N° 2192

ASSEMBLÉE NATIONALE

CONSTITUTION DU 4 OCTOBRE 1958

DIX-SEPTIÈME LÉGISLATURE

Enregistré à la Présidence de l'Assemblée nationale le 3 décembre 2025

RAPPORT D'INFORMATION

DÉPOSÉ

en application de l'article 145 du Règlement

PAR LA COMMISSION DES AFFAIRES ÉTRANGÈRES

en conclusion des travaux d'une mission d'information constituée le 15 janvier 2025

*sur l'***irruption de l'intelligence artificielle dans les ingérences étrangères**

ET PRÉSENTÉ PAR

M. ALAIN DAVID ET MME LAETITIA SAINT-PAUL,

Députés

SOMMAIRE

———

Pages

SYNTHÈSE DES PROPOSITIONS DES RAPPORTEURS 5

INTRODUCTION 7

I. L'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE REDÉFINIT LES OPÉRATIONS D'INGÉRENCES ÉTRANGÈRES EN MENAÇANT REDOUTABLEMENT NOS MODÈLES DÉMOCRATIQUES 11

A. L'ESSOR DE L'INTELLIGENCE ARTIFICIELLE CONSTITUE UNE RÉVOLUTION TECHNOLOGIQUE MAJEURE 11

1. Discipline et technologie, l'intelligence artificielle a connu des développements irréguliers depuis les années 1950 11

2. Les avancées récentes en matière d'intelligence artificielle consacrent un changement de paradigme en démocratisant le recours à cette technologie 13

3. Au cœur de logiques géopolitiques, le développement de l'intelligence artificielle s'inscrit dans une quête étatique de la puissance 14

B. LE RECOURS À L'INTELLIGENCE ARTIFICIELLE AMPLIFIE ET DIVERSIFIE LES ATTAQUES DANS LE CADRE D'UNE GUERRE HYBRIDE 18

1. De la guerre pour l'intelligence artificielle à la guerre par l'intelligence artificielle, les ingérences étrangères trouvent un terreau fertile dans un environnement géopolitique marqué par des tensions exacerbées 18

2. Afin de fragiliser les systèmes politiques et sociétaux, le recours à l'intelligence artificielle renforce les tentatives de déstabilisation du modèle occidental 22

a. De nombreuses attaques impliquant l'intelligence artificielle ont pour objectif la manipulation des opinions 22

b. Les attaques par empoisonnement des données font de l'intelligence artificielle à la fois un outil et une cible d'ingérence 28

c. L'utilisation de l'intelligence artificielle dynamise les cyberattaques en contribuant à intensifier la menace numérique 30

3. Les risques sécuritaires sont particulièrement accrus pour les systèmes démocratiques en raison des valeurs qu'ils défendent et de leurs processus électoraux 32

II. LA PERCEPTION FRANÇAISE ET EUROPÉENNE DU DANGER APPARAÎT ENCORE PARTIELLE ET PARFOIS NAÏVE AU REGARD DU POSITIONNEMENT DES AUTRES PUISSANCES 35

A. L'UNION EUROPÉENNE SOUFFRE D'UNE VULNÉRABILITÉ STRUCTURELLE RENDANT DIFFICILE L'ÉLABORATION D'UNE STRATÉGIE POUR L'INTELLIGENCE ARTIFIELLE 35

1. La souveraineté numérique européenne s'avère abstraite voire fantasmée, plaçant le vieux continent dans une situation de dépendance aux technologies étrangères 35

2. Cette absence d'autonomie se traduit par un retard dans le développement de l'intelligence artificielle en Europe 37

B. FACE AUX MENACES, UNE DOUBLE RÉPONSE RÉGLEMENTAIRE ET OPÉRATIONNELLE CONSTITUE UNE DÉFENSE PERFECTIBLE 41

1. Novateur, le cadre juridique européen érige un modèle de régulation protecteur mais demeure complexe et peu lisible 41

2. Les moyens opérationnels de contre-ingérence ont été renforcés mais paraissent encore insuffisants au regard de l'ampleur de la menace 47

a. L'Union européenne surveille les campagnes de désinformation grâce à son service pour l'action extérieure 47

b. En France, la lutte contre les ingérences étrangères mobilise plusieurs ministères et opérateurs, aussi bien dans le domaine informationnel que cyber 48

III. DIX-HUIT MESURES POURRAIENT ÊTRE PORTÉES AFIN DE SÉCURISER NOS DÉMOCRATIES EN INTÉGRANT LE RÔLE DE L'INTELLIGENCE ARTIFICIELLE DANS L'ÉVOLUTION DES MENACES 54

A. UNE RÉPONSE ADAPTÉE IMPLIQUE UNE MEILLEURE APPRÉHENSION DE LA TECHNOLOGIE DE L'IA, À LA FOIS EN TERMES DE PRODUCTION, DE DIFFUSION ET DE CONSOMMATION .. 54

1. Une maîtrise des technologies d'IA est un impératif afin de se prémunir des usages malveillants par des acteurs étrangers 54

2. Une population sensibilisée et préparée constitue l'élément central de la stratégie de défense d'une société résiliente face aux ingérences et aux dérives permises par l'IA 56

3. Le soutien à une production d'information fiable et de qualité est indispensable pour redonner du sens au doute et réapprendre à faire confiance 60

B. LES MOYENS DE CONTRE-INGÉRENCE DOIVENT ÊTRE RENFORCÉS AFIN DE DISPOSER D'OUTILS DÉFENSIFS ET OFFENSIFS PERFORMANTS 64

1. Les structures de détection et d'analyse des menaces mobilisant l'IA nécessitent un soutien accru pour entériner la montée en puissance de notre défense 64

2. La posture défensive doit se doubler d'une riposte offensive et proactive 66

EXAMEN EN COMMISSION 69

ANNEXE N° 1 : LISTE DES PERSONNES AUDITIONNÉES PAR LES RAPPORTEURS 91

SYNTHÈSE DES PROPOSITIONS DES RAPPORTEURS

Proposition : Rendre obligatoire le consentement aux algorithmes dans les suggestions de contenus sur les plateformes.

Proposition : Instaurer une « réserve algorithmique pré-électorale » sur les plateformes pendant une période définie en amont du scrutin afin d'éviter toute manipulation et de préserver la sincérité des votes.

Proposition : Mettre rapidement et effectivement en place, au niveau européen ou, à défaut, à l'échelle nationale, une majorité numérique conditionnant l'accès aux réseaux sociaux.

Proposition : Réinvestir l'espace public par le développement, avec Viginum, d'une communication active pour alerter sur la désinformation et l'usage malveillant de l'IA, en concentrant les efforts en période pré-électorale.

Proposition : Renforcer les moyens humains, matériels et financiers de Viginum afin d'accompagner sa montée en puissance et l'extension de son champ d'activité.

Proposition : Étendre le champ de la RSE à la protection de la démocratie et encourager, dans ce cadre, un financement vertueux de l'information par les entreprises privées et leurs pratiques publicitaires.

Proposition : Garantir à l'ANSSI les moyens nécessaires à la bonne exécution de ses missions dont le volume et la nature évoluent, tout en transposant au plus vite la directive européenne dite NIS 2 dans notre droit national.

Proposition : Imposer aux plateformes de réseaux sociaux d'étiqueter les contenus avec un score d'artificialité indiquant dans quelles proportions l'IA a été utilisée pour les générer.

Proposition : Envisager une labellisation informationnelle selon un modèle de nutri-score avec des critères objectifs et en sources ouvertes afin que chacun puisse vérifier l'impartialité du système.

Proposition : Renforcer les moyens financiers de l'audiovisuel public extérieur pour raffermir la réponse informationnelle française à l'étranger.

Proposition : Accroître les efforts en matière de communication stratégique et adopter une posture plus offensive pour contrer les narratifs adverses.

Proposition : Envisager l'élargissement du comité opérationnel de lutte contre les manipulations de l'information (COLMI) à d'autres ministères particulièrement mobilisés sur les questions d'ingérence et de désinformation.

Proposition : Nouer des partenariats avec des associations au niveau local pour atteindre davantage toutes les diasporas lors des opérations de sensibilisation aux dangers des ingérences étrangères.

Proposition : Promouvoir la création de *think-tanks* spécialisés au niveau national et, plus largement, soutenir la recherche scientifique afin d'objectiver les conséquences des ingérences étrangères et de diffuser une culture de la résilience dans la société.

Proposition : Encourager la conclusion de partenariats public-privé pour développer massivement les techniques de vérification de l'information.

Proposition : Soutenir les initiatives conjointes en matière de lutte contre la désinformation et d'utilisation de l'IA avec les pays amis afin de consolider un réseau de contre-ingérence, au niveau international et à l'échelle européenne dans le cadre du bouclier démocratique.

Proposition : Distinguer les investissements extra-communautaires et européens et favoriser les seconds notamment par des mécanismes de réduction des risques tels que des garanties européennes de financement, par une commande publique privilégiant l'origine européenne ou par un fléchage de l'épargne vers le développement du secteur de l'IA.

Proposition : Systématiser obligatoirement dans les maquettes des cursus en journalisme des cours dédiés à la vérification de l'information et à la lutte contre la désinformation.

INTRODUCTION

« *La violence s'arme des inventions des arts et des sciences pour combattre la violence* » [1]. Le constat dressé par le général Carl von Clausewitz après les guerres napoléoniennes s'impose aujourd'hui avec la même force qu'en 1830 : les sociétés, en quête de sécurité, ont cherché et aspirent à transformer les progrès technologiques en dispositifs de plus en plus efficaces pour surveiller les menaces, exercer une influence au-delà de leurs frontières et, en cas de guerre, permettre à la force de prévaloir. À l'ère numérique, ce triptyque s'exprime désormais par l'utilisation d'un nouveau canal, l'intelligence artificielle (IA).

Comprise comme l'ensemble des méthodes utilisées pour modéliser un phénomène avec l'objectif de résoudre un problème ou de répondre à une question [2], l'intelligence artificielle transforme en profondeur des pans entiers de l'activité humaine. Elle induit une rupture avec les modes de production antérieurs en promettant des gains de productivité substantiels tout en assurant le traitement de flux de données avec des volumes inédits. Opportunité du siècle, l'intelligence artificielle suscite également des craintes proportionnelles à l'enthousiasme généré. Pour la première fois, l'homme abandonne sa prérogative cognitive à un élément exogène. Ce basculement a ainsi pu justifier, selon certains, que soit acté le passage à une nouvelle époque, « l'âge de l'intelligence artificielle » [3]. Après l'âge de la religion, marqué par un monopole ecclésiastique de l'information, puis l'âge de la raison durant lequel l'humanité a pris conscience de sa capacité à penser le monde par le doute et la recherche de vérité, serait venu le temps d'une intelligence annexe remettant en cause la matrice cartésienne *cogito ergo sum*, « je pense donc je suis ».

Il s'agit donc désormais d'appréhender cette technologie pensante afin de fabriquer, contrôler, se défendre ou contre-attaquer. Ainsi, l'intelligence artificielle est en train d'acquérir une place centrale dans les rapports conflictuels interétatiques, au point de devenir un enjeu majeur de géopolitique. Alors que l'équilibre international se dégrade rapidement avec une multiplication des foyers de crise et une brutalisation des rapports de force sur tous les continents, le caractère peu probable d'un conflit direct conventionnel entre grandes puissances impose l'emploi de moyens pernicieux et détournés pour saper les fondations du modèle adverse. Cet impératif trouve naturellement une réponse dans l'usage de l'intelligence artificielle. Autrefois apanage des geeks, l'intelligence artificielle est aujourd'hui un instrument incorporé à la panoplie des outils de la guerre cognitive, informationnelle et cyber. Les acteurs hostiles y ont par conséquent de plus en plus

---

(1) Carl von Clausewitz, De la guerre*, 1832, Flammarion, édition 2014, p. 14.*

(2) Aurélie Jean, « Qu'est-ce que l'intelligence artificielle ? »*, Intelligence artificielle : quel progrès ?, Cahiers français, septembre – octobre 2024, p. 18.*

(3) Henry Kissinger, Eric Schmidt, Daniel Huttenlocher, The age of AI, éditions Hodder&Stoughton Libri 2022, 178 pages.

recours pour leurs opérations afin de toucher des publics élargis, avec une rapidité d'exécution augmentée et une réalisation sensiblement améliorée. L'utilisation de l'intelligence artificielle dans le cadre de campagnes de déstabilisation est en hausse, permettant, selon les mots du ministre de l'Europe et des affaires étrangères, à certains régimes autoritaires de « *pilonner par la désinformation* » (1) nos sociétés.

Cette irruption de l'intelligence artificielle dans les opérations d'ingérence étrangère constitue une révolution en raison des changements induits : un saut de productivité inédit, une explosion simultanée de la qualité et de la quantité des contenus créés, ainsi qu'un véritable effondrement du prix de production permettant une adoption rapide et massive de la technologie. Entendue en tant qu'« *agissement commis directement ou indirectement à la demande ou pour le compte d'une puissance étrangère et ayant pour objet ou pour effet, par tout moyen, y compris par la communication d'informations fausses ou inexactes, de porter atteinte aux intérêts fondamentaux de la Nation, au fonctionnement ou à l'intégrité de ses infrastructures essentielles ou au fonctionnement régulier de ses institutions démocratiques* » (2), l'ingérence étrangère est en effet profondément renouvelée par l'utilisation de l'intelligence artificielle. L'intégration croissante de cette technologie aux manœuvres d'ingérence tend à rendre la distinction avec les opérations d'influence encore plus ténue. En effet, « *l'ingérence est une politique d'influence masquée [qui] consiste, pour un État, à mener des actions visant à rendre la politique d'un autre pays structurellement favorable à la sienne, sans que l'on sache d'où parlent les personnes et les organisations auxquelles il a recours* » (3). L'intelligence artificielle permet de brouiller ce *continuum* (4) entre influence et ingérence, la seconde bénéficiant désormais du champ d'action large de la première tout en conservant ses caractéristiques clandestines et secrètes.

Si certains préfèrent évoquer une simple évolution d'un phénomène quasi immémorial, les rapporteurs sont convaincus que les mutations des pratiques malveillantes récemment observées constituent une véritable rupture avec le passé. La stratégie nationale du renseignement, mise à jour au début de l'année 2025, appelle ainsi à s'approprier l'intelligence artificielle et son pendant, les technologies quantiques. Toutes les deux sont clairement identifiées comme des innovations

---

(1) Propos de Jean-Noël Barrot, ministre de l'Europe et des affaires étrangères, dans le cadre du documentaire réalisé par Elsa Guiol, La fabrique du mensonge, Sur la piste des agents de Poutine, *France Télévisions, 2025.*

(2) Article L. 562-1 du code monétaire et financier modifié par la loi n° 2024-850 du 25 juillet 2024 visant à prévenir les ingérences étrangères en France.

(3) Propos de Nicolas Lerner, directeur général de la direction générale extérieure (DGSE) alors directeur général de la direction générale de la sécurité intérieure (DGSI), au cours de son audition par la commission d'enquête de l'Assemblée nationale relative aux ingérences politiques, économiques et financières de puissances étrangères – États, organisations, entreprises, groupes d'intérêts, personnes privées – visant à influencer ou corrompre des relais d'opinion, des dirigeants ou des partis politiques français, 2 février 2023.

(4) Frédéric Charillon, Guerres d'influence : les États à la conquête des esprits, *Odile Jacob, janvier 2022, 352 pages.*

disruptives duales, représentant à la fois une opportunité et une source de vulnérabilités pour le pays [1].

La France est, à cet égard, l'une des cibles privilégiées de tentatives d'ingérence étrangère assistées par l'intelligence artificielle. Cette menace, encore embryonnaire il y a quelques années, se matérialise aujourd'hui par des opérations d'influence numérique, de manipulation de l'information, et plus généralement par des stratégies de déstabilisation globale visant à altérer la confiance dans les institutions, à fragmenter l'opinion publique ou à influencer les processus électoraux. La France est deuxième pays en Europe le plus visé par ces agressions, après l'Ukraine [2], et ce pour trois raisons : « *Premièrement, la France reste, sur la scène internationale, une grande puissance dont la voix porte. Membre permanent du Conseil de sécurité des Nations unies, État doté, la France promeut par ailleurs un modèle démocratique* […]. *Deuxièmement, notre territoire accueille des communautés étrangères et des diasporas d'origines variées qui, depuis le territoire national, se livrent à une activité politique, parfois d'opposition au régime du pays dont ils sont originaires* […]. *Troisièmement, notre pays demeure une grande puissance dans le domaine de l'économie et de la recherche.* » [3].

Loin d'un discours moralisateur, ce rapport d'information s'attache à alerter sur la nature du changement de paradigme induit par le recours à cette nouvelle technologie dans le cadre d'opérations d'ingérence ainsi que sur la nécessité de prendre toute la mesure du danger qui pèse sur notre modèle de société. Il convient désormais de décliner cette nouvelle grammaire de l'intelligence artificielle évoquée par le président de la République [4] afin de pouvoir faire front. Notre réponse doit rester fidèle à notre identité : une démocratie vigilante et lucide qui assume ses fragilités en refusant de céder à la tentation de la dictature numérique.

---

(1) Coordination nationale du renseignement et de la lutte contre le terrorisme, Stratégie nationale du renseignement*, janvier 2025, p. 34.*

(2) Anne-Marie Descôtes, Secrétaire générale du ministère de l'Europe et des affaires étrangères, discours à l'occasion de l'évènement « Le Quai d'Orsay face à la guerre informationnelle », 9 septembre 2025.

(3) Nicolas Lerner, Op. Cit.

(4) Discours du président de la République Emmanuel Macron #AIforhumanity au Collège de France, 29 mars 2018.

I. L'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE REDÉFINIT LES OPÉRATIONS D'INGÉRENCES ÉTRANGÈRES EN MENAÇANT REDOUTABLEMENT NOS MODÈLES DÉMOCRATIQUES

A. L'ESSOR DE L'INTELLIGENCE ARTIFICIELLE CONSTITUE UNE RÉVOLUTION TECHNOLOGIQUE MAJEURE

1. Discipline et technologie, l'intelligence artificielle a connu des développements irréguliers depuis les années 1950

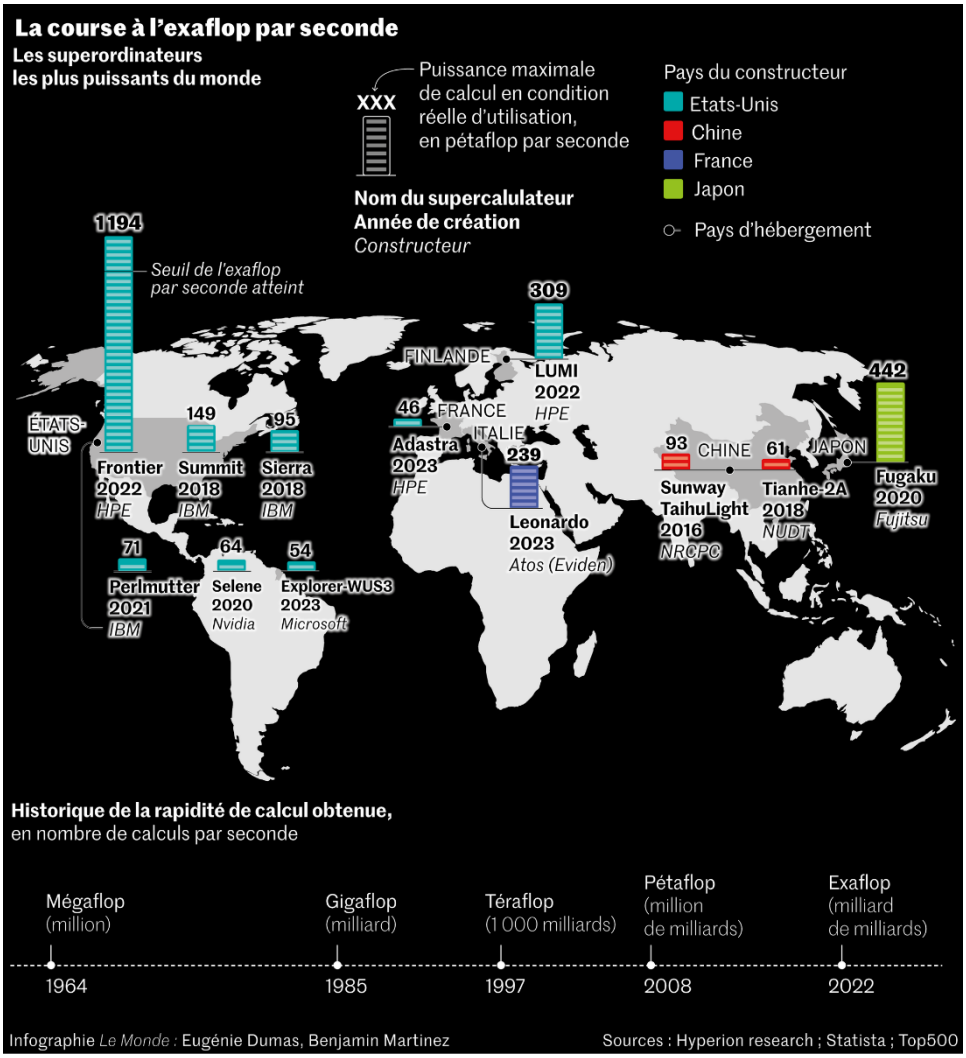
La genèse de l'intelligence artificielle remonte à la seconde guerre mondiale. Dès les années 1940, l'humanité assiste au premier affrontement tactique entre deux nations par machines interposées. Le britannique Alan Turing conçoit un automate pour décrypter les messages de l'appareil allemand Enigma. Le calculateur de Turing décompose un raisonnement logique en une série d'actions mécaniques pouvant être automatisées. C'est la première fois qu'un outil assiste l'humain, non pas dans la réalisation d'une tâche manuelle mais pour celle d'une activité intellectuelle. Par la suite, la première application grand public de l'intelligence artificielle prend la forme d'une calculatrice. Elle est ensuite mobilisée par des éditeurs de jeux vidéo et nourrit, déjà, des tensions commerciales entre l'entreprise californienne Atari et le développeur japonais Taito. La technologie est reprise avec l'essor d'internet pour alimenter les débuts de l'informatique, de l'ordinateur italien Olivetti à la machine américaine d'*International Business Machines* (IBM) en passant par les modèles du britannique Amstrad ou du français Thomson.

Le concept d'intelligence artificielle n'est toutefois formulé explicitement qu'en 1955 au cours des travaux préparatoires à un atelier scientifique, la conférence de Dartmouth à l'été 1956. Les chercheurs de l'époque, John McCarthy, Marvin Minsky, Claude Shannon et Nathaniel Rochester s'y retrouvent avec pour objectif la modélisation du comportement cérébral humain. Afin de faire évoluer cette discipline naissante, sont évoqués les sciences de l'information, l'apprentissage machine, le calcul computationnel (1), l'analyse sémantique du langage naturel, la modélisation par réseaux neuronaux, l'abstraction, l'algorithmique et la créativité (2). Les avancées de la fin du XX^{e} siècle demeurent bornées à la sphère expérimentale, faute de financements massifs et de percée majeure en termes de calcul. Les années 2000 relancent la course à l'intelligence artificielle grâce notamment au développement des supercalculateurs. La puissance de ces derniers a en effet doublé chaque année au cours des quatre dernières décennies. Les plus performants sont désormais capables de traiter plus d'un milliard de milliards d'opérations par seconde.

---

(1) Le calcul computationnel désigne l'ensemble des méthodes, modèles et procédures permettant à une machine d'effectuer automatiquement des opérations logiques ou mathématiques pour résoudre un problème.

(2) Aurélie Jean, Op. Cit.



En pratique, la mise en œuvre de l'intelligence artificielle se traduit par la construction d'un algorithme qui est entraîné et calibré avec un jeu de données en vue d'une future exécution sur de nouvelles informations fournies ultérieurement [(1)]. Parmi ces paramètres d'entraînement, l'algorithme identifie et capture des signaux forts et faibles, les *patterns*, qui correspondent à des caractéristiques du phénomène observé. Sont alors distingués deux types d'algorithmes au regard de la logique qui les sous-tend. Les algorithmes explicites comprennent une séquence d'opérations expressément définie par le concepteur. Ces algorithmes remontent à l'Antiquité et permettent par la répétition mécanique d'une action de déterminer un résultat. Une telle méthode d'identification des constantes apparaît cependant limitée si le but de l'opération est, par exemple, de reconnaître un animal sur une photographie. Là où la seule mathématisation est inefficace, une autre forme d'algorithme dit implicite est alors nécessaire. Elle suppose un apprentissage automatique de la machine sur un jeu de données préalable. Au demeurant, les modèles d'intelligence artificielle modernes mobilisent généralement les deux formes d'algorithmes selon un format hybride.

(1) Ibid.

2. Les avancées récentes en matière d'intelligence artificielle consacrent un changement de paradigme en démocratisant le recours à cette technologie

L'apprentissage automatique, indispensable pour faire tourner les algorithmes implicites et hybrides, a connu une avancée majeure au cours des années 2010. Les progrès de la recherche sur une nouvelle technologie appelée apprentissage profond, *deep learning*, révolutionnent le fonctionnement de l'intelligence artificielle. Jusqu'alors fondés sur l'apprentissage automatique, *machine learning*, par l'application de modèles mathématiques et du traitement des jeux de données précités, les appareils bénéficient désormais de réseaux de neurones artificiels permettant la résolution de problèmes complexes tels que la reconnaissance des formes ou le traitement du langage naturel (1). Fort de ce système, le projet de *deep learning* Google Brain de la firme américaine éponyme a étudié, en 2012, des millions de captures d'écran choisies aléatoirement. Au terme de l'analyse, l'algorithme a découvert le concept de chat lui-même, sans que personne ne lui a jamais dit ce qu'était un chat (2). Cette nouvelle capacité de traitement des données a permis au programme informatique *AlphaGo* de battre avec fracas, en 2016 et 2017, les meilleurs joueurs du monde du jeu de go. Si dès 1997, l'ordinateur *Deep Blue* avait battu le champion d'échecs Garry Kasparov, cette victoire de la machine au jeu de go revêt une signification différente au vu des conditions de succès plus hétérogènes (3) et du nombre de combinaisons possibles bien plus important. En outre, le style de jeu retenu par *AlphaGo* est entièrement le fruit de l'entraînement de l'intelligence artificielle et ne reprend aucun coup ou stratégie préprogrammée ou dérivée du jeu humain.

Ces développements ont permis l'essor d'un sous-segment applicatif du *deep learning*, les technologies d'intelligence artificielle générative (IAg). En rendant possible l'édition de textes, d'images ou de codes informatiques, elles actent un changement d'échelle dans la génération de contenus (4). L'arrivée sur le marché, en 2023, de l'agent conversationnel (*large language model*, LLM) conçu par l'américain OpenAI, ChatGPT, rend cette technologie accessible facilement et gratuitement pour le grand public. Ce *chatbot* a par ailleurs établi un record en devenant le logiciel grand public à la croissance la plus rapide de l'histoire, atteignant cent millions d'utilisateurs seulement soixante-quatre jours après sa sortie (5). L'IAg s'est ainsi imposée dans le quotidien d'environ un demi-milliard de personnes, dans 209 pays et pour près d'un travailleur sur huit (6). Cette diffusion

---

(1) Service de vigilance et de protection contre les ingérences numériques étrangères (Viginum), Défis et opportunités de l'intelligence artificielle dans la lutte contre les manipulations de l'information, Secrétariat général de la défense et de la sécurité nationale (SGDSN), 2025, p. 5.

(2) Morgane Tual, « Comment le « deep learning » révolutionne l'intelligence artificielle », *Le Monde, 24 juillet 2015.*

(3) Au jeu de go, il s'agit en effet de contrôler des zones et non d'éliminer spécifiquement une pièce du plateau.

(4) Viginum, Op. cit.

(5) Yan Liu, He Wang, Banque mondiale, « Who on Earth is using generative AI », *août 2024, p. 16.*

(6) Ibid, p. 24. – Sur les 218 économies nationales identifiées par la Banque mondiale, neuf pays n'ont donc pas été retenus, principalement des économies fragiles, en proie à des conflits et à la violence, pour lesquelles aucune donnée n'était disponible : Cuba, la République islamique d'Iran, la République populaire

reprend tous les éléments d'une révolution industrielle majeure : hausse de la quantité produite, augmentation notable de la qualité finale, prix bas assurant une distribution globale. L'intelligence artificielle n'a donc pas vocation à être une énième brique se superposant à l'amas numérique. Elle « *n'est pas seulement un outil ou une plateforme, c'est une « métatechnologie » transformatrice, la technologie derrière la technologie et tout le reste, elle-même productrice d'outils et de plateformes ; ce n'est pas seulement un système, c'est un générateur de systèmes de toute nature* » (1).

3. Au cœur de logiques géopolitiques, le développement de l'intelligence artificielle s'inscrit dans une quête étatique de la puissance

Tous les domaines d'activité où l'analyse de données permet d'automatiser des décisions – finances, gestion des flux urbains, énergie, médecine, entre autres – mobilisent désormais l'intelligence artificielle. Le secteur de la défense n'est pas en reste et a également pris la pleine mesure de la potentialité offerte par cette innovation : utiliser l'intelligence artificielle en tant qu'instrument d'aide à la décision voire comme un outil de planification d'opérations militaires. Ainsi, le recours à cette technologie peut notamment faciliter l'analyse en temps réel des zones de conflit, le guidage des missiles, la coordination des essaims de drones ou encore la conception d'armes intelligentes, létales et autonomes (2).

Cristallisant des enjeux financiers, scientifiques, sécuritaires et culturels, la maîtrise de l'intelligence artificielle et de ses composantes est par conséquent devenue un impératif géopolitique pour les États. Cette réalité fut clairement énoncée, dès 2017, par le président de la fédération de Russie, Vladimir Poutine : « *L'intelligence artificielle représente l'avenir, non seulement pour la Russie mais pour l'humanité toute entière* […] *La nation qui sera leader dans le domaine de l'intelligence artificielle dominera le monde* » (3). Annonçant une compétition mondiale, les mots de Moscou ne préfiguraient pas un développement pacifique et collaboratif de la ressource IA mais laissaient présager une approche davantage conflictuelle. La confrontation est d'ores et déjà en cours et s'exprime au travers d'un nouveau concept, celui de puissance en matière d'IA.

La puissance est la capacité d'imposer sa volonté aux autres (4). Couplée aux caractéristiques de l'IA, soit un répertoire de techniques visant à accroître, par-delà les capacités humaines, la productivité dans un nombre indéfini d'activités, par la substitution tendanciellement infinie de machines aux êtres humains, cette notion de puissance en matière d'IA désigne tout acteur, étatique ou non, en mesure de

---

démocratique de Corée, le Kosovo, la Libye, la Birmanie, la Namibie, le Soudan et la République arabe syrienne.

(1) Mustafa Suleyman, Michel Bhaskar, The coming wave: Technology, power and the twenty-first century's greatest dilemma*, Crown, 2023, 352 pages.*

(2) Bernard Benhamou, « Les nouveaux enjeux géopolitiques de l'intelligence artificielle », *Intelligence artificielle : quel progrès ?, Cahiers français, septembre – octobre 2024, p. 26.*

(3) Ibid.

(4) Raymond Aron, Paix et guerre entre les nations*, Paris, Calmann Lévy, 1962, édition 2004, 832 pages.*

modifier l'action, le comportement et la perception d'autres acteurs grâce à des gains de productivité liés au développement de dispositifs sociotechniques autonomes.

Pour établir une hiérarchie des puissances en matière d'IA, il faudrait donc différencier le niveau d'influence de chaque acteur. Si l'on se fie aux discours politiques et aux stratégies nationales pour l'intelligence artificielle – soixante-quinze pays en ont publié une à l'heure actuelle –, deux pays sont à nouveau au centre de l'attention : les États-Unis et la Chine. De fait, rien d'étonnant, ce sont les deux premières puissances militaires et économiques, et chacune a désigné l'autre comme son grand rival stratégique. De plus, dans la mesure où, d'un côté, le champ du pouvoir américain considère que la puissance sur la scène internationale repose éminemment sur sa domination technologique, et que, de l'autre, l'Exécutif chinois estime que la domination en IA de la Chine est la principale condition de l'ascension du pays à la tête du système international, il est aisément admis que ces deux États constituent aujourd'hui les deux premières puissances en la matière.

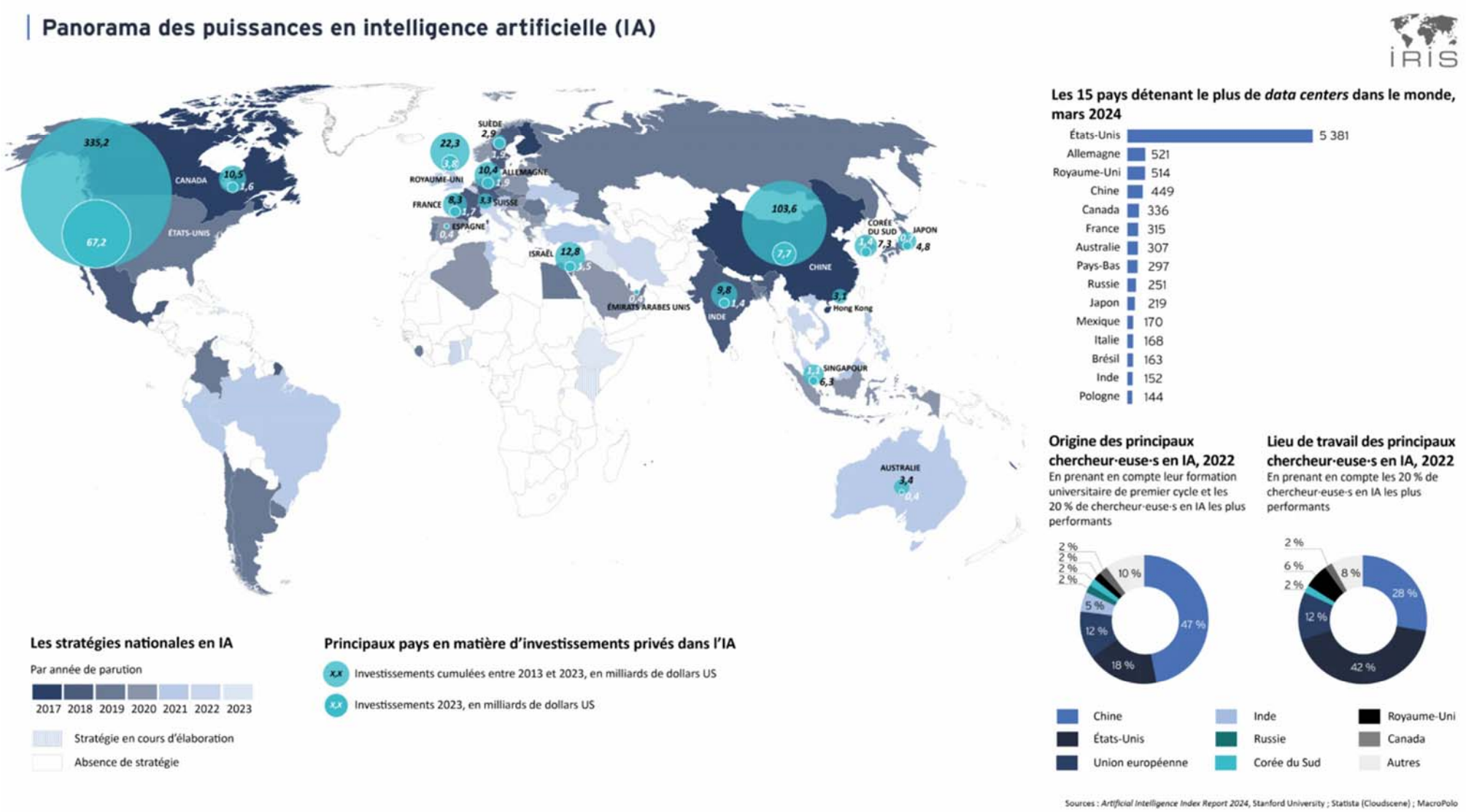
L'intelligence artificielle est ainsi devenue l'une des composantes majeures de l'affrontement entre Washington et Pékin. S'exprimant en premier lieu sur le plan économique, ce face-à-face géopolitique combine mesures offensives et stratégies proactives de développement. Dès 2022, l'administration Biden a limité les investissements américains dans les sociétés chinoises d'IA avant d'imposer un quasi-embargo sur les exportations à destination de la Chine comprenant des technologies liées à l'IA. En rétorsion, Pékin a réduit ses propres exportations de métaux critiques, le gallium et le germanium, dont l'utilisation est primordiale dans la fabrication de composants électroniques. Au demeurant, les limitations imposées aux Chinois ont paradoxalement encouragé Pékin à accélérer sa quête d'indépendance technologique. Privés des puces Nvidia, les ingénieux chinois ont dû être plus innovants. Leurs efforts semblent avoir été récompensés avec l'apparent succès de leur modèle conversationnel, Deepseek R1, aux capacités similaires à celles du concurrent américain, ChatGPT, le tout pour un coût d'entraînement présenté comme nettement inférieur (1).

En parallèle, chacun des deux pays s'est attaché à investir massivement pour soutenir l'essor de l'IA et apparaître comme le futur *leader* du secteur. Au lendemain de sa seconde investiture, Donald Trump a ainsi annoncé, le 21 janvier 2025, un investissement massif de 500 milliards de dollars au cours des quatre prochaines années et abondé par des acteurs privés. *Stargate*, ce projet au nom tiré de l'univers de la science-fiction, ambitionne de bâtir une infrastructure globale pour l'IA du futur avec notamment la construction d'immenses centres de données au Texas. De l'autre côté du Pacifique, Pékin n'est pas en reste. Au cours de la décennie 2013-2023, ce sont près de 104 milliards de dollars qui ont été injectés dans des firmes du secteur de l'IA par les fonds de capital-risque

(1) OpenAI aurait dépensé près de 3 milliards de dollars pour entrainer son agent conversationnel ChatGPT en 2024 ; Deepseek affirme avoir fait la même chose pour 6 millions de dollars. Des doutes sont toutefois émis sur ce résultat en raison du manque de transparence des autorités chinoises et de la probable acquisition de puces Nvidia par des réseaux secondaires.

appartenant au gouvernement chinois. Ce dernier vise à faire du pays le premier centre d'innovation mondiale en IA à l'horizon 2030. Cette stratégie repose avant tout sur des « plateformes d'innovation », sectoriellement différenciées, à la tête desquelles le gouvernement central chinois a placé certaines de ses principales firmes technologiques. Ainsi, Baidu est en charge de la conduite autonome ; Alibaba des *smart cities* ; Tencent de l'imagerie médicale ; iFlytek de « l'intelligence audio » ; SenseTime de la « vision intelligente ». Au total, ce sont quinze entreprises chinoises qui ont été placées à la tête de programmes de développement de l'IA à partir de 2017.

Afin de ne pas se faire happer par l'affrontement sino-américain, d'autres pays tentent de faire émerger des modèles IA concurrents. Toutefois les investissements apparaissent dérisoires face aux sommes engagées par Washington et Pékin. Les États-Unis et la Chine sont ainsi suivis lointainement par Israël, le Canada, l'Allemagne, l'Inde, la France, la Corée du Sud, le Japon et les pays du Golfe. Ces différences de moyens nourrissent une forte migration des talents orientée vers les États-Unis, où une majorité des ingénieurs et des chercheurs sont étrangers. Enfin, certains États ne se sont pas lancés dans la course au développement de l'IA, par la production de modèles génératifs ou une recherche intensive, mais l'utilisent tout de même massivement pour d'autres objectifs. La Russie, l'Iran et la Corée du Nord y ont ainsi recours à des fins purement militaires ou pour assurer un contrôle toujours plus poussé de leur population.



Source : Standford University, Artificial intelligence report 2024 – carte réalisée par l'Institut de relations internationales et stratégiques (IRIS), 13 février 2025.

B. LE RECOURS À L'INTELLIGENCE ARTIFICIELLE AMPLIFIE ET DIVERSIFIE LES ATTAQUES DANS LE CADRE D'UNE GUERRE HYBRIDE

1. De la guerre pour l'intelligence artificielle à la guerre par l'intelligence artificielle, les ingérences étrangères trouvent un terreau fertile dans un environnement géopolitique marqué par des tensions exacerbées

« *La guerre ne se déclare plus, elle se mène à bas bruit, insidieusement, elle est hybride* » (1). Quelques mois avant la présentation du projet de loi relatif à la programmation militaire pour les années 2024 à 2030 par le gouvernement, le président de la République affirmait avec force la nouvelle réalité stratégique imposée à la France et à ses partenaires. Cet état de guerre a en effet également été mentionné au cours de l'ensemble des auditions menées par les rapporteurs aux États-Unis et en Suède. Hybride, ce format synthétise les deux modalités classiques du concept : la guerre conventionnelle et la guerre irrégulière (2). Ce syncrétisme est protéiforme et inclut désormais des phénomènes comme le terrorisme ou la criminalité organisée mais aussi des moyens non liés à la violence physique comme la cyberguerre, la propagande ou la guerre économique (3). La guerre hybride permet ainsi de maximiser la déstabilisation politique et psychologique à un coût financier, militaire et diplomatique moindre. Pour ce faire, elle exploite les seuils de riposte – notamment le seuil de l'agression armée et celui de l'article 5 de l'Organisation du traité de l'Atlantique Nord (OTAN) – en transformant l'ambiguïté en instrument stratégique, freinant la prise de décision, divisant les opinions publiques et paralysant la réaction. En outre, elle rend plausible le déni d'attribution en multipliant les signaux contradictoires pour semer le doute (4).

Dans ce contexte, l'arrivée de l'intelligence artificielle démultiplie les menaces de chaque aspect de la guerre hybride. Sur le plan conventionnel, des systèmes embarqués sont d'ores et déjà équipés et utilisés sur le terrain. Les conflits ouverts ont ainsi été marqués par le recours à cette nouvelle technologie, au point que le magazine américain *Time* titrait récemment sa une « *La première guerre de l'IA* » (5) pour évoquer les combats en Ukraine. De même, à Gaza, l'armée israélienne s'appuie sur plusieurs systèmes pour guider ses frappes, en identifiant rapidement les lieux et cibles *via* notamment son logiciel IA *Lavender* (6).

En parallèle, depuis 2022 et l'invasion russe en Ukraine, les modes d'action observés sont coordonnés et mêlent des ingérences numériques aux activités cyber

---

(1) Emmanuel Macron, Vœux aux armées, Mont-de-Marsan, 20 janvier 2023.

(2) Voir notamment Elie Tennenbaum, Institut français des relations internationales (IFRI), « Le piège de la guerre hybride »*, Focus stratégique, octobre 2015, p. 13.*

(3) Ibid,, p. 8.

(4) Louise Souverbie, IRIS, « Incursions russes et guerre hybride : l'Europe sous pression aérienne »*, 26 septembre 2025.*

(5) Vera Bergengruen, « The first AI war »*, Time, 8 février 2024.*

(6) Laure de Roucy-Rochegonde,, « Guerre à Gaza : l'IA au service des frappes israéliennes »*, IFRI, 20 avril 2024.*

ou physiques. Les attaques informationnelles servent alors souvent à amplifier l'autre mode opératoire employé. Une nette intensification de ces campagnes et une diversification des acteurs impliqués, qui ne sont plus seulement des acteurs étatiques est également observée. L'emploi de proxies ou de groupes crapuleux se généralise ou, du moins, a tendance à être toléré voire encouragé par certains régimes. Le groupe cybercriminel nord-coréen *Lazarus* (APT38), leurs homologues chinois *Salt Typhoon* ou encore le gang russe *Black Basta* sont autant d'exemples de cette nouvelle réalité paraétatique. Certains sont plus retors et se cachent derrière des structures légales : en témoignent les campagnes d'ingérences menées contre la France et ses territoires ultra-marins par le *Baku Initiative Group*, un organe de désinformation structuré sous la forme d'une organisation non gouvernementale (ONG) et impliquant des individus proches du pouvoir politique azerbaïdjanais (1). L'intelligence artificielle permet à ces acteurs aux moyens limités d'opérer à un nouveau niveau sur une échelle plus importante. Au demeurant, la diffusion de cette technologie brouille encore davantage la démarcation entre les sphères civiles et militaires et contribue ainsi à une massification des théâtres d'opération.

En raison des prises de position françaises sur la scène internationale ces dernières années – soutien affirmé à l'Ukraine, défense d'un *statu quo* à Taïwan, promotion d'un modèle démocratique ouvert et inclusif – notre pays est particulièrement visé par des actions d'ingérence provenant principalement et vraisemblablement d'acteurs russes et chinois. Une influence agressive assumée, à la frontière de l'ingérence, est également relevée de la part de certains courants américains au travers de la mouvance *Make America great again* (MAGA). Première menace, la Russie a particulièrement redoublé ses efforts en déployant toute la palette de l'ingérence. Depuis 2022, ce sont ainsi plus de soixante opérations de guerre hybride (2) qui ont été menées par Moscou sur le vieux continent. L'intelligence artificielle a été utilisée intensivement pour amplifier notamment les attaques cyber et les tentatives de désinformation.

De la même manière, depuis les manifestations de mars 2019 à Hongkong, Pékin s'appuie sur un dispositif d'influence alimenté par IA dont l'objectif est de diffuser des narratifs favorables aux intérêts du Parti communiste chinois (PCC) auprès d'une audience internationale. Ce mode opératoire, appelé « *spamouflage* », repose sur des réseaux de comptes aux caractéristiques inauthentiques chargés de mener des manœuvres informationnelles sur une multitude de plateformes (3). Plusieurs actions semblent avoir été menées par ce biais notamment à l'occasion de la tenue des Jeux olympiques et paralympiques de 2024 en France. En outre, le fonctionnement obscur de l'application TikTok et ses liens avec l'État-parti, les ingérences manifestes du PCC lors des élections présidentielles taïwanaises de

---

(1) Viginum, UN-notorious BIG, Une campagne numérique de manipulation de l'information ciblant les DROM-COM et la Corse, *SGDSN, décembre 2024, p. 2.*

(2) Données Die Zeit « Wie Russland einen hybriden Krieg in Europa führt », *23 décembre 2024 ; carte réalisée par le Grand continent, 8 janvier 2025.*

(3) Viginum, Défis et opportunités de l'intelligence artificielle dans la lutte contre les manipulations de l'information, *SGDSN, 2025, p. 7.*

janvier 2024 [1] ainsi que les manœuvres tendancieuses du réseau diplomatique envers sa diaspora à l'étranger sont autant d'éléments troublants où l'intelligence artificielle a été mobilisée.

Enfin, sans être exhaustif, le retour aux affaires du président Trump s'est accompagné d'une hausse des actions offensives d'influence en flirtant avec l'ingérence ou en basculant dans le champ de la provocation directe. Sans atteindre le niveau de malveillance russe ou chinois, la sphère MAGA n'a eu de cesse de relayer de fausses informations et d'attaquer les modèles démocratiques français et européens. Le rachat du réseau social X, anciennement Twitter, par le techno-oligarque Elon Musk a démontré l'importance de l'intelligence artificielle et de la pratique des algorithmes dans la diffusion de ces narratifs hostiles regroupés sous la bannière *Make Europe great again* (MEGA) [2]. Dans ce contexte, la charge du vice-président américain, J.D Vance, contre les positions européennes lors du sommet de Paris pour l'action sur l'intelligence artificielle, n'est pas anodine.

Si le degré d'agressivité, les modalités et l'intensité de l'action de ces trois États varient, ils s'inscrivent dans une pratique décomplexée du « *sharp power* » [3]. Notion élaborée par deux chercheurs, Christopher Walker et Jessica Ludwig, le « pouvoir acéré » se situe dans une zone grise entre les concepts de *hard power*, la manifestation concrète de la puissance militaire pour imposer sa volonté à un autre corps politique, et celui de *soft power*, la capacité d'un État à influencer indirectement le cours des relations internationales [4]. Contrairement à ses deux *alter ego*, le *sharp power* ne vise pas à promouvoir un modèle et des valeurs ou à employer la force brute mais consiste, au contraire, à nuire au modèle adverse, en l'affaiblissant et en le décrédibilisant de l'intérieur. Le fer de lance du *sharp power* est ainsi l'ingérence numérique par le recours aux réseaux sociaux et, désormais, l'utilisation de l'IA.

Or, en parallèle, les moyens publics et privés destinés à lutter contre ces ingérences ont été réduits. Alors que les principaux réseaux sociaux utilisés dans le monde sont détenus par des entreprises américaines, le cas du groupe Meta est édifiant. Le 7 janvier 2025, son dirigeant, Mark Zuckerberg, a annoncé la fin du système de vérification de l'information (*fact checking*) en invoquant une réponse à la censure chinoise et européenne. Sous couvert de protection de la liberté d'expression, des centaines de modérateurs ont été licenciés et remplacés par des notes de contexte rédigées par les utilisateurs, copiant ainsi le modèle mis en place par Elon Musk sur la plateforme X. Au cours de son déplacement aux États-Unis en

---

(1) Florence de Changy, « Selon le ministre des affaires étrangères taïwanais, l'ingérence chinoise est de plus en plus sophistiquée »*, Le Monde, 2 janvier 2024.*

(2) Commission des affaires étrangères, Assemblée nationale, Table ronde sur les ingérences étrangères dans les processus démocratiques, avec la participation de Mme Claire Benoit, cheffe du bureau Coordination et stratégie de Viginum, M. David Colon, professeur agrégé d'histoire à l'Institut d'études politiques de Paris, et M. Bernard Benhamou, secrétaire général de l'Institut de la souveraineté numérique, 5 février 2025.

(3) Christopher Walker, Jessica Ludwig, « The meaning of sharp power: How authoritarian States project influence »*, Foreign Affairs, 16 novembre 2017.*

(4) Joseph Nye, Soft power : the means to success in world politics, Public affairs, avril 2005, 208 pages.

juillet 2025, la rapporteure s'est également inquiétée des conséquences des réformes mises en place par l'administration Trump sur ces sujets. Plusieurs institutions fédérales qui contribuaient à lutter contre les ingérences étrangères ont ainsi été récemment supprimées ou dépossédées d'une grande partie de leurs moyens :

- le *Global Engagement Center* (Centre pour la mobilisation mondiale), rattaché au département d'État, qui n'avait pas obtenu de financement fin 2024, a été officiellement supprimé le 16 avril 2025 [(1)]. Ce centre, créé en 2016, avait pour mission de détecter, de comprendre, d'exposer et de contrer les actions de désinformation et de propagande étrangères qui ciblaient les États-Unis et leurs alliés. Selon Sarah Rogers, auditionnée par la rapporteure à Washington avant d'être nommée officiellement sous-secrétaire à la diplomatie publique (*under-secretary for public diplomacy*), il pourrait être remplacé par une simple coopération interministérielle ;

- la *Foreign Influence Task Force* (FITF – Groupe de travail sur l'influence étrangère) du *Federal Bureau of investigation* (FBI), en février 2025. Le rôle de cette structure, créée en 2017, était de coordonner les enquêtes sur les opérations d'influence étrangère. Son avenir semble s'inscrire en pointillés et sa directrice, Jessica Brandt, auditionnée par la rapporteure, a été limogée en début d'année ;

- le *Foreign Malign Influence Center* (FMIC – Centre sur les influences étrangères malveillantes), créé en 2022 et rattaché au bureau de la directrice du renseignement national (*Office of the director of national intelligence* – ODNI), apparaissait en sursis lors du déplacement de la rapporteure à Washington du 8 au 11 juillet 2025. Le 20 août 2025, la directrice du renseignement national, Tulsi Gabbard, a finalement annoncé une réduction drastique des missions du Centre, ce dernier ne pouvant être supprimé qu'après accord du Congrès. Le FMIC coordonnait auparavant le renseignement en matière de lutte contre l'influence étrangère malveillante. Il avait notamment permis une surveillance continue des tentatives de piratage et des campagnes de désinformation en ligne, pour les élections américaines de 2022 et de 2024 ;

- la *Cybersecurity and Infrastructure Security Agency* (CISA – Agence de cybersécurité et de sécurité des infrastructures), rattachée au département de la sécurité intérieure (*Department of Homeland Security* – DHS), a vu son champ d'activité restreint, tout comme ses effectifs de personnel, au début de l'année 2025. Le CISA, créé en 2018, est en charge de la protection des infrastructures critiques, y compris les systèmes électoraux, contre les cybermenaces étrangères. En période électorale, il apportait une aide technique aux États fédérés et aux autorités locales.

---

(1) Entre-temps, il avait été rebaptisé Counter Foreign Information Manipulation and Interference Office *(R/FIMI – Bureau de lutte contre les manipulations de l'information et les ingérences étrangères).*

2. Afin de fragiliser les systèmes politiques et sociétaux, le recours à l'intelligence artificielle renforce les tentatives de déstabilisation du modèle occidental

Dans le contexte décrit précédemment, il ressort des auditions menées par les rapporteurs que l'intelligence artificielle est utilisée dans des proportions croissantes principalement pour manipuler l'opinion publique à coups d'attaques informationnelles, pour renforcer des opérations cyber et, enfin, pour empoisonner d'autres modèles d'intelligence artificielle. Le recours à l'IA démultiplie les menaces et engendre une asymétrie préjudiciable à la défense nationale. Il est en effet aisé de produire du contenu en masse tandis que l'identification et la démystification (*debunk*) de ce même contenu requièrent des moyens financiers et humains importants (*cf. infra*). À ce titre, les rapporteurs appellent à une prise de conscience générale quant au statut de « proie » associé à nos systèmes démocratiques.

a. De nombreuses attaques impliquant l'intelligence artificielle ont pour objectif la manipulation des opinions

La diffusion des modèles d'IAg dans la société s'est accompagnée, au cours des deux dernières années, d'une adoption tout aussi généralisée de cette innovation par les différents acteurs étrangers hostiles. Ainsi, trois tendances principales émergent et caractérisent les conséquences de l'usage de cette technologie sur la physionomie de la menace informationnelle actuelle :

- Un changement d'échelle dans la génération de contenus potentiellement inexacts ou trompeurs ;

- Des capacités décuplées pour la réplication et la publication coordonnée de contenus erronés à grande échelle ;

- Une aide pour la génération et la gestion de comptes inauthentiques sur les plateformes en ligne. (1)

Selon un service de renseignement auditionné, l'intelligence artificielle représente par conséquent un « *levier d'intensification des ingérences informationnelles*. En comparaison aux méthodes antérieures, elle facilite l'identification des victimes grâce au ciblage algorithmique. Au-delà, elle rend possible l'édition et la création de contenus de propagande crédibles à moindre coût – textes, sons, vidéos – vers un public large ou en les adaptant à une audience spécifique. Dans le second cas, l'intelligence artificielle produit un narratif individualisé et pertinent, en suivant l'instruction de commande (*prompt*) tout en mobilisant des données publiques sur Internet, notamment celles laissées par les utilisateurs des réseaux sociaux eux-mêmes. Ces contenus peuvent être aisément traduits dans d'autres langues, afin d'atteindre de nouveaux pays ou des communautés linguistiques spécifiques au sein d'un même territoire, notamment les

(1) Viginum, Op. Cit.

diasporas. Les modèles d'IAg facilitent également la reformulation de messages portant sur une thématique précise, afin d'amplifier leur visibilité tout en conservant une certaine discrétion, *via* la technique du « copy-pasta » (1).

Les campagnes de désinformation sont de fait moins faciles à détecter que par le passé. En parallèle, cette difficulté est amplifiée par la diffusion autorisée de contenus synthétiques, sous des formats humoristiques, appelés « *soft fake* ». Microsoft a ainsi constaté une nette augmentation de ce type de publications lors de l'élection présidentielle américaine de 2024.

Exemple de soft fake caricaturant le président de la République.

De manière générale, l'amélioration récente des IAg s'est traduite par la production d'hypertrucages (*deepfakes*) textuels, audios et vidéos (2). De plus en plus réalistes, certains *deepfakes* textuels sont aujourd'hui quasiment indétectables.

Deepfake relayé par des trolls et mettant en scène le président de la République (3).

---

(1) Le « copy-pasta » *consiste à imiter un texte ou un visuel en le copiant-collant à l'identique ou presque, généralement en le sortant de son contexte d'origine.*

(2) Le « deepfake » désigne « un contenu image, audio ou vidéo généré ou manipulé par l'IA qui ressemble à des personnes, des objets, des lieux, des entités ou des événements existants et qui semblerait faussement authentique ou véridique aux yeux d'une personne » – *Règlement (UE) 2024/1689 relatif à l'intelligence artificielle (AI Act), article 3.*

(3) Dounia Mahieddine, Agence France-Presse (AFP), « Non, Macron n'a pas lancé un programme offrant 1 800 euros mensuels aux Africains s'installant en France »*, 22 août 2025.*

Or, ces contenus parviennent souvent à créer un engouement, à engendrer du « *buzz* » et à générer de l'engagement. Progressivement, les personnes ciblées sont enfermées dans des « bulles informationnelles » artificielles, destinées à modifier leur opinion. Les victimes passent alors d'une perception de la réalité observée à celle d'une réalité relativisée voire d'une réalité alternative.

Ces productions générées par l'IA sont également diffusées grâce à cette technologie. L'intelligence artificielle permet en effet d'animer de faux comptes − voire même de les créer *ex nihilo* (1) − de manière automatisée et industrielle. Dans un tel contexte, l'exploitation de « fermes à trolls », qui emploient habituellement des personnes physiques pour diffuser des éléments de propagande, apparaissent de moins en moins nécessaires, réduisant le besoin en main d'œuvre humaine. L'identification d'un lieu physique de production, à l'instar de la première ferme à trolls russe, l'*Internet Research Agency* fondée par l'ancien chef du groupe Wagner, Evgueni Prigojine, est également plus complexe.

Au-delà, cette diffusion de fausses informations joue sur des temporalités opposées. En profitant de l'instantanéité des réseaux sociaux et de la rapidité de production assurée par l'IA, ces contenus permettent aux acteurs malveillants de regagner une maîtrise du temps long, la ressource temporelle humaine étant désormais préservée par l'automatisation. L'IA octroie alors une capacité d'érosion, extrêmement dommageable à terme et plus difficilement attribuable à un acteur précis. Les effets des campagnes d'ingérence ne se font ainsi ressentir que plusieurs mois voire années après leur réalisation. Ainsi, lors des échanges de la rapporteure à Taïwan, le *think-tank* DoubleThinkLab relevait que la perception des jeunes taïwanais tendait à évoluer vers une plus grande adhésion aux modèles chinois, et ce notamment chez les utilisateurs de la plateforme TikTok. Ce premier résultat positif pour le PCC vient justifier la théorisation, dès 2003, du concept des trois guerres non cinétiques (2) : la guerre de l'opinion publique, la guerre psychologique et la guerre du droit (3). Selon cette approche holistique, la Chine peut séduire et subjuguer tout en infiltrant et contraignant.

En parallèle, l'IA peut également être mobilisée en vase clos pour des phases de préparation préalable. Le *think-tank* taïwanais mentionné précédemment, DoubleThinkLab, et une équipe d'universitaires ont ainsi mis en évidence la possibilité d'anticiper des schémas de réponse à un fait donné dans un environnement hermétique et automatisé, appelé « *socioverse* ». Dans ce cas, l'utilisation de l'IA assure le traitement d'une multitude de scénarios potentiels, et dégage une issue probable. Au-delà, cette méthodologie permet d'analyser les conséquences d'actions prises au niveau micro dans l'émergence d'un modèle de

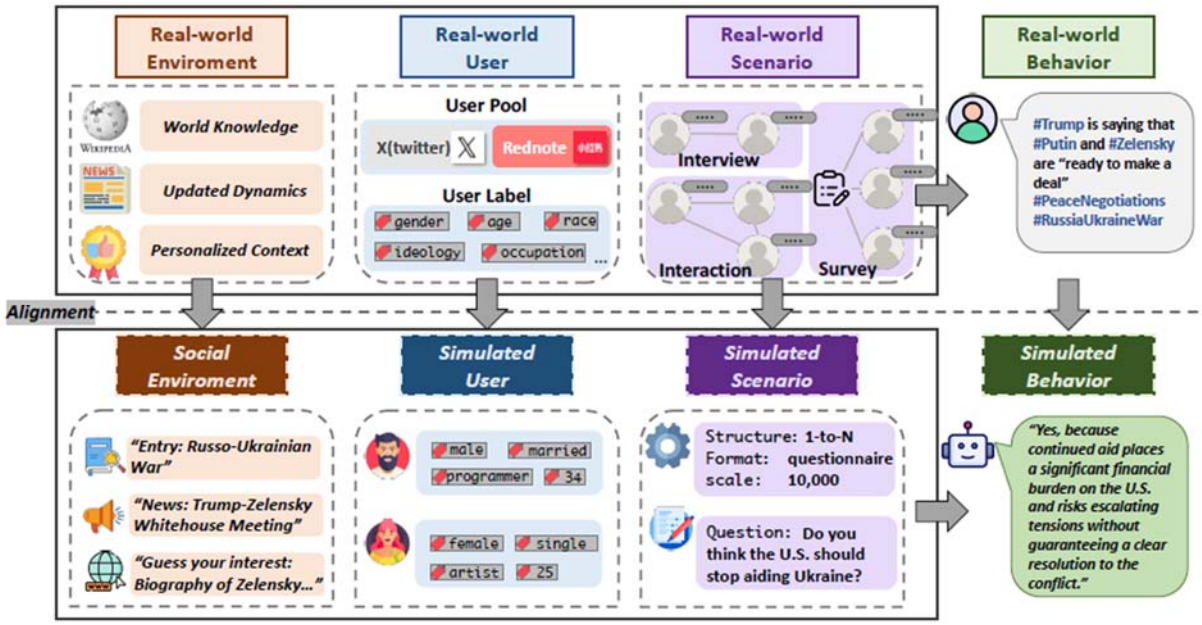
---

(1) Toutes les plateformes ne sont pas concernées. Certains réseaux sociaux, dont X, nécessitent encore une action humaine pour créer un compte utilisateur.

(2) Selon une grille de lecture chinoise, la guerre non cinétique n'occasionne ni morts ni destructions matérielles. Elle comprend notamment la lutte informationnelle et les actions cyber Par opposition, la guerre cinétique est une guerre conventionnelle ou nucléaire provoquant des dégâts matériels et humains.

(3) Nathalie Guibert, Brice Pedroletti, « Comment la Chine durcit sa guerre d'influence pour démontrer sa puissance »*, Le Monde, 3 septembre 2021.*

réponse à l'échelle macro. Cette prévisualisation pourrait dès lors affiner les campagnes d'ingérence afin de maximiser les chances de succès d'une opération.



Mise en production d'un socioverse appliqué à la guerre russo-ukrainienne (1)

Si l'objectif peut donc être de convaincre par le mensonge, la diffusion de publications fausses vise surtout et avant tout à miner la confiance du public dans les sources traditionnelles d'information et à favoriser la défiance à l'égard des médias ainsi que des institutions. Cette machination amplifiée par l'IA n'est pas nouvelle et répond aux visées anciennes des autoritarismes. Ainsi, les mots d'Hannah Arendt résonnent aujourd'hui avec une modernité nouvelle : « *le sujet idéal du règne totalitaire n'est ni le nazi convaincu ni le communiste convaincu, mais l'homme pour lequel la distinction entre fait et fiction, c'est-à-dire la réalité de l'expérience et la distinction entre vrai et faux, c'est-à-dire les normes de la pensée, n'existe plus* » (2). Selon Viginum, les effets de ces campagnes d'ingérence numérique sont avérés mais demeurent, pour le moment, modérés.

Les rapporteurs se sont intéressés spécifiquement aux dernières campagnes d'ingérence russes. Les modes opératoires utilisés et l'échelle des attaques sont en effet symptomatiques de la nouvelle forme de menace alimentée par l'IA. Formellement attribuées à Moscou, les opérations *Reliable recent news* (RNN) (3),

(1) Xinnong Zhang, Jiayu Lin, Xinyi Mou, Shiyue Yang, Xiawei Liu, Libo Sun, Hanjia Lyu, Yihang Yang, Weihong Qi, Yue Chen, Guanying Li, Ling Yan, Yao Hu, Siming Chen, Yu Wang, Xuanjing Huang, Jiebo Luo, Shiping Tang, Libo Wu,Baohua Zhou, Zhongyu Wei, « SocioVerse: A World Model for Social Simulation Powered by LLM Agents and A Pool of 10 Million Real-World Users », *2025.*

(2) Hannah Arendt, Les origines du totalitarisme – le système totalitaire, *Points, éditions 2025, 1951, 384 pages.*

(3) Viginum, RRN: une campagne numérique de manipulation de l'information complexe et persistante, *SGDSN, juin 2023, 5 pages.*

Portal Kombat [1], *Matriochka* [2] et *Storm-1516* [3] sont autant d'attaques informationnelles menées à grande échelle pour saper les fondements du système démocratique français.

Storm-1516, une opération d'ingérence numérique russe cible la France

Le 6 mai 2025, Viginum a révélé les détails d'une opération d'ingérence numérique étrangère russe appelée « Storm-1516 » et reliée au renseignement militaire russe (GRU).

Déployée, entre août 2023 et mars 2025, l'opération Storm-1516 a ciblé simultanément le territoire français, l'Ukraine ainsi qu'un ensemble d'États occidentaux. Cette opération a pris la forme de soixante-dix-sept actions coordonnées réparties en cinq vagues successives. Selon les données disponibles, cette campagne a généré plus de 55,8 millions de vues et 38 877 publications sur les réseaux sociaux [4]. À ce titre, Storm-1516 constitue, à ce jour, l'opération de désinformation étrangère la plus étendue jamais documentée sur le territoire national.

L'objectif affiché de l'opération Storm-1516 était de décrédibiliser le gouvernement ukrainien, avec pour finalité la suspension de l'aide occidentale à l'Ukraine, dans le contexte de l'invasion du pays par la Russie. En parallèle, le mécanisme opérationnel d'ingérence (MOI) utilisé cible directement des responsables politiques européens et leur entourage, notamment durant les périodes électorales en France, aux États-Unis et en Allemagne. En France, les attaques suivent mécaniquement les déclarations officielles de soutien à l'Ukraine ainsi que les scrutins électoraux, avec une intensification marquée en juin 2024.

(5)

Le schéma de diffusion de l'opération Storm-1516 présente un degré de complexité élevé et a connu une évolution progressive au fil du temps. L'utilisation de l'intelligence artificielle est observable au cours des cinq phases de l'opération [6].

---

(1) Viginum, Portal kombat – un réseau structuré et coordonné de propagande prorusse, *SGDSN, février 2024, 17 pages.*

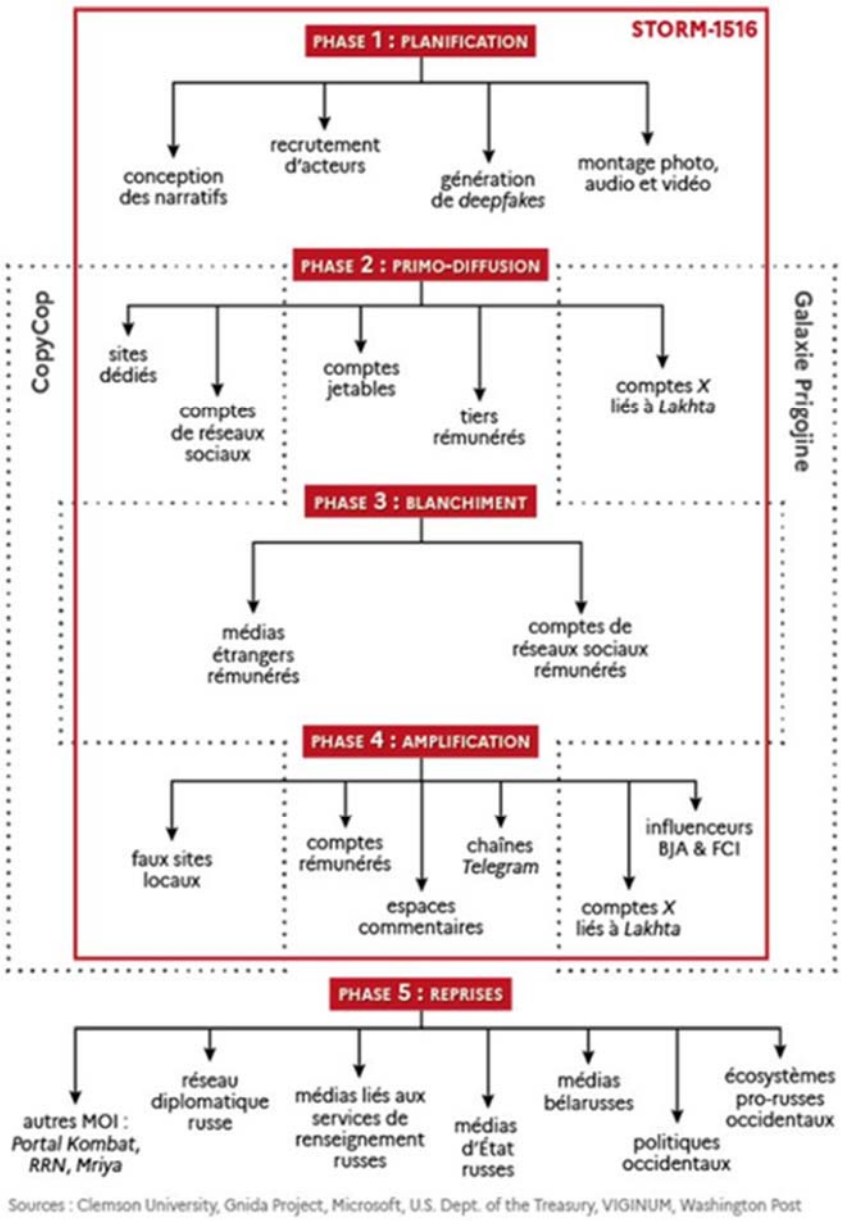
(2) Viginum, Matriochka – Une campagne prorusse ciblant les médias et la communauté des factcheckers », *SGDSN, juin 2024, 16 pages.*

(3) Viginum, Analyse du mode opératoire informationnel russe Storm-1516, *SGDSN, mai 2025, 41 pages.*

(4) Nathalie Huet, McKenzie Sadeghi, Chine Labbé, NewsGuard, « Une campagne de propagande russe cible la France avec des scandales générés par IA, générant 55 millions de vues sur les réseaux sociaux », *17 avril 2025.*

(5) Ibid.

(6) Ministère des armées et des anciens combattants, « Storm-1516 ou les dessous d'une opération d'influence russe », 14 mai 2025.



Sources : Clemson University, Gnida Project, Microsoft, U.S. Dept. of the Treasury, VIGINUM, Washington Post

Lors de la phase de planification, des *prompts* sont utilisés afin d'aider à la conception de narratifs ciblés. Les opérateurs emploient une grande diversité de contenus, tels que des montages photo et vidéo, des faux reportages, ainsi que des productions audio et vidéo générées à l'aide d'outils d'IA générative. Ces contenus sont disponibles en plusieurs langues, notamment en français, anglais, ukrainien, allemand, espagnol et arabe. Les auteurs de l'opération ont recours à des outils permettant de produire artificiellement des voix et des images, ce qui contribue à crédibiliser des profils se présentant comme des « lanceurs d'alerte » et à usurper l'identité de personnalités publiques ou d'internautes n'ayant aucun lien avec les narratifs propagés.

À cette première temporalité, succède une phase de primo-diffusion. L'IA est utilisée pour automatiser des comptes jetables, multiplier les sites et les profils sur les réseaux sociaux, et en renforcer la crédibilité par la création de contenus, d'articles ou de publications. Ce procédé se poursuit au cours de la phase d'amplification.

Enfin, au cours de la phase de blanchiment, un audit mené par l'entreprise NewsGuard a mis en évidence que, dans 32 % des cas, les principaux chatbots d'IA générative occidentaux reprenaient les récits de désinformation liés à Storm-1516, à la suite d'un processus de blanchiment stratégique reposant sur de faux sites d'information locaux et des vidéos de prétendus lanceurs d'alerte diffusées sur YouTube. Le 6 mai 2025, la France a officiellement condamné ces manœuvres par une déclaration conjointe émise par le ministère de l'Europe et des affaires étrangères et le secrétariat général de la défense et de la sécurité nationale (SGDSN).

Des attaques d'une ampleur plus restreinte ont également mobilisé partiellement l'IA pour « faire du bruit », fabriquer du doute et saturer l'espace médiatique français en diffusant un climat anxiogène. Loin d'être anecdotiques, elles ont permis de mettre en évidence le rôle multiplicateur de l'IA et sa capacité à réduire drastiquement les coûts pour un résultat maximal. À cet égard, la réalisation de centaines de tags d'étoiles de David bleues sur des murs éclairés de Paris a ému l'opinion publique et la classe politique, et ce avant la publication des conclusions de l'enquête officielle. Viginum a remonté la piste des quatre personnes de nationalité moldaves interpellées par la police jusqu'à un commanditaire affilié aux services de renseignement russes. Les clichés des photographies de ces tags, pris par l'un des individus sur place, ont par la suite été relayés sur les réseaux sociaux par 1 095 bots simultanément. La Russie combine le recours à l'IA à ses anciennes techniques de disruption *via* le recrutement à peu de frais d'individus pauvres issus de pays non russes en tant qu' « agents jetables ». Il devient alors économiquement rationnel d'investir quelques millions de roubles pour mener à bien ce type d'actions au vu du coût d'une seule journée de guerre conventionnelle sur la ligne de front.

En parallèle, les symboles utilisés ne sont pas choisis au hasard et servent à alimenter des peurs ou clivages déjà présents dans la société : en l'espèce, est ici repris un scénario utilisé par le Comité pour la sécurité de l'État (KGB) dans les années 1960 afin d'attiser la haine antisémite (1) . Le tout est par la suite confusément mêlé à la couleur bleue évocatrice du drapeau d'Israël pour cristalliser les tensions entre communautés autour du conflit à Gaza. Moscou recycle et modernise son concept de subversion en quatre phases tel que défini par le KGB pendant la guerre froide : « *Démoraliser un pays, c'est lui faire perdre foi en lui-même, faire en sorte que ses enfants ne sachent plus pourquoi leurs parents se sont battus, faire croire à une société qu'elle est fondamentalement mauvaise, raciste ou vendue à l'étranger, oppressante ou trop permissive, coloniale ou en perte d'empire, patriarcale ou décadente.* » Les mots de l'un des théoriciens, Yuri Bezmenov, définissent ainsi la première phase, la démoralisation, à laquelle succèdent la déstabilisation, la crise puis la normalisation (2).

b. Les attaques par empoisonnement des données font de l'intelligence artificielle à la fois un outil et une cible d'ingérence

Facilitateur de désinformation, l'intelligence artificielle se trouve également être l'une des victimes de ces manœuvres. Les acteurs hostiles étrangers ont ainsi identifié une nouvelle voie permettant de mobiliser cette technologie en tant que vecteur de fausses informations en la retournant contre elle-même par un phénomène d'empoisonnement des données. Les attaques par empoisonnement désignent, selon la Commission nationale de l'informatique et des libertés (CNIL), « [les actions visant] *à modifier le comportement du système d'IA en introduisant*

(1) En 1959, le KGB fait taguer dans des villes ouest-allemandes des swatiskas pour discréditer le bloc occidental. Les tags sont par la suite repris dans plusieurs pays occidentaux, au point qu'est évoquée une « épidémie de swatiskas ».

(2) Ariel Wizman, « Yuri Bezmenov, l'agent du KGB passé à l'Ouest », *podcast Une histoire truculente, France Culture, 9 novembre 2025.*

des données corrompues en phase d'entraînement ou d'apprentissage. Elles supposent que l'attaquant soit en mesure de soumettre des données à utiliser lors de l'entraînement du système d'IA ». La génération massive de faux contenus par l'IA pollue l'écosystème internet d'où sont extraits les jeux de données utilisés par la plupart des agents conversationnels (LLM) tels que ChatGPT. Un utilisateur peut donc se retrouver, sans en être conscient, face à une réponse générée à partir de fausses informations provenant de sites entièrement alimentés à dessein par l'IA. Les algorithmes qui assurent le fonctionnement de la plupart des LLM ont en effet tendance à sélectionner des éléments de réponse en fonction de leur fréquence d'apparition sur les moteurs de recherche.

Ce biais est parfaitement intégré par les acteurs malveillants, qui, à l'image de la Russie, n'hésitent pas à produire massivement des données erronées pouvant potentiellement être reprises par d'autres IA. Il s'agit d'une stratégie délibérée résultant d'un double constat : l'outil IA est vulnérable et ne sait ni distinguer le vrai du faux ni identifier les sources manifestement frauduleuses ; les utilisateurs lambda ont de plus en plus recours à cette technologie pour s'informer et suivre l'actualité.

À cet égard, Viginum a analysé l'activité d'un réseau organisé de 193 « portails d'information » numériques, appelé « réseau pravda » en raison des adresses web (URL) utilisées (1) ou réseau « Portal Kombat » en référence à la stratégie informationnelle offensive mise en place à partir du mois de février 2022 par les acteurs qui administrent ces portails numériques (2). Cet enchevêtrement de faux sites propose des publications pro-russes dans quarante-huit langues et produit en moyenne 9 500 articles par jour. Il ne crée aucun contenu original et regroupe des informations provenant de médias d'État, de responsables du Kremlin et d'influenceurs affiliés. Le réseau utilise des outils d'automatisation pour extraire les données des organes de presse nationaux, de Telegram et d'autres sources diverses avant de les agréger et de publier en masse. Les sites de ce réseau sont rarement visités par des humains et ont probablement été conçus pour cibler avant tout les robots d'indexation ainsi que les outils de moissonnage de données (*scraping*), afin de contaminer les réponses des chatbots d'IA. Les résultats de ce gavage intoxicant sont avérés. Selon l'entreprise spécialisée dans la lutte contre les infox, Newsguard, près de 3 600 000 articles de propagande russe ont été postés sur internet en 2024 et une grande part d'entre eux se retrouve désormais intégrée dans les résultats des systèmes occidentaux d'IA générative. Afin de tester leur propension à répéter ou au contraire à résister à des récits connus de désinformation, NewsGuard réalise, depuis juillet 2024, des audits mensuels des dix principaux chatbots d'IA générative (3). Mois après mois, une moyenne se dessine : ces chatbots répètent avec autorité de fausses informations dans environ 20 % des cas observés. D'après les enquêtes de NewsGuard, cette reprise des infox russes par des canaux réputés sans

---

(1) Le portail visant la France reprenait l'indicatif « -fr » en y adjoignant le mot « pravda *», signifiant la vérité en russe.*

(2) Viginum, Portal kombat – un réseau structuré et coordonné de propagande prorusse, *Op.Cit.*

(3) Isis Blachez, McKenzie Sadeghi, « Despite being exposed,massive Russian campaign continues to infect AI tools with pro-Kremlin propaganda *», NewsGuard, 9 juin 2025.*

danger se fait également par le biais de médias locaux dans des pays en développement, notamment africains, et permet au Kremlin d'entretenir une illusion de légitimité. La Russie s'est ainsi dotée d'un véritable mécanisme de blanchiment de l'information.

L'empoisonnement des données peut par ailleurs être plus subtil et amplifier des failles inhérentes aux agents conversationnels. Ces derniers peuvent en effet parfois être sujets à des « hallucinations », des erreurs manifestes de réponse résultant d'une défaillance technique de la machine. Au-delà, sans contrôle de la phase d'apprentissage, le *chatbot* peut également se mettre à copier les éléments de langage de ses utilisateurs. Tel fut le cas du prédécesseur de ChatGPT, le modèle développé par Microsoft en 2016, Tay, qui après quelques heures d'échanges avec des internautes, s'est mis à produire du contenu raciste, sexiste et réfutant l'Holocauste. De manière tout aussi problématique, en fonction des paramétrages de leur algorithme, certains modèles d'IAg tiennent délibérément des propos répréhensibles. Faisant déjà l'objet d'une enquête du parquet de Paris pour des faits similaires, l'agent conversationnel de la plateforme X, Grok, a ainsi récidivé le 19 novembre 2025 en tenant des propos ouvertement négationnistes, affirmant que les chambres à gaz n'étaient pas destinées à éliminer les juifs mais à les traiter contre le typhus [(1)].

Enfin, un pays peut également choisir de corrompre volontairement les données d'une intelligence générative nationale, dès sa conception. Ainsi, l'algorithme de l'IA générative chinoise DeepSeek adapte ses réponses en fonction des sujets considérés comme sensibles par le PCC. Une recherche sur la situation de la communauté ouïghours au Xinjiang, sur les manifestations à Hongkong en 2019 ou encore sur le statut de Taïwan n'apportera que des données lacunaires ou erronées.

c. L'utilisation de l'intelligence artificielle dynamise les cyberattaques en contribuant à intensifier la menace numérique

Au-delà de l'empoisonnement des données des IA génératives, les modèles d'intelligence artificielle peuvent aussi faire l'objet d'attaques plus conventionnelles par piratage (*hacking*). Une cyberattaque visant une infrastructure de gestion de données (*cloud*) pourrait permettre de prendre le contrôle de l'hébergeur du modèle, afin de réaliser ensuite d'autres actions malveillantes. Ces actions peuvent par ailleurs être grandement facilitées par l'assistance de l'IA. En rédigeant automatiquement du code sur demande, elle fait tomber certaines barrières d'accès au piratage. Dans le domaine de la cyberguerre, l'intelligence artificielle représente ainsi une menace croissante, identifiée par des acteurs publics comme privés.

À la date de rédaction de ce rapport, l'IA n'est pas encore capable de produire une cyberattaque complète de manière autonome. Elle est toutefois de plus

(1) Le Monde, AFP, « L'enquête sur la plateforme X étendue à des « propos négationnistes » publiés par son IA, Grok », *19 novembre 2025.*

en plus utilisée comme appoint par les hackeurs pour fluidifier et accroître l'efficacité des différentes phases de l'attaque, du ciblage de la victime à la conception du code. Elle permet alors de réduire le coût financier et humain d'une opération cyber. En conséquence, la ressource IA est de plus en plus intégrée dans les attaques. Google a ainsi identifié au cours des dernières années plusieurs activités de piratage impliquant son modèle IA, Gemini, et menées par des groupes paraétatiques provenant de près de vingt pays, avec une surreprésentation notable de la Chine (1). En juin 2025, l'entreprise californienne a également mis en évidence les activités criminelles du groupe Frozenlake (APT28), soutenu par le gouvernement russe et utilisant pour la première fois des logiciels malveillants (*malware*) combinés à un LLM. Ces programmes, « Promptsteal » et « Promptflux » utilisent un agent conversationnel pour générer dynamiquement des scripts agressifs et masquer leur propre code. Google y voit une évolution inquiétante vers une autonomisation croissante, une adaptation rapide et une prolifération exponentielle des programmes pirates (2).

Concrètement, ces attaques mobilisent souvent l'IA pour améliorer les techniques d'hameçonnage (*phishing*). Le pirate cherche à se faire passer pour un site ou un tiers de confiance afin de récupérer des données confidentielles de sa victime telles que des informations bancaires ou d'autres données personnelles sensibles. Les courriels ou messages produits sont nettement plus crédibles que par le passé grâce la reconnaissance orthographique et à l'accès à une traduction instantanée. En parallèle, la récupération de mot de passe est facilitée par la capacité de l'IA à tester des combinaisons multiples en un temps record. Ces vols de données ou l'introduction d'un cheval de Troie se traduisent ensuite par des attaques en déni de service (DDoS) mobilisant des raçongiciels permettant de bloquer l'accès à un équipement ou un système jusqu'au versement d'une rançon. Au-delà, les flux sont sensiblement augmentés et permettent de cibler un panel de victimes plus large, du simple particulier à une organisation gouvernementale en passant par un hôpital de province.

Enfin, les failles des systèmes d'intelligence artificielle ont fait émerger de nouvelles attaques. Lors du déplacement de la rapporteure à Washington, la société Recorded Future, l'a alertée sur le risque d'« *IA slopsquatting* », une attaque mobilisant les hallucinations des intelligences artificielles (*cf. supra*) lorsqu'elles produisent des codes à la demande d'un utilisateur. Un paquet de codes généré par IA peut en effet contenir des fragments de code inutiles ou des erreurs manifestes. En repérant ces tendances hallucinogènes, un pirate peut les exploiter pour renvoyer vers des sites corrompus et introduire, dès la conception d'un programme, des « portes dérobées » (*backdoors*) lui donnant accès à distance aux appareils utilisés. L'IA peut en effet suggérer l'utilisation d'un paquet de données qui n'existe pas. Un acteur malveillant crée alors un véritable progiciel reprenant le nom imaginé par

---

(1) Communément identifiés par métonymie, les advanced persistent threat (APT), *désignent les groupes de cybercriminels à l'origine des attaques informatiques – Selon les données de Google, la Chine en compte plus d'une cinquantaine https://cloud.google.com/security/resources/insights/apt-groups?hl=fr, consulté en ligne le 20 novembre 2025.*

(2) Google, GTIG AI threat tracker : advances in treat actor usage of AI tools, novembre 2025, p. 7.

l'IA, le télécharge sur un référentiel public tel que les gestionnaires pythons ou java, PyPI et npm, et y injecte un logiciel malveillant. Les développeurs qui s'appuient sur des outils IA finissent par installer à leur insu ce code contrefait directement au cœur de leurs projets.

3. Les risques sécuritaires sont particulièrement accrus pour les systèmes démocratiques en raison des valeurs qu'ils défendent et de leurs processus électoraux

Caractérisées par leur ouverture et leur pluralisme, les démocraties sont des cibles de choix pour des entités étrangères hostiles. Ces dernières perçoivent en effet nos valeurs sociétales comme autant de vulnérabilités pouvant être exploitées. L'absence d'une information centralisée et contrôlée par un parti unique, les divergences revendiquées d'opinions, les exigences de transparence de la vie économique et politique, notamment, sont des forces des modèles démocratiques instrumentalisées par des régimes autoritaires ou des organisations terroristes afin de nous nuire. Libre de douter, le citoyen peut ainsi être amené artificiellement à remettre en question les vertus d'une organisation lui garantissant pourtant des droits inexistants dans les régimes autocratiques à la manœuvre.

En sapant la confiance dans les institutions démocratiques, les acteurs malveillants s'attaquent au socle du contrat social et à l'acceptation citoyenne d'une aliénation des droits naturels au profit d'un intérêt commun. La Russie n'a ainsi eu de cesse d'attaquer les personnalités politiques européennes et françaises en leur prêtant des scandales et de fausses affaires de corruption. Dès 2017, en période de réserve électorale, des milliers de mails et documents de campagne du candidat à l'élection présidentielle, Emmanuel Macron, sont diffusés sur des sites internet obscurs avant d'être repartagés sur les réseaux sociaux par des bots sous le hashtag « #MacronLeaks ». Chaque élection est ainsi un moment à risque identifié par les services de renseignement en France et dans les pays visités par les rapporteurs. La Suède organisera notamment ses élections générales en 2026 et se prépare à faire face à des attaques visant à modifier le sens des scrutins. Si les effets des campagnes d'ingérence étrangère dans les processus électoraux sont encore rares, les cas

récents des scrutins en Roumanie [1] et en Moldavie [2] ont souligné l'augmentation des risques sous-jacents et leur démultiplication en raison du recours à l'intelligence artificielle.

Manipulations algorithmiques et ingérences russes lors de l'élection présidentielle roumaine de novembre 2024

Le 6 décembre 2024, le premier tour de l'élection présidentielle en Roumanie a été annulé par la Cour constitutionnelle, à la suite de la mise en évidence d'activités numériques ayant altéré la sincérité du scrutin. Le candidat d'extrême droite pro-russe, Călin Georgescu, peu présent dans la vie politique nationale et crédité d'environ 1 % d'intentions de vote début novembre, a obtenu 22,94 % des suffrages lors du premier tour de ces élections le 24 novembre. Ce résultat apparaît d'autant plus surprenant qu'aucune campagne physique n'avait été organisée, que ce candidat ne disposait pas de structure militante et qu'il n'avait pas déclaré de budget.

Cette progression fulgurante dans les urnes a été alimentée par une visibilité anormalement élevée sur les réseaux sociaux, essentiellement sur TikTok. Entre le 10 et le 24 novembre, le nombre d'abonnés et de vues associés au candidat y a été multiplié par trois. Les hashtags liés à sa campagne, notamment « #CalinGeorgescu », ont enregistré plus de 73 millions de vues en sept jours. Les services roumains ont identifié environ 25 000 comptes TikTok soutenant le candidat dont l'activité a fortement augmenté au cours des deux semaines précédant le scrutin, parmi lesquels plus de 700 anciens comptes auparavant inactifs. TikTok a indiqué que cette dynamique ne présentait pas les caractéristiques d'une croissance organique, mentionnant l'action de « *bénévoles coordonnés* ».

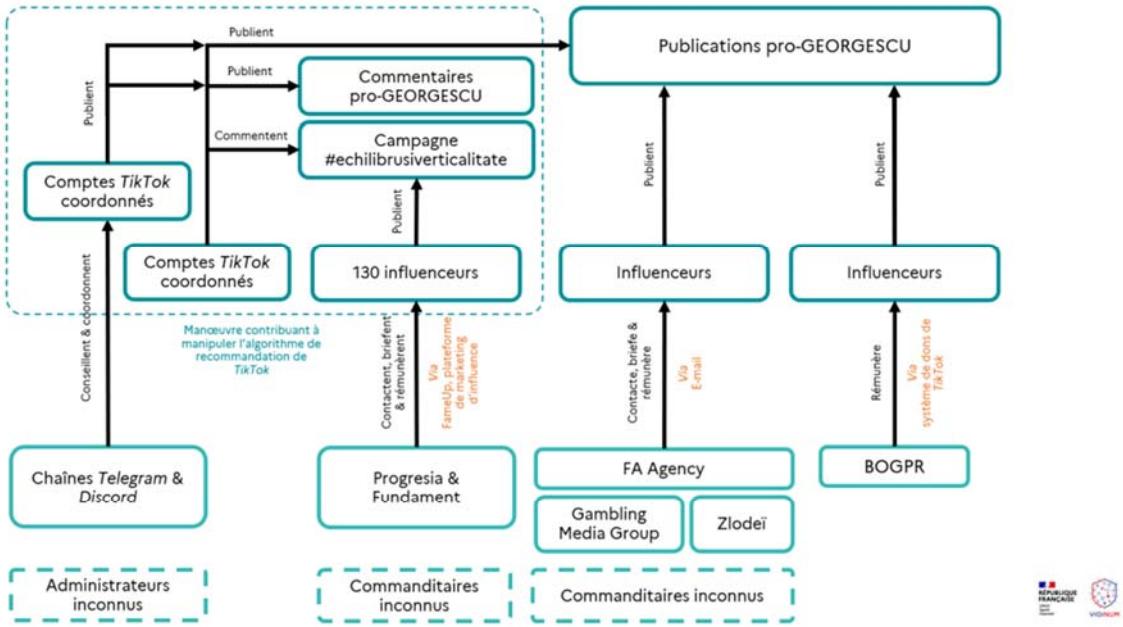


Schéma récapitulant les manœuvres observées sur TikTok visant à promouvoir le candidat GEORGESCU

(1) Viginum, Manipulation d'algorithmes et instrumentalisation d'influenceurs Enseignements de l'élection présidentielle en Roumanie & risques pour la France*, SGDSN, février 2025, 14 pages.*

(2) Le Monde, AFP, « Moldavie : l'UE dénonce « l'interférence sans précédent » de la Russie, après la victoire étriquée du oui au référendum *», 21 octobre 2024.*

Selon les documents déclassifiés par les services de renseignement roumains, ces éléments s'inscrivent dans une campagne structurée d'*astroturfing* [(1)], reposant sur l'usage coordonné de plusieurs dizaines de milliers de comptes et sur le recours à des réseaux d'influenceurs. La coordination s'appuyait sur des canaux des plateformes Discord et Telegram, dont certains étaient actifs depuis 2022, diffusant des instructions relatives à la production et au partage de contenus destinés à influer sur l'algorithme de recommandation de TikTok. Ces consignes comprenaient notamment l'adaptation de vidéos afin d'éviter la détection par les systèmes de modération et l'utilisation systématique de hashtags sous des publications populaires. Les autorités roumaines ont également signalé des comportements similaires sur les plateformes du groupe Meta, ainsi qu'un volume important de cyberattaques visant des infrastructures liées au processus électoral.

Parallèlement, la campagne a impliqué plus d'une centaine d'influenceurs totalisant environ huit millions d'abonnés. Ceux-ci ont publié des contenus généraux relatifs à l'élection, dont les espaces de commentaires ont ensuite été alimentés par des messages favorables au candidat. Plusieurs influenceurs ont indiqué après le scrutin avoir reçu une rémunération, parfois selon une grille proportionnelle à leur audience. Viginum mentionne également l'existence de financements non déclarés, dont au moins 381 000 dollars en dons *via* TikTok provenant d'un homme d'affaires roumain, ainsi que le rôle d'une entreprise étrangère dans le recrutement d'influenceurs. La Cour constitutionnelle a estimé que l'ensemble de ces irrégularités avait affecté les conditions d'égalité entre candidats et altéré le caractère libre et transparent du scrutin, conduisant, le 6 décembre 2024, à l'annulation du premier tour. À la suite de cette annulation, un nouveau processus électoral a été engagé et la Commission électorale a déclaré irrecevable la candidature de Călin Georgescu.

Les acteurs hostiles tendent également à viser des communautés plus spécifiquement. Les diasporas sont particulièrement vulnérables aux actions d'ingérence orchestrées par un État ou des groupes criminels. Les ressortissants des diasporas sont plus facilement atteignables par de la propagande étrangère diffusée dans leur langue maternelle. En outre, en raison de difficultés linguistiques, ils sont également généralement moins connectés aux médias locaux et à une production d'informations fiables. Il ressort ainsi des auditions menées que plusieurs régimes autoritaires tout comme certaines organisations terroristes ciblent prioritairement ces citoyens afin de relayer et blanchir des infox. En Suède, une campagne de désinformation organisée par l'Iran a par exemple résonné fortement dans les diasporas de confession musulmane en raison du narratif choisi : les services sociaux suédois auraient enlevé des enfants de familles musulmanes pour les placer dans des centres et les soustraire à l'islam. Cette opération faisait écho à des autodafés du Coran et visait à attiser la haine en brisant la promesse d'intégration des communautés immigrées.

*(1) D'après Viginum, l'*astroturfing *est un* « mode opératoire consistant à conférer de la visibilité à un sujet en faisant croire qu'il est un phénomène de masse alors même qu'il émane de la coordination de quelques comptes seulement qui produisent un volume important de publications sur un même sujet ».

II. LA PERCEPTION FRANÇAISE ET EUROPÉENNE DU DANGER APPARAÎT ENCORE PARTIELLE ET PARFOIS NAÏVE AU REGARD DU POSITIONNEMENT DES AUTRES PUISSANCES

A. L'UNION EUROPÉENNE SOUFFRE D'UNE VULNÉRABILITÉ STRUCTURELLE RENDANT DIFFICILE L'ÉLABORATION D'UNE STRATÉGIE POUR L'INTELLIGENCE ARTIFIELLE

1. La souveraineté numérique européenne s'avère abstraite voire fantasmée, plaçant le vieux continent dans une situation de dépendance aux technologies étrangères

« *C'est le plus important : l'Europe doit radicalement recentrer ses efforts collectifs sur la réduction du fossé qui la sépare des États-Unis et de la Chine en matière d'innovation, en particulier dans le domaine des technologies avancées.* […] *Alors que le monde est à l'aube d'une révolution de l'intelligence artificielle, l'Europe ne peut se permettre de rester bloquée dans les « technologies et industries intermédiaires » du siècle passé.* » (1). L'appel à un sursaut commun formulé en 2024 par l'ancien président de la Banque centrale européenne (BCE), Mario Draghi, est clair. Compte tenu du diagnostic alarmant établi, l'Union européenne doit se ressaisir pour combler un retard préoccupant sur le plan économique et technologique, tandis que sa croissance s'essouffle et que l'écart avec les autres grandes puissances mondiales se creuse. Ces fragilités sont accentuées par une pénurie relative de compétences résultant d'un double mouvement lié au vieillissement démographique et à une « fuite des cerveaux ». À cet égard, le départ pour les États-Unis, en 2013, des concepteurs français de Mindie, première application de défilement de vidéos à la verticale dont le concept a été repris par le chinois ByteDance pour créer TikTok, est symptomatique. Un tel manque d'autonomie stratégique sert directement les intérêts d'autres puissances étrangères.

Le rapporteur alertait déjà sur cette situation en 2021 dans le cadre de ses précédents travaux sur les géants du numérique (2). Au-delà du constat, des facteurs structurels ont abouti à ce décrochage européen : une absence de marché numérique commun, une hétérogénéité normative en matière de fiscalité des entreprises et un accès aux financements encore laborieux (3). Par ailleurs, le cycle de politiques publiques d'inspiration néolibérale enclenché au cours des années 1970 a conduit à rétrocéder l'objectif de puissance nationale, à tous niveaux, dans l'ordre des priorités. La fin des politiques volontaristes gaullistes a en effet mené à l'avènement d'un modèle théorique de pensée consistant à favoriser le développement des entreprises sur le sol national, sans distinction géographique. L'attraction des capitaux étrangers est devenue la pierre angulaire des décisions économiques.

---

(1) Mario Draghi, Commission européenne, L'avenir de la compétitivité européenne, *septembre 2024, p. 9.*

(2) Alain David, Marion Lenne, rapport d'information n° 4213 sur les géants du numérique, *Assemblée nationale, 2 juin 2021, XV^e législature, 189 pages.*

(3) Ibid.

Or, il n'existe pas d'exemple de pays qui soit parvenu au rang de puissance technologique – *a fortiori* dans le domaine de l'IA – en suivant une telle stratégie. Pour cause, les deux modèles qui ont donné les meilleurs résultats, la Chine et les États-Unis, sont tous deux fondés sur un modèle nationaliste. Si la France et l'Union européenne souhaitent devenir une puissance dans le secteur numérique, il ne semble pas possible de faire l'économie d'une telle logique. Cela impliquerait d'engager un rapport de force, non seulement avec certains voisins européens mais également avec les deux hégémons : les États-Unis et la Chine.

Au-delà, la Commission européenne s'est historiquement construite, sur un plan politique, en s'appuyant sur une relation partenariale privilégiée avec de grandes entreprises, tout spécialement américaines, pour affirmer une position dominante face aux États membres. Autrement dit, la Commission européenne est non seulement structurellement hostile à l'affirmation des puissances souveraines nationales mais elle s'est bâti une légitimité politique grâce à la pénétration des capitaux extra-européens sur le sol communautaire. Les intérêts étrangers ont un tel poids dans le processus décisionnel européen que la formalisation d'une « politique de puissance » européenne paraît aujourd'hui ardue.

Au demeurant, les auditions menées à Bruxelles par la rapporteure ont mis en lumière des blocages, ou du moins des lenteurs dans les processus, inhérents au fonctionnement communautaire. L'Union européenne est fondamentalement un système institutionnel inter-gouvernemental. Ses initiatives reposent *in fine* sur des rapports de force entre États membres et entre les systèmes productifs nationaux. Or les vingt-sept n'ont pas les mêmes intérêts, sont parfois en compétition et s'avèrent, pour certains, être insérés dans des sphères d'influence extra-européennes, américaines, chinoises ou russes. Trois éléments ont ainsi été rappelés : la perception du risque diffère en fonction des pays et de l'intensité de la menace – la France et la Suède sont, par exemple, plus mobilisées sur les questions de désinformation que leurs voisins italiens ou espagnols – ; certains gouvernements ont des conceptions divergentes selon une lecture opportuniste du problème, à l'image de la Hongrie ; d'autres objectifs européens peuvent entrer en contradiction, les règles de la politique de concurrence semblant ainsi compliquer l'émergence d'un géant du numérique et de l'IA en Europe.

Par conséquent, nous dépendons déjà d'entreprises étrangères pour la fourniture de la plupart des composants électroniques indispensables à la production des nouvelles technologies : 90 % des semi-conducteurs sont produits hors d'Europe, près de 70 % de l'informatique en nuage (*cloud*) utilisée en Europe est américain [(1)] et une grande majorité des Européens utilisent quotidiennement les suites de productivité de Microsoft avec le célèbre « pack Office ». Les puces pour cartes graphiques (GPU) utilisées sont également quasiment exclusivement fournies par deux entreprises étrangères, l'américain Nvidia et le taïwanais *Taïwan Semiconductor Manufacturing Company* (TSMC). La boutade de l'ancienne

(1) Antonin Bergeaud, « Si l'UE ne maîtrise pas les technologies qu'elle veut encadrer, elle deviendra une commentatrice impuissante », *Le Monde, 11 octobre 2025.*

présidente de la Confédération générale de l'industrie italienne (*Confindustria*), Emma Marcegaglia, est souvent répétée avec un sourire pincé par certaines personnes auditionnées : « *Les États-Unis innovent, la Chine copie, l'Europe réglemente* ». Cette situation aboutit à un véritable paradoxe communautaire. La régulation est prioritaire sans que des actions dynamiques soient entreprises en parallèle pour maîtriser la technologie, au risque de devenir rapidement inaudible : « *La souveraineté technologique, c'est d'abord la liberté de rester fidèle à soi-même, d'imposer ses règles quand elles sont contestées. C'est la possibilité de dire non à une économie de la démesure et de l'extraction des données, de refuser une innovation qui méprise la vie privée ou ignore les limites planétaires sans mettre en péril notre prospérité économique. Sans maîtrise des outils du progrès, nous ne pourrons pas influencer sa direction.* » (1).

2. Cette absence d'autonomie se traduit par un retard dans le développement de l'intelligence artificielle en Europe

La faiblesse de l'Europe dans le domaine des technologies numériques s'explique par « *une structure industrielle statique qui engendre un cercle vicieux de faibles niveaux d'investissement et d'innovation* » (2). Alors qu'aux États-Unis le profil des trois entreprises dépensant le plus pour la recherche a sensiblement évolué, passant de l'industrie automobile et pharmaceutiques aux compagnies de logiciels et d'informatique, avant de désormais correspondre à des sociétés du numérique depuis les années 2020, en Europe, ce sont toujours trois entreprises du même secteur de l'automobile qui caracolent en tête du classement des investisseurs en recherche et développement. Le virage technologique n'a donc pas été pris, ralentissant considérablement les avancées dans le domaine de l'IA sur le vieux continent. Depuis les années 2000, l'Union européenne a en effet échoué à s'imposer dans les moments clés, investissant principalement dans des technologies déjà matures, quand États-Unis et Chine misaient sur les technologies de rupture.

Plusieurs indicateurs objectifs démontrent ainsi une perte de vitesse de l'Europe dans la course mondiale au développement de l'IA. Entre 2017 et 2023, 70 % des modèles d'IA fondamentaux ont été développés outre-Atlantique, laissant les vingt-sept loin derrière dans la compétition internationale. En 2024, les États-Unis ont créé quarante modèles d'IA, surpassant largement la Chine et ses quinze modèles, ainsi que l'Europe qui n'en a produit que trois (3). Nerf de la guerre, la recherche n'est également pas encore suffisamment financée et approfondie. En 2024, le pays publiant le plus de travaux demeure la Chine avec 36 % des articles scientifiques sur le sujet, suivi par l'Inde et les États-Unis, représentant chacun environ 12 % du total. L'Europe serait certes virtuellement à la deuxième place avec près de 18 % des publications mais sa prise en compte en tant que force régionale est rarement retenue en raison du manque de cohérence apparent entre États et de la

(1) Ibid.

(2) Mario Draghi, Op. Cit.

(3) Standford University, Artificial Intelligence Index Report 2025*, 2025, p. 4.*

disparité des statuts, le Royaume-Uni étant par exemple généralement compris dans cet ensemble du classement [1].

Ce déficit en recherche aboutit à un nombre limité de brevets déposés. En 2024, près de 82 % des dépôts relatifs à l'IA provenaient de pays de la zone Asie de l'Est et Pacifique, suivis lointainement par ceux de la région Amérique du Nord avec environ 14 % du total des brevets [2] et, encore plus péniblement, par la région Europe et Asie centrale qui contribue pour à peine 3 % du total des dépôts de brevets sur des technologies liées à l'IA [3]. Rapportés à la population, ces chiffres démontrent que l'intensité de la recherche n'est pas uniquement due à la démographie. En effet, si la Chine et les États-Unis sont respectivement troisième et quatrième avec six et cinq dépôts de brevets d'IA par tranche de 100 000 habitants, c'est la Corée du Sud qui mène ce classement avec plus de dix-sept dépôts pour le même ratio. En comparaison, la France ne dépose que 0,43 brevet selon la même grille de lecture [4]. Hors du champ académique et expérimental, ce constat se traduit par une plus faible adoption de l'IA en Europe qu'en Asie et aux États-Unis. Ainsi, en 2024, 13,5 % seulement des entreprises situées sur le territoire de l'Union européenne déclaraient recourir à l'IA pour leurs activités [5].

Le développement d'un écosystème européen de l'IA est également encore rudimentaire. Si, en 2024, l'Union européenne comptait environ 6 300 *start-ups* spécialisées dans ce secteur, dont environ 10,6 % peuvent être classées comme *start-ups* spécialisées dans l'IA générative, des disparités importantes existent entre les vingt-sept pays membres. La plupart de ces petites entreprises se trouvent en Allemagne (19,9 %), en France (17,5 %), aux Pays-Bas (10,9 %) et en Suède (8,2 %) [6]. Ces *start-ups* se concentrent principalement sur le développement d'applications secondaires exploitant des systèmes d'IA déjà existants. Une proportion plus marginale travaille au développement de modèles de base et à la construction d'infrastructures pour les IA génératives. L'accès à la capacité de calcul nécessaire à la formation de leurs modèles ainsi que les financements correspondants font cependant encore défaut. La capacité de calcul européenne dédiée à l'IA représente en effet moins de 5 % du volume mondial, contre 75 % pour les États-Unis et 15 % pour la Chine [7].

---

(1) Ibid, p.33.

(2) Les résultats de la zone Asie de l'Est et Pacifique sont largement portés par la Chine, qui dépose près de 70 % de ce sous-total. Pour la région Amérique du Nord, les dépôts sont presque exclusivement américains.

(3) Standford University, Op. cit. p. 44.

(4) Ibid, p. 46.

(5) Organisation de coopération et de développement économiques (OCDE), Progress in Implementing the European Union Coordinated Plan on Artificial Intelligence (Volume 1) Member States' Actions, *10 novembre 2025, p. 70.*

(6) Ibid.

(7) Commission européenne, Communication de la Commission au Parlement européen et au Conseil : Une stratégie européenne relative à l'intelligence artificielle dans le domaine de la science – Poser les jalons du centre de ressources de la science pour et par l'IA en Europe (RAISE), *8 octobre 2025, p. 2.*

Si l'Union européenne est pourtant capable de produire des champions du numérique, à l'image du finlandais Nokia, du néerlandais Philips, du suédois Spotify ou encore du lituanien Vinted, la tâche apparaît plus complexe pour le secteur spécifique de l'IA. Alors que des licornes performantes telles que le français Mistral ou l'allemand Aleph Alpha ont développé des solutions innovantes, les rapporteurs ont été consternés d'entendre que plusieurs entreprises et gouvernements européens privilégient des partenariats avec les grandes firmes américaines pour équiper leurs infrastructures au cours des prochaines années. Une prise de conscience générale semble toutefois s'amorcer, comme en témoigne le récent sommet pour l'action sur l'intelligence artificielle organisé à Paris au début de l'année 2025.

Le sommet pour l'action sur l'intelligence artificielle : quel bilan pour l'industrie française et européenne ?

Du 6 au 11 février 2025, la France a accueilli à Paris le sommet pour l'action sur l'intelligence artificielle. Coprésidé par la France et l'Inde, cet événement international avait pour ambition de mettre en avant le savoir-faire européen et français dans le domaine tout en promouvant une utilisation raisonnée de la technologie afin de déterminer des pistes de gouvernance mondiale.

Le président de le République, Emmanuel Macron y a annoncé un grand plan d'investissement dans les technologies d'IA à hauteur de 109 milliards d'euros au cours des prochaines années. Une somme importante sur laquelle la présidente de la Commission européenne, Ursula von der Leyen, a renchéri en annonçant 200 milliards d'euros d'investissements européens supplémentaires. Les fonds seront toutefois principalement privés et dépendront donc en grande partie du bon vouloir des entreprises.

D'autres annonces sont venues compléter le tableau d'ensemble. La création d'une « fondation internationale sur l'IA d'intérêt général » et d'une « coalition pour l'IA durable sur le plan environnemental » [(1)] a été actée avant la signature de la déclaration finale sur une « intelligence artificielle inclusive et durable pour les peuples et la planète », par une soixantaine de pays et organisations internationales.

En France, ces discours ont commencé à se concrétiser au cours du sommet *Choose France,* le 19 mai 2025 [(2)]. Le fonds canadien Brookfield a confirmé un investissement de 10 milliards d'euros pour la création de centres de données de nouvelle génération, dont un site pilote à Cambrai avec une promesse de création de 4 000 emplois [(3)]. Le fonds émirati MGX, associé à Bpifrance, à Mistral AI et à Nvidia, a engagé 8,5 milliards d'euros pour établir en Île-de-France un campus d'envergure consacré à l'IA, intégrant supercalculateurs, infrastructures de formation, centres de données et laboratoires de recherche [(4)]. Le groupe américain Prologis a quant à lui annoncé un plan de 6,4 milliards d'euros pour la construction de quatre centres de données en région parisienne, représentant une capacité installée de 584 mégawatts et la création attendue de 3 400 emplois. Enfin, le spécialiste texan de la gestion de centres de données, Digital Realty, a confirmé la réalisation de deux projets structurants à Marseille et Dugny, pour un total de 2,3 milliards d'euros d'investissement et 750 emplois, venant renforcer les capacités françaises en matière de connectivité, de sécurité et d'hébergement de données stratégiques.

(1) Présidence de la République, Les actions de Paris pour l'intelligence artificielle*, 11 février 2025, 12 pages.*

(2) Ministère chargé de l'industrie et de l'énergie, Communiqué de presse « Sommet Choose France 2025 *», 20 mai 2025*, 2 pages.

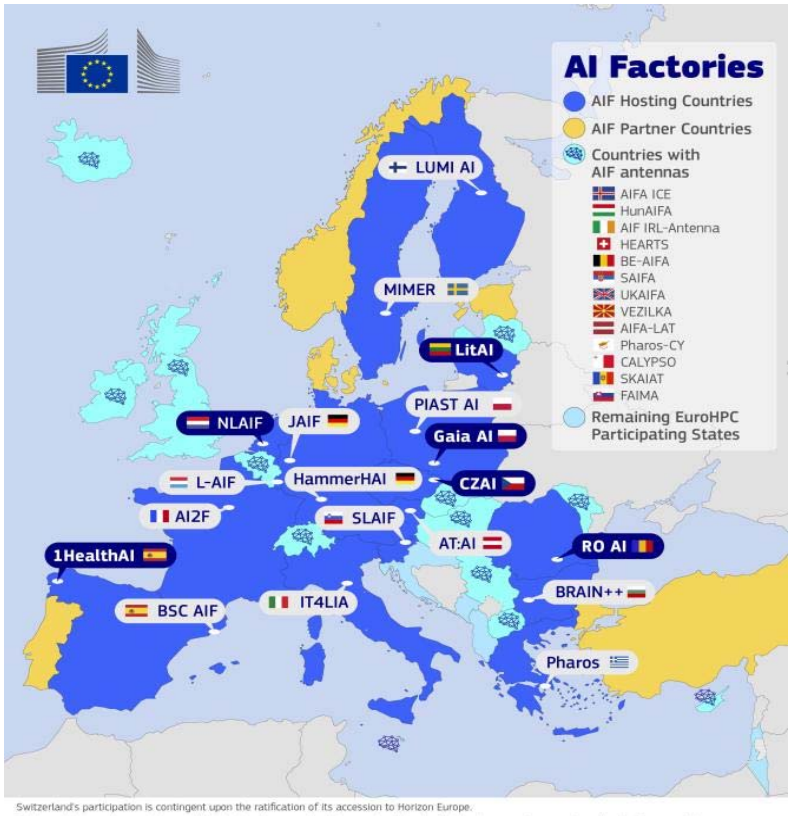
(3) Présidence de la République, Dossier de presse « Choose France, 8e édition du Sommet », 19 mai 2025, p. 9.

(4) Ibid, p. 8

Parallèlement, la France entend poursuivre la mise en œuvre de sa stratégie pour l'IA en misant sur la formation. L'objectif affiché est ainsi de former 100 000 jeunes par an aux technologies de l'IA [1].

Au niveau européen, le sommet pour l'action sur l'intelligence artificielle a constitué un catalyseur stratégique majeur pour la politique industrielle en matière d'IA. À la suite de cet évènement, la Commission européenne a dévoilé, en avril 2025, un « plan d'action pour le continent de l'IA », feuille de route audacieuse destinée à faire de l'Europe un chef de file mondial dans ce domaine [2].

Afin de donner corps à cette ambition, plusieurs initiatives concrètes ont été formalisées, parmi lesquelles figurent la création d'usines d'IA (*AI factory*) et de *gigafactories* [3] de l'IA, cofinancées par la Commission européenne et les États membres à hauteur de 2 milliards d'euros. La mise en production de ces centres devrait augmenter sensiblement la puissance de calcul disponible en Europe tout en favorisant une meilleure intégration des données et de l'énergie dans un maillage consolidé du territoire. Le premier supercalculateur exaflopique [4] européen, élaboré par le français Atos et financé par l'Union européenne et l'Allemagne [5], vient ainsi d'être inauguré près de Cologne.



(1) *Comité interministérielle de l'intelligence artificielle,* Faire de la France une puissance de l'IA*, 6 février 2025, p. 8.*

(2) *Commission européenne*, Communication de la Commission au Parlement européen et au Conseil : Le plan d'action pour le continent de l'IA, *9 avril 2025, 25 pages*.

(3) Gigafactory *est un anglicisme désignant une infrastructure industrielle à très grande échelle dédiée à la production, au déploiement et à l'exploitation de modèles d'intelligence artificielle.*

(4) *La puissance maximale de calcul est mesurée en pétaflop par seconde. Un exaflop équivaut à un milliard de milliards de calculs par seconde (cf. carte p. 12).*

(5) *Le Monde, AFP,* « L'Europe lance Jupiter, son supercalculateur, destiné à rattraper son retard dans l'IA », *5 septembre 2025.*

En complément, la Commission européenne souhaite faciliter l'accès aux financements pour les chercheurs et les *start-ups*. En mobilisant des fonds dans le cadre de l'initiative « GenAI4EU » [(1)], les programmes « Horizon Europe » et « Europe numérique » ainsi que le Conseil européen de l'innovation donneront par exemple accès à une enveloppe de 700 millions d'euros [(2)] pour soutenir de nouveaux projets.

B. FACE AUX MENACES, UNE DOUBLE RÉPONSE RÉGLEMENTAIRE ET OPÉRATIONNELLE CONSTITUE UNE DÉFENSE PERFECTIBLE

1. Novateur, le cadre juridique européen érige un modèle de régulation protecteur mais demeure complexe et peu lisible

À défaut de s'être entendus rapidement sur un financement commun des projets d'IA, les vingt-sept ont réagi sur le plan normatif. Ils ont ainsi conscience de la capacité de leur pouvoir régulateur à infuser au niveau mondial, selon un phénomène qualifié d' « effet Bruxelles » [(3)]. En définissant en premier un cadre normatif, l'Union européenne peut espérer une extériorisation et une diffusion *de facto* ou *de jure* de la législation communautaire au-delà de ses frontières *via* des mécanismes de marché.

Forte de ce constat, l'Union européenne est actuellement le seul acteur à avoir mis en place une règlementation encadrant l'usage des réseaux sociaux et de l'intelligence artificielle sur son territoire. Cette régulation s'est faite en deux temps, par l'adoption de deux textes majeurs : le règlement (UE) 2022/2065 du Parlement européen et du Conseil du 19 octobre 2022 relatif à un marché unique des services numériques et modifiant la directive 2000/31/CE (règlement sur les services numériques – *Digital Services Act*, DSA) ; le règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle – *AI Act*).

Le DSA modernise la directive sur le commerce électronique [(4)] dont le champ était devenu moins pertinent au regard de l'essor des écosystèmes algorithmiques. Ce règlement repose sur un principe absolu : « *ce qui est illégal hors ligne doit également être illégal en ligne* » [(5)]. Cette nouvelle maxime a pour objectif de préserver la continuité du cadre juridique entre l'espace physique et la

(1) GenAI4EU regroupe plus de trente appels à projets consacrés au développement de l'IA générative.

(2) Commission européenne, Op. cit. p. 17.

(3) Anu Bradford, The Brussels Effect: How the European Union Rules the World, *Oxford University Press Inc, 27 mars 2020, 424 pages.*

(4) Directive 2000/31/CE du Parlement européen et du Conseil du 8 juin 2000 relative à certains aspects juridiques des services de la société de l'information, et notamment du commerce électronique, dans le marché intérieur («directive sur le commerce électronique»).

(5) Conseil de l'Union européenne, *Communiqué de presse : Ce qui est illicite hors ligne devrait aussi l'être en ligne: le Conseil arrête sa position sur la législation sur les services numériques* », 25 novembre 2021, 2 pages.

sphère numérique. La Commission européenne supervise directement les obligations de contrôle et de transparence des très grandes plateformes en ligne (VLOP) et des très grands moteurs de recherche (VLOSE), avec l'appui du centre européen pour la transparence des algorithmes (ECAT), inauguré en 2023 et chargé d'évaluer les systèmes de recommandation, de classement et de modération [1].

Les obligations susmentionnées correspondent à plusieurs mesures visant à renforcer la protection des utilisateurs et à prévenir la diffusion de contenus illicites. Des mécanismes de notification et de retrait des publications manifestement illégales ont notamment été imposés aux plateformes afin de repérer et supprimer les diffusions liées aux sujets suivants : incitation à la haine, harcèlement, pédopornographie, terrorisme, commerce de produits illicites [2]. Les hébergeurs doivent également publier des rapports de transparence précisant le total de notifications reçues, leur délai de traitement et les actions de modération déployées. Ces obligations sont appliquées de manière proportionnée, en prenant en compte la taille, le rôle et la capacité d'influence des acteurs. Les plus grosses plateformes sont tenues d'évaluer chaque année les risques systémiques qu'elles génèrent – désinformation, manipulation électorale, contenus préjudiciables –, puis de mettre en œuvre des mesures correctives allant de l'adaptation des algorithmes à la mise en avant de sources d'information fiables. Il leur est également demandé de se soumettre annuellement à des audits indépendants sous la supervision de la Commission européenne et d'ouvrir l'accès à leurs données aux chercheurs,

Parallèlement, le DSA encadre plus strictement les pratiques publicitaires et les mécanismes d'influence en ligne. La publicité ciblée fondée sur des données sensibles – convictions politiques, orientation sexuelle, état de santé ou religion – est désormais interdite [3], de même que tout ciblage publicitaire à destination des mineurs. Les plateformes doivent par ailleurs rendre transparents leurs systèmes de recommandation et proposer une alternative non fondée sur le profilage [4]. Les pratiques manipulatoires (*dark patterns*), destinées à orienter l'utilisateur contre son intérêt, sont également prohibées [5].

En cas de manquement, la Commission peut faire usage de ses pouvoirs d'enquête et ordonner des changements techniques, infliger des sanctions financières ou aller jusqu'à restreindre l'accès au marché européen. Ainsi, en octobre 2024, la Commission européenne a rendu des conclusions préliminaires indiquant que Meta et TikTok n'avaient pas fourni un accès suffisant aux données, risquant en conséquence des amendes pouvant atteindre 6 % de leur chiffre

---

(1) Commission européenne, *« Le cadre de coopération au titre de la législation sur les services numériques »*, 12 février 2025.

(2) DSA, article 9.

(3) DSA, article 26.

(4) DSA, article 39.

(5) DSA, article 25.

d'affaires mondial [1]. Au-delà, les manipulations avérées de l'algorithme de TikTok lors de l'élection présidentielle roumaine (*cf. supra*) ont pu faire l'objet d'un contrôle par la Commission en application des dispositions du DSA.

Enfin, la supervision des très grandes plateformes est financée par une redevance annuelle plafonnée à 0,05 % de leur revenu net mondial. Le DSA marque donc une rupture dans la régulation mondiale du numérique, en attribuant aux plateformes une responsabilité proactive dans la réduction des risques sociaux liés à leurs architectures algorithmiques. Toutefois, certaines critiques s'élèvent, notamment aux États-Unis, dénonçant une forme de censure et une tentative de fragilisation des géants technologiques américains.

En complément, au vu de l'importance croissante de l'IA et de sa pénétration du monde numérique, les vingt-sept ont choisi de renforcer leur arsenal règlementaire en se dotant d'un texte spécifique sur la question. Adopté le 13 mars 2024, l'*AI Act* fixe des règles uniformes pour encadrer la conception, la commercialisation et l'usage des systèmes d'IA. Pour ce faire, l'Union européenne adopte une approche graduée, fondée sur la nature et l'intensité du risque généré par les systèmes d'intelligence artificielle.

L'*AI Act* distingue quatre niveaux de risques :

- le risque inacceptable – formellement interdit, il correspond aux IA utilisées pour des pratiques contraires aux valeurs de l'Union européenne telles que la manipulation subliminale, l'exploitation des vulnérabilités, la notation sociale et la catégorisation biométrique ;

- le haut-risque – il s'applique aux IA pouvant porter atteinte à la sécurité des personnes ou à leurs droits fondamentaux, ce qui justifie que leur développement soit soumis à des exigences renforcées ;

- le risque limité – il concerne les systèmes d'IA qui interagissent avec des personnes physiques, génèrent des contenus ou détectent des émotions ;

- le risque minimal – il recouvre les autres systèmes d'IA d'usage général et ne prévoit pas d'obligation particulière.

L'architecture du règlement repose également sur une responsabilisation accrue des fournisseurs de systèmes à haut risque. Ils sont tenus de mettre en œuvre des mécanismes de gestion des risques couvrant tout le cycle de vie du produit, d'assurer une gouvernance rigoureuse des données et de garantir la traçabilité des processus techniques [2]. Leurs obligations incluent en outre la constitution d'une documentation technique complète, ainsi que la garantie de niveaux élevés de robustesse, d'exactitude et de cybersécurité.

---

(1) Commission européenne, Communiqué de presse : « Selon les constatations préliminaires de la Commission, TikTok et Meta ne respectent pas leurs obligations de transparence au titre du règlement sur les services numériques »*, 24 octobre 2025, 2 pages.*

(2) AI Act*, article 9.*

Concomitamment, l'*AI Act* instaure un régime de sanctions, calqué sur le modèle du règlement général sur la protection des données (RGPD). Les amendes peuvent atteindre 7 % du chiffre d'affaires mondial ou sont déterminées par seuil avec un plafond fixé à 35 millions d'euros, selon la gravité du manquement et la taille de l'entreprise.

L'*AI Act* a suscité de nombreuses critiques à l'étranger mais aussi et surtout de la part d'entreprises européennes. Certains jugent la régulation trop rigide et susceptible de désavantager les acteurs émergents face à la concurrence internationale. Selon eux, les contraintes imposées sont coûteuses et risquent de freiner l'innovation dans un secteur en plein essor, au profit d'autres puissances étrangères. En juillet 2025, quarante-cinq entreprises européennes ont, à cet effet, signé une lettre ouverte appelant la Commission à suspendre le calendrier d'implémentation de certaines obligations, jugées prématurées ou excessives (1). Ces inquiétudes font écho à celles formulées dans le rapport Draghi sur la compétitivité européenne, qui relève des difficultés d'application des mesures d'encadrement de l'IA : « *les entreprises numériques sont dissuadées d'exercer leurs activités dans l'ensemble de l'UE par l'intermédiaire de filiales, étant donné qu'elles sont confrontées à des exigences hétérogènes, à une prolifération d'agences de régulation et à une « surtransposition » de la législation de l'UE par les autorités nationales. Les limitations du stockage et du traitement des données engendrent des coûts de mise en conformité élevés et entravent la création de grands ensembles de données intégrés pour l'entraînement des modèles d'IA. Enfin, l'existence de nombreuses règles nationales différentes en matière de marchés publics entraîne des coûts récurrents élevés pour les fournisseurs d'informatique en nuage. Cette charge réglementaire a pour conséquence que seules les grandes entreprises — souvent établies hors de l'UE — disposent de la capacité financière de supporter les coûts de mise en conformité et sont incitées à le faire. Les jeunes entreprises technologiques innovantes peuvent décider de ne pas du tout opérer dans l'UE.* » (2).

La mise en œuvre opérationnelle du règlement rencontre également des difficultés pratiques. En septembre 2025, la médiatrice européenne, Teresa Anjinho, a ouvert une enquête à la suite de signalements relatifs à des irrégularités de supervision par la Commission du processus de normalisation technique (3). Des ONG ont dénoncé un manque de transparence et d'inclusivité de la composition des comités compétents, où les grandes entreprises technologiques seraient surreprésentées au détriment de la société civile. Des préoccupations similaires émergent autour du code des bonnes pratiques de l'IA à usage général, publié en juillet 2025, critiqué car la rédaction en aurait été influencée par les grands acteurs américains du numérique. Les responsables auditionnés à Bruxelles par la

---

(1) Alexandre Piquard, « IA : 45 entreprises européennes demandent une « pause » dans l'application de l'AI Act », *Le Monde, 4 juillet 2025.*

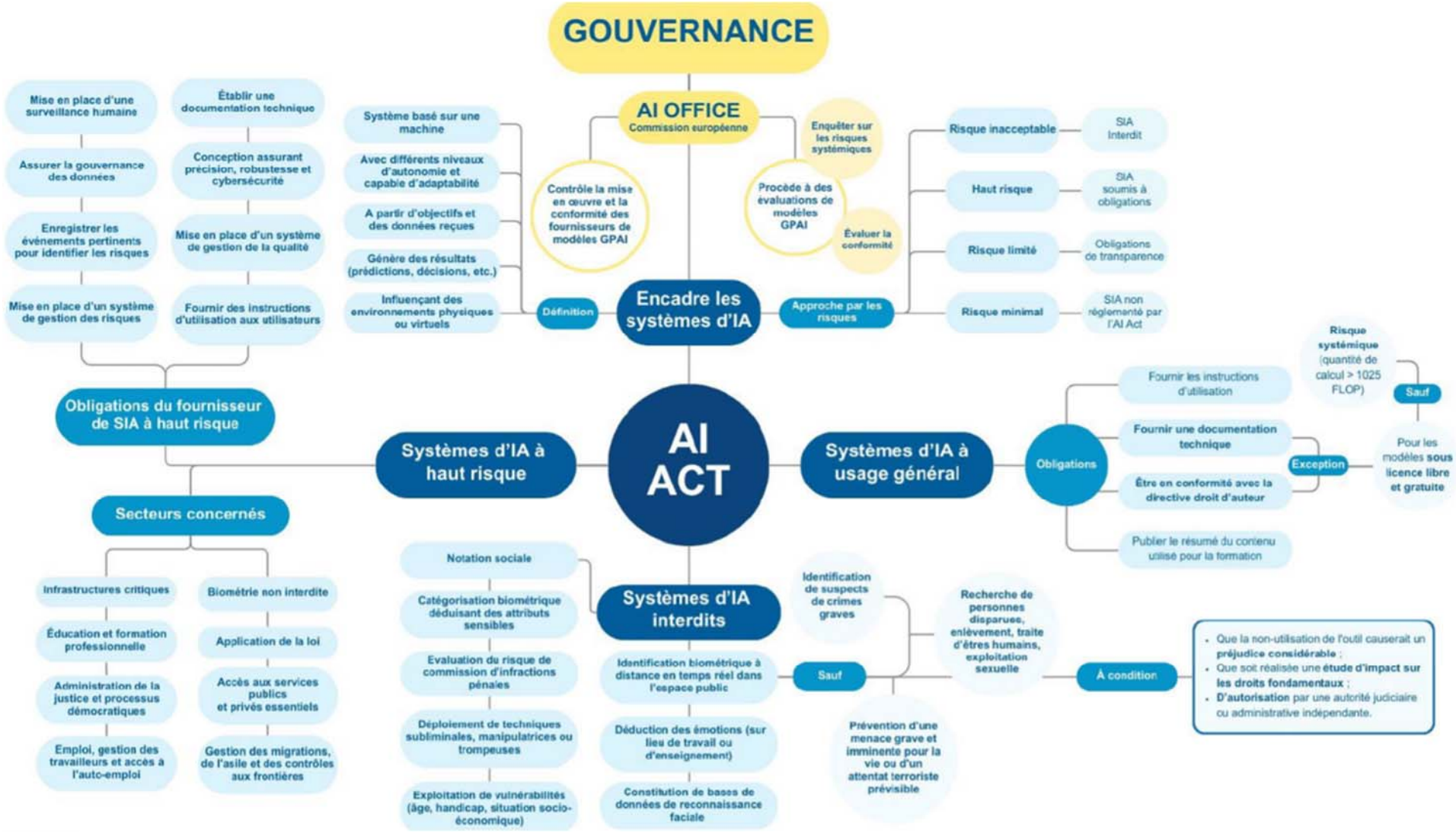
(2) Mario Draghi, Op. cit, p. 34.

(3) Médiateur européen, Affaire 1974/2025/MIK, 26 septembre 2025.

rapporteure n'ont pas relevé ces éléments mais ont mentionné des retards dans l'édiction des dispositions techniques permettant la pleine application du règlement.

La Commission européenne semble sensible à ces difficultés. Afin d'y répondre, elle a, entre autres, proposé, lors du sommet sur la souveraineté numérique européenne, le 18 novembre 2025, un report de l'entrée en vigueur des obligations applicables aux systèmes à haut risque à décembre 2027. Cette application différée s'accompagnerait d'aménagements spécifiques pour les petites et moyennes entreprises afin d'alléger leur charge administrative (1). La réalisation de ces annonces suppose toutefois qu'elles soient acceptées par les États-membres et le Parlement européen.

(1) Virginie Malingre, « La Commission européenne lance la simplification de la réglementation numérique », *Le Monde, 19 novembre 2025.*



Source : Laure de Roucy-Rochegonde, « Promesses artificielles ou régulation réelle ? Inventer la gouvernance mondiale de l'IA », Études de l'IFRI, IFRI, février 2025, p. 31.

2. Les moyens opérationnels de contre-ingérence ont été renforcés mais paraissent encore insuffisants au regard de l'ampleur de la menace

Les éléments juridiques mentionnés fournissent un cadre institutionnel règlementant l'utilisation de l'outil IA. Ce seul carcan n'est pas une panacée et doit être doublé d'une réponse opérationnelle face aux ingérences étrangères recourant à cette technologie. La structuration des moyens de contre-attaque se fait en partie au niveau européen et surtout à l'échelon national.

a. L'Union européenne surveille les campagnes de désinformation grâce à son service pour l'action extérieure

Une première réaction communautaire a été formulée par la Commission, en 2018, avec l'adoption d'un plan d'action contre la désinformation (1). Pilier de la stratégie européenne de sécurité et de résilience informationnelle, il comporte quatre axes prioritaires : le renforcement des capacités de détection et d'analyse, l'amélioration de la coordination entre les États membres, la mobilisation du secteur privé et des plateformes numériques, ainsi que le renforcement de la résilience des sociétés face aux manipulations de l'information (2). C'est dans ce contexte qu'a été pensé le *Rapid Alert System* (RAS), officiellement lancé en mars 2019 par le service européen pour l'action extérieure (SEAE). Conçu comme un dispositif de coordination interinstitutionnelle, le RAS vise à renforcer la coopération entre les États membres et les institutions de Bruxelles dans la lutte contre les campagnes de désinformation.

Le fonctionnement du RAS repose sur une approche en sources ouvertes (*open source*) (3). La plateforme, administrée par le SEAE, est alimentée par les États membres et les institutions européennes, ainsi que par des universités, des organisations de vérification de l'information, des centres de recherche et, dans certains cas, par les plateformes numériques elles-mêmes (4). Le système est également connecté au centre de crise du SEAE, au mécanisme de réaction rapide du G7, à l'OTAN, au réseau européen de coopération en matière électorale (ECNE) et à l'Observatoire européen des médias numériques (EDMO) (5). Outre le partage quotidien d'informations, la plateforme permet l'émission d'alertes prioritaires en cas d'incident majeur ou de campagne de désinformation à grande échelle. Le RAS a été mobilisé pour la première fois à la veille des élections européennes de 2019, puis durant la pandémie de Covid-19. De manière générale, le suivi des ingérences

---

(1) Commission européenne, Communication conjointe au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des régions : Plan d'action contre la désinformation, *5 décembre 2018, 15 pages.*

(2) Service européen pour l'action extérieure, « Fiche d'information: le système d'alerte rapide », 5 mars 2019.

(3) En open source*, le code informatique du logiciel concerné est accessible à tous. Chacun peut vérifier la manière dont il a été écrit et peut se l'approprier pour ses propres créations. Le choix de l'*open source *garantit une transparence aux utilisateurs.*

(4) James Pamment, « The EU's role in fighting disinformation: taking back the initiative », Carnegie endowment for international peace, 15 juillet 2020, 31 pages.

(5) Ibid.

étrangères dépend du SEAE. Ce dernier a structuré en son sein un groupe de travail (*taskforce*) dédié à la lutte informationnelle, la *East Stratcom Task Force* (ESTF). Se concentrant initialement sur les activités malveillantes russes, cette division a progressivement étendu sa veille à la surveillance d'une pluralité d'acteurs malveillants. Afin d'en assurer un suivi, elle a également créé une base de données, « EUvsDisinfo », recensant et démentant les cas de désinformation au niveau européen. À la date du 24 novembre 2025, 19 489 cas de désinformation ont été identifiés et réfutés par ce biais depuis 2016. Ces chiffres ont augmenté significativement au cours des trois dernières années, en raison notamment d'une utilisation intensive de l'IA par la Russie pour produire ses contenus fallacieux. En parallèle, l'ESTF propose aussi du contenu pédagogique à l'attention du public.

Lors des auditions menées par les rapporteurs, a également été mentionnée la création d'une future structure pour centraliser les bonnes pratiques en matière de lutte informationnelle. Annoncé par la présidente de la Commission européenne (1), ce « centre européen pour la résilience démocratique » serait proposé dans le cadre d'un dispositif intitulé « bouclier démocratique européen ». La commission spéciale éponyme du Parlement européen devra en définir les modalités d'action. Sans empiéter sur les prérogatives régaliennes des États membres, ce nouvel organe pourrait coordonner des actions de démystification et favoriser le partage de méthodes de travail entre services compétents.

b. En France, la lutte contre les ingérences étrangères mobilise plusieurs ministères et opérateurs, aussi bien dans le domaine informationnel que cyber

Au niveau national, la France a formalisé une structure plurielle pour se défendre face aux ingérences étrangères dans le cadre d'une approche globale de sécurité nationale (2). Une distinction est opérée au regard de la nature de la menace, informationnelle ou cyber.

S'agissant des attaques informationnelles, une approche interministérielle a été retenue en priorisant un aspect défensif. Elle prend la forme d'un comité opérationnel de lutte contre les manipulations de l'information (COLMI), présidé par le SGDSN et animé par le service de détection et de la caractérisation des ingérences numériques étrangères, Viginum. Le comité regroupe des représentants du ministère de l'Europe et des affaires étrangères (MEAE), du ministère des armées et des anciens combattants ainsi que du ministère de l'intérieur. Sont notamment associés la direction générale de la sécurité intérieure (DGSI), la direction générale de la sécurité extérieure (DGSE), la direction de la communication et de la presse (DCP) du MEAE, le commandement de la cyberdéfense (COMCYBER) et le pôle anticipation stratégique et orientation (ASO) des armées. Le COLMI permet ainsi de décloisonner le fonctionnement

(1) Philippe Jacqué, « La Commission européenne propose un « bouclier démocratique » communautaire », *Le Monde, 12 novembre 2025.*

(2) SGDSN, Revue nationale stratégique 2025, 14 juillet 2025, p. 73.

traditionnel en silos des différents acteurs mobilisés en assurant une coordination horizontale. En parallèle, la déclinaison verticale des impulsions décidées relève, dans un second temps, de chaque ministère en fonction de ses compétences (1). Le COLMI se réunit à un niveau stratégique chaque mois et à un niveau opérationnel chaque semaine (COLMI TECH). Il formule des orientations de travail pour réagir face aux manipulations de l'information, ainsi que des propositions de réponse à appliquer.

Au cœur du dispositif, Viginum a été créé par décret en juillet 2021 (2) afin d'identifier les ingérences numériques étrangères affectant le débat public en France. À cette fin, le service assure une veille des « phénomènes inauthentiques » caractérisés par :

- une atteinte potentielle aux intérêts fondamentaux de la Nation ;
- un contenu manifestement inexact ou trompeur ;
- une diffusion artificielle ou automatisée, massive et délibérée ;
- l'implication, directe ou indirecte d'un acteur étranger étatique, paraétatique ou non-étatique.

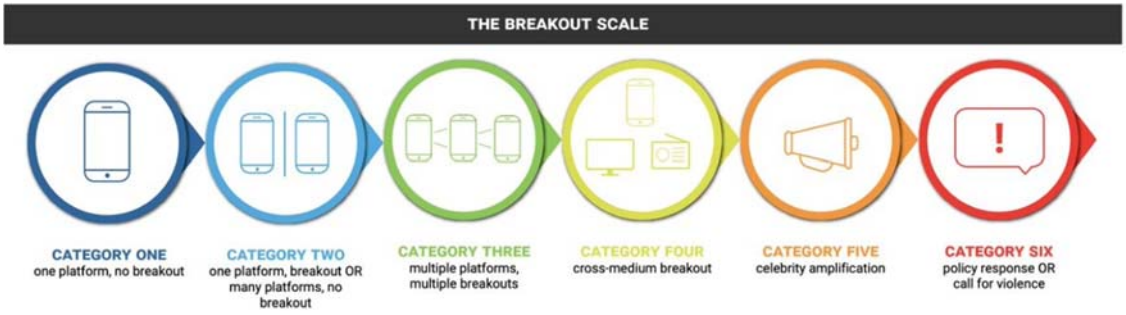
L'identification des campagnes de manipulation de l'information par Viginum

La détection d'une campagne de manipulation de l'information est un exercice complexe impliquant une méthodologie mouvante. Elle se fonde sur une approche par les conséquences ou, du moins, par une analyse du risque engendré. Il est en effet primordial de ne pas donner une visibilité plus grande à une action malveillante si elle peut être simplement ignorée. Pour ce faire, Viginum dispose de plusieurs outils :

- l'échelle de percée (*breakout scale*) du chercheur Ben Nimmo – ce modèle comparatif mesure la dynamique de diffusion des opérations d'influence à partir de données observables, réplicables et vérifiables. Il prend la forme d'une échelle comprenant six seuils de distribution : entre plusieurs plateformes, différentes communautés, et jusqu'aux médias, personnalités publiques et politiques. Cet outil est toutefois lié à une analyse ponctuelle. Il se prête donc moins à une étude de fond et à l'anticipation des phénomènes ;

(1) Rachid Temal, rapport n° 739 fait au nom de la commission d'enquête sur les politiques publiques face aux opérations d'influences étrangères visant notre vie démocratique, notre économie et les intérêts de la France sur le territoire national et à l'étranger afin de doter notre législation et nos pratiques de moyens d'entraves efficients pour contrecarrer les actions hostiles à notre souveraineté, Sénat, 23 juillet 2024, p. 116.

(2) Décret n° 2021-922 du 13 juillet 2021 portant création, auprès du secrétaire général de la défense et de la sécurité nationale, d'un service à compétence nationale dénommé « service de vigilance et de protection contre les ingérences numériques étrangères ».



- l'index du risque d'impact (*impact risk index*) – ce classement élaboré par l'ONG EU Disinfo Lab s'appuie sur un système de points, établi à partir de divers indicateurs liés à la viralité et à l'engagement généré par le contenu (1), pour déterminer quatre niveaux de risque. Par ailleurs, l'outil s'intéresse à la « surface de propagation », c'est-à-dire au nombre de plateformes et de langues utilisées, à la variété de formats déployés ou encore à la présence ou non d'un « appel à l'action » (2).

Viginum tente de synthétiser les apports de ces instruments en se détachant des métriques d'engagement, telles que les « *like* » ou les « vues » qui diffèrent selon les plateformes. L'analyse est ainsi enrichie par une description fine du phénomène observé et des techniques employées, par une évaluation de la dynamique de transmission et des moyens mobilisés, ainsi que par un examen de l'intentionnalité et du contexte de l'opération supposée.

Ce travail est réalisé par des analystes épaulés par un « *datalab* », une équipe d'experts en science des données. Ces derniers traitent les flux massifs d'informations à l'aide d'algorithmes et développent des dispositifs en sources ouvertes, accessibles publiquement sur internet. À titre d'exemple, a notamment été mis à disposition des acteurs de la société civile une méthodologie de détection des textes dupliqués de manière inauthentique *via* la publication d'un code python appelé « D3lta » (3). Concrètement, ce code permet de calculer des scores de similarité sémantique afin d'identifier les contenus reproduits par reformulation, par *copy-pasta* ou par simple traduction d'une langue à une autre.

L'action de Viginum est relayée par la direction de la communication et de la presse (DCP) du MEAE, qui en assure la traduction diplomatique. La décision finale d'agir ou non, prenant notamment la forme d'une dénonciation publique, revient donc au ministère. Une sous-direction spécifique « veille et stratégie » a été créée en 2022 au sein de la DCP afin d'affiner la détection et les réponses à apporter aux tentatives d'ingérences numériques étrangères. Cette sous-direction est structurée en deux pôles, l'un consacré à la veille des réseaux sociaux et à l'identification des tendances informationnelles émergentes ; l'autre dédié aux actions de riposte et à la communication d'influence. Cette « communication d'influence » complète utilement la communication traditionnelle et vise à structurer les opinions publiques en s'appuyant sur des tiers afin de délivrer des messages parfois inaudibles en cas de diffusion par les voies institutionnelles. Au-delà, le MEAE travaille également en étroite collaboration avec les partenaires

(1) L'engagement est ici entendu en tant qu'ensemble de réactions au contenu : nombre de partages, commentaires et vues.

(2) Viginum, « Mesurer le risque d'impact d'une campagne de manipulation de l'information : quelques réflexions », *10 juillet 2025.*

(3) Viginum, « L'IA au service de la lutte contre les manipulations de l'information : zoom sur le Datalab de Viginum », *27 février 2025.*

européens afin de faire appliquer la législation communautaire et d'affermir une prise de conscience collective de la menace.

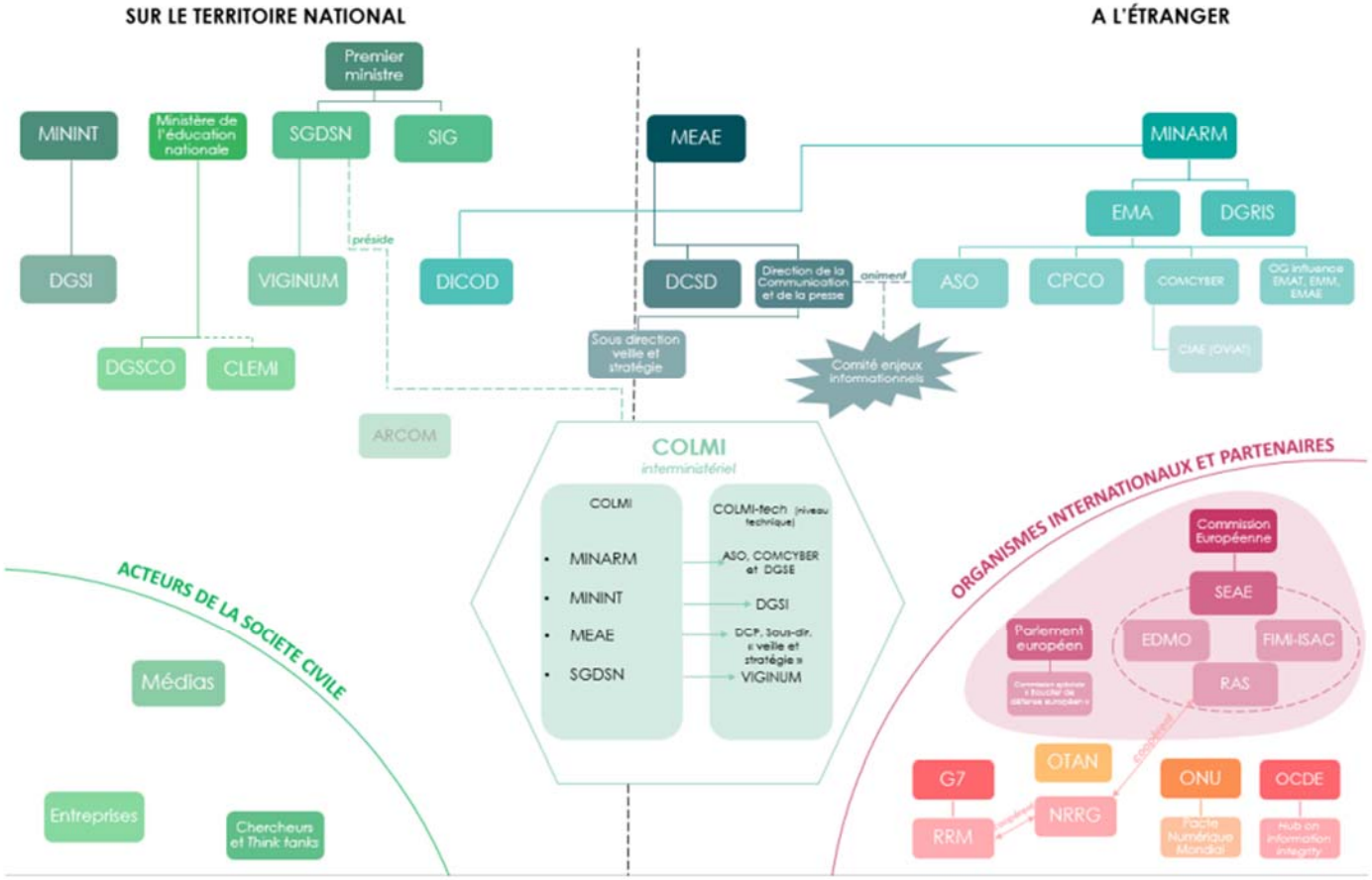
Si le MEAE assure le pilotage de la stratégie d'influence au niveau interministériel, une cellule de coordination similaire à la DCP existe aussi au sein du ministère des armées et des anciens combattants depuis 2022. Ce pôle ASO a pour objectif la structuration de la « fonction stratégique influence » (1) dans les armées. Son action s'inscrit dans le contexte plus global de la lutte informatique d'influence (L2I) (2). Dans ce cadre, des opérations militaires, commandées par le chef d'état-major des armées, qui en délègue le contrôle à l'officier général COMCYBER, répondent à un triple objectif de détection des attaques informationnelles, de caractérisation desdites attaques et, enfin, de contre-attaque.

Par leurs actions, la DGSE et la DGSI contribuent également à la défense face aux ingérences étrangères. Grâce à leur expertise, elles sont notamment capables d'attribuer formellement une attaque à des services de renseignement étrangers. Au-delà, lorsqu'une menace est caractérisée et que les personnes physiques ou morales qui en sont à l'origine sont identifiées, des moyens d'entrave préventive peuvent être déployés par la DGSI sur le sol français. Un agent d'influence est ainsi susceptible de se voir refuser l'accès au territoire national. En complément, la DGSI dispose de capacités répressives et peut procéder à des arrestations. Compte tenu du secret de la défense nationale, le détail des moyens dont disposent les services de renseignement en matière de prévention des menaces informationnelles ne pourra pas être communiqué dans ce rapport.

(1) SGDSN, Revue nationale stratégique 2022*, 9 novembre 2022, p. 10.*

(2) Ministère des armées et des anciens combattants, « Éléments publics de doctrine militaire de lutte informatique d'influence (L2I) », *2021, 14 pages.*

CARTOGRAPHIE DES ACTEURS DE L'INFLUENCE ET DE LA LUTTE CONTRE LES MANIPULATIONS DE L'INFORMATION



Source : Natalia Pouzyreff, Marie Récalde, rapport d'information n° 1661 sur le thème de « l'opérationnalisation de la nouvelle fonction stratégique influence », Assemblée nationale, XVII^e législature, 2 juillet 2025, p. 37.

Concernant la cyberdéfense, le centre de coordination des crises cyber (C4) a été créé en 2018, à la suite de la publication de la revue stratégique de cyberdéfense [1]. Présidé par le SGDSN, il comprend l'agence nationale de la sécurité des systèmes d'information (ANSSI), le COMCYBER, la direction générale de l'armement (DGA), la DGSI et la DGSE dans un format opérationnel, auxquels s'ajoute le MEAE dans une configuration stratégique. Le C4 a pour mission le partage de connaissances et de compétences en matière de cyberdéfense, afin de renforcer les capacités d'anticipation et de réaction face à la menace cyber. Sur un plan technique, il conduit des travaux de détection, de caractérisation et d'imputation des attaques. Au niveau stratégique, il propose aux autorités politiques des axes de réponse opérationnels.

Si l'IA décuple assurément la fréquence et l'intensité des attaques cyber (*cf. supra*), les interlocuteurs auditionnés par les rapporteurs ont tempéré ce constat. Selon eux, l'outil IA apparaît plus efficace pour la défense que pour l'attaque. Il est ainsi déjà mis à profit par l'ANSSI, qui utilise un algorithme d'apprentissage machine, « AnoMark », pour détecter des anomalies dans des lignes de commande au sein des journaux d'événements système [2]. De même, Google propose de recourir à l'IA générative dans le cadre de ses solutions de cybersécurité. Son « *Agentic SecOps* » fonctionne comme un assistant conversationnel en permettant d'identifier plus rapidement une compromission dans un réseau. Il correspond à une nouvelle déclinaison de l'IA, appelée « IA agentive » [3], et renforce la sécurité des systèmes informatiques en traitant en continu les tâches les plus chronophages et répétitives de détection. Cet aspect défensif de l'IA, loué notamment par plusieurs acteurs auditionnés aux États-Unis, présente néanmoins une limite majeure : pour en bénéficier, il apparaît nécessaire de disposer de capacités technologiques importantes et souveraines en matière d'IA.

---

(1) SGDSN, Revue stratégique de cyberdéfense*, 12 février 2018, 167 pages.*

(2) Alexandre Junius, ANSSI, « Détection d'Anomalies dans des lignes de commande à l'aide de chaînes de Markov »*, juin 2022, 13 pages.*

(3) L'IA agentive implique un « agent IA » qui, contrairement à l'IA traditionnelle, peut définir des objectifs, planifier et exécuter des tâches avec une intervention humaine limitée au minimum. Il ne se contente pas d'analyser mais peut décider et appliquer une action de manière autonome.

III. DIX-HUIT MESURES POURRAIENT ÊTRE PORTÉES AFIN DE SÉCURISER NOS DÉMOCRATIES EN INTÉGRANT LE RÔLE DE L'INTELLIGENCE ARTIFICIELLE DANS L'ÉVOLUTION DES MENACES

Les rapporteurs proposent dix-huit actions opérationnelles concrètes. La réalisation de ces solutions peut être envisagée rapidement, à moindre coût et permettrait de mieux protéger notre pays dans le contexte fluctuant de la guerre hybride.

A. UNE RÉPONSE ADAPTÉE IMPLIQUE UNE MEILLEURE APPRÉHENSION DE LA TECHNOLOGIE DE L'IA, À LA FOIS EN TERMES DE PRODUCTION, DE DIFFUSION ET DE CONSOMMATION

1. Une maîtrise des technologies d'IA est un impératif afin de se prémunir des usages malveillants par des acteurs étrangers

La dépendance en matière d'IA de l'Europe et de la France à l'égard d'acteurs étrangers est préoccupante. Elle fragilise notre défense face à de potentielles ingérences numériques, informationnelles et cyber. Une réaction est attendue au niveau de l'Union européenne, qui demeure le niveau idoine pour penser notre souveraineté technologique. En effet, avec 450 millions d'habitants et un produit intérieur brut (PIB) de près de 18 000 milliards d'euros en 2025 (1), l'Europe constitue un marché dynamique capable de rivaliser avec les géants chinois et américains. La remontée en puissance de l'Union européenne prendra du temps et des projets d'ampleur viennent tout juste d'être lancés pour amorcer une réindustrialisation dans le secteur de l'IA. Si ces initiatives sont positives, elles paraissent toujours alimentées majoritairement par des fonds étrangers (*cf. supra*). Une proportion plus importante d'investissements européens dans le cadre des appels à contributions devrait être encouragée pour assurer, à moyen terme, une plus grande autonomie du continent. S'agissant des structures nécessaires à ce redéploiement, il ne semble pas utile d'ajouter de nouvelles couches à un millefeuille institutionnel déjà extrêmement dense. Certains programmes ou organes existent déjà – *Invest EU*, le Fonds européen pour l'innovation, le Conseil européen de l'innovation – et pourraient catalyser les efforts si une impulsion politique est déterminée. Des garanties publiques sur les prêts bancaires aux entreprises développant l'IA, une commande publique liée à l'IA d'origine européenne ou encore des contrôles renforcés des acquisitions dans l'IA stratégique ainsi qu'un éventuel fléchage de l'épargne vers le secteur seraient autant de pistes à approfondir au niveau communautaire. Ce qui se fait pour favoriser le développement d'une économie verte devrait être transposable, en partie, pour le renforcement de la filière IA.

(1) Données Eurostat, 2025.

Proposition : Distinguer les investissements extra-communautaires et européens et favoriser les seconds notamment par des mécanismes de réduction des risques tels que des garanties européennes de financement, par une commande publique privilégiant l'origine européenne ou par un fléchage de l'épargne vers le développement du secteur de l'IA.

Si ces actions nécessitent un certain horizon temporel, des mesures peuvent être prises immédiatement pour agir sur le volet informationnel, en collaboration avec les plateformes recourant à l'IA. Au vu des effets préjudiciables observés, et sans alourdir démesurément les contraintes définies par l'*AI Act*, il apparaît opportun de mieux encadrer les paramètres de recommandation et les pratiques algorithmiques. Des obligations de transparence sont d'ores et déjà prévues par les textes européens mais ces exigences semblent encore trop faibles. Les manipulations sur TikTok observées pendant l'élection présidentielle en Roumanie l'ont lourdement rappelé. Afin de préserver ces moments charnières de la vie démocratique, les rapporteurs proposent d'appliquer une forme de « réserve pré-électorale » en suspendant temporairement l'utilisation des algorithmes pendant la période précédant le scrutin. Cette proposition, évoquée au cours des auditions, a fait l'objet d'un large consensus, à la fois auprès des personnes issues du monde académique ou administratif que d'élus américains. De manière plus générale, redonner du choix aux utilisateurs apparaît primordial. Alors que les États-Unis critiquent la régulation européenne en mettant régulièrement en avant les dispositions du premier amendement de leur Constitution relatif à la liberté d'expression, les rapporteurs proposent d'imposer aux plateformes de laisser l'usager décider s'il souhaite voir du contenu en fonction d'un algorithme. Cette obligation pourrait prendre la forme d'un consentement explicite sur le modèle de ce qui est appliqué pour les *cookies* (1) des sites internet. La liberté d'expression, intégralement préservée, se doublerait ainsi d'une liberté de consommation.

Proposition : Rendre obligatoire le consentement aux algorithmes dans les suggestions de contenus sur les plateformes.

Proposition : Instaurer une « réserve algorithmique pré-électorale » sur les plateformes pendant une période définie en amont du scrutin afin d'éviter toute manipulation et de préserver la sincérité des votes.

En parallèle, la responsabilisation des plateformes sur l'usage de l'IA doit être renforcée. Il appartient aux sociétés concernées de participer activement à la lutte contre la propagation de la désinformation. Cette implication pourrait être formalisée par des plans de stratégie de responsabilité sociale de l'entreprise (RSE) au sein d'une rubrique « protection de la démocratie ». Une telle RSE pourrait être généralisée à toutes les entreprises privées et trouverait à s'appliquer de manière concrète *via* leurs pratiques publicitaires. Lors des auditions, le rôle connexe des annonceurs a ainsi été mentionné. En effet, environ la moitié des sites générés grâce à l'IA et recensés par l'entreprise NewsGuard diffusent des publicités. Les revenus

(1) Cet anglicisme désigne, selon le dictionnaire Larousse, un petit bloc de données transmis par un site web, à l'insu de l'internaute, pour être stocké sur la machine et récupéré par la serveur à la connexion suivante.

qu'ils perçoivent par ce biais encouragent la naissance d'autres sites artificiels trompeurs (*slops*) et échappent aux médias crédibles. Si les entreprises s'engageaient à diffuser une part de leurs publicités sur des sites d'information crédibles, pour soutenir l'intégrité de l'information, tout en arrêtant de diffuser leurs annonces marketing sur des sites non vérifiés, un canal de financement non négligeable serait redirigé vers des sources fiables. En effet, la somme perçue au titre de la publicité par les sites de désinformation a été évaluée, en 2021, à 2,6 milliards de dollars de recettes annuelles (1).

Proposition : Étendre le champ de la RSE à la protection de la démocratie et encourager, dans ce cadre, un financement vertueux de l'information par les entreprises privées et leurs pratiques publicitaires.

Au-delà, dans le cadre des réglementations actuelles, une exigence de publication d'un « score d'artificialité » par les plateformes responsabiliserait à la fois l'hébergeur et les créateurs de contenus. Cet indicateur permettrait à l'utilisateur de savoir si la publication visionnée a été générée grâce à l'IA et dans quelles proportions. Il apparaît souhaitable que cet affichage soit réalisé par les plateformes elles-mêmes car elles ont une visibilité sur les métadonnées des contenus qu'elles diffusent. Cette mesure contraignante n'est pas antinomique des intérêts des plateformes elles-mêmes. En effet, d'après une récente étude du *Georges Washington Institute* (GWI), les principaux réseaux sociaux ont vu leur trafic diminué au cours des trois dernières années après un pic d'activité en 2022 (2). Cette attractivité réduite coïncide avec la diffusion des contenus générés par l'IA, stéréotypés et peu originaux. En parallèle, les rapporteurs enjoignent les acteurs du secteur à remettre en place des mécanismes opérants de vérification de l'information.

Proposition : Imposer aux plateformes de réseaux sociaux d'étiqueter les contenus avec un score d'artificialité indiquant dans quelles proportions l'IA a été utilisée pour les générer.

2. Une population sensibilisée et préparée constitue l'élément central de la stratégie de défense d'une société résiliente face aux ingérences et aux dérives permises par l'IA

Le citoyen est le point névralgique de la défense face aux ingérences étrangères mobilisant l'IA. Première cible de ces attaques, il n'est pas une victime impuissante et peut freiner ou empêcher la pénétration des idées disruptives instillées dans la société par les acteurs malveillants. Cette capacité nécessite un apprentissage en amont et une bonne connaissance des pratiques adverses. À cet égard, la Suède fait figure d'exemple au niveau mondial. Lors de son déplacement à Stockholm, la rapporteure a étudié le concept national de « défense totale ». Pour sa composante civile, cette notion recoupe trois éléments : un aspect

(1) Matt Skibinski, « Top brands are sending 2,6 billion $ to misinformation websites each year »*, Newsguard, 2021.*

(2) John Burn-Murdoch, « Have we passed peak social media ? »*, Financial Times, 2 octobre 2025.*

logistique, par la mise en place de stocks d'approvisionnement ; un volet psychologique, visant à contrer la propagande adverse ; une protection matérielle avec la construction d'abris anti-bombes (1). Née après la seconde guerre mondiale dans le contexte de la guerre froide, la « défense totale » requiert la participation active de la population. L'objectif de préservation du territoire national doit mobiliser le civil et le militaire car l'armée ne peut protéger seule le pays. Les Suédois sont ainsi conscients de ne pas vivre en paix et d'être attaqués régulièrement par la Russie, notamment dans la sphère informationnelle. Au cours des auditions, des mots éloquents ont ainsi été rapportés : « *L'ennemi a déjà franchi la frontière, les chars sont là mais on ne les voit pas* ». En réaction, le gouvernement a réédité, à la fin de l'année 2024, une brochure à destination de tous les citoyens et intitulée : « *En cas de crise ou de guerre* ». Traduit dans plusieurs langues, et outre la mention de règles élémentaires de survie, ce livret alerte spécifiquement sur le danger des fausses informations (2). Cette sensibilisation est vue par certaines personnes interrogées comme une « vaccination » face au virus informationnel. Pour filer la métaphore, des gestes barrières sont également recommandés en complément d'une méthode d'évaluation critique des sources d'information. Un corpus de six questions préliminaires est ainsi conseillé : « *S'agit-il d'une information factuelle ou d'une opinion ? Quel est le but de cette information ? Qui a diffusé ce contenu ? La source est-elle digne de confiance ? Cette information est-elle disponible ailleurs ? Ces informations sont-elles nouvelles ou anciennes et pourquoi sont-elles diffusées à ce moment précis ?* » (3).

Ces éléments sont relayés par les médias nationaux, à l'école et dans l'espace public. Si un guide similaire vient d'être publié par le service d'information du gouvernement français (4), le format imprimé souffre, dans notre pays, d'un déficit de visibilité. Les rapporteurs proposent, face à l'urgence de la menace, de réutiliser des moyens de communication percutants en diffusant des campagnes publicitaires massives, à l'image de ce qui a pu être fait pendant la pandémie de Covid-19. Des supports tels que des messages publicitaires télévisés ou des affiches dans les transports pourraient être réalisés. Dans ce cadre, Viginum aurait un rôle primordial en rassemblant les initiatives existantes pour sensibiliser les Français et renforcer la lisibilité d'ensemble de la politique publique.

Proposition : Réinvestir l'espace public par le développement, avec Viginum, d'une communication active pour alerter sur la désinformation et l'usage malveillant de l'IA, en concentrant les efforts en période pré-électorale.

La rapporteure a également pu mesurer l'importance d'une population pleinement sensibilisée à la question des ingérences lors de son déplacement à

---

(1) Anne-Françoise Hivert, « La Suède reconstruit son modèle de « défense totale », dans lequel chaque habitant doit se préparer »*, Le Monde, 24 mars 2025.*

(2) Agence suédoise de la protection civile (Myndigheten för civilt försvar), « Informations importantes pour la population de Suède : en cas de crise ou de guerre »*, version française, novembre 2024, p. 5.*

(3) Ibid.

(4) Service d'information du gouvernement (SIG), « Tous responsables, face aux risques, agissons : bien préparés, bien protégés, tous engagés »*, 15 pages.*

Taïwan, où les habitants sont grandement mobilisés face aux intimidations répétées de la République populaire de Chine. Afin d'éviter les critiques portant sur une définition étatique de dogmes incontestables, le gouvernement se détache de l'image du ministère de la vérité orwellien (1) en s'assurant que la société civile soit neutre et indépendante. Cette garantie s'articule autour d'une coopération solide entre le secteur privé, les ONG, les *think-tanks* et le monde éducatif. L'approche taïwanaise souligne la nécessité de pouvoir compter sur des acteurs non gouvernementaux, engagés et investis pour la défense des intérêts nationaux. Cette philosophie fait pour le moment défaut en France et serait une ressource précieuse dans le cadre de la lutte informationnelle.

La résilience cognitive taïwanaise

Taïwan a développé un véritable écosystème de participation citoyenne en conjuguant réactivité gouvernementale et engagement de la société civile. La population et les ONG locales jouent un rôle crucial dans la détection des manœuvres informationnelles malveillantes ainsi que dans l'élaboration d'un contre-discours. Des plateformes collaboratives de vérification de l'information, telles que *Cofacts*, un *chatbot* intégré à la messagerie LINE, permettent aux citoyens d'éprouver la véracité des messages reçus et d'obtenir rapidement une évaluation objective des faits. En permettant de cibler des discussions personnelles, ce type d'outil est particulièrement innovant. Son utilisation dépend *in fine* toujours de la décision de l'utilisateur de consentir à l'ajout d'un robot conversationnel dans un groupe privé. En parallèle, le site indépendant *MyGoPen*, créé en 2015, recense les rumeurs en circulation et fournit au public des éléments de démystification.

Par ailleurs, Taïwan s'appuie sur les compétences du secteur privé et de *think tanks*, tels que *Doublethink Lab*, *AI Labs* et le *Taiwan Information Environment Research Center* (IORG) qui produisent des rapports détaillés sur les campagnes étrangères de désinformation ou de propagande, contribuent à la détection des *bots* à grande échelle et conseillent les acteurs publics sur les nouvelles techniques d'attaque mobilisant l'IA.

L'éducation à l'information et aux médias, aussi appelée littératie (2), est l'autre pilier de la résilience taïwanaise. Le gouvernement l'a intégrée dans les directives du programme scolaire national, les « *108 curriculum guidelines* ». La compréhension de la sphère médiatique et informationnelle est perçue comme une compétence clé à développer par les élèves tout au long du cycle primaire et secondaire (3). L'objectif est d'apprendre aux jeunes générations à vérifier les sources consultées, à repérer les biais cognitifs et à contribuer efficacement à la définition d'une information fiable. L'école forme ainsi des acteurs éclairés et non des consommateurs passifs. En outre, des ateliers ciblés sont organisés pour les individus les plus exposés. À cet effet, le programme communautaire *Fake News Cleaner* (4) organise des sessions de formation dans les collectivités locales, notamment en zones rurales, auprès des seniors.

Au niveau gouvernemental, un ministère des affaires numériques parachève cette organisation en assurant une fonction centrale dans la stratégie nationale de résilience informationnelle. Il supervise un mécanisme institutionnel de réaction rapide, qui fixe le délai de réponse officielle

(1) George Orwell, 1984, Folio, édition 2015, 1949, 418 pages.

(2) OCDE, « Les adultes possèdent-ils les compétences nécessaires pour s'épanouir dans un monde en mutation ? Évaluation des compétences des adultes 2023 », *Études de l'OCDE sur les compétences, 10 décembre 2024, p. 4.*

(3) William Hung, « Media literacy education: Taiwan's Key to combating disinformation », *Global Taiwan Institute*, *6 mars 2024.*

(4) Huizhong Wu, « In Taiwan, a group is battling fake news one conversation at a time — with a focus on seniors », *Associated Press,* 1er avril 2024.

face à une attaque de désinformation massive à soixante minutes (1). L'ancienne ministre responsable de cette administration, Audrey Tang, avait aussi mis en place des mesures de communication à destination de l'ensemble des Taïwanais. Son programme phare, « *Humour against rumour* » reposait sur une règle dite du « 2-2-2 » : dès la détection d'une infox, une publication était diffusée dans un délai de vingt minutes, rédigée en deux-cent mots maximum et accompagnée de deux visuels humoristiques (2). Ce format permettait d'offrir une réponse claire et rapide, tout en s'assurant qu'elle soit suffisamment captivante pour être diffusée massivement par les citoyens eux-mêmes. La carte de l'humour a ainsi permis d'endiguer « l'infodémie » en période de crise par la création de messages plus viraux que le contenu fallacieux initial.

Proposition : Promouvoir la création de *think-tanks* spécialisés au niveau national et, plus largement, soutenir la recherche scientifique afin d'objectiver les conséquences des ingérences étrangères et de diffuser une culture de la résilience dans la société.

Ces exemples, suédois et taïwanais, ont pour point commun un granulage des actions de sensibilisation, en prévoyant des mesures ciblées pour certaines catégories de personnes. En Europe, deux publics méritent une attention particulière en raison de leurs caractéristiques et de la fréquence des attaques les visant : les jeunes et les diasporas.

S'agissant de la jeunesse, l'exposition à des contenus malveillants, dès l'enfance ou l'adolescence, constitue une vulnérabilité manifeste pour les individus et la société. Afin de se laisser la possibilité de pouvoir former un esprit critique mature, il est important de donner du temps aux écoles, aux familles et aux institutions en préservant les enfants de données potentiellement parasitaires. Pour ce faire, les rapporteurs s'associent à des initiatives déjà portées par des collègues parlementaires afin de limiter l'accès aux réseaux sociaux avant un certain âge, fixé à 15 ans. En l'espèce, le Parlement a déjà adopté, en 2023, une proposition de loi, portée par M. Laurent Marcangeli, visant à instaurer une majorité numérique et à lutter contre la haine en ligne, qui comprenait cette disposition (3). Toutefois, les deux décrets nécessaires à l'application du texte n'ont toujours pas été pris en raison de réserves émises par la Commission européenne sur la faisabilité technique des dispositifs de vérification de l'âge, dans le cadre de l'application du DSA, et sur le manque d'homogénéité des pratiques nationales (4). Afin de relancer le sujet au niveau communautaire, un programme pilote avec un prototype d'application

(1) Duncan Barron, « Taiwan's model for digital defense of democracy goes global », *The Diplomat, 28 juin 2025.*

(2) Arwa Mahdawi, « Humour over rumour? The world can learn a lot from Taiwan's approach to fake news », *The Guardian, 17 février 2021.*

(3) Loi n° 2023-566 du 7 juillet 2023 visant à instaurer une majorité numérique et à lutter contre la haine en ligne, article 4.

(4) L'article 8 du RGPD prévoit que « le traitement des données à caractère personnel relatives à un enfant est licite lorsque l'enfant est âgé d'au moins seize ans. Lorsque l'enfant est âgé de moins de 16 ans, ce traitement n'est licite que si, et dans la mesure où, le consentement est donné ou autorisé par le titulaire de la responsabilité parentale à l'égard de l'enfant. Les États membres peuvent prévoir par la loi un âge inférieur pour ces finalités pour autant que cet âge inférieur ne soit pas en-dessous de 13 ans. » *En Allemagne ou aux Pays-Bas, cette majorité numérique pour le consentement a ainsi été fixée à 16 ans. La France a choisi 15 ans. De nombreux États, comme la Belgique ou les pays nordiques ont opté pour13 ans.*

européenne pour vérifier l'âge a été annoncé en juillet 2025 par la Commission européenne. En complément, des lignes directrices pour la protection des mineurs en ligne viennent d'être publiées [1]. La France devrait ainsi être capable de pouvoir appliquer prochainement les dispositions législatives déjà promulguées.

Proposition : Mettre rapidement et effectivement en place, au niveau européen ou, à défaut, à l'échelle nationale, une majorité numérique conditionnant l'accès aux réseaux sociaux.

Concernant les diasporas, la méthode suédoise est opportune. Les opérateurs de l'État, notamment l'agence de défense psychologique, travaillent de concert avec des associations locales [2] permettant un relais efficace et de confiance auprès des publics concernés.

Proposition : Nouer des partenariats avec des associations au niveau local pour atteindre davantage toutes les diasporas lors des opérations de sensibilisation aux dangers des ingérences étrangères.

Enfin, ces évolutions justifient une plus grande imbrication des acteurs de la contre-ingérence au sens large. Une intégration au comité opérationnel de lutte contre les manipulations de l'information (COLMI – *cf. supra*) de représentants d'autres ministères, selon un format non limité aux seuls domaines régaliens, pourrait dès lors être nécessaire. En raison de leur implication croissante, le ministère de la justice et celui de l'éducation nationale auraient notamment vocation à être associés plus directement aux délibérations.

Proposition : Envisager l'élargissement du comité opérationnel de lutte contre les manipulations de l'information (COLMI) à d'autres ministères particulièrement mobilisés sur les questions d'ingérence et de désinformation.

3. Le soutien à une production d'information fiable et de qualité est indispensable pour redonner du sens au doute et réapprendre à faire confiance

L'efficacité des opérations d'ingérence repose avant tout sur la crise de légitimité des autorités informatives traditionnelles, c'est-à-dire les médias installés et les institutions politico-administratives. Ce constat rappelé au cours des auditions par le chercheur Charles Thibout est étayé par les résultats des multiples baromètres de la confiance. Par exemple, celui réalisé chaque année pour le journal *La Croix* indique que 62 % des sondés se méfient de ce que disent les médias sur les grands sujets d'actualité (+ 8 points depuis 2022) [3]. En parallèle, dans le cadre du

(1) Commission européenne, communication de la Commission : « Lignes directrices concernant des mesures visant à garantir un niveau élevé de protection de la vie privée, de sûreté et de sécurité des mineurs en ligne, conformément à l'article 28, paragraphe 4, du règlement (UE) 2022/2065 *», 10 octobre 2025, 38 pages.*

(2) À titre d'exemple, l'agence de défense psychologique a conclu un partenariat intitulé « Vi Behövs *» avec l'ONG Fryshuset et l'agence suédoise de la protection civile (MSB). Ce programme vise à former les jeunes, issus de toutes les communautés, à la préparation aux crises.*

(3) Laure Salvaing, Guillaume Caline, Alexandre Vassas, Baromètre La Croix – Verian – La Poste : « La confiance des Français dans les médias », *janvier 2025, 43 pages.*

baromètre du centre de recherches politiques de Sciences Po (CEVIPOF) 26 % des sondés affirment ne pas avoir confiance dans la politique [1]. Dans le détail, les institutions qui suscitent le plus la méfiance sont, en premier, le gouvernement (77 %) et l'Assemblée nationale (76 %), devançant de peu l'institution présidentielle (74 %) [2].

Autrement dit, les grands régulateurs sociaux de l'information, les institutions qui traditionnellement font autorité en la matière, doivent composer avec une situation marquée par une défiance importante. Cet état de fait rend la population davantage perméable à d'autres sources d'information ou de désinformation, et, en particulier, aux messages qui renforcent le sentiment de méfiance à l'égard des institutions publiques. Ce sujet majeur n'appelle pas une réponse univoque et doit faire l'objet de réflexions multiples portant notamment sur les modalités de garantie et de contrôle de la probité des acteurs publics, sur les conditions d'amélioration du pluralisme informationnel dans le secteur audiovisuel et la presse, sur l'indépendance des journalistes et des rédactions, sur la protection juridique des journalistes et sur la qualité de la formation des professionnels. Les hommes et femmes politiques, de par leurs fonctions, ont également une responsabilité conséquente et doivent être singulièrement attentifs aux questions d'ingérence et de désinformation. Une forme de pacte de « non-agression » entre élus, *a minima,* sur les sujets liés à l'actualité, serait louable et redonnerait une hauteur de vue aux débats tout en évitant des relais faciles pour des puissances étrangères.

Redonner confiance aux citoyens doit donc être la boussole guidant les décideurs et les fournisseurs d'informations. Cette tâche de fond n'empêche pas d'avancer par étape et de penser la place des médias dans la question des ingérences étrangères et du recours à l'IA. À cet effet, les rapporteurs ont notamment entendu des représentants de France Médias Monde (FMM) et de l'agence de presse américaine Bloomberg. En France comme aux États-Unis, les rédactions se sont organisées pour renforcer la vérification de l'information. L'utilisation de l'IA augmente drastiquement les volumes de données à traiter et donc le risque d'erreurs en sortie de production. Afin d'y faire face, les médias sont obligés de mobiliser de nouvelles ressources, technologiques et humaines. Compte tenu des flux engendrés par l'IA, Bloomberg utilise désormais un outil automatisé, le « *summary by Bloomberg AI* », pour le traitement préliminaire des communiqués de presse. L'implication de la machine impose cependant toujours une relecture par un humain.

Certains groupes peuvent manquer de moyens pour mettre en place ce type d'instrument ou pour formaliser des processus de vérification de l'information. De même, de nombreuses administrations ne bénéficient pas encore de l'expertise en interne pour analyser efficacement le contexte informationnel. Dans l'attente d'une montée en puissance générale, une meilleure intégration des sphères publique

(1) CEVIPOF, « En qu[o]i les Français ont-ils confiance aujourd'hui ? » *p. 27.*

(2) Ibid, p. 29.

et privée pourrait en partie remédier à ce problème et fournir immédiatement des moyens opérationnels. Les compétences de sociétés spécialisées seraient une ressource précieuse pour l'écosystème, en appui ou en partenariat direct. À titre d'exemple, l'entreprise NewsGuard recense déjà les principales fausses informations circulant en ligne *via* son catalogue « *False claim fingerprints* ». Les données recueillies sont exploitées par d'autres entreprises privées dans des secteurs divers ainsi que par certains moteurs de recherche. Les équipes en charge des vérifications comprennent des journalistes et des éditeurs expérimentés qui produisent en parallèle des évaluations de fiabilité des sites d'information au regard de neuf critères journalistiques et apolitiques (1). La généralisation de ce type de notation, sous la forme d'une véritable « étiquette nutritionnelle » pourrait apporter une plus grande clarté aux consommateurs et responsabiliser davantage les producteurs.

À terme, le savoir-faire en matière de lutte contre la désinformation doit être intégré dans un socle commun d'apprentissage au vu de l'augmentation drastique des risques liés à l'utilisation de l'IA. Dans d'autres pays, comme en Suède, des modules spécifiques à la désinformation sont ainsi notamment prévus dans les programmes des écoles de journalisme.

Proposition : Encourager la conclusion de partenariats public-privé pour développer massivement les techniques de vérification de l'information.

Proposition : Envisager une labellisation informationnelle selon un modèle de nutri-score avec des critères objectifs et en sources ouvertes afin que chacun puisse vérifier l'impartialité du système.

Proposition : Systématiser obligatoirement dans les maquettes des cursus en journalisme des cours dédiés à la vérification de l'information et à la lutte contre la désinformation.

Les médias ont également un rôle majeur dans la diffusion de contre-narratifs, notamment à l'étranger. Pour ce faire, il ne faut pas seulement riposter en diffusant un démenti, il est nécessaire de façonner l'environnement informationnel dans lequel évoluent les auditeurs et lecteurs. Cette fonction est principalement assurée par l'audiovisuel public extérieur (2). À cet effet, France Médias Monde dispose d'équipes référentes à l'international dédiées à la vérification de l'information, à la veille et à l'investigation numérique ainsi qu'au suivi de la désinformation. En audition, la présidente-directrice générale de France Médias Monde a souligné les résultats de deux services référents en la matière :

(1) Le site d'information évalué l'est au regard des neuf critères suivants : 1. ne publie pas de contenus faux ou manifestement trompeurs de manière répétée ; 2. recueille et présente l'information de façon responsable ; 3. dispose de procédures efficaces pour corriger les erreurs ; 4. gère de manière responsable la différence entre informations et opinions, 5. évite les titres trompeurs ; 6. le site indique à qui il appartient et comment il est financé ; 7. identifie clairement la publicité ; 8. indique qui est responsable des contenus et tous conflits d'intérêt possibles ; 9. le site fournit des informations sur les créateurs de contenu.

(2) L'audiovisuel public extérieur comprend les chaînes du groupe France Médias Monde (FMM) – France 24, Radio France internationale (RFI) et Monte-Carlo Doualiya (MCD) – et la chaîne multilatérale TV5 Monde.

- à France 24, la rédaction de l'émission « *Les Observateurs* » s'appuie sur un réseau de 5 000 « observateurs » partout dans le monde et produit plusieurs programmes à l'image de « *Info ou Intox* », en français, en anglais, en arabe et en espagnol ;

- à Radio France Internationale (RFI), la cellule « InfoVérif » créée en 2023, produit une chronique et une émission hebdomadaires « *Les dessous de l'infox* ». Cette cellule travaille en complémentarité avec France 24 et les rédactions locales.

Les équipes dédiées à la lutte contre les infox, en particulier celles liées à l'intelligence artificielle, ne se limitent pas à ces deux structures. De nombreux journalistes y contribuent également au sein des autres rédactions, dans les vingt-et-une langues de diffusion du groupe. En Afrique notamment, ils se mobilisent pour lutter contre les manipulations de l'information par la production de programmes locaux tels que « *Ukweili au Uongo* » (vrai ou faux en kiswahili), qui s'attache chaque semaine à décrypter et analyser les publications mensongères qui essaiment sur le numérique en swahili. Au total, une quinzaine de programmes est proposée par les médias de France Médias Monde pour lutter contre la désinformation. La diffusion dans les langues locales est primordiale car certaines sont extrêmement stratégiques pour démystifier des contenus – l'arabe, le russe, l'ukrainien ou encore le chinois – et d'autres, en particulier panafricaines, sont très exposées aux manipulations informationnelles à l'image du mandenkan, du fulfulde, du kiswahili ou encore de l'haoussa.

Alors que l'environnement informationnel se dégrade, en partie à cause du recours à l'IA, le maintien d'une production de qualité implique un soutien accru des pouvoirs publics. Les fonds alloués à France Médias Monde, et de manière générale à l'audiovisuel public extérieur, sont trop limités au regard de l'évolution géopolitique et technologique. Ces dernières années, les grands pays démocratiques disposant d'un audiovisuel public extérieur ont tous conforté les moyens de leurs services, à l'instar des britanniques de la *BBC World Service*, qui disposaient en 2025 d'un budget de 479 millions d'euros, ou des allemands de *Deutsche Welle* avec une enveloppe de 420 millions d'euros pour l'année. Au-delà, la décision de l'administration américaine de cesser le financement des médias internationaux américains – 813 millions d'euros en 2023 répartis entre *Radio Free Europe / Radio Liberty* (RFE/RL), *Radio Free Asia*, *Radio Marti* et le *Middle East Broadcasting* – en démantelant l'*United States Agency for Global Media* (USAGM) pourrait créer un « appel d'air » laissant une place importante aux médias russes et chinois. Dès 2023, Moscou aurait ainsi déjà investi plus d'1 milliard de dollars dans sa stratégie d'influence et d'ingérence (1) en s'appuyant en parallèle sur deux médias, *Russia Today* (RT) et *Sputnik*. Comme l'a récemment rappelé le rapporteur dans son avis sur les crédits du compte de concours financier « *Avances à l'audiovisuel*

(1) Sébastien Seibt, « Les « *Kremlin Leaks* » révèlent une propagande inédite du candidat Poutine », *France 24, 26 février 2024.*

public » du projet de loi de finances pour 2026 [1], les 387 millions d'euros prévus pour le financement de l'audiovisuel public extérieur français au titre de l'exercice budgétaire à venir ne sont pas à la hauteur des enjeux décrits. Dans le cadre d'une guerre hybride, la hausse du budget militaire annoncée par le gouvernement ne peut se limiter aux armées et doit aussi prendre en compte la composante informationnelle civile.

Proposition : Renforcer les moyens financiers de l'audiovisuel public extérieur pour raffermir la réponse informationnelle française à l'étranger.

B. LES MOYENS DE CONTRE-INGÉRENCE DOIVENT ÊTRE RENFORCÉS AFIN DE DISPOSER D'OUTILS DÉFENSIFS ET OFFENSIFS PERFORMANTS

1. Les structures de détection et d'analyse des menaces mobilisant l'IA nécessitent un soutien accru pour entériner la montée en puissance de notre défense

La France dispose aujourd'hui d'une structure défensive performante et peut compter sur l'implication d'agents compétents au professionnalisme remarquable. La qualité de nos services de renseignement n'est plus à démontrer et les opérateurs civils tels que Viginum et l'ANSSI sont régulièrement cités en modèle par nos partenaires étrangers. Toutefois, la menace évolue rapidement et ne nous permet pas de nous reposer sur nos acquis. L'objectif des rapporteurs n'est pas d'appeler à des versements superfétatoires dans un contexte budgétaire contraint. Il est cependant indispensable d'allouer des montants plus conséquents à la défense numérique de notre pays, dans un secteur encore méconnu par nos concitoyens et pourtant au cœur des stratégies adverses. Outre les bénéfices en termes de sécurité, cet investissement serait vertueux sur le plan financier car le coût de l'inaction en cas de réussite d'une attaque informationnelle ou cyber peut s'avérer bien plus important. Cet effort pécuniaire devrait faire l'objet d'une réflexion dans le cadre des débats budgétaires en actant soit des augmentations de crédits soit un redéploiement des moyens depuis d'autres missions du projet de loi de finances.

Les éléments de réponse formulés par les rapporteurs se limiteront, sur le plan législatif, à ce sujet. Une loi relative à la prévention des ingérences étrangères en France [2] a déjà été adoptée en 2024 afin de créer notamment un registre des activités d'influence étrangère, d'autoriser, à titre expérimental, les services de renseignement à utiliser la technique algorithmique pour détecter des connexions susceptibles de révéler des ingérences étrangères et d'étendre la procédure de gel des avoirs financiers aux affaires d'ingérences étrangères. Il apparaît désormais

(1) Alain David, avis n° 1990, au nom de la commission des affaires étrangères, sur le projet de loi de finances pour 2026, tome VIII *« Avances à l'audiovisuel public – Audiovisuel extérieur », Assemblée nationale, 22 octobre 2025, 71 pages.*

(2) Loi n° 2024-850 du 25 juillet 2024 visant à prévenir les ingérences étrangères en France.

opportun de ne pas « sur-légiférer » et de renforcer les outils existants tout en impliquant davantage la population.

Concernant les besoins de la DGSE et de la DGSI, la nature et les moyens actuellement mobilisés n'ont pas été révélés. Les rapporteurs appellent tout de même à flécher spécifiquement des ressources substantielles pour appréhender les risques informationnels et s'adapter aux nouveaux formats des cyberattaques.

S'agissant de l'ANSSI, un projet de loi relatif à la résilience des infrastructures critiques et au renforcement de la cybersécurité est actuellement en cours de discussion à l'Assemblée nationale [(1)]. En transposant dans le droit national la directive (UE) 2022/2555 du Parlement européen et du Conseil du 14 décembre 2022 concernant des mesures destinées à assurer un niveau élevé commun de cybersécurité dans l'ensemble de l'Union, modifiant le règlement (UE) no 910/2014 et la directive (UE) 2018/1972, et abrogeant la directive (UE) 2016/1148 (directive SRI 2), dite directive NIS 2, il prévoit, entre autres, d'augmenter drastiquement le nombre d'entités à superviser par l'agence. La préparation des futures échéances nationales avec l'accueil du prochain sommet du G7 à Évian en 2026 ainsi que la perspective des Jeux olympiques d'hiver 2030 alourdissent également la charge de travail de l'ANSSI. En raison de la dispersion des attaques cyber à une pluralité de niveaux et de structures, le besoin de soutien en province est aussi plus important que par le passé et accapare davantage les équipes. En outre, l'impératif de « cyberhygiène » incombant à chacun ne peut être respecté que si les individus et les structures connaissent les bonnes méthodes, ce qui suppose une multiplication des actions de formation et de prévention. Pour ces raisons, a été demandé dès l'an dernier une augmentation des effectifs à hauteur de 70 équivalents temps plein (ETP) supplémentaires. Non satisfaite en 2025, cette requête a été à nouveau formulée au cours des auditions. Compte tenu de l'extension des missions de l'agence, les rapporteurs soutiennent un renforcement des moyens attribués à l'ANSSI. Ils appellent également à une poursuite rapide de la navette parlementaire pour le projet de loi mentionné précédemment et relatif à la résilience des infrastructures critiques et au renforcement de la cybersécurité.

Proposition : Garantir à l'ANSSI les moyens nécessaires à la bonne exécution de ses missions dont le volume et la nature évoluent, tout en transposant au plus vite la directive européenne dite NIS 2 dans notre droit national.

Concernant Viginum, le service est placé au cœur de notre système de surveillance des ingérences étrangères numériques. Son expertise et sa flexibilité lui permettraient d'assurer les opérations de sensibilisation mentionnées précédemment. Alors qu'en Suède l'agence de défense psychologique estime que 75 % de la population mesure l'étendue de ses missions, Viginum n'est pas encore un acteur connu par nos concitoyens. L'association aux campagnes publicitaires proposées et l'établissement d'une connexion avec d'autres institutions civiles

(1) Projet de loi n° 1112 relatif à la résilience des infrastructures critiques et au renforcement de la cybersécurité, XVII^e législature, 13 mars 2025.

devraient pouvoir y remédier rapidement. L'identification d'une « structure repère » par les Français faciliterait la prise de conscience générale de la menace. À ce titre, Viginum a d'ores et déjà amorcé un développement de ses activités annexes à destination du public. Un partenariat a été lancé avec la direction générale de l'enseignement scolaire (Dgesco) afin d'insérer dans les programmes scolaires des éléments de littératie pour les élèves ainsi que pour les enseignants. En coopération avec le centre de liaison de l'enseignement et des médias d'information (Clemi), opérateur du ministère de l'éducation nationale dédié à l'éducation aux médias et à l'information, Viginum produit également des *podcasts* visant la tranche d'âge des collégiens et lycéens.

Outre le volet technico-opérationnel de ses rapports d'analyse, Viginum publie en parallèle de nombreux guides pour les entreprises (1), les personnels politiques (2) et les médias (3). Ces derniers sollicitent par ailleurs régulièrement les équipes de Viginum pour vérifier des points particuliers dans le cadre de reportages. En complément, des partenariats avec la presse quotidienne régionale (PQR) sont conclus afin de toucher à la fois des territoires qui peuvent se trouver isolés de certaines informations nationales et des tranches d'âges plus difficile d'accès. Pour 2026, Viginum a également pour objectif de formaliser ses activités de formation par la création d'une académie de lutte contre les manipulations de l'information ainsi qu'en accompagnant la montée en puissance d'un centre d'excellence sur l'intelligence artificielle et la désinformation. Ces projets ainsi que l'extension du champ d'activité de Viginum nécessitent toutefois une revalorisation des ressources associées. Le service n'est aujourd'hui pas dimensionné pour réaliser pleinement ces tâches. Comprenant aujourd'hui une soixantaine d'agents, une augmentation des effectifs pour porter la structure à 90 ETP apparaît souhaitable. Cette taille est celle retenue par l'agence de défense psychologique suédoise et permet de rayonner dans la société en conservant une agilité indispensable aux missions de surveillance et de réaction.

Proposition : Renforcer les moyens humains, matériels et financiers de Viginum afin d'accompagner sa montée en puissance et l'extension de son champ d'activité.

2. La posture défensive doit se doubler d'une riposte offensive et proactive

La France ne s'abaissera pas à recourir aux méthodes crapuleuses décrites précédemment et utilisées par certains acteurs étrangers. Nos valeurs démocratiques guident notre action extérieure et sont la force de notre modèle sociétal. Cette profession de foi n'équivaut pas à une position de faiblesse, passive et

---

(1) Viginum, Guide de sensibilisation à la menace informationnelle – Écosystème des acteurs économiques associés aux Jeux olympiques et paralympiques de Paris 2024, *SGDSN, 20 juin 2024, 8 pages.*

(2) Viginum, Protéger le débat public numérique en contexte électoral : guide de sensibilisation à l'attention des équipes de campagne, SGDSN, *mars 2024, 8 pages.*

(3) Viginum, Guide de sensibilisation à la menace informationnelle pendant les Jeux olympiques et paralympiques à destination des médias et journalistes fact-checkeurs, *SGDSN, 25 juillet 2024, 8 pages.*

apathique. Notre pays, par ses médias et ses institutions, est capable de muscler sa réponse et de renforcer sa culture de l'influence.

La communication offensive est un exercice délicat jonglant entre réactivité et pragmatisme. La difficulté réside dans la nature même du métier de diplomate, fondé sur la parole et la centralité de certains éléments du discours officiel. Il nous faut désormais apprendre à articuler finement les signaux d'alerte avec les nuances propres à chaque langage diplomatique ou digital. Afin de mieux les appréhender, un important travail de consolidation des structures de communication stratégique est en cours au sein du MEAE. Ces cadres regroupent les grands narratifs par acteurs, permettant de définir clairement notre positionnement vis-à-vis de chacun d'eux. Par exemple, à l'égard de la Chine, il s'agit de maintenir un dialogue constant tout en adoptant une posture ferme, notamment avec l'Union européenne, sur les questions commerciales.

La France souffre toutefois d'un certain retard historique sur les questions d'influence digitale, en partie car le marketing a longtemps été considéré dans notre pays comme un objet vulgaire, manquant de finesse et de rondeur. Or, dans le contexte actuel, le simple rayonnement culturel ou diplomatique ne suffit plus. L'approche anglo-saxonne du sujet semble aujourd'hui plus adaptée et commence à être davantage assumée. Les britanniques ont notamment lancé, en 2012, avant les Jeux olympiques de Londres, une campagne massive « *Great Britain is Great* » pour promouvoir à l'international les forces de leur pays. Cette initiative a depuis fait l'objet de rééditions, notamment en 2021 au moment de l'accueil du sommet du G7 en Cornouailles et de la conférence des Nations unies sur les changements climatiques (COP26) à Glasgow. La France commence progressivement à faire de même avec notamment des campagnes telles que « *Make it iconic* » et la valorisation du « *Choose France* ». En réponse directe aux manœuvres de désinformation, le Quai d'Orsay a lancé, en septembre 2025, un compte de riposte sur la plateforme *X* qui inclut une part d'automatisation. Intitulé « *French Response* », ce profil reprend les codes des réseaux sociaux et démystifie rapidement les attaques informationnelles menées par des acteurs étrangers. Plusieurs publications de trolls russes ont par exemple été exposées, ainsi que certains messages d'Elon Musk.

Cette approche plus offensive gagnerait à être relayée largement par nos postes diplomatiques à l'étranger. L'ambassade de France aux États-Unis repartage régulièrement les publications de *French Response* et crée également directement sur *X* des contenus pour contrer les narratifs erronés de la sphère MAGA. Des ressources spécifiques doivent être déployées pour développer cette communication stratégique (*stratcom*) et ridiculiser les attaques informationnelles adverses. L'objectif est d'outiller les postes diplomatiques en moyens humains et technologiques afin d'exploiter pleinement les réseaux disponibles. Ces actions peuvent par ailleurs être réalisées en coopération avec d'autres ambassades de pays amis partageant nos valeurs. Les sommets mondiaux pour l'action sur l'intelligence artificielle sont à cet égard des forums intéressants pour échanger entre partenaires sur les bonnes pratiques et, au-delà des effets d'annonces, intensifier nos

coopérations. Après l'édition 2024 à Séoul, un réseau international des instituts de sécurité de l'IA a ainsi été créé pour former une coalition de pays autour d'une vision commune. Ce réseau regroupe l'Australie, le Canada, la France, le Japon, le Kenya, la Corée du Sud, Singapour, le Royaume-Uni, les États-Unis et la Commission européenne. Dans ce cadre, la France s'est dotée d'un institut national pour l'évaluation et la sécurité de l'intelligence artificielle (INESIA). Cet institut n'est pas un énième organe satellite mais fédère, sans nouvelle structure juridique, les acteurs nationaux de l'évaluation et de la sécurité, et tout particulièrement l'ANSSI, l'Institut national de recherche en sciences et technologies du numérique (INRIA), le laboratoire national de métrologie et d'essais (LNE), et le pôle d'expertise de la régulation numérique (PEReN). En format restreint, la création du centre européen pour la résilience démocratique devrait également permettre de renforcer notre capacité à faire front commun.

Proposition : Accroître les efforts en matière de communication stratégique et adopter une posture plus offensive pour contrer les narratifs adverses.

Proposition : Soutenir les initiatives conjointes en matière de lutte contre la désinformation et d'utilisation de l'IA avec les pays amis afin de consolider un réseau de contre-ingérence, au niveau international et à l'échelle européenne dans le cadre du bouclier démocratique.

EXAMEN EN COMMISSION

Au cours de sa réunion du mercredi 3 décembre 2025 à 9 heures 30, la commission a examiné le rapport d'information sur l'irruption de l'intelligence artificielle dans les ingérences étrangères.

Mme Constance Le Grip, présidente. Le président Fuchs conduit actuellement une délégation de notre commission à la 80e Assemblée générale des Nations unies. Il m'a demandé de le suppléer.

Nous allons entendre ce matin la présentation du rapport d'information de Mme Laetitia Saint-Paul et M. Alain David sur l'irruption de l'intelligence artificielle (IA) dans les ingérences étrangères. Notre commission apporte ainsi sa contribution à un débat crucial, alors que le président de la République a lancé, fin octobre, une réflexion sur la démocratie à l'épreuve des réseaux et des algorithmes.

L'irruption de l'intelligence artificielle dans les processus démocratiques constitue une transformation révolutionnaire, porteuse à la fois de promesses et de menaces, cette technologie, surtout l'IA générative, pouvant influencer la manière dont les citoyens s'informent, débattent et votent.

Au premier chef, l'IA peut générer à bas coûts, très rapidement, de fausses informations, propagées notamment *via* les *deepfakes* et les *chatbots* automatisés. En Ukraine, aux États-Unis, des vidéos truquées ont semé la confusion. Ainsi, en 2024, plusieurs vidéos synthétiques imitant la voix et le visage de responsables politiques ont été diffusées à la veille d'élections, incitant les électeurs à s'abstenir ou à voter pour un candidat précis. Ces manipulations difficiles à détecter par le grand public fragilisent le principe fondamental de transparence sur lequel repose toute démocratie.

L'IA facilite également les ingérences étrangères par la désinformation. L'analyse de données massives permet d'identifier des groupes d'électeurs vulnérables, leurs peurs et leurs aspirations, puis de diffuser des messages calibrés pour influencer leur comportement électoral. L'affaire *Cambridge Analytica*, utilisant en 2018 des technologies moins avancées, illustrait déjà ce microciblage politique avec pour slogan : *Data drive all we do* – « *Les données sont au cœur de tout ce que nous faisons* ». Désormais, l'IA peut se déployer à grande échelle, brouillant les frontières entre information, propagande et manipulation.

Dans un contexte de guerre hybride, les démocraties, dont la France, sont particulièrement ciblées. Madame et monsieur les rapporteurs, vous allez présenter vos travaux et vos préconisations, nous dire comment la réglementation européenne sur l'IA et peut-être des mesures au plan national peuvent renforcer notre résilience numérique ainsi que notre capacité à résister à l'érosion de nos principes démocratiques.

Mme Laetitia Saint-Paul, rapporteure. C'est lorsque des étudiants m'ont parlé d'une fausse nouvelle en disant « mais c'est vrai, je l'ai vu sur TikTok » que je me suis rendue compte que l'intelligence artificielle pouvait être utilisée pour effectuer des ingérences étrangères. Lorsque nous avons décidé en janvier 2025 de créer cette mission d'information, le sujet semblait de nature technique. L'actualité récente montre qu'il est très politique. Je remercie Alain David, qui fut rapporteur sur les géants du numérique, d'avoir accepté d'en être co-rapporteur.

Cristallisant des enjeux financiers, scientifiques, sécuritaires et culturels, la maîtrise de l'intelligence artificielle est devenue un impératif géopolitique pour les États. Ainsi, dès 2017, Vladimir Poutine déclarait que l'IA représentait l'avenir non seulement pour la Russie mais pour l'humanité et que la nation qui serait leader dans ce domaine dominerait le monde. De tels propos ne préfiguraient pas une collaboration pacifique mais une approche conflictuelle. La confrontation est en cours et, selon l'Organisation de coopération et de développement économiques (OCDE), la France est la première victime d'ingérences étrangères de l'Union européenne, la seconde après l'Ukraine sur le continent européen. Nous l'avons constaté, nous sommes une proie. Forts de notre puissance nucléaire et conventionnelle, nous ne nous pensons pas comme tel, contrairement à des pays qui ont identifié clairement un responsable de ces ingérences : la Russie pour la Suède ou les États baltes, la Chine continentale pour Taïwan.

En France, les ingérences sont protéiformes, étatiques ou non-étatiques. En premier lieu, selon *NewsGuard*, la Russie a, en 2024, publié 3,6 millions d'articles de propagande et on estime qu'elle consacre 1 milliard d'euros chaque année à la désinformation et à la guerre cognitive. Dans le cadre de l'opération Storm-1516, l'une des plus importantes campagnes d'ingérence russe, Moscou a mené 77 actions de déstabilisation qui ont généré 55 millions de vues ; ont aussi eu lieu 38 000 publications, des usurpations d'identité de présentateurs, le montage de faux sites d'information dans 48 langues, des achats de pages de journaux en Afrique pour blanchir l'information et empoisonner les modèles d'IA. En France, derrière les étoiles de David bleues, il y avait quatre Moldaves et surtout 1 095 *bots* qui ont relayé l'information pour démoraliser notre société. Nous sommes aussi une proie de la sphère *Make America Great Again* (MAGA) – car les États-Unis détestent le règlement européen sur les services numériques (DSA) et interprètent de façon très particulière le premier amendement de leur Constitution sur la liberté d'expression – , une proie pour l'Iran dans le débat public sur le nucléaire ou le voile et pour la Chine, bien sûr, qui diffuse des narratifs favorables au Parti communiste chinois selon un mode opératoire dit « spamouflage », fondé sur des réseaux de comptes aux caractéristiques inauthentiques opérant sur une multitude de plateformes. Citons encore l'Azerbaïdjan en ce qui concerne nos outremers, ainsi que le proto-État islamiste – à ce propos, je mène conjointement, au nom de l'Union interparlementaire (UIP), des travaux sur l'IA dans la lutte contre le terrorisme : on estime qu'en 2000 il fallait seize mois pour être radicalisé, en 2010 quelques mois seulement et, désormais, quelques semaines ; tout a lieu en ligne, y compris l'auto-radicalisation : au Royaume-Uni, 25 % des personnes arrêtées pour radicalisme sont

mineures. Citons enfin dans cette énumération Israël et la Hongrie, au gré de l'actualité.

Espionnage, propagande, désinformation ont toujours existé. L'IA ne serait qu'un mode d'action permettant de démultiplier les effets de méthodes telles que l'*astroturfing*, sorte de raid numérique pour créer une tendance, la création de faux comptes et de médias alternatifs, le financement d'influenceurs. Mais j'irai plus loin : l'IA constitue une révolution industrielle des ingérences. En effet, elle fait partie de notre quotidien. La quantité de données utilisées a explosé, leur qualité rend impossible de discerner le vrai du faux et les prix de revient se sont effondrés. La combinaison de ces trois facteurs a toujours provoqué une révolution industrielle. Cependant, pour ceux qui n'ont pas les clés de la technologie, l'IA met en quelque sorte de l'encre dans un stylo qui écrit tout seul, parfois avec des données empoisonnées, et plus nous croyons apprendre, plus nous nous trompons. On en arrive au point qu'en Corée du Sud, sur le modèle du *machine learning* – l'apprentissage automatique –, il est question aujourd'hui de *machine unlearning*.

On peut parler désormais de cyberguerre car, grâce à l'IA, on peut être hackeur sans savoir coder et déstabiliser nos entreprises, nos collectivités, nos hôpitaux. Je reviendrai sur les préconisations de responsabilisation que nous faisons, en particulier, comme pour les cookies, sur le consentement à donner par chaque utilisateur à l'accès au contenu proposé par algorithme, et non plus selon l'itinéraire imposé dans le DSA, pratiquement impossible à mettre en œuvre, les plateformes l'ayant transformé en labyrinthe. En 2023, l'Assemblée nationale a voté à l'unanimité en faveur de la majorité numérique. Or cette dernière n'est pas appliquée, pour des raisons que je ne m'explique pas. Le DSA n'était qu'un cadre, pas une contrainte, auquel Thierry Breton lui-même pensait que l'on pouvait connecter de nouveaux éléments. Nous pouvons donc avoir des exigences envers la Commission européenne. Nous pouvons également présenter d'autres propositions, par exemple sur le mandat de Viginum, le service de l'État de vigilance et de protection contre les ingérences numériques étrangères qui procède à des investigations en sources ouvertes, que nos concitoyens connaissent mal. Ces propositions peuvent, je le crois, faire l'unanimité dans notre Assemblée.

M. Alain David, rapporteur. Nous avons commencé nos travaux peu avant le sommet mondial pour l'action sur l'intelligence artificielle de février dernier à Paris. Depuis, il ne se passe pas une semaine sans que l'on parle d'intelligence d'artificielle dans des termes plus ou moins élogieux. L'IA n'est pas un outil bon ou mauvais en soi mais cette révolution technologique amplifie de manière très inquiétante la portée et la nocivité des attaques portées par des puissances hostiles.

Notre rapport d'information n'a pas pour objectif d'alarmer vainement les pouvoirs publics et la population mais de souligner une rupture majeure avec les pratiques passées. Il faut l'appréhender à l'échelon national et surtout au niveau communautaire. Tant que l'Europe sera dépendante sur le plan technologique, elle restera vulnérable aux ingérences étrangères qui mobilisent de plus en plus l'IA.

Aussi, prendre ce sujet à bras le corps n'est pas seulement une question économique ou technologique : c'est une question de souveraineté, de sécurité nationale et de défense de notre système démocratique.

Or l'Europe ne perçoit le danger que de façon partielle, parfois naïve. Pendant que les États-Unis et la Chine s'affichent en puissances assumées, nous pensons encore notre modèle comme un « grand marché ouvert » en misant sur notre attractivité pour les capitaux étrangers et la concurrence libre et non faussée.

Le résultat, c'est que nous sommes devenus tributaires de technologies que nous ne contrôlons pas : 90 % des semi-conducteurs que nous utilisons sont produits hors d'Europe ; l'immense majorité du *Cloud* européen est américain ; les puces indispensables pour l'IA proviennent de deux entreprises non-européennes, Nvidia et Taïwan Semiconductor Manufacturing Company (TSMC) ; nos propres talents, dans l'innovation numérique, partent travailler ailleurs. Dans ce contexte, croire que nous pourrions lutter efficacement contre les ingérences étrangères d'acteurs de plus en plus virulents relève de la naïveté ou, pire, d'une forme d'*hubris*.

Notre dépendance technologique est en effet une faille stratégique majeure. Ce constat, je le dressais déjà au cours de mes travaux sur les géants du numérique. Nous avons, collectivement, raté le virage technologique de l'an 2000. Nous ne pouvons nous permettre de manquer celui de l'intelligence artificielle.

L'Union européenne dispose pour cela d'atouts non négligeables. Avec 450 millions d'habitants et un produit intérieur brut (PIB) de près de 18 000 milliards d'euros, elle représente un marché capable de rivaliser avec les géants chinois et américains. Surtout, « l'effet Bruxelles » n'a rien d'un mythe : nous avons une capacité unique à influencer les normes mondiales. Pour preuve, en adoptant le DSA et le règlement établissant des règles harmonisées concernant l'intelligence artificielle, l'AI Act, nous avons défini un ensemble ambitieux et protecteur qui encadre les grandes plateformes et interdit les pratiques d'IA contraires à nos valeurs : manipulation subliminale, notation sociale, catégorisation biométrique.

Mais la régulation ne peut pas être notre seule stratégie. Sans maîtrise technologique, nous finirons par devenir inaudibles, voire marginalisés. Notre tissu de start-up européennes est dynamique mais ces entreprises renoncent trop souvent à opérer en Europe en raison de coûts réglementaires qu'elles estiment excessifs. Pendant ce temps, les grands acteurs étrangers disposent des moyens nécessaires pour s'adapter à ces règles et renforcer leur domination.

À court terme, des mesures défensives ont certes été prises. Ce sont, au niveau européen, des plans d'action contre la désinformation ; le *Rapid Alert System* du service européen pour l'action extérieure ; la *task force East Stratcom* ; la base *EUvsDisinfo* qui recense et démonte les campagnes de désinformation.

En France, Viginum détecte les ingérences numériques étrangères affectant le débat public. Citons aussi le comité opérationnel de lutte contre les manipulations

de l'information, le Centre de coordination des crises cyber, l'Agence nationale de la sécurité des systèmes d'information (ANSSI) et divers services de renseignement. Leur travail est remarquable et traduit une réelle prise de conscience nationale. Toutefois, ils agissent sur les conséquences du problème, non sur ses causes.

Nous surveillons des flux hébergés sur des plateformes étrangères, tentons de détecter des opérations informationnelles menées *via* des outils que nous ne contrôlons pas, essayons d'endiguer un flot dont la source nous échappe. Lorsque nous utilisons l'intelligence artificielle pour nous défendre, nous sommes de même limités par notre manque de capacités souveraines en IA. Nous ne défendrons pas durablement nos démocraties contre les ingérences étrangères sans une reconquête de notre autonomie technologique.

Il ne s'agit pas de rêver d'une autarcie numérique mais de se fixer des priorités : réduire la dépendance aux infrastructures critiques étrangères – *Cloud*, semi-conducteurs, supercalculateurs ; investir massivement dans les capacités de calcul, les modèles d'IA de base, les centres de données, les supercalculateurs, ce qui commence à être fait avec les *AI factories* et les *gigafactories* mais reste insuffisant ; privilégier des solutions d'IA européennes afin de bâtir une architecture communautaire indépendante ; orienter davantage les financements européens et l'épargne vers l'IA ; surtout, assumer une logique de puissance européenne.

Pour le moment, nous commentons les technologies des autres, nous les encadrons, nous les critiquons mais nous ne mettons pas le même niveau d'énergie à les produire. Face à une guerre informationnelle permanente, des opérations malveillantes massives mêlant la réalité et l'artificiel, ainsi que des cyber-attaques, cette attitude n'est plus tenable. Si nous voulons que nos démocraties résistent aux ingérences étrangères à l'heure de l'IA, nous devons accepter cette évidence : la souveraineté technologique n'est pas un énième gadget futuriste, elle constitue, *in fine*, une condition de la souveraineté politique. Sinon, nous resterons vulnérables, quelles que soient la sophistication de nos lois et la qualité de nos rapports.

Réguler, oui. Mais, sans la puissance technologique en complément, nous ne faisons que mettre des mots sur une réalité que d'autres façonnent à notre place.

Mme Constance Le Grip, présidente. Je vous remercie. Vous êtes d'une certaine manière des lanceurs d'alerte et c'est à juste titre que vous insistez sur la souveraineté technologique européenne.

Nous en venons aux interventions des orateurs des groupes politiques.

M. Bertrand Bouyx (HOR). L'irruption de l'intelligence artificielle dans les opérations d'ingérence redéfinit la guerre hybride. L'opération Storm-1516 que vous avez mentionnée constitue, selon Viginum, une montée en puissance sans précédent dans la stratégie russe, laquelle vise moins à convaincre qu'à fissurer de l'intérieur les sociétés démocratiques. Votre rapport montre amplement combien cette capacité à fracturer l'espace public et à brouiller la frontière entre influence et ingérence rend nos démocraties vulnérables. L'Union européenne a commencé à y

répondre avec l'AI Act et le DSA, le renforcement du rôle de Viginum, le Bouclier démocratique européen. Vous l'avez dit, cette défense reste perfectible, notamment face à la Russie. Le Conseil de l'Europe a également engagé des travaux sur la lutte contre la manipulation de l'information. En visite en Moldavie, j'ai constaté que la Russie avait consacré 320 millions d'euros à y déstabiliser les dernières élections législatives. Dans ce rapport, vous insistez sur la nécessité de renforcer les moyens opérationnels des structures luttant contre ces ingérences, au premier rang desquelles Viginum. Comment garantir que ce renforcement s'inscrive dans une stratégie européenne coordonnée ?

Mme Laetitia Saint-Paul, rapporteure. Thierry Breton, l'architecte des cinq piliers de l'espace informationnel européen, m'a dit lui-même que la bonne échelle était forcément l'Union européenne car chaque loi votée à l'échelle nationale crée des frontières. Mais il a eu l'humilité de reconnaître qu'il ne pouvait pas penser à tout et que ce cadre européen était conçu justement pour que l'on puisse y insérer des initiatives nationales pour l'enrichir. Au-delà, Nathalie Loiseau préside une commission spéciale sur le Bouclier démocratique européen, qui rendra ses travaux début 2026. L'idée est, pour ce projet, de s'inspirer de Viginum mais, en France, Viginum, qui exploite les sources ouvertes, travaille aussi avec la direction générale de la sécurité extérieure (DGSE), la direction générale de la sécurité intérieure (DGSI) et d'autres services de renseignement. De ce fait, créer un « Viginum européen » n'est pas nécessairement souhaitable mais le Bouclier peut s'adosser au service européen pour l'action extérieure (SEAE).

Actuellement la Commission européenne ne fait pas respecter l'espace informationnel européen, notamment face aux campagnes agressives de Shein. C'est notre rôle que de peser sur la Commission européenne pour que ce cadre solide soit respecté. Au niveau national, nous ne devons pas nous laisser tyranniser quand nos initiatives législatives sont retoquées pour des raisons que je ne m'explique pas.

M. Arnaud Le Gall (LFI-NFP). Des raisons inavouées !

M. Jean-Paul Lecoq (GDR). Le lobbying, tout simplement.

M. Alain David, rapporteur. Les Vingt-sept ont un PIB de 18 000 milliards d'euros. Cet énorme potentiel nous permet de lancer des opérations de recherche et d'entraîner des entreprises dans ce processus pour lutter contre la guerre déclenchée contre nous depuis une dizaine d'années. Malheureusement, sur le plan européen, nous ne parvenons pas à engager suffisamment les grandes entreprises, alors qu'aux États-Unis leurs homologues financent cette aventure.

Mme Laetitia Saint-Paul, rapporteure. Personnellement, je m'inquiète des tentatives actuelles et des pressions subies à Bruxelles contre l'AI Act et le DSA alors que nous n'avons les clés ni de la production de l'IA ni de sa diffusion, et à peine les moyens d'encadrer son utilisation. On vient de repousser certaines dispositions de l'AI Act ; je ne le comprends pas. Au contraire, il faut absolument

accélérer car, si on ouvre les vannes, nous allons nous faire siphonner toutes nos données. Il faut agir collectivement en ce sens.

M. Alain David, rapporteur. Il faut vraiment provoquer une prise de conscience.

Mme Constance Le Grip, présidente. Votre rapport y contribuera. L'AI Act et le DSA, comme d'autres régulations européennes, sont l'objet d'un rapport de forces entre l'Union européenne et les États-Unis d'Amérique que nous devons mesurer sans naïveté. Pour avoir suivi l'élaboration de ce règlement au sein des institutions européennes, je peux dire que le débat entre impératif de régulation et impératif d'innovation était très vif.

M. Jean-Paul Lecoq (GDR). Nous avons tous mesuré l'utilité de l'intelligence artificielle et je peux en témoigner en ce qui concerne le port du Havre. Ses capacités sont démesurées, illimitées même. Le lecteur de science-fiction que je suis se disait même qu'elle pourrait être incontrôlable. On me répondait que non, qu'il faudrait toujours une intervention humaine, mais il en allait de même quand j'ai parlé de l'influence des éruptions solaires sur l'informatique ; et voilà que les Airbus sont cloués au sol. Il nous faut vraiment militer pour instaurer des espaces de régulation au plan international. Par exemple, lorsque la recherche a rendu possible le clonage des êtres humains, nous avons réagi car nous ne voulions pas de cette société.

Pour l'IA aussi, il va falloir fixer des lignes rouges. L'humanité est en danger ; il faut réguler. Nous examinons ce rapport, très bien. Mais est-ce que, en France, le gouvernement ou d'autres entités utilisent l'IA pour procéder à des ingérences dans d'autres pays ?

Mme Laetitia Saint-Paul, rapporteure. Nous avons posé la question au Quai d'Orsay. Il nous a répondu que non.

M. Jean-Paul Lecoq. Nous avons la chance que soient présents ce matin un ancien président de la République et une ancienne première ministre, je suppose qu'ils pourraient apporter leur éclairage avisé sur mon interrogation.

Mme Laetitia Saint-Paul, rapporteure. S'ils souhaitent intervenir en complément sur ce sujet, je n'y vois aucune objection. En tout état de cause, vous avez raison de dire que toute l'humanité est menacée.

Dans le cadre de cette mission d'information, j'ai rencontré aux États-Unis des membres de l'ancienne administration Biden qui ont été licenciés, d'autres de l'administration Trump et des représentants des géants du numérique. Pour eux, l'IA est un nouveau Far West et leur obsession est de ne pas être dépassés par la Chine. Ils interprètent le premier amendement de la Constitution américaine sur la liberté d'expression comme le rejet de toute interdiction. L'Union européenne réglemente lorsqu'un problème risque de se poser ; aux États-Unis, on laisse prospérer les problèmes et l'on voit ensuite si l'on réagit. J'ai objecté aux représentants de

l'administration Trump qu'il y a des dérives, par exemple en ce qui concerne les mineurs – l'enfance reste un sujet dont on peut discuter. J'ai aussi fait observer que l'on ne peut pas mettre constamment en avant la liberté individuelle quand les citoyens ne sont pas libres du contenu auquel ils accèdent, puisque les plateformes produisent un contenu clivant créant une dépendance et leur procurant de la publicité. Je leur ai soumis ma proposition de permettre à chacun de dire qu'il accepte les algorithmes, comme on le fait pour les cookies. Ce consentement à l'utilisation des algorithmes figure à l'article 38 du DSA mais n'est pas appliqué. Sur Instagram ou Facebook, vous pouvez effectivement modifier les réglages mais cela suppose des manipulations compliquées que, de plus, il faut recommencer à zéro à chaque connexion ! L'administration Trump était au moins réceptive à l'idée que toute consommation d'un produit suppose le consentement.

On jugera peut-être cela prétentieux mais j'espère que notre rapport aura un écho plus large hors de ces murs. Si vous le souhaitez aussi, travaillons ensemble pour que chacun ait le choix et qu'une machine à gagner de l'argent avec de la publicité ne nous impose plus ce à quoi penser à force de gavage par ses algorithmes.

M. Alain David, rapporteur. Outre les États-Unis, la Russie et la Chine sont absolument opposées à toute régulation.

Mme Élisabeth Borne (EPR). Pas chez eux, en tout cas !

M. Alain David, rapporteur. Chinois et Américains n'entendent pas partager la suprématie mondiale qu'ils ont acquise dans ce secteur. S'agissant de l'usage de l'énergie atomique, on a réussi à instaurer une régulation internationale. C'est bien plus difficile avec l'IA, compte tenu de la nature même du produit.

M. Jean-Paul Lecoq. Pourquoi ne pas inscrire dans les conclusions de ce rapport l'idée d'une conférence des parties (COP) numérique ? J'entends bien que les grands acteurs ne sont pas d'accord mais ce serait une manière de pousser l'idée. Après tout, sur l'environnement, les États-Unis n'étaient pas d'accord au départ. Plantons la graine et aidons-la à grandir.

Mme Constance Le Grip, présidente. Un sommet pour l'action internationale en matière d'IA s'est tenu à Paris en février dernier à l'initiative du président de la République. Il a rassemblé de nombreux acteurs qui ont essayé d'envisager collectivement des modalités de gouvernance de l'internet. Il y a donc des ébauches mais le chantier est gigantesque, en effet. C'est un début.

M. Guillaume Bigot (RN). Tout en se tournant vers l'Union européenne comme solution de tous nos maux numériques, votre rapport souligne à juste titre la vulnérabilité, l'impuissance même de l'Europe en la matière. Une de vos formules dit tout : « *Les États-Unis innovent, la Chine copie, l'Europe réglemente* ». J'ajouterai « *et la Russie pirate* ». La souveraineté européenne dans le numérique relève du fantasme, écrivez-vous. L'Europe est telle une caverne de Platon où les images projetées sur les murs sont américaines, chinoises, parfois russes. L'Union est entre leurs mains, parfois aussi celles de l'Ukraine ou du Qatar ;

la corruption sévit dans les couloirs de Bruxelles où toutes les puissances étrangères sont chez elles, et non les peuples européens puisqu'il n'y a pas de démocratie directe. Vous-même concluez qu'il faut réglementer mais que l'échelle européenne n'est pas la bonne. Dans ce cas, pourquoi produit-on des normes inappliquées et inapplicables ? Pour faire diversion, faire croire qu'on fait quelque chose quand on ne fait rien.

Par télépathie – ou par Bluetooth ? – avec le président de la République, l'une de vos recommandations les plus spectaculaires est la labellisation informationnelle sous forme de nutri-score. Mais le nutri-score alimentaire n'a pas supprimé la malbouffe. Pensez-vous que votre nutri-score informationnel empêcherait l'abrutissement et les ingérences étrangères ? Le seul effet serait de créer un marché noir de l'information libre. Proposer d'emboîter le pas à la Russie et à la Chine en censurant et en infantilisant les citoyens, c'est paradoxal.

Mme Laetitia Saint-Paul, rapporteure. Je m'attendais à cette question.

M. Guillaume Bigot (RN). Je me doute, elle est d'actualité.

Mme Laetitia Saint-Paul, rapporteure. L'espace de l'information se doit d'être européen. C'est le Parlement européen, où tous les partis politiques sont représentés, qui a adopté à 90 % – et donc presque par consensus de l'extrême gauche à l'extrême droite – les cinq piliers de cet espace européen. Que la Commission européenne ne fasse pas respecter ce cadre, je le reconnais et le martèle. Mais il est nécessaire et pensé pour évoluer. Dans mes propositions, j'insiste pour que la majorité numérique que notre Assemblée a adoptée il y a deux ans, à l'unanimité, soit appliquée. Il est inacceptable qu'elle ne le soit pas alors même que la situation a plutôt empiré. Je suis toute disposée à travailler avec chacun d'entre vous pour y parvenir.

Quant à la labellisation informationnelle, il se trouve que ce rapport qui, initialement, paraissait technique, est maintenant atteint par le mauvais *buzz* qui accable le président de la République. Mais c'était une préconisation des états généraux de l'information. On peut s'interroger en constatant que depuis 1986 il n'y a pas eu de grande loi sur l'espace de l'information, alors qu'en quarante ans il a énormément évolué. Cela étant, à aucun moment le président Macron n'a dit que l'État devait labelliser les moyens d'information. On entend beaucoup citer George Orwell ces derniers temps – à propos de la vidéo-surveillance pendant les Jeux olympiques par exemple, et aussitôt qu'il est question de sécurité numérique dans les démocraties. À Taïwan, dès que j'avançais une proposition visant à lutter contre les ingérences étrangères, on me renvoyait à *Big Brother*. Je commence à me méfier d'Orwell s'il y est fait référence pour empêcher de penser la sécurité numérique, mais en disant cela je risque de me voir décerner le prix de la surveillance de masse généralisée.

M. Guillaume Bigot (RN). Certes, plutôt qu'à Orwell, c'est à Aldous Huxley qu'il faudrait faire référence puisque les algorithmes fonctionnent avec

l'assentiment des individus consommateurs. Quant à votre analyse du cadre européen, j'ai l'impression qu'on redécouvre l'eau chaude. Ce n'est pas de la nature humaine qu'il faut attendre l'absence de corruption mais de la séparation des pouvoirs – nous dit Montesquieu – ou de la démocratie directe – nous dit Rousseau – : tout ce qui manque à la machine qu'est l'Europe, dont la Commission n'est pas une instance démocratique. Dès lors, la corruption est normale.

Mme Constance Le Grip, présidente. Je m'exprimerai brièvement ici au nom du groupe EPR, d'abord pour remercier les rapporteurs de leur excellent travail. Je « n'achète » pas forcément l'ensemble de leurs dix-huit préconisations, dont certaines suscitent le débat et d'autres provoquent un choc. Pour en rester aux ingérences extérieures, notre sujet, il est temps en effet d'appeler à un réveil citoyen et d'augmenter les capacités de résistance, de résilience, comme les font les États baltes, la Suède, Taïwan. C'est aussi grâce à des citoyens informés, au-delà des structures d'État et de la réglementation, nationale ou européenne, que l'on pourra contrer les ingérences malveillantes.

Mme Laetitia Saint-Paul, rapporteure. Quelles sont les propositions que vous « n'achetez pas » ? Le savoir nous permettrait de mieux argumenter.

Mme Constance Le Grip, présidente. Je poursuivrai ce débat avec grand plaisir en dehors de cette séance mais je dois continuer à donner la parole aux orateurs des groupes.

M. Arnaud Le Gall (LFI-NFP). Ce rapport d'information a le mérite de traiter d'un sujet majeur : l'usage de l'IA dans les ingérences étrangères. Mais il convient de rappeler d'abord que pour l'essentiel, l'IA est au service des GAMAM (Google, Apple, Meta, Amazon et Microsoft) états-uniens, acteurs ultra-dominants en Europe. Le rapport détaille les ingérences les plus visibles, russes et autres, dont il faut évidemment se prémunir, mais il ne fait qu'effleurer, édulcorer la présence dans la pièce des éléphants que sont les GAMAM, alors même qu'en février dernier, à Paris et à Munich, le vice-président Vance, en leur nom, attaquait violemment les velléités européennes de réglementation. Vous soulignez à juste titre les manipulations russes des élections en Roumanie et en Moldavie mais pas la mise du réseau X au service de l'AfD en Allemagne par Elon Musk. Proposer de rendre obligatoire le consentement des usagers aux algorithmes est utile mais insuffisant. C'est aux algorithmes même qu'il faut avoir accès. Protéger un espace public démocratique de ces ingérences ne concerne pas seulement l'IA mais toute la filière numérique.

Pour cela, il faut cesser de déréglementer sous les injonctions des seigneurs de la Tech pour, comme y vise la loi sur la simplification économique, attirer des centres qui accumulent des milliards de milliards de données sur les individus pour nourrir les logiciels d'IA et qui, même stockées en France, sont régies pour l'essentiel par les lois des États-Unis. C'est accepter une colonisation numérique et engloutir dans la bulle de l'IA des milliards d'euros qui ne seront pas investis pour notre souveraineté.

Vous appelez à développer une puissance européenne mais l'Europe fait tout pour ne pas déplaire aux GAMAM et à la présidence des États-Unis. La Commission européenne ne fait pas appliquer la réglementation existante ; pire, l'omnibus numérique présenté le 19 novembre l'affaiblit. Selon l'organisation non gouvernementale *None of Your Business* (NOYB), c'est la pire attaque contre les droits numériques des Européens depuis des années. Emmanuel Macron a déroulé le tapis rouge aux GAMAM lors du sommet sur l'IA en février dernier, montrant qu'il ne croyait pas en un développement indépendant en ce domaine. La France a pourtant des atouts, comme des ingénieurs parmi les meilleurs du monde et une diplomatie en pointe pour réglementer l'IA à l'Organisation des Nations unies. Encore faut-il proposer une stratégie numérique alternative assurant la maîtrise des infrastructures clés et la constitution d'un domaine public des données, ainsi que le développement de modèles d'IA visant l'intérêt général et non la course au profit. Pourquoi ne pas prendre acte de la nécessité de changer de cadre ?

M. Alain David, rapporteur. Si nous nous tournons vers l'Europe, c'est que c'est notre cadre de référence.

M. Arnaud Le Gall (LFI-NFP). C'est bien le problème.

M. Alain David, rapporteur. Mais sinon, dans quel cadre s'inscrire ? Nous ne devons certainement pas nous arrimer encore plus aux États-Unis.

M. Arnaud Le Gall (LFI-NFP). Il faut croire en la France !

M. Alain David, rapporteur. Il faut acquérir notre indépendance sur le plan français, bien sûr, et sur le plan européen. Il n'y a pas d'autre solution pour être une force capable de rivaliser avec les États-Unis, la Chine, la Russie. Dans ce domaine, la France est un petit pays.

Mme Laetitia Saint-Paul, rapporteure. Nous traitons des effets de l'utilisation de l'intelligence artificielle dans les ingérences étrangères. Ce qu'a dit M. Vance n'entre pas dans ce cadre, le vice-président américain assumant ses propos en s'exprimant publiquement. Cela étant, le rapport n'occulte en rien l'attitude des États-Unis. Nous soulignons qu'ils détiennent la clé de la production et la clé de la diffusion et qu'ils sabordent notre clé d'utilisation. Néanmoins, ayant rencontré de très nombreux interlocuteurs, nous avons noté des résistances. Aux États-Unis, tous les services traitant des ingérences étrangères ont été supprimés. Le terme est presque devenu une grossièreté en soi – on m'avait indiqué qu'il me fallait peut-être l'éviter lors de ma rencontre avec mes interlocuteurs américains mais je n'en ai rien fait – et le travail sur la désinformation a été éteint. Facebook a ainsi supprimé son système de vérification des faits ou l'a réduit au strict minimum. Je n'ai pas parlé de Far West sans raison et, non, je ne passe pas à côté de l'éléphant dans la pièce en prétendant ne pas le voir.

Mme Pascale Got (SOC). Je remercie nos collègues pour la qualité de leur rapport, dans lequel ils documentent précisément la modification de la nature des ingérences contre nos démocraties induite par l'intelligence artificielle. Ils mettent

aussi en lumière un phénomène moins connu, l'empoisonnement volontaire des données servant à entraîner nos propres intelligences artificielles. Cette redoutable ingérence silencieuse est « *un véritable mécanisme de blanchiment de l'information* », écrivez-vous. Vous rappelez aussi que nos démocraties sont par nature plus vulnérables parce qu'elles croient au débat, à la transparence, à la diversité des opinions et que ceux qui veulent nous affaiblir utilisent ces mêmes notions pour les pervertir. Le concept de *sharp power* ou « pouvoir acéré » évoqué, qui consiste à attaquer notre modèle de l'intérieur en sapant progressivement la confiance des citoyens, résume cette stratégie.

Nous devons agir sur le plan technologique et sécuritaire mais la première ligne de protection est le citoyen. Sa capacité à repérer une manipulation et à garder l'esprit critique est devenue une condition de résilience de la démocratie. Les plus jeunes, les plus précaires, les moins familiers du numérique sont plus exposés que d'autres ; cela renforce la nécessité d'une action publique ambitieuse et équitable. Votre proposition visant à systématiser la formation à la vérification de l'information et à la lutte contre la désinformation dans le cursus des étudiants en journalisme montre qu'une réponse structurelle est possible. Mais quelles sont vos pistes pour renforcer l'éducation aux médias et à l'information ? Comment donner aux jeunes les bons réflexes face aux contenus manipulés, en particulier ceux qui ont été créés ou amplifiés par l'intelligence artificielle ? La Suède s'est dotée d'une Agence de défense psychologique. Quels dispositifs pourraient, en France, structurer une stratégie nationale de lutte contre la désinformation, contribuant ainsi à protéger le débat démocratique ?

M. Alain David, rapporteur. Actuellement, nous subissons. Nous préconisons d'agir plutôt que de subir mais pour agir il faut s'armer et progresser, et nous n'en sommes pas au stade où nous recherchons des solutions pour parvenir à la fois à financer ce secteur et à travailler sur la recherche. Des mesures visant à aiguiser l'esprit critique de la jeunesse ont déjà été prises et des enseignants s'y emploient mais une politique générale devrait être définie en liaison avec l'éducation nationale et d'autres services. Ce travail de longue haleine est indispensable.

Mme Laetitia Saint-Paul, rapporteure. Je me suis inspirée du modèle de l'Agence de défense psychologique suédoise. Cet organisme estime que 75 % de la population suédoise connaît ses missions. En France, Viginum n'est pas connu du grand public. Ce service nous fournit des rapports d'une extrême qualité et nous donne des consignes de politiques publiques mais il n'informe encore que très peu nos concitoyens de manière directe. J'ai évoqué l'esprit dans lequel agit l'Agence de défense psychologique suédoise devant mes interlocuteurs de Viginum, qui se sont dits capables de le suivre et volontaires pour le faire s'ils disposent des effectifs voulus. Viginum, qui travaille déjà avec l'Agence suédoise, serait l'organe pertinent de restitution des informations compilées à nos concitoyens. À l'éducation nationale, un partenariat existe entre la direction générale de l'enseignement scolaire et le centre de liaison de l'enseignement et des médias d'information visant à insérer dans les programmes scolaires des éléments d'éducation aux médias et à

l'information pour les élèves et pour les enseignants. Le mouvement est donc déjà amorcé.

M. Michel Herbillon (DR). Je vous remercie pour la présentation de ce rapport sur un sujet particulièrement complexe et sensible qui doit impérativement retenir notre attention. Je partage votre diagnostic : nous assistons à une augmentation inquiétante des ingérences étrangères dans notre espace numérique et votre rapport met en lumière le rôle de l'intelligence artificielle dans ce phénomène. Il n'est plus question d'artisanat mais d'une industrialisation de la déstabilisation, l'intelligence artificielle permettant de produire des faux au kilomètre et de saturer nos espaces de débat. Face à cette guerre hybride, la France a raison d'engager de nouveaux moyens pour lutter contre les ingérences étrangères, et je salue la récente création par le ministère de l'Europe et des affaires étrangères du compte *French Response* sur le réseau social X pour exposer les manœuvres adverses, les démentir et y répondre.

Néanmoins, certaines de vos propositions inquiètent et, contrairement à madame Le Grip qui a été d'une pudeur de gazelle, je dirai celles avec lesquelles je ne suis pas d'accord. Le développement de l'intelligence artificielle ne saurait devenir le prétexte à une régulation excessive de la liberté d'expression. La « labellisation informationnelle » selon le modèle du nutri-score suscite mon extrême scepticisme et vos arguments ne m'ont pas convaincu. Ce n'est pas parce que la proposition est issue des états généraux de l'information qu'elle est forcément bonne. La vitalité démocratique repose sur la pluralité des expressions permettant à chacun de se forger un avis éclairé. Quelle instance aurait la légitimité de distribuer les bons et les mauvais points de l'information sans dériver vers une forme de censure ?

Vous préconisez d'autre part l'instauration d'une « *réserve algorithmique préélectorale* ». Soyons réalistes : le fonctionnement des réseaux sociaux repose sur leur algorithme. Est-ce à dire que vous souhaitez les suspendre durant les campagnes électorales ? Cela ne me semble ni possible ni souhaitable.

Enfin, vous proposez d'imposer aux plateformes d'étiqueter les contenus avec un « *score d'artificialité* ». Une fois encore, il faut se garder de stigmatiser l'outil plutôt que l'usage qui en est fait. Ce n'est pas parce qu'un contenu a été produit ou aidé par une intelligence artificielle qu'il est nécessairement faux, manipulateur ou malveillant. L'intelligence artificielle peut être une arme à combattre dans le cadre des ingérences étrangères mais il faut se garder de sacrifier nos libertés numériques de manière inconséquente.

Mme Laetitia Saint-Paul, rapporteure. Labelliser l'information ne vise pas à s'en prendre aux médias d'opinion ou à entraver la liberté d'expression mais à contrer la génération de faux journalistes et de faux présentateurs par l'intelligence artificielle. Les représentants de France Médias Monde auditionnés nous ont expliqué que des sortes d'hologrammes de leurs journalistes sont créés, à qui l'on fait dire n'importe quoi. Nous ne sommes plus dans les années 1980, époque de la

dernière loi sur les médias. Les neuf critères objectifs retenus par *NewsGuard* pour évaluer la fiabilité d'un site d'information sont les suivants : « *ne publie pas de contenus faux ou manifestement trompeurs de manière répétée ; recueille et présente l'information de façon responsable ; dispose de procédures efficaces pour corriger les erreurs ; gère de manière responsable la différence entre informations et opinions ; évite les titres trompeurs ; le site indique à qui il appartient et comment il est financé ; identifie clairement la publicité ; indique qui est responsable des contenus et tout conflit d'intérêt possible : le site fournit des informations sur les créateurs de contenu* ».

Alors que nous vivons un chaos informationnel, devons-nous tout nous interdire au motif que parler de sécurité numérique serait entrer *de facto* dans un système de contrôle orwellien ? La question mérite d'être débattue ; peut-être est-elle entrée dans le débat public très récemment par la mauvaise porte.

Il est très facile d'instaurer, dans le cadre de l'article 38 du DSA, une réserve algorithmique préélectorale et de spécifier si l'on accepte ou non les algorithmes, en suivant la procédure *ad hoc*. Le problème, c'est l'existence même de cette procédure, ainsi programmée que pour l'instant on ne peut échapper par défaut aux algorithmes. Que se passe-t-il lors d'une journée sans algorithme sur vos réseaux sociaux ? Le fil d'actualité des personnes que vous suivez demeure disponible et vous renouez avec un contenu de qualité que vous avez choisi. Mais outre que la démarche doit être volontaire, puisque les réseaux sociaux ne le proposent pas par défaut, elle doit être répétée à chaque nouvelle connexion.

La proposition de réserve algorithmique préélectorale, extrêmement simple à mettre en œuvre, découle de ce qui s'est passé en Roumanie. Quand des influenceurs sont payés sans que, bien sûr, personne le sache, et que des dizaines de milliers de faux comptes sont créés en moins d'un mois, cela fausse le résultat du scrutin dans les pays démocratiques où les élections ne se jouent pas à 80 % des voix contre 20 %. Combien d'entre nous ont été élus par 60 % des voix contre 40 %, ou à 51 % contre 49 % ? Quand la majorité est de 50 % plus une voix, de telles interférences peuvent fausser un scrutin. C'est pourquoi je pense que la réserve algorithmique préélectorale est souhaitable pendant le mois précédant les élections, pour accéder aux contenus que l'on a choisis, *Libération*, *Le Monde*, CNews…

M. Jean-Paul Lecoq (GDR). *L'Humanité*…

Mme Laetitia Saint-Paul, rapporteure. J'ai oublié *L'Humanité*, pardonnez-moi. Quand on s'informe par ces médias, on sait à quoi s'attendre. Avec les contenus obtenus par algorithmes, que l'on ne paye pas puisque c'est la publicité qui finance, on ne sait pas ce que l'on achète et on ignore ce sur quoi l'on va tomber, mais il est certain qu'ils sont conçus pour nous garder captifs et nous diviser de manière que le temps passé sur les réseaux sociaux s'allonge et que les plateformes gagnent plus d'argent. La réserve algorithmique préélectorale est techniquement faisable et je suis convaincue qu'elle est souhaitable.

Le score d'artificialité est factuel. Il faut au moins connaître le pourcentage d'intelligence artificielle par métadonnées. Il est nécessaire de savoir que la présentatrice que l'on voit et écoute n'existe pas.

M. Alain David, rapporteur. En Roumanie, le candidat qui a été aidé par l'intelligence artificielle était crédité de 2 ou 3 % des voix au début de la campagne électorale ; huit jours plus tard, influenceurs et intelligence artificielle aidant, la proportion était supérieure à 20 %. Ces manipulations faussent complètement une élection, au point que le premier tour de l'élection présidentielle roumaine a dû être annulé.

M. Michel Herbillon (DR). Nous sommes tous conscients des conséquences de ces ingérences sur le résultat des élections. La présidente de la Moldavie m'a expliqué ce qui se passait dans son pays et les exemples que vous avez cités sont éloquents. Mais, comme l'a dit aussi madame Pascale Got, il faut trouver le bon équilibre entre la régulation nécessaire et la liberté du citoyen, et ce n'est pas simple.

M. Jean-Louis Roumégas (EcoS). Je remercie les rapporteurs pour ce travail de qualité. Il met en évidence les menaces que fait peser sur nos démocraties la guerre menée par le biais de l'intelligence artificielle. Elle permet de pilonner notre société par une désinformation d'une ampleur inédite mais, si cet outil pose problème, il constitue aussi une partie de la solution et nous ne pouvons pas laisser d'autres puissances en avoir la maîtrise. L'Europe doit impérativement développer des capacités souveraines en matière d'intelligence artificielle. Actuellement, notre dépendance technologique est trop forte ; il faut donc investir pour garantir notre autonomie et faire face aux ingérences étrangères.

Or les ressources de l'audiovisuel public ne cessent de diminuer, ce qui contredit l'objectif que nous visons. Comment lutter contre la submersion de *fake news* générées par l'intelligence artificielle si l'on ne se donne pas les moyens de renforcer l'indépendance et le pluralisme des médias et les processus de vérification de l'information ? La Commission européenne propose, avec le Bouclier européen pour la démocratie, de combattre la concentration excessive des médias et de renforcer le soutien financier au secteur, la première arme contre la désinformation étant bien sûr l'information vérifiée. Or, en France, quelques grands groupes détenus par des milliardaires se partagent de plus en plus le gâteau – la presse, les maisons d'édition, l'audiovisuel – et en profitent parfois pour diffuser des idées univoques, souvent nauséabondes. Il en résulte que les rédactions perdent leur indépendance, ce qui accroît encore la défiance vis-à-vis des médias traditionnels, dont 62 % des Français déclarent se méfier aujourd'hui, contre 54 % en 2022. Face aux ingérences étrangères, il nous faut mettre au point une intelligence artificielle souveraine et éthique mais aussi garantir des médias indépendants et de qualité.

Enfin, comment développer une intelligence artificielle compétitive mais compatible avec nos impératifs écologiques ?

M. Alain David, rapporteur. Une commission d'enquête sur la neutralité, le fonctionnement et le financement de l'audiovisuel public vient d'être lancée, dont nous examinerons attentivement les conclusions. S'agissant de l'audiovisuel extérieur de la France, nous avons insisté auprès du gouvernement pour que l'entité France Médias Monde soit suffisamment financée. Nous n'avons pas été forcément entendus mais il est nécessaire qu'au cours des prochaines années cette expression de la France à l'étranger soit privilégiée. Le groupe France Médias Monde, qui contribue à la lutte contre les *fake news*, à la transmission de nos valeurs et au rayonnement international de notre pays, doit être respecté et aidé dans sa mission.

Mme Laetitia Saint-Paul, rapporteure. Lors de son audition, Mme Marie-Christine Saragosse, présidente-directrice générale de France Médias Monde, a exposé les programmes mis au point par l'entreprise pour mettre à jour les infox et décrypter les manipulations. Elle bénéficiait aussi, à cette fin, de contrats privés avec des plateformes mais, bien sûr, ces contrats sont en perdition. Il est nécessaire de maintenir des médias publics forts, d'autant que le budget de la *U.S. Agency for Global Médias*, le pendant américain de France Médias Monde, a été passé à la tronçonneuse ; on peut s'inquiéter de savoir qui occupera le vide qui va être laissé par les médias publics américains. Régulièrement en voyage, je peux témoigner que France Médias Monde, qui diffuse dans de nombreuses langues, porte haut nos couleurs en donnant à entendre des programmes de qualité, reconnus comme tels. On notera aussi que les chaînes de France Médias Monde ne sont pas sur le banc des accusés dans les procès faits actuellement aux médias du chef de manque de pluralisme. Nous devons faire attention à ce que nous détruisons nous-mêmes.

M. Alain David, rapporteur. J'appelle l'attention sur les financements respectifs des médias publics selon les pays : environ 200 millions d'euros pour France Médias Monde, quelque 479 millions pour BBC World, 420 millions pour Deutsche Welle, plus de 1 milliard pour les Chinois. On voit les différences de force de frappe.

M. Lionel Vuibert (NI). Le sujet relève de la sécurité nationale et je suis surpris que nous soyons aussi peu nombreux pour en débattre. Élu d'une circonscription qui compte 110 000 habitants, je rencontre beaucoup de monde et je suis effaré par le nombre d'âneries que j'entends. On crée des bulles informationnelles artificielles dans lesquelles les gens se perdent faute de préparation à cette exposition malveillante. Certes, on ne pouvait pas les y préparer puisque l'outil n'existait pas mais, à présent, il faut éduquer à l'intelligence artificielle comme on éduque à l'utilisation des outils en tous domaines.

Peut-on se préserver de ces usages abusifs ? Nous devons en finir une fois pour toutes avec la naïveté et agir. Je suis plutôt favorable à le faire dans le contexte européen mais l'urgence est telle que s'il est techniquement possible de rendre obligatoire le consentement aux algorithmes dans les suggestions de contenus sur les plateformes, ne peut-on légiférer en ce sens ? Et puis, si des outils numériques

permettent de marquer les images ou les contenus générés en tout ou en partie par l'intelligence artificielle, il faut l'imposer aux plateformes de réseaux sociaux.

L'urgence est absolue ; il en va de notre sécurité nationale intérieure et extérieure. Nous dansons sur un volcan et pour éviter les manipulations à l'avenir, nous devons donner à nos concitoyens les outils leur permettant de débusquer les fausses informations et les informations trafiquées. Ne pas le faire, c'est s'exposer à de grandes déconvenues lors des élections qu'il est devenu très facile de faire basculer, avec des conséquences qui peuvent être dramatiques – sans même parler de ce qui peut advenir quand on se trouve dans un conflit comme celui qui a lieu aux portes de l'Union européenne. En résumé, il y a urgence à agir dans un contexte européen et même international, comme l'a suggéré monsieur Lecoq, tout en prenant des mesures immédiates tels que le réglage des logiciels par défaut et le marquage des contenus modifiés par l'intelligence artificielle. C'est un appel au secours : nous devons nous saisir du sujet.

Mme Laetitia Saint-Paul, rapporteure. Nous l'avons écrit dans le rapport : les sociétés démocratiques sont des proies pour les régimes autoritaires. Ils visent à les détruire par leurs ingérences parce que ce modèle social leur fait horreur et le moment où les sociétés ouvertes sont les plus vulnérables est celui des campagnes électorales. Les élections municipales auront lieu en mars 2026, dans peu de temps. Que nous ne pensions pas en termes de temps long aggrave notre sort. Pour les Chinois, l'art suprême de la guerre est de soumettre l'ennemi sans combattre, peu importe le temps qu'il y faudra ; et pour eux, le temps long n'est pas le temps d'une vie, c'est un millénaire. En Russie, le KGB (Comité pour la sécurité de l'État) avait pour doctrine de base la démoralisation de l'ennemi ; là aussi, on œuvre sur le temps long. Aux États-Unis, les plateformes numériques ont fondé leur modèle économique sur le clivage pour allonger le temps passé en ligne par les consommateurs et, de la sorte, accroître leurs bénéfices. Nous sommes la proie de ces différentes puissances et nous ne voulons pas le voir. La prise de conscience est un bon début. La population des États baltes, de la Norvège, de la Suède et de la Finlande est éclairée parce qu'elle accepte de voir ce qui a été décrit ce matin et ce à quoi nous sommes aveugles.

Nous en appelons donc à la vigilance citoyenne et c'est pourquoi la plupart de nos propositions partent du citoyen, et non du gouvernement. Tous les services de l'État que j'ai rencontrés ont pour hantise d'être accusés de devenir le « ministère de la Vérité », le *Miniver* orwellien ; tous y ont fait allusion devant moi. Ils abordent donc ces sujets avec une très grande prudence.

Aussi, nos propositions visent à renforcer la vigilance du citoyen, à faire qu'il choisisse ce qu'il consomme et reprenne la main. Mais pour reprendre la main, encore faut-il savoir qu'on est manipulé. La question des algorithmes est apparue depuis très peu de temps dans le débat public. Les contenus proposés sont divertissants et on ne s'interrogeait pas spécialement sur leur substance. L'appel au discernement est assez nouveau et nous ne sommes qu'au début d'un long processus. Aussi, madame la présidente, je vous propose de faire un point d'étape

dans un an : on saura alors quelle suite aura été donnée à la proposition de la Commission européenne visant à créer le Bouclier démocratique européen en instaurant un centre européen de résilience démocratique.

M. Jean-Paul Lecoq (GDR). Entre-temps, combien de mairies seront tombées ?

M. Alain David, rapporteur. De fait, on voit en France qu'en matraquant un « produit politique » à longueur de journées sur les plateaux de télévision et dans les médias, on en fait en quelque sorte un produit commercial qui peut finalement émerger dans l'opinion publique. Je ne pense à personne, bien entendu, mais cela en dit long sur la force de persuasion des médias sur l'opinion publique et, par là même, sur la force de persuasion des ingérences étrangères qui veulent notre perte.

Mme Constance Le Grip, présidente. Je prends note de la suggestion de madame la rapporteure. La commission des affaires européennes va lancer un travail de six mois sur le Bouclier démocratique européen. Le président Fuchs et le bureau de notre commission établiront comment articuler les travaux au mieux.

Nous en venons à présent aux interventions et questions formulées à titre individuel.

M. Pierre Pribetich (SOC). Je félicite les rapporteurs. Il faut, avez-vous dit, reprendre la main. Cela demande la maîtrise des algorithmes et donc des constructions d'arbres de décision, d'une part, des réseaux neuronaux, d'autre part. Pour parvenir, comme vous le proposez, au consentement obligatoire à l'algorithme ou à l'instauration d'une réserve algorithmique préélectorale, encore faut-il pouvoir maîtriser les composants dont nous ne disposons pas, l'intégralité des puces servant à l'intelligence artificielle étant fabriquées à Taïwan. Que préconisez-vous à ce sujet ?

Mme Laetitia Saint-Paul, rapporteure. Le rapport peut donner l'impression que nous sommes désarmés mais nous ne le sommes pas, pour la raison qu'aucun grand acteur économique dans le monde ne peut se passer du marché européen et de ses 18 000 milliards d'euros de produit intérieur brut. Aussi devons-nous exiger de la Commission européenne qu'elle fasse respecter le cadre établi pour l'espace informationnel européen.

Imposer un consentement algorithmique est possible techniquement mais il faut la volonté politique de l'imposer à tous. Cela se résume en une phrase : « Vous voulez vendre sur notre marché ; c'est possible, à condition de respecter nos règles ». Cela vaut pour les produits vendus sur les plateformes chinoises comme pour les géants du numérique. Nous, Européens, devons cesser de considérer l'Europe comme un nain alors que la force de notre marché nous donne le pouvoir de faire plier quiconque.

Mme Véronique Besse (NI). Votre rapport montre que l'intelligence artificielle n'est pas seulement un nouvel outil d'ingérence mais un multiplicateur

de puissance à bas coûts structurant dans la guerre hybride contemporaine. Vous soulignez l'effondrement du prix de production des contenus, la massification des opérations et la montée en puissance d'acteurs hybrides ou para-Étatiques. L'exemple de l'opération Storm illustre ce basculement : une capacité d'impact massive obtenue avec des ressources humaines et financières limitées, et surtout avec une attribution politique volontairement opaque. Notre appareil de contre-ingérence est-il suffisamment structuré pour répondre à une menace à très faible coût d'entrée devenue persistante, ainsi conçue que la saturation de l'espace informationnel prime désormais la persuasion directe ?

Mme Laetitia Saint-Paul, rapporteure. Nous proposons de renforcer les moyens de Viginum. Son équivalent suédois, l'Agence de défense psychologique, compte au moins un tiers d'agents en plus, alors même que la Suède est un plus petit pays que la France. Il ne s'agit pas de faire grandir Viginum exagérément, car le service ne doit pas perdre son agilité et sa réactivité, mais de lui permettre d'élargir le spectre de ses missions, comme je l'ai évoqué précédemment.

D'autre part, nous devons être plus offensifs. Jusqu'à présent, nous nous le sommes interdit, considérant que le remède serait pire que le mal, mais c'est possible et la communication visant à démontrer la fausseté de certaines informations *(debunking)* que le Quai d'Orsay commence à mettre en œuvre est plus offensive, comme il est nécessaire. Face aux nouvelles formes de menaces, nous devons revoir nos pratiques.

M. Alain David, rapporteur. De plus, la DGSE et d'autres services de sécurité extérieure et intérieure surveillent étroitement tout ce qui peut s'apparenter à des délits pour débusquer et localiser ceux qui contreviennent à nos intérêts.

M. Arnaud Le Gall (LFI-NFP). L'Union européenne serait donc la seule échelle à laquelle nous pouvons développer notre indépendance numérique, et vous insistez sur l'importance du marché européen, qui ferait de l'Europe une puissance. Mais il ne suffit pas d'être un marché pour être une puissance. Dans les années 1990 déjà, l'Union européenne disait : « Laissons faire le marché en matière numérique ». Au même moment, les États-Unis et la Chine planifiaient ce développement chacun à leur manière et mettaient au service d'une stratégie numérique les moyens de la puissance publique, soit par la commande publique, soit par les subventions, soit par les investissements publics directs ou indirects ; on voit où cela nous a menés.

On nous dit maintenant que, désormais, l'Europe réglemente. Mais il suffit que l'administration américaine s'en mêle pour que, immédiatement, on commence à tout détricoter. La Commission européenne est totalement gangrenée par les lobbies des GAMAM, plusieurs orateurs l'ont dit. On pourra donc voter toutes les réglementations européennes que l'on voudra, les GAMAM feront ce qu'il faut pour les détricoter. L'Omnibus numérique proposé ne fait pas autre chose.

Idéalement, l'échelle européenne est, bien sûr, la meilleure mais cela ne se fera pas. Admettons donc qu'il convient de changer de cadre, sans courir derrière le

gigantisme des acteurs chinois ou américains car cela n'a de sens ni écologiquement ni économiquement. Nous sommes dans un capitalisme de prédation tributaire qui ne peut être que colonialiste, puisque partager les données c'est perdre l'essentiel de la richesse en matière d'intelligence artificielle. Il faut changer d'angle, déterminer des besoins précis d'intérêt général en matière de renseignement militaire, de santé publique, de recherche, et développer nos propres solutions. Nous avions, historiquement, un immense savoir-faire en matière de numérique, et des ingénieurs capables de le faire ; on doit pouvoir y arriver. Dire en permanence que tout se fera au niveau européen, c'est de la poudre aux yeux, puisqu'il le faudrait mais que cela ne se fera pas.

Mme Laetitia Saint-Paul, rapporteure. Que tout se pense au niveau européen n'est pas la conclusion à laquelle nous parvenons. Beaucoup de nos propositions se situent à l'échelle nationale. D'autre part, l'espace informationnel européen a été pensé de manière que l'on puisse connecter des directives nationales qui peuvent infuser au niveau européen. C'est pourquoi il m'est insupportable que la loi Marcangeli visant à instaurer une majorité numérique, votée ici à l'unanimité, soit retoquée par la Commission européenne. En fait, ce n'est pas possible. J'ai épluché son rapport à ce sujet, long de 60 pages. Entre acronymes et notes de bas de page, tout est fait pour perdre le lecteur mais la Commission européenne ne présente aucun argument digne de ce nom. De plus, depuis le 14 juillet 2025, les dernières réticences sur la majorité numérique étant levées, je n'accepte pas cette situation. Oui, un atlantisme historique prévaut chez la plupart de nos partenaires européens qui dépendent des États-Unis pour leur défense, c'est un fait. Mais le renoncement en faveur de la stabilité, comme celui que nous a fait subir Ursula von der Leyen au sujet des droits de douane trumpistes, cela ne marche jamais, prenons-en acte.

Enfin, l'Omnibus numérique annoncé par la Commission doit encore être accepté par les États membres et le Parlement européen. Tous nos groupes ont des députés européens…

M. Jean-Paul Lecoq (GDR). Non.

Mme Laetitia Saint-Paul, rapporteure. C'est notre rôle de nous y opposer et toutes les initiatives qui vous viendraient à l'esprit seront les bienvenues.

M. Arnaud Le Gall (LFI-NFP). L'administration française n'est pas la dernière à pousser en faveur de la déréglementation en ce domaine.

M. Alain David, rapporteur. En visant à mobiliser l'Europe dans la lutte contre les ingérences étrangères, nous visons l'idéal. Y parviendra-t-on ? De même, face à l'attitude des États-Unis, on vise le réarmement de l'Europe. Y parviendrons-nous ? Là encore, c'est l'idéal auquel il faut tendre.

Mme Constance Le Grip, présidente. Il a beaucoup été question de Viginum, le service de vigilance et de protection contre les ingérences numériques étrangères. Je précise que cette agence de l'État est placée sous l'égide du secrétariat

général de la défense et de la sécurité nationales (SGDSN) ; cela dit que les questions abordées aujourd'hui sont des questions de sécurité nationale. Viginum est en quelque sorte le pendant de l'ANSSI, l'autre agence placée sous l'autorité du SGDSN, elle-même compétente en matière de lutte contre les cyber-attaques. Viginum est chargé de prévenir, de détecter et de contrer les campagnes d'ingérence numérique étrangères malveillantes et toxiques. Ce n'est pas un service de renseignement. Je souligne aussi qu'aucun de nos services de renseignement ne pratique une surveillance électorale ou politique.

Le Quai d'Orsay comme le ministère des armées ont engagé des stratégies de riposte et une réflexion de fond s'élabore visant à passer d'une stratégie défensive à une stratégie proactive et offensive. C'est le débat permanent qui se pose à notre démocratie : nous ne souhaitons ni utiliser les instruments des puissances illibérales et autoritaires ni rester une proie, et nous voulons être des combattants de la liberté.

Conformément à l'article 145 du Règlement de l'Assemblée nationale, à l'issue des échanges, la commission ***autorise*** *la publication du rapport d'information qui lui a été présenté.*

Mme Constance Le Grip, présidente. Je félicite nos rapporteurs et les remercie pour ce travail passionnant.

ANNEXE N° 1 :
LISTE DES PERSONNES AUDITIONNÉES PAR LES RAPPORTEURS

- M. Bernard Benhamou, secrétaire général de l'Institut de la Souveraineté numérique ;
- Mme Marie-Doha Besancenot, conseillère communications stratégiques, cabinet du ministre de l'Europe et des affaires étrangères ;
- Mme Céline Berthon, directrice générale de la sécurité intérieure (DGSI) ;
- M. Marc-Antoine Brillant, chef du service de vigilance et de protection contre les ingérences numériques étrangères (Viginum) ;
- M. Nicolas Curien, professeur, membre fondateur de l'Académie des technologies et pilote du groupe de travail « IA générative et mésinformation » ;
- Mme Anne-Sophie Dhiver, chef de service adjointe de Viginum ;
- Mme Chine Labbé, rédactrice en chef et vice-présidente senior en charge des partenariats, Europe et Canada, NewsGuard ;
- M. Christophe Lemoine, directeur de la communication et de la presse, porte-parole, ministère de l'Europe et des affaires étrangères ;
- M. Corentin Masclet-Andrieu, responsable relations institutionnelles et communication transverse, France médias monde ;
- M. Benjamin Pajot, chercheur associé du centre Géopolitique des technologies de l'Institut français des relations internationales (IFRI) ;
- Mme Laure de Roucy-Rochegonde, directrice du centre Géopolitique des technologies de l'IFRI ;
- Mme Marie-Christine Saragosse, présidente-directrice générale de France médias monde ;
- M. Vincent Strubel, directeur général de l'Agence nationale de la sécurité des systèmes d'information (ANSSI) ;
- M. Charles Thibout, chercheur associé à l'Institut de relations internationales et stratégiques (IRIS) ;
- M. Henri Verdier, ambassadeur pour le numérique ;
- Direction générale de la Sécurité extérieure (DGSE).

Déplacement de la rapporteure aux États-Unis (6-11 juillet 2025)

À New York (7 et 8 juillet)

- M. Jérôme Brunet-Possenti, chef de service adjoint, pôle « enjeux globaux », mission permanente de la France auprès des Nations Unies à New York ;

- M. Mark Dimont, chef de produit « actualités, web et applications des médias sociaux », Bloomberg ;
- Mme Angela Moon, directrice exécutive « actualités mondiales », Bloomberg ;
- Mme Jennifer P., chargée de programme dans l'équipe « intégrité de l'information », département de la communication globale des Nations Unies ;
- M. Adam Segal, directeur du programme politiques digitales et du cyberespace, *Council on Foreign Relations* (CFR) ;
- Mme Charlotte Scaddan, conseillère senior sur l'intégrité de l'information, département de la communication globale des Nations Unies ;
- M. Vikram Sura, spécialiste senior des affaires internationales, département de la communication globale des Nations Unies.

À Washington (8 au 11 juillet)

- Mme Vanessa Audemard d'Alançon, responsable de la communication stratégique et de l'intégrité de l'information, service presse et communication, ambassade de France aux États-Unis ;
- Mme Rita Balogh, responsable des affaires publiques et gouvernementales en Europe, Google
- M. Eric Berthillier, attaché de sécurité intérieure adjoint, ambassade de France aux États-Unis ;
- Mme Jessica Brandt, chercheuse principale sur la technologie et la sécurité nationale, *Council on Foreign Relations* (CFR) ;
- M. Marc Brocheton, conseiller économique, chef du pôle entreprise et innovation, ambassade de France aux États-Unis ;
- M. Emerson T. Brooking, directeur de la stratégie et chercheur principal résident, DFRLab, *Atlantic Council* ;
- M. Valentin Châtelet, chercheur associé sur la sécurité, DFRLab, *Atlantic Council* ;
- M. Damien Cristofari, premier conseiller, chef de la section politique, ambassade de France aux États-Unis ;
- M. Marc Dagorn, conseiller douanier pour les États-Unis, le Canada et le Mexique, direction générale des douanes et droits indirects (DGDDI), ambassade de France aux États-Unis ;
- M. Jean Baptiste Descollonges, ingénieur en chef de l'armement, section coopération armement, ambassade de France aux États-Unis ;
- Mme Renée Di Resta, enseignante et chercheuse, université Georgetown ;
- M. Michael Flossman, chercheur, OpenAI ;
- Mme Kim Gajewski, consultante principale, Google ;

- M. Jean-Olivier Grall, capitaine de vaisseau, attaché de défense adjoint marine, ambassade de France aux États-Unis ;
- M. Tim Harper, analyste politique senior et chef de projet élections et démocratie, *Center for Democracy and Technology* (CDT) ;
- M. Max Katz, staffer démocrate, Sénat américain ;
- Mme Megan Keeling, analyste senior en renseignement sur les menaces stratégiques, *Recorded Future* ;
- Mme Julie Lavet, responsable des relations et partenariats avec les États membres de l'Union européenne, OpenAI ;
- M. Dave Leichtman, directeur mondial élections, Microsoft ;
- Mme Isabel Linzer, chercheuse sur les élections et la démocratie, CDT ;
- M. Brian Liston, analyste senior en renseignement sur les menaces stratégiques, *Recorded Future* ;
- M. Ben Nimmo, enquêteur principal sur les menaces, équipe renseignement et enquêtes, OpenAI ;
- Jay Obernolte, représentant républicain de Californie, coprésident de la *task force* bipartisane sur l'intelligence artificielle (« *House bipartisan task force on artificial intelligence* ») en 2024 ;
- Mme Fanta Orr, directrice de l'analyse du renseignement, Microsoft ;
- M. Philippe Ribière, attaché pour la science et la technologie, ambassade de France aux États-Unis ;
- Mme Sarah Rogers, future under secretary for public diplomacy ;
- M. David Salvo, directeur exécutif, *Alliance for Securing Democracy*, *German Marshall Fund* (GMF) ;
- M. Bret Schafer, chercheur principal sur la désinformation dans les médias et le numérique, *Alliance for Securing Democracy*, GMF ;
- Mme Kenton Thibaut, chercheuse résidente principale, Chine, DFRLab, *Atlantic Council* ;
- Mme Océane Thiériot, conseillère pour la cybersécurité, le numérique et les technologies émergentes, ambassade de France aux États-Unis.

<u>Déplacement de la rapporteure à Bruxelles (14 octobre 2025)</u>

- M. Xavier Coget, conseiller diplomatique de Mme Henna Virkkunen, commissaire européenne et vice-présidente exécutive en charge de la souveraineté technologique, de la sécurité et de la démocratie ;
- M. Thomas Schmitz, chargé des médias et du bouclier démocratique, cabinet de Mme Henna Virkkunen, commissaire européenne et vice-présidente exécutive en charge de la souveraineté technologique, de la sécurité et de la démocratie ;

- M. Thierry Boulangé, adjoint au chef d'unité, division règlement et conformité en matière d'IA du bureau de l'intelligence artificielle, direction générale des réseaux de communication, du contenu et des technologies de la Commission européenne ;
- Mme Mélanie Gornet, experte technique, division règlement et conformité en matière d'IA du bureau de l'intelligence artificielle, direction générale des réseaux de communication, du contenu et des technologies de la Commission européenne ;
- M. Albin Birger, chef d'unité, division convergence des médias et médias sociaux direction générale des réseaux de communication, du contenu et des technologies de la Commission européenne ;
- M. Mickaël Bazin, expert technique, division surveillance et coopération « *digital service act* ».
- M. David Ringrose, chef de la division connectivité et transition digitale, direction des enjeux globaux, service européen pour l'action extérieure ;
- M. Jacob Tamm, adjoint au chef de la division de l'intégrité de l'information et de la lutte contre la manipulation et l'ingérence étrangères en matière d'information, direction de la planification politique et de la communication stratégique, service européen pour l'action extérieure ;
- M. Julien Chassereau, directeur associé pour l'IA et la data policy, *Digital Europe* ;
- M. Sacha Baudinet, conseiller cyber-hybride et manipulations de l'information, représentation permanente de la France auprès de l'Union européenne ;
- Mme Marie-Léa Rols, conseillère télécommunication, numérique et postes, représentation permanente de la France auprès de l'Union européenne.

<u>Déplacement de la rapporteure à Stockholm (12-14 novembre 2025)</u>

- Mme Hanna Gunnarsson, députée suédoise, membre de la commission de la Défense du Riksdag ;
- M. Lars Püss, député suédois, membre de la commission de la Défense du Riksdag ;
- M. Antonin Gauthier, premier secrétaire et chef du service presse de l'ambassade de France en Suède ;
- Mme Linn Mardstam, attachée au service presse de l'ambassade de France en Suède ;
- M. Steven Lloyd Wilson, chercheur et chef du projet de société digitale du *V-dem Institute* ;
- M. Fredrick Bynander, chef du secrétariat de coordination de crise, Conseil national de sécurité de Suède ;
- M. Vincent Boulanin, directeur du programme « Gouvernance de l'IA » au *Stockholm International Peace Research Institute* (SIPRI) ;
- M. Tor-Björn Åstrand, chef du département recherche et analyse de l'Agence de défense psychologique (Myndigheten för psykologiskt försvar – MPF) ;
- Mme Karin Klasa, analyste, MPF ;
- Mme Sara Sörensen, analyste, MPF ;

- Mme Karina Shyrokykh, chercheuse sur la formation des populations civiles face à la manipulation de l'information et experte en matière de propagation de législation illibérale, Université de Stockholm.

Dans le cadre d'un déplacement organisé par le groupe d'études à vocation internationale de l'Assemblée nationale à Taïwan, du 6 au 13 juin 2025, la rapporteure Laetitia Saint-Paul a également rencontré le ministère des affaires étrangères de Taïwan ; le ministère de la Justice et son Bureau d'investigation ; la Commission des affaires continentales ; le ministère des affaires numériques ; le Bureau français de Taipei ; l'Institut de recherche pour la démocratie, la société et les nouvelles technologies ; l'Institut de recherche sur la sécurité et la défense nationale ; le Centre de recherche et de développement d'applications innovantes en matière d'IA pour drones en Asie ; le DoubleThink lab.