ASSEMBLÉE NATIONALE

17ème législature

Impact des « deepfakes » sur la cybersécurité

Question écrite n° 1720

Texte de la question

Mme Virginie Duby-Muller appelle l'attention de Mme la secrétaire d'État auprès du ministre de l'enseignement supérieur et de la recherche, chargée de l'intelligence artificielle et du numérique, sur l'impact de l'utilisation des *deepfakes* sur la cybersécurité. Un *deepfake* (abréviation de *deep learning* et *fake*) est une vidéo manipulée à l'aide de techniques d'intelligence artificielle (IA), où le visage, les mouvements et la voix d'une personne sont superposés sur une autre, donnant l'illusion qu'elle réalise des actions ou prononce des paroles qu'elle n'a jamais faites ou dites en réalité. Si ces nouvelles intelligences artificielles peuvent représenter de véritables innovations, il convient toutefois de rappeler les dangers de celles-ci sur la cybersécurité en fonction de l'utilisation qui en est faite. Ces manipulations vidéo peuvent semer la désinformation politique, entraîner des fraudes financières, le vol d'identité, voire influencer les marchés financiers. En Allemagne, le Gouvernement a exprimé une vive inquiétude face aux *deepfakes*, allant jusqu'à lancer une campagne de sensibilisation pour alerter les parents sur les dangers de ces technologies. Le 10 avril 2024 a été voté le projet de loi dit « SREN » pour mieux réguler l'espace numérique et protéger les internautes, notamment les plus jeunes, ainsi que les entreprises. Cette loi prend bien en compte les dangers liés à publication en ligne d'hypertrucages ou *deepfake* qui seront mieux réprimées. Néanmoins, selon un sondage IFOP, seulement un tiers des citoyens français estiment avoir la capacité de repérer un *deepfake* et à peine 6 % en sont totalement sûrs, illustrant ainsi le niveau élevé d'incertitude qui prévaut chez eux. Les jeunes et les hommes se montrent plus confiants : 55 % des 18-24 ans pensent pouvoir le faire, contre 28 % des plus de 35 ans, tandis que 40 % des hommes le croient possible, comparé à 28 % des femmes (source : Les Français et les jeunes face aux *deepfakes* - sondage IFOP). Les *deepfakes* utilisant le *machine learning* pour s'améliorer, ils risquent d'être de moins en moins détectables par une majorité de personnes. Récemment, les visages influenceurs et de personnalités publiques ont été utilisés pour la promotion de casinos en ligne, applications mobiles frauduleuses et cryptomonnaies douteuses. Les arnaqueurs usurpent l'identité de ces personnalités publiques, car elles ont une notoriété forte auprès du grand public. Des fausses vidéos sont ainsi propagées sur les réseaux sociaux, notamment TikTok, qui cible majoritairement les jeunes et se retrouvent à la merci de ces arnaques. Ainsi, elle lui demande ce que le Gouvernement compte mettre en place pour renforcer la sensibilisation et l'encadrement des *deepfakes*, cette problématique sérieuse étant de l'ordre de l'atteinte au droit à l'image, de l'usurpation d'identité et de l'escroquerie.

Texte de la réponse

Le Gouvernement est pleinement engagé dans la lutte contre les hyper-trucages (deep fakes), pour lesquels les sanctions ont été significativement renforcées dans le cadre de la loi du 21 mai 2024 visant à sécuriser et réguler l'espace numérique (loi « SREN »). La sanction des délits d'atteintes à la représentation de la personne via le recours aux hyper-trucages a été notablement renforcée par une série d'articles de la loi. L'article 15 modifie l'article 226-8 du code pénal, sanctionnant le partage de montages falsifiés de représentations sonores ou visuelles d'une personne, réalisés sans son consentement, afin d'inclure les contenus générés par un traitement algorithmique (deep fakes). Le partage sur les réseaux sociaux constitue une circonstance

aggravante, pouvant être puni de deux ans d'emprisonnement et de 45 000 euros d'amende. L'article 16 de la loi SREN dispose également que le juge peut prononcer, dans les cas de cyberharcèlement (notamment via des deep fakes), une peine de bannissement numérique, interdisant à l'auteur du délit l'utilisation des réseaux sociaux pendant une période maximale de six mois (portée à un an en cas de récidive). Le juge peut également prononcer, à titre de peine complémentaire ou alternative, une obligation de stage de sensibilisation au respect des personnes dans l'espace numérique et à la prévention des infractions commises en ligne. L'article 21 crée l'article 226-8-1 dans le même code, interdisant les montages et deep fakes à caractère sexuel, punis de deux ans d'emprisonnement et de 60 000 euros d'amende. Le partage de tels montages sur les services de communication au public en ligne constitue une circonstance aggravante, punie de trois ans d'emprisonnement et de 75 000 euros d'amende. La loi SREN, en ses articles 7 et 8, renforce les mesures de sensibilisation aux risques des outils et contenus générés par intelligence artificielle, incluant les deep fakes, à destination des élèves et de leurs représentants légaux, des étudiants de l'enseignement supérieur et des membres du personnel enseignant et d'éducation. Ces dispositions visent à sensibiliser ce public aux risques d'atteinte à la personne, de manipulation commerciale, d'escroquerie, de harcèlement ou de violences sexistes et sexuelles dans l'espace numérique. Enfin, il convient de rappeler qu'en vertu du règlement européen du 13 juin 2024 sur l'intelligence artificielle, les déployeurs d'un système d'intelligence artificielle qui génère ou manipule des images ou des contenus audio ou vidéo constituant un hyper-trucage, sont tenus à une obligation de transparence et d'étiquetage (là-dessus, se référer à l'article 50 relatif à l'obligation de mentionner qu'un contenu est généré par intelligence artificielle). Au-delà de dispositions existantes pour encadrer ces pratiques, une sensibilisation plus gobale de la population aux enjeux de l'intelligence artificielle apparait effectivement indispensable. A cet égard le gouvernement à annoncé dans le cadre du Sommet pour l'action sur l'IA qui s'est tenu à Paris qu'à compter de la rentrée 2025, la plateforme PIx proposera un parcours de formation à l'IA qui sera obligatoire pour tous les élèves de quatrième et de seconde. De même, le déploiement sur l'ensemble du territoire du dispositif Café IA porté par le Conseil national du Numérique (CNNum) et par l'Agence Nationale de la Cohésion des Territoires (ANCT) avec l'appui du réseau des conseillers numériques présents sur tout le territoire, permettra de sensibiliser 2 millions de Français à ces technologies d'ici 2027.

Données clés

Auteur : Mme Virginie Duby-Muller
Circonscription : Haute-Savoie (4e circonscription) - Droite Républicaine
Type de question : Question écrite
Numéro de la question : 1720
Rubrique : Numérique
Ministère interrogé : Intelligence artificielle et numérique
Ministère attributaire : Intelligence artificielle et numérique

Date(s) clée(s)

Question publiée au JO le : 5 novembre 2024, page 5809
Réponse publiée au JO le : 3 juin 2025, page 4582