



Vers une réglementation spécifique des deepfakes ?

Le *deepfake*, ou hypertrucage en français, peut être défini comme « une technique de synthèse mono-ou multimédia reposant sur l'intelligence artificielle. Elle peut servir à superposer ou fusionner des images, des fichiers vidéo ou audio existants sur d'autres images ou fichiers vidéo ou audio, à créer un contenu artificiel sur une personne cible à partir du comportement d'une personne source, ou même à créer artificiellement des contenus ressemblants à partir de commandes textuelles »¹.

Désinformation, constitution de fausses preuves, *deepfakes* pornographiques, etc., le champ des usages et risques associés aux *deepfakes* est particulièrement vaste et relève de plusieurs régimes juridiques. En particulier, la protection de la personne ciblée par un *deepfake* comporte un volet civil (I) et un volet pénal (II), ce dernier ayant vocation à être renforcé par le projet de loi visant à sécuriser et réguler l'espace numérique (SREN)² en cours d'examen à l'Assemblée Nationale.

I. Une protection civile protéiforme

La personne victime d'un *deepfake* a la possibilité de mobiliser plusieurs catégories de droits au plan civil.

- **Les droits de la personnalité**

La jurisprudence a développé, sur le fondement de l'article 9 du Code civil consacrant le droit à la vie privée, un droit à l'image et à la voix. Toute personne peut ainsi autoriser ou interdire l'utilisation de son image ou de sa voix par autrui. Un *deepfake*, même peu réaliste, est susceptible de porter atteinte au droit à l'image et/ou à la voix d'une personne puisque de simples traits et/ou des intonations peuvent être protégés dès lors qu'ils permettent d'identifier la personne. Une personne victime d'un *deepfake*, qu'elle n'aurait pas autorisé et qui serait identifiable, même par un nombre limité de personnes, pourrait donc agir en responsabilité civile sur le fondement de l'article 9 du code civil et obtenir réparation.

¹ Amendement n°127 du 3 juillet 2023 présenté par le Gouvernement au Projet de loi « Sécuriser et réguler l'espace numérique » : https://www.senat.fr/amendements/2022-2023/778/Amdt_127.html

² Projet de loi n°1514, adopté par le Sénat visant à sécuriser et réguler l'espace numérique, enregistré à la Présidence de l'Assemblée nationale le 7 juillet 2023.

- **Le droit de la propriété intellectuelle**

Les titulaires de droits de propriété intellectuelle tels que des droits d'auteur ou des droits voisins, sont susceptibles de se défendre en cas de *deepfakes* réalisés en violation de leurs droits. Les concepteurs d'un *deepfake* se doivent d'obtenir l'accord des titulaires de droits s'ils utilisent des œuvres protégées par le droit d'auteur ou les droits voisins. A titre d'illustration, si le concepteur d'un *deepfake* superpose une image avec une musique dont il ne possède pas les droits, il réalise alors un acte de contrefaçon. En outre, s'il y a une modification substantielle de la musique, il peut être porté atteinte au respect de l'œuvre ou bien encore à l'interprétation de l'artiste ce qui occasionne en sus une atteinte au droit moral des titulaires. Toutefois, le Code de la propriété intellectuelle prévoit des exceptions³ dont les concepteurs de *deepfakes* pourraient se prévaloir afin de se dédouaner de toute atteinte. Il pourrait s'agir de l'exception de parodie ou de courte citation. Ces exceptions sont néanmoins soumises à des conditions strictes⁴.

- **Le droit de la protection des données personnelles**

L'image ainsi que la voix d'une personne constituent des données à caractère personnel au sens du Règlement Général sur la Protection des Données (RGPD). A ce titre, ces données bénéficient d'une protection et leur traitement doit reposer sur une base légale telle que le consentement de la personne ou bien l'intérêt légitime du responsable de traitement qui doit être justifié par la nécessité du traitement.

II. Une protection pénale potentiellement accrue

- **Cadre pénal actuel**

Les articles 226-1 et suivants du Code pénal sanctionnent d'un an d'emprisonnement et de 45 000 euros d'amende la captation et la diffusion d'images d'un individu sans son consentement en ce que cette capture ou cette diffusion porte atteinte à l'intimité de sa vie privée.

L'article 226-8 du Code pénal punit quant à lui d'un an d'emprisonnement et de 15 000 euros d'amende « le fait de publier, par quelque voie que ce soit, le montage réalisé avec les paroles ou l'image d'une personne sans son consentement, s'il n'apparaît pas à l'évidence qu'il s'agit d'un montage ou s'il n'en est pas expressément fait mention ».

- **Un renforcement des dispositions pénales**

Le 5 juillet dernier, le Sénat a adopté à l'unanimité le projet de loi SREN. Ce dernier a été transmis à l'Assemblée Nationale qui l'examine depuis le 19 septembre dernier.

³ Article L122-5 du Code de la propriété intellectuelle.

⁴ L'exception de parodie est limitée aux « lois du genre ». La Cour de cassation a rappelé récemment que « pour être qualifiée de parodie, l'œuvre seconde doit revêtir un caractère humoristique et éviter tout risque de confusion avec l'œuvre parodiée » et ne pas porter « une atteinte disproportionnée aux intérêts légitimes de l'auteur et de son ayant droit » (Cour de cassation, civile, Chambre civile 1, 22 mai 2019, 18-12.718).

L'exception de courte citation doit quant à elle être justifiée par le caractère critique, polémique, pédagogique, scientifique ou d'information de l'œuvre à laquelle elles sont incorporées et est permise sous réserve que soient indiqués clairement le nom de l'auteur et la source.

Les apports de ce projet sont multiples et donnent lieu à de vives discussions. Parmi eux, deux amendements déposés par le gouvernement concernent spécifiquement les *deepfakes*.

Le premier amendement propose d'ajouter deux alinéas à l'article 226-8 du Code pénal⁵ :

- Le premier de ces aliénas étend la notion de montage aux *deepfakes*, que ceux-ci soient visuels ou sonores : « Est assimilé à l'infraction mentionnée au présent alinéa et puni des mêmes peines le fait de publier, par quelque voie que ce soit, un contenu visuel ou sonore généré par un traitement algorithmique et reproduisant l'image ou les paroles d'une personne, sans son consentement, s'il n'apparaît pas à l'évidence qu'il s'agit d'un contenu généré algorithmiquement ou s'il n'en est pas expressément fait mention ». L'introduction de cet alinéa permettrait de confirmer l'intuition d'une partie de la doctrine qui pressentait déjà l'application de l'article 226-8 aux *deepfakes*⁶. Le critère d'« évidence » n'est cependant pas précisé et pourrait donner lieu à des appréciations divergentes. La recommandation pourrait donc être « de préciser systématiquement cette utilisation de l'outil informatique »⁷, même si cela soulève la question de l'opportunité et des modalités d'affichages de l'information en ligne.
- Le second alinéa a pour but de répondre à l'enjeu de « viralité » propre aux réseaux sociaux en introduisant une circonstance aggravante lorsque le *deepfake* a été publié sur un service de communication au public en ligne. Les peines peuvent alors être portées à deux ans d'emprisonnement et 45.000 euros d'amende.

Le gouvernement a également proposé d'enrichir la législation relative aux *deepfakes* en introduisant un nouvel article 226-8-1 au Code pénal qui condamne la diffusion sans consentement du montage ou de l'hypertrucage (*deepfake*) d'une personne présentant un caractère sexuel : « Est puni de deux ans d'emprisonnement et de 60 000 euros d'amende le fait de publier, sans son consentement, par quelque voie que ce soit, le montage réalisé avec les paroles ou l'image d'une personne, et présentant un caractère sexuel⁸. Est assimilé à l'infraction mentionnée au présent alinéa et puni des mêmes peines le fait de publier, par quelque voie que ce soit, un contenu visuel ou sonore généré par un traitement algorithmique et reproduisant l'image ou les paroles d'une personne, sans son consentement, et présentant un caractère sexuel ». Un sous-amendement⁹ précisant l'article 226-8-1 prévoit également une circonstance aggravante lorsque la publication du *deepfake* a été réalisée au moyen d'un service de communication au public en ligne et porte les peines à trois ans d'emprisonnement et 75.000 euros d'amende. Cet article 226-8-1 répondrait à un objectif de protection de l'ordre public et de lutte contre le sexisme, puisque 99% des *deepfakes* pornographiques concernent des femmes.

⁵ Amendement n°127.

⁶ V. notamment : C. Langlais-Fontaine, « Démêler le vrai du faux : étude de la capacité du droit actuel à lutter contre les deepfakes », La Revue des droits de l'homme [En ligne], 18 | 2020, mis en ligne le 12 juin 2020.

⁷ J. Groffe-Charrier, Maître de conférences HDR, membre du CERDI ; Vers un encadrement légal des deepfakes ; Dalloz actualité, 12 juillet 2023.

⁸ Amendement n°129 rect. bis

⁹ Amendement n°129 rect. Bis

A noter qu'à ce jour, il n'existe pas de cadre juridique européen unifié applicable aux *deepfakes*, même si différents textes épars les mentionnent directement – à commencer par l'IA Act¹⁰. Au vu de l'engouement suscité par le phénomène des *deepfakes* mais surtout des risques associés, ce mouvement devrait s'amplifier et peut-être un jour se matérialiser par une réglementation européenne spécifique, comme le recommande un récent rapport du Parlement européen¹¹.

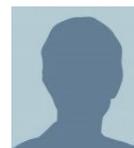
Auteurs



Charles Bouffier
Associé
cbouffier@racine.eu



Victoire Danès
Avocate
vdanes@racine.eu



Armin Rahmani
Juriste
arahmani@racine.eu

¹⁰ Proposition de Règlement du Parlement Européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union : l'article 52 énonce que « Les utilisateurs d'un système d'IA qui génère ou manipule du texte, des contenus audio ou visuels pouvant être perçus à tort comme authentiques ou véridiques et représentant des personnes semblant tenir des propos qu'elles n'ont pas tenus ou commettre des actes qu'elles n'ont pas commis («hypertrucage») précisent, de manière appropriée, claire, visible et en temps utile que les contenus ont été générés ou manipulés artificiellement et, si possible, indiquent le nom de la personne physique ou morale qui les a générés ou manipulés. À cet égard, les contenus sont étiquetés d'une manière qui signale clairement leur caractère non authentique aux personnes auxquelles ils sont destinés. Aux fins de l'étiquetage des contenus, les utilisateurs prennent en considération l'état de la technique généralement reconnu et les normes harmonisées et spécifications pertinentes. »

¹¹ European Parliamentary Research Service; Tackling deepfakes in European policy; July 2021, p.58-86.