

# Les Lumières à l'ère numérique



RAPPORT DE LA COMMISSION  
JANVIER 2022



# Sommaire

<b>Lettre de mission du Président de la République</b> .....	4
<b>Présentation de la commission d'experts</b> .....	6
<b>Synthèse</b> .....	12
<b>Glossaire</b> .....	16
<b>Introduction</b> .....	17
CHAPITRE 1	
<b>Les mécanismes psychosociaux de la désinformation</b> .....	29
CHAPITRE 2	
<b>La logique algorithmique</b> .....	39
CHAPITRE 3	
<b>L'économie des infox</b> .....	55
CHAPITRE 4	
<b>Les ingérences étrangères</b> .....	65
CHAPITRE 5	
<b>Droit et numérique</b> .....	76
CHAPITRE 6	
<b>Esprit critique et Éducation aux médias</b> .....	88
<b>Conclusion</b> .....	102
<b>Recommandations</b> .....	106
<b>Annexes</b> .....	116

Paris, le 29 septembre 2021

Cher Gérald Bronner,

En quelques années, la façon dont nous nous informons s'est radicalement métamorphosée. Là où l'information était depuis des décennies structurée par quelques grands rendez-vous collectifs très encadrés – les journaux télévisés, la lecture de la presse écrite, elle est aujourd'hui éclatée avec, d'un côté, des médias traditionnels en déclin et, de l'autre, la multiplication de contenus autoproduits rencontrant leur public au gré des affinités idéologiques ou des hasards algorithmiques.

Cette révolution souvent virtuelle a des impacts bien réels : positifs, quand ils permettent à des communautés de passionnés de se retrouver pour échanger sur des sujets que le marché n'investissait pas ; négatifs lorsque l'éclatement des sources d'information aboutit à la division de la société en groupes construits autour de postulats déconnectés de tout fondement rationnel. Avec de premières conséquences concrètes : aux États-Unis, l'attaque en janvier 2021 du Capitole par des émeutiers convaincus que l'élection présidentielle avait été truquée, qui a provoqué des morts et des blessés ; en France, la montée des mouvements antivaccins qui, s'ils avaient prospéré, auraient entraîné une aggravation de la situation sanitaire.

Ma conviction est que ces événements ne sont que les premiers d'une longue série. Que, si nous ne faisons rien, ce phénomène de déstructuration du paysage de l'information, que l'accès aux réseaux sociaux et plateformes de partage de vidéos dès le plus jeune âge ne cesse d'amplifier, peut emporter ce que nous avons de plus précieux : notre cohésion nationale, notre système démocratique hérité des Lumières. Garant, comme Président de la République, de l'unité de la Nation et de la pérennité de nos institutions, je souhaite donc qu'un large débat s'engage sur le sujet, de nature à provoquer un sursaut collectif.

Par vos travaux, vous avez été parmi les pionniers à alerter sur les dangers liés aux bouleversements de l'ère Internet. Au moment où la plupart des observateurs mettaient exclusivement en avant les progrès liés à cette technologie – diffusion du savoir et de la culture, multiplication des liens sociaux, développement de l'innovation, vous vous attachiez, sans jamais renier ces progrès, à donner à voir leur part sombre, à souligner les risques qu'ils entraînaient, avec lucidité. Nous y sommes.

Pour faire face aux dangers que l'ère numérique fait peser sur notre démocratie et pour permettre à chacun de se libérer des bulles algorithmiques, je vous demande donc de présider une commission de haut niveau qui aura les objectifs suivants :

- Établir un état de l'art qui puisse éclairer grand public, médias, acteurs de la société civile sur l'impact d'Internet dans nos vies de citoyens : notre information, notre rapport à l'autre, notre représentation du monde et de nous-mêmes, notre exposition à des biais cognitifs qui peuvent enfermer.

– Formuler des propositions dans les champs de l'éducation, de la prévention, de la régulation et de la judiciarisation des diffuseurs de haine, afin de libérer la société des bulles de filtre, qui enferment une partie de nos concitoyens et nourrissent les extrémismes, la discorde, la violence, les dérives sectaires et les obscurantismes.

– Proposer de nouveaux espaces communs de la démocratie, de la citoyenneté et du collectif, qui puissent trouver leur place dans le monde numérique, servir de repaire comme de repère à des citoyens isolés.

– Développer une analyse historique et géopolitique de l'exposition de la France aux menaces qui pèsent sur notre démocratie et notre société au travers d'Internet et des recommandations sur les mesures à prendre dans les cadres français, européen et international.

Votre commission pourra s'appuyer sur les services du Premier ministre concernés par le sujet, notamment le secrétariat général de la défense et de la sécurité nationale et le service d'information du Gouvernement, ainsi que sur les services des autres ministères compétents, notamment au sein du ministère de l'éducation nationale, du ministère de l'enseignement supérieur, de la recherche et de l'innovation, du ministère de l'économie, des finances et de la relance, le ministère de l'intérieur et le ministère de la culture.

Vous pourrez également mobiliser les autorités et les organismes dont les décisions et les avis dans ce domaine les qualifient tout particulièrement, notamment le Conseil supérieur de l'audiovisuel, la Commission nationale de l'informatique et des libertés, la Commission nationale consultative des droits de l'homme et le Conseil national du numérique.

Je souhaite que vos travaux puissent être éclairés par les réflexions d'experts nationaux et internationaux de ces questions, notamment les équipes de recherche et les associations qui s'y consacrent. Votre commission pourra utilement interroger les acteurs du secteur des médias, notamment le service public de l'audiovisuel.

La remise complète de vos travaux est attendue pour janvier 2022.

Je sais pouvoir compter sur vous pour faire vivre un débat qui, loin des discours maniérés, saura s'adresser aux Français de toutes générations et de tous milieux, entraîner tous les engagés pour protéger le droit à une information libre et de qualité.

L'ambition est grande : faire en sorte que la société française sorte grandie de ces 25 ans de vie à l'ère numérique. Faire qu'Internet continue à être une chance pour notre Nation et notre démocratie. Une chance pour tous.

*avec ma confiance,*  
*Macron*

Emmanuel MACRON



## Présentation de la commission d'experts

Lancée par le Président de la République Emmanuel Macron le 29 septembre 2021, la commission « Les Lumières à l'ère Numérique » a été présidée par le sociologue Gérard Bronner. Cette commission a regroupé 14 experts de spécialités différentes, historiens, politologues, juristes, journalistes, enseignants, sociologues, acteurs du monde universitaire ou de la société civile, réunis pour mesurer et comprendre les dangers que le numérique fait peser sur la cohésion nationale et notre démocratie afin de mieux y faire face.



### Gérald Bronner, président de la commission

Spécialiste de sociologie cognitive, il est professeur de sociologie à l'Université de Paris, membre de l'Académie nationale de médecine, de l'Académie des technologies, de l'Institut universitaire de France, ainsi que du comité de rédaction de L'Année Sociologique. Ses ouvrages, *L'empire des croyances* (2003), *La Démocratie des crédules* (2013) ou encore *Apocalypse cognitive* (2021) portant sur la formation et la disparition des croyances collectives, rumeur, idéologie, religion ou magie, et sur la cognition humaine. Ils ont joué un rôle important dans la prise de conscience sociologique des dangers qui guettent la démocratie à l'heure où Internet fait le lit du relativisme.



### Roland Cayrol

Politologue dont les travaux portent sur les médias et leur influence politique, les structures et l'évolution de l'opinion publique ainsi que sur les comportements politiques et électoraux comparés, en France et en Europe. Fondateur-directeur de l'institut Harris France (1977-1986), il a contribué à fonder le CESA, qu'il a dirigé de 1986 à 2008. Il collabore régulièrement avec France 5, RTL, RTBF ou France 24h pour commenter l'actualité médiatique. Directeur de recherche associé à la Fondation Nationale des Sciences Politiques, directeur et conseiller de la revue Régions Magazine, Roland Cayrol est également directeur de sa société de conseil aux entreprises, le Centre d'Études et d'Analyse.



### **Laurent Cordonier**

Directeur de la recherche à la Fondation Descartes à Paris, il y conduit des études sur les questions relatives à l'information, à la désinformation et au débat public à l'heure d'Internet et des réseaux sociaux. Titulaire depuis 2016 d'un doctorat en sciences sociales de l'Université de Lausanne, il y demeure rattaché en tant que collaborateur scientifique externe. Ses travaux portent sur les théories du complot, les déterminants de la confiance et les mécanismes sociocognitifs d'affiliation sociale, notamment *La nature du social – L'apport ignoré des sciences cognitives* (2018).



### **Frédérick Douzet**

Spécialiste des enjeux géopolitiques du cyberspace et professeure à l'Université Paris 8. Directrice de son laboratoire de recherche (IFG Lab) et du projet GEODE - Géopolitique de la Datasphère (geode.science), elle est membre du comité d'éthique de la Défense depuis janvier 2020 et a participé au Comité de rédaction de la Revue stratégique de défense et de sécurité nationale en 2017. Elle a fait partie de la Global Commission on the Stability of Cyberspace (2017-2020) et dirigé la Chaire Castex de cyberstratégie de l'Institut des Hautes Etudes de Défense Nationale (IHEDN) de 2013 à 2018. Elle a reçu plusieurs prix scientifiques nationaux et internationaux pour ses recherches.



### **Rose-Marie Farinella**

Ex-journaliste de la presse écrite et enseignante, elle anime des ateliers d'éducation aux médias et à l'information. Elle a créé un scénario pédagogique « Info ou intox, comment faire la différence sur internet dès l'école primaire » qu'elle met en pratique depuis 2014 auprès d'élèves de CM2. Son travail a été primé cinq fois, notamment à l'échelon international par l'Unesco et la Commission Européenne. Elle a participé à la rédaction d'un ouvrage collectif *Des têtes bien faites* chez Puf et a co-écrit avec une journaliste d'Okapi *Stop à la manipulation*, qui est paru en octobre.



### Aude Favre

Journaliste web spécialiste des fake news, elle tient depuis 2017 une chaîne Youtube, WTFake, spécialisée dans le débusquage de fake news, afin de lutter contre la désinformation et d'ouvrir le monde du journalisme au grand public. Elle y interpelle les plus grands désinformateurs du web, ce qui a permis d'effacer nombre de contenus complotistes. Auteure de documentaires et reporter depuis 10 ans, elle travaille au sein de la société Zebra Production, et fonde l'association FAKE OFF qui vise à lutter contre les fake news en formant les plus jeunes au recul critique vis-à-vis des médias



### Jean Garrigues

Historien spécialiste d'histoire politique de la France contemporaine. Professeur émérite à l'Université d'Orléans, président du Comité d'histoire parlementaire et politique, il a publié une trentaine d'ouvrages consacrés notamment aux institutions, aux acteurs, aux valeurs, aux rites et aux mythologies républicaines. Derniers ouvrages publiés : *Les Scandales de la République. De Panama à Benalla, 2019*; *La République incarnée. De Gambetta à Macron, 2019*; *Les Perdants magnifiques. De 1958 à nos jours, 2020*; *Charles de Gaulle, l'homme providentiel, 2020*.



### Rahaf Harfoush

Anthropologue canadienne de l'innovation et de la culture numérique qui s'intéresse en particulier à l'utilisation harmonieuse des technologies émergentes en entreprise, à l'éthique de l'intelligence artificielle, à la revitalisation numérique de nos zones rurales et au renforcement de la cybersécurité en France. Membre du Conseil National du Numérique, elle a fondé une société de conseil numérique nommée Red Thread Institute of Digital Culture, et enseigne à Sciences Po Paris. Auparavant, elle était directrice associée du programme des technology pioneers au Forum économique mondial.





### Rachel Khan

Juriste, actrice, écrivaine, elle a d'abord été athlète de haut niveau durant sa jeunesse, avant d'étudier le droit public et le droit international des droits fondamentaux. Elle a été conseillère à la culture de Jean-Paul Huchon en Île-de-France de 2009 à 2015, directrice de l'association 1000 visages qui agit en faveur de l'accès des jeunes aux métiers du cinéma, et est actuellement co-directrice de La Place, centre culturel Hip Hop de la ville de Paris. À partir de 2013, elle commence une carrière d'actrice et publie plusieurs livres, notamment un roman d'inspiration autobiographique paru en 2016, et un essai de 2021 intitulé *Racée*, qui prend ses distances avec la pensée décoloniale.



### Anne Muxel

Sociologue et politologue spécialisée dans la compréhension des formes du lien des individus à la politique et plus largement du système démocratique, au travers de l'analyse de leurs attitudes et de leurs comportements (nouvelles formes d'expression politique, rapport au vote, modes de socialisation et de construction de l'identité politique). Elle a mené de nombreux travaux sur la transmission des valeurs dans la dynamique générationnelle et est une spécialiste reconnue des études sur la jeunesse. Directrice de recherches en sociologie et en science politique au CNRS (CEVIPOF/Sciences Po), elle a aussi la responsabilité du domaine Défense et société de l'Institut Stratégique de l'École Militaire (IRSEM) au ministère de la Défense.



### Rudy Reichstadt

Fondateur et directeur de Conspiracy Watch, service de presse en ligne consacré à l'approche critique des théories du complot, et expert associé à la Fondation Jean-Jaurès où il a coordonné plusieurs enquêtes d'opinion sur le conspirationnisme dans la société française. Co-auteur du documentaire « Complotisme : les alibis de la terreur » et auteur d'un essai sur la question complotiste publié chez Grasset, il co-anime le podcast « Complorama » sur France Info. Il est également membre de l'Observatoire de la haine en ligne du CSA.



### Iannis Roder

Historien spécialiste de la Shoah et enseignant dans un collège à Saint-Denis. Il est aussi responsable des formations au mémorial de la Shoah, directeur de l'Observatoire de l'éducation de la Fondation Jean-Jaurès, et membre du conseil des sages de la laïcité. Il collabore régulièrement avec le supplément Éducation du Monde, et a écrit plusieurs ouvrages sur l'enseignement dans les espaces de relégation sociale et sur l'enseignement de l'histoire de la Shoah, notamment *Allons z'enfants... la République vous appelle* en 2018 et *Sortir de l'ère victimaire, Pour une nouvelle approche de la Shoah et des crimes de masse* en 2020.



### Bertrand Warusfel

Professeur à l'Université Paris 8, avocat au barreau de Paris et vice-président de l'Association française de droit de la sécurité et de la défense (AFDSD). Croisant la recherche universitaire avec une expérience de praticien du droit, ses travaux se situent principalement à la charnière entre le droit public et le droit privé et autour des questions de droit de l'information et de l'immatériel. Ayant développé une expertise particulière en droit public de la défense et de la sécurité, il intervient également dans les domaines de la propriété industrielle et du droit des nouvelles technologies et du numérique.



### Annette Wieviorka

Historienne spécialiste de la Shoah et de l'histoire des Juifs au XX<sup>e</sup> siècle depuis la publication en 1992 de sa thèse, *Déportation et génocide : entre la mémoire et l'oubli*. Agrégée d'histoire, longtemps enseignante en Chine, directrice de recherche au CNRS, elle a été membre de la mission d'étude sur la spoliation des Juifs de France, dite mission Mattéoli. Son essai *L'heure d'exactitude : Histoire, mémoire, témoignage*, paru en 2011, est une synthèse de la mémoire de la Shoah et ses évolutions marquantes, qui montre combien « l'ère du témoin » constitue un tournant mémoriel, mais aussi historiographique.



## Synthèse

La révolution numérique bouleverse nos modes de vie, nos économies et nos pratiques sociales. Elle transforme aussi en profondeur notre rapport à l'information. En effet, nous sommes aujourd'hui confrontés à une masse inédite d'informations disponibles et à une concurrence généralisée des points de vue, qui s'expriment sans filtre et selon une logique peu intelligible pour les utilisateurs du web et des réseaux sociaux. Cette saturation et cette dérégulation du marché de l'information en ligne mettent à rude épreuve nos capacités de vigilance épistémique, ce qui nous rend davantage perméables aux fausses informations.

Désinformation, mésinformation, infox, *fake news*, théories du complot... Les vocables se multiplient pour désigner ces fausses nouvelles qui circulent en ligne et sont susceptibles d'influencer nos attitudes, nos comportements, mais aussi notre représentation du monde environnant, au risque de faire émerger des réalités parallèles incommensurables et disparaître l'espace épistémique commun nécessaire à la confrontation des opinions, des idées et des valeurs, autrement dit, à la vie démocratique. Certaines de ces désinformations, nous le verrons, relèvent d'ailleurs d'authentiques ingérences numériques étrangères, émanent d'acteurs qui cherchent à manipuler nos opinions, encourager la violence et la haine ou déstabiliser notre société à des fins stratégiques.

Notre commission avait pour mandat, premièrement, d'établir de manière synthétique l'état des connaissances sur les désordres informationnels à l'ère numérique et sur les perturbations de la vie démocratique qu'ils engendrent et, deuxièmement, de proposer des recommandations pour y faire face. Chercher à agir contre la désinformation comporte le risque de porter atteinte à des valeurs essentielles de notre démocratie, telles que les libertés d'expression, d'opinion ou d'information. C'est avec le souci de la préservation de ces libertés que notre commission a travaillé. Nos recommandations ne visent dès lors pas à éradiquer les désordres informationnels – ce qui ne serait bien entendu ni possible, ni souhaitable –, mais à limiter la propagation des contenus qui nuisent à la vie démocratique, dissuader les comportements malveillants, sanctionner les pratiques illicites, améliorer la prévention des risques et renforcer la vigilance des utilisateurs.

La compréhension des mécanismes psychosociaux (chapitre I) qui nous rendent perméables aux fausses informations nous éclaire sur les leviers permettant d'en limiter les effets. Les fausses informations sont minoritaires parmi les contenus informationnels en circulation sur Internet et les réseaux sociaux et nous sommes généralement capables de les distinguer des informations fiables. Pourtant, certaines d'entre elles parviennent à se frayer un chemin dans les esprits et sont alors susceptibles d'entraîner des conséquences dommageables tant pour les individus concernés que pour la société. La configuration des réseaux sociaux, où les informations sont noyées dans une masse de contenus de divertissement, ne nous encourage nullement à la vigilance cognitive, rempart pourtant essentiel à la crédulité. D'où notre recommandation de développer la formation à l'esprit critique (R27 et R29). En effet, la recherche

académique montre que faire preuve d'un esprit analytique capable de résister à certaines de nos intuitions immédiates est une compétence centrale pour distinguer le vrai du faux, en particulier sur Internet et les réseaux sociaux. Nous recommandons par ailleurs d'investir dans la recherche scientifique (R1) et de pousser les plateformes numériques à ouvrir leurs données aux chercheurs (R20), car les connaissances sur la prévalence de la désinformation en ligne (particulièrement en France), sur ses effets ainsi que sur les mécanismes par lesquels elle affecte les individus demeurent lacunaires. Finalement, nous attirons l'attention sur le fait que la lutte contre la désinformation dans notre pays ne pourra porter ses fruits que si médias et institutions, en tant qu'autorités épistémiques, travaillent à retisser un lien de confiance avec l'ensemble des citoyens.

Certaines logiques algorithmiques (chapitre II), sans être responsables de nos croyances ou nos comportements, contribuent toutefois à les façonner. Nous nous sommes penchés sur trois de ces logiques en particulier : l'éditorialisation algorithmique, qui désigne la manière dont les algorithmes organisent à la fois l'ordre et la fréquence d'apparition des informations selon leur capacité à capter l'attention ; le calibrage social, ou la façon dont les réseaux sociaux altèrent la perception de la représentativité et de la popularité de certains points de vue ; enfin, l'influence asymétrique rendant possible la prévalence de certains discours extrêmes et minoritaires. Nous proposons donc une série de mesures visant à améliorer le design des interfaces utilisateurs (R2) et lutter contre le biais de popularité (R3) pour sortir d'une logique algorithmique fondée sur un modèle strictement commercial ; responsabiliser les influenceurs (R4) dotés d'une grande visibilité numérique ; mettre en avant les compétences (R5) et encourager le dialogue entre plateformes et scientifiques (R6) pour mieux refléter l'état des connaissances ; enfin, se prémunir contre le risque de sur-modération (R7) en analysant plus finement les signalements d'utilisateurs.

L'un des moteurs majeurs de la désinformation est le profit. L'étude de l'économie des infox (chapitre III) démontre que la publicité programmatique constitue une source de revenus substantielle pour les artisans de la désinformation. Et ce, bien souvent à l'insu des entreprises ayant recours à des agences pour diffuser leurs campagnes et dont les publicités se retrouvent sur des sites propageant des contenus haineux, conspirationnistes ou susceptibles de troubler la paix publique. C'est pourquoi nous proposons de responsabiliser les acteurs de la publicité programmatique (R8). Les plateformes participatives ou les chaînes YouTube monétisées permettent également de récolter des financements, d'où la proposition d'encourager les bonnes pratiques mises en œuvre par les plateformes pour éviter de participer indirectement au financement de projets se compromettant avec l'incitation à la haine ou la propagation de la désinformation (R9). Enfin, les sites de presse généraliste ont fréquemment recours à des liens sponsorisés renvoyant vers des sites « pièges à clics », bien souvent pourvoyeurs de fausses informations, notamment en matière de santé.

L'autre moteur majeur de la désinformation est la compétition stratégique. Le durcissement du contexte géopolitique mondial entraîne une logique d'affrontement permanent qui caractérise la conflictualité à l'ère numérique. Cette logique se traduit par des opérations d'ingérences numériques étrangères (chapitre IV). Elle conduit à l'émergence de menaces de plus en plus

hybrides qui ont perturbé les campagnes présidentielles aux États-Unis depuis 2016 et ont aussi touché la France. D'où l'importance d'analyser les campagnes de désinformation passées, afin de protéger l'intégrité des prochains processus électoraux (R10). Ces manœuvres informationnelles se sont internationalisées au cours des deux dernières années avec la crise sanitaire, appelant à la création d'un mécanisme de gestion de crise à l'échelle européenne (R14). Ces menaces impliquent une grande diversité d'acteurs et de modes opératoires, ce qui en complique la compréhension, la détection et la prévention. Leur analyse exige un accès aux données des plateformes par les chercheurs (R20) et un partage de données structurées entre les acteurs qui étudient ces phénomènes (R11). Le droit international est un levier peu efficace pour y répondre. C'est pourquoi nous recommandons un régime de coopération exigeante avec les plateformes (R15) et la création d'un groupe de travail à l'OCDE dans un esprit de co-régulation. Enfin, la militarisation de l'espace numérique se traduit par une prolifération des opérations informationnelles. Or l'espace numérique est ultra-dynamique et partagé entre tous les acteurs. Il en résulte d'importantes interactions entre les mondes civil, économique et militaire qui brouillent les notions de théâtre intérieur/extérieur et produisent des effets qui, à leur tour, alimentent la menace. Pour ces raisons, la commission recommande de solliciter l'avis du Comité d'éthique de la défense sur la doctrine française de lutte informatique d'influence (R13) et la création d'un mécanisme de gouvernance numérique interministérielle qui tiennent compte des interactions multiples propres à ce domaine partagé (R12).

En matière de droit et numérique (chapitre V), l'étude des dispositions juridiques pouvant potentiellement être utiles pour prévenir ou sanctionner les différentes formes de désinformation (au sens de la diffusion de mauvaise foi de fausses nouvelles) incitent à ne pas modifier ou remplacer l'actuel article 27 de la loi de 1881 sur la liberté de la presse (R16 et R17). En revanche, la sanction pénale pourrait être complétée par un mécanisme de mise en cause de la responsabilité civile des diffuseurs de mauvaise foi de fausses nouvelles pouvant porter préjudice à autrui, responsabilité qui pourrait notamment être proportionnée au niveau de viralité de la diffusion et de la popularité numérique de son auteur (R18). Les délais de procédure judiciaire, en particulier pour obtenir une décision définitive au fond, demeurent largement inadaptés à la réaction rapide qu'exige la diffusion virale de certaines fausses nouvelles. Le Conseil supérieur de l'audiovisuel, qui va devenir l'Autorité de régulation de la communication audiovisuelle et numérique au 1<sup>er</sup> janvier 2022, sera chargé de veiller au respect par les plateformes de leurs obligations de retrait rapide de certains contenus illicites graves et dispose d'ores et déjà d'une compétence plus générale de lutte contre la diffusion de fausses nouvelles. On peut estimer qu'il manque au moins une procédure formalisée de signalement auprès de la future ARCOM qui soit ouverte à tout citoyen (R19) afin de faire connaître a posteriori à l'ARCOM les difficultés rencontrées dans la prise en compte de la réclamation par la plateforme ou, au contraire, les cas de retrait unilatéral d'un contenu qui ne justifiait pas une mesure aussi radicale, afin que la plateforme lui apporte une réponse appropriée. Enfin, dans le cadre de la loi européenne sur les services numériques, pour responsabiliser les plateformes, la commission propose d'introduire explicitement une disposition qui reconnaît que les fausses nouvelles susceptibles de troubler l'ordre public constituent des contenus répréhensibles (R21), de mettre en place un organe d'expertise extérieur pour coopérer avec les plateformes (R22) et de créer un régime de co-régulation entre plateformes, régulateurs et société civile (R23).

Enfin, face aux perturbations informationnelles si complexes à juguler, la meilleure réponse est sans doute la modération individuelle, puisque tout un chacun est devenu un opérateur sur le marché en ligne de l'information. Les connaissances en matière d'éducation aux médias et à l'information (EMI) et de pédagogie de l'esprit critique (chapitre VI) ouvrent des pistes pour nous aider à mieux évaluer cette cacophonie d'information, avec une indépendance de jugement retrouvée. L'Éducation nationale, en ce sens, a un rôle essentiel à jouer. Les initiatives en la matière sont cependant disparates, d'où la nécessité de créer une cellule interministérielle dédiée au développement de l'esprit critique et d'une EMI tout public (R24). Une meilleure compréhension des difficultés cognitives expérimentées par les élèves permettrait également de mieux adapter les contenus pédagogiques (R25). L'éveil à ces enjeux pourrait passer par la création d'une Grande Cause nationale pour le développement de l'esprit critique et l'EMI (R26), la systématisation de la formation en milieu scolaire (R27) et des actions de sensibilisation menées auprès des autorités scolaires au sein des établissements et des rectorats, ainsi qu'auprès des élus locaux, des collectivités locales et des responsables de bibliothèques (R28). Enfin, il est important de créer un continuum entre le temps scolaire, l'université, le monde culturel, le monde du travail et la société civile (R29). *In fine*, la formation à la vigilance intellectuelle doit être un objectif partagé pour toute société attachée à faire vivre l'héritage du siècle des Lumières et les espoirs qu'il portait.

En conclusion, une réflexion prospective nous fait entrevoir de nouvelles questions qui surgiront demain. Le concept de métavers, notamment, esquisse un univers où nous serons immergés dans une confusion croissante entre les mondes réels et virtuels, et nécessite une réflexion éthique (R30).

Notre rapport avait pour seule ambition de penser, dans l'urgence, des solutions pour juguler un problème amplifié, voire transformé par le numérique. Ce travail ne nous exonère en rien de la réflexion collective que nous devons mener en parallèle pour penser quelle société et quelle démocratie nous souhaitons construire dans ce monde numérique en devenir.

## Glossaire

**Fausse information (ou mésinformation) :** Contenu d'information faux ou inexact, ayant ou non été délibérément créé et diffusé pour induire les gens en erreur.

Dans le présent rapport, le terme de fausse information est aussi utilisé comme terme générique pour désigner ensemble les mésinformations, les désinformations, les infox, les informations hyperpartisanes, les théories du complot et les informations pièges à clics.

**Désinformation :** Contenu d'information ou ensemble de contenus d'information faux ou inexact(s), créé(s) avec l'intention délibérée d'induire les gens en erreur.

**Infox (fake news) :** Contenu d'information fabriqué de toute pièce ou extrêmement inexact publié sur Internet et mise en forme de manière à ressembler à un contenu d'information grand public légitime.

**Information hyperpartisane :** Contenu d'information couvrant des événements s'étant réellement produits, mais avec un très fort parti pris partisan le rendant potentiellement trompeur.

**Information piège à clics :** Contenu d'information sensationnaliste, souvent faux, inexact ou trompeur, créé dans le seul but d'attirer l'attention des internautes en vue de générer du trafic sur la page l'hébergeant.

**Théorie du complot :** Récit tendant à attribuer abusivement, et au détriment d'une explication plus plausible, l'origine d'un événement ou d'un phénomène à l'action occulte d'un groupe généralement restreint d'individus poursuivant un but légalement ou moralement répréhensible. Outre qu'elle témoigne d'une préférence pour des explications de type intentionnaliste, une théorie du complot a le plus souvent pour caractéristiques de contester sans véritables preuves l'explication communément admise d'une circonstance donnée et d'accuser ceux qui y auraient réellement ou supposément intérêt.

**Ingérence étrangère :** Intervention d'un État ou d'acteurs agissant pour le compte d'un État dans la politique d'un autre État par des moyens numériques.  
Cette définition varie selon les plateformes et les institutions. Celle retenue par Viginum est : « *Opérations structurées et coordonnées par des acteurs étrangers, destinées à propager via les plateformes numériques des contenus manifestement trompeurs et hostiles, cherchant à porter atteinte aux intérêts fondamentaux de la Nation.* »

**Influence numérique étrangère :** Opération informationnelle menée dans l'espace numérique (web, réseaux sociaux) par un acteur ou un groupe d'acteurs étrangers à des fins d'influence.



# ➔ Introduction

Dans son opuscule *Qu'est-ce que les Lumières?* (1793), le philosophe Emmanuel Kant interpelait ses contemporains par une formule célèbre : «*Ose savoir! Aie le courage de te servir de ton propre entendement! Voilà donc la doctrine des Lumières.*» Cette doctrine portait l'espoir d'un siècle : l'avènement prochain, grâce aux progrès de l'éducation et de la disponibilité de l'information, d'une société éclairée, fondée sur la raison et la connaissance.

Les débuts du XXI<sup>e</sup> siècle ne paraissent pas avoir réalisé tout à fait cet espoir, et cette «doctrine des Lumières» mérite d'être réexaminée à l'ère de la révolution numérique. La rupture majeure qu'elle apporte bouleverse nos modes de vie, nos économies, et nos pratiques sociales. Elle interroge aussi profondément les notions de pouvoir et de démocratie. Son avènement s'inscrit dans un contexte de montée des populismes, d'exacerbation des conflits religieux et des tensions géopolitiques entre grandes puissances, de défiance

populaire envers les élites et les institutions et de formidables défis pour l'avenir de l'humanité, tels que le dérèglement climatique et les pandémies. La révolution numérique offre une opportunité inédite de repenser les cadres de la démocratie représentative en tirant le meilleur parti des systèmes dynamiques complexes qui peuvent permettre, entre autres, une diffusion massive de la connaissance, un niveau d'interactions sociales inédit, une participation citoyenne accrue<sup>1</sup>. Elle offre aussi de nouveaux modes de gouvernance et d'intelligence collective, qu'il reste néanmoins, pour la plupart, à inventer.

Nous sommes encore à l'aube de cette révolution dont nous commençons tout juste à prendre la mesure. Elle nous appelle à définir nos ambitions pour un monde en devenir, dans lequel nous avons encore du mal à nous projeter collectivement. Et pourtant, nous devons d'ores et déjà répondre aux nombreux défis que cette révolution nous pose.

---

<sup>1</sup> Par exemple, l'utilisation des médias sociaux dans la campagne américaine de 2008 illustre la manière dont les plateformes technologiques peuvent permettre aux citoyens de se mobiliser autour des processus démocratiques – voir Harfoush, R. (2009). *Yes We Did: An inside Look at How Social Media Built the Obama Brand, New Riders.*

## Le chaos informationnel contemporain

L'un des faits contemporains les plus marquants est la dérégulation massive du marché de l'information, accélérée par le développement d'Internet. Celle-ci peut être décrite par deux faits importants au moins. D'une part, par la masse extraordinaire – et inédite dans l'histoire de l'humanité – des informations disponibles et, d'autre part, par le fait que chacun peut verser sa propre représentation du monde sur ce marché devenu proliférant.

Ceci a toutes sortes de conséquences, mais la plus évidente est l'éclosion d'une concurrence généralisée de tous les modèles intellectuels qui prétendent décrire le monde, des plus frustes aux plus sophistiqués. Aujourd'hui, quiconque dispose d'un compte sur un réseau social peut directement apporter une contradiction, sur la question des vaccins par exemple, à un professeur de l'Académie nationale de médecine. Le premier peut même avoir plus d'audience que le second. Cette concurrence généralisée des points de vue, non hiérarchisés selon les compétences et le savoir de ceux qui les font valoir, peut-elle faire advenir ce monde de connaissance auquel aspiraient nos ancêtres du siècle des Lumières? Peut-on espérer que les énoncés les mieux argumentés et les plus solidement démontrés s'imposent grâce à cette libre concurrence face aux productions de la crédulité que sont les superstitions, les légendes urbaines et autres théories complotistes?

Un regard même superficiel sur la situation actuelle permet d'en douter. En effet, si Internet et les réseaux sociaux autorisent l'accès à un volume jamais atteint de connaissances et d'information fiables, ils ont également ouvert la voie au partage d'une grande quantité de fausses informations, dont les conséquences restent rarement confinées aux réseaux sociaux. L'invasion du Capitole en 2021 aux États-Unis illustre de façon exemplaire combien les théories du complot, telles que celles qui ont abondamment circulé sur les réseaux sociaux des supporters de Donald Trump, peuvent catalyser la violence politique. La désinformation en ligne durant la période pandémique a exacerbé les peurs relatives aux vaccins conduisant parfois, en France, jusqu'au saccage de centres de vaccination. Toujours en France, un certain nombre d'actions criminelles ont été fomentées, et parfois même menées à bien, au nom de théories du complot diffusées sur Internet – on peut par exemple penser à l'enlèvement de la jeune Mia par des individus s'inspirant des thèses conspirationnistes de Rémy Daillet ou aux actions violentes, déjouées par la Direction générale de la sécurité intérieure (DGSI), qu'avaient planifiées les membres d'une mouvance complotiste d'extrême droite à l'encontre du ministre de la santé, d'une loge maçonnique et de centres de vaccination.

Il serait bien entendu naïf de penser que de tels événements ne sont que le produit du fonctionnement d'Internet et des réseaux sociaux. D'une part, la manipulation des faits et des informations n'a pas attendu Internet pour exister. D'autre part, la désinformation en ligne n'est pas le cœur du problème, mais le révélateur et le catalyseur des maux de nos sociétés. Elle contribue

toutefois bien souvent à les exacerber. À ce titre, les théories du complot sont emblématiques de ces narrations fantasmées qui ont toujours accompagné l'histoire des sociétés humaines et font leur miel de la défiance à l'égard des autorités, des institutions et des médias ou encore du sentiment d'anomie<sup>2</sup>. En France comme ailleurs, et bien avant l'apparition d'Internet, les récits conspirationnistes ont embrasé les esprits tout au long de notre histoire contemporaine<sup>3</sup>. Les théories sur le complot des juifs, des jésuites ou encore des francs-maçons ont pollué le débat public au XIX<sup>e</sup> siècle et durant une bonne partie du XX<sup>e</sup>. Elles avaient en commun de proposer une lecture simpliste de la société de l'époque, prétendument menacée par une organisation occulte et puissante aspirant à dominer le monde. Cet «*âge d'or de la Conjuración*», selon Raul Girardet<sup>4</sup>, était l'expression d'un profond malaise social, d'une sourde angoisse collective face à un monde en pleine transformation vers la démocratie, la révolution industrielle et le capitalisme. Cette analyse, replacée dans le contexte actuel, trouverait à se renouveler dans la séduction qu'exercent ces récits conspirationnistes simplistes et vengeurs à l'heure de la mondialisation et du sentiment de dépossession qu'elle implique, du sentiment d'éloignement de la décision politique ou encore celui de la perte de maîtrise de son environnement.

Il est donc important de le rappeler : les théories du complot prospèrent aussi sous l'effet de conditions sociales (dé)favorables. En effet, les études témoignent d'un niveau de complotisme en moyenne plus élevé dans les pays au sein desquels les gens se sentent socialement menacés (taux de chômage élevé, par exemple) et où les institutions et les autorités sont perçues comme indignes de confiance<sup>5</sup>. Si l'on ajoute à cela que certains gouvernements ne sont pas toujours innocents de tentatives de manipulations de l'opinion publique par la diffusion de fausses informations, on comprend que de nombreux facteurs sont réunis pour assurer un certain succès aux théories conspirationnistes.

Ces récits fantasmés offrent une intelligibilité politique du monde. Pour cette raison, ils peuvent être paradoxalement socialisateurs et mobilisateurs<sup>6</sup>, permettant à certains de trouver de nouvelles coalitions, de nouvelles intégrations sociales et même une nouvelle façon de faire de la politique. Ces nouveaux cadres de socialisation influencent les attitudes et les comportements en terme de vie personnelle et sociale, mais aussi de représentation du monde environnant. Ainsi, il a été montré que l'exposition à des thèses conspirationnistes décourage la participation à la vie démocratique par le vote, alimente les préjugés, voire la violence envers certaines catégories de la population et peut conduire au rejet du consensus scientifique sur diverses questions, telles que le dérèglement climatique ou l'efficacité des vaccins<sup>7</sup>.

2 Voir p. ex. Wagner-Egger, P. (2021) *Le bruit de la conspiration : Psychologie des croyances aux théories du complot*, PUG

3 Girardet, R. (1986) *Mythes et mythologies politiques*, Le Seuil.

4 Ibid.

5 Cordonier, L., Caferio, F., et Bronner, G. (2021). «Why are conspiracy theories more successful in some countries than in others? An exploratory study on Internet users from 22 Western and non-Western countries.» *Information sur les Sciences Sociales / Social Science Information*, 60(3), 436-456.

6 Cordonier, L., et Dieguez, S. (2021). «Le complotisme, un outil de mobilisation dangereux.» *TANGRAM* (Commission fédérale contre le racisme), 45, 55-56

7 Jolley, D., Mari, S., et Douglas, K. M. (2020). Consequences of Conspiracy Theories. In: *Routledge Handbook of Conspiracy Theories*. Routledge, London, 231-241.

Le succès de ces récits est donc profondément enchâssé dans certaines réalités sociales largement indépendantes du monde numérique. Cependant, et au-delà des seules croyances complotistes, certaines propriétés d'Internet accroissent le potentiel de nuisance des fausses informations. L'instantanéité et l'ubiquité des réseaux sociaux, notamment, permettent à des contenus nuisibles d'apparaître et de se diffuser simultanément aux événements sur lesquels elles portent. Ainsi, toutes sortes de thèses complotistes sur l'incendie de Notre-Dame de Paris se sont mises à fleurir sur les réseaux sociaux tandis que le feu faisait encore rage. Certaines de ces théories, très partagées et commentées, ont rapidement acquis une visibilité telle qu'elles ont dû faire l'objet de démentis dans les médias, condamnés à pourchasser la désinformation.

Enfin, les outils numériques décuplent les forces d'acteurs, notamment étatiques, qui cherchent à s'ingérer dans un processus électoral, manipuler l'opinion publique, tromper l'adversaire, discréditer les dissidents politiques, escroquer des victimes ou harceler des personnes vulnérables. Des services étatiques, des criminels ou de simples individus peuvent à moindre coût organiser une viralité artificielle de contenus, masquer leurs traces et leur identité, fabriquer de fausses images ou de fausses vidéos quasiment impossibles à distinguer des vraies dans le but de nuire, faire du profit, avancer leurs intérêts ou encore déstabiliser des sociétés démocratiques.

# Endiguer la propagation de la désinformation

Au vu des nuisances potentielles de la désinformation<sup>8</sup>, il semble pertinent de chercher à freiner sa propagation sur Internet. Cependant, tout projet d'intervention volontaire sur ce marché de l'information, plus encore s'il est d'origine politique, pose la question de la préservation des libertés, notamment celle d'opinion, qui est au cœur de la Déclaration des droits de l'homme et du citoyen. Pour autant, la situation actuelle de cacophonie informationnelle ne garantit nullement la pleine expression de cette liberté. En effet, l'information sur Internet est en réalité pré-éditorialisée selon des logiques algorithmiques qui paraissent parfois échapper à leurs créateurs mêmes<sup>9</sup> et nous asservissent alors qu'elles devaient nous servir. Ainsi, sur YouTube par exemple, 120 000 ans de temps de vidéos sont visionnés chaque jour. Parmi elles, 70% sont regardées en raison de la recommandation de l'intelligence artificielle de la plateforme<sup>10</sup>. Il s'agit là d'une illustration parmi beaucoup d'autres de la puissance de prescription éditoriale des grands opérateurs du web. L'information est ainsi organisée dans un monde numérique dé-régulé : elle est régie par des logiques algorithmiques qui échappent à notre regard et peuvent dès lors contribuer à orienter la formation de nos opinions à notre insu.

Qui plus est, ce type de prescription n'est pas toujours favorable à l'information la plus sincère et la mieux argumentée. Par exemple, une étude de 2019 a montré que les recherches du terme « climat » sur YouTube ont majoritairement (54%) orienté les internautes vers des vidéos climatosceptiques<sup>11</sup>. Si les réseaux sociaux deviennent un mode de plus en plus important d'information sur l'actualité, notamment pour les jeunes générations<sup>12</sup>, toutes les enquêtes montrent qu'ils sont perçus en même temps comme le moyen le moins fiable de consulter l'actualité<sup>13</sup>. Ce paradoxe n'est pas sans rappeler Ovide : « *Video meliora proboque, deteriora sequor* » (« Je vois le meilleur et je l'approuve, je fais le pire. »)

Les réseaux sociaux ne favorisent pas non plus un débat démocratique serein. Une analyse portant sur Twitter montre par exemple qu'un message a 17% de chances de plus d'être ré-partagé pour chaque mot d'indignation qu'il contient<sup>14</sup>. Le constat pour Facebook n'est pas

8 Voir p. ex., Jolley, D., Mari, S., et Douglas, K. M. (2020). Consequences of Conspiracy Theories. In: *Routledge Handbook of Conspiracy Theories*. Routledge, London, 231-241 ; Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., et Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature human behaviour*, 5(3), 337-348.

9 [https://www.lemonde.fr/pixels/article/2021/10/26/comment-l-algorithme-de-facebook-echappe-au-controle-de-ses-createurs\\_6099888\\_4408996.html](https://www.lemonde.fr/pixels/article/2021/10/26/comment-l-algorithme-de-facebook-echappe-au-controle-de-ses-createurs_6099888_4408996.html)

10 Solsman, J. E. (10/01/2018) Ever Get Caught in an unexpected hour long YouTube Binge? Thank YouTube AI for That. CNET.

11 Allgaier, J. (2019). Science and environmental communication on YouTube: Strategically distorted communications in online videos on climate change and climate engineering. *Frontiers in Communication*, 4:36. doi: 10.3389/fcomm.2019.00036.

12 Watson, A. (2021) Social media as a news source worldwide 2021. Statista; Watson, A. (2021) Frequency of using selected news sources among Millennials in the United States as of May 2021. Statista; Watson, A. (2021) Media used for national political news in EU countries 2019, by age.

13 Watson, A. (2021) Share of adults who trust selected news sources worldwide in 2018, by region. Statista.

14 Brady et al. (2017) Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*.

plus reluisant, puisque le célèbre réseau social a favorisé algorithmiquement les messages incitant à la colère plutôt que ceux exprimant la tempérance et l'approbation<sup>15</sup>. Ceci contribue à faire des plateformes sociales des lieux d'expression conflictuels, plutôt que des espaces de partage et de discussion raisonnée des points de vue. On sait d'ailleurs que les prescriptions algorithmiques des réseaux sociaux peuvent participer à la radicalisation des esprits. Selon un rapport interne de Facebook, par exemple, les individus ayant intégré un groupe extrémiste sur le réseau social l'ont fait dans deux tiers des cas suite à une recommandation de l'algorithme<sup>16</sup>.

Les algorithmes façonnent ainsi notre rapport à l'information d'une manière qui reste encore souvent trop opaque pour l'utilisateur comme pour le législateur. Pourtant, la compréhension de la manière dont l'information se fabrique et se diffuse, mais aussi de la façon dont les utilisateurs se l'approprient et la partagent est l'un des premiers piliers de la résilience de nos sociétés.

---

15 <https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/>

16 Horwitz, J., Seetharaman, D. (26 mai 2020) «Facebook Executives Shut Down Efforts to Make the Site Less Divisive.» *Wall Street Journal*.

## Renforcer la résilience de la société

On sait qu'en matière de désinformation et de théories du complot, la prévention est plus efficace que la correction<sup>17</sup>. Une étude a d'ailleurs montré que la première impression induite par une fausse information perdure souvent, même lorsque l'individu qui y a été confronté apprend qu'elle est bel et bien fausse<sup>18</sup>. Le démenti ne suffit donc pas toujours à faire disparaître cette empreinte qui a marqué son esprit et le conduit à avoir ensuite une interprétation erronée de toute nouvelle donnée sur le même sujet. On le comprend, l'instantanéité des réseaux sociaux confère dès lors un certain avantage concurrentiel aux fausses informations, rapides à générer et à diffuser, face aux informations fiables qui nécessitent, elles, du temps pour être vérifiées et recoupées.

Un autre aspect du fonctionnement d'Internet peut jouer en faveur de notre crédulité. Les psychologues ont montré de longue date que, dans de nombreuses situations, nous avons tendance à favoriser les informations nouvelles qui vont dans le sens de nos croyances établies par rapport à celles qui pourraient les contredire (surtout lorsque les croyances en question sont en lien avec nos valeurs). Il s'agit là du célèbre « biais de confirmation », aussi appelé « *congeniality bias* » par les chercheurs<sup>19</sup>. Ce biais de confirmation nous incline ainsi à rechercher avant tout des informations qui renforceront nos points de vue. Internet facilite l'expression de ce biais, dans la mesure où la quantité d'informations disponibles y est telle que chacun pourra y trouver en quelques clics celles qui le satisferont, indépendamment de leur adéquation au réel.

Cela signifie non pas que nous soyons moins exposés sur Internet à des points de vue divergents que dans la vie hors ligne mais que nous pouvons aisément y trouver quantité d'éléments venant soutenir nos croyances, y compris quand ces dernières sont contraires à l'état de la connaissance sur un sujet donné. Des recherches ont montré qu'un tel mécanisme de renforcement de nos croyances sur Internet opère bel et bien en ce qui concerne les théories du complot, et qu'il peut même inciter certains individus concernés à s'entourer sur les réseaux sociaux de personnes qui partagent leurs croyances complotistes, formant ainsi des « chambres d'écho » au sein desquelles les positions se radicalisent progressivement<sup>20</sup>.

L'état de saturation du marché de l'information en ligne met à rude épreuve notre capacité de vigilance épistémique. En effet, le nombre de contenus auxquels nous sommes confrontés est

17 Voir, par exemple, Jolley, D., et Douglas, K. M. (2017). « Prevention Is Better Than Cure: Addressing Anti-Vaccine Conspiracy Theories. » *Journal of Applied Social Psychology*, 47(8), 459-469 ; Bonetto, E., Troian, J., Varet, F., Lo Monaco, G., et Girandola, F. (2018). « Priming Resistance to Persuasion Decreases Adherence to Conspiracy Theories. » *Social Influence*, 13(3), 125-136.

18 De Keersmaecker J. et Roets A. (2017). « Fake news: Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions », *Intelligence*, 65, pp.107-110.

19 Pour une méta-analyse sur le « *congeniality bias* », voir Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., et Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin*, 135(4), 555.

20 Voir, par exemple, Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., et Quattrociocchi, W. (2016). The Spreading of Misinformation Online. *Proceedings of the National Academy of Sciences*, 113(3), 554-559 ; Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., et Quattrociocchi, W. (2016). « Echo Chambers: Emotional Contagion and Group Polarization on Facebook. » *Scientific Reports*, 6, 37825.

tel que nous ne pouvons accorder que peu de temps à l'évaluation de la crédibilité de chacun d'entre eux, ce qui nous rend davantage perméables aux fausses informations<sup>21</sup>. La répétition en ligne d'une information erronée peut de plus renforcer son pouvoir de persuasion, puisque plus nous rencontrons le même argument, le même post ou le même tweet, plus nous avons l'impression qu'il est vrai<sup>22</sup>.

Par conséquent, il existe un risque que les individus évoluent dans des réalités parallèles où le consensus sur des faits documentés de manière empirique par des professionnels de l'information ou des théories étayées par les expériences et la littérature scientifique ne soit plus possible.

---

21 Bago, B., Rand, D. G. et Pennycook, G. (2020), « Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines », *Journal of Experimental Psychology: General*, 149(8), 1608.

22 Pennycook, G., Cannon, T. D. et Rand, D. G. (2018), « Prior exposure increases perceived accuracy of fake news », *Journal of Experimental Psychology: General*, 147-12, pp. 1865-1880.



## La nécessité d'un espace épistémique commun

Entre la disponibilité des fausses informations sur Internet et la polarisation des réseaux sociaux, c'est la possibilité même d'un espace épistémique et de débat commun qui est menacée, c'est-à-dire d'un monde où il est possible d'échanger, de se contredire, de réformer son jugement, un monde où les points de vue peuvent diverger mais demeurent toujours commensurables. Donald Trump, avec ses 89 millions de followers sur Twitter avant son éviction de ce réseau social, est devenu une figure emblématique de cette menace. Aux États-Unis, certains citoyens vivent dans la même société mais pas forcément dans le même monde. C'est ainsi que l'on peut interpréter les déclarations de l'ancienne directrice de campagne de Trump, Kellyanne Conway. Celle-ci avait soutenu l'idée que davantage de personnes avaient assisté à l'investiture de Trump qu'à celle de n'importe quel président auparavant, alors même que les faits lui donnaient manifestement tort. Elle aurait pu admettre qu'elle s'était fourvoyée, mais elle avait préféré parler de « fait alternatif », comme s'il était possible de donner de la même réalité deux interprétations contradictoires et pourtant de même valeur.

Cette déclaration officialisait la rupture d'un espace de débat commun aux États-Unis. Le désaccord est normal en démocratie, mais la dispute présuppose que les arguments échangés soient commensurables, et c'est ce principe fondamental qui est menacé aujourd'hui. Si la France n'est certes pas les États-Unis, une récente étude de l'université de Stanford<sup>23</sup> fait néanmoins apparaître qu'au sein de notre pays le niveau de « polarisation affective » – à savoir, le niveau des sentiments hostiles des citoyens à l'égard d'autres partis que le leur – augmente régulièrement depuis quarante ans, atteignant aujourd'hui l'une des valeurs les plus élevées parmi les 12 pays de l'OCDE étudiés.

L'existence d'un espace épistémique commun est au fondement de la vie sociale, en particulier de la vie démocratique. Sans un tel espace, aucun problème collectif ne peut trouver de solutions acceptables malgré les divergences d'opinion. Or, les problèmes auxquels nous sommes confrontés sont considérables – et l'on pense par exemple au dérèglement climatique – mais pour les résoudre ils demandent un préalable : pouvoir tirer le meilleur de l'intelligence collective.

23 <https://www.brown.edu/Research/Shapiro/pdfs/cross-polar.pdf>

On le voit, Internet est un formidable progrès qui permet la circulation de l'information et de la connaissance à une rapidité et dans des proportions inédites, tout comme il rend possible l'échange public entre citoyens en transcendant les distances géographiques. Revers de la médaille, cette technologie facilite également la diffusion d'informations fausses ou trompeuses, aux conséquences parfois bien réelles, et pourrait favoriser la polarisation de notre société plutôt que l'échange argumenté des points de vue. C'est sur la question de telles perturbations de la vie démocratique par l'univers numérique que notre commission a été invitée à réfléchir par le Président de la République.

## Objectifs et méthodes de travail de la commission

Cette commission s'est donné pour tâche de faire un état de lieux de la recherche et des connaissances accumulées sur le sujet, tant par la consultation de la littérature scientifique et des rapports existants que par l'audition et la consultation écrite de chercheurs et d'acteurs publics et privés en lien avec le monde numérique. Elle a dû le faire en un temps très contraint (100 jours) et a, dans ces conditions, abandonné immédiatement l'objectif d'exhaustivité.

Ce n'est bien entendu pas une question que nous découvrons, puisque des institutions comme l'OMS, l'ONU, le Conseil de l'Europe et bien d'autres encore ont produit des analyses du phénomène. Des discussions sont d'ailleurs en cours au sein du Conseil de l'Union Européenne et du Parlement européen sur la nouvelle législation sur les services numériques (DSA) visant à garantir un environnement en ligne sûr et responsable.

Pour traiter le sujet des perturbations de la vie démocratique par l'univers numérique, les membres de notre commission ont considéré qu'il pouvait être analytiquement découpé en sept sous-thèmes, qui organisent le plan de ce rapport.

Le premier sous-thème porte sur les mécanismes psychosociaux susceptibles de nous rendre perméables aux fausses informations et de diminuer notre capacité à les identifier comme erronées ou trompeuses. Que dit la science sur les variables impliquées dans ces phénomènes ?

Le deuxième interroge les possibilités d'altération des logiques algorithmiques du marché en ligne de l'information. Est-il possible de modifier certaines règles de visibilité et de viralité qui président à ce marché pour en atténuer les effets négatifs ?

Le troisième explore les logiques économiques de la diffusion des fausses informations et de la haine en ligne. L'écosystème de l'information sur Internet est assujéti à une économie de l'attention, dont le fonctionnement dépend des grandes entreprises du numérique (réseaux sociaux, moteurs de recherche, plateforme de vidéos en ligne...). Celles-ci ne sont pas toujours opposées aux efforts nécessaires pour réguler les externalités négatives de ce marché, mais une partie de leurs intérêts économiques (qui tiennent notamment à l'engagement de leurs utilisateurs) ne va pas nécessairement de pair avec un souci de la qualité de l'information qui se diffuse dans le monde numérique.

L'autre grande menace qui pèse sur l'équilibre de la vie démocratique vient de l'existence d'ingérences numériques d'acteurs étrangers, étatiques ou privés, qui servent leurs intérêts dans le monde numérique. Ces tentatives sont documentées et feront l'objet d'un éclairage par le quatrième sous-thème du présent rapport.

Le cinquième sous-thème répondra à certaines des questions précédentes en interrogeant la question de la régulation de ce marché par le droit. Cette question est à la fois sensible et essentielle. Elle est d'ailleurs à l'agenda de toutes les réflexions portant sur les perturbations de la vie démocratique par les flux numériques – et notamment, au moment où ce rapport a été rédigé, prise en compte dans les travaux préparatoires à la législation sur les services numériques (Digital Services Act).

La meilleure réponse aux perturbations informationnelles induites par le monde numérique est sans doute la modération individuelle, puisque tout un chacun est devenu un opérateur sur le marché en ligne de l'information. Ce sera donc l'objet du sixième sous-thème que d'établir l'état de la connaissance en matière d'EMI (Éducation aux médias et à l'information) et de pédagogie de l'esprit critique. Comment évaluer l'information, suspendre son jugement, combattre les raisonnements captieux ? Les compétences nécessaires aux bonnes pratiques en ce domaine peuvent être proposées à tous les niveaux et à tous les moments de notre formation intellectuelle. À ce titre, l'Éducation nationale est une institution essentielle pour doter chacun de nos concitoyens d'outils qui lui permettront de retrouver son indépendance de jugement dans cette cacophonie d'informations. Au-delà, la formation à la vigilance intellectuelle doit être un objectif partagé pour toute société qui conserve vivant l'héritage et les espoirs portés par le siècle des Lumières. Pour aider à cet effort, les médias traditionnels (presse, radio, télévision) ont un rôle essentiel à jouer car ils demeurent la principale source de production de contenus. Cependant, ils ne sont pas épargnés par les externalités négatives de cette dérégulation du marché. La façon dont une certaine logique numérique contamine l'exercice professionnel des journalistes et restreint leur liberté éditoriale mérite d'être interrogée.

Le septième sous-thème, qui tiendra lieu de conclusion à ce rapport, posera la question d'une nouvelle forme de citoyenneté numérique. L'implication éclairée de chacun de nos concitoyens est une des pistes envisagées pour contrebalancer la prévalence sur les réseaux sociaux des propositions les plus radicales et les plus conflictuelles. Si certaines idées acquièrent dans le monde numérique une visibilité qui n'est pas proportionnée à leur représentativité, c'est parce qu'elles sont portées (c'est notamment le cas avec les anti-vaccins) par des communautés plus motivées que les autres à faire valoir leur point de vue. Cette asymétrie ne doit évidemment pas nous inciter à la censure mais à une réflexion sur l'implication de chacun dans ce nouvel espace citoyen que sont devenus les mondes numériques.

Ceux-ci offrent par ailleurs les conditions techniques idéales pour créer des espaces de débats démocratiques renouvelés. Il reste à penser les formes que ces espaces pourraient prendre pour éviter certains écueils observés et s'assurer que c'est bien la sagesse des foules qui pourra s'y exprimer plutôt que sa déraison.

## CHAPITRE 1

# → Les mécanismes psychosociaux de la désinformation

Une grande part de ce que nous savons, ou pensons savoir, ne nous vient pas de nos propres sens et expériences, mais du témoignage d'autrui. En effet, dès notre plus jeune âge, nous sommes en permanence exposés à des informations communiquées par les personnes qui nous entourent – parents, amis, enseignants, etc. –, et les médias nous rapportent des nouvelles sur l'état du monde que nous ne pourrions acquérir par nous-mêmes. L'être humain se trouve ainsi dans un état de profonde dépendance épistémique à l'égard de ses semblables.

Si cette situation nous permet d'élargir de beaucoup notre volume de connaissances par rapport à ce que nous pourrions savoir seuls, elle nous expose aussi au risque d'être involontairement induits en erreur, voire délibérément trompés par les autres. L'existence d'un tel risque ne nous empêche pas d'adopter une forme de confiance par défaut envers les informations rencontrées. En effet, la recherche montre que nous avons en moyenne tendance à accepter, plutôt qu'à rejeter les propositions nouvelles qui nous sont

soumises<sup>24</sup>. Il nous arrive même de nous fier à des informations inexactes que nous devrions pourtant reconnaître comme telles sur la base de nos connaissances préalables<sup>25</sup>.

Notre propension à prendre pour vrai ce qui nous est communiqué n'est pas en soi une attitude irrationnelle. En effet, en situation normale, la plupart des informations qui nous sont transmises par les membres de notre entourage sont vraies – il s'agit généralement d'informations banales sans grands enjeux épistémiques. D'un point de vue statistique, il est donc plutôt rationnel de se montrer a priori confiant à l'égard de ce qui nous est rapporté, et de ne rejeter que ce qui est très invraisemblable ou manifestement faux (ce que nous faisons effectivement la plupart du temps)<sup>26</sup>. Cependant, dans un monde où quantité d'informations nous proviennent désormais d'Internet et des réseaux sociaux, une telle confiance minimale par défaut demeure-t-elle raisonnable? Là encore, tout dépend de la part relative que vrai et faux occupent en ligne.

24 Pour une revue de la littérature, voir Brashier, N. M., et Marsh, E. J. (2020). «Judging truth.» *Annual review of psychology*, 71, 499-515.

25 P. ex. Fazio, L. K., Barber, S. J., Rajaram, S., Ornstein, P. A., et Marsh, E. J. (2013). «Creating illusions of knowledge: Learning errors that contradict prior knowledge» *Journal of Experimental Psychology: General*, 142(1), 1-5.

26 Sur ces différents points, voir Brashier, N. M., et Marsh, E. J. (2020), *op. cit.*

# La représentation des fausses informations sur Internet

La recherche académique est aujourd'hui incapable de fournir une estimation précise de la part de la désinformation sur les réseaux sociaux et, plus largement, sur Internet<sup>27</sup>. En réalité, il s'agit là d'une estimation extrêmement difficile à produire, dont le résultat serait de toute façon largement fluctuant dans le temps et selon les régions linguistiques ou pays considérés. On sait par exemple que les périodes électorales dans les pays démocratiques sont particulièrement propices à la diffusion en ligne de fausses informations.

Une étude<sup>28</sup> portant sur l'élection présidentielle américaine de 2016 permet de l'illustrer. Ses auteurs ont recherché les principales infox en circulation avant le scrutin. Ils en ont identifié 115 en faveur de Donald Trump (ou anti Hillary Clinton) et 41 en faveur d'Hillary Clinton (ou anti Donald Trump). Les chercheurs ont ensuite mesuré leur diffusion sur Facebook au cours des 3 mois précédant l'élection. Il en ressort que les infox pro-Trump ont été partagées sur ce seul réseau social 30,3 millions de fois durant cette période, et celles en faveur de Clinton l'ont été 7,6 millions de fois.

Si ces chiffres sont impressionnants, les infox sont pourtant minoritaires parmi l'ensemble des contenus d'actualité auxquels s'exposent les internautes américains, y compris en période électorale. C'est ce que montrent les études qui se sont penchées sur les sources d'information consultées par les Américains : les sites Internet connus pour publier de fausses informations ne composent qu'une faible part de leur régime informationnel en ligne<sup>29</sup>. Pour ce qui est de la France, les données sur la consommation médiatique effective des internautes sont rares. Une récente étude<sup>30</sup> de la Fondation Descartes montre cependant que, dans l'ensemble, les Français s'informent eux aussi majoritairement sur des sources web fiables.

Les auteurs de cette étude ont suivi durant 30 jours consécutifs le parcours informationnel sur Internet de 2 372 personnes majeures résidant en France, sélectionnées de manière à composer un panel représentatif de la population. Il en ressort que si 39% de ces personnes ont consulté au moins une fois une source d'information non fiable au cours de cette période, elles n'y ont passé en moyenne que 11% de leur temps quotidien d'information en ligne (ce qui représente 0,4% de leur temps total de connexion à Internet). Cette valeur moyenne fluctue

27 Voir p. ex. Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F.,... et Zittrain, J. L. (2018). «The science of fake news.» *Science*, 359(6380), 1094-1096 ; Pennycook, G., et Rand, D. G. (2021). The psychology of fake news. *Trends in cognitive sciences*, 25(5), 388-402.

28 Allcott, H., et Gentzkow, M. (2017). «Social media and fake news in the 2016 election» *Journal of economic perspectives*, 31(2), 211-36.

29 P. ex., Guess, A. M., Nyhan, B., et Reifler, J. (2020). «Exposure to untrustworthy websites in the 2016 US election.» *Nature human behaviour*, 4(5), 472-480.

30 Cordonier, L., et Brest, A. (2021). *Comment les Français s'informent-ils sur Internet ? Analyse des comportements d'information et de désinformation en ligne*. Étude de la Fondation Descartes, [www.fondationdescartes.org/wp-content/uploads/2021/03/Etude\\_Information\\_Internet\\_FondationDescartes\\_2021.pdf](http://www.fondationdescartes.org/wp-content/uploads/2021/03/Etude_Information_Internet_FondationDescartes_2021.pdf)

bien entendu entre les individus, certains d’entre eux ayant consulté ces sources non fiables plus régulièrement et plus longtemps que les autres.

Il faut préciser que, de l’aveu même de ses auteurs, cette étude principalement axée sur la fréquentation de sites Internet d’information et de désinformation sous-estime l’exposition des individus aux fausses informations en circulation sur les réseaux sociaux. Il en va de même pour les études conduites aux États-Unis qui mobilisent une méthodologie approchante<sup>31</sup>. Dès lors, si l’on peut affirmer que les sites web publiant des infox sont dans l’ensemble beaucoup moins consultés par les Français que les sites des médias traditionnels, les données manquent pour estimer le niveau d’exposition moyen de nos concitoyens aux fausses informations sur les réseaux sociaux.

On sait toutefois qu’en France, des infox bénéficient régulièrement d’une certaine viralité sur les réseaux sociaux<sup>32</sup> et que leurs usagers sont plus susceptibles que les autres de se rendre sur des sites d’information non fiables<sup>33</sup>. Ce fait, qui s’observe également aux États-Unis<sup>34</sup>, permet de conclure que les réseaux sociaux constituent une porte d’entrée importante vers la désinformation<sup>35</sup>, quand bien même elle est probablement minoritaire dans l’ensemble des contenus d’actualités qui y circulent<sup>36</sup>.

31 Voir Pennycook, G., et Rand, D. G. (2021), *op. cit.*

32 Voir p. ex. CSA (2020). *La propagation des fausses informations sur les réseaux sociaux : étude de la plateforme Twitter*. <https://www.csa.fr/Informer/Collections-du-CSA/Focus-Toutes-les-etudes-et-les-comptes-rendus-synthetiques-proposant-un-zoom-sur-un-sujet-d-actualite/La-propagation-des-fausses-informations-sur-les-reseaux-sociaux-etude-de-la-plateforme-Twitter>

33 Cordonier, L., et Brest, A. (2021), *op. cit.*

34 P. ex., Fournay, A., Racz, M. Z., Ranade, G., Mobius, M., et Horvitz, E. (2017). «Geographic and Temporal Trends in Fake News Consumption During the 2016 US Presidential Election.» *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, 2071–2074* ; Guess, A. M., Nyhan, B., et Reifler, J. (2020). *Op. cit.*

35 Lazer, D. M., et al. (2018), *op. cit.*

36 P. ex. Guess, A., Nagler, J., et Tucker, J. (2019). «Less than you think: Prevalence and predictors of fake news dissemination on Facebook.» *Science advances*, 5(1), eaau4586.

## Effets de la désinformation

Une désinformation massive n'est pas nécessaire pour influencer négativement les personnes qui y sont exposées : un petit nombre de fausses informations peut déjà avoir des effets mesurables sur les croyances ou les attitudes des individus. C'est ce que permet d'illustrer une étude conduite au Royaume-Uni et aux États-Unis visant à quantifier l'impact de la désinformation sur l'intention de se faire vacciner contre la covid-19.

Début septembre 2020, les auteurs de cette étude<sup>37</sup> ont exposé 3 000 Britanniques et autant d'Américains à 5 fausses informations sur le vaccin – il s'agissait de messages trompeurs circulant beaucoup à cette époque sur les réseaux sociaux. Parallèlement, dans chacun des deux pays, 1 000 participants ont quant à eux été exposés à 5 informations correctes sur le vaccin. Les chercheurs ont mesuré l'intention des participants de se faire vacciner avant et après les avoir exposés soit aux 5 fausses informations (groupes tests de 3 000 personnes dans chaque pays), soit aux 5 informations correctes (groupes de contrôle de 1 000 personnes dans chaque pays).

Lors de la mesure initiale, 54,1% des participants britanniques et 42,5% des participants américains affirmaient avoir « définitivement » l'intention de se faire vacciner contre la covid-19. Dans les groupes tests, après exposition aux 5 fausses informations sur le vaccin ces proportions tombaient respectivement à 48,6% et 39,8%, soit environ 6 points de moins que dans les groupes de contrôle après exposition aux 5 informations correctes. Ces résultats montrent clairement que l'exposition à un petit nombre de posts de réseaux sociaux trompeurs est suffisante pour produire un effet négatif (au moins à court terme) sur la disposition des individus à l'égard de la vaccination.

Comme tout un chacun peut le constater avec l'infodémie<sup>38</sup> qui accompagne la crise de la covid-19 depuis ses débuts, et qui semble être particulièrement virulente en France<sup>39</sup>, la désinformation en ligne peut prendre diverses formes, dont celle de théories du complot plus ou moins élaborées et détaillées<sup>40</sup>. Mais les chercheurs n'ont pas attendu cette infodémie pour étudier les effets négatifs des théories du complot sur les croyances et les attitudes des individus<sup>41</sup> : des études antérieures avaient ainsi déjà permis d'établir que l'exposition à des théories du complot sur les vaccins diminue l'intention de se faire vacciner ou de faire vacciner ses enfants<sup>42</sup>.

37 Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., et Larson, H. J. (2021). «Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA.» *Nature human behaviour*, 5(3), 337-348.

38 <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%99infodemic%E2%80%99-misinformation-and-cybercrime-covid-19>

39 Islam, M. S., Sarkar, T., Khan, S. H., Kamal, A. H. M., Hasan, S. M., Kabir, A., ... et Seale, H. (2020). «COVID-19–related infodemic and its impact on public health: A global social media analysis.» *The American journal of tropical medicine and hygiene*, 103(4), 1621.

40 *Idem*

41 Pour une revue de la littérature sur le sujet, voir Jolley, D., Mari, S., et Douglas, K. M. (2020). «Consequences of Conspiracy Theories.» In: *Routledge Handbook of Conspiracy Theories*. Routledge, London, 231-241.

42 P. ex., Jolley, D., et Douglas, K. M. (2014a). «The effects of anti-vaccine conspiracy theories on vaccination intentions.» *PLoS one*, 9(2), e89177.



Les théories du complot en circulation sur les réseaux sociaux remettent en question le consensus scientifique sur bien d'autres sujets que les seuls vaccins. Certaines d'entre elles soutiennent par exemple que, contrairement à ce que scientifiques, gouvernements et médias voudraient nous faire croire, le dérèglement climatique ne serait pas une réalité établie ou ne serait pas causé par l'activité humaine. Or, il a été montré que le fait d'être exposé à des thèses complotistes de ce type réduit l'intention d'adopter des comportements en faveur de la cause climatique<sup>43</sup>.

Plus généralement, l'exposition à des théories du complot en tout genre favorise la défiance à l'égard des autorités et des institutions<sup>44</sup>, décourage la participation à la vie démocratique par le vote<sup>45</sup> et alimente les préjugés négatifs<sup>46</sup>, voire les attitudes hostiles<sup>47</sup> à l'égard de diverses catégories de la population. Plus inquiétant encore, certaines thèses complotistes sont fortement soupçonnées de participer à la radicalisation des esprits au sein de groupes extrémistes (islamistes ou d'extrême droite, par exemple) et de faciliter ainsi le passage de ces groupes à des actions violentes ou terroristes<sup>48</sup>. Plusieurs études récentes ont par ailleurs observé l'existence d'un lien statistique significatif entre le fait d'adhérer à des théories du complot sur la covid-19 et le fait de manifester l'intention de commettre des actes violents<sup>49</sup>.

On le voit, la désinformation peut avoir toutes sortes de conséquences délétères sur les individus et la société. Que sait-on des mécanismes psychosociaux qui permettent aux fausses informations d'exercer leurs effets néfastes sur les esprits ?

43 Jolley, D., et Douglas, K. M. (2014b). «The social consequences of conspiracism : Exposure to conspiracy theories decreases intentions to engage in politics and to reduce one's carbon footprint.» *British Journal of Psychology*, 105(1), 35-56.

44 Einstein, K. L., et Glick, D. M. (2015). „Do I think BLS data are BS? The consequences of conspiracy theories.» *Political Behavior*, 37(3), 679-701.

45 Jolley, D., et Douglas, K. M. (2014b). *Op. cit.*

46 Jolley, D., Meleady, R., et Douglas, K. M. (2020). «Exposure to intergroup conspiracy theories promotes prejudice which spreads across groups.» *British Journal of Psychology*, 111(1), 17-35.

47 Bilewicz, M., Winiewski, M., Kofta, M., et Wójcik, A. (2013). «Harmful Ideas, The Structure and Consequences of Anti-Semitic Beliefs in Poland.» *Political Psychology*, 34(6), 821-839.

48 Bartlett, J., et Miller, C. (2010). *The power of unreason: Conspiracy theories, extremism and counter-terrorism*. London: Demos, <http://westernvoice.net/Power%20of%20Unreason.pdf> ; voir aussi p. ex. Amarasingam, A., et Argentino, M. A. (2020). The QAnon conspiracy theory: A security threat in the making. *CTC Sentinel*, 13(7), 37-44.

49 P. ex. Levinsson, A., Miconi, D., Li, Z., Frounfelker, R. L., et Rousseau, C. (2021). «Conspiracy theories, psychological distress, and sympathy for violent radicalization in young adults during the CoViD-19 pandemic: a cross-sectional study.» *International journal of environmental research and public health*, 18(15), 7846 ; Jolley, D., et Paterson, J. L. (2020). «Pylons ablaze: Examining the role of 5G COVID-19 conspiracy beliefs and support for violence.» *British journal of social psychology*, 59(3), 628-640.

## Distinguer le vrai du faux sur Internet

La désinformation possède souvent un caractère politique, dans le sens où elle vise à jeter le discrédit sur les membres des partis adverses ou sur leurs positions ou, inversement, à valoriser le camp duquel elle provient. Dans ce contexte, il a été avancé qu'un mécanisme motivationnel inconscient pourrait nous rendre particulièrement susceptibles de prendre pour vraies des informations politiques en réalité fausses ou hyperpartisanes : nous voudrions croire aux informations idéologiquement concordantes avec notre propre positionnement politique, quel que soit leur statut de vérité. Cette hypothèse est cependant remise en question à la lumière de données récentes<sup>50</sup>.

Si les individus ont effectivement tendance à croire davantage aux informations allant dans le sens de leur positionnement politique, des études montrent que, pour autant, « la politique ne l'emporte pas sur la vérité<sup>51</sup> ». En effet, des informations vraies mais politiquement discordantes sont en moyenne davantage crues que des infox politiquement concordantes<sup>52</sup>. À eux seuls, les biais partisans ne suffisent donc pas à donner du crédit à certaines désinformations politiques rencontrées sur Internet ou les réseaux sociaux.

La raison pour laquelle il arrive aux individus de se fier à de fausses informations relève probablement moins d'une motivation à y croire que de la simple incapacité à les identifier comme fausses. Nous évaluons généralement la véracité d'une information nouvelle à l'aune de nos connaissances préalables. Les informations compatibles avec nos connaissances, ou qui les confortent, seront facilement acceptées, tandis que nous aurons tendance à rejeter celles qui les contredisent<sup>53</sup>. Il n'est dès lors pas surprenant que nous soyons plus à risque de prendre pour vraie une infox lorsque nous manquons de connaissances ou avons des connaissances erronées sur le sujet en question<sup>54</sup>. Une étude a par exemple montré que les personnes ayant un faible niveau de connaissances scientifiques sont plus susceptibles que les autres de croire à de fausses informations sur la covid-19<sup>55</sup>.

Cependant, la connaissance n'immunise pas systématiquement contre le risque d'accorder du crédit à de fausses informations<sup>56</sup>. Ces dernières peuvent pénétrer l'esprit des individus en

50 Voir Pennycook, G., et Rand, D. G. (2021), *op. cit.*

51 «politics does not trump truth», *ibid.*, p. 390.

52 Pennycook, G., et Rand, D. G. (2019). «Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning.» *Cognition*, 188, 39-50.

53 Voir Brashier, N. M., et Marsh, E. J. (2020), *op. cit.*

54 À noter qu'en matière de rétablissement de la vérité, les connaissances erronées posent davantage problème que l'absence de connaissances : les personnes non-informées sur un sujet sont plus susceptibles de mettre à jour leurs croyances lorsqu'elles sont confrontées à des informations fiables que les personnes mésinformées. Voir Li, J., et Wagner, M. W. (2020). «The value of not knowing: Partisan cue-taking and belief updating of the uninformed, the ambiguous, and the misinformed.» *Journal of Communication*, 70(5), 646-669.

55 Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., et Rand, D. G. (2020). «Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention.» *Psychological science*, 31(7), 770-780.

56 Voir p. ex. Fazio, L. K., Barber, S. J., Rajaram, S., Ornstein, P. A., et Marsh, E. J. (2013), *op. cit.*

profitant de leur manque de vigilance, de leur distraction, voire d'une certaine forme de paresse intellectuelle. En effet, soupeser et analyser une information nouvelle avant de l'intégrer ou de la rejeter demande un effort cognitif plus important que de se fier à la première impression qu'elle nous laisse<sup>57</sup>. Or, nous nous comportons généralement en « avares cognitifs », préférant minimiser nos efforts mentaux<sup>58</sup>.

Il existe néanmoins des différences interindividuelles dans la propension à se contenter ou non de suivre sa seule intuition face à une information ou donnée nouvelle. La recherche sur la manière dont l'être humain raisonne montre que nous sommes tous équipés de deux systèmes de traitement de l'information : le premier, rapide et intuitif ; le second, plus lent et réflexif, susceptible de nous faire revenir sur une évaluation du premier<sup>59</sup>. Cependant, certaines personnes qualifiées de « réflexives » ou « d'analytiques » sont davantage enclines que d'autres, dites « intuitives », à faire appel à leur second système de traitement de l'information et, partant, à réviser si besoin une première impression erronée. Ces différences de style de pensée entre les individus peuvent être mesurées au moyen de divers tests cognitifs<sup>60</sup>.

Une série de recherches empiriques<sup>61</sup> recourant à ces tests montre que les personnes plus « réflexives » parviennent mieux à distinguer les infox des informations fiables, et qu'elles sont moins susceptibles d'y croire. Une étude expérimentale<sup>62</sup> a par ailleurs établi que si l'on entrave la vigilance des individus face à des informations nouvelles, les poussant donc à se fier à leur seule intuition, leur capacité à identifier les infox diminue. Il semblerait donc que la crédulité résulte dans bien des cas d'un défaut de vigilance cognitive. Or, les réseaux sociaux n'incitent assurément pas à une telle vigilance, dans la mesure où les contenus d'information sérieux y sont souvent noyés parmi les contenus de divertissement. De plus, de nombreuses infox sont partagées sur les réseaux sociaux sous la forme d'images sans hyperliens vers une quelconque source, ce qui rend difficile pour les utilisateurs de vérifier la solidité des faits avancés<sup>63</sup>.

La configuration des réseaux sociaux aurait également des conséquences négatives sur la propension de leurs utilisateurs à y partager eux-mêmes de fausses informations. Il arrive en effet aux individus de décider de partager sur les réseaux sociaux des informations qu'ils

57 Voir Brashier, N. M., et Marsh, E. J. (2020), *op. cit.*

58 P. ex., Fiske, S. T., et Taylor, S. E. (1991), *Social Cognition*. New York: McGraw-Hill.

59 Pour une présentation des différentes versions de ce modèle dual et une discussion des critiques qui leur sont adressées, voir p. ex. Evans, J. S. B., et Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on psychological science*, 8(3), 223-241.

60 Le plus connu de ces tests étant le *Cognitive Reflection Test*. Voir Frederick, S. (2005). «Cognitive reflection and decision making.» *Journal of Economic perspectives*, 19(4), 25-42 ; Toplak, M. E., West, R. F., et Stanovich, K. E. (2011). «The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks.» *Memory et cognition*, 39(7), 1275-1289.

61 Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., et Cannon, T. D. (2019). «Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking.» *Journal of applied research in memory and cognition*, 8(1), 108-117 ; Pehlivanoglu, D., Lin, T., Deceus, F., Heemskerk, A., Ebner, N. C., et Cahill, B. S. (2021). «The role of analytical reasoning and source credibility on the evaluation of real and fake full-length news articles.» *Cognitive research: principles and implications*, 6(1), 1-12 ; Pennycook, G., et Rand, D. G. (2019). *op. cit.* ; Ross, R. M., Rand, D. G., et Pennycook, G. (2021). «Beyond «fake news»: Analytic thinking and the detection of false and hyperpartisan news headlines.» *Judgment et Decision Making*, 16(2), 484-504.

62 Bago, B., Rand, D. G., et Pennycook, G. (2020). «Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines.» *Journal of experimental psychology: general*, 149(8), 1608.

63 Fazio, L. (2020). «Out-of-context photos are a powerful low-tech form of misinformation.» *The Conversation*, 14.02.2020, <https://theconversation.com/out-of-context-photos-are-a-powerful-low-tech-form-of-misinformation-129959>

ne considèrent pourtant pas comme vraies quand on leur demande de les évaluer<sup>64</sup>. Or, ce comportement relèverait moins d'une volonté de tromper les autres que de la distraction et de la recherche de «likes». En effet, deux études expérimentales<sup>65</sup> ont établi que si l'on attire discrètement l'attention des individus sur la question de la véracité des contenus, leur intention de partager des informations qu'ils sont capables de reconnaître comme fausses diminue significativement.

En plus des effets d'un manque de connaissances préalables et d'un défaut de vigilance, la littérature scientifique a identifié d'autres mécanismes susceptibles de brouiller la distinction entre vraies et fausses informations dans l'esprit des individus, particulièrement sur Internet et les réseaux sociaux<sup>66</sup>. L'un d'entre eux est l'effet de répétition d'un message. De nombreuses études montrent que plus une information – qu'elle soit vraie ou fausse – est répétée à un individu, plus ce dernier aura tendance à la croire vraie<sup>67</sup>. Une seule exposition préalable à un contenu peut déjà suffire à augmenter sa crédibilité lorsqu'il est vu une seconde fois.

Ce phénomène est renforcé par le fait que si l'on se souvient généralement du message en question, on tend à en oublier la source<sup>68</sup>. Ainsi, une fausse information qui nous avait initialement paru douteuse en raison de sa source peu fiable pourra nous apparaître comme vraie lorsque nous la croiserons à nouveau dans un autre contexte ; nous la trouverons d'autant plus vraie que nous l'avons déjà rencontrée auparavant. Les réseaux sociaux favorisent probablement un tel mécanisme, puisque certaines infox qui y circulent sont repartagées par de nombreux comptes et peuvent ainsi réapparaître régulièrement sur le « mur » ou le « fil » d'actualité des utilisateurs – un processus amplifié par les algorithmes d'engagement, dont le travail consiste à présenter aux utilisateurs des contenus similaires à ceux avec lesquels ils ont déjà interagi. De manière plus pernicieuse, les opérations de *fact checking*, en donnant de la visibilité aux infox auxquelles elles s'attaquent, pourraient bien elles aussi contribuer à les crédibiliser par effet de répétition<sup>69</sup>.

Enfin, la défiance à l'égard des médias, des institutions et du gouvernement est un facteur corrélé tant avec la fréquentation sur Internet de sources d'information non fiables<sup>70</sup> qu'avec l'adhésion aux théories du complot<sup>71</sup>. Cela s'explique probablement par le fait qu'une telle

64 Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., et Rand, D. G. (2021). «Shifting attention to accuracy can reduce misinformation online.» *Nature*, 592(7855), 590-595 ; Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., et Rand, D. G. (2020), *op. cit.*

65 *Idem*

66 Pour des revues de la littérature, voir Brashier, N. M., et Marsh, E. J. (2020), *op. cit.* ; Pennycook, G., et Rand, D. G. (2021), *op. cit.* ; Rapp, D. N. (2016). The consequences of reading inaccurate information. *Current Directions in Psychological Science*, 25(4), 281-285.

67 Pour une méta-analyse de ces études, voir Dechêne, A., Stahl, C., Hansen, J., et Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, 14(2), 238-257.

68 Voir p. ex. Rapp, D. N. (2016), *op. cit.* ; Brashier, N. M., et Marsh, E. J. (2020), *op. cit.*

69 Voir p. ex. Lazer et al. (2018), *op. cit.*

70 Cordonier et Brest (2021), *op. cit.*

71 Pour deux revues récentes de la littérature sur les facteurs d'adhésion aux théories du complot, voir Wagner-Egger, P. (2021). *Psychologie des croyances aux théories du complot : Le bruit de la conspiration*. PUG ; Delouée, S., et Dieguez, S. (2021). *Le complotisme : Cognition, culture, société*. Mardaga. À noter que l'on retrouve également au niveau international un lien statistique entre défiance à l'égard des autorités et complotisme : Cordonier, L., Cafiero, F., et Bronner, G. (2021). « Why are conspiracy theories more successful in some countries than in others? An exploratory study on Internet users from 22 Western and non-Western countries. » *Social Science Information*, 60(3) 436-456.

défiance conduit les personnes concernées à rechercher des informations sur des sources « alternatives » aux médias traditionnels, qu'elles considèrent comme biaisés, corrompus ou à la solde du pouvoir. Ces sources d'information font la part belle aux théories du complot. Ces dernières seront alors susceptibles de séduire les individus se défiant des médias et des autorités, précisément parce qu'elles contestent les explications d'événements historiques et d'actualité provenant d'acteurs médiatiques et institutionnels. Exposés à de telles narrations complotistes, la défiance de ces individus se verra alors encore renforcée.

On sait par ailleurs que les personnes affectées par un sentiment ou une crainte de précarisation, de stigmatisation ou de déclassement sont particulièrement à risque de céder aux théories du complot<sup>72</sup>. Si elles le sont, c'est certainement parce que les théories du complot leur offrent une grille de lecture du monde susceptible de conférer un sens à leur situation et de désigner une cause univoque aux injustices et menaces sociales dont elles se considèrent victimes<sup>73</sup>.

---

72 Voir p. ex., Wagner-Egger, P., Adam-Troian, J., Cordonier, L., Cafero, F., et Bronner, G. (in press). The Yellow Vests in France: Psychosocial determinants and consequences of the adherence to a social movement in a representative sample of the population. *International Review of Social Psychology* ; DiGrazia, J. (2017). The social determinants of conspiratorial ideation. *Socius*, 3, 1-9 ; Goertzel, T. (1994). Belief in conspiracy theories. *Political Psychology*, 15, 731-742 ; Uscinski, J. E. et Parent, J. M. (2014). *American Conspiracy Theories*. Oxford, Oxford University Press ; Mazzocchetti, J. (2012). « Sentiments d'injustice et théorie du complot. Représentations d'adolescents migrants et issus des migrations africaines (Maroc et Afrique subsaharienne) dans des quartiers précaires de Bruxelles. » *Brussels Studies* [en ligne], 63.

73 Pour une illustration de cette situation, voir Mazzocchetti, J. (2012). *op. cit.*

## Conclusion du chapitre 1

Au regard de ce que nous apprend la recherche, la vigilance cognitive et le développement de notre esprit analytique constituent probablement les meilleurs remparts individuels face à aux fausses informations. Dès lors, la piste d'action qui semble la plus prometteuse pour lutter contre les effets délétères de la désinformation est celle du renforcement de la formation à l'esprit critique et de l'éducation aux médias et à l'information (EMI) (R. 27 et 29). La formation à l'esprit critique doit impérativement se faire sur la base de contenus pédagogiques dont l'efficacité aura été scientifiquement évaluée. Il s'agit donc de mettre en place des procédures scientifiques et une structure de recherche permettant une telle évaluation (R. 24). Nous reviendrons sur ces recommandations dans le chapitre 6 du présent rapport, spécifiquement consacré à l'esprit critique et à l'EMI.

Par ailleurs, la recherche scientifique sur la prévalence de la désinformation en ligne, sur ses effets ainsi que sur les mécanismes par lesquels elle affecte les individus doit être soutenue et renforcée dans notre pays (R. 1). En effet, les données étudiées dans la littérature scientifique portent trop rarement sur la France, et les conclusions de travaux basés sur des données issues d'autres pays – États-Unis, principalement – ne sont pas nécessairement transposables chez nous.

La France, via l'Union européenne, devrait en outre exiger des plateformes numériques qu'elles ouvrent plus largement l'accès à leurs données aux chercheurs afin qu'ils puissent étudier les phénomènes de désinformation en ligne sous leurs différents aspects. Les modalités de cet accès pourraient être celles que propose la Commission européenne dans le cadre du *Digital Services Act*<sup>74</sup>, qui fait actuellement l'objet de négociations (R. 20).

Pour finir, il importe de souligner que la lutte contre la désinformation dans notre pays ne pourra passer uniquement par des mesures visant à inciter les individus à la vigilance sur Internet ou à améliorer le fonctionnement algorithmique des réseaux sociaux. C'est plus profondément le lien de confiance entre les citoyens et les médias et les institutions qu'il s'agit de retisser.

---

74 Article 31 de la proposition de règlement du Parlement européen et du Conseil relatif à un marché intérieur des services numériques (Législation sur les services numériques – Digital Services Act – DSA) et modifiant la directive 2000/31/CE, COM(2020) 825 final

## CHAPITRE 2

# → Logiques algorithmiques

L'actualité récente - et l'on songe notamment à l'affaire des Facebook files<sup>75</sup> - nous rappelle le rôle que l'environnement numérique et les logiques algorithmiques peuvent jouer dans les dynamiques de diffusion de la fausse information et de radicalisation des opinions. Cet environnement, comme nous l'avons souligné au chapitre précédent, ne saurait être considéré comme seul facteur de la perturbation de la vie démocratique mais la façon dont il modifie et modèle les opinions mérite qu'on lui consacre un chapitre de ce rapport.

Nous allons d'abord souligner que l'influence des effets algorithmiques ne doit pas être exagérée, avant d'établir que les spécificités du monde numérique font néanmoins peser de nouveaux risques sur la vie démocratique et d'expliquer pourquoi il est urgent de prendre des dispositions supplémentaires pour y faire face. En conclusion, enfin, nous verrons que, malgré les limites de leurs actions, les plateformes ne restent pas totalement passives face aux dangers qu'elles engendrent.

---

<sup>75</sup> Frances Haugen, une ancienne employée de Facebook a rendu publics de nombreux documents internes de l'entreprise qui ont révélé des scandales relatifs à la modération des échanges sur le réseau social. Ces documents ont été en partie publiés et sont appelés les *Facebook files*. Frances Haugen a par ailleurs été auditionnée par la commission.

## De la nécessité de la nuance

La connaissance scientifique sur la manière dont les algorithmes façonnent nos croyances ou nos comportements, notamment politiques, n'est pas encore stabilisée et fait parfois valoir des données et des arguments qui peuvent paraître contradictoires. Certaines recherches par exemple ont montré que les réseaux sociaux tendent à nous enfermer dans des chambres d'écho idéologiques, au sein desquelles nous rencontrons avant tout des arguments allant dans le sens de nos propres opinions.<sup>76</sup> D'autres travaux exposent pourtant que la contradiction y est habituelle sur les réseaux sociaux<sup>77</sup> et, d'une façon générale, que les interactions avec des individus aux opinions divergentes y sont plus fréquentes qu'on ne le croit souvent<sup>78</sup> : une situation qui est de nature à générer des échanges virulents entre internautes peut même aboutir à l'expression de propos haineux. De même, certaines recherches montrent que les réseaux sociaux exposent leurs usagers à des sources d'information plus diverses que celles qu'ils consultent dans la vie hors ligne<sup>79</sup>. Cependant, ce point est partiellement trompeur. En effet, l'offre des médias traditionnels sur les réseaux sociaux (sous forme de partages d'articles, par exemple) est par nature morcelée. Ceux qui consomment des informations venant de la presse par ce truchement ne lisent le plus souvent qu'un article<sup>80</sup>, et sont moins susceptibles de lire l'intégralité du journal que ceux qui y ont accès dans d'autres conditions. De ce fait, ce seront souvent les thématiques plutôt que les supports qui domineront l'éditorialisation numérique. Dans ces conditions, cette diversité peut être factice, car la lecture se fera sur des thèmes préférentiels traités par des médias transverses, plutôt sur une réelle diversité de sujets.

En ce qui concerne les informations d'actualité, s'il est bien une promesse non tenue d'Internet, c'est celle de l'élargissement qualitatif de l'offre tout autant que de la demande. En effet, quand on observe les flux des échanges d'actualités en ligne, on constate que le marché cognitif y est animé par des effets de concentration d'attention brefs, soudains et massifs<sup>81</sup>. Cette concentration temporelle de l'attention constitue ce que certains appellent le buzz. Cela prend une forme très tangible lorsqu'on observe à grande échelle la façon dont notre attention collective est orientée vers une histoire qui fera l'actualité un bref instant pour nous conduire à une autre, qui n'aura pas plus d'espérance de vie.

76 Matteo, C., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W. and Starnini, M. (2021), «The echo chamber effect on social media», *Proceedings of the National Academy of Sciences* 118-9.

77 Barnidge, M. (2017), «Exposure to political disagreement in social media versus face-to-face and anonymous online settings Political Communication», 34-2, pp.302–321. Silver, L., Huang, C., et Taylor, K. (2019), «In Emerging Economies, Smartphone and Social Media Users Have Broader Social Networks», Pew Research Center report.

78 Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A. et Bonneau, R. (2015). «Tweeting from left to right: Is online political communication more than an echo chamber?» *Psychological Science*, 26(10), 1531–1542.

79 Fletcher, R. et Nielsen, R. K. (2018), «Are people incidentally exposed to news on social media? A comparative analysis», *New Media and Society*, 20-7, pp.2450–2468.

80 Cordonier, L., et Brest, A. (2021). *Comment les Français s'informent-ils sur Internet ? Analyse des comportements d'information et de désinformation en ligne*. Étude de la Fondation Descartes, <https://hal.archives-ouvertes.fr/hal-03167730/document>

81 Beauvisage, T., Beuscart, J.S., Couronne, T. et Mellet, K. (2013), «Le succès sur Internet repose-t-il sur la contagion? Une analyse des recherches sur la viralité», *Tracés. Revue de Sciences humaines* : <http://journals.openedition.org/traces/5194> ; DOI : <https://doi.org/10.4000/traces.5194>



Trois informaticiens<sup>82</sup> ont analysé durant trois mois 90 millions d'articles publiés tant sur les sites web de médias grand public que sur des blogs. Leur analyse des cycles de vie de l'actualité montre combien la compétition fait rage pour attirer l'attention et combien fugace – quelques jours tout au plus – est l'attrait que nous pouvons avoir collectivement pour un sujet. Leur modèle confirme aussi bien la dispersion gigantesque des sources (un million six cent mille, en l'occurrence) que la convergence des thèmes. Autrement dit, l'augmentation importante du nombre des sources et du volume des flux d'information engendrée par le développement d'Internet n'a pas inversé la tendance à l'homogénéisation des thématiques d'actualité captant massivement l'attention du public.

On entend souvent qu'Internet et les réseaux sociaux sont envahis par les fausses informations, idée qu'il faut fortement nuancer, comme nous l'avons vu dans le chapitre précédent : plusieurs études, conduites tant aux États-Unis qu'en France, soulignent que la désinformation constitue probablement une part minoritaire du volume globale des actualités consultées sur les réseaux sociaux, et plus largement sur Internet<sup>83</sup>. Pour autant, il faut se garder d'en conclure que la désinformation en ligne ne constituerait dès lors pas un problème. En effet, la question du seuil à partir duquel des effets tangibles de la désinformation peuvent être observés est laissée en suspens par les études qui se focalisent sur la part de la population exposée à ces informations<sup>84</sup>. Par ailleurs, si les sites de fausses informations n'ont pas toujours l'influence directe qu'on leur prête, une étude<sup>85</sup> montre que les médias traditionnels ont tendance à reprendre certaines informations de ces sources douteuses lorsqu'elles vont dans le sens de leur inclination partisane, participant ainsi à leur mise en avant.

De même que la désinformation en ligne ne doit pas être surestimée, gardons-nous d'exagérer son influence sur les grands événements sociaux<sup>86</sup>. La polarisation politique par exemple ne peut que partiellement s'expliquer par le contexte numérique, et on ne trouve pas dans la littérature scientifique de réponse ferme à la question du rôle qu'y jouent les réseaux sociaux et Internet<sup>87</sup>. L'impact de la désinformation sur les résultats électoraux mériterait d'ailleurs aussi d'être scientifiquement mieux établi<sup>88</sup>.

82 Leskovec J., Backstrom L. et Kleinberg J. (2009), « Meme-tracking and the dynamics of the news cycle », *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, ACM, pp. 497-506.

83 Voir le chapitre précédent de ce rapport : « Les mécanismes psychosociaux de la désinformation ».

84 Fletcher, R., Cornia, A., Graves, L., et Nielsen, R. K. (2018). Measuring the Reach of « Fake News » and Online Disinformation in Europe. Reuters Institute factsheet, <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/Measuring%20the%20reach%20of%20fake%20news%20and%20online%20distribution%20in%20Europe%20CORRECT%20FLAG.pdf> Cordonier, L. et Brest, A. (2021). op. cit. ; voir chapitre précédent.

85 Vargo, C. J., Guo, L., et Amazeen, M. A. (2018), « The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016 », *New Media and Society*, 20-5, pp.2028–2049.

86 Mercier, H. (2020), *Not Born Yesterday: The Science of Who We Trust and What We Believe*, Princeton University Press.

87 A. Guess, B. Nyhan, B. Lyons, and J. Reifler, « Avoiding the echo chamber about echo chambers », Technical report, Knight Foundation, 2018. J. Tucker, A. Guess, P. Barbera, C. Vaccari, A. Siegel, S. Sanovich, D. Stukal, et B. Nyhan (2018), « Social media, political polarization, and political disinformation: A review of the scientific literature », *Technical report*, Hewlett Foundation.

88 Aral, S. et etEckles, D. (2019), « Protecting elections from social media manipulation », *Science*, 365(6456), pp.858–861. <https://science.sciencemag.org/content/365/6456/858>

Le plus probable est que de tels phénomènes soient plurifactoriels et que leur explication ne puisse donc reposer sur la seule influence du monde numérique et des perturbations qu'il provoque.

Plus essentiel encore, il faut souligner que nous ne sommes pas incompetents quand il s'agit de détecter les fausses informations et que, dans l'ensemble, nous les trouvons moins plausibles que les nouvelles authentiques<sup>89</sup>. Nous possédons en nous-mêmes les ressources pour prévenir certains dangers des fausses nouvelles, point que développera le chapitre portant sur l'esprit critique et l'EMI.

---

89 Allen, J., Arechar, A. A., Pennycook, G. et Rand, D.G. (2020), « Scaling up fact-checking using the wisdom of crowds », <https://psyarxiv.com/9qdza/>

## Les bouleversements algorithmiques

La prudence s'impose donc lorsque l'on aborde le lien entre les algorithmes et la configuration des plateformes avec les phénomènes sociaux négatifs qu'ils sont accusés de produire. Pour autant, les réseaux numériques ont des spécificités<sup>90</sup> qui contribuent de manière inédite à l'amplification de ces nuisances. D'abord, la taille des réseaux numériques, le nombre de contacts que nous pouvons y avoir ou encore la visibilité potentielle des messages que nous y faisons circuler sont plus importants que jamais. Ensuite, dans les interactions hors ligne, la proximité spatiale entre les individus les incite généralement à éviter l'insulte ou l'invective ; les réseaux sociaux n'offrent pas cette caractéristique pacificatrice. Les échanges sur Internet incitent souvent à l'intolérance<sup>91</sup> et à ce que l'on appelle la désinhibition numérique<sup>92</sup>. Enfin, la multiplicité des sources d'informations a tendance à favoriser une balkanisation des perceptions de la réalité, comme nous l'avons évoqué dans l'introduction de ce rapport.

Les principaux effets de la révolution algorithmique en matière d'organisation de l'information peuvent être subdivisés en trois axes, que nous allons explorer successivement dans le reste de ce chapitre :

- ◆ l'éditorialisation algorithmique : la manière dont les algorithmes régissent à la fois l'ordre et la fréquence d'apparition des informations, selon une logique de captation de l'attention ;
- ◆ le calibrage social : la façon dont les réseaux sociaux altèrent la perception de la représentativité et de la popularité de certains points de vue ;
- ◆ l'influence asymétrique : le fait qu'Internet permette à des individus motivés d'accéder à une visibilité numérique qui excède de beaucoup leur représentativité, rendant ainsi possible la prévalence de certains discours extrêmes qui profitent des conditions numériques pour sortir de leur espace de radicalité et diffuser leurs argumentaires.

90 Kozyreva, A., Lewandowsky, S. et Hertwig, R. (2020), «Citizens versus the internet: Confronting digital challenges with cognitive tools», *Psychological Science in the Public Interest*, 21-3, pp.103-156.

91 Rossini, P. (2020), «Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk.» *Communication Research*, 10.1177/0093650220921314.

92 Suler, J. (2004), « The Online Disinhibition Effect », *CyberPsychology and Behavior*, 7-3, pp.321-326.

L'un des rôles des médias est d'éditorialiser les informations d'actualité, c'est-à-dire de les sélectionner et de les hiérarchiser pour son public. En effet, la masse de données disponibles est telle qu'il nous est impossible de l'embrasser et d'en prendre connaissance d'un seul regard, a fortiori depuis le développement d'Internet. Pour un journal papier traditionnel, par exemple, c'est la salle de rédaction et son rédacteur en chef qui choisissent et organisent l'information jugée pertinente, idéalement suivant les normes déontologiques de la profession. Chacun comprend qu'un article en première page, occupant plus de colonnes ou assorti de photos, gagnera davantage de visibilité. De la même façon, les plateformes éditorialisent l'information. Elles le font cependant selon une logique algorithmique qui reste opaque pour les utilisateurs<sup>93</sup>. Lorsqu'une recherche est faite sur Google ou YouTube, ou encore qu'un fil Facebook est ouvert, certaines informations sont présentées préférentiellement et ont de ce fait plus de chances d'être prises en compte par l'internaute.

Il se trouve que les requêtes adressées par les utilisateurs à leur moteur de recherche sont susceptibles d'amplifier leurs biais, notamment politiques<sup>94</sup>, parce que les intelligences artificielles sont sensibles aux préférences partisans des individus telles qu'elles sont révélées par les mots-clés qu'ils utilisent<sup>95</sup>. Ces recherches peuvent altérer la perception de certains sujets<sup>96</sup>, notamment parce que les premiers résultats d'une requête possèdent une prévalence cognitive<sup>97</sup>. Le discret travail d'éditorialisation de l'information des algorithmes pourrait même, dans certaines circonstances, influencer sur les préférences électorales de leurs utilisateurs<sup>98</sup>.

Ce que nous pourrions penser relever de notre liberté de choix se révèle ainsi, parfois, le produit d'architectures numériques influençant nos conduites. Cette question architecturale nous encourage à considérer, en plus de la logique algorithmique, celle du design. Le terme inquiétant de «dark patterns» a ainsi cristallisé ces dernières années les inquiétudes autour de la capacité du design des plateformes à se jouer des régularités de notre système cognitif, jusqu'à nous faire prendre des décisions malgré nous.

---

93 Burrell, J. (2016), «How the machine 'thinks': Understanding opacity in machine learning algorithms», *Big Data et Society*, 3:205395171562251.

94 Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., et Wilson, C. (2018) «Auditing partisan audience bias within google search.» *Proceedings of the ACM on Human-Computer Interaction*, 2:1–22,

95 Mustafaraj, E., Lurie, E. et Devine, C. (2020) «The case for voter-centered audits of search engines during political elections.» *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 559–569, 20. Trielli, D., et Diakopoulos, N. (2020). «Partisan search behavior and Google results in the 2018 U.S. midterm elections. *Information, Communication et Society*, 10.1080/1369118X.2020.1764, 2020.

96 Allam, A., Schulz, P. J., et Nakamoto, K. (2014) The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: Two experiments manipulating google output. *Journal of Medical Internet Research*, 16:e100.

97 Novin, A., et Meyers, E., «Making sense of conflicting science information: Exploring bias in the search engine result page.» *Proceedings of the 2017 Conference on Human Information Interaction and Retrieval*, 175–184. ACM, 2017.

98 Epstein R. et Robertson, R. E., «The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections.» *Proceedings of the National Academy of Sciences*, 112:E4512–E4521, 2015.



La question des darks patterns<sup>99</sup> (interfaces conçues dans le but de manipuler ou tromper les utilisateurs) et celle de la possibilité de leur régulation exigent de prendre en compte la problématique du design des interfaces utilisateurs. En effet, même au-delà de toute intention maligne, les choix de design ont nécessairement une influence sur le comportement des utilisateurs sur les plateformes numériques. Ils définissent le contexte au sein duquel les individus sont amenés à exercer leur pouvoir décisionnel. C'est pourquoi Cass Sunstein propose de concevoir les designers comme des «architectes de choix»<sup>100</sup>, soulignant ainsi la responsabilité qui est la leur.

Ce pouvoir d'orienter les choix individuels soulève d'importantes questions sociales, éthiques et politiques, notamment celle de la collecte des données personnelles. Ces pratiques de design sont-elles conformes aux normes démocratiques de nos sociétés? Peut-on considérer par exemple qu'un consentement éclairé, au partage des données personnelles a été recueilli si l'option de refus était difficile d'accès ou invisibilisée? Est-il tolérable que certains de nos biais cognitifs soient instrumentalisés pour capter notre attention et dégager un profit? D'une façon générale, en quoi les choix des individus relèvent-ils vraiment de leur décision personnelle dans ce contexte?

La résolution de ces questions ne peut être laissée aux seules plateformes, car elles dépassent le cadre strictement technique ou juridique et nécessitent l'implication du régulateur et de la société civile. Il apparaît dès aujourd'hui nécessaire de travailler à une grille d'analyse rigoureuse des pratiques de design et de leurs conséquences sur les individus et la société. Cette tâche, qui permettra de caractériser ce qui relève d'un design abusif ou trompeur, suppose de recourir aux compétences et savoirs des professionnels du design, mais aussi des spécialistes des sciences humaines et sociales (psychologues, sociologues et philosophes).

99 CNIL (janvier 2019) «La forme des choix. Données personnelles, design et frictions désirables», Cahiers Innovation et Prospective n°6.

100 Sunstein, C. (2015) *Choosing not to choose*



**Les interfaces utilisateurs étant appelées à évoluer et à prendre de plus en plus d'importance dans nos relations sociales, il est crucial de se doter de moyens pérennes d'analyser leur influence et de la réguler. Il faut ainsi encourager une montée en compétence du régulateur, en normalisant le recours à des experts de l'intrication entre design, psychologie et éthique. À ce titre, un dialogue avec le monde de la recherche doit être favorisé dans le but de s'assurer de la réactivité et de l'effectivité de l'action publique.**

## **Recommandation**

**Ouvrir une réflexion en vue d'une régulation sur l'importance de la question du design des interfaces utilisateurs (R2).**

La logique qui oriente la conception de ces architectures numériques ne répond, la plupart du temps, qu'à un objectif économique : il s'agit pour les plateformes numériques de retenir aussi longtemps que possible l'attention de leurs utilisateurs, afin de pouvoir la convertir en ressources financières au moyen d'espaces publicitaires payants ; ou encore de les inciter à partager plus de données, in fine monétisables, qu'il n'est nécessaire au strict fonctionnement du service. Pour cela, toutes les tactiques sont bonnes - aussi longtemps qu'elles ne sont pas interdites. Les plateformes s'ajustent ainsi sans arrêt à nos conduites et aux traces que nous laissons dans l'univers numérique. Ces ajustements visent à satisfaire nos inclinations cognitives naturelles, nous faisant alors creuser de plus en plus profondément ces sillons numériques qui finissent par nous enfermer.

Le problème est que cet enfermement n'est pas seulement dommageable pour les individus – ce qui serait en soi une raison suffisante pour s'en inquiéter – mais produit également des effets collectifs négatifs. Mark Zuckerberg lui-même a reconnu, en 2018, que les algorithmes de classement fondés sur l'engagement pouvaient être dangereux. Facebook a par exemple remarqué qu'un émoticône « colère » suscitait généralement plus d'engagements sur un post du réseau social qu'un simple « like ». Pour profiter de cet effet d'engagement, l'entreprise a calibré son algorithme de manière à attribuer 5 fois plus de poids à ces expressions d'indignation, assurant ainsi aux contenus concernés une visibilité maximale dans les fils d'actualité<sup>101</sup>. Dans ces conditions, comment s'étonner des effets de polarisation affective que l'on observe ?

101 Série d'articles relatifs aux *Facebook files* dans le *Wall Street Journal* et audition de Frances Haugen.

De la même façon, l'introduction de la modification MSI (Meaningful Social Interactions) devait permettre de corriger la survalorisation du contenu le plus viral, en introduisant un critère de « santé du réseau » mesurant la probabilité qu'un post soit aimé ou repartagé. Or, cette modification qui devait favoriser les interactions avec le cercle social restreint a eu pour effet pervers de mettre en avant les contenus les plus extrêmes. Cela s'expliquerait par le fait que les gens sont généralement peu sensibles aux repartages de contenus, sauf s'ils viennent de leurs cinq amis les plus proches, ou si le contenu est suffisamment extrême pour attirer leur attention. Pour autant, une fois cet effet pervers repéré, Facebook n'a pas désactivé le MSI<sup>102</sup>.

Ces remarques sont bien entendu valables pour d'autres plateformes numériques, et notamment YouTube qui cherche, elle aussi, à maximiser le temps de présence de ses utilisateurs au moyen d'un algorithme de recommandations personnalisées<sup>103</sup>. Une étude menée par Guillaume Chaslot, ancien ingénieur de Google<sup>104</sup>, a montré que l'algorithme de YouTube conduisait vers des contenus de plus en plus extrêmes, pavant ainsi la voie à la radicalisation<sup>105</sup>. Cet algorithme a été jugé responsable d'une partie de la progression des extrêmes-droites allemandes et états-uniennes<sup>106</sup>.

102 Audition de Frances Haugen.

103 Covington, P., Adams, J. and E. Sargin, E. (2016), «Deep neural networks for YouTube recommendations», *Proceedings of the 10th ACM Conference on Recommender Systems—RecSys*, 16.

104 <https://www.wsj.com/articles/how-YouTube-drives-viewers-to-the-internets-darkest-corners-1518020478>

105 Horta Ribeiro M., Ottoni, R. West V., Almeida, A.F. et Meira W. (2020), «Auditing radicalization pathways on YouTube», *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 131-141. ACM.

106 Kaiser, J. et Rauchfleisch, A. (2018), «Unite the right? How YouTube's recommendation algorithm connects the U.S. far-right», *Medium*, 11-04.

## Perturbation du calibrage social

Les mécanismes cognitifs de coalition et d'affiliation sociale sont profondément inscrits dans notre nature<sup>107</sup>. De ce fait, notre opinion naissante sur une question donnée peut largement être influencée par la visibilité de celle que les autres ont exprimée à son sujet, en particulier s'ils font partie de notre réseau d'amis ou nous paraissent socialement semblables. La numérisation des relations sociales et la multiplication des producteurs de contenus d'information perturbent beaucoup notre calibrage social<sup>108</sup>, c'est-à-dire l'accès raisonné que nous avons à l'avis des autres. L'altération de notre perception de la prévalence des opinions des autres peut prendre au moins deux formes.

D'une part, elle conduit à accorder une prime aux contenus que la métrique numérique a rendus populaires. Ainsi, les algorithmes qui organisent la visibilité de l'information ont pour but de maximiser l'attention des utilisateurs et leur engagement plutôt que de proposer des sources fiables<sup>109</sup> et équilibrées<sup>110</sup>. Ils le font par exemple en mettant particulièrement en avant les contenus bénéficiant du plus de commentaires, de « likes » ou de partages. Cette disposition pourrait paraître raisonnable en partant du principe qu'une intelligence collective a plus de chances de faire émerger les points de vue solides et argumentés. Il n'en est rien, en raison de l'existence de ce que l'on nomme le biais de popularité<sup>111</sup> qui, comme le montre la recherche<sup>112</sup>, réduit la qualité globale du contenu de l'information. À un certain niveau de popularité, la diffusion d'un article, par exemple, ne cessera de s'amplifier : plus une personne est exposée à une idée, plus les chances seront grandes qu'elle la fasse sienne et finisse par la diffuser à son tour. Le passage de l'information au crible des métriques numériques affecte donc notre calibrage social.

107 Cordonier, L. (2018), *La nature du social. L'apport ignoré des sciences cognitives*, Paris, Puf.

108 Tufekci, Z.. Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colorado Technology Law Journal*, 13:203–218, 2015.

109 Dechêne, A., Stahl, C., Hansen, et J., Wänke, M. (2010), «The truth about the truth: A meta-analytic review of the truth effect», *Personality and Social Psychology Review*, 14-2, pp. 238-257.

110 Borah, P. et Xiao, X. (2018), «The importance of "likes": The interplay of message framing, source, and social endorsement on credibility perceptions of health information on Facebook», *Journal of Health Communication*, 23-4, pp. 399-411.

111 Zhu, Z., He, Y., Zhao, X. et Caverlee, J. (2021), «Popularity Bias in Dynamic Recommendation» In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery et Data Mining (KDD '21)*. Association for Computing Machinery, New York, NY, USA, 2439–2449. DOI: <https://dl.acm.org/doi/10.1145/3447548.3467376>

112 Nematzadeh, A., Ciampaglia, G. L., Menczer, F., et Flammini, A. (2017), «How algorithmic popularity bias hinders or promotes quality», arXiv e-prints.



## ➔ Recommandation

**Permettre aux utilisateurs de mieux se représenter l'état du réseau et la prévalence réelle des opinions en désactivant par défaut les métriques de popularité et l'éditorialisation algorithmique, et en mettant en avant des métriques permettant de juger de la qualité épistémique des contenus (historique de partage notamment) (R3).**

D'autre part, nous avons tendance à nous associer sur les réseaux sociaux (comme dans la vie réelle) à des personnes qui nous ressemblent et partagent nos points de vue, et à nous éloigner de celles qui nous paraissent au contraire trop dissemblables (par exemple, en les supprimant de notre liste d'amis ou en les bloquant). Cette tendance à l'homophilie est banale<sup>113</sup>, mais elle est facilitée sur les réseaux sociaux, car les points de vue des individus, ainsi que certaines de leurs caractéristiques psychologiques et sociales (goûts, préférences, appartenances de groupe, etc.) y sont souvent plus immédiatement visibles et mesurables que dans la vie hors ligne. En nous entourant progressivement, sans nous en rendre compte, de personnes qui nous ressemblent, partagent nos opinions et le manifestent en « likant » nos posts et en publiant des contenus allant dans le sens de nos idées, nous risquons d'avoir l'impression que ces dernières sont très majoritaires. Nous pouvons en effet facilement oublier que notre environnement numérique n'est nullement représentatif de l'ensemble de la population. Ainsi peuvent se former des communautés épistémiques au sein desquelles émergent des consensus illusoires et où s'opère un renforcement mutuel des opinions<sup>114</sup>.

113 Cordonier, L. (2018), *La nature du social. L'apport ignoré des sciences cognitives*, Paris, Puf.

114 Leviston, Z., Walker, I., et Morwinski, S.,. Your opinion on climate change might not be as common as you think. *Nature Climate Change*, 3:334–337, 2013.

## Influences asymétriques et radicalisation

Très tôt, plusieurs études<sup>115</sup> ont montré que, sur Internet, un petit nombre de personnes motivées pouvait influencer l'opinion. Internet a en effet permis l'émergence de ce que certains nomment, en référence aux thèses de l'école de Columbia, des « super leaders d'opinion<sup>116</sup> ». Les audiences colossales de certains internautes affaiblissent l'idée qu'Internet permettrait de « démocratiser la démocratie » : en réalité, dans la cité numérique, certains ont beaucoup plus voix au chapitre que les autres<sup>117</sup>. Dans la logique du winner takes all, leur audience est accrue par le système de recommandation des plateformes<sup>118</sup>. Cela ne serait pas nécessairement problématique s'il n'était démontré que cette configuration favorisant les influenceurs numériques les plus visibles constitue un facteur essentiel de la propagation virale des fausses informations<sup>119</sup>. Ceux-ci ne sont pas nécessairement producteurs ou relais de fausses informations, mais lorsqu'ils cèdent à la tentation d'en partager, ils deviennent les causes principales de cascades de désinformation<sup>120</sup>.

### → Recommandation

**Encourager les plateformes à une modération plus attentive des influenceurs afin de les responsabiliser. Les conséquences des informations produites ou diffusées par un compte doté d'une grande visibilité numérique sont potentiellement plus importantes que pour des comptes de faible audience (R4 + cf. aussi le chapitre « Droit numérique »).**

D'une façon générale, la motivation des acteurs sur ce marché cognitif peut leur conférer une visibilité qui dépassera leur représentativité. Il se trouve que, pour le meilleur et pour le pire, certains groupes motivés ont montré qu'ils étaient capables de préempter une part déséquilibrée de la visibilité numérique. Ainsi, sur Facebook, les mouvements anti-vaccins ont réussi – avant la pandémie – à occuper une place dominante face aux pro-vaccins<sup>121</sup>.

115 Watts D. J. et Strogatz S. (1998) « Collective dynamics of «Small-World» Networks », *Nature*, vol. 393, no 6684, p. 440-442.

116 Gladwell, M. (2002) *The Tipping Point. How Little Things can make a Big Difference*. Boston - New York : Little, Brown and Company.

117 Lance Bennett, W. et Manheim, J.B. (2006), «The one-step flow of communication», *The ANNALS of the American Academy of Political and Social Science*, 608, pp. 213–232.

118 Gupta, P., Goel, A., Lin, J., Sharma, A., Wang, D., et Zadeh, R. (2013), «WTF: The Who to Follow service at Twitter», *Proceedings of the 22<sup>nd</sup> International Conference on World Wide Web—WWW '13*. ACM.

119 Pastor-Satorras, R. et Vespignani, A. (2001), «Epidemic spreading in scale-free networks», *Physical Review Letters*, 86, pp. 3200–3203.

120 Watts, D.J. (2002), «A simple model of global cascades on random networks», *Proceedings of the National Academy of Sciences*, 99, pp. 5766–5771.

121 Wadman, M. (2020), <https://www.science.org/content/article/vaccine-opponents-are-gaining-facebook-battle-hearts-and-minds-new-map-shows>

Certaines analyses proposent de généraliser ces observations en montrant que la tendance des réseaux sociaux est d’invisibiliser les modérés au profit des opinions extrêmes <sup>122</sup>.

## ➔ Recommandations

- ◆ **Donner plus d’influence à la compétence en mettant en avant les comptes d’experts et en amplifiant leur contenu (lorsqu’ils portent sur un sujet relatif à leur expertise) (R5).**
- ◆ **Veiller à ce que, sur certains sujets fermement établis, le classement algorithmique n’induisse pas en erreur le public sur l’état réel des connaissances. Pour cela, encourager un dialogue entre les plateformes et les institutions scientifiques afin que l’existence d’un consensus soit reflétée dans la visibilité accordée aux diverses opinions (R6).**

En dehors des réseaux sociaux, le classement proposé par un moteur de recherche comme Google, par exemple, peut être influencé par l’activité plus ou moins coordonnée de certains réseaux militants. Ainsi le spam de requêtes et de publications peut-il modifier la hiérarchie des résultats d’un moteur de recherche <sup>123</sup>. Une technique bien identifiée <sup>124</sup> consiste pour certains mouvements - notamment les suprémacistes blancs <sup>125</sup> - à exploiter les data voids. Il s’agit de thèmes qui produisent peu de résultats lorsqu’on les recherche et sont donc facilement préemptables par une manipulation coordonnée. C’est notamment le cas lorsqu’un fait d’actualité vient de se produire (par exemple, un attentat) et n’a pas encore généré beaucoup d’articles. Dès lors, si un groupe motivé par la manipulation de l’opinion agit rapidement, il peut, au moins provisoirement, détourner les premières recherches vers des interprétations idéologisées de l’événement.

122 Bail, C. (2021), *Breaking the Social Media Prism*, Princeton University Press.

123 Takis Metaxas, P. (2009), «Web spam, social propaganda and the evolution of search engine rankings», *International Conference on Web Information Systems and Technologies*, pp.170–182. Springer.

124 Cf. audition Google.

125 Golebiewski, M. and Boyd, D. (2019), «Data voids: Where missing data can easily be exploited», *Technical report, Data et Society*.

L'action de tels groupes motivés peut contribuer à produire des bulles épistémiques<sup>126</sup>, espaces numériques à l'intérieur desquels l'esprit critique peine à faire valoir ses droits<sup>127</sup>. Dans ces communautés virtuelles, les fausses informations peuvent se diffuser sans rencontrer beaucoup de contradictions. Il est documenté qu'elles alimentent l'extrémisme et la polarisation affective<sup>128</sup>. Ces groupes peuvent par ailleurs agir de façon plus ou moins coordonnée pour signaler en essaim des comptes qui contrarient leur combat idéologique et obtenir des suspensions ou des bannissements.

## ➔ Recommandation

**Se prémunir contre le risque de sur-modération en analysant plus finement les signalements d'utilisateurs (signalements en essaim) (R7).**

126 Pour la différence entre la notion de chambre d'écho et de bulle épistémique cf. Nguyen, C. (2020), « Echo chambers and epistemic bubbles », *Episteme*, 17(2), pp.141-161. doi:10.1017/epi.2018.32

127 Duffy, R. (2018), *The perils of perception*, Londres, Atlantic Books.

128 Lau, R. R., Andersen, D. J., Ditonto, T. M., Kleinberg, M. S., et Redlawsk, D. P. (2017). Effect of media environment diversity and advertising tone on information search, selective exposure, and affective polarization. *Political Behavior*, 39(1), 231–255. Tsfaty, Y., et Nir, L. (2017). Frames and reasoning: Two pathways from selective exposure to affective polarization. *International Journal of Communication*, 11, 22. Suhay, E., Bello-Pardo, E., et Maurer, B. (2018). The polarizing effects of online partisan criticism: Evidence from two experiments. *The International Journal of Press/ Politics*, 23(1), 95–115.

## Conclusion du chapitre 2

Les grandes plateformes numériques ne restent donc pas totalement inertes face au danger des fausses informations. Facebook a ainsi promu des bandeaux incitant ses utilisateurs à la vigilance au sujet des échanges impliquant le vaccin ou la Covid-19. La plateforme de vidéos en ligne YouTube a aussi officiellement fait savoir que « Les contenus qui prétendent faussement que les vaccins approuvés sont dangereux et provoquent des effets chroniques sur la santé, qui affirment que les vaccins ne réduisent pas la transmission des maladies ou le fait de contracter des maladies ou qui contiennent des informations erronées sur les substances contenues dans les vaccins seront supprimés ». À ce titre, plus de 130 000 vidéos ont été effacées depuis un an<sup>129</sup>. D'une autre façon, Twitter impose désormais une fenêtre surgissante à tout utilisateur qui souhaite partager un lien sans en avoir lu le contenu. Or 59% des personnes qui partagent des articles sur les réseaux sociaux n'en ont lu que le titre et rien du contenu<sup>130</sup>. Audrey Herblein-Stoop, directrice des affaires publiques de Twitter, affirme que cette disposition aboutit bien à ce qu'un grand nombre d'utilisateurs renoncent à partager un article sans l'avoir lu<sup>131</sup>. Des dispositifs de cette nature ont aussi été mis en avant par les représentants de Tik Tok lors de leur audition auprès de la commission.

Depuis plusieurs années, de grandes plateformes numériques ont adhéré au Global Network Initiative<sup>132</sup>, qui les engage à la fois sur le thème de la défense des droits de l'Homme et de la transparence. En 2010, Google a introduit un rapport annuel de transparence<sup>133</sup>, notamment sur la question épineuse de la suppression de contenus et de profils (YouTube, Google), suivi par Twitter en 2012 puis Facebook en 2013 et bien d'autres.

La réponse la plus radicale des entreprises du numérique dans ce domaine est la suppression des comptes jugés problématiques, mesure que l'on nomme désormais la « déplateformisation ». Est-ce une façon efficace de lutter contre la désinformation ? Aux États-Unis, ce sont les membres de QAnon, les suprémacistes blancs ou encore les conspirationnistes qui ont fait les frais de cette politique. En France, il s'est produit la même chose pour des personnages comme Alain Soral ou Dieudonné, aussi bien sur Facebook que sur YouTube où ils bénéficiaient d'une audience importante. La littérature scientifique sur les effets de la « déplateformisation », qui va croissant, tend à montrer son efficacité globale. Bien sûr, ceux qui sont bannis des grands réseaux cherchent à se retrouver sur des plateformes alternatives, comme Telegram ou Parler, mais, partout où on a pu l'observer, ce mouvement de migration a abouti à une fragmentation

129 [https://www.lemonde.fr/pixels/article/2021/09/29/YouTube-renforce-sa-politique-de-lutte-contre-les-vidéos-antivaccins\\_6096493\\_4408996.html](https://www.lemonde.fr/pixels/article/2021/09/29/YouTube-renforce-sa-politique-de-lutte-contre-les-vidéos-antivaccins_6096493_4408996.html)

130 Gabielkov, M., Ramachandran, A., Chaintreau, A. et Legout, A. (2016), « Social Clicks: What and Who Gets Read on Twitter? », *ACM SIGMETRICS Performance Evaluation Review*, 44-1, pp.179-192

131 Audition du 15/11/21.

132 Maclay, C. M. (2010). Protecting privacy and expression online: Can the Global Network Initiative embrace the character of the net. In R. J. Deibert, J. Palfrey, R. Rohozinski, et J. Zittrain (Eds.), *Access Controlled: The Shaping of Power, Rights, and Rule in Cyberspace* (pp. 87–108). Cambridge, MA: MIT Press.

133 Drummond, D. (2010). Greater transparency around government requests. Google (blog), April 20.

des communautés qui les affaiblit, même si le risque est qu'elles s'y radicalisent plus encore<sup>134</sup>. Quelles que soient les mesures qui ont été utilisées pour évaluer l'efficacité de la « déplateformisation », on observe à chaque fois une baisse de l'influence des personnes bannies. Ainsi, 11 000 comptes supprimés de YouTube ayant migré vers la plateforme BitChute ont connu une nette baisse d'audience<sup>135</sup>.

D'une autre façon, l'analyse de 49 millions de tweets a montré que l'exclusion de comptes conspirationnistes comme celui d'Alex Jones avait nettement amoindri la toxicité de leurs soutiens sur les réseaux sociaux<sup>136</sup>.

Les réseaux sociaux n'attirent que si les utilisateurs n'ont pas l'impression de s'y isoler. Sur ce point, le bannissement de Donald Trump des réseaux sociaux doit faire réfléchir. L'ex-président demeure une personnalité incontournable aux États-Unis, et son compte Twitter était suivi par 89 millions d'individus. Dans ces conditions, son intention de créer son propre réseau social, Truth Social, en 2022, n'est pas anodine. Sur la forme, son design ressemblera beaucoup à celui de Twitter mais, sur le fond, on peut craindre des règles de modération si permissives qu'elles permettront un épanouissement inédit des expressions de la radicalité. Après s'être servi des réseaux sociaux comme d'un moyen de désintermédiation entre les électeurs et lui, il prétend à présent lutter contre la « tyrannie des géants des technologies » et pourrait réussir son pari. De tous les acteurs de la société de « post-vérité », Donald Trump est celui qui bénéficie du plus grand capital de visibilité sociale, ce qui lui permettra peut-être de briser le plafond qu'aucune plateforme alternative n'a réussi à franchir jusqu'à présent. Si ce pari était gagné, le fossé entre les deux rives de la société américaine risquerait de se creuser davantage.

---

134 Urman, A. et Stefan Katz, S. (2020), « What they do in the shadows: examining the far-right networks on Telegram, *Information, Communication et Society*, DOI: 10.1080/1369118X.2020.1803946.

135 Rauchfleisch, A. et Kaiser, J. (2021), « Deplatforming the Far-right: An Analysis of YouTube and BitChute », <https://ssrn.com/abstract=3867818> or <http://dx.doi.org/10.2139/ssrn.3867818>

136 Jhaver, H., Boylston, C. Yang, D. et Bruckman, A. (2021), « Evaluating the Effectiveness of Deplatforming as a Moderation Strategy on Twitter », *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, 381, DOI: <https://doi.org/10.1145/3479525>

## CHAPITRE 3

# ➔ Économie des infox

La circulation des infox et des contenus à caractère conspirationniste est amplifiée par la visibilité et la viralité sans précédent qu'ont acquis la désinformation et la mésinformation au cours des dernières années. Les infox occasionnent des dépenses considérables pesant sur l'ensemble de l'économie.

En matière de santé publique, les dépenses engagées aux États-Unis pour soigner des personnes atteintes de maladies comme la rougeole, que la vaccination parvient pourtant à prévenir efficacement, s'élèvent à environ 9 milliards de dollars. Un coût qui concerne essentiellement des individus non-vaccinés influencés par des contenus hostiles aux vaccins<sup>137</sup>. C'est ce que met en exergue une étude publiée par la société de cybersécurité CHEQ et réalisée par Roberto Cavazos, professeur d'économie à l'université de Baltimore. Ce dernier estime que le coût des infox pour les acteurs économiques américains approche les 78 milliards de dollars en 2019. Les fausses informations infligeraient ainsi des dommages aux marchés boursiers mondiaux, se traduisant par des pertes qui pourraient aller jusqu'à 0,05% de la valeur totale des marchés soit 39 milliards de dollars<sup>138</sup>. Les dépenses consenties par les grandes en-

treprises pour prévenir le risque réputationnel et démentir les fausses informations qui les visent devrait en outre dépasser, selon la même étude, les 9,5 milliards de dollars d'ici 2022.

Même s'ils ne reposent que sur des estimations, ces chiffres montrent que la désinformation provoque un affaiblissement substantiel de nos économies. Cet état de fait n'est rendu possible que par les revenus que la désinformation parvient à générer, par de multiples canaux : vente de produits (livres et DVD conspirationnistes, vêtements, appareils électriques, cryptomonnaies...) et de services (formations, stages, polices d'assurances...), collecte de dons, levées de fonds participatives ou encore revenus publicitaires, qui permettraient à de nombreux médias de désinformation de prospérer. Comme l'affirme Roberto Cavazos, «la prolifération des fake news est liée au développement d'un marché publicitaire en ligne ultra-lucratif et ultra-compétitif. Tout ce qui est extrême et sensationnaliste attire les clics, donc gonfle les revenus. Du coup, des myriades de supports non identifiés produisent des contenus à la chaîne, et ces mauvaises informations entraînent de mauvaises prises de décision<sup>139</sup>.»

137 Myers, K., (10 avril 2019) «Anti-vaxxers are costing Americans billions each year», *Yahoo Finance*, <https://finance.yahoo.com/news/antivaxxers-costing-americans-billions-each-year-191839191.html>

138 Cavazos, R. (2019) «The Economic Cost of Bad Actors on the Internet: Fake News in 2019», *CHEQ*, <https://s3.amazonaws.com/media.mediapost.com/uploads/EconomicCostOfFakeNews.pdf>

139 Berthelot, B. et Eliakim, P. (20 avril 2021), «Fake news : comment les théories du complot fragilisent notre économie», *Capital*, <https://www.capital.fr/economie-politique/fake-news-comment-les-theories-du-complot-fragilisent-notre-economie-1400650>

## La publicité programmatique, source de revenus substantielle pour la désinformation

L'une des principales manières de générer des revenus en ligne est, pour des éditeurs de sites web ou de blogs, de monétiser leur audience en installant sur leurs plateformes des emplacements publicitaires pouvant prendre la forme de bannières (ou « bandeaux »), de skycrapers (pour les formats verticaux) ou d'habillage en arrière-plan.

Deux types de prestations de publicité digitale peuvent être distinguées : la publicité classique, qui consiste en l'achat d'espace publicitaire, et la publicité dite « programmatique ».

La publicité programmatique est aujourd'hui utilisée par de nombreuses entreprises pour s'adresser de manière spécifique à un grand nombre d'internautes, moyennant un coût financier et humain modéré. Selon la société d'évaluation des sites d'actualité NewsGuard, elle représente aujourd'hui « plus de 85% de toute la publicité digitale, pour un total de 80 milliards de dollars de dépenses annuelles aux États-Unis en 2020 <sup>140</sup> ».

Son originalité réside dans le fait que les campagnes ne consistent pas à afficher une publicité sur un espace publicitaire précis (tel ou tel site) à laquelle sont uniformément exposés tous les visiteurs pendant une période donnée, mais à s'adresser à une audience bien déterminée. Pour ce faire, la publicité programmatique recourt au système des enchères. Les annonceurs peuvent ainsi acheter de l'espace publicitaire de façon automatisée (il s'écoule 120 à 150 millisecondes en moyenne entre la proposition d'une enchère et la réponse à cette proposition à raison d'environ 15 à 20 milliards d'enchères par jour en France) tout en ciblant les utilisateurs en fonction de leurs centres d'intérêts, de leur âge, de leur genre ou encore de leur localisation géographique. Ces critères sont induits algorithmiquement grâce aux données personnelles et traces numériques laissées par les utilisateurs lors de leur navigation en ligne. Comme le définit le décret n° 2017-159 du 9 février 2017 relatif aux prestations de publicité digitale, ces campagnes « s'appuient sur des méthodes d'achat de prestations en temps réel sur des espaces non garantis, notamment par des mécanismes d'enchères, pour lesquelles les critères déterminants de l'achat sont le profil de l'internaute et l'optimisation de la performance du message. »

Cependant, il est apparu depuis plusieurs années que ces annonces programmatiques se retrouvent fréquemment sur des sites propageant notoirement et souvent de manière répétée des contenus haineux, conspirationnistes, attentatoires à la dignité humaine et à l'égalité femmes-

---

140 « NewsGuard lance l'outil « Publicité responsable dans les médias » (« RANS »), pour permettre aux marques de ne plus financer la désinformation, tout en ramenant des publicités vers les éditeurs de qualité », *NewsGuard*, 30 mars 2021. <https://www.newsguardtech.com/fr/press/newsguard-rans-annonce/>



hommes, ou encourageant des dérives sectaires, relevant de la désinformation caractérisée ou susceptibles de troubler la tranquillité publique. Les revenus publicitaires que ces sites en tirent constituent une manne financière considérable qui entretient la pollution informationnelle.

À l'heure actuelle, les prestataires de publicité programmatique sont invités à communiquer à l'annonceur « *toutes les mesures mises en œuvre [...] pour éviter la diffusion de messages publicitaires sur des supports illicites ou dans des univers de diffusion signalés par l'annonceur comme étant préjudiciables à l'image de sa marque et à sa réputation*<sup>141</sup> ». Rien ne les oblige toutefois à communiquer la liste exhaustive des sites sur lesquels leurs publicités peuvent se retrouver.

Les budgets publicitaires d'une fondation œuvrant pour la recherche contre le cancer ont ainsi pu contribuer, de fait, aux revenus d'un site proposant par ailleurs des traitements « alternatifs » pour soigner le cancer. Une célèbre ONG de protection de l'environnement a pu, de la même manière, participer au financement d'un site proposant des contenus climatosceptiques<sup>142</sup>. Et des géants du numérique consacrent d'un côté des budgets pour lutter contre les fausses informations tout en contribuant, de l'autre, via notamment leurs prestations de technologie publicitaire (« Ad Tech »), au financement de certains des sites qui les propagent.

Outre les aberrations auxquelles peut conduire un tel système, il permet à une myriade de sites toxiques de prospérer sur la captation d'une source de revenus pratiquement intarissable. Selon NewsGuard en effet, nombreux sont les pourvoyeurs de désinformation « qui ne recevraient aucun soutien financier sans ces publicités involontaires ».

Or ce type de campagne publicitaire est en développement<sup>143</sup>. D'après une étude de Integral Ad Science (IAS), 52% des annonceurs déclarent que plus de la moitié de leurs dépenses publicitaires sont dorénavant effectuées en achats programmatiques. Pour 80%, cela représente un tiers ou plus de leurs dépenses. Selon la même étude, 42% des annonceurs estiment que la publicité programmatique accuse un manque de transparence qui ne leur permet pas d'être en situation de savoir sur quels espaces leurs campagnes sont diffusées ou de connaître l'identité de ceux qu'elles contribuent de ce fait à rémunérer.

141 Article 3 du décret n° 2017-159 du 9 février 2017 relatif aux prestations de publicité digitale.

142 Grossin, B. (29 août 2021) « Le financement des fake news : « Un Far West auquel participent des marques et des institutions publiques » », *France Culture*, <https://www.franceculture.fr/medias/le-financement-des-fake-news-un-far-west-auquel-participent-des-marques-et-des-institutions>

143 « Perfecting Your Supply Path: The Expansion of SPO in Programmatic », Integral Ad Science, 17 novembre 2021. <https://integralads.com/apac/insider/research-report-perfecting-your-supply-path/>

Dans la plupart des cas, il apparaît que les marques sollicitent les services d'une agence de publicité assurant le paramétrage et la diffusion de leurs campagnes en ligne. Ces agences leur proposent régulièrement d'utiliser des outils de *brand safety* («sûreté de la marque»), afin que leurs annonces ne s'affichent pas sur des sites qui pourraient nuire à leur image de marque (sites pornographiques, sites de ventes d'armes, etc.). Or, les sites proposant des contenus préjudiciables et néfastes, classifiés dans une sorte de zone «grise» (les contenus qui ne sont pas manifestement illicites et qui n'ont pas donné lieu à une décision de justice) sont très largement absents de ces outils de *brand safety*, si bien que nombreuses sont les marques qui se retrouvent à payer pour de la *brand safety* tout en finançant malgré elles des contenus conspirationnistes ou trompeurs.

Une étude de NewsGuard, en collaboration avec l'entreprise américaine d'analyse publicitaire Comscore, soutient que le secteur de la désinformation est «*en plein essor – avec 2,6 milliards de dollars de revenus publicitaires estimés versés aux diffuseurs de désinformation et de désinformation chaque année par les annonceurs programmatiques, y compris des centaines de millions de dollars de revenus soutenant les fausses allégations sur la santé, les mythes anti-vaccins, les infox électorales, la propagande partisane et d'autres fausses informations*<sup>144</sup>».

Certains «super désinformateurs», dont le trafic se chiffre en plusieurs millions d'utilisateurs par mois, attirent un grand nombre d'annonceurs. Ainsi, le site conspirationniste américain The Gateway Pundit (environ 30 millions de visites mensuelles<sup>145</sup>) aurait touché en 2020 l'équivalent de 200 000 euros par mois en moyenne de revenus issus de la publicité programmatique.

Les attitudes des annonceurs face à ce problème sont diverses : certaines marques mettent un point d'honneur à ne pas apparaître sur des sites de désinformation. D'autres semblent ignorer le problème, n'y ayant pas été sensibilisées. Certains annonceurs ne veulent pas savoir si leurs annonces finissent par apparaître sur des sites de désinformation. Enfin, une dernière catégorie d'annonceurs assume de financer des sites de désinformation en toute connaissance de cause.

Le collectif Sleeping Giants France, qui applique en France les méthodes de sensibilisation développées outre-Atlantique, alerte les annonceurs de la présence – le plus souvent involontaire – de leurs annonces sur des sites extrémistes et/ou consacrés à la diffusion massive d'infoc et de théories du complot. En quatre ans d'activité, le collectif a reçu près de 2 000 réponses positives à ses alertes, de la part d'annonceurs et d'agences de publicité. Plusieurs autres milliers d'annonceurs auraient, dans le même temps, procédé au retrait de leurs annonces de ces sites toxiques sans pour autant communiquer publiquement sur le sujet.

Une entreprise spécialisée dans l'offre de solutions contre la désinformation en ligne a contacté en 2018 plus de 200 annonceurs concernés par le problème de la rémunération d'ac-

144 Skibinski, M., «Rapport : La publicité sur les sites de désinformation», NewsGuard, septembre 2021. <https://www.newsguardtech.com/fr/special-reports/marques-milliards-sites-mesinformation-newsguard-comscore/>

145 SimilarWeb. <https://www.similarweb.com/fr/website/thegatewaypundit.com/#overview>

teurs toxiques par la publicité programmatique (parmi lesquelles des marques de la grande distribution, de la téléphonie mobile ou de l'industrie automobile, disposant pour la plupart d'entre elles de départements RSE, c'est-à-dire promouvant la « responsabilité sociale des entreprises ») pour leur proposer un audit gratuit de leurs campagnes publicitaires. Moins de 10% des entreprises contactées ont donné suite.

Le panorama du secteur de la publicité programmatique ne serait pas complet si l'on ne citait pas également, parmi ses acteurs, les fournisseurs de technologie publicitaire permettant l'implantation des annonces programmatiques tels que Google Ads (leader dans le domaine), Xandr (filiale d'AT&T), Taboola ou encore Criteo. Chacune de ces entreprises touche une commission lorsqu'un utilisateur est exposé à l'une de ses publicités.

En mars 2020, l'ONG américaine Global Disinformation Index (GDI), dont le but est d'assécher les financements des sites de désinformation, estimait à 76 millions de dollars le montant des revenus publicitaires fléchés « par inadvertance », sur le territoire de l'Union européenne, vers ces sites par des marques comme Amazon Prime, Burger King, Mercedes Benz, Samsung, Spotify ou Volvo<sup>146</sup>. En septembre 2019, GDI estimait à 235 millions de dollars les revenus publicitaires versés aux 20 000 sites de désinformation qui constituent sa base de données mondiale, au moyen de la publicité programmatique<sup>147</sup>. Selon plusieurs interlocuteurs contactés, les sites à plus haut risque de désinformation - tous pays confondus - seraient en réalité assez peu nombreux: environ 1200. Ainsi, si des fournisseurs de technologie publicitaire tels que Google ou Criteo décidaient de ne plus permettre la monétisation de tels sites, l'impact sur la société serait considérable. Il est par ailleurs à souligner que les règlements de ces entreprises (« publisher policies ») interdisent déjà, pour la plupart, la monétisation de tels sites mais ces règlements demeurent souvent inappliqués.

En mars 2021, NewsGuard a lancé un label « Publicité responsable dans les médias » (« RANS »)<sup>148</sup> prenant en compte non seulement le non-financement de la désinformation mais également la réorientation de ces dépenses publicitaires vers des sites journalistiques de qualité. L'obtention du label est conditionnée à la réalisation d'audits réguliers (au moins deux par an) vérifiant notamment l'actualisation des listes d'inclusion et d'exclusion utilisées par les annonceurs. En effet, un site a priori fiable peut, en un laps de temps limité, se muer en une plateforme toxique proposant des contenus de désinformation.

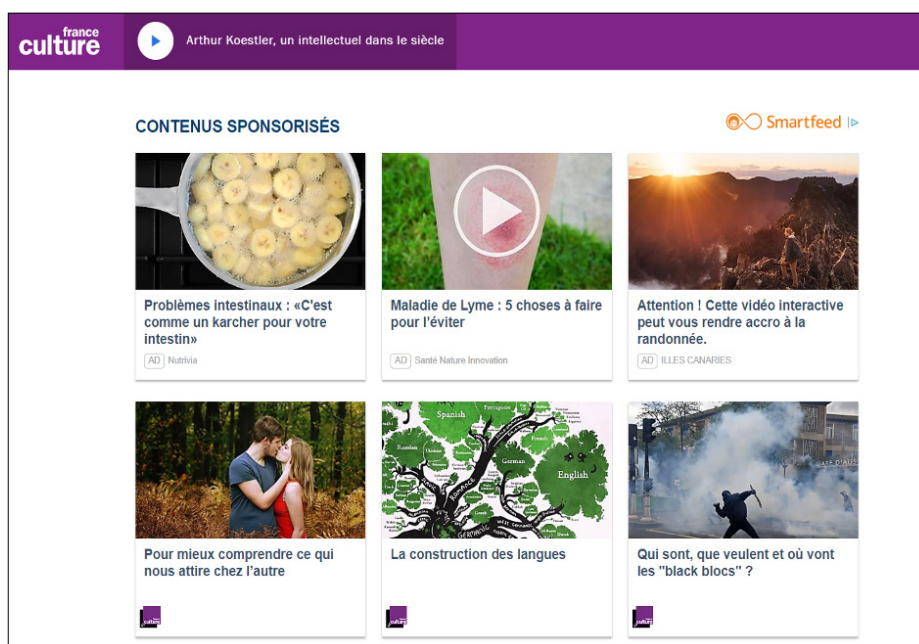
146 « Why is ad tech giving millions to EU disinformation sites? », *Global Disinformation Index*, 17 mars 2020. <https://disinformationindex.org/2020/03/why-is-ad-tech-giving-millions-to-eu-disinformation-sites/>

147 « The Quarter-Billion Dollar Question for Ad Tech », *Global Disinformation Index*, 22 septembre 2019. <https://disinformationindex.org/2019/09/the-quarter-billion-dollar-question-for-ad-tech/>

148 « NewsGuard lance l'outil « Publicité responsable dans les médias » (« RANS »)... », *art. cit.*, 30 mars 2021. <https://www.newsguardtech.com/fr/press/news-guard-rans-annonce/>

## Le trafic indirect généré depuis des sites de presse généralistes vers des sites « pièges à clics »

Plusieurs sites grand public se financent en partie par l'intégration de modules de recommandation de liens sponsorisés comme Outbrain (qui apparaît aussi parfois sous la mention SmartFeed). Il est fréquent que des sites d'information généralistes aient recours à ce type de dispositif. Or, ces liens sponsorisés peuvent renvoyer vers des sites « pièges à clics » (« clickbaits ») qui proposent parfois des contenus douteux, notamment en matière de santé.



Capture d'écran du site web de France Culture (9 avril 2019).

Ainsi, on a pu constater qu'un article douteux, issu d'un site (Santé Nature Innovation) qui, selon Le Monde, « diffuse parfois des informations fausses, exagérées ou dépourvues de sources, par exemple sur des aliments miracles ou sur de supposés dangers de la vaccination, démenties par l'immense majorité des spécialistes<sup>149</sup> », figurait parmi les contenus sponsorisés recommandés par un module Smartfeed installé sur une page du site de France Culture.

Un assainissement de l'environnement numérique, notamment de la sphère de la presse généraliste, passe donc par un découragement des redirections vers des sites « pièges à clics ».

La commission estime nécessaire de responsabiliser les différents acteurs de la publicité programmatique en mettant en œuvre les recommandations suivantes (R9) :

- ◆ Promouvoir l'investissement publicitaire responsable des entreprises en encourageant le recours, par les annonceurs, les régies, les agences publicitaires et surtout les fournisseurs de technologie publicitaire, à des « listes d'exclusion et d'inclusion de sites web » dynamiques, telles que celles élaborées par exemple par NewsGuard, Global Disinformation Index ou Storyzy.
- ◆ Engager un dialogue avec les fournisseurs de technologie publicitaire afin qu'ils aient recours à ce système, qui permettrait d'assécher de manière considérable l'économie des infox.
- ◆ Veiller à l'exemplarité en la matière des administrations et des entreprises publiques ayant recours à la publicité programmatique en généralisant pour elles le recours à des listes d'inclusion dynamiques.
- ◆ Envisager une obligation pour les entreprises engagées dans une démarche RSE à réaliser un audit indépendant annuel complet de leurs campagnes publicitaires programmatiques permettant de dresser la liste complète des adresses web (url) des sites sur lesquels ces campagnes sont placées, et à rendre publique cette liste.
- ◆ Encourager les organismes de labellisation comme l'AFNOR à tenir compte du problème du financement de la désinformation dans l'attribution de labels « responsables », en imposant des audits réguliers pour les entreprises candidates à ces labels.

149 Décodex, notice « Santé Nature Innovation », *Le Monde*, [https://www.lemonde.fr/verification/source/sante-nature-innovation/?xtor=AL-33281008-\[extension\]](https://www.lemonde.fr/verification/source/sante-nature-innovation/?xtor=AL-33281008-[extension])

- Envisager une obligation pour les fournisseurs de technologie publicitaire à sensibiliser leurs clients au risque de financer des sites toxiques dans le cas où ces derniers omettraient d'avoir recours à des listes d'exclusion dynamiques.
- Recommander aux sites de presse généralistes de bannir de leurs espaces publicitaires les liens sponsorisés renvoyant vers des sites pièges à clics de désinformation. Les encourager à ne plus faire appel aux services des sociétés publicitaires les associant à de tels liens sponsorisés.

### **LA MONÉTISATION D'UNE CHAÎNE YOUTUBE**

Les contenus de désinformation ou complotistes sont nombreux sur la plateforme de vidéos en ligne YouTube. Certains sont publiés par des chaînes monétisées par le biais de la publicité.

L'accord permettant de générer des revenus sur YouTube est le « Programme Partenaire YouTube ». L'éligibilité à ce programme suppose de satisfaire un certain nombre de conditions : compter au minimum 1000 abonnés au moins et 4000 heures de visionnage, ne pas faire l'objet d'un avertissement « actif » (en cours) pour non-respect des règles de la plateforme relatives aux contenus.

Une fois la chaîne monétisée, YouTube se réserve la possibilité de supprimer un contenu qui enfreindrait ses règles et d'envoyer au titulaire de la chaîne concernée un e-mail de mise en garde sans sanction. Si l'enfreinte aux règles d'utilisation se répète, le titulaire de la chaîne peut recevoir un avertissement. L'accumulation de trois avertissements en l'espace de 90 jours conduit à la suppression de la chaîne. Dans des cas exceptionnels (commission, même une seule fois, d'un abus grave relatifs aux règles d'utilisation de la plateforme), YouTube se réserve la possibilité discrétionnaire de clôturer la chaîne de l'utilisateur.



## L'alimentation de la désinformation par le financement participatif

Les plateformes de financement participatif permettent à des entreprises, des associations ou des particuliers de lever des fonds auprès de leur communauté pour financer une cagnotte, un projet ou un événement. Elles se rémunèrent à la commission ou en sollicitant les dons volontaires des donateurs (c'est le cas d'HelloAsso).

Il faut distinguer les plateformes de dons récurrents telles que Tipeee ou Patreon des plateformes dites de « cagnoting » comme Leetchi ou HelloAsso. Viennent enfin les plateformes de financement participatif avec contreparties telles que Ulule ou KissKissBankBank.

Certaines de ces plateformes ont été médiatisées pour avoir offert ou pour continuer à offrir à des projets douteux une solution de levée de fonds<sup>150</sup>. D'autres ont mis en œuvre des procédures internes visant à éviter de financer des projets pouvant se compromettre avec des contenus de désinformation, complotistes ou incitant à la haine.

La commission considère qu'il convient d'encourager les bonnes pratiques mises en œuvre par les plateformes de financement participatif (R10) :

- ◆ Envisager une obligation pour les plateformes de financement participatif à indiquer explicitement à leurs utilisateurs toutes les mesures mises en œuvre pour éviter de participer indirectement au financement de projets se compromettant avec l'incitation à la haine ou la propagation de la désinformation.
- ◆ Inciter les plateformes de financement participatif à faire appel à des sociétés fournissant des évaluations de crédibilité de sites internet ou à obtenir un label reconnu intégrant la préoccupation de ne pas financer des sites toxiques. Cette incitation pourrait prendre la forme d'un droit à un abattement fiscal sur le bénéfice imposable de ces entreprises.

150 Labrunie, F., « Fiscalité obscure, haine et conspirations : les ombres de Tipeee, le meilleur ami des YouTubeurs », *Numerama.com*, 25 juin 2018. <https://www.numerama.com/politique/380227-fiscalite-obscure-haine-et-conspirations-les-ombres-de-tipeee-le-meilleur-ami-des-youtubeurs.html>

## Le financement public de médias de désinformation en ligne

Il arrive que certains titres de presse régulièrement cités comme diffusant des infox ou incitant à la haine soient reconnus d'information politique et générale (IPG). Cet agrément IPG leur est accordé par la Commission paritaire des publications et agences de presse (CPPAP) et leur ouvre droit au bénéfice du régime économique de la presse.

Ce régime consiste en des tarifs postaux et fiscaux préférentiels (taux super réduit de TVA à 2,1% notamment) et en l'accès à des aides financières directes pour les titres reconnus d'IPG. Une immatriculation à la CPPAP permet donc à un titre de presse de bénéficier d'un financement indirect par le contribuable.

La CPPAP est une instance indépendante composée à parité de représentants de l'administration (ministères de la culture et des finances en particulier) et de membres représentant la profession. Son secrétariat est assuré par le ministère de la culture et de la communication.

Le code des postes et des communications électroniques et le code général des impôts posent des conditions d'admission au régime économique de la presse relatives au respect de la dignité humaine. De fait, ne peuvent être admises en principe « les publications négationnistes, incitant à la haine raciale, à la xénophobie et celles portant atteinte à la dignité de la personne humaine<sup>151</sup> ».

Un rapport remis à la ministre de la culture par Laurence Franceschini, présidente de la CPPAP, propose diverses modifications réglementaires, portant tant sur les publications de presse que sur les services de presse en ligne, avec des contraintes renforcées pour les titres d'information politique et générale les plus aidés. La réforme des textes réglementaires régissant l'accès au régime économique de la presse est en cours de discussion avec les organisations professionnelles de la presse et les syndicats de journalistes. Elle pourrait constituer un outil de contrôle renforcé de la CPPAP sur l'accès au régime économique de la presse.

151 <http://www.cppap.fr/contenus-contraires-a-lordre-public/>



## CHAPITRE 4

# → Ingérences et influences étrangères

En moins de deux décennies, l'espace numérique est devenu un champ privilégié de confrontation et de compétition stratégique entre les États, voire, pour la France comme pour d'autres pays, un nouveau domaine militaire. Dans l'arsenal des combattants numériques figurent désormais les opérations informationnelles<sup>152</sup>. Les guerres de l'information ne sont certes pas nouvelles et sont même inhérentes à la stratégie militaire, qu'il s'agisse de convaincre l'opinion de la légitimité d'un combat, de contrer l'influence de l'adversaire ou de ruser pour tromper l'ennemi afin d'obtenir un avantage tactique<sup>153</sup>. Mais leur transposition dans le monde numérique pose des problèmes inédits qui perturbent la vie démocratique<sup>154</sup>. En témoigne l'annulation de la possibilité de vote électronique pour les Français de l'étranger aux élections présidentielles de 2017, à cause des opéra-

tions d'ingérence russe dans la campagne électorale américaine de 2016<sup>155</sup>.

Les raisons de ces perturbations et des difficultés à y faire face sont multiples. D'une part, l'évolution du contexte géopolitique mondial entraîne une logique d'affrontement permanent qui caractérise la conflictualité à l'ère numérique. Cette logique conduit à l'émergence de menaces de plus en plus hybrides qui impliquent une grande diversité d'acteurs et de modes opératoires et en compliquent la compréhension, la détection et la prévention. D'autre part, l'espace numérique est par nature dual<sup>156</sup> et ultra-dynamique. Il en résulte d'importantes interactions entre les mondes civil, économique et militaire qui brouillent les notions de théâtre intérieur/extérieur et produisent des effets qui à leur tour alimentent la menace.

152 Clack, T., et Johnson, R. (Éds.). (2021). *The World Information War: Western Resilience, Campaigning, and Cognitive Effects*. London: Routledge. <https://doi.org/10.4324/9781003046905>.

153 Rid, T. (2020). *Active measures : The secret history of disinformation and political warfare*. (First edition.). New York: Farrar, Straus and Giroux.

154 Benkler, Y., Faris, R., et Roberts, H. (2018). *Network Propaganda : Manipulation, Disinformation, and Radicalization in American Politics*. New York: Oxford University Press. <https://doi.org/10.1093/oso/9780190923624.001.0001> ; National Intelligence Council. (2021). *Foreign Threats to the US 2020 Federal Elections* (Intelligence Community Assessment 2020-00078D). <<https://www.dni.gov/files/ODNI/documents/assessments/ICA-declass-16MAR21.pdf>>

155 Rees, M. (11 mars 2017). «Retour sur la suppression du vote électronique pour les élections législatives 2017». [Billet de blog]. <https://www.nextinpact.com/article/25851/103636-avant-suppression-vote-electronique-dysfonctionnements-pointes-lors-tests>

156 À usage civil et militaire

## L'émergence de menaces de plus en plus hybrides

À partir de la fin des années 2000, les grandes puissances ont fait du cyberspace une priorité stratégique et investi massivement dans les capacités de lutte informatique offensive et défensive pour affirmer leur puissance et se protéger d'une menace initialement perçue comme essentiellement technique et militaire<sup>157</sup>. Or la vague d'attentats djihadistes du milieu des années 2010 a suscité une double prise de conscience. D'abord, la menace cyber pouvait être de nature informationnelle. L'utilisation experte des réseaux sociaux par l'État islamique pour diffuser sa propagande, pousser à la radicalisation, lever des fonds et organiser les départs vers la Syrie a créé une véritable surprise stratégique, même si des précédents étaient observables en Irak dès 2004<sup>158</sup>. Mais surtout, les États européens ont réalisé qu'ils avaient peu de pouvoir pour contraindre les plateformes, initialement dans le déni de leur propre responsabilité, à empêcher la diffusion de ces contenus.

Malgré cette expérience, les opérations d'ingérence menées par la Russie lors de l'élection présidentielle de 2016 ont complètement pris au dépourvu l'administration américaine comme les plateformes, et ajouté une couche de complexité au problème<sup>159</sup>. En combinant le recours à des cyberattaques (registres d'électeurs, piratage de la messagerie du parti démocrate), la publication des emails sur Wikileaks, l'amplification artificielle (botnets, fermes à trolls) de messages cliquants (contrôle des armes à feu, violences policières, racisme) sur les réseaux sociaux ou encore de publicités ciblées, à des moyens d'influence plus traditionnels (médias d'État, réseaux humains), elles ont signé l'émergence d'une menace plus hybride, protéiforme, difficile à appréhender et plus encore à contrer<sup>160</sup>.

Ces pratiques ont aussi touché la France, avec la publication des MacronLeaks à la veille du deuxième tour de l'élection présidentielle en 2017<sup>161</sup>. Elles se sont également exportées sur ses territoires d'intérêt stratégique, notamment en Afrique où la France a été prise pour cible par des campagnes de dénigrement<sup>162</sup>. La commission recommande donc de protéger l'inté-

157 Dunn Cavelty, M. (2012). *The militarisation of cyberspace : Why less may be better*. 2012 4th International Conference on Cyber Conflict (CYCON 2012), 113. Talinn.

158 Martin, D. A., Shapiro, J. N., et Nedashkovskaya, M. (2019). Recent Trends in Online Foreign Influence Efforts. *Journal of Information Warfare*, 18(3), 1548.

159 François, C., et Lin, H. (2016). « Cartographeur un angle mort : La surprise stratégique des opérations informationnelles russes sur les réseaux sociaux en 2016 ». *Herodote*, 177178(2-3), 3357.

160 Dawson, A., et Innes, M. (2019). « How Russia's Internet Research Agency Built its Disinformation Campaign ». *The Political Quarterly*, 90(2), 245256. <https://doi.org/10.1111/1467-923X.12690>

161 Jeangène Vilmer, J.-B., Escorcía, A., Guillaume, M., et Herrera, J. (2019). *Les manipulations de l'information : Un défi pour nos démocraties*. Paris: CAPS du ministère de l'Europe et des Affaires étrangères et IRSEM du ministère des Armées. [https://www.diplomatie.gouv.fr/IMG/pdf/les\\_manipulations\\_de\\_l\\_information\\_2\\_\\_cle04b2b6.pdf](https://www.diplomatie.gouv.fr/IMG/pdf/les_manipulations_de_l_information_2__cle04b2b6.pdf)

162 Centrafrique : Le Drian relève la présence de mercenaires russes. (23 janvier 2019). *Le Figaro*. <https://www.lefigaro.fr/flash-actu/2019/01/23/97001-20190123FILWWW00322-centrafrique-le-drian-releve-la-presence-de-mercenaires-russes.php>

grité du processus électoral par une coopération accrue avec les plateformes et les chercheurs (R10).

Enfin, ces manœuvres informationnelles se sont internationalisées au cours des deux dernières années avec le durcissement du contexte stratégique et la montée en puissance des tensions géopolitiques liées à la crise sanitaire. Des publications et des déclarations publiques ont pointé les opérations d'influence de la Russie<sup>163</sup>, la Turquie<sup>164</sup>, l'Iran ou encore la Chine<sup>165</sup>.

Dans le but d'éviter l'escalade des conflits et de répondre à des situations d'urgence, la commission recommande la création d'un mécanisme et d'exercices de gestion de crise sur les menaces informationnelles au sein de l'Union européenne (R14).

Les menaces hybrides permettent de créer de l'ambiguïté dans un contexte géopolitique où les limites entre guerre et paix sont de plus en plus floues, donnant lieu à une zone grise plus justement définie par les notions de compétition, contestation et affrontement<sup>166</sup>. Elles sont aussi caractérisées par une diversité croissante d'acteurs –étatiques et non-étatiques–, de modes opératoires et d'effets produits qui entraînent une grande confusion sémantique et compliquent la compréhension des phénomènes et la réponse à y apporter.

163 Innes, M., Grinnell, D., Innes, H., Harmston, D., et Roberts, C. (2020). Normalisation et domestication de la désinformation numérique : Les opérations informationnelles d'interférence et d'influence de l'extrême droite et de l'État russe en Europe. *Herodote*, 177178 (2-3), 101123. ; Graphika. (2020). *Step into My Parler : Suspected Russian Operation Targeted Far-Right American Users on Platforms Including Gab and Parler, Resembled Recent IRA-Linked Operation that Targeted Progressives*. <https://graphika.com/reports/step-into-my-parler/>

164 Twitter safety. (12 juin 2020). Disclosing networks of state-linked information operations we've removed. [Billet de blog]. [https://blog.twitter.com/en\\_us/topics/company/2020/information-operations-june-2020](https://blog.twitter.com/en_us/topics/company/2020/information-operations-june-2020) ; Macron : « Il y aura des tentatives d'ingérence » de la Turquie dans l'élection présidentielle. (2021, mars 23). *Le Figaro*. <https://www.lefigaro.fr/flash-actu/macron-il-y-aura-des-tentatives-d-ingerence-de-la-turquie-dans-l-election-presidentielle-20210323>

165 Nimmo, B., Hubert, I., et Cheng, Y. (2021). *Spamouflage Breakout : Chinese Spam Network Finally Starts to Gain Some Traction*. Graphika. <https://graphika.com/reports/spamouflage-breakout/>; Charon, P., et Jeangène Vilmer, J.-B. (2021). *Les opérations d'influence chinoises—Un moment machiavélien*. Paris: Institut de recherche stratégique de l'École militaire (IRSEM). <https://www.irsem.fr/rapport.html>; Timberg, C., et Harris, S. (2020a, août 12). «Chinese network of fake accounts targets Trump with English-language videos.» *The Washington Post*. <https://www.washingtonpost.com/technology/2020/08/12/china-video-network-trump/>

166 NATO Strategic Communications Centre of Excellence. (2016). *Social Media as a Tool of Hybrid Warfare*. NATO. [https://stratcomcoe.org/pdfs/?file=/publications/download/public\\_report\\_social\\_media\\_hybrid\\_warfare\\_22-07-2016-1.pdf?zoom=page-fit](https://stratcomcoe.org/pdfs/?file=/publications/download/public_report_social_media_hybrid_warfare_22-07-2016-1.pdf?zoom=page-fit)

## Une grande diversité d'acteurs, de stratégies et de modes opératoires

La diversité des appellations (désinformation, manipulations de l'information, opérations ou manœuvres informationnelles, ingérences numériques) témoigne de cette confusion sémantique mais aussi de la difficulté à étudier, comprendre et caractériser ces phénomènes. La recherche s'est fortement développée au cours des cinq dernières années mais se heurte à la difficulté d'accès aux données des plateformes et aux biais liés aux aléas de la collecte de données en sources ouvertes, limitée par les restrictions techniques et juridiques visant à protéger les droits des utilisateurs<sup>167</sup>. Certaines études s'appuient donc sur des jeux de données, de plus ou moins bonne qualité<sup>168</sup> que des plateformes acceptent de partager<sup>169</sup>. Mais elles n'offrent, au mieux, qu'une vue partielle du problème, alors même que les opérations hybrides se déploient sur de multiples canaux. D'autres études se fondent sur des campagnes spécifiques (élections, pandémies) dans certains pays à travers l'analyse de différents vecteurs, mais elles se heurtent aux difficultés d'identifier les auteurs des opérations, leurs intentions et leurs éventuels liens avec des États<sup>170</sup>. Car la relative impunité qui règne en la matière a encouragé une multitude d'acteurs privés (entrepreneurs d'influence, mercenaires, criminels) à lancer leurs propres campagnes, rendant plus complexe encore cet écosystème<sup>171</sup>.

Des acteurs étatiques, académiques et privés étudient ces questions à partir de points de vue extrêmement différents, sans qu'il y ait d'unité d'analyse ou de grille de lecture commune, ni de mécanisme de partage d'information institutionnalisé. Cela encourage la focalisation sur les aspects tactiques de ces opérations au détriment d'une compréhension globale de leurs objectifs stratégiques, de leur ampleur et de leurs effets réels sur nos sociétés.

D'où la nécessité d'exiger des plateformes l'accès aux données par les chercheurs (R20) et d'organiser un partage de données structurées de manière cohérente entre des acteurs qui étudient ces phénomènes (R11).

---

167 Bateman, J., Hickok, E., Courchesne, L., Thange, I., et Shapiro, J. N. (2021, juin 28). *Measuring the Effects of Influence Operations : Key Findings and Gaps From Empirical Research*. Carnegie Endowment for International Peace. <https://carnegieendowment.org/2021/06/28/measuring-effects-of-influence-operations-key-findings-and-gaps-from-empirical-research-pub-84824>

168 Timberg, C. (2021, septembre 10). «Facebook made big mistake in data it provided to researchers, undermining academic work». *The Washington Post*. <https://www.washingtonpost.com/technology/2021/09/10/facebook-error-data-social-scientists/>

169 Nimmo, B., François, C., Eib, C. S., et Ronzard, L. (2020). *IRAAgain : Unlucky Thirteen. Facebook Takes Down Small, Recently Created Network Linked to Internet Research Agency* (p.38). Graphika. <https://graphika.com/reports/ira-again-unlucky-thirteen/>

170 Crime and Security Research Institute. (2021). *How a Kremlin-Linked Influence Operation is Systematically Manipulating Western Media to Construct and Communicate Disinformation - Part 1 Detection Report*. Cardiff University. [https://www.cardiff.ac.uk/\\_\\_data/assets/pdf\\_file/0008/2560274/OSCAR-report-September-2021.pdf](https://www.cardiff.ac.uk/__data/assets/pdf_file/0008/2560274/OSCAR-report-September-2021.pdf)

171 Laruelle, M., et Limonier, K. (2021). Beyond «hybrid warfare» : A digital exploration of Russia's entrepreneurs of influence. *Post-Soviet Affairs*, 37(4), 318335. <https://doi.org/10.1080/1060586X.2021.1936409>

Face à la confusion sémantique, Camille François propose un cadre d'analyse en forme d'ABC de la désinformation<sup>172</sup>: A pour Acteurs (manipulative Actors) qui s'engagent délibérément dans des campagnes de manipulation en ligne en cherchant à masquer leur identité et leurs intentions; B pour comportements trompeurs (deceptive Behavior), qui comprennent une grande diversité de techniques et de vecteurs (plateformes, sites webs, blogs) pour amplifier la portée, la viralité et l'impact des campagnes en ligne ; C pour contenus nocifs ou malveillants (harmful Content), critère le plus subjectif et complexe à définir. On considère qu'il s'agit d'une ingérence étrangère lorsque l'acteur manipulateur est une puissance étrangère, ou agit pour son compte. Or cela n'est pas toujours aisé à établir avec certitude : un acteur étranger peut instrumentaliser un acteur national pour relayer ses contenus malveillants ; un entrepreneur d'influence peut mener une campagne pour s'attirer les faveurs d'une puissance étrangère sans pour autant en être l'agent ; un acteur manipulateur peut recourir à des comportements transparents (non trompeurs) pour diffuser du contenu politiquement désagréable (mais licite voire légitime) et bénéficier d'une viralité organique (non artificielle), parce que ses contenus plaisent et sont repris par d'autres<sup>173</sup>.

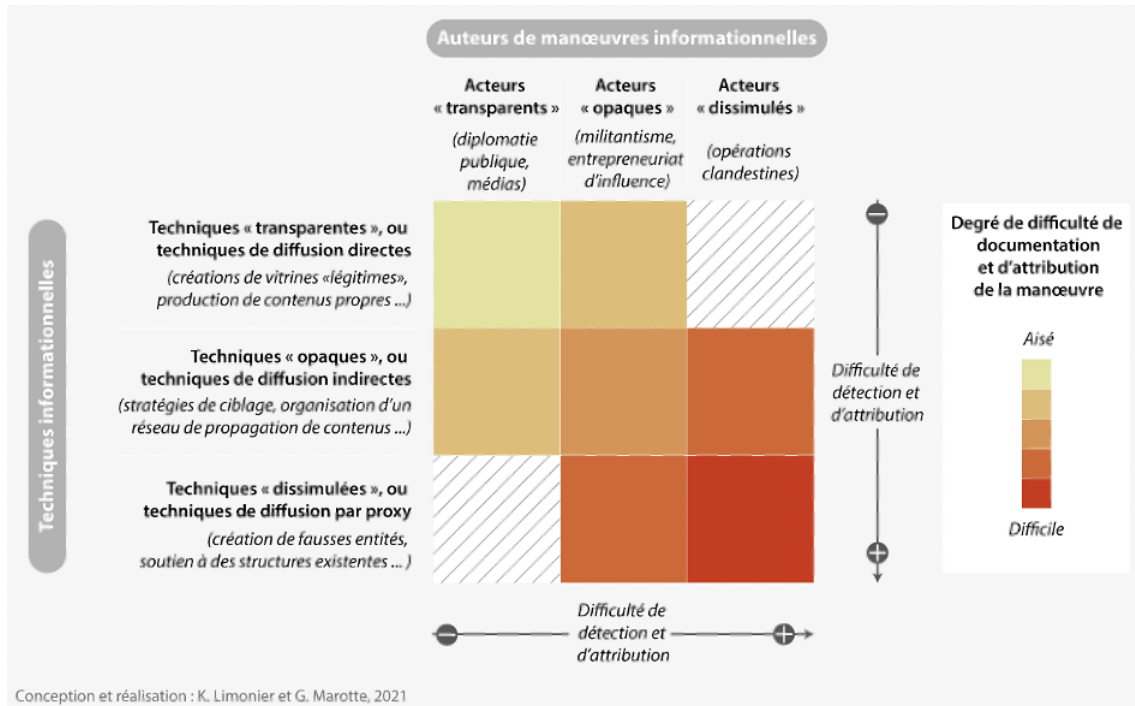
Dès lors, il faut comprendre le problème comme un spectre dans lequel une pluralité d'acteurs, plus ou moins manipulateurs, utilise une palette de techniques, plus ou moins trompeuses, pour diffuser une diversité de contenus, plus ou moins nuisibles. Kevin Limonier propose une matrice pour représenter la diversité des situations en fonction d'une typologie des acteurs et des techniques informationnelles russes, qu'il regroupe en trois catégories : transparents, opaques et dissimulés<sup>174</sup>. En fonction de la combinaison de ces techniques et des acteurs qui les utilisent, les opérations sont plus ou moins faciles à détecter et à attribuer à leurs auteurs.

172 François, C. (2019). *Actors, Behaviors, Content : A Disinformation ABC. Highlighting Three Vectors of Viral Deception to Guide Industry et Regulatory Responses*. Santa Monica, California : Working Paper of the Transatlantic Working Group on Content Moderation Online and Freedom of Expression. <[https://www.ivir.nl/publicaties/download/ABC\\_Framework\\_2019\\_Sept\\_2019.pdf](https://www.ivir.nl/publicaties/download/ABC_Framework_2019_Sept_2019.pdf)>

173 Douzet, F., Limonier, K., Mihoubi, S., et René, É. (2020). Cartographier la propagation des contenus russes et chinois sur le Web africain francophone. *Herodote*, 177178(2-3), 7799.

174 Limonier, K., et Laruelle, M. (s. d.). « Typologie des manoeuvres informationnelles russes à l'égard des pays francophones. » *Annuaire français des relations internationales*. à paraître.

Matrice des opérations informationnelles russes



La compréhension de ce spectre est indispensable pour adapter la réponse et éviter les pièges qui conduisent à faire le jeu des manipulateurs.

## Des réponses complexes

Les opérations informationnelles d'origine étrangère qui visent la France et l'Europe sont par nature transfrontalières et utilisent comme vecteur les plateformes les plus fréquentées, qui sont pour la plupart basées aux États-Unis. En conséquence, la réponse nécessite une coopération internationale avec les acteurs régaliens pour ce qui relève de la mise en œuvre du droit, et avec les acteurs privés des plateformes qui détiennent à la fois les données et de puissants moyens d'action. Selon qu'on cherche à cibler les acteurs malveillants, les comportements trompeurs ou les contenus, la réponse et les acteurs impliqués ne seront pas les mêmes.

### LA COOPÉRATION ENTRE LES ÉTATS ET LA MOBILISATION DU DROIT INTERNATIONAL

Pour lutter contre les acteurs étatiques malveillants, les États peuvent mobiliser les outils existants du droit international. En 2016 par exemple, l'administration Obama a publiquement accusé la Fédération de Russie d'ingérence dans l'élection présidentielle et renvoyé ses ambassadeurs en signe de protestation. La mise en œuvre du principe de non-intervention en la matière n'est cependant pas aisée (difficultés d'attribution, de qualification, de choix de la réponse) et à double-tranchant, car c'est l'instrument privilégié des régimes autoritaires pour justifier la censure institutionnalisée en ligne<sup>175</sup>. In fine, c'est un levier peu efficace pour contrer le phénomène et dissuader les acteurs.

Sur les contenus, la lutte contre le terrorisme a paradoxalement ouvert la voie à un consensus international, alors même que le contrôle des contenus terroriste divise les États de longue date. Mais ce consensus tient surtout à la désignation commune d'un ennemi, l'État islamique, dont l'action a été qualifiée de menace à la paix et la sécurité internationale. Outre l'absence d'un consensus sur la définition du problème, les protections offertes par le droit international aux droits humains et la liberté d'expression rendent peu probable une régulation internationale sur les manipulations de l'information<sup>176</sup>. Les initiatives visent plutôt à réguler les comportements, en coopération avec le secteur privé. C'est pourquoi la commission recommande plutôt un régime de co-régulation prévoyant une coopération exigeante avec les plateformes dans le cadre de la législation sur les services numériques(R23).

175 L'exemple de la loi américaine sur les «agents de l'étranger» et son instrumentalisation par la Russie va en ce sens. Voir Rebo, S. (2021), FARA in Focus: What Can Russia's Foreign Agent Law Tell Us About America's? *Journal of National Security Law et Policy*, 12(2).

176 Douzet, F., et Géry, A. (2021). La régulation pour contrer les manipulations de l'information en ligne : L'impossible consensus. Dans Marangé C., et Quessard M., (Édit.), *Les guerres de l'information à l'ère numérique* (p. 395-418). Paris, Presses universitaires de France.

## LA COOPÉRATION ENTRE ÉTATS ET PLATEFORMES

La plupart des plateformes étaient initialement réticentes à travailler avec les gouvernements par crainte de perdre la confiance de leurs utilisateurs, déjà mise à mal par les révélations d'Edward Snowden, mais aussi d'avoir à défendre leurs actions vis-à-vis des régimes autoritaires qui exercent une pression constante pour supprimer des contenus et fermer des comptes. Leur modèle économique fondé sur une conception maximaliste de la liberté d'expression ne les préparait pas à ce type de contrainte étatique.

En 2015 toutefois, la prolifération de vidéos de décapitation et la multiplication de départs de jeunes vers la Syrie a suscité une forte pression des utilisateurs et des gouvernements pour qu'elles assument leur part de responsabilité et se donnent les moyens d'endiguer la propagande djihadiste. Faute de mécanismes de coopération judiciaires internationaux efficaces, elles ont mis en place des processus de coopération avec les gouvernements et la société civile, basés sur leur charte des utilisateurs, afin de faciliter le signalement et le retrait des contenus terroristes<sup>177</sup>. Ce n'est qu'à partir de 2017 que les plateformes se sont réunies pour lutter ensemble contre l'extrémisme violent en ligne avec la création du Global Internet Forum to Counter Terrorism (GIFTC)<sup>178</sup>. Avec des résultats encore limités, comme en a témoigné la circulation de la vidéo des attentats de Nouvelle-Zélande en 2019, qui a conduit au lancement de « l'Appel de Christchurch pour supprimer les contenus terroristes et extrémistes violents en ligne » par la Première ministre Ardern et le président Macron (2019)<sup>179</sup>.

L'émergence des manipulations de l'information visant directement la démocratie américaine en 2016 a mis les plateformes face à leur propre pouvoir et responsabilités. Leurs efforts se sont principalement portés sur le ciblage des comportements trompeurs<sup>180</sup>. Les acteurs manipulateurs s'appuient sur la technologie et le modèle économique qu'elles ont construits en les détournant de leur usage initial : facilité d'élaboration de multiples comptes, publicités ciblées, algorithmes de recommandation, facilité et rapidité de partage. Les définitions des grandes plateformes convergent autour des notions de « comportements inauthentiques coordonnés » (Facebook), « fausses campagnes d'influence » (Twitter) et « comportements trompeurs » (Google); dès 2018, le gouvernement américain entame une collaboration étroite et publique avec les grandes plateformes, qui prend notamment la forme de signalements du FBI aux plateformes aussitôt que de nouvelles opérations sont détectées.

177 Biddle, S. (12 octobre 2021). «Revealed : Facebook's Secret Blacklist of «Dangerous Individuals and Organizations.»» *The Intercept*. <https://theintercept.com/2021/10/12/facebook-secret-blacklist-dangerous/>

178 Global Internet Forum to Counter Terrorism. <https://gifct.org/>

179 *Appel de Christchurch pour agir contre le terrorisme et l'extrémisme violent en ligne*, (2019). <https://www.appeldechirstchurch.com/index.html>

180 Stiglitz, A. (2020). *Mis- and Disinformation online : A Taxonomy of Solutions* [Thèse de doctorat]. Universidad de Navarra, Pamplona. <https://www.sipa.columbia.edu/sites/default/files/documents/Anyas%20thesis%20July%202020%20pdf.pdf> ; Faesen, L., Klimburg, A., van Hoeve, S., et Sweijs, T. (2021). *Redlines et Baselines : Towards a European Multistakeholder Approach to Counter Disinformation*. La Haye: The Hague Centre for Strategic Studies. <https://hcsc.nl/report/red-lines-baselines/>



Malgré de réels efforts pour combattre la menace et offrir plus de transparence sur leurs pratiques, les initiatives restent fragmentées, aussi bien entre les plateformes que parfois entre les différentes plateformes d'un même groupe<sup>181</sup>.

C'est pourquoi la commission propose la création d'un groupe de travail au sein de l'OCDE pour établir des normes minimales communes à toutes les plateformes et harmoniser les législations nationales sur leurs obligations (R15).

Les choix que les plateformes effectuent en la matière restent relativement opaques et échappent à tout encadrement juridique européen. Et surtout, leurs efforts tendent à s'essouffler au fur et à mesure que la pression de Washington retombe<sup>182</sup>. Car l'élection présidentielle de 2020 a montré que la menace informationnelle sur le processus électoral venait cette fois de l'intérieur : de l'extrême droite, des conspirationnistes (QAnon<sup>183</sup>), voire de la Maison Blanche<sup>184</sup>.

Les stratégies de réponse aux manipulations de l'information étrangères reposent souvent sur l'exposition médiatique du démantèlement de réseaux de comptes et de campagnes identifiées par les États ou les plateformes. Cette logique de dénonciation publique (naming and shaming) envoie un message diplomatique tout en sensibilisant l'opinion publique aux risques et techniques de désinformation. Elle porte toutefois le risque d'amplifier des opérations ou des acteurs qui n'avaient jusque-là qu'une visibilité limitée, d'ailleurs difficile à mesurer. Ces mises en accusation publiques peuvent être habilement instrumentalisées à des fins politiciennes, par les manipulateurs comme par leurs accusateurs (affaire Benalla<sup>185</sup>; crise des gilets jaunes).

Enfin, la montée en puissance des manœuvres informationnelles a conduit à une forme de militarisation de l'espace informationnel, dont certains aspects risquent en retour d'alimenter la menace.

181 Timberg, C. (7 octobre 2020). «Parler and Gab, two conservative social media sites, keep alleged Russian disinformation up, despite report.» *The Washington Post*. <<https://www.washingtonpost.com/technology/2020/10/07/russian-trolls-graphika-parler-gab/>> ; Isaac, M., et Roose, K. (2018, octobre 19). Disinformation Spreads on WhatsApp Ahead of Brazilian Election. *The New York Times*. <<https://www.nytimes.com/2018/10/19/technology/whatsapp-brazil-presidential-election.html>>

182 François, C., et Douek, E. (2021). The Accidental Origins, Underappreciated Limits, and Enduring Promises of Platform Transparency Reporting about Information Operations. *Journal of Online Trust and Safety*, 1(1). <https://doi.org/10.54501/jots.v1i1.17>

183 Rothschild, M. (2021). *The Storm is Upon Us : How QAnon Became a Movement, Cult, and Conspiracy Theory of Everything..*

184 Center for an Informed Public, Digital Forensic Research Lab, Graphika, et Stanford Internet Observatory (2021). *The Long Fuse: Misinformation and the 2020 Election*. Stanford Digital Repository: Election Integrity Partnership. v1.3.0. <https://stacks.stanford.edu/file/druid:tr171zs0069/EIP-Final-Report.pdf>

185 Gérard, C., et Marotte, G. (2020). «#AffaireBenalla : Déconstruction d'une polémique sur le rôle de la communauté Twitter « russophile » dans le débat politique français.» *Herodote*, 177178(2-3), 125147.

## La militarisation de l'espace informationnel

En France, le commandement de la cyberdéfense s'est trouvé en première ligne pour contrer la menace informationnelle terroriste, perçue comme une véritable rupture stratégique<sup>186</sup>. Les militaires ont mené des opérations pour endiguer la diffusion de la propagande djihadiste<sup>187</sup>, à défaut de pouvoir l'éradiquer.

Depuis, on observe une prolifération des manœuvres informationnelles qui pousse les gouvernements à considérer le champ informationnel comme un enjeu de sécurité nationale et à développer des capacités pour se défendre, mais aussi contre-attaquer<sup>188</sup>. Le 20 octobre 2021, Florence Parly, ministre des Armées, annonçait clairement que la France se dotait d'une doctrine de lutte informatique d'influence pour « détecter, caractériser et contrer les attaques », mais aussi « faire de la déception<sup>189</sup>, de façon autonome ou en combinaison avec d'autres opérations<sup>190</sup> ».

Cette dynamique s'inscrit dans la continuité de la course aux armes numériques et pose les mêmes problèmes. D'une part, il est impossible de contenir les effets recherchés à la seule sphère militaire, parce que l'espace numérique informationnel est partagé entre le monde civil, économique et militaire. La propagation des contenus est difficilement maîtrisable et les actions sont potentiellement observables par une multitude d'acteurs. Elles peuvent contribuer à affaiblir la confiance dans l'information numérique et envers les institutions<sup>191</sup>.

D'autre part, les opérations informationnelles permettent aux différents acteurs d'apprendre les uns des autres. Des États comme des criminels peuvent exploiter les mêmes vulnérabilités, copier les modes opératoires et les réutiliser. Lors de l'élection présidentielle américaine de 2020, de jeunes militants pro-Trump ont été accusés de réutiliser les méthodes d'usine à trolls pour soutenir leur candidat<sup>192</sup>. Leurs comptes ont été fermés.

186 Douzet, F. (2016). « Le cyberspace, troisième front de la lutte contre Daech. » *Herodote*, 160-161(1-2), 223238.

187 Commission de la défense nationale et des forces armées. ( 4 mars 2020). *Compte-rendu d'audition du général de division aérienne Didier Tisseyre, général commandant la cyber défense sur le thème « le cyber, nouvel espace de conflictualité »*. Assemblée Nationale, [https://www.assemblee-nationale.fr/dyn/15/comptes-rendus/cion\\_def/115cion\\_def1920040\\_compte-rendu](https://www.assemblee-nationale.fr/dyn/15/comptes-rendus/cion_def/115cion_def1920040_compte-rendu)

188 Marangé, C., et Quessard-Salvaing, M. (2021). *Les guerres de l'information à l'ère numérique*. Paris, France: Presses universitaires de France.

189 Mesures visant à induire l'ennemi en erreur, grâce à des truquages, des déformations de la réalité, ou des falsifications, en vue de l'inciter à réagir d'une manière préjudiciable à ses propres intérêts (définition OTAN).

190 Ministère des armées. (2021, octobre 20). Florence Parly présente la doctrine militaire de lutte informatique d'influence. Consulté 29 novembre 2021, à l'adresse <https://www.defense.gouv.fr/actualites/articles/florence-parly-presente-la-doctrine-militaire-de-lutte-informatique-d-influence>

191 Graphika et Stanford Internet Observatory. (2020). *More-Troll Kommat : French and Russian Influence Operations Go Head to Head Targeting Audiences in Africa*. <<https://graphika.com/reports/more-troll-kommat/>>

192 Stanley-Becker, I. (15 septembre 2020). « Pro-Trump youth group enlists teens in secretive campaign likened to a 'troll farm,' prompting rebuke by Facebook and Twitter. » *The Washington Post*. [https://www.washingtonpost.com/politics/turning-point-teens-disinformation-trump/2020/09/15/c84091ae-f20a-11ea-b796-2dd09962649c\\_story.html](https://www.washingtonpost.com/politics/turning-point-teens-disinformation-trump/2020/09/15/c84091ae-f20a-11ea-b796-2dd09962649c_story.html); Stanford Internet Observatory, Cyber Policy Center. (2020). *Reply-Guys Go Hunting : An Investigation into a U.S. Astroturfing Operation on Facebook, Twitter, and Instagram*. <https://cyber.fsi.stanford.edu/io/news/oct-2020-fb-rally-forge>

Pour ces raisons, la commission recommande l'avis du Comité d'éthique de la défense sur la doctrine de lutte informatique d'influence (R13).

Un certain nombre d'acteurs l'ont d'ailleurs bien compris et instrumentalisent cette dualité de l'espace numérique en utilisant des acteurs civils comme vecteur d'influence numérique<sup>193</sup>, brouillant un peu plus les pistes entre ingérence étrangère et menace interne, dans une logique d'hybridité qui complique plus encore la réponse des démocraties. Cela crée un climat de tension dans lequel les États redoutent en permanence qu'une information soit l'instrument ou le fruit d'une manœuvre dans un but d'influence stratégique<sup>194</sup> et que la situation leur échappe, ce qui renforce la course aux moyens.

L'approche des risques numériques doit ainsi être globale car les menaces sont de plus en plus hybrides et transversales. D'où la nécessité de créer un mécanisme de gouvernance numérique interministérielle afin de définir des stratégies, des politiques publiques et des réponses fortement coordonnées en matière de défense, sécurité et diplomatie, qui tiennent compte des interactions multiples propres à ce domaine partagé (R12).

---

193 Vitkine, B. (14 septembre 2022). «Une officine d'influence russe s'intéresse aux violences policières en France.» *Le Monde*. [https://www.lemonde.fr/police-justice/article/2021/09/14/une-officine-d-influence-russe-s-interesse-aux-violences-policieres-en-france\\_6094599\\_1653578.html](https://www.lemonde.fr/police-justice/article/2021/09/14/une-officine-d-influence-russe-s-interesse-aux-violences-policieres-en-france_6094599_1653578.html)

194 L'instrumentalisation par la Turquie des déclarations du président Macron sur les caricatures utilisées par l'enseignant Samuel Paty, sauvagement assassiné en 2020, a donné lieu à une campagne informationnelle qui a notamment abouti à la création de Viginum, service de vigilance et de protection contre les ingérences numériques étrangères.

## CHAPITRE 5

# → Droit et numérique

La prévention et la lutte contre la diffusion de fausses informations nécessitent la mise en œuvre coordonnée de différents moyens qui, pour la plupart, relèvent plus de l'incitation politique ou de l'autorégulation des acteurs que de normes juridiques contraignantes. Néanmoins, dans un État de droit, il est indispensable que la loi prévoie quelques instruments pour contrecarrer et sanctionner certaines formes graves de cette diffusion, en particulier sur les réseaux numériques.

L'étude des différentes dispositions juridiques pouvant être utiles pour prévenir ou sanctionner les différentes formes de désinformation (au sens de la diffusion de mauvaise foi de fausses nouvelles) incite à ne pas modifier ni remplacer l'actuel article 27 de la loi de 1881 sur la liberté de la presse. En revanche, la sanction pénale pourrait être complétée par un mécanisme de mise en cause de la responsabilité civile des diffuseurs de mauvaise foi de fausses nouvelles pouvant porter préjudice à autrui, responsabilité qui pourrait notamment être proportionnée à la viralité de leur diffusion et à la popularité numérique de leur auteur.

À côté des dispositions légales assurant la sanction de la désinformation, il paraît indispensable de développer les mécanismes de modération et de régulation et d'y soumettre les plateformes numériques qui sont au cœur de la diffusion virale de contenus de désinformation. Ainsi le renforcement des maigres prérogatives accordées ces dernières années au Conseil supérieur de l'audiovisuel (future ARCOM) semble nécessaire pour s'assurer de la collaboration des plateformes numériques dans la détection et la suppression rapide des fausses informations susceptibles de troubler l'ordre public et superviser leurs actions en la matière, voire les sanctionner. À terme, c'est au niveau européen qu'il est essentiel que la future législation sur les services numériques (Digital Services Act) puisse imposer aux plateformes la mise en œuvre d'une modération effective des fausses nouvelles susceptibles de troubler l'ordre public, quitte à s'appuyer sur une expertise indépendante pour évaluer l'opportunité de retrait des contenus ou de déréférencement en prenant par ailleurs en considération la préservation de la liberté d'expression.

## Définition juridique et sanction des fausses nouvelles pénalement répréhensibles

Diffuser une nouvelle qui se révèle partiellement ou totalement fausse n'est pas en soi, dans un système libéral, un acte répréhensible. Tout au contraire, la jurisprudence de la Cour européenne des droits de l'homme (CEDH) estime que la possibilité d'exprimer publiquement des propos ou des informations inexacts fait partie intégrante de l'exercice de la liberté d'expression, protégée par l'article 10 de la Convention européenne de sauvegarde des droits de l'homme et des libertés fondamentales (ConvEDH) et que seules des exceptions légalement justifiées par d'autres intérêts majeurs pouvaient permettre d'y déroger. L'arrêt *Handyside* de la CEDH a notamment affirmé que « *la liberté d'expression vaut non seulement pour les informations ou idées accueillies avec faveur ou considérées comme inoffensives ou indifférentes mais aussi pour celles qui heurtent, choquent ou inquiètent une personne ou une partie de la population* »<sup>195</sup>.

En droit français, le Conseil constitutionnel a rappelé dans sa décision du 18 juin 2020<sup>196</sup> que s'applique en la matière l'article 11 de la Déclaration des droits de l'homme et du citoyen de 1789 qui affirme que « *la libre communication des pensées et des opinions est un des droits les plus précieux de l'homme : tout citoyen peut donc parler, écrire, imprimer librement, sauf à répondre de l'abus de cette liberté dans les cas déterminés par la loi* ». Il a en a déduit que seules certaines catégories particulièrement préjudiciables de fausses nouvelles pouvaient être soumises à des procédures administratives restrictives.

Il serait donc contraire tant à la Constitution qu'aux engagements internationaux de la France de vouloir sanctionner juridiquement toutes les formes de diffusion de fausses nouvelles, ce qui en particulier confondrait gravement la « *mésinformation* » avec de la « *désinformation* ». Mais il serait également dangereux d'un point de vue juridique de retenir pour définir les actes de « *désinformation* » des critères trop larges ou susceptibles d'une interprétation trop extensive. Tout au contraire convient-il dans ce domaine très sensible de ne soumettre à un encadrement juridique qu'une catégorie restreinte et particulièrement nuisible de fausses nouvelles.

Cette catégorie de fausses nouvelles que l'on pourra qualifier ci-après de « *répréhensible* » pour les distinguer de celles dont la diffusion est par principe libre (sauf à répondre de leur abus dans les conditions civiles de droit commun) a déjà été définie en droit français en 1881 lors de l'adoption de la loi relative à la liberté de la presse<sup>197</sup>. C'est son article 27 que nous proposons

195 CEDH, n° 5493/72, *Handyside c. Royaume-Uni*, 7 décembre 1976, paragraphe 49.

196 Conseil constitutionnel, Décision DC n° 2020-801 du 18 juin 2020 (Loi visant à lutter contre les contenus haineux sur internet).

197 Loi du 29 juillet 1881 sur la liberté de la presse.

de conserver comme l'un des principaux piliers du dispositif juridique d'accompagnement de la politique nationale et européenne de prévention et de lutte contre la désinformation pouvant affecter gravement le débat démocratique et ce que certains auteurs n'hésitent pas à dénommer « l'ordre public numérique ».

Dans sa version actuellement applicable, cet article 27 de la loi du 29 juillet 1881 sanctionne

*« La publication, la diffusion ou la reproduction, par quelque moyen que ce soit, de nouvelles fausses, de pièces fabriquées, falsifiées ou mensongèrement attribuées à des tiers lorsque, faite de mauvaise foi, elle aura troublé la paix publique, ou aura été susceptible de la troubler. »*

Le caractère répréhensible d'une fausse nouvelle implique donc une triple condition :

- ♦ qu'elle ait fait l'objet d'une communication au public (de quelque manière que ce soit, y compris via un service en ligne),
- ♦ qu'elle trouble ou soit de nature à troubler l'ordre public,
- ♦ et que sa diffusion ait été réalisée de mauvaise foi.

L'interprétation de ces conditions ayant déjà été effectuée par la jurisprudence, nous disposons ici d'une base juridique restreinte mais solide, permettant de déterminer la frontière entre les excès de la liberté d'expression qui ne sont pas – en eux-mêmes – répréhensibles de ceux qui, au contraire, remplissent les conditions pour pouvoir être sanctionnés pénalement voire faire l'objet, le cas échéant, de mesures administratives contraignantes.

Ainsi ne peut être sanctionnée que la diffusion d'une « nouvelle », au sens d'une « *annonce d'un événement arrivé récemment, faite à quelqu'un qui n'en a pas encore connaissance* » (et non d'un commentaire d'une information déjà rendue publique<sup>198</sup>). Cette nouvelle doit être « *fausse, c'est-à-dire mensongère, erronée ou inexacte dans la matérialité du fait et dans les circonstances*<sup>199</sup> ». Quant au trouble à l'ordre public, il peut recouvrir différents cas de désordre collectif, qu'il s'agisse des risques de trouble sur la voie publique, d'influence sur les relations internationales, mais aussi sur les risques de tensions entre les citoyens. De plus, il n'est pas nécessaire de prouver l'existence d'un trouble déjà effectif mais simplement de démontrer que la diffusion concernée serait susceptible de créer un tel trouble.

À ceci s'ajoute que l'État est particulièrement protégé contre la fausse nouvelle résultant d'une ingérence informationnelle étrangère. En effet l'article 411-10 du code pénal punit de peines sensiblement plus lourdes « *le fait de fournir, en vue de servir les intérêts d'une puissance étrangère, d'une entreprise ou organisation étrangère ou sous contrôle étranger, aux autorités*

198 Cass. Crim., 13 avril 1999, Bulletin criminel 1999 n° 78, p. 214.

199 CA Paris, 11ème ch., 18 mai 1988.

*civiles ou militaires de la France des informations fausses de nature à les induire en erreur et à porter atteinte aux intérêts fondamentaux de la nation*».

D'autres catégories de fausses nouvelles sont visées par des dispositions spéciales :

- les fausses informations faisant croire à un incident imaginaire (faux sinistre, faux accident, fausse dégradation ou détérioration dangereuse) (article 322-14 du code pénal) ;
- « *l'indication fausse ou trompeuse* » pouvant affecter les cours des marchés financiers, ou de leurs indices (articles L.465-3-1 à L.465-3-2 Code monétaire et financier),
- les « *allégations ou imputations inexactes ou trompeuses d'un fait de nature à altérer la sincérité* » d'un scrutin électoral (article L. 163-2 Code électoral) ;

Tout au plus peut-on remarquer qu'aucune disposition spécifique ne vient sanctionner la diffusion d'une fausse nouvelle affectant uniquement une ou plusieurs personnes privées. C'est généralement sur le seul terrain de la « diffamation » (art. 29 de la loi du 29 janvier 1881) que la justice peut sanctionner le fait de nuire à une personne privée par la publication d'un fait qui soit faux ou qui fasse l'objet d'une présentation fallacieuse dans le but de lui porter préjudice. Par ailleurs, la jurisprudence a admis la possibilité d'agir (y compris en référé) pour faire cesser la diffusion d'informations préjudiciables à la vie privée (en application de l'article 9 du code civil). Mais une disposition civile plus spécifique pourrait s'avérer utile, comme on l'évoquera plus loin, sans affecter la portée du dispositif pénal de la loi de 1881.

On peut donc constater – comme l'a relevé le Conseil d'État dans son avis du 19 avril 2018 – que « *la lutte contre les fausses informations est une préoccupation ancienne et récurrente du législateur, à laquelle répondent déjà, quoique de façon éparse, de nombreuses dispositions* ».

Par ailleurs, et d'un point de vue plus général, il faut se féliciter que la loi du 29 juillet 1881, bien que prévue pour sanctionner les infractions de presse écrite, soit devenue également le cadre légal de la communication publique d'informations sur tous les supports numériques.

Dès lors et compte tenu également du risque d'encombrement législatif – voire de contrariété avec la liberté d'expression – qui pourrait résulter de l'adoption d'une nouvelle disposition réprimant la désinformation, il est recommandé de conserver l'article 27 de la loi du 29 juillet 1881 comme pierre angulaire du dispositif pénal de sanction de la diffusion de mauvaise foi de fausses nouvelles répréhensibles (comme le rapport du Sénat de 2016 l'avait plutôt conseillé<sup>200</sup>).

Tout au plus, l'ouverture du droit d'action en la matière à des associations – en leur permettant de se constituer partie civile – pourrait renforcer le dispositif car cela permettrait d'accroître l'utilisation de cette incrimination pénale et de son effet dissuasif tout en favorisant le développement d'une jurisprudence détaillée et adaptée en particulier aux cas de désinformation sur les réseaux numériques.

## **Recommandations**

### **Conserver dans sa rédaction actuelle l'article 27 de la loi du 29 juillet 1881 (R16) comme :**

**– fondement des actions pénales visant la diffusion publique de fausses nouvelles sur les réseaux de communication numériques et les plateformes,**

**– ainsi que comme définition de référence pour la détermination de ce qui constitue une fausse nouvelle répréhensible dont le retrait ne porterait pas une atteinte injustifiée à l'exercice de la liberté d'expression.**

### **Compléter l'article 48-1 de la loi du 29 juillet 1881 afin de permettre à des associations ayant pour objet de lutter contre les fausses nouvelles de nature à susciter une atteinte à l'ordre public de pouvoir exercer les droits reconnus à la partie civile en ce qui concerne l'infraction prévue par l'article 27 de la même loi (R17).**

<sup>200</sup> « Vos rapporteurs ne recommandent pas d'abandonner le cadre de la loi de 1881 : il est préférable au contraire de réformer ce texte, en l'adaptant au développement d'Internet, qui remet en cause un certain nombre de ses mécanismes », Rapport sur l'équilibre de la loi du 29 juillet 1881 à l'épreuve d'Internet, Sénat, document n° 767, 6 juillet 2016.



## Sanction civile proportionnée de la diffusion de fausses nouvelles

Même si la répression pénale est un instrument essentiel de la lutte contre les phénomènes de désinformation du fait de leurs impacts collectifs très forts, il ne faut pas sous-estimer l'efficacité potentielle des actions civiles. Dans plusieurs domaines comme la lutte contre la contrefaçon ou la protection de la vie privée, les actions civiles ont montré leur efficacité à côté des actions pénales. Ainsi l'article 9 du code civil (créé par la loi du 17 juillet 1970) permet de mettre en cause la responsabilité civile de celui qui commet quelque atteinte que ce soit à la vie privée d'autrui.

L'un des avantages de ce dispositif complémentaire serait notamment de pouvoir faciliter la prise en compte par la juridiction de l'influence ou de la popularité numérique de celui qui a diffusé sciemment la fausse nouvelle. Outre le préjudice moral et économique de la victime, la loi peut demander que le juge civil prenne aussi en compte pour proportionner sa sanction deux variables : celle de la viralité de la diffusion et celle de l'influence relative de celui qui diffuse ou relaye la diffusion fautive.

Le rapport parlementaire du Sénat de 2016 déjà cité, avait d'ailleurs proposé de « permettre une réparation des préjudices nés des abus de la liberté d'expression sur le fondement de la responsabilité civile de droit commun<sup>201</sup> ».

Par ailleurs, la jurisprudence civile qui pourrait se développer sur la base d'une telle disposition permettrait de prévoir une protection plus large que celle assurée, au pénal, par l'article 27 de la loi de 1881 en ne visant plus uniquement les fausses nouvelles risquant de troubler l'ordre public mais plus largement toutes les diffusions préjudiciables de fausses nouvelles.

---

201 Proposition n° 18, rapport n° 767, Sénat, 6 juillet 2016.

Alors que la Cour de cassation limite la compétence du juge pénal français pour sanctionner un contenu publié en ligne à l'étranger<sup>202</sup>, le juge civil se reconnaît plus facilement compétent à l'encontre d'une diffusion étrangère pour autant que le contenu litigieux est accessible en ligne depuis le territoire national et qu'au moins une partie du dommage qu'il cause est localisé en France.

## ➔ Recommandation (R18)

**Introduire dans la loi sur la confiance dans l'économie numérique (LCEN) un nouvel article engageant la responsabilité civile du diffuseur de mauvaise foi d'une fausse nouvelle préjudiciable, laquelle pourrait être rédigée comme suit :**

*« Toute diffusion par voie numérique d'une nouvelle que l'on sait être inexacte et qui porte préjudice à autrui engage la responsabilité civile de celui qui la commet ainsi que de toute personne qui la rediffuse en connaissance de cause.*

*Pour fixer les dommages et intérêts, la juridiction prend en considération distinctement :*

- 1. Les conséquences économiques négatives de la diffusion ;*
- 2. Le préjudice moral causé par cette dernière ;*
- 3. L'étendue et la vitesse de propagation de la diffusion ;*
- 4. Et le niveau d'audience et de popularité numérique de son auteur».*

---

202 Voir notamment l'arrêt de sa chambre criminelle : Cass. crim., n° 15-86645, 12 juillet 2016.

## S'agissant de l'intervention d'une instance de contrôle indépendante

Même si l'action judiciaire fondée sur les infractions de la loi de 1881 peut désormais donner lieu à comparution immédiate (depuis la loi 24 août 2021 confortant le respect des principes de la République), les délais de procédure judiciaire (en particulier pour obtenir une décision définitive au fond) demeurent largement inadaptés à la réaction rapide face à la diffusion virale de certaines fausses nouvelles.

Il est donc pertinent d'encourager les premières initiatives prises ces dernières années pour donner compétence à une autorité administrative indépendante nationale et lui permettre d'intervenir d'office ou sur signalement en vue d'enjoindre aux services numériques concernés de prendre des mesures rapides de prévention ou de retrait de contenus.

C'est ce rôle de supervision que la récente loi du 4 août 2021 confortant le respect des principes de la République a déjà confié au Conseil supérieur de l'audiovisuel (qui va devenir l'Autorité de régulation de la communication audiovisuelle et numérique au 1<sup>er</sup> janvier 2022) en le chargeant (par l'article 42 de cette loi de 2021) de veiller au respect par les plateformes de leurs obligations de retrait rapide de certains contenus illicites graves (parmi lesquels ne figurent cependant pas les fausses nouvelles réprimées par l'article 27 de la loi de 1881).

Mais, en matière de désinformation, la loi du 22 décembre 2018 - par ailleurs dédiée à la lutte contre les fausses nouvelles de nature à fausser les opérations électorales - a également donné au même CSA (future ARCOM) une compétence plus générale de lutte contre la diffusion de fausses nouvelles.

L'article 12 de cette loi de 2018 affirme en effet sa compétence pour lutter contre les fausses informations pouvant porter atteinte à la sincérité d'un scrutin, mais plus largement contre toutes celles dont la diffusion serait de nature à troubler l'ordre public (c'est-à-dire, les fausses nouvelles considérées comme potentiellement répréhensibles au sens de l'article 27 précité).

Il est notamment prévu que la future ARCOM puisse «en cas de nécessité» adresser aux grandes plateformes des «recommandations visant à améliorer la lutte contre la diffusion de telles informations» et s'assurer par ailleurs du respect par ces plateformes des mesures préventives qu'elles doivent adopter en particulier pour lutter «contre les comptes propageant massivement de fausses informations» (articles 11 et 12 de la loi du 22 décembre 2018).

Dans son premier rapport publié en juillet 2020 à ce sujet, le CSA s'est d'ailleurs prononcé en faveur d'une «régulation prescriptive et ciblée sur la responsabilisation des réseaux sociaux mise en œuvre par une autorité administrative indépendante» (rôle qu'il entend assumer). Mais pour autant la lecture de ce premier rapport ainsi que des dispositions de la loi du 22 décembre 2018 montre que cette intervention possible de la future ARCOM auprès des grandes plateformes demeure encore une prérogative aux contours trop vagues pour être véritablement efficace et qu'il convient donc de renforcer.

En effet, on peut estimer qu'il manque au moins une procédure formalisée de signalement auprès de la future ARCOM qui soit ouverte à tout citoyen. Cette procédure de signalement ne devrait pas avoir pour but de demander le retrait d'un ou de plusieurs contenu(s) particulier(s) pouvant constituer une fausse nouvelle susceptible de troubler l'ordre public. En revanche, il s'agirait de faire connaître a posteriori à l'ARCOM soit les difficultés rencontrées par un requérant pour que la plateforme considérée prenne en compte sa réclamation à l'encontre d'un contenu qu'il estimerait préjudiciable, soit au contraire la contestation par l'auteur d'un contenu retiré par la plateforme et que ce dernier estimerait injustifiée. Dans les deux cas, l'ARCOM pourrait intervenir auprès de ladite plateforme pour s'assurer que le point de vue du requérant a bien été pris en compte et a fait l'objet d'une réponse appropriée par la plateforme.

## → Recommandation (R19)

Compléter l'article 17-2 de la loi du 30 septembre 1986 pour prévoir :

- ♦ d'une part, que l'ARCOM peut être saisie par toute personne ayant rencontré une difficulté pour obtenir l'intervention et la coopération d'une plateforme afin de prévenir ou stopper la diffusion massive d'un contenu susceptible de véhiculer une fausse nouvelle pouvant troubler l'ordre public, ou pour contester une décision affectant l'un de ses contenus ;
- ♦ et d'autre part, lui permettre d'enjoindre à la plateforme concernée – une fois avisée par l'ARCOM - de lui fournir à bref délai un état des mesures prises par ses soins dans le cas considéré et de coopérer avec l'ARCOM pour rechercher et mettre en place les mesures préventives ou curatives appropriées au traitement d'un tel cas.

## S'agissant de la responsabilisation des plateformes pour prévenir la diffusion massive de fausses nouvelles répréhensibles

Le caractère mondialisé de l'espace numérique et des principales plateformes qui y sont actives ne permet pas d'imaginer que des mesures juridiques purement nationales puissent avoir un effet suffisant sur le phénomène de la diffusion des « infox ». C'est pourquoi il apparaît particulièrement pertinent d'encourager, à l'occasion de l'adoption de la prochaine législation sur les services numériques (DSA), la mise en place de règles contraignantes imposées aux dites grandes plateformes pour lutter contre la diffusion de fausses nouvelles.

Comme l'a dit justement le CSA, la lutte contre les contenus préjudiciables est une « *politique publique qui doit reposer sur un équilibre entre une politique répressive et une logique de responsabilisation accrue des acteurs par la régulation ex ante*<sup>203</sup> ».

Pour ce faire, il conviendrait que l'obligation de « modération des contenus » qui devrait être imposée à ces plateformes vise de manière suffisamment explicite les fausses nouvelles susceptibles de troubler l'ordre public. Le projet mis en discussion en 2020 se référait simplement, concernant les plateformes et les fournisseurs de services intermédiaires, aux « *contenus illicites* » ainsi qu'aux « *informations incompatibles avec leurs conditions générales* ».

Certes, la définition proposée du « *contenu illicite* » couvre toute information qui « *n'est pas conforme au droit de l'Union ou au droit d'un État membre, quel qu'en soit l'objet précis ou la nature précise* », ce qui pourrait recouvrir la fausse nouvelle répréhensible au sens de l'article 27 de la loi de 1881. Mais plutôt que de laisser la porte ouverte à des débats d'interprétation du droit français de la presse par les plateformes et les juridictions des différents États-membres, il serait préférable que les obligations de modération des contenus imposées par le futur DSA vise, en particulier, les « fausses nouvelles susceptibles de troubler l'ordre public ».

À défaut d'une inscription explicite dans le nouveau règlement, on pourrait pour le moins rechercher que les critères de sensibilité et de viralité des fausses nouvelles répréhensibles puissent être discutés et convenus en amont dans le cadre d'un mécanisme de co-régulation qui serait engagé entre les autorités européennes, les régulateurs nationaux et les principales plateformes.

<sup>203</sup> Rapport CSA, 2020, op.cit., p.14

Par ailleurs, on pourrait proposer de compléter la mise en œuvre des mécanismes de modération des contenus par la création – sur une base initiale nationale (mais qui pourrait avoir vocation à s'européaniser en cas de succès) – d'un mécanisme volontaire de recours à un organe d'expertise indépendant qui – avec l'accord du plaignant – pourrait être saisi à très bref délai par la plateforme pour recueillir son avis consultatif sur la qualification d'un contenu comme constituant ou non une fausse nouvelle susceptible de troubler l'ordre public.

Un tel mécanisme pourrait être encore renforcé par l'adjonction, dans les conditions générales de la plateforme qui le souhaite, d'une disposition contractuelle spécifique par laquelle chaque usager qui signalerait une fausse information potentiellement répréhensible serait réputé avoir accepté par avance la possibilité de la saisine de l'organe d'expertise extérieur et s'interdirait de saisir une quelconque juridiction tant que l'organe d'expertise n'aurait pas rendu son avis.

## ➔ Recommandations

◆ Introduire explicitement dans le DSA une disposition qui reconnaît que les fausses nouvelles susceptibles de troubler l'ordre public constituent des contenus répréhensibles devant être pris en compte par le mécanisme de modération des contenus imposés aux plateformes (R21).

◆ Mettre en place un organe d'expertise extérieur indépendant avec lequel les plateformes pourraient signer une convention leur permettant, face à la demande de retrait d'un contenu argué d'être une fausse nouvelle répréhensible, de soumettre le cas à ces experts extérieurs dont elles s'engageraient à suivre l'avis (R22).

◆ Dans une version plus poussée de cette recommandation, les conditions générales de la plateforme pourraient prévoir que le plaignant serait réputé accepter contractuellement le principe de cette expertise extérieure et s'engager à ne pas engager d'action contentieuse avant d'en connaître le résultat.

◆ Créer un régime de co-régulation entre plateformes, régulateurs, et société civile dans le cadre de la législation sur les services numériques (Digital Services Act). Mettre en place un mécanisme de coopération exigeante avec les plateformes pour concevoir, mettre en œuvre et évaluer les mesures appliquées par les plateformes pour modérer les contenus en préservant la liberté d'expression et les droits humains (R23).

## CHAPITRE 6

# → Une opportunité démocratique : développer l'esprit critique et l'Éducation aux Médias et à l'Information (EMI)

Comme le soulignent certains chercheurs<sup>204</sup>, beaucoup de phénomènes imputés aux algorithmes sont en fait initiés par nos comportements en ligne puis amplifiés par les modèles algorithmiques, ce qui n'est pas forcément une mauvaise nouvelle. En effet, cela signifie que, d'une certaine façon, nous avons encore notre destin entre les mains à condition de comprendre les processus mentaux qui conduisent à ces comportements et d'y remédier. Nous avons en nous les ressources pour éviter de tomber dans le piège que constituent les fausses informations et les erreurs de raisonnement. Le développement de ces ressources est devenu un des enjeux essentiels dans un monde numérique où chacun a le droit d'intervenir – par un blog, un compte Facebook ou encore un commentaire sur site d'information – dans l'espace public. Ces ressources consistent, en premier lieu, à tenter de raisonner librement et aussi juste

que possible, autrement dit à développer l'esprit de méthode, qu'on peut aussi appeler l'esprit critique. Comme le rappelait Descartes dans *Discours de la méthode* : « *Ce n'est pas assez d'avoir l'esprit bon, mais le principal est de l'appliquer bien.* »

Deux choses à retenir avant de définir ce qu'est l'esprit critique et ce que la science contemporaine nous en dit. D'une part, il ne saurait se confondre avec le fait de douter de tout par principe. Ce doute, souvent revendiqué par la pensée conspirationniste, prétend exister pour lui-même et sans aucune contrainte. Or ce scepticisme sans limites peut facilement devenir une forme de nihilisme. Lorsque l'on cherche, sans respecter les canons de la méthode et de l'administration de la preuve, des versions alternatives aux réalités historiques ou à des faits d'ac-

204 Contributions écrites dans les archives du rapport.



tualité, on aboutit le plus souvent à des récits sans consistance épistémique.

D'autre part, l'esprit critique ne peut pas non plus se résumer aux exercices de démystification des fausses informations que proposent les fact-checkers. L'état de la science nous permet de savoir que ces efforts ne sont pas vains et proposent une réponse possible à la diffusion de fausses informations<sup>205</sup>. Malgré tout, ceux qui sont le plus susceptibles d'être séduits par la mésinformation sont aussi ceux qui sont le moins réceptifs aux exercices de fact-checking<sup>206</sup>. Pire encore, les tentatives de rétablissement des faits peuvent affermir leurs convictions<sup>207</sup> surtout si celles-ci sont de nature à mettre en péril leur système de représentations<sup>208</sup>. Ces effets de renforcement paradoxaux sont bien connus de la littérature sous le nom d'effet « boomerang »<sup>209</sup> ou « contrecoup »<sup>210</sup>. Nous pouvons avoir

la tentation d'aller chercher des informations qui vont dans le sens de nos croyances pour nous rassurer<sup>211</sup> ou, au contraire – mais cela aboutit aussi à une défense de nos convictions – rechercher avec une certaine mauvaise foi des contre-arguments aux faits qui nous sont exposés<sup>212</sup>.

Il est utile de remarquer que la contradiction rationnelle est moins susceptible d'échouer lorsqu'elle est émise par un membre de notre groupe social<sup>213</sup> ou politique<sup>214</sup>. D'une façon générale, la contradiction a moins de chances d'être repoussée a priori si elle est endogène. En poussant cette constatation jusqu'à son terme logique, on aboutit à la conclusion que la critique la plus efficace est celle qui vient... de nous-mêmes ; ce qui est une première esquisse de ce qu'est l'esprit critique. Il faut cependant pousser l'enquête un peu plus loin pour en cerner avec précision les contours.

205 Walter, N., Cohen, J., Lance Holbert, R., and Morag, Y. (2019) «Fact-checking: A meta-analysis of what works and for whom. *Political Communication*», 37, 350–375.

206 Lyons, B., Mérola, V., Reifler, J. et Stoeckel, F. (2020), «How politics shape views toward fact-checking: Evidence from six European countries», *The International Journal of Press/Politics*, 25, pp.469–492.

207 Nyhan, B. et Reifler, J. (2010), «When corrections fail: The persistence of political misperceptions», *Political Behavior*, 32-2, pp.303–330.

208 Ecker, U. K. H., et Ang, L. C. (2019), «Political attitudes and the processing of misinformation corrections», *Political Psychology*, 40-2, pp.241–260.

209 Hart, P. S., et Nisbet, E. C. (2012), «Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies» *Communication Research*, 39-6, pp.701-723.

210 Guess, A., et Coppock, A. (2018). Does counter-attitudinal information cause backlash? Results from three large survey experiments. *British Journal of Political Science*, 1– 19.

211 Khanna, K. et Sood, G. (2018), «Motivated responding in studies of factual learning», *Political Behavior*, 40-1, pp. 79–101.

212 Chan, M. S., Jones, C. R., Hall Jamieson, K. et Albarracín, D. (2017), «Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation», *Psychological Science*, 28-11, pp.1531–1546.

213 Margolin, D. B., Hannak, A., et Weber, I. (2018), «Political fact-checking on Twitter: When do corrections have an effect?», *Political Communication*, 35-2, pp.196–219. Vraga, E. K., et Bode, L. (2017), «Using expert sources to correct health misinformation in social media», *Science Communication*, 39-5, pp.621–645.

214 Swire, B., Berinsky, A. J., Lewandowsky, S., et Ecker, U. K. H. (2017), «Processing political misinformation: Comprehending the Trump phenomenon», *Royal Society Open Science*, 4-3.

## Éléments de définition de l'esprit critique

Il existe plusieurs définitions de ce que l'on peut entendre par « esprit critique » dans la littérature. Leur dénominateur commun est de définir l'esprit critique comme la capacité à évaluer correctement les contenus et les sources des informations à notre disposition<sup>215</sup> afin de mieux juger, mieux raisonner, ou prendre de meilleures décisions. Évaluer la qualité épistémique d'une information consiste à déterminer si l'information a de bonnes chances de correspondre à la réalité, donc si elle mérite notre confiance. Nous pouvons donc définir l'esprit critique comme la capacité à faire confiance à bon escient, après évaluation de la qualité des informations, opinions, connaissances à notre disposition, y compris les nôtres. Il se trouve que cette compétence est une prédisposition chez l'être humain.

Par exemple, les enfants de trois ans sélectionnent leurs informateurs en fonction de leur lien de proximité avec eux. Les adultes familiers étant moins susceptibles d'avoir des motifs de les tromper<sup>216</sup>, les enfants prennent de préférence des informations auprès d'adultes bienveillants ou qui montrent du respect pour les normes socio-morales<sup>217</sup>, tout en écartant les informateurs qui ont été décrits comme méchants ou menteurs par d'autres<sup>218</sup>. Ces enfants de trois ans préfèrent également suivre les opinions d'individus qui font preuve d'une certaine forme de connaissance générale<sup>219</sup> ou qui ont un accès direct et perceptif à l'information<sup>220</sup>.

Par conséquent, les êtres humains sont très tôt dotés de dispositifs de vigilance épistémique qui permettent de détecter une partie des informations trompeuses diffusées par ruse ou par incompetence. Cependant, dans l'environnement numérique, ces dispositifs qui nous permettent de raisonner et d'écarter les sources d'informations douteuses entrent en compétition avec d'autres qui nous incitent à croire trop facilement et à nous laisser tromper. Comme nous l'avons vu dans le chapitre portant sur les mécanismes psychosociaux de la désinformation, notre esprit est fortement tenté d'accepter des idées vraisemblables et qui n'entraînent pas des processus analytiques coûteux en termes intellectuels. Notre disposition à être méinformé vient en partie d'une forme d'avarice cognitive.

215 Lai, E. R. (2011), «Critical thinking: A literature review», *Pearson's Research Reports*, 6, pp.40-41.

216 Harris, P. L. et Corriveau, K. H. (2011), «Young children's selective trust in informants», *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), pp.1179–1187 ou Koenig, M. A. (2010), «Selective Trust in Testimony: Children's Evaluation of the Message, the Speaker, and the Speech Act» In T. Szabo-Gendler and J. Hawthorne, *Oxford studies in epistemology* (pp. 3-253), Oxford University Press.

217 Vanderbilt, K. E., Liu, D. et Heyman, G. D. (2011), «The development of distrust», *Child development*, 82(5), pp. 1372-1380.

218 Mascaro, O. et Sperber, D. (2009), «The moral, epistemic, and mindreading components of children's vigilance towards deception», *Cognition*, 112(3), pp.367-380.

219 Pasquini, E. S., Corriveau, K. H., Koenig, M. et Harris, P. L. (2007), «Preschoolers monitor the relative accuracy of informants», *Developmental psychology*, 43(5), pp.1216-1226.

220 Numrsoo, E. et Robinson, E. J. (2009), «Identifying unreliable informants: Do children excuse past inaccuracy?», *Developmental science*, 12(1), pp.41-47.

Par ailleurs, les mécanismes de la vigilance épistémique peuvent être trompeurs en particulier sur les réseaux sociaux qui perturbent notre calibrage social – nous l’avons vu dans le chapitre portant sur la régulation algorithmique. La confiance que nous pouvons avoir dans le jugement des autres, tel qu’il peut être évalué par sa visibilité sociale, est profondément affectée par le basculement de la vie sociale en ligne. Nous ne pouvons plus compter seulement sur nos dispositions naturelles pour évaluer intuitivement l’information mais devons cultiver de nouvelles aptitudes mentales, principalement par l’éducation et la formation à la pensée critique.

## L'espoir raisonnable du développement de l'esprit critique

Plusieurs études nous permettent de penser, au-delà d'un doute raisonnable, que l'esprit critique et la pensée analytique permettent de mieux résister aux fausses informations<sup>221</sup> et notamment aux théories du complot<sup>222</sup>. Ils nous rendent aussi plus aptes à réformer notre jugement lorsque cela est nécessaire<sup>223</sup>. Ces résultats encourageants ne sont pas seulement le fait d'études de laboratoire, ils peuvent aussi être observés dans des dispositifs pédagogiques<sup>224</sup> qui ont montré les effets positifs de l'éducation à la pensée critique avec un transfert de compétences, en particulier si l'intervention est spécifiquement conçue pour encourager ce transfert (par exemple via une pratique répétée, l'utilisation d'exemples de différents élèves et l'explicitation de règles générales à appliquer dans une variété de contextes et de contenus). En d'autres termes, les compétences analytiques acquises dans un exercice peuvent être mobilisées dans d'autres types d'exercices si l'enseignement de l'esprit critique, se fait de façon adéquate.

De même, une méta-analyse<sup>225</sup> de la littérature scientifique a mis en évidence un effet bénéfique global des interventions pédagogiques visant à développer la pensée critique lorsque ces interventions incluent un dialogue et un échange entre les élèves, des problèmes concrets, localisés et réalistes sur lesquels travailler, un mentorat adapté du côté de l'enseignant, et des interventions métacognitives, c'est-à-dire permettant au sujet apprenant de prendre conscience des mécanismes de sa propre pensée.

Un travail important reste néanmoins à faire pour rendre opérationnelles les initiatives concernant le développement de l'esprit critique. En effet, ce que l'on appelle « enseignement de la pensée critique » recouvre des réalités très hétérogènes. Dans les deux méta-analyses récentes citées ci-dessus, les auteurs indiquent que la difficulté de leur tâche réside dans la diversité énorme entre les études en termes de durée d'interventions, d'intensité, de contenu, d'âge ciblé, de méthodes de mesure d'impact et de qualité. Les interventions peuvent être limitées à quelque leçons visant à doter les élèves de compétences argumentatives<sup>226</sup>, ou bien s'étendre à l'échelle d'un pays sur une longue durée<sup>227</sup>, tandis que les effets de transfert de compétences

221 Ross, R. M., Rand, D. G., et Pennycook, G. (2021), «Beyond «fake news»»: Analytic thinking and the detection of false and hyperpartisan news headlines», *Judgment et Decision Making*, 16(2), pp.484-504.

222 Swami, V., Voracek, M., Stieger, S., Tran, U. S. et Furnham, A. (2014), «Analytic thinking reduces belief in conspiracy theories», *Cognition*, 133(3), pp.572-585.

223 De keersmaecker, J. et Roets, A. (2017), ««Fake news»: Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions», *Intelligence*, 65, pp.107-110.

224 Halpern, D. F. (2013), *Thought and knowledge: An introduction to critical thinking*. Psychology Press.

225 Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., et Persson, T. (2015), «Strategies for teaching students to think critically: A meta-analysis», *Review of Educational Research*, 85(2), pp.275-314.

226 Kuhn, D., Zillmer, N., Crowell, A. et Zavala, J. (2013), «Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentative competence», *Cognition and Instruction*, 31(4), pp.456-496.

227 Herrnstein, R. J., Nickerson, R. S., de Sanchez, M., et Swets, J. A. (1986), «Teaching thinking skills», *American Psychologist*, 41(11), 1279.

à long terme et à distance sont rarement évalués. De plus, l'enseignement de l'esprit critique a des objectifs si larges qu'on peut regrouper sous ce vocable des actions visant à améliorer aussi bien le raisonnement, la lecture et l'interprétation textuelle que les compétences scientifiques ou argumentatives.

De très nombreuses initiatives sont prises par l'Éducation nationale, les associations et les écoles de journalisme pour développer l'esprit critique et l'Éducation aux médias et à l'information (EMI). Or lorsqu'elles évaluent leurs effets pédagogiques et produisent des données sur leur travail, ce qui est loin d'être systématique, ces données restent souvent éparses et hétérogènes, rendant difficile la constitution d'un socle de connaissances et d'un programme d'actions.

## → Recommandation (R24)

**Mandater un organisme dont le but serait de réunir l'ensemble des données produites, de penser des protocoles standardisés et de mettre en place l'évaluation du matériel pédagogique et des dispositifs de formation. Pour mener à bien ce projet, il faudrait créer une délégation interministérielle dédiée, associant les principaux protagonistes (ministères, associations, médias, bibliothèques...), chargée de se concerter, de mutualiser, de valoriser les ressources à disposition.**

## → Recommandation (R25)

**Solliciter l'expérience des enseignants pour qu'ils recensent les aspects des programmes les plus contre-intuitifs pour les élèves et les erreurs les plus fréquentes qui en découlent, notamment en termes de raisonnement. Cette cartographie des difficultés cognitives permettrait de poser les jalons d'une pédagogie de la métacognition.**

Ces erreurs de raisonnement types qu'il s'agit de repérer peuvent surgir dans tous les domaines de la connaissance (physique, biologie, mathématiques, sciences économiques et sociales, histoire, philosophie...), raison pour laquelle il ne s'agirait pas de créer de nouveaux cours d'esprit critique mais de souligner qu'apprendre à raisonner est aussi important qu'apprendre à lire, écrire ou compter, et d'en tirer les leçons pour l'ensemble du parcours pédagogique.

Les programmes scolaires sont parsemés de ces difficultés cognitives qui n'ont pas encore été systématiquement repérées. Pour n'en prendre qu'un exemple, la théorie de l'évolution se heurte chez les élèves à des obstacles cognitifs spontanés<sup>228</sup>. L'enjeu ici n'est pas seulement de comprendre la théorie de Darwin, mais également de montrer aux élèves pourquoi celle-ci leur résiste. Ainsi, ils apprendraient à développer leur propre pensée en entrant dans un apprentissage de métacognition.

Pour prendre un autre exemple classique de biais cognitif<sup>229</sup>, la confusion fréquente entre corrélation et causalité pourrait, elle aussi, susciter une attention pédagogique toute particulière que ce soit en mathématiques, en physique, en sciences économiques et sociales, en histoire ou même en philosophie. On pourrait aussi réfléchir de façon critique sur l'argument du *is fecit cui prodest* (à qui profite le crime), prologue à toutes les théories conspirationnistes. Ce point pourrait être abordé tout aussi bien en histoire, en sciences économiques et sociales qu'en philosophie. Les exemples et les occasions pédagogiques ne manquent pas où l'on pourrait utilement introduire un éclairage par l'esprit critique. Les recherches s'accordent à considérer que cette initiation à la pensée analytique peut être faite dès le plus jeune âge, en plein accord avec la théorie de l'inoculation<sup>230</sup> qui consiste à confronter avant l'heure les individus aux arguments trompeurs qu'ils pourront rencontrer plus tard sur les réseaux sociaux. Cette anticipation renforce en quelque sorte le système immunitaire intellectuel<sup>231</sup> des individus, afin qu'ils soient plus aptes à identifier la fausse information et ses arguments, à la rejeter ou du moins à s'en méfier. Ce type de techniques est particulièrement adapté aux jeunes esprits en formation car elles peuvent se décliner en jeu<sup>232</sup> qui vont constituer des initiations aux techniques de désinformations et d'exploitation de nos illusions mentales<sup>233</sup>.

228 Bronner G. (2007), « La résistance au darwinisme : croyances et raisonnements », *La Revue Française de Sociologie* et Bronner, G. (2014), « Cognition et formation académique - Les professeurs de SVT face au « problème des éléphants », *Revue Européenne des Sciences Sociales*, 52-1, pp. 139-161.

229 Tversky, A., Kahneman, D. et Slovic P. (Eds.) (1984), *Judgment under uncertainty : Heuristics and biases*, Cambridge, Cambridge University Press.

230 J. Cook, S. Lewandowsky, et Ecker, U. K. H. (2017), « Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence ». *PLOS ONE*, 12:e0175799.

231 Papageorgis, D. et McGuire, W. J. (1961), « The generality of immunity to persuasion produced by pre-exposure to weakened counterargument », *The Journal of Abnormal and Social Psychology*, Vol. 62-3, p. 475-481.

232 Roozenbeek, J. et van der Linden, S. (2019) « Fake news game confers psychological resistance against online misinformation ». *Palgrave Communications*, 5:65.

233 Basol, M., Roozenbeek J., et Van der Linden, S. (2020) « Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news. » *Journal of Cognition*, 3:1-9.

## L'EMI et l'esprit critique : deux démarches complémentaires

Parallèlement au développement de l'enseignement de l'esprit critique, il sera utile et complémentaire d'améliorer l'éducation aux médias et à l'information. En France, l'Éducation nationale la conçoit comme un enseignement qui permet aux élèves de se construire une culture médiatique, informationnelle, numérique et civique. L'EMI est inscrite dans la Loi d'Orientation pour la Refondation de l'École du 8 juillet 2013 et constitue un des éléments du « parcours citoyen » (2016) pour les élèves du premier et du second degré.

Ses effets bénéfiques sur la stimulation de notre système immunitaire intellectuel, largement prouvés en Finlande<sup>234</sup> et dans les pays du nord de l'Europe en général<sup>235</sup>, ont également été mesurés en France, où une étude a montré son influence positive sur la consommation de l'actualité par les jeunes<sup>236</sup>.

Le développement de l'EMI est d'autant plus crucial que l'écosystème médiatique devient de plus en plus complexe et que l'on observe une contamination des supports traditionnels (journaux, radio et télévision) par la logique numérique. Ainsi, les considérations éditoriales tiennent de plus en plus compte des mécanismes d'économie de l'attention, cherchant à optimiser la visibilité de leurs produits et à s'adapter au design des plateformes numériques. Comme l'explique Jean-François Dumas, président de l'agence *Influence communication*, qui scrute quantitativement le fonctionnement du monde de l'information professionnelle, « le problème, c'est que les médias traditionnels sont en train d'agir et de se comporter comme les médias sociaux le font. Il se transpose une culture de médias sociaux dans le média traditionnel<sup>237</sup>. » Les médias sont donc tentés de promouvoir les hameçons attentionnels que sont notamment les informations fondées sur la peur ou la conflictualité. L'histoire des médias a toujours donné l'exemple de cet ajustement réciproque entre l'offre et la demande, mais Internet a ouvert la voie à sa massification<sup>238</sup>. Il semble d'ailleurs que cette logique de contamination des médias traditionnels par le monde numérique soit très vive en France<sup>239</sup>.

Et celle-ci est d'autant plus prégnante qu'une partie importante de la manne publicitaire a migré des médias conventionnels vers les grands opérateurs du net. Ainsi, aux États-Unis, en

234 Standish, R. (1er mars 2017) « Why is Finland able to fend off Putin's information war? » Foreign Policy. Retrieved ; <https://foreignpolicy.com/2017/03/01/why-is-finland-able-to-fend-off-putins-information-war/>

235 <https://www.rcmediafreedom.eu/Publications/Reports/Common-sense-wanted.-Resilience-to-post-truth-and-its-predictors-in-the-new-Media-Literacy-Index-2018>

236

237 <https://files.influencecommunication.com/bilan/bilan-2016-qc.pdf>

238 Bronner, G. (2021), *Apocalypse cognitive*, Paris, Puf.

239 Christin, A. (2020), *Metrics at Work: Journalism and the Contested Meaning of Algorithms*, Princeton, Princeton University Press.

2016, 85% des recettes publicitaires ont été absorbées par Google et Facebook. En dix ans, les journaux traditionnels ont perdu la moitié des milliards de dollars qu'ils récoltaient chaque année en recette publicitaire, tandis que celle de Google était multipliée par cinquante. Cette situation a eu pour effet direct la disparition des emplois dans le secteur de la presse. Les journalistes de presse écrite qui étaient 71 000 en 2008 aux États-Unis n'étaient plus que 39 000 en 2017, soit une baisse de 45% des effectifs, selon les chiffres fournis par le Département américain du Travail. Dans ces conditions, la qualité de l'information et de son éditorialisation ne peut que s'en trouver affectée.

La pression concurrentielle sur le marché de l'information réduit mécaniquement le temps de vérification de l'information en augmentant les risques d'effets de cascade. Le délai pour interroger des experts s'en trouve aussi réduit. Comment s'assurer que le recrutement des experts, par exemple en période de pandémie, suive un processus rationnel plutôt que les injonctions de l'urgence et la facilité de carnets d'adresses peu souvent renouvelés ? Les réponses apportées lors des auditions menées par la commission à ce sujet n'étaient guère convaincantes. À ce sujet, il manque clairement une intermédiation entre le monde de la science et celui des médias. Ces éléments invitent à réfléchir aux moyens de garantir la liberté éditoriale des journalistes, happés à leur corps défendant par des formes de logiques numériques qui peuvent altérer la qualité de leur travail.

Pour toutes ces raisons, l'EMI est devenue fondamentale car elle permet à tout un chacun d'être initié aux réalités complexes de l'écosystème médiatique qui demeure l'un des piliers de la vie démocratique.

Au sein de l'Éducation nationale, en France, cet enseignement est prévu tout au long de la scolarité : des entrées EMI figurent comme domaine de compétences transversales même, si de selon le directeur du CLEMI (Centre de Liaison de l'Enseignement et des Médias d'Information) : « *l'enseignement à l'EMI manque de lisibilité et de continuité dans le parcours de l'élève* »<sup>240</sup>.

L'EMI est assurée de façon très hétérogène<sup>241</sup>, avec des disparités territoriales<sup>242</sup> : « *Bien que supposé concerner les professeurs de toutes les disciplines, le parcours citoyen, dont l'EMI est le pilier, est rarement formalisé, cohérent et évalué. Celui-ci souffre d'un manque d'appropriation par l'ensemble des acteurs de la chaîne éducative* »<sup>243</sup>.

L'EMI est le plus souvent assurée par les professeurs d'histoire-géographie, notamment dans le cadre de l'EMC (Enseignement moral et civique) et par les professeurs documentalistes dont la mission consiste aussi à permettre « *l'acquisition pour tous les élèves d'une culture de*

240 Rapport de Serge Barbet, directeur du CLEMI au ministre de l'Éducation nationale en mai 2021 <https://www.education.gouv.fr/remise-du-rapport-sur-le-renforcement-de-l-education-aux-medias-et-l-information-et-de-la-323927>

241 *Op. cit.*

242 Rapport Barbet, *op. cit.* Audition d'E. Geffray.

243 *Ibid.*



*l'information et des médias* ». Mais ces enseignants interviennent le plus souvent dans le cadre des horaires de leur propre discipline, ce qui prend du temps sur leur propre programme. Ils ne peuvent aujourd'hui enseigner et mettre en œuvre l'EMI à la hauteur des besoins. Le CLEMI a un rôle stratégique puisqu'il assure la formation des enseignants, la création et la mise à disposition de ressources pédagogiques. Chaque académie compte en effet des coordonnateurs du CLEMI, professeurs détachés, qui forment un réseau et dont c'est la mission. Le manque de moyens grève les capacités du CLEMI qui se voit en même temps de plus en plus sollicité<sup>244</sup>.

Pendant, on note des évolutions récentes au sein du MENJS (Ministère de l'Éducation Nationale de la Jeunesse et des sports) qui travaille à l'élaboration d'un guide sur l'enseignement de l'EMI dès l'école primaire à destination des professeurs, décliné par niveau de scolarité<sup>245</sup>.

Des rendez-vous d'évaluation des compétences numériques des élèves de troisième et de terminale ont été mis en place, via PIX<sup>246</sup>.

1<sup>er</sup> degré : Il est prévu la mise en place d'une évaluation en fin de CM2 et d'un « permis internet ».

2<sup>nd</sup> degré : En classe de Seconde, un cours de Sciences Numériques et Technologie comprenant quelques chapitres dédiés à l'EMI vient d'être mis en place, et des compétences EMI ont été introduites dans l'épreuve du grand oral du Bac.

La formation continue est assurée au sein des académies et par le biais des Plans académiques de formation, mais ces stages sont trop peu nombreux et la formation paraît encore insuffisante à de nombreux acteurs.

Par ailleurs, un réseau d'Écoles Académiques de Formation Continue (EAFC) doit être mis en place dans chaque académie en janvier 2022<sup>247</sup> pour proposer des formations et des activités pédagogiques avec mise à disposition de matériel.

La formation initiale n'existe pas partout, et si certains Instituts Nationaux Supérieurs du Professorat et de l'Éducation (INSPE) chargés de la formation initiale des enseignants ont commencé à introduire des modules d'EMI dans leur cursus, cet enseignement est loin d'être généralisé<sup>248</sup>. La mise en place de la certification PIX permet aux élèves de 3e et de Terminale de valider 16 compétences numériques, intégrant l'usage des réseaux sociaux, la connaissance des phéno-

244 Rapport du Conseil Économique, Social et Environnemental (CESE), *Les défis de l'Éducation aux médias et à l'information*, décembre 2019. Pour l'ensemble du territoire, le CLEMI compte 19 ETP, une directrice scientifique et pédagogique 2 jours par semaine et 2 apprentis. Le réseau des coordonnateurs académiques du CLEMI est composé de 33 personnels, avec une hétérogénéité de statuts.

245 Prévu en janvier 2022. Audition d'Édouard Geffray Directeur général de l'enseignement scolaire, 22/11/2021.

246 Le projet PIX est développé conjointement par le MENJS, le MESRI, le CNED, l'UOH, l'Université de Strasbourg et le CNAM et permet d'obtenir une validation des compétences numériques. Le 30 août 2019 est paru un décret concernant le Cadre de Référence des Compétences Numériques, inspiré du référentiel européen DIGCOMP. La plateforme d'évaluation et de certification en ligne Pix n'est pas cantonnée à la sphère de l'Éducation Nationale, mais accessible gratuitement à tous les citoyens ([www.pix.fr](http://www.pix.fr))

247 Audition d'E. Geffray du 22 novembre 2021

248 Contribution de Laurent Petit, chargé de mission numérique INSPE Paris.

mènes de mésinformation et de désinformation. Cette certification n'est cependant pas une appréciation scientifique de l'efficacité des innovations pédagogiques proposées. Comme on le constate, dans ce domaine comme dans celui de l'esprit critique, l'offre est florissante mais manque de coordination, de standardisation et d'évaluation.

## ➔ Recommandation (R27)

**Systématiser la formation à l'esprit critique et à l'EMI, d'une part pour les élèves et ce dès l'école primaire et jusqu'après le secondaire et d'autre part pour les enseignants en formation initiale et continue. Il faut renforcer pour ce faire de manière significative le réseau de référents et coordinateurs académiques dans ces domaines.**

En outre, à la manière de l'enseignement de l'esprit critique, l'EMI ne doit pas être pensée qu'à la seule destination des élèves, étant donné que le problème posé par la mésinformation et la désinformation concerne tous les citoyens<sup>249</sup>. À ce titre, le Conseil supérieur de l'Audiovisuel préconise de renforcer les actions d'éducation aux médias auprès des adultes<sup>250</sup>.

Les acteurs qui interviennent dans le champ de l'EMI sont très nombreux (institutions, collectivités territoriales, animateurs, éducateurs, médiateurs, bibliothécaires, professionnels de l'information, médias, acteurs du numérique... ) et proposent des activités dans de nombreuses structures en contact avec le grand public (associations, centres sociaux, centres aérés, bibliothèques et médiathèques...). Mais il n'existe pas non plus de recensement des dispositifs EMI mis en place sur le territoire (en dehors de l'Éducation Nationale) ni donc, encore une fois, d'évaluation de ces dispositifs hétérogènes.

249 Rapport du CESE de décembre 2019. [https://www.lecese.fr/sites/default/files/pdf/Avis/2019/2019\\_30\\_defi\\_education\\_medias\\_information.pdf](https://www.lecese.fr/sites/default/files/pdf/Avis/2019/2019_30_defi_education_medias_information.pdf)

250 [https://www.lemonde.fr/actualite-medias/article/2021/11/18/le-csa-preconise-de-renforcer-les-actions-d-education-aux-medias-aupres-des-adultes\\_6102554\\_3236.html](https://www.lemonde.fr/actualite-medias/article/2021/11/18/le-csa-preconise-de-renforcer-les-actions-d-education-aux-medias-aupres-des-adultes_6102554_3236.html)

## ➔ Recommandation (R29)

**Créer un continuum entre le temps scolaire, l'université, le monde culturel et le monde du travail et considérer que la formation à l'esprit critique et à l'EMI concerne tous les citoyens en identifiant tous les moments sociaux propices à proposer ce type de formations.**

De ce point de vue, l'EMI est un sujet trans-ministériel qui, outre le MENSJS, concerne notamment les Ministères de la Culture<sup>251</sup>, de l'Enseignement Supérieur, de la recherche et de l'Innovation<sup>252</sup>, de la Cohésion des territoires et des Relations avec les collectivités territoriales. De nombreuses initiatives sont prises, mais de manière un peu anarchique et non coordonnée.

Un autre réseau à solliciter dans le domaine de l'EMI est celui des bibliothèques et médiathèques. En effet, 63% des Français identifient la médiathèque comme le premier lieu ressources numériques<sup>253</sup>. Les 12 429 bibliothèques et 480 bibliothèques universitaires<sup>254</sup> constituent des relais privilégiés pour toucher tous les publics, enfants comme adultes.

Si les ressources d'EMI en bibliothèque sont très riches, l'implication des établissements est inégale et les initiatives peu visibles ou peu identifiées. Le ministère de la culture a donc introduit un volet EMI au « Plan Bibliothèque » en 2018, permettant la mise en place des stages de formation en région et une formation en ligne<sup>255</sup>. Les organismes de formation des bibliothécaires ont commencé à intégrer l'EMI dans leur cursus<sup>256</sup> et de nombreuses ressources à destination des bibliothécaires ont été créées par la BNF, la BPI<sup>257</sup> et Bibliothèque sans Frontières.

Pour compléter ce panorama, ajoutons que l'implication de certains médias et du CSA va crescendo<sup>258</sup>. Près de 1 700 médias ont participé à la dernière édition de la Semaine de la Presse et des Médias à l'École en mars 2021 et tout au long de l'année à travers des collectifs ou associations (Entre les lignes, Cartooning for peace, Lumières sur l'info, Globe reporters, Fake off, etc.<sup>259</sup>)

251 Mars 2018 avec un budget 3 millions d'euros. <https://www.culture.gouv.fr/Aides-demarches/Appels-a-projets/Lancement-d-un-appel-a-projets-national-sur-l-education-aux-medias-et-a-l-information>

252 Contribution de Jean-Michel Jolion, conseiller de la Ministre de l'Enseignement Supérieur

253 C'est ce que révèle le Baromètre du numérique 2018, publié le 3 décembre 2018 dans une étude réalisée par le Credoc et pilotée par l'Arcep, le Conseil général de l'économie et l'Agence du Numérique <https://www.euronews.com/next/2021/11/10/seoul-to-become-the-first-city-to-enter-the-metaverse-what-will-it-look-like>

254 Recensement 2018 du Ministère de la Culture

255 Séminaire « EMI en bibliothèque » organisé à la BPI le 25 novembre 2019 : <https://pro.bpi.fr/seminaire-emi-en-bibliotheques-atelier-integrer-lemi-dans-le-projet/>

256 Contribution écrite de l'ENSSIB (Ecole Nationale Supérieure des Sciences de l'Information des Bibliothèques).

257 Contribution de la BPI.

258 Rapport remis en novembre 2021 sur l'EMI pour l'exercice 2020/2021, le Conseil Supérieur de l'Audiovisuel (CSA)

259 Contributions d'Entre les lignes, Fake Off, We Report, la plateforme d'EMI de France Info, Spicée.

L'EMI apparaît dans l'enseignement des principales écoles de journalisme, de manière plus structurée pour certaines que pour d'autres<sup>260</sup>, signe d'une prise de conscience du problème des fausses informations.

Quant au vaste monde des associations (La ligue de l'enseignement<sup>261</sup>, CEMEA<sup>262</sup>, le réseau des MJC, la Fondation La main à la pâte, l'UNAF<sup>263</sup>, etc.), il est très actif en matière de propositions sur l'esprit critique et l'EMI. Des associations d'éducation populaire s'y investissent depuis plusieurs années, introduisant des modules d'EMI dans la formation initiale et continue de leurs intervenants. Mais l'EMI demeure encore en phase de maturation et d'invention quant à ses contenus<sup>264</sup>.

Pour finir, il reste un autre territoire social de conquête pour l'EMI et pour la formation à l'esprit critique : celui des entreprises privées. On note que certaines fondations de grandes entreprises (GAFA, fondations AXA ou EDF) participent au financement d'actions éducatives en matière d'EMI mais que la visibilité sur leurs initiatives, notamment en termes de formation continue, demeure relativement faible.

## ➔ Recommandation (R28)

**Sensibiliser les chefs d'établissements scolaires, les inspecteurs de l'Éducation Nationale, les recteurs aux enjeux de l'EMI et de la formation à la pensée critique, ainsi que les élus locaux, les responsables de ressources humaines des collectivités locales et les responsables de bibliothèques.**

La formation professionnelle continue permet d'acquérir de nouvelles compétences durant sa vie active pour le retour ou le maintien dans l'emploi, et pour sécuriser ou optimiser les parcours professionnels. Elle constitue une obligation légale (article L6311-1 du Code du Travail). Elle pourrait être un moment idéal pour promouvoir la formation à l'esprit critique et à l'EMI.

260 Contribution écrite de la Conférence des Écoles de Journalisme (CEJ).

261 Contribution de la Ligue de l'enseignement.

262 *Ibid.*

263 *Ibid.*

264 Contribution d'Olivier Magnin Responsable mission nationale déléguée EMI de la Ligue de l'Enseignement

## Conclusion du chapitre 6

Permettre à chaque citoyen de développer son autonomie intellectuelle par la formation (initiale ou continue) à l'esprit critique et à l'EMI doit être une Grande Cause nationale, un objectif prioritaire des démocraties confrontées aux perturbations engendrées par le monde numérique. Cela pourrait se faire en augmentant sa visibilité par la diffusion de messages d'intérêt général dans les médias (R26).

Tout d'abord, il s'agit de la façon la moins liberticide de réguler ce marché de l'information devenu chaotique. Comme chacun d'entre nous est devenu un opérateur sur ce marché, il peut décider de partager ou ne pas partager, d'apprécier ou de ne pas apprécier une information. Pour cette raison, la bonne santé de notre démocratie implique que chaque citoyen améliore sa vigilance intellectuelle.

D'autre part, cette démarche est une façon pour une nation de reprendre une partie de son destin en main. En effet, nous l'avons vu, certaines recommandations essentielles que ce rapport propose dépendent du bon vouloir des grands opérateurs numériques ou d'un rapport de force qu'il s'agit d'engager avec eux. Le développement de l'esprit critique et de l'EMI, au contraire, ne dépend que de la volonté ferme et coordonnée d'une politique nationale. La meilleure façon de nous libérer des asservissements algorithmiques est sans doute de nous en remettre aux ressources puissantes de notre cerveau.

Cet objectif est fondamental, enfin, car il permet de transformer une situation inquiétante en formidable opportunité politique : former des citoyens autonomes dans leur jugement grâce au développement des compétences de métacognition. Ce choix pourrait nous aider à prendre la bonne direction face à ce carrefour sociétal auquel nous sommes confrontés, valorisant une démocratie de la connaissance.

# → Conclusion finale

Ce rapport n'a pas vocation à se poser en juge de vérité ou à éradiquer la désinformation ou la mésinformation en ligne, mais à réfléchir aux moyens techniques, juridiques, sociétaux d'en limiter les conséquences négatives sur la vie démocratique. Une manière possible d'atteindre cet objectif serait d'agir à la fois en amont de la diffusion de fausses informations par des propositions visant à responsabiliser les plateformes et les diffuseurs de publicité en ligne et en aval. Cela implique, d'une part, de renforcer l'éducation aux médias et la vigilance critique à l'égard des contenus en circulation et, d'autre part, de permettre aux chercheurs de comprendre la nature et l'ampleur exacte du phénomène. À ce titre, les données détenues par les entreprises du monde numérique doivent être considérées, à terme, comme un bien commun.

Ce rapport s'est écrit, d'abord, avec l'ambition de tenir compte de l'état de la connaissance et des différentes et nombreuses initiatives qui ont déjà été proposées ou mises en place. Il s'est écrit, ensuite, en intégrant la situation de dépendance stratégique dans laquelle l'Europe se trouve à l'égard de grandes plateformes américaines sur lesquelles elle n'a pas autorité. Il s'est écrit, enfin, avec la conviction qu'il est essentiel de préserver la libre expression des points de vue.

Nos réflexions se sont développées dans un contexte propice, puisque qu'au moment même où nous terminons ce rapport, la commission du marché intérieur du Parlement

Européen (IMCO) vient de voter le texte de la législation sur les services numériques (Digital Services Act), renforçant considérablement, par rapport au texte initial de la Commission européenne, les obligations de transparence et la responsabilité des très grandes plateformes, afin de mieux protéger les utilisateurs et leurs droits fondamentaux en ligne. La prochaine étape sera le vote final du Parlement européen début 2022.

Au terme de notre travail, nous avons la conviction que la révolution numérique, dont nous sommes les contemporains, entraîne une escalade de bouleversements que nous avons encore du mal à appréhender. Une réflexion à peine prospective nous fait entrevoir certains éléments qui, demain, susciteront de nouvelles questions. L'annonce de la création d'un métavers par Mark Zuckerberg, le créateur de Facebook, en fait partie. La question qui se pose de façon lancinante pour les réseaux sociaux se posera derechef pour ce nouveau « graal des interactions sociales », comme aime à le nommer Zuckerberg, qui risque d'envahir rapidement nos vies. Cet univers alternatif, dans lequel nous serions immergés sous forme d'avatar pour rencontrer nos amis, jouer, travailler, ou encore faire du shopping, n'existe pas encore. Mais l'enjeu de la modération y sera plus essentiel encore que sur les réseaux sociaux, car cette technologie étant immersive, on ne peut qu'imaginer l'ampleur des dégâts qui pourraient être occasionnés si la haine en ligne ou le harcèlement y faisaient valoir leur loi.

Cette inquiétude n'est pas vaine : elle est même exprimée directement par Andrew Bosworth, le directeur technique de Meta (ex-Facebook) qui, dans une note interne divulguée le 12 novembre 2021<sup>265</sup>, souligne que des erreurs de modération à répétition pourraient mettre en péril l'existence même de l'entreprise.

Une ultime suggestion que nous pourrions formuler concerne donc la vigilance prospective dont nous devrions faire preuve à l'égard de ces innovations qui sont annoncées et dont les effets pourraient bouleverser notre rapport au réel et à l'information. D'autant que l'initiative de Meta n'est pas isolée. La ville de Séoul a annoncé en novembre la création de « Metaverse Seoul », un univers en trois dimensions, issu des technologies de réalité augmentée et virtuelle, qui deviendra le premier centre de services publics virtuel dans lequel les citoyens seront reçus par des avatars. Le maire de Séoul a pour ambition de faire de la capitale de la Corée du Sud la première grande ville du métavers pour en faire « une ville de coexistence, un leader mondial, une ville sûre, future ville émotionnelle<sup>266</sup> ». Le métavers symbolise notre immersion progressive dans un univers où les mondes réels et virtuels iront jusqu'à se confondre.

À ce titre, il nous paraît pertinent de mener une réflexion sur les enjeux éthiques de ces univers d'immersion numérique, qui repoussent toujours plus loin les limites du monde physique et promettent des interactions sociales d'une nouvelle nature. Cette réflexion pourrait être menée par le Comité national pilote d'éthique du numérique, dans le prolongement de son avis sur les agents conversationnels adopté en septembre 2015 (R30). À terme, elle pourrait ouvrir sur une réflexion plus large à l'échelle internationale, impliquant des experts de l'industrie du numérique, de la communauté académique, de la société civile et des gouvernements.

Au-delà de ces défis, nous sommes aussi conscients des formidables avancées permises par la révolution numérique et des possibilités qu'elle ouvre. La pandémie de COVID-19 a précipité la transformation numérique de nos sociétés, grâce à laquelle nous avons pu basculer du jour au lendemain dans le télétravail, suivre, en temps réel et partout dans le monde, la propagation du virus et de ses variants, créer un vaccin en un temps record et organiser des campagnes de vaccination massives.

D'ores et déjà, de nouveaux modes de travail collaboratif ont émergé qui laissent entrevoir les promesses d'une intensité nouvelle des échanges scientifiques. Ainsi, l'initiative Tela Botanica<sup>267</sup>, par exemple, permet la mise en réseau de dizaines de milliers de botanistes, les uns professionnels, les autres amateurs pour réviser de façon efficace l'intégralité de la nomenclature des plantes existantes en France.

265 <https://www.ft.com/content/d72145b7-5e44-446a-819c-51d67c5471cf>

266 <https://www.euronews.com/next/2021/11/10/seoul-to-become-the-first-city-to-enter-the-metaverse-what-will-it-look-like>

267 <https://www.tela-botanica.org>

De telles plateformes collaboratives facilitent aussi l'identification d'espèces menacées<sup>268</sup> ou la mise en commun de données permettant de reconnaître poissons, champignons, plantes, oiseaux...<sup>269</sup>

Ce travail collaboratif peut aller beaucoup plus loin encore : le jeu *Foldit*, développé par les départements d'informatique et de biochimie de l'Université de Washington, propose aux internautes de résoudre des énigmes scientifiques par l'exploration collective du possible<sup>270</sup>. Il s'agit, par exemple, de tenter librement des combinaisons moléculaires pour mieux cerner la façon dont les protéines peuvent se déployer dans l'espace : déplacer des portions ici, en ajouter là ou même de détruire des liaisons. Ce jeu collectif de construction a permis de résoudre en trois semaines un problème sur lequel les scientifiques planchaient depuis dix ans : la forme réelle d'une enzyme liée au sida chez le macaque rhésus.

Dans le même ordre d'idées, l'une des expressions les plus abouties de l'« intelligence collective » est sans doute l'encyclopédie en ligne Wikipédia, qui n'est pas exempte de critiques mais qui a démontré qu'elle pouvait concurrencer sans difficulté les meilleures encyclopédies traditionnelles<sup>271</sup>.

Le numérique offre donc aussi des ressources pour que notre intelligence collective puisse s'exprimer d'une façon plus efficace que jamais et devienne le support d'une revitalisation démocratique. À partir des années 1960, de nombreux auteurs ont pressenti l'avènement de ce que l'on nomme aujourd'hui la crise démocratique, qui se traduit notamment par des taux de méfiance record au sein de la population à l'égard des médias ou des politiques. Ces auteurs ont appelé à un renouvellement de la vie démocratique sous une forme plus participative<sup>272</sup>. Des théoriciens comme Carole Pateman<sup>273</sup> et Benjamin Barber<sup>274</sup> notamment font de l'implication de chacun dans les affaires publiques la condition de toute liberté politique authentique. Jusqu'à récemment, cette implication de tous se heurtait à des obstacles techniques que peuvent lever les outils numériques.

L'espoir d'une revitalisation de la vie démocratique par le numérique trouve des expressions très concrètes, comme les expérimentations menées à Taïwan<sup>275</sup> sous l'égide d'Audrey Tang<sup>276</sup>, ministre du numérique, qui promeut l'utilisation de plates-formes de délibération citoyenne et

268 <http://www.edgeofexistence.org/instantwild>

269 <http://www.ispot.org.uk/> Dans le fond, cette mutualisation de l'information n'est pas différente de celle qu'avaient permis au XVIII<sup>e</sup> et XIX<sup>e</sup> siècles les amateurs de botanique ou d'entomologie mais, une fois de plus, l'outil technologique permet de donner à cette mutualisation une ampleur sans précédent.

270 <https://fold.it>

271 Audition de Pierre-Yves Beaudoin, président de Wikimedia France.

272 Girard, C. et Legoff, A. (2010), *La démocratie délibérative - Anthologie de textes fondamentaux*, Paris, Hermann.

273 Pateman, C. (1970), *Participation and democratic theory*, Cambridge England, Cambridge University Press, 1970.

274 Barber, B. (1984), *Strong Democracy: Participatory Politics for a New Age*, Berkeley, University of California Press.

275 <https://osf.io/jnq8u/>

276 Audrey Tang s'est fait connaître en tant qu'hacktiviste en aidant les jeunes taïwanais à mener leur mouvement des Tournesols en 2014. Le terme de « Tournesol » est d'ailleurs repris par Lê Nguyễn Hoang, chercheur en informatique et vulgarisateur scientifique à l'École polytechnique fédérale de Lausanne, pour désigner un projet innovant. Il propose de développer un algorithme vertueux qui suggérerait des recommandations – de vidéos par exemple – en intégrant une évaluation collective de la pertinence et de l'utilité d'un contenu.



de co-élaboration des lois. Ces nouveaux dispositifs de consultation démocratique ont permis de trancher des sujets difficiles, comme la vente d'alcool en ligne ou la régulation d'Uber – même si les observateurs ont noté que la participation citoyenne, fondée sur le volontariat, reste limitée et demeure l'apanage d'un public averti.

Ces initiatives sont loin de refléter l'étendue des possibles qu'offre Internet. Nous vivons à l'ère de la datasphère, dans laquelle la plupart de nos activités humaines s'appuient sur des technologies et laissent des traces numériques, produisant ainsi un nouvel espace, une sphère des données qui interagit avec le monde physique<sup>277</sup>. La croissance exponentielle de la datasphère ouvre le champ de la gouvernance numérique pour répondre aux grands défis de notre siècle, à commencer par la dégradation de l'environnement et le changement climatique qui, au-delà de nos démocraties, menacent l'humanité tout entière<sup>278</sup>.

Notre rapport avait pour seule ambition de penser, dans l'urgence, des solutions pour juguler un problème amplifié, voire transformé par le numérique : la désinformation. Ce travail ne nous exonère en rien de la réflexion collective que nous devons mener pour penser le monde de demain. Le numérique est un formidable moyen. Reste à savoir quelle société et quelle démocratie nous souhaitons construire dans ce monde numérique en devenir.

277 Douzet, F. (2020), « Du cyberspace à la datasphère. Enjeux stratégiques de la révolution numérique », *Hérodote*, 177-178, pp. 3-16.

278 Grumbach, S. (2020), « Gouvernance numérique et changement climatique », *Hérodote*, 177-178, pp. 17-32.

# ➔ Recommandations

La désinformation est un phénomène en grande partie licite et relève de la protection de la liberté d'expression au sein de nos démocraties. Nos recommandations ne visent donc pas à l'éradiquer car ce n'est ni possible, ni souhaitable. Elles visent à limiter la propagation des contenus qui nuisent à la vie démocratique, dissuader les comportements malveillants, sanctionner les pratiques illicites, améliorer la prévention des risques et renforcer la vigilance des utilisateurs.

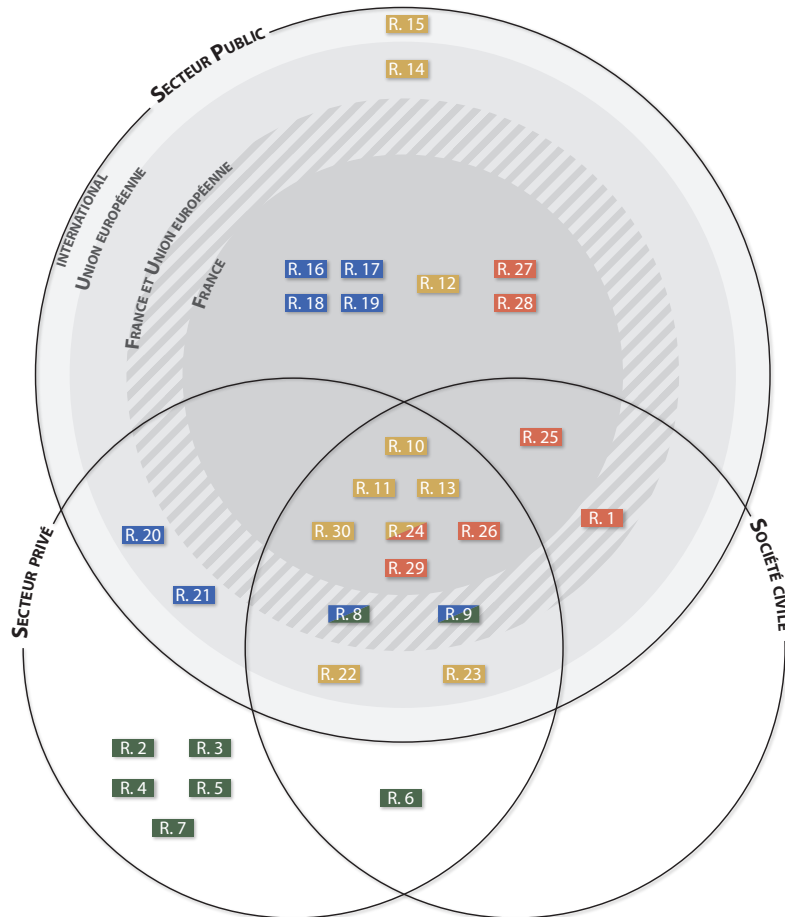
Il n'y a pas de solution miracle. La désinformation en ligne prend de multiples formes, s'appuie sur des techniques en perpétuelle évolution et produit des effets d'ordres différents sur une diversité de publics. Elle relève dès lors d'interventions de différentes natures, que nous avons regroupées en 4 grandes catégories : la régulation, les bonnes pratiques, la gouvernance numérique et l'éducation.

Elle circule au sein d'un écosystème numérique dont la gouvernance est complexe et implique une multitude d'acteurs (plateformes, gouvernements, société civile) qui partagent en commun ce problème, quelles que soient les rivalités qui les animent ou les différends

qui les opposent. Aucun d'entre eux ne peut agir seul. C'est la raison pour laquelle nous présentons dans le graphique ci-dessous les différentes sphères d'action (publique, privée, société civile) qui s'entrecroisent. Nombre des mesures que nous proposons nécessitent une coopération ou une co-régulation entre ces acteurs et se situent à l'intersection de ces ensembles. Enfin, nous avons distingué plusieurs niveaux de gouvernement, car la France ne peut agir seule.

Notre réflexion s'est construite autour de grands thèmes qui constituent autant de chapitres, dont la lecture est indispensable pour donner sens à nos recommandations. Nous avons choisi de conserver cette structure par chapitre pour présenter nos recommandations, par souci de clarté et de cohérence. Pour autant, nombre de nos recommandations sont transversales et dépassent donc cette organisation. C'est notamment le cas des recommandations qui relèvent du droit numérique ou de la recherche. C'est pourquoi, dans un certain nombre de chapitres, nous proposons des renvois vers les recommandations d'autres chapitres.

**30 recommandations pour répondre aux désordres de l'information**  
 Un problème partagé, une réponse globale



**Comment lire cette infographie**

La position d'une recommandation dans l'infographie indique le ou les différents acteurs impliqués dans sa mise en œuvre. Pour le secteur public, nous avons distingué plusieurs niveaux de pouvoir : France, Union européenne, International.

**Type de recommandation**

- R. 1 numéro de recommandation
- R. XX gouvernance numérique
- R. XX recherche et éducation
- R. XX régulation
- R. XX bonnes pratiques

## Mécanismes psychosociaux

### 1. FAVORISER LA RECHERCHE PUBLIQUE

Soutenir et renforcer en France la recherche scientifique sur la désinformation en ligne et sur les ingérences numériques étrangères. Ce soutien pourrait passer par le fléchage de crédits de recherche et la création de postes de recherche.

La France devrait encourager l'Union Européenne à soutenir à son niveau la recherche scientifique sur ces sujets.

## Logiques algorithmiques

### 2. PENSER LA RÉGULATION DU DESIGN DES INTERFACES UTILISATEURS

Ouvrir une réflexion en vue d'une régulation sur l'importance de la question du design des interface utilisateurs

### 3. LUTTER CONTRE LE BIAIS DE POPULARITÉ

Permettre aux utilisateurs de mieux se représenter l'état du réseau et la prévalence réelle des opinions en désactivant par défaut les métriques de popularité et l'éditorialisation algorithmique, et en mettant en avant des métriques permettant de juger de la qualité épistémique des contenus (historique de partage notamment).

### 4. RESPONSABILISER LES INFLUENCEURS

Encourager une modération accrue des influenceurs par les plateformes afin de les responsabiliser.

### 5. METTRE EN AVANT LA COMPÉTENCE

Donner plus d'influence à la compétence en mettant en avant les comptes d'experts et en amplifiant leur contenu (lorsqu'il porte sur un sujet relatif à leur expertise).

### 6. REFLÉTER L'ÉTAT DES CONNAISSANCES

Veiller à ce que, sur certains sujets fermement établis, le classement algorithmique n'induisse pas en erreur le public sur l'état réel des connaissances. Pour cela, encourager un dialogue entre les plateformes et les institutions scientifiques afin que l'existence d'un consensus soit reflétée dans la visibilité accordée aux diverses opinions.

## 7. PRÉVENIR LE RISQUE DE SUR-MODÉRATION

Se prémunir contre le risque de sur-modération en analysant plus finement les signalements d'utilisateurs (signalements en essaim).

# Économie des infox

## 8. RESPONSABILISER LES ACTEURS DE LA PUBLICITÉ PROGRAMMATIQUE

Promouvoir l'investissement publicitaire responsable des entreprises en encourageant le recours, par les annonceurs, les régies, les agences publicitaires et surtout les fournisseurs de technologie publicitaire, à des « listes d'exclusion et d'inclusion de sites web » dynamiques, telles que celles élaborées par exemple par NewsGuard, Global Disinformation Index ou Storyzy. Engager un dialogue avec les fournisseurs de technologie publicitaire afin qu'ils aient recours à ce système, qui permettrait d'assécher de manière considérable l'économie des infox.

Veiller à l'exemplarité en la matière des administrations et des entreprises publiques ayant recours à la publicité programmatique en généralisant pour elles le recours à des listes d'inclusion dynamiques.

Envisager une obligation pour les entreprises engagées dans une démarche RSE à réaliser un audit indépendant annuel complet de leurs campagnes publicitaires programmatiques permettant de dresser la liste complète des adresses web (url) des sites sur lesquels ces campagnes sont diffusées, et à diffuser publiquement cette liste.

Encourager les organismes de labellisation comme l'AFNOR à prendre en compte le problème du financement de la désinformation dans l'attribution de labels « responsables », en imposant des audits réguliers pour les entreprises candidates à ces labels.

Envisager une obligation pour les fournisseurs de technologie publicitaire à sensibiliser leurs clients au risque de financer des sites toxiques dans le cas où ces derniers omettraient d'avoir recours à des listes d'exclusion dynamiques.

Recommander aux sites de presse généraliste de bannir de leurs espaces publicitaires les liens sponsorisés renvoyant vers des sites pièges à clics de désinformation. Les encourager à ne plus faire appel aux sociétés publicitaires les associant à de tels liens sponsorisés

## 9. ENCOURAGER LES BONNES PRATIQUES MISES EN ŒUVRE PAR LES PLATEFORMES DE FINANCEMENT PARTICIPATIF

Réfléchir à une obligation pour les plateformes de financement participatif à indiquer explicitement à leurs utilisateurs toutes les mesures mises en œuvre pour éviter de participer indirectement au financement de projets se compromettant avec l'incitation à la haine ou la propagation de la désinformation.

Inciter les plateformes de financement participatif à faire appel à des sociétés fournissant des évaluations de crédibilité de sites internet ou à obtenir un label reconnu intégrant le souci de ne pas financer des sites toxiques. Cette incitation pourrait prendre la forme d'un droit à un abattement fiscal sur le bénéfice imposable de ces entreprises.

## Ingérences numériques étrangères

### Niveau national

#### 10. PROTÉGER L'INTÉGRITÉ DES PROCESSUS ÉLECTORAUX

Analyser les données de campagnes d'ingérence étrangère ayant visé la conduite de la vie démocratique française afin de mieux anticiper le risque.

Rassembler les données des réseaux sociaux et méta-données récoltées par une diversité de chercheurs et d'institutions et les analyses existantes

Effectuer une analyse approfondie pour mieux comprendre et anticiper les menaces

Mettre en place un dispositif de coopération entre plateformes, institutions et communauté académique pour réagir rapidement aux opérations détectées

#### 11. PERMETTRE LE PARTAGE DES DONNÉES ENTRE ACTEURS DE CONFIANCE

Adapter la plateforme publique, open source Open CTI pour le partage des données de la désinformation entre gouvernement, chercheurs, plateformes, journalistes :

Créer des modules techniques manquants

Engager la réflexion d'une communauté d'acteurs sur la modélisation de la menace

Définir une doctrine d'usage respectueuse de la protection des données personnelles en partenariat avec la CNIL

Encourager la structuration d'une communauté d'utilisateurs qui travaillent sur l'analyse des ingérences numériques, dont les centres de recherche en sciences humaines et sociales

#### 12. CRÉER UN MÉCANISME DE GOUVERNANCE NUMÉRIQUE INTERMINISTÉRIELLE

L'approche sur les risques numériques doit être globale (risques cyber et informationnels associés) car les menaces sont de plus en plus hybrides, et transversales (pan étatique) dans cet espace partagé.

L'enjeu est de développer une culture de la sécurité du numérique qui inclut le risque informationnel et associe tous les acteurs de l'Etat et du gouvernement

L'enjeu est aussi de comprendre les effets involontaires et les interactions entre les différents domaines et de mieux définir les réponses

### **13. SAISIR LE COMITÉ D'ÉTHIQUE DE LA DÉFENSE SUR LA DOCTRINE DE LUTTE D'INFLUENCE INFORMATIQUE DU MINISTÈRE DES ARMÉES.**

Les opérations informatiques d'influence doivent être strictement encadrées d'un point de vue éthique afin d'évaluer au mieux l'équilibre entre les avantages stratégiques et les risques éthiques liés à ces opérations informationnelles. Le comité d'éthique pourra s'intéresser, entre autres, aux audiences ciblées, aux modes opératoires choisis ou encore aux types de discours et narratifs proposés.

#### **Niveau européen**

### **14. CRÉER UN MÉCANISME DE GESTION DE CRISE À L'ÉCHELLE DE L'UNION EUROPÉENNE ET CRÉER DES EXERCICES DE GESTION DE CRISE**

- ◆ pour réagir rapidement à des opérations informationnelles massives
- ◆ pour mieux préparer la gestion de la dimension informationnelle de crises mondiales (sanitaires ou sécuritaires)
- ◆ pour mieux faire face à des menaces informationnelles

#### **Niveau international**

### **15. PROPOSER LA CRÉATION D'UN GROUPE DE TRAVAIL À L'OCDE**

Travailler à la création de normes minimales communes applicables par toutes les plateformes

Construire sur le code de bonnes pratiques existant de l'Union européenne, les bonnes pratiques éprouvées par les plateformes et les résultats de la recherche académique en matière de : community guidelines, fact checking, labellisation, bot takedowns, modération algorithmique, publicité politique, procédures de vérification, transparence et remédiation.

Travailler à une harmonisation au niveau international de la législation sur les obligations des plateformes

## Droit et numérique

### 16. CONSERVER DANS SA RÉDACTION ACTUELLE L'ARTICLE 27 DE LA LOI DU 29 JUILLET 1881, COMME:

- fondement des actions pénales visant la diffusion publique des fausses nouvelles sur les réseaux de communications numériques et les plateformes,
- ainsi que comme définition de référence pour la détermination de ce qui constitue une fausse nouvelle répréhensible dont le retrait ne porterait pas une atteinte injustifiée à l'exercice de la liberté d'expression.

### 17. COMPLÉTER L'ARTICLE 48-1 DE LA LOI DU 29 JUILLET 1881

Ceci afin de permettre à des associations ayant pour objet de lutter contre les fausses nouvelles de nature à susciter une atteinte à l'ordre public de pouvoir exercer les droits reconnus à la partie civile en ce qui concerne l'infraction prévue par l'article 27 de la même loi.

### 18. INTRODUIRE UN NOUVEL ARTICLE DANS LA LOI SUR LA CONFIANCE DANS L'ÉCONOMIE NUMÉRIQUE (LCEN)

Introduire un nouvel article engageant la responsabilité civile du diffuseur de mauvaise foi d'une fausse nouvelle préjudiciable, lequel pourrait être rédigé comme suit :

*« Toute diffusion par voie numérique d'une nouvelle que l'on sait être inexacte et qui porte préjudice à autrui engage la responsabilité civile de celui qui la commet ainsi que de toute personne qui la rediffuse en connaissance de cause.*

*Pour fixer les dommages et intérêts, la juridiction prend en considération distinctement :*

- 1. Les conséquences économiques négatives de la diffusion ;*
- 2. Le préjudice moral causé par cette dernière ;*
- 3. L'étendue et la vitesse de propagation de la diffusion ;*
- 4. Et le niveau d'audience et de popularité numérique de son auteur».*



**19. COMPLÉTER L'ARTICLE 17/2 DE LA LOI DU 30 SEPTEMBRE 1986, POUR PRÉVOIR :**

- d'une part, que l'ARCOM peut être saisie par toute personne ayant rencontré une difficulté pour obtenir l'intervention et la coopération d'une plateforme afin de prévenir ou stopper la diffusion massive d'un contenu susceptible de véhiculer une fausse nouvelle pouvant troubler l'ordre public ;
- et d'autre part, lui permettre d'enjoindre à la plateforme concernée – une fois avisée par l'ARCOM - de lui fournir à bref délai un état des mesures prises par ses soins dans le cas considéré et de coopérer avec l'ARCOM pour rechercher et mettre en place les mesures préventives ou curatives appropriées au traitement d'un tel cas.

**20. EXIGER DES PLATEFORMES QU'ELLES OUVRENT AUX CHERCHEURS L'ACCÈS À LEURS DONNÉES**

Veiller à ce que, dans la version finale du Digital Services Act (DSA), les modalités entourant l'obligation pour les plateformes de donner accès à leurs données (DSA article 31<sup>279</sup>) forment un cadre optimal permettant aux chercheurs de procéder à des recherches contribuant à l'identification et à la compréhension des risques systémiques (dont fait partie la désinformation - DSA article 26) dans les meilleures conditions.

**21. INTRODUIRE DANS LE DIGITAL SERVICES ACT UNE DISPOSITION SUR LES FAUSSES NOUVELLES**

Introduire explicitement dans le DSA une disposition qui reconnaisse que les fausses nouvelles susceptibles de troubler l'ordre public constituent des contenus répréhensibles devant être pris en compte par le mécanisme de modération des contenus imposés aux plateformes.

**22. METTRE EN PLACE UN ORGANE D'EXPERTISE EXTÉRIEUR INDÉPENDANT**

Mettre en place un organe indépendant avec lequel les plateformes pourraient signer une convention leur permettant, face à la demande de retrait d'un contenu argué d'être une fausse nouvelle répréhensible, de soumettre le cas à ces experts extérieurs dont elles s'engageraient à suivre l'avis.

Dans une version plus poussée de la recommandation précédente, les conditions générales de la plateforme pourraient prévoir que le plaignant serait réputé accepter contractuellement le principe de cette expertise extérieure et s'engager à ne pas engager d'action contentieuse avant d'en connaître le résultat.

<sup>279</sup> Article 31 de la proposition de règlement du Parlement européen et du Conseil relatif à un marché intérieur des services numériques (Législation sur les services numériques – Digital Services Act – DSA) et modifiant la directive 2000/31/CE, COM(2020) 825 final.

### **23. CRÉER UN RÉGIME DE CO-RÉGULATION ENTRE PLATEFORMES, RÉGULATEURS, ET SOCIÉTÉ CIVILE DANS LE CADRE DU DIGITAL SERVICES ACT**

Mettre en place un mécanisme de coopération exigeante avec les plateformes pour concevoir, mettre en œuvre et évaluer les mesures appliquées par les plateformes pour modérer les contenus en préservant la liberté d'expression et les droits humains

- ♦ les régulateurs établissent un cadre général avec les grands principes
- ♦ les corégulateurs traduisent ces principes en normes applicables
- ♦ les plateformes mettent en œuvre les normes dans le cadre du respect de leurs obligations fixées par le DSA
- ♦ les régulateurs veillent à la mise en œuvre des normes et évaluent l'efficacité des mesures prises par les plateformes

## **Esprit critique et EMI**

### **24. CRÉER UNE CELLULE INTERMINISTÉRIELLE DÉDIÉE AU DÉVELOPPEMENT DE L'ESPRIT CRITIQUE ET D'UNE EMI TOUT PUBLIC**

Créer une cellule interministérielle dédiée au développement de l'esprit critique et d'une EMI tout public, associant les principaux protagonistes (ministères, associations, médias, bibliothèques...); une délégation sous l'égide de Matignon chargée de se concerter, de mutualiser, de valoriser les ressources et de mandater un organisme ou de créer une structure pour mettre en place l'évaluation du matériel pédagogique et des dispositifs de formation avec des protocoles scientifiques standards.

### **25. CARTOGRAPHIER LES DIFFICULTÉS COGNITIVES DES ÉLÈVES**

Solliciter l'expérience des enseignants pour établir une cartographie des difficultés cognitives les plus fréquemment rencontrées chez les élèves, en vue d'engager une réflexion sur la pédagogie de la métacognition.

### **26. FAIRE DU DÉVELOPPEMENT DE L'ESPRIT CRITIQUE ET DE L'EMI UNE GRANDE CAUSE NATIONALE.**

Augmenter leur visibilité par la diffusion de messages d'intérêt généraux dans les médias.

### **27. SYSTÉMATISER LA FORMATION À L'ESPRIT CRITIQUE ET À L'EMI EN MILIEU SCOLAIRE**

Systematiser la formation pour les élèves dès l'école primaire et jusqu'après le secondaire et pour les enseignants en formation initiale et continue et renforcer de manière significative le réseau de référents et coordinateurs académiques dans ces domaines.

### **28. SENSIBILISER LES AUTORITÉS SCOLAIRES AUX ENJEUX DE L'EMI**

Sensibiliser les chefs d'établissements scolaires, les inspecteurs de l'Éducation Nationale, les recteurs aux enjeux de l'EMI et de la formation à la pensée critique, ainsi que les élus locaux, les responsables de ressources humaines des collectivités locales et les responsables de bibliothèques.

### **29. DÉVELOPPER LA FORMATION À L'ESPRIT CRITIQUE ET À L'EMI DANS LA SOCIÉTÉ CIVILE**

Il est important de créer un continuum entre le temps scolaire, l'université, le monde culturel et le monde du travail. La formation à l'esprit critique et l'EMI doit donc être systématisée aussi bien dans des projets éducatifs de territoires et des cités éducatives, que dans des dispositifs d'insertion, des jeunes volontaires en service civique jusqu'aux retraités et à la formation continue.

### **30 .SAISIR LE COMITÉ NATIONAL PILOTE D'ÉTHIQUE DU NUMÉRIQUE DE LA QUESTION DES MONDES NUMÉRIQUES ET DE LA RÉALITÉ VIRTUELLE ET AUGMENTÉE.**

L'immersion croissante des utilisateurs dans des mondes numériques où la distinction entre le réel et le virtuel s'efface progressivement peut entraîner des risques éthiques. Le projet de Métavers annoncé par Meta (ex-Facebook) ou le projet de Metaverse Seoul pourraient accélérer ce phénomène. Une première réflexion au niveau national pourrait conduire à la constitution d'un groupe international multi-acteurs pour prévoir un encadrement éthique au développement de ces environnements numériques.

## Listes des personnes et institutions sollicitées

### **Andler Daniel**

Directeur du projet TESaCo, Académie des sciences morales et politiques, Institut de France

### **Angaud Bernard**

Bureau du Conseil de déontologie journalistique et de médiation

### **Arata Fabienne**

Country Manager et Senior LTS Director chez LinkedIn

### **Astolfi Charles-Pierre**

Conseiller Régulations et communs numériques Secrétariat d'État chargé de la Transition numérique et des Communications électroniques

### **Audinet Maxime**

Chercheur à Institut de Recherche Stratégique de l'École Militaire (IRSEM)

### **Avia Laetitia**

Députée de Paris (12<sup>e</sup> - 20<sup>e</sup> arrondissements)

### **Bail Chris**

Professeur à l'Université Duke et directeur du Polarization Lab

### **Barbet Serge**

Directeur délégué du CLEMI

### **Battesti Anton'Maria**

Responsable des Affaires publiques Facebook

### **Beaudouin Pierre-Yves**

Président du Conseil d'administration de Wikimedia France

### **Belin Celia**

Docteure en science politique de l'université Panthéon-Assas, actuellement chercheuse invitée au Centre sur les États-Unis et l'Europe au sein de la Brookings Institution (Washington DC)

### **Benabou Valérie Laure**

Professeure de droit privé à l'Université Paris Saclay

### **Benard Yohann**

Directeur de la Stratégie d'Amazon

### **Benoualid Shani**

Conseillère pour le numérique et les réseaux sociaux DILCRAH

### **Berkouk Hannah**

Directrice générale d>HelloAsso

### **Bienaimé Besse Carole**

Membre du Collège du Conseil supérieur de l'audiovisuel

### **Blanchot Guillaume**

Directeur Général du Conseil supérieur de l'audiovisuel

### **Blanquer Jean-Michel**

Ministre de l'Éducation Nationale, de la Jeunesse et des Sports

**Borry-Estrade Elisa**

Chargé des affaires publiques Facebook

**Borst Grégoire**

Professeur de psychologie à l'Université de Paris

**Bothorel Éric**

Député de la 5<sup>e</sup> circonscription des Côtes-d'Armor

**Bouillon Stéphane**

Préfet hors classe, Secrétaire général de la défense et de la sécurité nationale (SGDSN)

**Bourdet Julienne**

Chargée de formation à Médiat, Rhône-Alpes

**Cabannes Laurent**

Professeure de technologie à l'académie de Créteil

**Cardon Dominique**

Directeur du médialab de Sciences Po

**Caroti Denis**

Chargé de mission Esprit critique au rectorat d'Aix-Marseille

**Cathelineau Yolaine**

Doctorante à GEODE, Université Paris 8

**Cattan Jean**

Secrétaire Général CNNum

**Charon Paul**

Directeur du domaine « Renseignement anticipation et menaces hybrides, » Institut de Recherche Stratégique de l'Ecole Militaire (IRSEM)

**Chérel Ronan**

Professeur d'Histoire-Géographie au collège Rosa Parks (Rennes)

**Chirouze Aline**

Professeure des écoles intervenant en milieu carcéral

**Claerr Thierry**

Chef du bureau de la Lecture publique

**CNews**

N'a pas répondu à nos sollicitations

**Cointet Jean-Philippe**

Chercheur au médialab de Sciences Po

**Colrat Philippine**

Responsable des Affaires publiques d'Amazon

**Conférence des Ecoles de Journalisme (CEJ)****Corbin Noël**

Délégué général à la transmission, aux territoires et à la démocratie culturelle (DG2TDC)

**d'Aubert François**

Président de l'ARPP

**Dagnaud Monique**

Directrice de recherche CNRS-EHESS

**Duguin Stéphane**

Chief Executive Officer, CyberPeace Institute

**Daviet Emmanuelle**

Médiatrice à Radio France

**de la Chapelle Bertrand**

Directeur et co-fondateur du Internet et Jurisdiction Policy Network

**Deloire Christophe**

Secrétaire Général, Journaliste à Reporters sans frontières

**Delouée Sylvain**

Chercheur en psychologie sociale à l'Université Rennes II

**Dey Aurélie**

Lieutenant-Colonel, commandant de la division de lutte contre les crimes et délits haineux (DLCH) à l'Office Central de Lutte contre les Crimes contre l'Humanité, les génocides et les crimes de guerre (OCLCH)

**di Palma Cyril**

Délégué général de l'association « Génération Numérique »

**Dieguez Sébastien**

Chercheur en neurosciences au Laboratoire de sciences cognitives et neurologiques de l'Université de Fribourg (Suisse)

**Doucet-Bon Pascal**

Directeur délégué de l'information en charge de la stratégie chez France Télévisions

**Dufour Mathias**

Président de #Leplusimportant

**Durand Pascale**

Directrice de la vie fédérale, FCPE (Fédération des Conseils de Parents d'élèves)

**Durand-Viel Laure**

Déléguée à la régulation de plateformes numériques au Ministère de la Culture

**Élizéon Sophie**

Préfète, Déléguée Interministérielle DILCRAH

**Eveno Patrick**

Bureau du Conseil de déontologie journalistique et de médiation

**Ferriol Gabriel**

Chef du Service de vigilance et de protection contre les ingérences numériques étrangères (Viginum)

**Fogiel Marc-Olivier**

Directeur Général BFMTV

**Forestier Florian**

Philosophe, #Leplusimportant

**Forteza Paula**

Députée de la 2<sup>e</sup> circonscription des Français établis hors de France, à l'origine du rapport « Liberté, Égalité... Vérité »

**Franceschini Laurence**

Présidente de la CPPAP

**François Camille**

Chief Innovation Officer GRAPHIKA

**Frau-Meigs Divina**

Professeure en sciences de l'information et la communication à l'Université Paris III

**Freyssinet Éric**

Général de brigade commandant en second de la gendarmerie dans le cyberspace. Docteur en informatique membre associé du LORIA

**Froissard Laureline**

Directrice des affaires publiques et juridiques à l'Union des Marques

**Garandeau Éric**

Directeur des Relations Institutionnelles et des Affaires Publiques France chez Tik Tok

**Garnier Marie Caroline**

Directrice Agence CORPCOM

**Gautellier Christian**

Responsable du conseil d'orientation du Pôle « Médias, Numérique, Éducation critique » à l'association nationale des Ceméa

**Gayraud Jean-François**

Conseiller à la Coordination Nationale du Renseignement et de la Lutte contre le Terrorisme Centre National de Contre-Terrorisme

**Geffray Edouard**

Directeur général à la Direction générale de l'enseignement scolaire (DGESCO)

**Gérard Colin**

Doctorant à GEODE, Université Paris 8

**Gérard Olivier**

Coordonnateur Médias-Usages du Numérique à l'UNAF (Union Nationale des associations familiales)

**Gery Aude**

Post-doctorante à GEODE, Université Paris 8

**Giret Vincent**

Directeur de l'information et des sports Radio France

**Gourdin Jean-Baptiste**

Directeur général des médias et des industries culturelles au Ministère de la Culture

**Grosset Kathleen**

Bureau du Conseil de déontologie journalistique et de médiation

**Grumbach Stéphane**

Directeur de recherche à l'INRIA, responsable de l'équipe Données de l'Internet au Coeur de l'Économie (Dice) et directeur de l'Institut Rhône-Alpin des Systèmes Complexes

**Guiroy Thibault**

Responsable des relations institutionnelles chez GOOGLE

**Haugen Frances**

Ingénieure, scientifique des données et gestionnaire de projets, ancienne employée de Facebook

**Hecketsweiler Jean-Philippe**

Président de la Fondation Descartes

**Herblin-Stoop Audrey**

Directrice des affaires publiques Twitter

**Huchon Thomas**

Journaliste à Spicée et à LCI

**Innes Martin**

Professeur, directeur de l'Institut de recherche sur la criminalité et la sécurité à Cardiff University

**Jacquier Sarah**

Chargée de mission à la sous-direction des affaires juridiques au Ministère de la Culture

**Jean Aurélie**

Docteure en mécanique numérique

**Jeangene Vilmer Jean-Baptiste**

Directeur de l'Institut de Recherche Stratégique de l'Ecole Militaire (IRSEM)

**Jézéquel Gwénaël**

Conseiller pour les relations institutionnelles et la communication à la SGDSN

**Jolion Jean-Michel**

Conseiller auprès du cabinet de la Ministre en charge de l'Enseignement supérieur, de la Recherche et de l'Innovation

**Jounot Olivier**

Responsable du département RSE, groupe AFNOR

**Khemis Sarah**

Responsable Senior des Relations Institutionnelles et des Affaires Publiques France chez TIK TOK

**Klein Olivier**

Professeur de psychologie à l'ULB

**Koutchouk Alexandre**

Sous-directeur de la presse écrite et des métiers de l'information au Ministère de la Culture

**Labbé Chine**

Rédactrice en chef et vice-présidente en charge des partenariats Europe de News-Guard

**Laboulais François**

Chargé de mission « Éducation aux médias » pour l'association nationale des CEMEA

**Laffont Sandra**

Journaliste, présidente d'« Entre les Lignes »

**Larrieu Mathilde**

Coordinatrice missions « Éducation aux médias » pour l'ENSSIB (formation des bibliothécaires)

**Le Monde**

N'a pas répondu à nos sollicitations

**Le Roux Yann**

Directeur Europe du Sud à Integral Ad Science (IAS)

**Lee Bouygues Helen**

Fondatrice de la Fondation reboot ; Présidente de Conforama

**Lesage François**

Directeur de la communication Twitter

**Limonier Kevin**

Maître de conférences à GEODE, Université Paris 8

**Loutrel Benoît**

Membre du Collège de Conseil supérieur de l'audiovisuel

**Machet Julien**

Membre du Groupe de Ressources transdisciplinaire Esprit Critique du CSEN (Conseil Scientifique de l'Éducation Nationale)

**Macron Brigitte**

Présidente de la Fondation Hôpitaux de Paris-Hôpitaux de France

**Magnin Olivier**

Directeur du Pôle Éducation à l'image, aux médias et à l'information de la Ligue de l'enseignement

**Maistre Roch-Olivier**

Président du Conseil supérieur de l'audiovisuel

**Melford Clare**

Co-Fondatrice et Directrice exécutive à The Global Disinformation Index

**Mercadal Delasalles Françoise**

Co-Présidente du CNNum

**Mercier Hugo**

Chargé de Recherche au CNRS

**Mercier Arnaud**

Professeur en sciences de l'information et la communication à l'Université Paris II



**Missoffe Sébastien**

Directeur Général de GOOGLE France

**Morel François**

Direction territoriale Auvergne Rhône Alpes  
du Réseau Canopé

**Motte Stanislas**

Directeur et co-fondateur de STORYZY

**Moukheiber Albert**

Docteur en neurosciences cognitives

**Nathan Michaël**

Service d'information du Gouvernement  
(SIG)

**Ndior Valère**

Professeur de droit public à l'université de  
Bretagne Occidentale

**Nguyen Huang Lé**

Mathématicien enseignant à polytechnique  
Lausanne et animateur de la chaîne Science  
4all

**Nicolas Laurent**

Directeur d'Implicit

**Novel Anne-Sophie**

Collectif « Informer n'est pas un délit »

**Novel Catherine**

Présidente de l'Association des professeurs  
documentalistes de l'Éducation Nationale  
(APDEN)

**Nuñez Laurent**

Préfet, Coordinateur National du Rensei-  
gnement et de la Lutte contre le Terrorisme,  
ancien secrétaire d'Etat auprès du Ministre  
de l'Intérieur

**Ouvrard Béatrice**

Manager chargée des Affaires Publiques  
chez Facebook

**Ohayon Esther**

Group Manager, Corporate Communications  
chez LinkedIn

**Orphelin Matthieu**

Député de la 1<sup>ère</sup> circonscription  
de Maine-et-Loire, à l'origine du rapport  
« Liberté, Égalité... Vérité »

**Pasquinelli Elena**

Chercheuse en philosophie à l'institut  
Jean Nicod

**Petit Laurent**

Chargé de mission numérique à l'Institut  
Universitaire de formation du professorat  
et de l'éducation (INSPE) de Paris

**Picquet Gautier**

Président de Publicis Media, président de  
l'Union des Entreprises de conseil et d'Achat  
Media (UDECAM)

**Pigalle Céline**

Directrice de rédaction chez BFMTV

**Pospisil Marek**

Senior Lead Public Policy et Economic Graph  
chez LinkedIn

**Quattrociocchi Walter**

Professeur en sciences informatiques  
à l'Université de Rome

**Renard Yves**

Directeur de l'École Supérieure de Journa-  
lisme de Lille

**Robin Valérie**

Chargée de mission « éducation artistique et culturelle - EMI » à la BPI-Centre Pompidou

**Rolle Pierre-Louis**

Directeur des programmes « société numérique », « nouveaux lieux, nouveaux liens, » à l'ANCT (Agence Nationale de la Cohésion des Territoires)

**Ruquier Pierre-Albert**

Directeur marketing et co-fondateur de Storyzy

**Schapiro Jacob**

Professeur de politique et affaires internationales et Directeur du projet études empiriques des conflits à Princeton University

**Schiffrin Anya**

Directrice de la spécialisation Technologie, médias et communications à la School of International and Public Affairs de Columbia University

**Schmidt Philipp**

Directeur exécutif de Prisma Media Solutions

**Schwartz Arnaud**

Directeur de l'Institut du journalisme de Bordeaux Aquitaine (IJBA)

**Séjourné Stéphane**

Président de Renew Europe

**Servan-Schreiber Emile**

Docteur en psychologie cognitive, fondateur d'Hypermind

**Signoux Martin**

Chargé des affaires publiques chez Facebook

**Simon-Nahum Perrine**

Directrice de recherches au CNRS et professeure attachée au département de philosophie de l'ENS

**Simonet Vincent**

Engineering Director chez GOOGLE

**Sleeping Giants France**

Collectif de lutte contre le financement des discours haineux

**Taguieff Pierre-André**

Politologue, historien, directeur de recherche au CNRS

**Théobalt Jean-Christophe**

Chargé de mission Médiation numérique et EMI

**Tisseyre Didier**

Général, commandant la cyberdéfense des Armées

**Vachet Carole**

Directrice de cabinet Secrétariat d'État chargé de la Transition numérique et des Communications électroniques

**van Prooijen Jan-Willem**

Chercheur en psychologie sociale à Maastricht University

**Verdier Henri**

Ambassadeur pour le Numérique

**Vincent Emmanuel**

Responsable d'édition multisupport aux Éditions de l'EHESS

**Wagner-Egger Pascal**

Chercheur en psychologie sociale à l'Université de Fribourg

**Watrin Laurent**

Adjoint au Maire délégué à l'innovation des politiques publiques et au numérique, ancien journaliste du service public audiovisuel à la Mairie de Nancy

**We Report**

Collectif de journalistes indépendants

**Yesilaltay Sacha**

Doctorant en sciences cognitives au CNRS

**Zuckerman Ethan**

Professeur associé, directeur de l'initiative sur l'infrastructure publique numérique à l'UMass Amherst

Les membres de la commission tiennent également à adresser leurs remerciements à tous ceux qui ont participé au bon déroulement de cette commission :

Margot Godefroi, Izabela Luniak, Samuel Leenhardt, Guilhem Marotte, Sasha Morinière, Astrid Roucher

