L'incertitude comme ressource stratégique : deepfakes et recomposition des rapports de force

Angèle Billaud

Analyste en droit international, sécurité internationale, cybersécurité et défense, diplômée de l'Université Grenoble Alpes, France.

28 mai 2026

ISSN : 2739-3283

Comment citer cette publication :

Angèle Billaud, *L'incertitude comme ressource stratégique : deepfakes et recomposition des rapports de force,* Institut d'études de géopolitique appliquée, Paris, 28 mai 2026.

66 avenue des Champs-Élysées, 75008 Paris
Courriel : secretariat@institut-ega.org
Site internet : www.institut-ega.org

SOMMAIRE

Introduction

Avec l'émergence de nouvelles technologies sont apparues de nouvelles ressources stratégiques qui contribuent à remettre en question les dynamiques de puissance, en temps de paix comme en temps de guerre. La capacité d'un acteur à orienter l'environnement stratégique à son avantage ne repose plus uniquement sur des moyens militaires ou économiques classiques, mais s'exerce aussi dans l'espace informationnel, devenu un champ de confrontation à part entière. Les conflits contemporains ont ainsi renforcé l'enjeu que constitue la maîtrise de certaines technologies clé, au premier rang desquelles figure l'intelligence artificielle.

L'IA permet d'automatiser un certain nombre de tâches et d'accélérer le traitement de volumes massifs de données. Elle promet, en théorie, une analyse plus complète du champ de bataille, une prise de décision plus fluide et contribuerait à prévenir la surcharge cognitive du soldat. Toutefois, ses capacités stratégiques dépassent largement le seul traitement de données, notamment dans le domaine de la génération de médias synthétiques.

Parmi les vidéos, audios et images générés par des systèmes d'IA, une catégorie de contrefaçons numériques particulièrement réalistes retient l'attention : les « deepfakes ». Ces hypertrucages sont capables d'imiter la voix d'une personne ou de modifier une image de manière presque imperceptible. Un deepfake peut, par exemple, usurper l'identité d'un dirigeant d'entreprise lors d'une visioconférence, permettant de transmettre des consignes contraires aux intérêts de l'organisation tout en bénéficiant d'une présomption de crédibilité liée à la fois au canal de communication et au réalisme de la falsification[1].

Les deepfakes se sont d'abord retrouvés au cœur des débats politiques et juridiques en raison de leur capacité à faciliter les fraudes ou à détourner des images à des fins pornographiques. Toutefois, le conflit russo-ukrainien a mis en lumière une autre possibilité d'utilisation du deepfake. Cette nouvelle dimension intéresse particulièrement la sphère militaire en raison de ses effets sur le déroulement des hostilités. En effet, les deepfakes peuvent également servir à manipuler l'information, ce qui constitue un levier central dans les conflits dits hybrides, mêlant affrontement conventionnel et exploitation stratégique de l'espace numérique.

Les campagnes de désinformation ne sont pas nouvelles. Elles visent depuis longtemps à affaiblir l'adversaire par la confusion, la polarisation de la société ou l'érosion de la confiance d'un peuple envers son État et envers les alliances interétatiques[2]. L'intelligence artificielle introduit cependant un effet d'industrialisation de ces pratiques[3] en renforçant le réalisme des contenus, en augmentant la vitesse et la quantité de production, et en automatisant leur diffusion à grande échelle. De nombreuses études montrent d'ailleurs que les fausses informations

[1] NAFFI Nadia, *Deepfakes and the crisis of knowing,* UNESCO, Octobre 2025, https://www.unesco.org/
[2] BYMAN Daniel, GAO Chongyang, MESEROLE Chris, SUBRAHMANIAN V.S, *Deepfakes and international conflict,* Foreign Policy at Brookings, Janvier 2023, https://www.brookings.edu/
[3] EH&A Consulting, *Intelligence artificielle et manipulation de l'information*, mars 2023, https://www.eha-consulting.com/

tendent à se propager plus rapidement que les vraies, et que les tentatives de rétablissement de la vérité peuvent parfois renforcer la dynamique de désinformation[4]. Les deepfakes permettent en outre de dissimuler l'origine d'une ingérence tout en préservant ses effets : accentuation des divisions internes, remise en cause de l'appareil démocratique, perte de confiance envers les institutions étatiques[5].

Cette massification de la production et de la diffusion de deepfakes accroît leur impact émotionnel et psychologique sur les publics ciblés. Elle contribue à l'installation d'un doute généralisé, où la désinformation semble omniprésente et où plus aucune source, même officielle, ne paraît totalement fiable. Cette logique, susceptible de se diffuser tant dans les sphères civiles que militaires, constitue une menace croissante en raison de l'avantage stratégique qu'elle confère : celui de l'incertitude. Dès lors, la capacité des États à maîtriser ces outils ou à s'en prémunir apparaît comme un facteur déterminant dans la recomposition des rapports de force à l'œuvre depuis plusieurs années.

Les deepfakes s'imposent en tout état de cause comme de nouveaux instruments de puissance dans les conflits contemporains. Leur émergence fait apparaître différents défis stratégiques et normatifs. Alors qu'ils sont capables de bousculer les équilibres informationnels et stratégiques, leur maîtrise et leur encadrement apparaissent comme des défis majeurs pour les acteurs étatiques et internationaux.

[4] BYMAN Daniel, GAO Chongyang, MESEROLE Chris, SUBRAHMANIAN V.S, *Deepfakes and international conflict, op., cit*.
[5] GALICIA Bailey, *In the fight against foreign information manipulation, the US can't afford to disarm,* Atlantic Council, Juillet 2025, https://www.atlanticcouncil.org/

Les deepfakes comme instruments de puissance : recomposition des équilibres informationnels et stratégiques

Les deepfakes ne constituent pas seulement une innovation technologique. Ils participent aussi à une transformation structurelle des rapports de pouvoir dans l'espace informationnel. En altérant les conditions de production, de circulation et de vérification de l'information, ils fragilisent les fondements de la confiance collective, redéfinissent les modalités de la conflictualité et s'imposent progressivement comme un outil stratégique à part entière. Ils constituent dès lors un levier de puissance, dont les effets se déploient aussi bien en temps de paix que dans la conduite des hostilités, et dont les démocraties doivent appréhender l'ambivalence.

Les deepfakes en temps de paix : érosion diffuse de la confiance sociale et institutionnelle

L'IA est désormais capable de générer des deepfakes, sous forme d'images, de vidéos ou d'enregistrements audio, à partir de volumes de données limités, parfois même en temps réel et à un coût très faible[6]. Cette évolution technologique s'accompagne d'une diffusion rapide des outils nécessaires à leur production. Certains modèles d'hypertrucage sont aujourd'hui disponibles en source ouverte, permettant à un large public de générer de tels contenus synthétiques[7] : une technologie autrefois réservée à des acteurs disposant de ressources importantes est désormais accessible à un grand nombre d'utilisateurs, y compris à des individus cherchant à mener des actions malveillantes. Les conséquences économiques de ces usages frauduleux sont déjà mesurables. Aux États-Unis les pertes liées à des fraudes par IA générative pourraient atteindre 40 milliards de dollars d'ici 2027 (contre 12,3 milliards en 2023)[8].

Dans ce contexte, tant que la société restera largement incapable de distinguer avec certitude une image réelle d'une image synthétique, les deepfakes constitueront un outil particulièrement efficace pour générer des profits illicites. Pour de nombreux experts, la principale source de préoccupation ne réside donc pas seulement dans l'existence de cette technologie, mais dans

[6] WILLIAMS Rhiannon, *AI is already making online crimes easier. It could get much worse*, MIT Technology review, Février 2026, https://www.technologyreview.com/
[7] JOHNSON Derek, From fake nudes to fake quotes: AI deepfakes plagued Olympic athletes, Cyberscoop, Mars 2026, https://cyberscoop.com/
[8] NAFFI Nadia, *Deepfakes and the crisis of knowing,* UNESCO, *op., cit*.

son accessibilité et sa progression à venir[9]. La simplicité d'utilisation de ces outils abaisse considérablement les barrières à l'entrée : des individus sans compétences techniques avancées peuvent désormais produire des contenus falsifiés convaincants. Les deepfakes ne seraient ainsi plus uniquement un outil d'usurpation d'identité ou de fraude ponctuelle, mais pourraient devenir, à plus grande échelle, une véritable fabrique de cybercriminels.

Dans une société de plus en plus numérisée, cette évolution constitue une menace non négligeable. La facilité croissante avec laquelle il devient possible de perturber le fonctionnement d'une organisation (en imitant la voix d'un dirigeant, en falsifiant une instruction ou en manipulant des communications internes) soulève notamment des interrogations quant à la sécurité et à la résilience des infrastructures critiques d'un pays. Ces actions ne sont en effet pas uniquement motivées par un gain financier. Elles peuvent également servir des objectifs politiques, idéologiques ou sociétaux, en cherchant à perturber ou à désorganiser les infrastructures essentielles au fonctionnement d'un État.

Parallèlement, la digitalisation croissante de la société transforme également la sphère informationnelle. Les plateformes de diffusion de l'information se diversifient, tandis que l'essor des réseaux sociaux permet une circulation mondiale, rapide et massive de contenus issus de sources extrêmement variées. Dans cet environnement informationnel fragmenté, la fiabilité des sources n'est pas toujours garantie. Les deepfakes peuvent ainsi se diffuser au sein de ces nouveaux circuits d'information sans se distinguer clairement des contenus authentiques, remettant progressivement en cause la valeur probante de l'ensemble des preuves audiovisuelles.

Cette évolution s'inscrit dans un contexte plus large de suspicion croissante vis-à-vis de l'utilisation de l'IA. À mesure que les deepfakes deviennent plus réalistes, ils contribuent à installer une dynamique de doute généralisé. Les mécanismes traditionnels par lesquels les individus identifient la « vérité » et construisent la « connaissance » se trouvent ainsi fragilisés[10]. Si les deepfakes parviennent à remettre en question la fiabilité de sources auparavant considérées comme des « autorités informationnelles », telles que les institutions publiques ou certains médias, il devient possible d'envisager un phénomène de banalisation du doute. Dans ce contexte, toute information pourrait potentiellement être contestée.

Un tel processus pourrait progressivement placer les individus dans une forme d'isolation cognitive, où la perception de la vérité dépendrait de plus en plus du récit proposé par certaines sources qui ne sont pas nécessairement fiables. La société risquerait alors de se fragmenter davantage, les individus se regroupant selon leur degré de confiance envers certaines sources d'information et selon leur propre définition de ce qui constitue la vérité.

Cette banalisation du doute favorise également un phénomène connu sous le nom de « liar's dividend », ou « dividende du menteur ». Celui-ci désigne la situation dans laquelle une information pourtant authentique peut être discréditée en invoquant la possibilité (difficilement

[9] WILLIAMS Rhiannon, *AI is already making online crimes easier. It could get much worse*, MIT Technology review, *op., cit*.
[10] NAFFI Nadia, *Deepfakes and the crisis of knowing, op., cit*.

réfutable) qu'il s'agisse d'un contenu falsifié, et notamment d'un deepfake[11]. Même en temps de paix, un État peut être fragilisé si ces technologies parviennent à décrédibiliser à la fois l'autorité informationnelle des institutions et la valeur de la preuve photographique ou audiovisuelle.

L'érosion progressive de la confiance à la fois entre les individus et envers les institutions peut ainsi produire des effets profonds sur la stabilité économique, politique et sociale d'un pays. Elle peut également constituer une faille stratégique majeure dans sa capacité de défense. Dans ce contexte, les deepfakes ne représentent pas seulement un risque technologique ou criminel : ils peuvent aussi devenir un levier indirect de déstabilisation, susceptible d'offrir un avantage significatif à tout compétiteur stratégique potentiel.

Les deepfakes dans la conduite des hostilités : outil de guerre informationnelle et cognitive

Dans un contexte de conflit armé, les deepfakes peuvent devenir un outil stratégique susceptible de perturber la capacité d'un acteur étatique à maintenir la supériorité informationnelle et la cohérence décisionnelle dans un environnement stratégique. En effet, les mécanismes de déstabilisation observés en temps de paix peuvent aisément être transposés à une logique militaire. Là où l'usurpation de l'identité d'un dirigeant d'entreprise lors d'une visioconférence peut provoquer une fraude financière, un scénario comparable dans un contexte militaire pourrait viser un membre de la chaîne de commandement. La diffusion de faux ordres crédibles serait alors susceptible de perturber la prise de décision, voire de paralyser temporairement la capacité opérationnelle des forces. À plus long terme, ce type de manipulation pourrait également réduire la confiance des troupes envers l'ensemble des ordres reçus et provoquer des erreurs tactiques ou stratégiques.

L'effet stratégique recherché dans l'utilisation de deepfakes s'inscrit plus largement dans une stratégie de guerre cognitive ou psychologique. Ces deux stratégies visent à agir directement sur le moral des forces et sur leur volonté de combattre, en passant par le champ informationnel et par la capacité à influencer la perception du conflit par les différents acteurs. Certaines puissances ont d'ailleurs explicitement intégré ces dimensions dans leur doctrine. La Chine, par exemple, suit depuis 2003 une approche reposant sur trois formes de guerre : la guerre de l'opinion publique, la guerre psychologique et la guerre juridique[12].

L'efficacité de ces opérations est renforcée par ce que Clausewitz qualifiait de « brouillard de la guerre ». La rapidité des opérations, la fragmentation des sources d'information et la pression décisionnelle rendent en effet particulièrement difficile la vérification systématique des contenus auxquels sont exposés les décideurs, les militaires ou les populations civiles. Dans cet environnement informationnel saturé et incertain, un contenu falsifié mais plausible peut

[11] *Ibid*.
[12] NASU Hitoshi, *Deepfake technology in the age of information warfare,* Lieber Institute, Mars 2022, Articles of War, https://lieber.westpoint.edu/

circuler et produire des effets avant même que sa véracité ne soit contestée. Les deepfakes pourraient ainsi constituer un levier stratégique majeur, précisément parce qu'ils exploitent cette difficulté structurelle à distinguer rapidement le vrai du faux en situation de crise.

Leur impact potentiel s'étend à plusieurs niveaux stratégiques. Au niveau de la population civile, ils pourraient par exemple être utilisés pour discréditer les dirigeants politiques afin d'affaiblir sa confiance en les institutions et sa volonté de défense[13]. Au niveau militaire, ils pourraient viser la cohésion des forces en attaquant le respect envers la chaîne de commandement, par exemple à travers la diffusion de deepfakes montrant des chefs militaires se moquant de soldats blessés, et ainsi fragiliser la capacité des unités à opérer de manière coordonnée et structurée[14]. Ces technologies pourraient enfin également être utilisées pour fragiliser les alliances entre États en semant le doute quant à la fiabilité des partenaires, à l'image d'éventuels propos falsifiés d'un chef d'État à propos de pertes alliées, remettant en cause un engagement ou insinuant des accords avec des puissances adverses[15].

En pratique, la guerre russo-ukrainienne offre déjà certains exemples de l'utilisation de deepfakes dans une logique stratégique. Plusieurs tentatives d'usurpation de l'identité du président ukrainien ont été menées afin de diffuser de faux messages appelant les forces ukrainiennes à se rendre ou à déposer les armes. Des contenus manipulés ont également été utilisés dans la bataille informationnelle entourant le conflit, comme lorsque la Russie a diffusé un deepfake d'une attaque ukrainienne pour justifier l'invasion comme un acte de légitime défense[16]. Bien que ces opérations n'aient pas produit les effets escomptés, notamment en raison de la qualité encore limitée de certains deepfakes, elles illustrent néanmoins une volonté claire d'attaquer l'environnement informationnel de l'adversaire. À mesure que les technologies d'IA progressent, ce type d'opération pourrait devenir nettement plus crédible et potentiellement plus efficace.

L'enjeu des deepfakes dépasse ainsi la simple manipulation de l'opinion publique ou la déstabilisation politique d'un adversaire. Il touche également à la dynamique même des conflits contemporains. En créant de l'incertitude et en exploitant la difficulté à vérifier l'information dans des contextes de crise, les deepfakes peuvent contribuer à modifier la perception des événements et, par conséquent, influencer les décisions stratégiques. Dans cette perspective, ils constituent moins un simple outil de désinformation qu'un instrument susceptible d'influer sur l'équilibre informationnel et stratégique d'un conflit.

---

[13] BYMAN Daniel, GAO Chongyang, MESEROLE Chris, SUBRAHMANIAN V.S, *Deepfakes and international conflict,* op.cit.
[14] *Ibid*.
[15] *Ibid*.
[16] NASU Hitoshi, *Deepfake technology in the age of information warfare*, *op., cit*.

Les démocraties face à l'ambivalence stratégique des deepfakes

Les démocraties sont confrontées à une tension entre les avantages stratégiques que pourraient offrir les deepfakes et les risques qu'ils font peser sur la confiance publique et la stabilité politique. Dans la mesure où l'étendue réelle des capacités de l'IA générative et des conséquences associées aux deepfakes demeure encore largement incertaine, les États démocratiques doivent avancer avec prudence. Ils sont ainsi contraints d'envisager les usages potentiels de ces technologies tout en tenant compte de leurs effets néfastes sur les équilibres politiques et sociaux, afin d'en prévenir l'apparition.

Dans le cadre d'une utilisation nationale, le recours aux deepfakes par le gouvernement semble à première vue risqué. Les bénéfices stratégiques qu'ils pourraient offrir semblent relativement limités dans un régime démocratique, dans la mesure où leur efficacité repose principalement sur des stratégies de manipulation de l'information. À l'inverse, les risques associés sont clairement identifiables : remise en question des libertés fondamentales, érosion de la confiance envers les institutions publiques et fragilisation des processus démocratiques. En d'autres termes, l'usage de deepfakes par un gouvernement pourrait contribuer à décrédibiliser les principes mêmes sur lesquels repose un régime démocratique.

L'utilisation de deepfakes par l'administration Trump fournit un exemple récent de ces risques. Le gouvernement a diffusé à plusieurs reprises des images et vidéos synthétiques, en précisant parfois leur caractère artificiel, mais en omettant de le faire dans d'autres cas. Sur les réseaux sociaux de la Maison-Blanche et de certaines institutions fédérales, ces contenus ont progressivement été intégrés comme un outil de communication politique[17]. Les exemples vont de la modification des émotions d'une manifestante à la diffusion d'un deepfake montrant un joueur de hockey se moquant de l'adversaire vaincu[18]. Si le gouvernement américain a défendu ces productions comme de simples plaisanteries, cette intention ludique n'est pas unanimement perçue comme telle, notamment par les individus dont l'image est manipulée et qui peuvent en subir les conséquences en dehors des réseaux sociaux. Quelle qu'en soit l'intention, cette pratique participe à l'érosion progressive de la confiance de la population envers les représentants de l'État, et donc de leur crédibilité en tant qu'autorité informationnelle.

Le doute sur les intentions de l'État et les limites qu'il s'impose dans l'utilisation de tels outils se retrouve également dans le contexte des conflits armés. À cet égard, l'utilisation d'une fausse campagne de vaccination contre la polio par les États-Unis dans le cadre de l'opération visant à localiser Oussama ben Laden constitue un parallèle éclairant[19]. Cette opération avait suscité un débat important sur les limites éthiques des stratégies de dissimulation. Masquer une opération militaire derrière une initiative humanitaire peut profondément heurter l'opinion publique, qui en vient à s'interroger sur les moyens que l'État est prêt à employer pour obtenir un avantage stratégique : jusqu'où peut-il aller, au détriment de qui, et dans quels domaines ?

---

[17] JOHNSON Derek, From fake nudes to fake quotes: AI deepfakes plagued Olympic athletes, *op., cit*.
[18] *Ibid*.
[19] BYMAN Daniel, GAO Chongyang, MESEROLE Chris, SUBRAHMANIAN V.S, *Deepfakes and international conflict, op., cit*.

Transposée au cas des deepfakes, cette problématique renvoie à l'utilisation de la manipulation informationnelle comme instrument stratégique. Une telle approche suppose de trouver un équilibre délicat entre le gain potentiel et les risques politiques ou moraux associés. La perte progressive du soutien de l'opinion publique au sein des États impliqués dans les guerres au Vietnam, en Afghanistan ou en Irak témoigne de l'importance de cette dernière dans la continuité des opérations. Or, si l'utilisation faite des deepfakes est perçue comme trop cruelle, perfide ou génératrice de pertes civiles importantes, cela pourrait entraîner une désolidarisation de la population vis-à-vis du conflit.

La décision d'utiliser un deepfake dans un contexte stratégique implique donc de considérer ses conséquences concrètes sur le terrain. Un message synthétique appelant des combattants ennemis à déposer les armes pour faciliter une attaque pourrait également entraîner la présence de civils dans la zone d'affrontement, augmentant ainsi le risque de dommages collatéraux. L'utilisation de ces outils devrait donc systématiquement prendre en compte ses implications humanitaires au cas par cas, et être encadrée par des mécanismes robustes de surveillance et de responsabilité[20].

Leur emploi pourrait fragiliser la crédibilité normative des États qui se présentent comme défenseurs du droit international. Si les démocraties recouraient elles-mêmes à des pratiques de manipulation informationnelle sophistiquées, leur capacité à promouvoir une régulation internationale de ces technologies pourrait s'en trouver affaiblie, ce qui constituerait un désavantage diplomatique et politique majeur.

Tant que l'usage des deepfakes ne sera pas encadré par des normes et des processus clairs, les démocraties auraient probablement intérêt à limiter leur utilisation, que ce soit en temps de paix ou en temps de guerre. Cela ne signifie toutefois pas que leurs avantages stratégiques soient totalement inaccessibles. Il est possible d'en tirer parti sans compromettre la confiance de la population ni les principes démocratiques.

Le conflit russo-ukrainien illustre bien cette diversité d'approches. La Russie a choisi une approche offensive, qui repose sur une longue histoire de propagande, axée sur l'abondance des informations véhiculées pour déstabiliser la population du compétiteur stratégique[21]. L'Ukraine, quant à elle, a une approche plus défensive, et utilise les deepfakes essentiellement à des fins de sensibilisation, en mettant en évidence les trucages afin de démontrer les capacités de manipulation de l'IA et, par conséquent, de décrédibiliser les contenus de désinformation attribués à la Russie[22].

Dans les deux cas, l'objectif est de décrédibiliser le chef d'État adverse auprès de sa population et sur la scène internationale[23]. L'approche ukrainienne présente toutefois une dimension supplémentaire. En exposant ouvertement les mécanismes de manipulation, elle

[20] *Ibid*.
[21] BARTHELEMY Léo-Paul, *Guerre en Ukraine et deepfakes : un outil discursif, stratégique et systémique,* Revue du Centre de recherche Analyse du discours, 2025, Discours et communication stratégique en situation de crise, Hors-série, pp. 117-127, https://hal.science/
[22] *Ibid*.
[23] *Ibid*.

contribue à renforcer la résilience de la population face à la désinformation et facilite l’obtention d’un soutien international en démontrant une certaine transparence et une volonté de bonne foi. Cet exemple souligne que l’avantage stratégique associé aux deepfakes ne réside pas uniquement dans leur utilisation active, mais également dans la capacité d’un acteur à y résister et ainsi à neutraliser l’avantage qu’ils pourraient conférer à l’adversaire.

En somme, qu’ils soient mobilisés en temps de paix ou dans la conduite des hostilités, les deepfakes participent à l’émergence d’un climat d’incertitude informationnelle susceptible de fragiliser les États et d’influencer les rapports de puissance. Cette capacité à déstabiliser un adversaire en fait une ressource stratégique à part entière. La question de leur encadrement est dès lors déterminante : limiter leurs effets, en particulier dans leur dimension stratégique, suppose de développer des mécanismes de régulation, de détection et de gouvernance adaptés.

Encadrer et maîtriser les deepfakes : vers une gouvernance stratégique et juridique

Si les deepfakes apparaissent désormais comme un instrument potentiel de déstabilisation stratégique, la question centrale devient celle de leur maîtrise. La rapidité de leur diffusion dans les écosystèmes numériques contemporains, conjuguée à leur accessibilité croissante et à l’amélioration de leur qualité, complique leur détection et renforce leurs effets sur l’espace informationnel. Face à ces risques, les réponses s’organisent progressivement autour de plusieurs leviers complémentaires : le renforcement des capacités de résilience face aux dynamiques de propagation, l’adaptation des cadres juridiques à ces nouvelles formes de manipulation et l’émergence d’initiatives de coopération internationale visant à structurer une régulation collective des deepfakes.

Les vecteurs de propagation des deepfakes et les capacités de résilience

La menace que représentent les deepfakes implique, en premier lieu, la mise en place d’une gouvernance stratégique. Pour les démocraties qui s’abstiennent d’utiliser ces technologies à des fins de manipulation, cette gouvernance consiste principalement à comprendre les mécanismes de production et de diffusion de l’hypertrucage, à anticiper ses effets et à en limiter les impacts. Elle suppose également la coordination d’acteurs multiples (institutions publiques, entreprises technologiques, chercheurs et société civile) afin d’assurer une résilience à l’échelle nationale.

Dans cette optique, il est d’abord nécessaire d’identifier les principaux vecteurs de propagation des deepfakes. L’un des premiers points de tension réside dans la rapidité d’évolution de ces technologies, qui dépasse largement la capacité d’adaptation des cadres normatifs et, dans une certaine mesure, des dispositifs techniques de détection. Cette asymétrie confère un avantage significatif aux utilisateurs de deepfakes, qu’ils soient civils ou militaires, car ils évoluent dans un environnement juridique encore flou et dans un contexte technique souvent favorable à l’innovation offensive.

Les progrès réalisés dans la détection des deepfakes contribuent paradoxalement à améliorer les techniques de génération. Chaque avancée dans les outils de détection permet d’identifier les failles des modèles existants, lesquelles peuvent ensuite être corrigées à faible coût par les développeurs d’hypertrucages[24]. À l’inverse, les acteurs chargés de la défense (institutions publiques, entreprises technologiques ou équipes de recherche) sont souvent contraints par des ressources humaines et financières limitées, ainsi que par des procédures et standards qui ralentissent le développement de solutions opérationnelles. Cette dynamique d’adaptation permanente, comparable à une course technologique, implique que le modèle le plus performant est celui qui parvient à anticiper et à contourner les capacités de l’adversaire.

Les réseaux sociaux jouent un rôle central dans la diffusion de deepfakes. Ils ne constituent pas seulement de simples plateformes de circulation de l’information, mais également des outils de valorisation des deepfakes. Avec une compréhension relativement basique du fonctionnement de leurs algorithmes, un créateur peut accroître significativement la visibilité d’un deepfake, et par là même, l’illusion d’une légitimité de l’information. Plusieurs techniques sont couramment utilisées à cette fin. L’une d’elles consiste à mobiliser des bots, qui sont des programmes conçus pour imiter un utilisateur humain, parfois de manière complètement autonome. Les bots peuvent être des « abonnés fantômes » dont l’objectif est de donner l’illusion qu’un compte est fiable grâce à son nombre d’abonnés, contribuant ainsi à légitimer le contenu qui y est publié auprès du public cible. Les bots peuvent également agir comme des utilisateurs classiques, capables notamment de dialoguer avec d’autres utilisateurs ou de repartager massivement un contenu, donnant ainsi à l’algorithme l’impression qu’il suscite un fort intérêt et qu’il doit être mis en avant.

Une autre stratégie, utilisée à la fois par des bots et par des humains, repose sur le phénomène de « trolling ». Celui-ci consiste à publier des contenus ou des commentaires volontairement provocateurs ou controversés, incitant d’autres utilisateurs à réagir. L’algorithme interprète alors cet afflux de réactions comme un signe d’engagement élevé, ce qui conduit à une mise en avant accrue du contenu. Si celui-ci est un deepfake, le trolling permet de jouer avec les mécanismes algorithmiques du réseau social afin qu’il contribue lui-même à valoriser cet élément de désinformation et à le diffuser auprès d’un plus grand nombre de personnes.

[24] BYMAN Daniel, GAO Chongyang, MESEROLE Chris, SUBRAHMANIAN V.S, *Deepfakes and international conflict, op., cit*.

Les deepfakes bénéficient ainsi non seulement des réseaux sociaux comme vecteurs de diffusion, mais également de leur logique algorithmique comme outil de légitimation et d'amplification. Dans certains cas, la diffusion peut également être restreinte à des cercles spécifiques, ce qui modifie la dynamique de propagation et peut renforcer le potentiel de manipulation[25]. Ces restrictions de diffusion peuvent être volontaires, lorsqu'un acteur cherche à cibler un groupe précis d'utilisateurs partageant une caractéristique commune, par exemple un groupe linguistique, politique ou communautaire. Elles peuvent aussi être involontaires, résultant de décisions de modération ou de restrictions d'accès imposées à certaines plateformes.

Un exemple notable concerne les restrictions appliquées à Telegram depuis le déclenchement de la guerre en Ukraine. L'accès à plusieurs canaux russes a été limité dans certains pays, ce qui a eu pour conséquence de compliquer le travail des analystes chargés de suivre le cycle de vie des deepfakes et d'évaluer les capacités informationnelles russes[26]. Cette limitation réduit également les opportunités de démenti : lorsque les contenus circulent dans des espaces numériques restreints, les personnes qui y sont exposées ne disposent pas nécessairement des outils ou des ressources nécessaires pour vérifier l'authenticité des informations[27]. Cette situation met en évidence une tension stratégique majeure : faut-il restreindre l'accès à certaines plateformes afin de limiter la diffusion des deepfakes, au risque de renforcer leur impact au sein des publics qui y restent exposés ? Faut-il au contraire privilégier la liberté de circulation de l'information afin de maximiser les possibilités de vérification et de démenti ?

Face à ces dynamiques de propagation, différents moyens de détection ont été développés. Lors du salon international Eurosatory 2024, le COMCYBER a présenté une solution reposant sur l'intelligence artificielle pour identifier les deepfakes[28]. Cette solution propose de charger le contenu dans un logiciel qui déterminera par lui-même son authenticité et fournira ses critères d'analyses pour une vérification humaine[29]. L'IA permet ainsi d'automatiser différentes techniques d'analyse des médias, aussi effectuées manuellement par des analystes humains. Ces derniers recherchent notamment des erreurs dans le processus de génération des contenus, comme des incohérences sur les caractéristiques techniques de l'appareil supposé avoir capturé l'image ou la vidéo, telles que les signatures propres aux capteurs des caméras[30]. Ils examinent également les propriétés visuelles de l'image elle-même, comme les incohérences entre le visage et l'arrière-plan, ou encore les anomalies dans les points de repère faciaux, qui diffèrent souvent entre l'image authentique et le contenu synthétique[31].

---

[25] BARTHELEMY Léo-Paul, *Guerre en Ukraine et deepfakes : un outil discursif, stratégique et systémique,* op.cit.
[26] *Ibid*.
[27] *Ibid*.
[28] COMCYBER, *Désinformation : l'intelligence artificielle au service de la détection de deepfake*, Juin 2024, https://www.defense.gouv.fr/
[29] *Ibid*.
[30] BYMAN Daniel, GAO Chongyang, MESEROLE Chris, SUBRAHMANIAN V.S, *Deepfakes and international conflict, op., cit*.
[31] *Ibid*.

D’autres pratiques moins techniques de vérification ont émergé. L’une d’elles consiste, par exemple, à demander à l’interlocuteur d’effectuer un mouvement brusque ou inattendu afin de révéler d’éventuelles failles dans un deepfake généré en temps réel. Ces pratiques s’inscrivent dans une approche plus large désignée par le concept d’« agentivité épistémique », c’est-à-dire la capacité des individus à produire, transmettre ou utiliser le savoir. Dans un contexte où les informations qui constituent le savoir deviennent incertaines, l’agentivité épistémique consiste pour les individus à développer leurs propres critères d’évaluation de la crédibilité et de la véracité d’une information[32]. Dans les milieux militaires et du renseignement, ces logiques de vérification ne sont pas nouvelles : l’utilisation de messages codés, de mots de passe ou de procédures d’authentification constitue depuis longtemps un élément central de la sécurisation des communications. Des mécanismes similaires, quoique allégés, pourraient être mobilisés dans d’autres secteurs de la société afin de réduire la probabilité de succès des opérations de manipulation audiovisuelle.

À l’instar de la double authentification, la vérification systématique de l’identité d’un interlocuteur à travers des procédures établies pourrait, par exemple, être un prérequis pour toute action sensible. Ces procédures sont d’autant plus cruciales pour les organisations appartenant à des secteurs essentiels ou importants de la société, comme la santé, l’administration publique ou la finance, puisque l’effet déstabilisateur d’une attaque par deepfake serait plus conséquent. De plus, l’application obligatoire de ces processus permet de limiter l’exploitation du facteur humain par des acteurs malveillants, notamment dans le cadre d’attaques de type « fraude au président » qui reposent sur l’exploitation des relations hiérarchiques.

Plus largement, dans une perspective de prévention et de résilience nationale, l’éducation numérique de la population constitue un levier essentiel. La capacité d’une société à reconnaître et à résister aux campagnes de manipulation de l’information, parfois désignée sous le terme de « défense psychologique », renforce la solidité de l’État face aux opérations informationnelles en temps de paix comme en temps de guerre. À ce titre, l’approche ukrainienne de l’utilisation des deepfakes à des fins de sensibilisation offre à la population un avantage sur le long terme puisqu’elle leur donne les clés pour se protéger en amont plutôt que de se concentrer exclusivement sur la réparation des dommages causés. C’est une approche qui est d’ailleurs partagée par de nombreux États européens dans d’autres domaines du numérique, comme en témoignent les stratégies nationales de cybersécurité, qui incluent fréquemment de telles mesures éducatives.

Lorsqu’une telle résilience est développée dès le plus jeune âge, à travers l’apprentissage des mécanismes de désinformation, des menaces informationnelles et des bonnes pratiques de vérification, les effets potentiels des campagnes de déstabilisation en temps de paix peuvent être significativement réduits. Cette préparation présente un avantage supplémentaire en cas de conflit : les combattants eux-mêmes disposent alors de réflexes acquis en amont, limitant ainsi

[32] NAFFI Nadia, *Deepfakes and the crisis of knowing, op., cit.*

l'efficacité stratégique des opérations informationnelles. À terme, cette capacité collective de discernement pourrait également produire un effet dissuasif envers les adversaires potentiels.

L'encadrement juridique des deepfakes : avancées et limites

Si la gouvernance stratégique constitue un levier important pour se prémunir des effets des deepfakes, l'établissement d'un cadre juridique demeure tout aussi essentiel. À l'échelle nationale comme internationale, la mise en place de normes permet en effet de sanctionner les acteurs malveillants, de dissuader certains usages de ces technologies et d'établir des limites éthiques communes quant à leur emploi.

Au niveau national, plusieurs États ont déjà intégré les deepfakes dans leur dispositif juridique. Le Royaume-Uni, par exemple, impose aux fournisseurs de réseaux sociaux d'identifier et de gérer les risques associés aux contenus illégaux[33]. Depuis l'adoption du « Online Safety Act », la diffusion non consentie de deepfakes à caractère « intime » fait explicitement partie de ces contenus[34]. La France a également introduit cette problématique dans son droit pénal, qui interdit désormais la diffusion de deepfakes à des fins de désinformation ou lorsqu'elle concerne des contenus pornographiques réalisés sans le consentement de la personne représentée[35].

Aux États-Unis, l'évolution normative est plus ambivalente. Si l'administration Trump a démantelé un certain nombre d'organismes et de programmes destinés à lutter contre la manipulation de l'information[36], invoquant notamment la protection de la liberté d'expression, la tendance législative observée ces dernières années va néanmoins dans le sens d'un encadrement plus strict de ces technologies. Certains textes criminalisent ainsi la production de vidéos synthétiques destinées à faciliter des activités délictueuses[37]. Par ailleurs, le directeur du renseignement national est chargé d'assurer une veille sur l'utilisation des deepfakes par des gouvernements étrangers et d'évaluer leurs conséquences potentielles pour la sécurité nationale[38].

Malgré ces avancées, l'encadrement juridique des activités numériques ne peut se limiter à une approche strictement nationale. Par nature, les activités dans le cyberespace échappent largement à la notion traditionnelle de territoire étatique. Cette contrainte explique notamment pourquoi la majorité des législations nationales se concentrent sur la protection des individus, en particulier les enfants ou les victimes d'images à caractère sexuel, plutôt que sur les opérations informationnelles menées par des acteurs étrangers visant à déstabiliser un État.

---

[33] KUŹNICKA-BŁASZKOWSKA Dominika, KOSTYUK Nadiya, *Emerging need to regulate deepfakes in international law: the Russo–Ukrainian war as an example*, *op., cit*.
[34] *Ibid.*
[35] BARANES Yankel, *Deep Fake : Régulation française et européenne,* Laboratoire de cyberjustice, 2024, https://www.cyberjustice.ca/
[36] GALICIA Bailey, *In the fight against foreign information manipulation, the US can't afford to disarm, op., cit*.
[37] KUŹNICKA-BŁASZKOWSKA Dominika, KOSTYUK Nadiya, *Emerging need to regulate deepfakes in international law: the Russo–Ukrainian war as an example*, *op., cit*.
[38] *Ibid*.

Dans ce domaine, seuls le droit international ou les accords multilatéraux permettent véritablement d'encadrer le comportement d'acteurs étatiques.

En temps de paix, l'un des rares textes susceptibles de s'appliquer indirectement aux deepfakes est la Convention internationale concernant l'emploi de la radiodiffusion dans l'intérêt de la paix, signée en 1936[39]. Ce traité engage les 28 États signataires à « interdire et, le cas échéant, à faire cesser sans délai sur leurs territoires respectifs toute émission qui, au détriment de la bonne entente internationale, serait de nature à inciter les habitants d'un territoire quelconque à des actes contraires à l'ordre intérieur ou à la sécurité d'un territoire d'une Haute Partie contractante. »[40] Bien que ce texte ait le mérite d'exister, il reste toutefois largement insuffisant pour encadrer efficacement les pratiques contemporaines de désinformation dans le cyberespace, un théâtre de conflictualité qui n'existait pas au moment de sa rédaction.

En temps de guerre, la littérature scientifique s'intéresse davantage à la qualification juridique des deepfakes au regard du droit international humanitaire. Le débat porte notamment sur la possibilité de considérer leur utilisation comme un acte de perfidie, au sens de l'article 37 du Protocole additionnel I aux Conventions de Genève[41]. Selon cet article, la perfidie désigne « les actes faisant appel, avec l'intention de la tromper, à la bonne foi d'un adversaire pour lui faire croire qu'il a le droit de recevoir ou l'obligation d'accorder la protection prévue par les règles du droit international applicable dans les conflits armés. »[42] Dans cette perspective, un deepfake mettant en scène une troupe se rendant ou simulant une évacuation médicale afin d'inciter l'adversaire à cesser les hostilités avant de l'attaquer constituerait un acte de perfidie et, par conséquent, une violation du droit international humanitaire[43].

Ce même article précise cependant que les ruses de guerre ne sont pas interdites. Contrairement à la perfidie, ces pratiques « ont pour but d'induire un adversaire en erreur ou de lui faire commettre des imprudences, mais […] n'enfreignent aucune règle du droit international applicable dans les conflits armés. »[44] Parmi ces ruses figurent notamment « l'usage de camouflages, de leurres, d'opérations simulées et de faux renseignements ». Appliquée aux deepfakes, cette distinction devient particulièrement difficile à établir. La diffusion de fausses informations sur des mouvements de troupes pourrait par exemple relever d'une ruse de guerre[45], comparable à un leurre ou à un faux renseignement, s'il mène

[39] NASU Hitoshi, *Deepfake technology in the age of information warfare, op., cit*.
[40] Convention internationale concernant l'emploi de la radiodiffusion dans l'intérêt de la paix, Article 1, Genève, 23 septembre 1936.
[41] AKKUŞ Berkant, *Deepfakes and the Geneva Conventions: Does Deceptive AI Generated Misinformation Directed at an Enemy During Armed Conflict Violate International Humanitarian Law? A Critical Discussion*, School of Law, Department of Public International Law, Inonu University, 2025, https://doi.org/
[42] Article 37, Protocole additionnel 1 à la Convention de Genève, 1977.
[43] AKKUŞ Berkant, *Deepfakes and the Geneva Conventions: Does Deceptive AI Generated Misinformation Directed at an Enemy During Armed Conflict Violate International Humanitarian Law? A Critical Discussion*, *op., cit*.
[44] Article 37, Protocole additionnel 1 à la Convention de Genève, 1977
[45] AKKUŞ Berkant, *Deepfakes and the Geneva Conventions: Does Deceptive AI Generated Misinformation Directed at an Enemy During Armed Conflict Violate International Humanitarian Law? A Critical Discussion*, *op., cit*.

l’adversaire à abandonner une position stratégique. Si ces mouvements artificiels de troupes les plaçaient dans un hôpital ou une école contenant des civils, l’adversaire cesserait les hostilités en vertu de ses obligations internationales. En tirer un avantage reviendrait-il alors à tromper la bonne foi de l’adversaire ? Serait-ce donc un acte de perfidie ?

Cette incertitude juridique sur l’application du droit international humanitaire à l’utilisation des deepfakes en temps de guerre permet potentiellement à leurs créateurs de l’exploiter pour se soustraire à leurs obligations. Or, partir de ce postulat reviendrait à renforcer un climat de doute généralisé, selon lequel des images d’évacuation médicale de civils ou de troupes qui se rendent pourraient être fausses. Cette logique pourrait entraîner des violations des obligations internationales lors de situations réelles, affaiblissant progressivement les principes humanitaires fondamentaux. L’existence même des deepfakes, et l’incertitude qu’ils provoquent, possède ainsi la capacité de fragiliser l’application du droit international humanitaire.

Il apparaît dès lors nécessaire de réinterpréter le droit à la lumière des nouvelles capacités technologiques de manipulation, afin de définir des limites éthiques claires et des règles d’engagement auxquelles les acteurs ne pourraient se soustraire sans encourir de sanctions. Une telle évolution contribuerait notamment à renforcer l’effectivité du droit international dans les conflits contemporains et à dissuader l’usage des deepfakes en réduisant l’avantage stratégique lié à l’incertitude juridique.

Les transformations rapides du champ informationnel appellent également à une réponse normative à la fois rapide et anticipatrice. En effet, si les deepfakes n’ont pas produit les effets escomptés dans le conflit russo-ukrainien, les progrès récents observés dans le domaine suggèrent que leur qualité pourrait rapidement s’améliorer et révéler leur potentiel stratégique, auquel le droit demeure encore partiellement préparé. Compte tenu du caractère transnational de ces phénomènes, une réponse multilatérale apparaît de plus en plus nécessaire, fondée sur un consensus aussi large que possible entre les États.

Alliances interétatiques et standards communs : vers une régulation collective des technologies de manipulation

Malgré la place croissante de la manipulation de l’information dans les débats internationaux, l’émergence d’un cadre normatif global demeure limitée. Cette difficulté s’explique en partie par trois facteurs principaux : la valeur stratégique que représentent des outils comme les deepfakes, la tension persistante entre la lutte contre la manipulation de l’information et la protection de la liberté d’expression, et enfin l’absence de consensus sur la définition de l’information légitime et de l’influence jugée acceptable. Dans ce contexte, la formation de coalitions d’États partageant des principes communs pourrait favoriser une évolution progressive des pratiques et des normes encadrant ces technologies.

Sur le plan stratégique, il apparaît peu réaliste d’espérer convaincre les grandes puissances qui utilisent déjà les deepfakes à des fins d’influence ou de déstabilisation de renoncer à ces pratiques. En revanche, une initiative regroupant les États qui s’en abstiennent pourrait créer

un cadre de coopération alternatif. Une telle initiative pourrait offrir une plateforme d'échange sur les bonnes pratiques de résilience face aux manipulations informationnelles, rassemblant des institutions publiques, des entreprises technologiques ou des représentants de la société civile. Elle pourrait également encourager des programmes de coopération internationale dans le développement d'outils de détection des deepfakes.

Dans cette perspective, l'Union européenne apparaît particulièrement bien placée pour jouer un rôle moteur. Elle dispose déjà de mécanismes de partenariat entre secteurs public et privé ainsi que de cadres de coopération avec des États extérieurs à l'Union. Ces instruments pourraient servir de base à la mise en place d'une plateforme internationale consacrée à la sécurité informationnelle et à la détection des contenus synthétiques.

Une coopération élargie autour de bonnes pratiques, de standards de sécurité informationnelle et de technologies de détection permettrait également à chaque État d'améliorer sa propre résilience nationale, en fonction des menaces spécifiques auxquelles il est confronté. Une telle approche présente aussi l'avantage de contourner en partie la tension entre régulation de l'information et liberté d'expression. Plutôt que de reposer sur des mécanismes de censure étatique, ces pratiques privilégieraient l'autonomie des individus et des organisations dans l'identification des contenus manipulés. En développant des outils de détection des médias synthétiques et en diffusant les compétences nécessaires pour reconnaître les tentatives de manipulation, les risques pour la liberté d'expression peuvent ainsi être limités.

L'amélioration rapide de la qualité des deepfakes implique que ces mécanismes de détection et de résilience devront continuellement évoluer. Les États devront donc poursuivre leurs investissements dans les technologies de détection, tout en renforçant les capacités d'attribution des opérations de manipulation de l'information. La capacité à identifier l'auteur d'une opération et à corriger rapidement les fausses informations constitue en effet un élément central de la réponse stratégique[46].

L'élaboration de définitions communes constitue, enfin, une étape indispensable à la construction de normes internationales spécifiques aux deepfakes et à la manipulation de l'information. Ces définitions représentent le socle sur lequel pourraient être élaborés de futurs traités internationaux, seuls capables d'encadrer un phénomène qui dépasse par nature les frontières nationales. Si la complexité du domaine peut donner l'impression que la tâche est hors de portée, l'histoire du droit international montre pourtant le contraire. Les protocoles additionnels aux Conventions de Genève pour compléter le droit des conflits armés, l'interprétation du principe de participation directe aux hostilités par le CICR pour tenir compte du déplacement des affrontements en zone urbaine, ou la Convention de Budapest sur la cybercriminalité pour réguler, ensemble, un nouvel espace, sont autant d'exemples qui illustrent la capacité d'adaptation du droit international. Cette évolution progressive des normes témoigne de la capacité de la communauté internationale à s'organiser et à coopérer pour encadrer de nouveaux espaces de conflictualité.

[46] EH&A Consulting, *Intelligence artificielle et manipulation de l'information*, *op., cit*.

Dans cette perspective, la coopération internationale permet non seulement de définir des limites éthiques communes, mais aussi de mobiliser des ressources techniques et institutionnelles pour en assurer l’application. Elle constitue enfin le fondement nécessaire à l’élaboration de mécanismes de sanction en cas de violation de ces normes.

Conclusion

Les deepfakes s’inscrivent désormais dans les dynamiques contemporaines de compétition informationnelle et participent à une recomposition des rapports de puissance dans l’espace numérique. En exploitant l’incertitude et en brouillant les repères de véracité, ils offrent aux acteurs étatiques comme non étatiques un levier de déstabilisation susceptible d’affecter la confiance sociale, les processus décisionnels et la conduite des hostilités. À ce titre, ils tendent à s’imposer non seulement comme un outil de manipulation, mais comme une ressource stratégique à part entière.

Cette évolution soulève inévitablement la question de leur place dans les cadres normatifs existants, en particulier dans le droit des conflits armés. Si l’adaptation de ces cadres demeure un processus lent, tributaire de dynamiques de négociation et de recherche de consensus entre États, l’absence d’un régime juridique pleinement stabilisé ne signifie pas pour autant l’absence de réponses. Des mécanismes techniques, institutionnels et sociaux allant des outils de détection aux pratiques de vérification, en passant par le renforcement des capacités de résilience informationnelle, constituent déjà des moyens concrets de limiter les effets déstabilisateurs de ces technologies. Lorsqu’ils sont largement diffusés et partagés, ces instruments contribuent également à réduire l’intérêt stratégique des deepfakes. Un adversaire conscient de l’existence de mécanismes de détection, de procédures de vérification ou de standards communs peut être dissuadé d’y recourir, dans la mesure où l’efficacité de la manipulation repose précisément sur l’effet de surprise et l’incertitude qu’elle génère.

Si cette incertitude ne pourra sans doute jamais être totalement éliminée, au même titre que la désinformation qui la précède, il demeure possible d’en atténuer la portée stratégique en rendant son exploitation plus coûteuse, plus risquée ou plus facilement contestable. La réponse la plus pertinente semble résider dans une approche collective. Parce que les deepfakes cherchent précisément à fragmenter la connaissance partagée et à diviser les sociétés afin d’exploiter leurs failles informationnelles, leur neutralisation passe par le renforcement des coopérations entre États, institutions publiques, acteurs industriels et société civile. À la fois pragmatique et symbolique, cette dynamique de coopération contribue à restaurer des cadres communs de vérification et de confiance, condition essentielle pour limiter l’impact stratégique de l’incertitude dans l’espace informationnel contemporain.

Bibliographie

AKKUŞ Berkant, *Deepfakes and the Geneva Conventions: Does Deceptive AI Generated Misinformation Directed at an Enemy During Armed Conflict Violate International Humanitarian Law? A Critical Discussion*, School of Law, Department of Public International Law, Inonu University, 2025, https://doi.org/

BARANES Yankel, *Deep Fake : Régulation française et européenne,* Laboratoire de cyberjustice, 2024, https://www.cyberjustice.ca/

BARTHELEMY Léo-Paul, *Guerre en Ukraine et deepfakes : un outil discursif, stratégique et systémique,* Revue du Centre de recherche Analyse du discours, 2025, Discours et communication stratégique en situation de crise, Hors-série, pp.117-127, https://hal.science/

BYMAN Daniel, GAO Chongyang, MESEROLE Chris, SUBRAHMANIAN V.S, *Deepfakes and international conflict,* Foreign Policy at Brookings, Janvier 2023, https://www.brookings.edu/

COMCYBER, *Désinformation : l'intelligence artificielle au service de la détection de deepfake*, Juin 2024, https://www.defense.gouv.fr/

Digital Forensic Research Lab, *Russian War Report: Hacked news program and deepfake video spread false Zelenskyy claims,* Atlantic Council, Mars 2022, https://www.atlanticcouncil.org/

EH&A Consulting, *Intelligence artificielle et manipulation de l’information*, mars 2023, https://www.eha-consulting.com/

GALICIA Bailey, *In the fight against foreign information manipulation, the US can’t afford to disarm,* Atlantic Council, Juillet 2025, https://www.atlanticcouncil.org/

JOHNSON Derek, *From fake nudes to fake quotes: AI deepfakes plagued Olympic athletes*, Cyberscoop, Mars 2026, https://cyberscoop.com/

KUŹNICKA-BŁASZKOWSKA Dominika, KOSTYUK Nadiya, *Emerging need to regulate deepfakes in international law: the Russo–Ukrainian war as an example*, Journal of Cybersecurity, 2025, Volume 11, Issue 1, https://doi.org/10.1093/cybsec/tyaf008

Ministère des armées, *La désinformation, une arme de guerre,* Mars 2025, https://www.defense.gouv.fr/

NAFFI Nadia, *Deepfakes and the crisis of knowing,* UNESCO, Octobre 2025, https://www.unesco.org/

NASU Hitoshi, *Deepfake technology in the age of information warfare,* Lieber Institute, Mars 2022, Articles of War, https://lieber.westpoint.edu/

WILLIAMS Rhiannon, *AI is already making online crimes easier. It could get much worse*, MIT Technology review, Février 2026, https://www.technologyreview.com/

Institut
EGA

ISSN : 2739-3283

Institut d'études de géopolitique appliquée
66 avenue des Champs-Élysées, 75008 Paris
Courriel : secretariat@institut-ega.org
Site internet : www.institut-ega.org