

BENCHMARK DES OUTILS DE LUTTE CONTRE LA DÉSINFORMATION

Commission « Manipulations de l'Information »



Association des auditeurs en
intelligence économique de
l'Institut des Hautes Etudes
de Défense Nationale

RÉALISATION

MEMBRES DE LA COMMISSION :

Supervision :

Thibault RENARD, Expert Intelligence Economique, senior advisor Cybercercle

Analyse et rédaction :

Boubacar DIALLO, Analyste IE, Académie de l'Intelligence Economique

Contributions :

Salima LAMERRE, Responsable commerciale, Spécialiste Intelligence Économique

Delphine PICAUVET, Responsable de projets recherche & innovation

RELECTURE ET CONTRIBUTIONS EXTÉRIEURES :

Laurent CORDONIER, Docteur en sciences sociales, Directeur de la recherche de la Fondation Descartes

Institution partenaire dans la collecte de données (Ecole d'Ingénieur EPITA) :

Eloi HONNET, étudiant ingénieur

Marie CHAUVIRE, étudiante ingénieur

L'ASSOCIATION DES AUDITEURS EN INTELLIGENCE ECONOMIQUE DE L'IHEDN (AAIE-IHEDN)

Fondée en 1999, l'AAIE-IHEDN contribue aux missions de l'Institut des Hautes Etudes de Défense Nationale (IHEDN).

Elle rassemble près de deux cents praticiens, cadres et dirigeants d'entreprises, fonctionnaires de l'Etat et relais d'opinion représentant un éventail très large de métiers et de secteurs.

La vocation de l'AAIE-IHEDN est de :

- maintenir et développer les liens entre les auditeurs ayant participé aux formations en Intelligence Economique et Stratégique
- promouvoir l'intelligence économique, sociétale et stratégique dans tous les secteurs d'activités de la société et en diffuser une véritable culture au plan national
- apporter son concours à l'IHEDN pour l'accomplissement de sa tâche dans le domaine de l'Intelligence Economique, de la Défense et de la Sécurité Nationale
- apporter des méthodes et outils d'aide à la décision aux instances gouvernementales, notamment en matière de sensibilisation et de formation
- contribuer à la réflexion sur les enjeux liés à la défense, la sécurité et la souveraineté nationale.

Au travers de son réseau professionnel, l'AAIE-IHEDN développe des pôles de compétence et d'expertise dans son domaine.

Commissions, colloques et manifestations diverses contribuent à enrichir la réflexion nationale et européenne ainsi qu'à formaliser les démarches et méthodes de mise en œuvre de l'intelligence économique, sociétale et stratégique.

<https://www.ie-ihedn.org/>

SOMMAIRE

I. Contexte et objectifs	5
II. Désinformation : de quoi parle-t-on ?	7
III. Tendances amplificatrices de la désinformation	8
IV. Typologie des outils de lutte contre la désinformation	11
V. Sources bibliographiques	18

I. CONTEXTE ET OBJECTIFS

Depuis plus d'une décennie, l'information et plus généralement la donnée se diffusent de façon exponentielle grâce à l'interconnectivité numérique. Cet accroissement est rendu possible à travers, d'une part, le développement et la popularisation des outils technologiques (notamment les téléphones smartphones et autres objets connectés) et d'autre part, à travers les larges possibilités de diffusion et de partage d'informations sur les réseaux sociaux (Tik-tok, WhatsApp, Instagram, Snapchat, Twitch, Tweeter, Telegram, YouTube, etc.), les plateformes numériques, l'intelligence artificielle, sans oublier certains médias traditionnels. Ce phénomène constitue une réelle avancée démocratique en matière de liberté d'expression, d'accès « gratuit » et à priori facile à l'information, de partages et d'interactions entre toutes les couches sociales sur internet.

Cependant, cette propagation techno-numérique engendre également une menace grandissante pour les sociétés démocratiques, du fait de la diffusion intentionnelle de contenus nuisibles non contrôlés tels que les *fake news* (fausses informations), les discours haineux, les photos et vidéos retouchées voire fausses, les propagandes informationnelles, les théories du complot, etc. Aujourd'hui, du fait de la surexposition directe de la population aux contenus de toutes natures (vraies et/ou fausses) sur les réseaux sociaux, la désinformation devient une préoccupation centrale dans les grandes démocraties. Il s'ajoute aussi du hors contexte, de l'incomplet, de l'ultra-partial, etc, qui peuvent composer de la mésinformation aux conséquences potentiellement néfastes. C'est le cas de certaines rumeurs ou d'informations erronées en santé.

La désinformation, ainsi que la mésinformation renforcent potentiellement tous les grands phénomènes déstabilisateurs du 21^{ème} siècle (covid-19, dérèglement climatique, etc.) car elle fragilise la possibilité de tenir un débat public serein et rationnel sur des enjeux sociétaux. Dans son dernier rapport en janvier 2024, le Forum économique mondial de Davos alertait ainsi sur « *le risque majeur de désinformation qui affecte les sociétés mondiales, dans un contexte de crise* ».

En décembre 2020, la Commission européenne a publié dans son rapport intitulé « *plan d'actions pour démocratie européenne* »¹ une série d'actions visant à renforcer la sécurisation de ses États membres. Parmi elles, la lutte contre la manipulation d'informations et les ingérences étrangères en imposant des sanctions financières aux auteurs.

¹ Plan d'action pour la démocratie européenne : renforcer les démocraties de l'UE : https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan_fr

Cela a abouti notamment à un groupe de travail *Radicalisation Awareness Network* (RAN YOUNG Platform)². Il s'agit d'une plateforme qui réunit des praticiens de première ligne notamment des jeunes de toute l'Europe pour partager des connaissances, des bonnes pratiques et des expériences dans le domaine de la protection de la vie privée, de la lutte contre l'extrémisme violent, mais aussi contre la manipulation en ligne. En collaboration avec la Commission européenne, ce groupe de travail a tenu une réunion avec 17 jeunes (étant probablement plus présents sur les réseaux sociaux) pour discuter des conséquences de la désinformation en ligne et réfléchir sur des solutions numériques pour lutter contre ce fléau.

En octobre 2022, un nouveau règlement européen appelé *Digital Service Act*³ applicable à toute les plateformes en ligne et moteurs de recherche est entré en vigueur dans le but de renforcer la transparence des acteurs numériques, de protéger les droits de consommateurs européens contre toute diffusion de contenu illégal (haines, manipulations, désinformations, contrefaçons, attaques racistes, images pédopornographiques, ...) en ligne.

En outre, selon la Fondation Descartes en 2023⁴ des études ont montré que ce sont les aînés qui relaient le plus de *fakes news* sur Internet. Ils sont aussi particulièrement touchés par les sites pièges à cliques. La même étude affirme que les femmes sont davantage ciblées par la désinformation en santé.

À l'échelle française, le 22 décembre 2018, la loi contre la manipulation d'information⁵ appelée « *loi infox* » ou « *loi fake news* » a vu le jour. Son principal objectif est de protéger la démocratie contre les formes intentionnelles de diffusion de fausses informations sur internet. C'est notamment dans le cadre de cette loi que les réseaux sociaux et les plateformes numériques ont une obligation de transparence⁶ en signalant les contenus sponsorisés et leurs auteurs.

Afin de renforcer ce dispositif réglementaire, le Secrétariat Général de la Défense et de la Sécurité Nationale (SGDSN) a créé, en septembre 2021, une agence de lutte contre les manipulations d'informations. Selon Stéphane Bouillon, Secrétaire Général du SGDSN, « *avec cette agence, il ne s'agit pas de corriger ou rétablir la vérité, mais d'arriver à détecter les attaques quand elles viennent de l'étranger.* »⁷.

² Inciter les jeunes à lutter contre la désinformation en ligne – Document de consultation : https://home-affairs.ec.europa.eu/system/files/2021-09/ran_young_galvanising_youth_07-08_122020_fr.pdf / https://home-affairs.ec.europa.eu/networks/radicalisation-awareness-network-ran/ran-young-platform_en

³ Le règlement européen sur les services numériques (DSA) vise une responsabilisation des plateformes : <https://www.vie-publique.fr/eclairage/285115-dsa-le-reglement-sur-les-services-numeriques-ou-digital-services-act>

⁴ INFORMATION ET SANTÉ : <https://www.fondationdescartes.org/2023/10/information-et-sante/>

⁵ LOI n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information : <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000037847559>

⁶ <https://www.gouvernement.fr/action/contre-la-manipulation-de-l-information>

⁷ Fakes news : la France se dote d'une agence de surveillance des réseaux : <https://www.lesechos.fr/politique-societe/gouvernement/fake-news-la-france-se-dote-dune-agence-de-surveillance-des-reseaux-1320268>

Tous ces dispositifs réglementaires visent à renforcer la lutte contre la désinformation. En complément de ces dispositifs, il existe également d'autres outils techniques (plateformes en ligne, logiciels, cours en ligne, etc.) de lutte contre la manipulation en ligne. Ces derniers peuvent être classés selon leur mode de fonctionnement, la cible, et leur efficacité. L'intérêt de ce travail de comparaison est d'identifier les outils numériques qui permettent de combattre la désinformation.

L'objectif de ce rapport est, dans un premier temps, de mettre en évidence les tendances amplificatrices de la désinformation en ligne, puis de dresser un benchmark analytique des outils existants de lutte contre la désinformation.

II. DÉSINFORMATION : DE QUOI PARLE-T-ON ?

Pour parler de benchmark des outils, il faut, au préalable, définir les concepts utilisés ; à savoir distinguer la manipulation de l'information, la désinformation, la mésinformation et la malinformation ainsi que leur fonctionnement. Le rapport « *Guide sur la manipulation de l'information : Un conseil pratique pour les élections et au-delà* » réalisé par l'International Republican Institute (IRI), le National Democratic Institute (NDI) et la Stanford Internet Observatory en 2021⁸, propose d'éclairantes définitions.

- **La manipulation d'information** : ensemble de tactiques impliquant la collecte et la diffusion d'informations afin d'influencer ou de perturber la prise de décision démocratique. Autrement dit, elle fait usage de contenus à des fins de manipulation ou de déstabilisation.
- **La désinformation** : création, diffusion et/ou amplification délibérée d'informations fausses, inexactes ou trompeuses avec l'intention de tromper.
- **La mésinformation** : information fautive, inexacte ou trompeuse diffusée, indépendamment de l'intention de tromper. C'est-à-dire que la personne qui diffuse ce type d'information ne sait pas qu'elle fautive.
- **La malinformation**⁹ : utilisation d'informations véridiques ou factuelles à des fins de persuasion. Il peut s'agir, par exemple, de contenus publiés dans le cadre d'une opération de piratage et

⁸ Guide sur la manipulation de l'information : Un conseil pratique pour les élections et au-delà : https://www.ndi.org/sites/default/files/infomanip_playbook_french_final.pdf

⁹ Tools of disinformation : Inauthentic content-CISA Cybersecurity and Infrastructure Security Agency (CISA)

fuitage (hack-and-leak), où des messages privés sont partagés publiquement dans le but de nuire à un adversaire.

- **L'ingérence informationnelle**¹⁰ : opération structurée et coordonnée par des acteurs étrangers, destinée à propager via les plateformes numériques des contenus manifestement trompeurs et hostiles, cherchant à porter atteinte aux intérêts fondamentaux de la Nation.

Ces définitions tirées du rapport permettent de vulgariser les concepts et expressions utilisés pour expliquer les différentes situations de manipulation de l'information.

III. TENDANCES AMPLIFICATRICES DE LA DÉSINFORMATION

L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE

Du fait de sa capacité à créer des images et contenus réalistes facilement détournables à diverses fins, le développement de l'intelligence artificielle générative provoque à la fois des opportunités économiques et des risques sociétaux en matière de désinformation en ligne. En effet, il est aujourd'hui quasiment impossible pour un humain de différencier le vrai du faux tant les algorithmes d'IA générative sont capables de générer des images très crédibles et acceptables aux yeux de tous. Des exemples très médiatisés l'illustrent : la fausse [photo virale du Pape François générée avec une IA et publiée](#) sur twitter en mars 2023 dont la portée était humoristique et qui a rapidement fait l'objet de différentes interprétations. Le premier enseignement à tirer de ce phénomène est le niveau de réalisme des images générées par l'IA. Elle peut engendrer une extrême viralité sur les réseaux sociaux (Tik-Tok, X, Facebook, Instagram, etc...), et devient, à ce titre, un terrain propice à la propagation de la désinformation.

d'autres cas concrets tels que la [supposée arrestation de Donald Trump](#) suite à des affaires judiciaires en cours ou les [fausses images d'Emmanuel Macron](#) en pleine manifestation face à la police, en sont des cas très frappants notamment sur la vitesse de propagation de la désinformation. Le cas de Donald Trump est un cas d'école intéressant du fait du croisement deux composantes : une vraie information (ses problèmes judiciaires en cours) combiné à une fausse information (son arrestation). Or, les biais

¹⁰ Les lumières à l'ère du numérique (2022). Rapport de la commission.
<https://www.vie-publique.fr/files/rapport/pdf/283201.pdf>

cognitifs ont de fortes répercussions sur nos croyances, nos décisions et notre capacité à maintenir une lucidité afin de juger les données reçues.

Finalement, l'automatisation de la création et de diffusion de contenus va contribuer inéluctablement à intensifier la désinformation dans les médias. L'IA générative, du fait de sa capacité à fabriquer des photos et vidéos manipulées très réalistes, en impliquant, déformant ou détournant les individus, les contextes, les événements pour donner une fausse facette de la réalité contribue à accroître la désinformation dans les flux auxquels nous sommes confrontés. L'avantage « concurrentiel » de certains médias traditionnels par rapport aux réseaux sociaux réside dans leur travail de vérification en amont de leurs informations. Dans ce contexte, l'IA rend ces médias encore plus importants.

PERSONNALISATION DE LA DÉSINFORMATION

Les algorithmes d'intelligence artificielle intégrés dans les médias sociaux peuvent personnaliser les contenus en fonction des caractéristiques démographiques, des préférences politiques et des comportements en ligne de chaque internaute. Cette personnalisation accessible à tous augmente l'efficacité de la désinformation en ciblant spécifiquement les vulnérabilités et les biais de chaque utilisateur (Par exemple, des publicités politiques diffusées exclusivement à des électeurs potentiellement influençables). L'affaire Cambridge Analytica de 2016¹¹ est un excellent exemple en ce qui concerne l'usage de la manipulation en ligne pendant la campagne électorale de Donald Trump. Les messages sur l'immigration, la covid-19, l'écologie ou la guerre en Ukraine, conçus pour faire résonner leurs inquiétudes spécifiques, exploitent les failles individuelles, transformant l'algorithme en un outil de manipulation de masse à des fins politiques.

Dans un tel contexte, l'internaute se trouve dans une bulle « informationnelle » qui enferme dans ses préjugés. Ce phénomène peut s'amplifier avec l'automatisation de la désinformation. En effet, à l'aide d'outils automatisés comme les bots sur les réseaux sociaux, la diffusion à grande échelle de fausses informations ou de contenus trompeurs est devenu un jeu d'enfant. Et comme disait Éric Sadin, philosophe français spécialiste du numérique, « *la bulle devient votre réalité quotidienne, un écho constant de vos propres idées, une chambre où la dissidence est étrangère* ». La désinformation automatisée, en exploitant cette quête algorithmique de la familiarité, empêche de prendre le temps d'analyser la qualité des informations auxquelles nous sommes confrontés et devient à ce titre une menace pour la vérité, la cohésion sociale et la confiance démocratique.

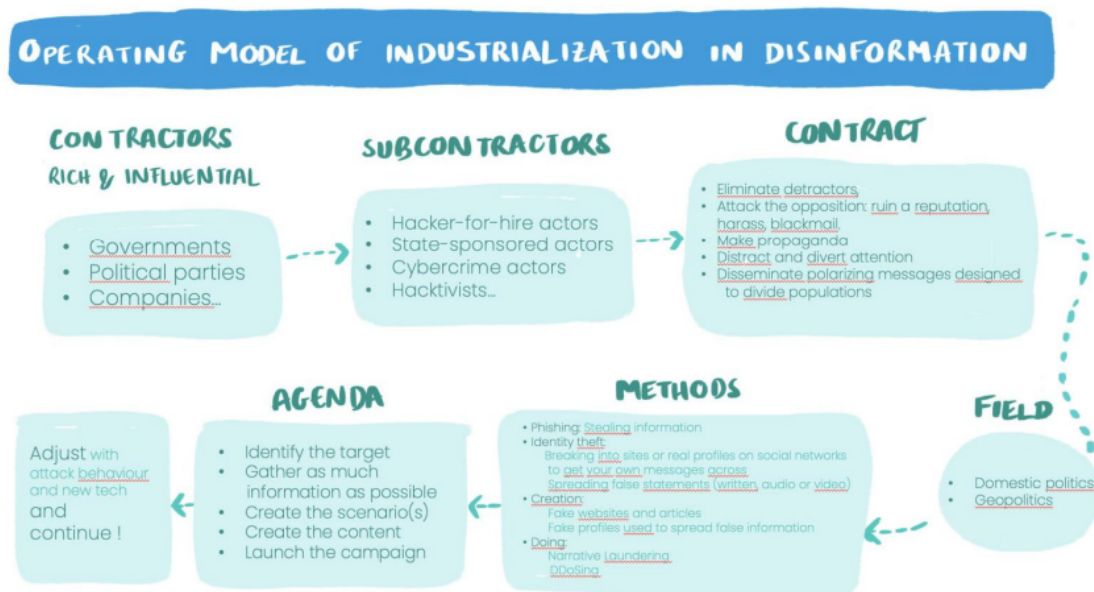
¹¹ Ce qu'il faut savoir sur Cambridge Analytica, la société au cœur du scandale Facebook : https://www.lemonde.fr/pixels/article/2018/03/22/ce-qu-il-faut-savoir-sur-cambridge-analytica-la-societe-au-c-ur-du-scandale-facebook_5274804_4408996.html

L'INDUSTRIALISATION DE LA DÉSINFORMATION

Certaines organisations ont bien compris ce phénomène de société et ont donc industrialisé la désinformation. Cette industrialisation renvoie à la mise en œuvre de méthodes et de systèmes à grande échelle pour créer et diffuser de la de fausses informations de manière organisée et systématique.

Plusieurs acteurs, notamment des entreprises spécialisées dans le marketing et la communication, des groupes d'activistes cyber hackers, des politiciens, des gouvernements sont au cœur de cette industrie. Les motivations sont multiples et variées : les diffuseurs s'intéressent à l'appât du gain, les commanditaires, notamment des entreprises, cherchent à ternir l'e-réputation de leurs concurrents, les gouvernements étrangers visent quant à eux à déstabiliser leurs ennemis, faire de l'ingérence dans les affaires du pays ciblé, manipuler son opinion publique, créer de l'instabilité politico-sociale. D'où le rôle crucial de l'intelligence économique dans l'analyse stratégique de l'information à l'ère de la guerre (dés-)informationnelle pour éclairer les décideurs publics et économiques.

Cet écosystème industrialisé est un marché extrêmement lucratif à l'échelle mondiale. En effet, Emma Briant, chercheuse au Bard College affirme que : « *la manipulation de la vérité s'est transformée en un commerce de plusieurs milliards de dollars, qui s'appuie sur des centaines, voire des milliers de prestataires dans le monde entier* »¹². Le schéma ci-dessous résume le modèle des relations entre acteurs (activistes, cyberattaquants, gouvernements, organisations, entreprises, ...) de l'industrie de la désinformation dans le rapport sur l'industrialisation de la désinformation de l'université d'Oxford (2021)¹³ :



Source : Fact checking Trends & patterns (p41), france.tv, 21 avril 2023

¹² À l'ère des fausses informations : une journaliste, un meurtre et la poursuite d'une enquête inachevée en inde : <https://forbiddenstories.org/fr/story-killers/gauri-lankesh-ere-fausses-informations/>

¹³ Industrialized Disinformation. Global Inventory of Organized Social Media Manipulation : <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/02/CyberTroop-Report20-Draft9.pdf>

IV. TYPOLOGIE DES OUTILS DE LUTTE CONTRE LA DÉSINFORMATION

L'analyse des différents outils montre plusieurs éléments : d'abord il existe une diversité de dispositifs techniques disponibles gratuitement (ou non) pour la lutte contre la désinformation en ligne. Certains fonctionnent à l'aide d'algorithmes d'intelligence artificielle, d'autres nécessitent l'implémentation humaine voire mixte (humain et IA).

Une **base de données non exhaustive** d'outils web de lutte contre la désinformation a été construite suite à une analyse approfondie de leur fonctionnalité. Ce travail de recherche et collecte des outils en ligne a été réalisé en deux phases. Phase 1 : deux étudiants de l'école d'ingénieurs EPITA ont collecté les différents outils existant et ont produit une base de données ; Phase 2, le traitement et l'analyse du contenu de cette base de données d'outils a été complétée par les membres de la commission notamment l'équipe porteuse du projet.

L'analyse du graphique ci-dessous montre une répartition par typologie de tâches spécialisées de chaque outil. Il en ressort que plus de **43%** des outils sont **spécialisés dans la vérification** de faux sites web, d'informations générées par des bots intelligents ou de l'IA, du contenu manipulé, de propagande informationnelle, d'authenticité d'une photo/vidéo en ligne, de faits publiés en ligne et des déclarations des hommes politiques, etc.

En matière de **sensibilisation, de formation et de notation de crédibilité, 15 outils sur 45 en sont spécialisés. Certains (8 d'entre eux)** permettent de construire une **notation de crédibilité** d'une source d'information et du contenu, en fonction (I) de l'étendue et de la qualité des sources, (II) de l'expertise de l'auteur du contenu, (III) de la nature du langage utilisé et (IV) de la réputation historique du média en question.

En ce qui concerne l'évaluation d'une source, des chercheurs comme Jacques Breillat, expert en IE, proposent une méthode en deux phases : la première phase renvoie à l'analyse des éléments clés du document (date de publication, auteur ou organisation) et la deuxième phase consiste à définir une échelle du renseignement permettant de « noter » la source et évaluer l'information¹⁴.

¹⁴ Jacques Beillat : intelligence économique, Information et influence : <http://jacques.beillat.fr/methode/evaluer-ses-sources>

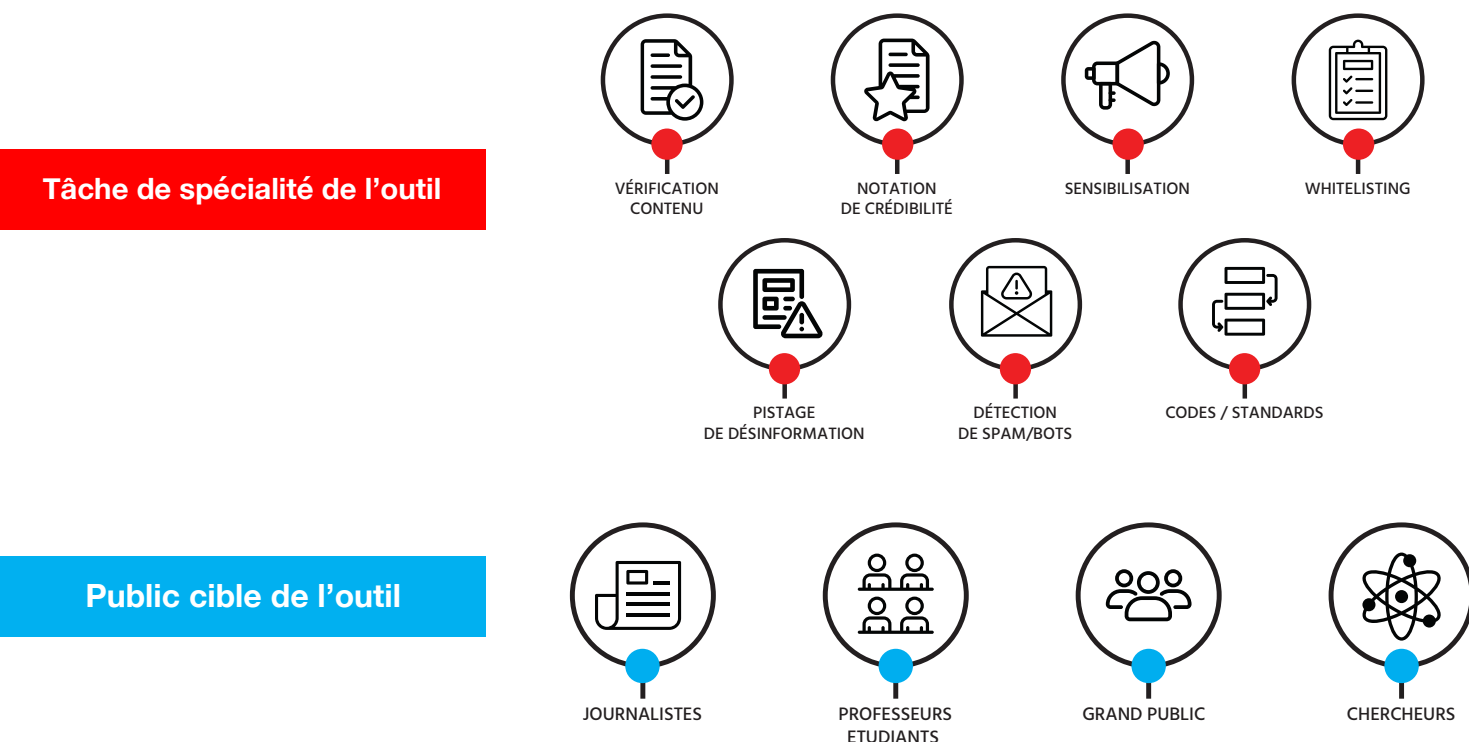
Selon J. Breillat, il existe des limites à la notation d'une source d'information que sont : I) une information importante ou des « signaux faibles » peuvent parfois passer inaperçus et II) une fausse information peut être surestimée au point de se laisser « intoxiquer ».

Dans le domaine de la défense, la lutte contre de la désinformation est vue sous l'angle du cyberspace. Selon Aurélie Laizé, Responsable de la structuration de l'activité de lutte informatique d'influence pour la France chez Airbus Defence and Space et Directrice du site de Rennes, dans le domaine de la défense, l'approche de lutte contre la désinformation est multicanale. Pour elle, deux défis majeurs se présentent dans le domaine de la défense : I) comment tirer profit de tous ces outils existants pour apporter des réponses opérationnelles et II) comment faire collaborer des profils aux compétences pluridisciplinaires (intelligence économique, sociologie, sciences politiques, marketing et sciences dures) afin faire converger les efforts de lutte contre les attaques informationnelles.

D'autres (7 outils) dans cette base de données sont dédiés à faire de la sensibilisation/formation en ligne en proposant des modules captivants sur des sujets tels que la partialité des médias, la désinformation, la pensée conspirationniste, etc.

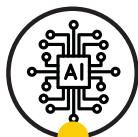
L'un d'eux, [Bad News](#), est un jeu vidéo en ligne dont l'accès est gratuit et il aide les internautes à mieux connaître, détecter les « fake news » pour mieux les ignorer/contrer.

Ci-dessous, la légende du tableau des outils existants permet de se repérer facilement, à travers des icônes sur le public cible, le type de tâche de l'outil et le mode de fonctionnement.





BLOCKCHAIN



MACHINE LEARNING / IA



COURS OU
JEU EN LIGNE



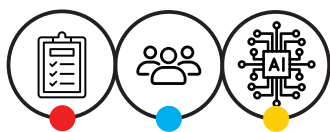
IMPLÉMENTATION
HUMAINE



CROWDSOURCING

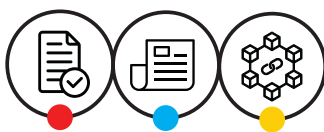
TABLEAU DES OUTILS WEB IDENTIFIÉS DANS LE CADRE DE CE BENCHMARK.

Dans les pages suivantes, le tableau recense l'ensemble des outils web analysés collectés, traités et analysés.



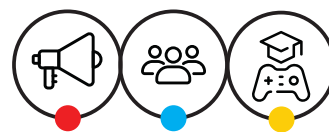
Domain Whitelist

Outil d'extension chrome qui détecte les faux éléments dans les sites web et identifie le processus de fabrication et diffusion de l'information.



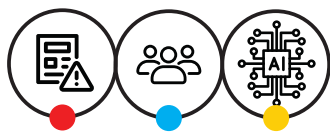
Snopes

Snopes : plateforme en ligne de vérification de contenu visuel manipulé, d'informations générées par des bots d'intelligents ou encore la propagande informationnelle et censure les fausses informations.



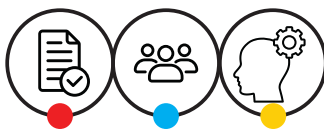
Truth Goggles 2.0

Outil qui met à disposition des utilisateurs à une diversité de sources d'information et de points de vue sur des affirmations spécifiques afin de permettre une meilleure vision plus précise des preuves sur un sujet donné.



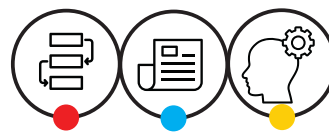
WoTargets Me

Outil à la fois axé sur le contenu et sur le processus. Il authentifie des photos, renseigne sur le ciblage en ligne sur la manière dont l'information est produite et diffusée. Il piste les informations manipulées, notamment les discours politiques, etc.



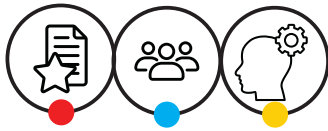
Factcheck.org

Outil identifiant l'exactitude des faits, en explorant et corrigeant les fausses informations, en donnant au public une ressource pour localiser des informations basées sur des faits. Il se concentre sur les déclarations faites par les hommes politiques américains notamment.



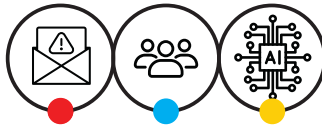
IFCN Codes and Principles

Outil qui vise à établir des normes, des méthodes pour garantir la vérification rigoureuse et objective des faits en promouvant la valeur et crédibilité des outils.



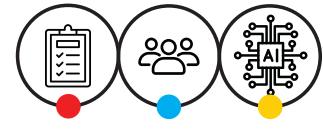
Disinformation Index

Outil en ligne qui évalue la crédibilité des organes d'information en fonction de la probabilité de désinformation. Il note sur la qualité des sources (exactitude, fiabilité, objectivité, etc.).



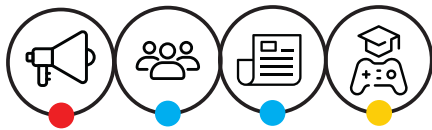
Botometer

Outil en ligne qui chasse et classe les faux comptes Twitter, les bots en examinant les caractéristiques d'un profil (amis, structure du réseau social, activité temporelle, langage et le sentiment). Il réalise un score global de bot (0-5) et d'autres scores qui mesurent la probabilité que le compte soit un bot.



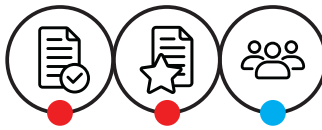
Website whitelist

Extension navigateur qui permet aux utilisateurs d'identifier les sites web fiables à mettre sur liste blanche, de bloquer les sites web non fiables et de réduire également le micro ciblage (moyen de désinformation également).



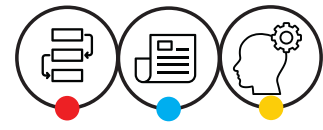
Trust and Verification

Cours en ligne gratuit qui sensibilise sur la construction de confiance en tant que journaliste/créateur de contenu à l'ère de la désinformation. Il met l'accent sur la manière d'instaurer la confiance, sur l'évolution du paysage médiatique et sur la manière de vérifier la fiabilité en ligne.



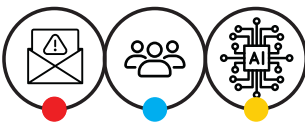
Claim Review

ClaimReview est une étiquette de source attachée au contenu d'un site web qui fournit aux moteurs de recherche des informations sur le contenu du site web. Il peut s'agir d'une étiquette "Fact Check" pour les informations qui sont vérifiées par des sources et des organisations dans le but de fournir des informations plus précises aux utilisateurs.



Certified Content Coalition

Outil qui vise à réduire la désinformation en améliorant les pratiques journalistiques et les normes relatives à la qualité de l'information et à la transparence parmi les journalistes, les éditeurs et les autres acteurs de l'espace médiatique. L'outil encourage l'adoption de normes par les éditeurs de médias en ligne et à certifier les éditeurs qui respectent ces normes.



Bot Sentinel

Plateforme gratuite en ligne dédiée à la détection, au suivi de trollbots et de faux comptes Twitter (X). Il combat la désinformation en identifiant, étiquetant et suivant les bots et en classant les comptes indignes de confiance. Cela permet aux internautes de faire la distinction entre un contenu automatisé et un contenu humain.



InVid

Boîte à outils conçue pour aider les journalistes mais aussi le grand public à gagner du temps et à être plus efficaces dans leurs tâches de vérification des faits et de démystification sur les réseaux sociaux, en particulier les images et vidéos.



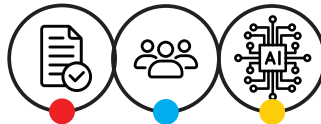
Checkology

Plateforme de formation en ligne gratuite qui propose des leçons captivantes, fiables sur des sujets tels que la partialité des médias d'information, la désinformation, la pensée conspirationniste et bien d'autres encore. Elle outille les apprenants à développer leur capacités à identifier, qualifier des informations et sources crédibles et faire preuve d'esprit critique pour distinguer les vrais et faux contenus.



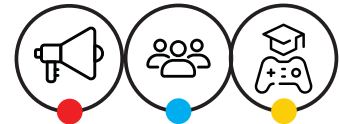
Trust Project Indicators

Une initiative de collaboration entre plusieurs entreprises de presse qui développent les indicateurs de confiance suivants : Meilleures pratiques, expertise de l'auteur/journaliste, type de travail, citations et références, méthodes, sources locales, voix diverses, retour d'information exploitable. Les organismes de presse afficheront les indicateurs sur leurs propres articles. Cela permet d'aider les utilisateurs à éviter les sources de moindre qualité susceptibles de contenir des informations erronées ou de la désinformation.



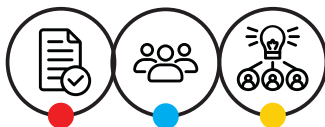
ClaimBooster

Outil de vérification des faits automatisés et en direct, basé sur le web. L'outil s'appuie sur le traitement du langage naturel et l'apprentissage supervisé (sur la base d'un ensemble de données) pour identifier les informations factuelles et fausses. Une application est également disponible pour Slack. Cet outil vise à lutter contre la désinformation en détectant et corrigeant les fausses informations, en donnant aux internautes une ressource pour localiser des informations basées sur des faits.



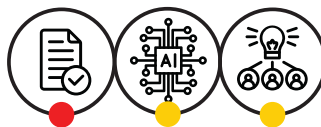
Crashcourse

Série de vidéos éducatives diffusées sur une chaîne YouTube. Une série de six vidéos est consacrée à l'éducation aux médias. Cet outil vise à lutter contre la désinformation en créant une cohorte de consommateurs d'informations mieux à même d'identifier et d'éviter les fausses informations en ligne.



CrossCheck

Outil de collaboration en ligne permettant de signaler avec précision les allégations fausses, trompeuses et déroutantes qui ont circulé en ligne. Son but est de permettre de publier davantage d'articles de fact-checking pour démonter les rumeurs circulant en ligne, grâce à une coordination des moyens de détection et de vérification des informations.



Exiftools

Application informatique qui fournit des métadonnées sur la source, l'horodatage, la création et les modifications. Cet outil permet de lutter contre la désinformation en vérifiant et en authentifiant les images.



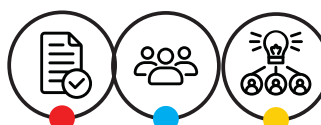
Share the Facts

Widget web et une application Slack qui fournit aux lecteurs un résumé d'un article vérifié, mettant en évidence l'affirmation vérifiée et la conclusion. En utilisant le widget, les vérifications de faits peuvent être mises en évidence dans les résultats de recherche Google par exemple.



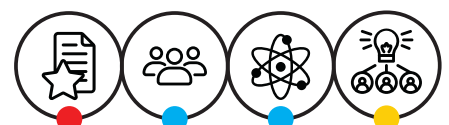
Exitdata

Outil en ligne qui fournit des informations sur la source, l'horodatage, la création et la modification. Axé sur la vérification de contenu l'outil permet d'évaluer directement les informations telles que l'authenticité d'une photo. Il suffit d'importer le document, l'image pour qu'il puisse l'authentifier.



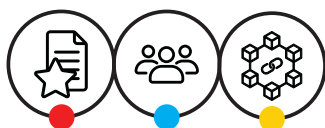
Rbutr

Plateforme collaborative en ligne qui développe une base de données de pages web, pour lutter contre la désinformation et de réduire les "fake news". Il ne fait pas de distinction entre le vrai et le faux, mais identifie simplement les réfutations afin de fournir aux utilisateurs un ensemble de perspectives plus diversifiées.



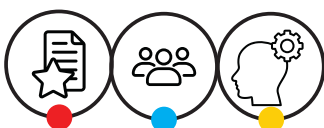
Climate Feedback

Outil de notation de contenu basé sur le web qui permet aux scientifiques d'annoter des articles afin de fournir un contexte supplémentaire et d'attirer l'attention sur les inexactitudes. Le processus aboutit à un score de crédibilité. Cet outil vise à lutter contre la désinformation en créant une plateforme pour la diffusion d'informations vérifiées et fondées sur des preuves.



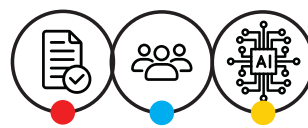
Misinformation Detector

Outil qui permet d'évaluer la crédibilité des informations. Il analyse le contenu des médias ainsi que la manière dont l'information se propage. Il attribue ensuite à l'organisme d'information un classement de confiance. Il permet également aux journalistes et aux éditeurs de travailler avec l'organisation et d'en être les partenaires. L'outil fournit aux lecteurs des informations qui les aident à distinguer les sources d'information de qualité de celles qui ne le sont pas.



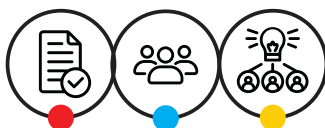
KnowNews

Extension de navigateur développée qui classe les sites d'information en fonction de leur crédibilité (crédibles, douteux, non classés). Cet outil utilise des évaluations et des notations de la qualité des sources (exactitude, objectivité, etc.) pour guider les lecteurs vers des informations de meilleure qualité plutôt que vers des informations de mauvaise qualité.



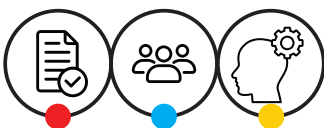
Duke VideoFastchecking Tools

Extension de navigateur qui permet de vérifier en direct les informations diffusées à la télévision. Il a été évalué par les développeurs à l'aide d'expériences basées sur les utilisateurs. Cet outil permet de lutter contre la désinformation en vérifiant et en authentifiant les images vidéo.



Captain Fact

Outil basé sur le web et conçu pour la vérification de contenu de l'internet. Il comprend une extension de navigateur qui permet de superposer aux vidéos Internet des sources et des informations contextuelles, ainsi que des icônes indiquant la crédibilité sur la base des votes des utilisateurs. Il a un moteur de recherche dans lequel, il est possible d'écriture le sujet recherché.



TV New Fact Check

Cet outil vise à lutter contre la désinformation en mettant à la disposition du public davantage d'informations de première main et exactes dans une "bibliothèque numérique" de pages web, de livres, d'enregistrements audio, de vidéos, d'images et de logiciels.



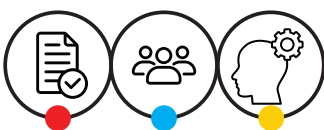
Youtube Data Viewer

Outil de vérification vidéo en ligne créé par Amnesty International. Les utilisateurs saisissent l'URL de YouTube et l'outil fournit des informations sur la vidéo qui sont utiles pour la vérifier. Il s'agit notamment de l'heure de téléchargement et des vignettes qui peuvent être utilisées pour la recherche inversée d'images. Cet outil permet de lutter contre la désinformation en vérifiant et en authentifiant le contenu des vidéos.



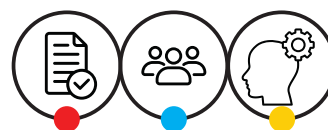
Verification Tool

Outil français qui permet de vérifier l'authenticité et l'intégrité des documents publiés et protéger contre la fraude en ligne ou la manipulation d'information à toutes fins.



Wiztrust

Outil français qui permet de vérifier, certifier, tracer, sécuriser et authentifier des documents et rapports stratégiques diffusés par les entreprises et les médias afin de les protéger de la désinformation en ligne.



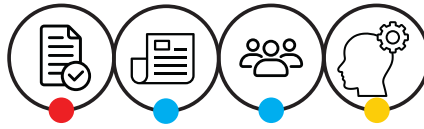
Politifact

Site web qui vérifie les faits concernant des déclarations "dignes d'intérêt et significatives" et leur attribue les mentions "vrai", "presque vrai", "à moitié vrai", "presque faux", "faux" et "sans fondement". Le processus comprend l'examen d'autres sources de vérification des faits, des recherches sur Google, des recherches dans des bases de données en ligne, des consultations d'experts etc.



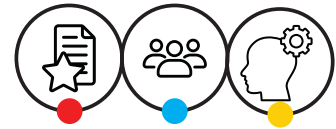
Video Verification Plugin (InVid)

Boîte à outils plug-in conçue pour faciliter la vérification des faits par le biais de la vérification vidéo. L'outil fournit aux utilisateurs des informations contextuelles sur les vidéos, une recherche inversée d'images, des métadonnées vidéo, des informations sur les droits d'auteur des vidéos, ainsi que d'autres fonctions permettant de vérifier le contenu.



NewsGuard

Extension navigateur en ligne qui évalue à l'aide de neuf critères, les sites d'informations en fonction de leur fiabilité. L'outil donne aux médias une note, attribuée par des équipes de journalistes en fonction de ses critères de fiabilité et de transparence. Ceux remplissent obtiennent un symbole en forme de bouclier vert, les recalés écopent d'un bouclier rouge et les sites satiriques sont signalés par un symbole jaune.



Media Bias Fact Check

Plateforme en ligne qui permet d'évaluer la partialité des médias en fonction du niveau et de l'orientation de la partialité d'un site web, et d'identifier les sources douteuses et conspirationnistes. Cet outil permet de lutter contre la désinformation en fournissant aux utilisateurs davantage d'informations sur la partialité des sources, ce qui leur permet de prendre des décisions en connaissance de cause et de rechercher des informations objectives.

En définitive, ce benchmark nous permet d'avoir un aperçu de ce que pourrait-être un outil type de lutte contre la désinformation. Il apparaît que **trois principales fonctionnalités** demeurent **essentielles** dans un outil :

- I) la **fonction-tâche de l'outil** (vérification de contenu, détection de bots, notation de crédibilité),
- II) la **fonction-moyen ou fonctionnement** (le machine learning, l'IA, le deep learning permettent plus efficacement de détecter et distinguer automatiquement et systématiquement la vraie de la fausse information, contenu de l'infobésité à laquelle les internautes sont confrontés quotidiennement)
- III) la **fonction-cible** (grand public, journalistes, chercheurs, étudiants, enfants, etc.).

Ce benchmark et son tableau des outils n'est pas exhaustif. Il pourra être complété par d'autres travaux de recherches sur le sujet. Néanmoins, l'idée était d'analyser et de partager quelques outils en ligne gratuits qui peuvent aider le grand public, les journalistes, chercheurs, étudiants, professeurs et chefs d'entreprises à s'armer pour mieux lutter contre la désinformation dans une ère guerre informationnelle systémique. D'où le rôle central des professionnels de l'intelligence économique dans la compréhension et le décryptage du paysage informationnel.

V. SOURCES BIBLIOGRAPHIQUES

France.tv & European Commission. (2023). Fact checking trends & patterns on disinformation. Dans <https://www.meta-media.fr/>. <https://www.meta-media.fr/wp-content/uploads/sites/33/2023/05/trends--patterns-on-disinformation-21april2023.pdf>

HENIN, N. (2023). Disinformation landscape in France. Dans Word Economic Forum. EU Disinfo Lab. https://www.disinfo.eu/wp-content/uploads/2023/03/20230224_FR_DisinfoFS.pdf

Digital Services Act (DSA): <https://www.vie-publique.fr/eclairage/285115-dsa-le-reglement-sur-les-services-numeriques-ou-digital-services-act>

FURNEMONT, J., OUSMANE, A., & DE LA BROSSE, R. (2022). La lutte contre la désinformation dans les politiques publiques francophones : état des lieux comparatifs : Etat des lieux comparatif. Dans <https://www.francophonie.org>. ORGANISATION INTERNATIONALE DE LA FRANCOPHONIE. https://www.francophonie.org/sites/default/files/2022-10/OIF_Rapport%20politiques%20publiques_Lutte%20contre%20la%20de%CC%81sinformation.pdf

Le règlement européen sur les services numériques (DSA) vise une responsabilisation des plateformes : <https://www.vie-publique.fr/eclairage/285115-dsa-le-reglement-sur-les-services-numeriques-ou-digital-services-act>

Les lumières à l'ère du numérique (2022). Rapport de la commission présidée par Gérald Bronner à la demande de la Présidence de la République. Vie-publique. <https://www.vie-publique.fr/files/rapport/pdf/283201.pdf>

HELMUS, Todd. C. (2022). Artificial intelligence, deepfakes, and Disinformation. Dans <https://www.rand.org>. A Primer RAND Corporation. https://www.rand.org/content/dam/rand/pubs/perspectives/PEA1000/PEA1043-1/RAND_PEA1043-1.pdf

POLDI, F., ALAPHILIPPE, A., JONUSAITE, R., & SESSA, M. G. (2022). Disinformation on Telegram : Research and content moderation policies. Dans EU Disinfo Lab (Éd.), <https://www.disinfo.eu>. European Commission. https://www.disinfo.eu/wp-content/uploads/2022/12/20221220_TD_Telegram.pdf

Tools that fight disinformation online. (2022). Rand Organisation. <https://www.rand.org/research/projects/truth-decay/fighting-disinformation/search.html>

The four key ways disinformation is spread online. (2022, août 9). World Economic Forum. <https://www.weforum.org/agenda/2022/08/four-ways-disinformation-campaigns-are-propagated-online/>

BRADSHAW, S., ALAPHILIPPE, A., BAILEY, H., & HOWARD, P. N. (2021). Industrialized disinformation. Global inventory of organized social media manipulation. Dans <https://demtech.oii.ox.ac.uk>. University of Oxford. <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/01/CyberTroop-Report-2020-v.2.pdf>

National Democratic Institute, International Republican Institute, & Stanford Internet Observatory cyber policy center. (2021). Guide sur la manipulation de l'information : un conseil pratique pour les élections et au-delà. Dans <https://www.ndi.org/>. https://www.ndi.org/sites/default/files/infomanip_playbook_french_final.pdf

Fake news : La France se dote d'une agence de surveillance des réseaux sociaux. (2021, 2 juin). <https://www.lesechos.fr>. <https://www.lesechos.fr/politique-societe/gouvernement/fake-news-la-france-se-dote-dune-agence-de-surveillance-des-reseaux-1320268>

Contre la manipulation de l'information. (2021, 15 juin). <https://www.gouvernement.fr>; Les Echos. <https://www.gouvernement.fr/action/contre-la-manipulation-de-l-information>

Radicalisation Awareness Network. (2020). Inciter les jeunes à lutter contre la désinformation en ligne : Document de conclusion. Dans <https://home-affairs.ec.europa.eu>. https://home-affairs.ec.europa.eu/document/download/3ed371fb-4044-4e69-8f17-ae3f4ad392ff_fr

Commission européenne. (2020, décembre). Plan d'action pour la démocratie européenne : Renforcer les démocraties de l'UE. <https://ec.europa.eu>; A Primer RAND Corporation. https://ec.europa.eu/commission/presscorner/detail/fr/qanda_20_2251

LOI n° 2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information. (2018, 22 décembre). <https://www.gouvernement.fr>; Légifrance. <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000037847559/>

BRADSHAW, S., & HOWARD, P. N. (2017). Troops, Trolls and Troublemakers : a global inventory of organized social media manipulation. Dans <https://demtech.oii.ox.ac.uk> University of Oxford. <https://ora.ox.ac.uk/objects/uuid:cef7e8d9-27bf-4ea5-9fd6-855209b3e1f6/files/m3ca8c455852611e82d0fb182445a471f>



PRÉSENTATION DE LA COMMISSION « MANIPULATIONS DE L'INFORMATION » DE L'AAIE-IHEDN

PHILOSOPHIE ET OBJECTIFS :

Jamais l'information n'a été si accessible, abondante et partageable. Accéder à la « vérité », ou du moins vérifier par soi-même la véracité d'une information ainsi que l'intention de son émetteur, est dans la majorité des cas à la portée de n'importe quel citoyen motivé, comme en témoigne la démocratisation de l'OSINT. Néanmoins, certaines « fake news », théories du complot ou « faits alternatifs » persistent, séduisent et convainquent, tout en rendant leurs victimes inaudibles.

La manipulation s'applique sur l'information, mais également sur le « champs informationnel » dans lequel elle évolue. D'avantage qu'une connaissance, l'information y fait office de munition dans une « guerre de l'information » où il s'agit de prendre le contrôle de la perception et de l'attention de sa cible. L'attention est-elle une ressource rare. Elle ne se partage pas : elle se conquiert et se garde. Une bataille de l'attention implique désormais entreprises, états et citoyens. Ses enjeux dépassent le cadre de la pédagogie, de l'éthique et de l'addiction aux écrans.

Accaparer l'attention, changer les perceptions, c'est non seulement se donner les moyens de tisser un lien privilégié avec ses interlocuteurs pour mieux communiquer et persuader, mais consiste aussi à en priver la "concurrence", qu'elle soit politique, économique, sociale, qu'il s'agisse de nos vies personnelles et professionnelles. Cet affrontement informationnel et attentionnel est planétaire via internet. Sans début ni fin, la bataille ne connaît pas de répit, rythmée par les technologies persuasives de nos outils. Elle ne se limite pas aux contenus et aux écrans, mais se déroule aussi dans notre cerveau, et bientôt notre environnement via nos objets connectés.

L'UE qualifiait la désinformation de phénomène « *sous surveillance mais pas sous contrôle* ». L'objectif de la commission « Manipulations de l'information » de l'AAIE-IHEDN, créée au premier semestre 2023, est donc certes de comprendre les manipulations de l'information, à travers deux axes, la « *guerre de l'information* » et « *l'économie de l'attention* », mais aussi d'identifier des réponses à adopter.

FONCTIONNEMENT :

La commission « Manipulations de l'information » est composée de 20 personnes maximum, à parité homme femme. Ses réunions, qui font le point sur les projets en cours, se déroulent uniquement en présentiel.

Elle organise des diners débats, où un intervenant extérieur vient présenter un sujet, l'idée étant de constituer un réseau d'experts complémentaire à celui de l'AAIE IHEDN. Elle a organisé fin 2023 à l'Ecole Militaire la conférence AAIE IHEDN « *La guerre cognitive* », ouverte au public. Enfin, la commission réalise des livrables, dont ce « *Benchmark des outils de lutte contre la désinformation* ».

Thibault Renard



Association des auditeurs en
intelligence économique de
l'Institut des Hautes Etudes
de Défense Nationale