distinction essentielle est faite entre les groupes qui peuvent payer pour accéder à des informations de qualité et ceux qui ne le peuvent pas. Une autre distinction est celle qui existe entre les jeunes et les utilisateurs plus âgés des médias numériques : les premiers ont tendance à posséder des compétences techniques plus élevées mais à s'intéresser moins à l'actualité et à la politique, tandis que l'inverse s'applique aux seconds. D'autres divisions sociétales (fondées sur la religion, l'appartenance ethnique, le sexe et l'éducation, entre autres) façonnent également les conditions dans lesquelles les individus peuvent bénéficier des possibilités offertes par l'écosystème de l'information. Ces complexités sont à l'origine de certains défis majeurs pour l'autonomisation des utilisateurs face à la mésinformation et à la désinformation.

[70] Humprecht, E., Esser, F., & Van Aelst, P. (2020), "Resilience to Online Disinformation : A Framework for Cross-National Comparative Research", *The International Journal of Press/Politics*, 25(3), pp. 493-516.

Premièrement, l'autonomisation devrait être considérée comme un objectif à moyen et long terme, avec une dépendance limitée à l'égard des incitations et des solutions immédiates ou à court terme en cas de crise et d'élections litigieuses en situation d'urgence. Deuxièmement, les politiques visant à renforcer l'autonomie des utilisateurs nécessitent des investissements à long terme de la part des gouvernements, une réforme des programmes d'enseignement (y compris l'apprentissage tout au long de la vie et la formation continue), ainsi qu'une vaste collaboration entre diverses parties prenantes. Troisièmement, toute initiative visant à renforcer l'autonomie doit partir des utilisateurs et des communautés, en s'adaptant à leurs besoins spécifiques et aux conditions contextuelles, en particulier en ce qui concerne les groupes les plus vulnérables et les plus déconnectés (à commencer par les enfants et les jeunes adultes). Les programmes qui établissent des liens significatifs avec les communautés qu'ils visent à servir, notamment par l'élaboration en commun des programmes et des interventions, ont plus de chances de réussir que les programmes qui adoptent une approche unique et descendante. Le recours à des systèmes de type ombudsman pour traiter de manière indépendante les griefs des utilisateurs de manière flexible et impartiale pourrait également contribuer à sensibiliser les différentes communautés et à sauvegarder leurs droits. Quatrièmement, de nombreuses initiatives visant à renforcer l'autonomie, telles que les programmes d'alphabétisation, n'atteignent pas certains des groupes les plus vulnérables, tels que les personnes âgées et celles qui n'ont pas accès aux technologies numériques.[71] L'engagement limité des principales plateformes et les inégalités substantielles entre les États membres en ce qui concerne la qualité, la diversité et l'accessibilité des programmes d'éducation en la matière sont d'autres défis majeurs qui doivent être relevés.[72] Enfin, comme les attitudes, les compétences et les comportements humains sont complexes et difficiles à modifier, les effets des programmes et des interventions d'autonomisation des utilisateurs sont difficiles à mesurer et à démontrer. Ce problème est aggravé par le fait que les initiatives dans ce domaine, par exemple les programmes d'éducation au numérique et aux médias, renoncent souvent à l'étape cruciale de l'évaluation de leurs effets, en partie à cause du manque de financement et de la brièveté des délais. En outre, les programmes qui entreprennent des évaluations solides de leurs effets ne diffusent pas toujours les résultats de ces exercices, ce qui rend difficile le développement de connaissances cumulatives sur ces questions.[73]

L'un des éléments clés de l'autonomisation des utilisateurs est l'esprit critique, c'est-à-dire la tendance à remettre en question et à chercher à vérifier les informations que l'on rencontre, et à être prêt à toujours réfléchir et remettre en question ses croyances à la lumière de preuves nouvelles et fiables. L'esprit critique est une caractéristique souhaitable de la citoyenneté démocratique car il favorise la vigilance et la responsabilité des utilisateurs, réduit le risque que la désinformation se propage rapidement dans les réseaux sociaux en ligne et hors ligne et, en fin de compte, dissuade les élites politiques et médiatiques de produire et de diffuser des faussetés, étant donné qu'elles sont plus susceptibles d'être signalées à l'ordre et sanctionnées.

[71] Voir par exemple l'aperçu des programmes basés au Royaume-Uni dans le document du ministère britannique du numérique, de la culture, des médias et du sport intitulé Stratégie d'éducation aux médias en ligne.
[72] Cabrera Blázquez, F., Cappello M., Talavera Milla J., Valais S. (2022), *User empowerment against disinformation online,* Observatoire européen de l'audiovisuel .
[73] L'Office des Communications du Royaume-Uni a développé une boîte des outils pour évaluer les interventions en matière d'éducation aux médias, disponible à l'adresse suivante : https://www.ofcom.org.uk/research-and-data/media-literacy-research/approach/evaluate/toolkit, consulté le 28 septembre 2023.

D'autre part, le discours public soulignant l'ampleur de la désinformation et les initiatives visant à sensibiliser les citoyens au problème peuvent également cultiver le cynisme, en particulier parmi certains groupes déjà déconnectés et méfiants à l'égard des institutions publiques et des médias.[74] Le cynisme est la croyance généralisée que les autres ne sont pas dignes de confiance et qu'il est impossible, voire indésirable, de faire la différence entre le vrai et le faux. Globalement, cette attitude est démocratiquement dysfonctionnelle parce qu'elle encourage une mentalité de "tout est permis" où personne ne se sent responsable des informations qu'il diffuse et de la correction des informations partagées par d'autres. Le cynisme est donc à l'opposé de l'autonomisation, car il encourage le manque de responsabilité et rejette l'idée que les utilisateurs ont un pouvoir sur l'environnement dans lequel ils communiquent. Il est important de noter que le sentiment que le cynisme est endémique dans une population peut également générer ce que les spécialistes appellent le "dividende du menteur",[75] qui permet aux acteurs politiques et autres accusés de mensonge de se défendre de ces accusations en affirmant qu'il est impossible d'établir la vérité de manière crédible. Exploiter et cultiver cette forme de cynisme est un principe clé de l'approche des régimes autoritaires en matière de propagande au XXIe siècle.[76]

Il est important que toute initiative visant à promouvoir l'éducation au numérique et aux médias et d'autres formes d'autonomisation des utilisateurs cultive l'esprit critique tout en évitant de renforcer le cynisme des citoyens. Le fait d'être exposé à un contenu inexact, ne serait-ce que pour montrer comment il peut être identifié comme trompeur,[77] ou à des discussions sur le problème et l'ampleur de la désinformation dans les écosystèmes médiatiques contemporains,[78] ou à la couverture médiatique de campagnes de désinformation à grande échelle,[79] peut conduire à une diminution des niveaux de confiance dans l'information et le processus démocratique, ce qui suggère le cynisme plutôt que l'esprit critique.

c. Relever les défis

[74] Boyd, D. (2017), Did media literacy backfire, *Journal of Applied Youth Studies*, 1(4), pp. 83-89.
[75] Chesney, B., & Citron, D. (2019), "Deep Fakes : A Looming Challenge for Privacy, Democracy, and National Security", *California Law Review*, 107(6), pp. 1753-1820.
[76] Pomerantsev, P. (2017), *Rien n'est vrai et tout est possible : Aventures dans la Russie moderne*. Faber & Faber.
[77] Vaccari, C., & Chadwick, A. (2020), "Deepfakes and disinformation : Exploring the impact of synthetic political video on deception, uncertainty, and trust in news ", *Social Media + Society*, 6(1), disponible sur : https://journals.sagepub.com/doi/full/10.1177/2056305120903408, consulté le 28 septembre 2023.
[78] Van Duyn, E., & Collier, J. (2019), "Priming and fake news : Les effets du discours de l'élite sur les évaluations des médias d'information", *Mass Communication and Society*, *22*(1), pp. 29-48.
[79] Ross, A. R., Vaccari, C., & Chadwick, A. (2022), "Russian meddling in US elections : How news of disinformation's impact can affect trust in electoral outcomes and satisfaction with democracy", *Mass Communication and Society*, 25(6), pp. 786-811.

Il existe au moins trois grandes catégories d'initiatives susceptibles de renforcer la capacité des utilisateurs à lutter contre la désinformation. Premièrement, les plateformes numériques peuvent fournir des outils et des services qui donnent aux utilisateurs les moyens de contrôler le contenu mis à leur disposition et recommandé,[80] de vérifier les sources et leur fiabilité,[81] et d'obtenir une réparation rapide, équitable et efficace s'ils estiment que leurs droits de l'homme ont été limités par les plateformes ou d'autres utilisateurs. Ces initiatives améliorent l'expérience des utilisateurs de plateformes numériques spécifiques en mettant à leur disposition des informations qu'ils n'auraient pas rencontrées autrement et qui peuvent les aider à prendre de meilleures décisions ou à protéger leurs droits. Toutefois, en l'absence de formes strictes de régulation ou de corégulation, ces initiatives dépendent des décisions des plateformes et la transparence sur la manière dont elles sont mises en œuvre est limitée.[82]

[80] Par exemple, Facebook offre aux utilisateurs divers moyens de contrôler le contenu qu'ils voient sur leur fil d'actualité (disponible à l'adresse : https://www.facebook.com/help/964154640320617, consulté le 28 septembre 2023) et des explications sur les raisons pour lesquelles une publicité particulière leur a été montrée (disponible à l'adresse : https://www.facebook.com/help/562973647153813?helpref=faq_content, consulté le 28 septembre 2023).

[81] Par exemple, dans certaines régions et langues, YouTube enrichit certains résultats de recherche de panneaux d'information présentant le contenu de vérificateurs de faits (disponible à l'adresse : https://support.google.com/youtube/answer/9229632?hl=en, consulté le 28 septembre 2023).

[82] L'engagement 22 du code de bonnes pratiques renforcé 2022 de l'Union européenne en matière de désinformation, qui a été signé par la plupart des entreprises technologiques qui ont adhéré au code, mais pas toutes, oblige les plateformes à fournir aux utilisateurs ce type d'outils. L'engagement 24 du code exige des plateformes qu'elles fournissent des informations adéquates sur les mesures prises à leur encontre et sur les mécanismes de recours contre ces décisions.

Deuxièmement, divers acteurs publics et privés peuvent concevoir des outils - y compris des plug-ins de navigateur,[83] des jeux éducatifs,[84] des guides pratiques,[85] et des chatbots d'assistance (agents conversationnels)[86] - qui aident les utilisateurs à comprendre la dynamique de la circulation de l'information en ligne, à vérifier le contenu qu'ils voient et à s'engager de manière constructive avec d'autres utilisateurs.[87] Ces outils peuvent aider les utilisateurs dans leur vie numérique de tous les jours, mais il faut que les utilisateurs soient conscients de leur existence et qu'ils soient prêts à s'engager avec eux. C'est pourquoi il est essentiel que les principales plateformes numériques collaborent pour promouvoir et diffuser ces outils auprès de leurs utilisateurs. L'Union européenne a soutenu le développement de plusieurs outils de ce type dans le cadre de ses programmes de lutte contre la désinformation.[88] L'un de ces outils est EDMO, dont l'approche holistique comprend la vérification des faits dans toutes les langues de l'Union européenne, la protection des événements sensibles dans l'UE tels que les élections, ainsi que la recherche et la vérification des engagements pris par les plateformes. Les acteurs publics peuvent également soutenir un écosystème de l'information sain et renforcer l'autonomie des utilisateurs en mettant à disposition, de leur propre initiative, les informations qu'ils détiennent, conformément aux exigences de la Recommandation Rec(2002)2 du Comité des Ministres aux États membres sur l'accès aux documents publics et de la Convention du Conseil de l'Europe sur l'accès aux documents publics (STCE n° 205 - la Convention de Tromsø).

[83] Par exemple, le plugin de vérification InVID aide les utilisateurs à vérifier le contenu de divers médias sociaux (disponible à l'adresse : https://www.invid-project.eu/tools-and-services/invid-verification-plugin/, consulté le 28 septembre 2023).

[84] Par exemple, "Bad News" (disponible à l'adresse : https://www.getbadnews.com/books/english/, consulté le 28 septembre 2023) aide les utilisateurs à comprendre les facteurs clés qui facilitent la propagation de la désinformation et "Go Viral" (disponible à l'adresse : https://www.goviralgame.com/en, , consulté le 28 septembre 2023) se concentre sur la désinformation liée à Covid.

[85] La Fédération internationale des associations de bibliothécaires et institutions des bibliothèques a élaboré des infographies qui résument les principales stratégies que les utilisateurs peuvent employer (disponible à : https://repository.ifla.org/handle/123456789/167, consulté le 28 septembre 2023). First Draft propose des ressources aux utilisateurs et aux journalistes pour la vérification des informations (disponible à l'adresse : https://firstdraftnews.org/training/, consulté le 28 septembre 2023). L'UNESCO a publié un manuel pour l'éducation et la formation au journalisme sur la manière d'identifier la désinformation. Ireton (C.), Posetti (J.), (Eds.), (2018), *Journalism, fake news & disinformation : handbook for journalism education and training*, disponible sur : https://unesdoc.unesco.org/ark:/48223/pf0000265552, consulté le 28 septembre 2023.

[86] Le "Fake News Immunity Chatbot" (disponible à l'adresse : http://fni.arg.tech/, consulté le 28 septembre 2023) aide les individus à identifier les arguments non valables dans différents types d'informations. Divers vérificateurs de faits, comme Maldita en Espagne, ont expérimenté des chatbots qui répondent automatiquement aux questions des utilisateurs sur la véracité des informations qu'ils rapportent (disponible à l'adresse : https://www.europeanpressprize.com/article/maldita-es-whatsapp-chatbot/, consulté le 28 septembre 2023).

[87] La RAND Corporation tient à jour une liste d'outils de lutte contre la désinformation en ligne (disponible à l'adresse suivante : https://www.rand.org/research/projects/truth-decay/fighting-disinformation/search.html, consulté le 28 septembre 2023).

[88] Une liste et une description des projets financés sont disponibles à l'adresse suivante : https://commission.europa.eu/strategy-and-policy/coronavirus-response/fighting-disinformation/funded-projects-fight-against-disinformation_en, consulté le 28 septembre 2023.

Troisièmement, l'éducation aux médias et à l'information est essentielle pour renforcer l'autonomie des utilisateurs. Il existe plusieurs définitions des différents types d'éducation : par exemple, le Comité des ministres, dans sa recommandation CM/Rec(2022)4 sur la promotion d'un environnement favorable à un journalisme de qualité à l'ère numérique, déclare que l'éducation aux médias et à l'information implique le développement de compétences et de capacités cognitives, techniques et sociales qui permettent aux personnes: (a) d'accéder efficacement au contenu des médias et d'analyser l'information de manière critique, ce qui leur permet de comprendre comment le contenu des médias est produit, financé et réglementé, ainsi que d'avoir la confiance et la compétence nécessaires pour prendre des décisions éclairées sur les médias qu'ils utilisent et la manière dont ils les utilisent ; (b) de comprendre les implications éthiques des médias et de la technologie ; et (c) de communiquer efficacement, y compris en interprétant, en créant et en publiant du contenu.[89] L'Office britannique des communications définit l' « éducation aux médias » comme « la capacité d'utiliser, de comprendre et de créer des médias et des communications dans divers contextes ».[90] L'UNESCO définit l' « éducation aux médias » et à comme « un ensemble de compétences interdépendantes qui aident les gens à maximiser les avantages et à minimiser les inconvénients dans les nouveaux paysages de l'information, du numérique et de la communication ».[91] Les universitaires ont défini la « maîtrise de l'information » comme une prise de conscience des cinq « C » relatifs à la manière dont l'information est produite, distribuée et comprise : le contexte (l'environnement dans lequel l'information existe), la création (le processus de production de l'information), le contenu (ce que l'information dit), la circulation (la manière dont l'information est distribuée) et la consommation (si le public y prête attention et comment il lui donne un sens).[92] Étant donné que tous ces types d'alphabétisation sont essentiels à la capacité des citoyens à se protéger contre la désinformation et à exercer leurs droits, leur inclusion dans les programmes d'enseignement à tous les niveaux est une première étape nécessaire.[93] Toutefois, à un moment donné, la plupart des membres d'une société ne sont pas inscrits dans un établissement d'enseignement formel. Il est donc tout aussi important que des possibilités d'améliorer ces compétences soient offertes aux personnes de tous horizons, en particulier à celles qui sont plus susceptibles d'être vulnérables et moins susceptibles d'entrer en contact avec des établissements d'enseignement.[94] Les médias de service public peuvent jouer un rôle crucial en proposant à ces utilisateurs des contenus permettant d'améliorer leur niveau de maîtrise en la matière.

[89] Recommandation CM/Rec(2022)4 sur la promotion d'un environnement favorable à un journalisme de qualité à l'ère numérique, annexe, paragraphe 9.

[90] Disponible à l'adresse : https://www.ofcom.org.uk/research-and-data/media-literacy-research, consulté le 28 septembre 2023.

[91] Disponible à l'adresse : https://www.unesco.org/en/media-information-literacy/about, consulté le 28 septembre 2023.

[92] Vraga, E. K., Tully, M., Maksl, A., Craft, S., & Ashley, S. (2021), "Theorizing News Literacy Behaviors", *Communication Theory*, 31(1), pp. 1-21.

[93] L'Union européenne a élaboré un ensemble de "Lignes directrices à l'intention des enseignants et des éducateurs sur la lutte contre la désinformation et la promotion de la culture numérique par l'éducation et la formation". Disponible à l'adresse suivante : https://education.ec.europa.eu/news/guidelines-for-teachers-and-educators-on-tackling-disinformation-and-promoting-digital-literacy-through-education-and-training, consulté le 28 septembre 2023.

[94] La recommandation CM/Rec(2022)13 sur l'impact des technologies numériques sur la liberté d'expression propose des suggestions sur les aspects que les programmes d'éducation numérique devraient inclure. La stratégie d'éducation aux médias en ligne du Ministère britannique du numérique, de la culture, des médias et du sport énumère cinq principes clés qui soutiennent de solides capacités d'éducation aux médias.

Il est important de souligner que toute initiative visant à renforcer l'autonomie des utilisateurs ne doit pas être considérée isolément, mais dans le cadre d'une stratégie plus large et à plusieurs volets visant à améliorer la santé des écosystèmes d'information et la résistance à la désinformation dans les sociétés démocratiques. Non seulement parce que, comme nous l'avons vu plus haut, il y a des limites à ce qu'une initiative peut accomplir à court ou moyen terme, mais aussi parce que mettre l'accent uniquement ou principalement sur les compétences des citoyens peut involontairement évoluer vers des récits qui "blâment la victime", c'est-à-dire les cibles de la désinformation, pour leur incapacité à se défendre contre les dommages causés par sa diffusion, et risque de négliger l'importance de promouvoir des conditions structurelles qui rendent un contenu de haute qualité facilement accessible à tous les membres de la société.

Il existe de nombreux exemples à travers le monde d'outils numériques et d'initiatives visant à promouvoir l'autonomisation des utilisateurs dans différents secteurs de la population. L'Alliance pour l'éducation aux médias et à l'information de l'UNESCO[95] promeut la coopération internationale indispensable sur ces questions et fournit de l'expertise et des ressources. La Digital Future Society a résumé les réussites internationales en la matière.[96] L'Observatoire européen de l'audiovisuel a publié un rapport complet qui met en lumière les principales politiques et initiatives aux niveaux national et européen.[97] La Plate-forme européenne des instances de régulation (EPRA) héberge des documents et des présentations issus d'un événement sur l'éducation de l'information organisé en 2022 et réunissant de nombreuses parties prenantes.[98]

Comme il s'agit d'un domaine qui évolue rapidement et qui est très diversifié, il serait impossible de résumer de manière exhaustive et précise les principales initiatives dans ce domaine et leurs résultats, mais la mise en commun des informations et des ressources - en particulier les évaluations des effets des différentes actions - serait très bénéfique pour toutes les parties prenantes impliquées. C'est pourquoi, au paragraphe 34 de la Note d'orientation, nous recommandons que les principaux acteurs impliqués dans la promotion d'un écosystème informationnel sain collaborent à la création d'une ressource informationnelle publique qui rassemble des données et des informations étendues, standardisées et actualisées sur toutes les initiatives visant à promouvoir l'autonomisation des utilisateurs dans les États membres.

Enfin, il est nécessaire de poursuivre les recherches sur ce qui fonctionne et ce qui ne fonctionne pas dans ce domaine. Comme indiqué précédemment, nos connaissances en la matière sont limitées par le manque d'évaluations solides et systématiques des interventions existantes, mais aussi par le fait que la recherche scientifique sur ces sujets en est encore à ses débuts et se limite à un sous-ensemble restreint de démocraties libérales riches. Il est particulièrement urgent de mener des recherches sur les facteurs qui permettent de prédire les résultats positifs au sein de différents groupes sociaux, notamment les plus vulnérables de la population (par exemple les enfants, les personnes âgées,[99] , les personnes vivant dans des écosystèmes d'information médiocres et les non-utilisateurs des technologies numériques), et dans des contextes où les institutions et les normes démocratiques sont historiquement plus faibles ou sont mises à rude épreuve par des acteurs nationaux et étrangers qui cherchent à tirer profit des perturbations démocratiques.

---

[95] Disponible à l'adresse : https://www.unesco.org/en/media-information-literacy/alliance, consulté le 28 septembre 2023.

[96] Société numérique du futur (2020), Faire face à la désinformation : Stratégies d'autonomisation des citoyens numériques.

[97] Cabrera Blázquez, F., Cappello M., Talavera Milla J., Valais S. (2022), *User empowerment against disinformation online*, Observatoire européen de l'audiovisuel.

[98] Les événements de l'EPRA sont disponibles à l'adresse suivante : https://www.epra.org/attachments?category=mil-taskforce&page=1, consulté le 28 septembre 2023.

[99] Voir Hermans, A. (2022), *L'ère numérique ? L'éducation aux médias et à l'information : une clé pour garantir les droits des personnes âgées à participer à l'ère numérique*, Département de la société de l'information, DGI (2022)03, Conseil de l'Europe.

5. Conclusion

Cet exposé des motifs a pour but d'expliquer le contexte, le raisonnement et les principes qui sous-tendent les recommandations contenues dans la note d'orientation. En résumé, en tant que pratique et profession, la vérification des faits doit s'adapter et innover pour lutter contre la prolifération de la mésinformation et de la désinformation. Des études indiquent que les normes internationales en matière de vérification des faits sont encore en cours d'élaboration et que l'identité professionnelle des vérificateurs de faits évolue également. Plus important encore, le paysage de la vérification des faits doit tenir compte des preuves empiriques concernant l'efficacité et les effets de leur travail. De plus en plus d'ouvrages indiquent que les citoyens n'acceptent pas toujours les informations corrigées, mais que leur réaction à la vérification des faits dépend fortement de l'idéologie politique et de la manière dont les informations vérifiées sont présentées. En outre, en ce qui concerne les solutions de conception des plateformes, il y a des défis à relever pour lutter contre la désinformation, en raison du manque de transparence et de responsabilité des systèmes algorithmiques des plateformes. Heureusement, il existe d'importantes normes européennes et internationales sur lesquelles la Note d'orientation s'appuie et qui peuvent contribuer à garantir un environnement en ligne favorable à un débat public pluraliste, qui peut servir de contrepoids crucial à la désinformation et à la mésinformation. Enfin, l'autonomisation des utilisateurs peut faciliter un cercle vertueux dans lequel les utilisateurs non seulement évitent de tomber dans le piège des contenus de mauvaise qualité tels que la désinformation, mais aussi protègent et alertent les autres utilisateurs qui pourraient le faire, tout en améliorant la responsabilité et la capacité des plateformes à lutter contre la propagation de la désinformation.