
Démêler le vrai du faux : étude de la capacité du droit actuel à lutter contre les *deepfakes*

Claire Langlais-Fontaine



Édition électronique

URL : <http://journals.openedition.org/revdh/9747>

DOI : [10.4000/revdh.9747](https://doi.org/10.4000/revdh.9747)

ISSN : 2264-119X

Éditeur

Centre de recherches et d'études sur les droits fondamentaux

Référence électronique

Claire 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, consulté le 11 février 2021. URL : <http://journals.openedition.org/revdh/9747> ; DOI : <https://doi.org/10.4000/revdh.9747>

Ce document a été généré automatiquement le 11 février 2021.

Tous droits réservés

Démêler le vrai du faux : étude de la capacité du droit actuel à lutter contre les *deepfakes*

Claire Langlais-Fontaine

- ¹ Aujourd'hui, le célèbre adage, attribué à Saint Thomas, « je ne crois que ce que je vois » est mis à mal par l'usage de nouvelles technologies aux fins de brouiller les limites entre le vrai et le faux. L'abandon du candidat aux élections municipales du parti présidentiel à la mairie de Paris, Benjamin Griveaux, le 14 février 2020¹ a ainsi mis sur le devant de la scène médiatique et politique l'utilisation d'outils digitaux et numériques aux fins de déstabilisation politique. Si les vidéos diffusées du candidat ne sont finalement pas des *deepfakes*², contrairement à ce qu'avaient d'abord supposé certains élus comme Marine Le Pen, cet événement interroge cependant sur les conséquences de l'utilisation de ces outils et sur la capacité du droit français à poursuivre les auteurs de la diffusion de telles images.
- ² Le terme *deepfakes* – qui a fait son apparition en 2017³ – renvoie à l'utilisation de logiciels d'intelligence artificielle (IA) en vue de truquer des contenus audios et vidéos. La technique permet de créer de fausses vidéos – les *deepfakes* ou « permutation intelligente de visages », dites également vidéos truquées ou vidéos hypertruquées – en manipulant des images et des sons à l'aide de la technique de « l'apprentissage profond » ou *deeplearning*. L'objectif de telles vidéos est alors de pouvoir faire faire ou faire dire n'importe quoi à n'importe qui. Cette technique est relativement facile d'accès – certains logiciels de *deeplearning* dédiés aux vidéos hypertruquées étant présents sur le web en *open source* c'est-à-dire libres de droits – et relativement difficile à contrer. En effet, les logiciels de création de telles vidéos fonctionnent grâce à un mécanisme propre à berner les algorithmes de détection et reposant sur la mise en concurrence de deux algorithmes⁴ : le premier recopie une vidéo à l'identique une multitude de fois en y important un visage externe, le second détecte la qualité des vidéos créées par le premier algorithme afin d'écarter les moins crédibles. La technique

des *deepfakes* tend ainsi à créer des vidéos hypertruquées très réalistes tout en se prémunissant contre la détection rapide de faux.

- 3 Des vidéos hypertruquées de Barack Obama, insultant son successeur Donald Trump en avril 2018, ou encore de Mark Zuckerberg, célèbre dirigeant de l'entreprise Facebook, avouant manipuler ses utilisateurs et leurs données en juin 2019, ont par exemple fait la une des journaux⁵. Les *deepfakes* ne sont toutefois pas restreintes au monde politique ou économique. Leur première apparition s'est en réalité faite sur les grandes plateformes web de vidéos pornographiques dès 2017. Le *deepfake* permet alors d'importer des visages de célébrités (par exemple d'actrices ou d'acteurs connus) sur les corps des personnes qui tournent la vidéo pornographique. De nombreuses actrices hollywoodiennes apparaissent désormais dans des vidéos pornographiques sur Internet sans jamais en avoir tourné et, surtout, sans avoir donné leur consentement à la diffusion de telles images.
- 4 Le phénomène est assez important. Selon le rapport 2019 de *Deeptrace Lab*⁶, une entreprise hollandaise qui étudie les *deepfakes*, 14 000 vidéos hypertruquées étaient mises en ligne en 2019, ce qui représente une hausse de 84% par rapport à l'année précédente. Plus de 850 personnes ont été identifiées comme victimes de ces vidéos hypertruquées, et 96% des vidéos *deepfake* sont des vidéos pornographiques. Le phénomène a pris une telle ampleur que des plateformes spécialement dédiées aux vidéos *deepfake* ont été créées, démontrant l'existence d'un marché pour les sites Internet hébergeant ce type de vidéos, et invitant à réfléchir à l'encadrement de telles pratiques.
- 5 Si le détournement de vidéos à des fins commerciales ou politiques n'est pas encore très pratiqué en France, les vidéos pornographiques truquées sont en revanche une réalité. Il convient alors de s'interroger sur la capacité des normes françaises actuelles à appréhender ces deux phénomènes, d'autant que, pour l'instant, le terme n'apparaît que dans deux rapports d'information parlementaires⁷ et dans certains comptes-rendus en commission. Le législateur ne semble donc pas s'être pleinement saisi de la question. Néanmoins, une députée a interpellé le gouvernement le 5 février 2019 afin que l'exécutif propose un plan pour « lutter efficacement » contre ces vidéos⁸, jugeant l'état du droit positif actuel insuffisant.
- 6 Quelles garanties existent déjà au sein de l'arsenal législatif français pour appréhender et lutter contre les *deepfakes* ? Le droit français actuel comporte-t-il des carences face à cette pratique ?
- 7 Les deux domaines principaux d'utilisation de vidéos hypertruquées doivent être distingués aux fins de la présente étude, du fait des conséquences différentes qu'elles emportent. Nous verrons donc dans un premier temps l'insuffisance du droit pénal français pour sanctionner l'utilisation de *deepfakes* pornographiques (I) et, dans un second temps, le renforcement du cadre législatif aux fins de lutter contre la déstabilisation économique ou politique que les *deepfakes* peuvent engendrer (II).

I. La lutte contre les *deepfakes* pornographiques : les insuffisances du cadre législatif français actuel

- 8 L'encadrement juridique de cette pratique est un enjeu considérable puisque les vidéos hypertruquées pornographiques représentent 96% des *deepfakes*, et que 99% de ces

vidéos hypertruquées visent des femmes, dont 81% des personnalités publiques (actrices, musiciennes, etc.). La part de ces vidéos visant des femmes inconnues du grand public correspond principalement à une logique de *revenge porn* ou vengeance pornographique. Cette pratique consiste à diffuser publiquement un contenu sexuellement explicite sans le consentement de la personne apparaissant dans ce contenu, dans un but de vengeance. Bien que le législateur se soit précisément intéressé à cette question, la pénalisation récente de cette pratique ne permet toutefois pas de pénaliser en totalité les vidéos hypertruquées, ce qui constitue une première limite du droit positif actuel face aux *deepfakes* (A). Reste alors la perspective d'une appréhension de ce phénomène par l'intermédiaire des infractions au droit à l'image (B) ou encore par l'intermédiaire des infractions de presse (C), tant pour les femmes inconnues du grand public que pour les personnalités publiques. Mais là encore, la législation en vigueur est imparfaite pour lutter contre les *deepfakes* pornographiques.

A L'imparfaite appréhension des *deepfakes* par les infractions pénales relatives au *revenge porn*

- 9 La pénalisation du *revenge porn* peut, de prime abord, apparaître comme un bon moyen de lutter contre les vidéos hypertruquées à caractère sexuel ou pornographique par la poursuite au pénal des auteurs de ces vidéos ou des personnes impliquées dans leur diffusion. Cette voie de droit permet alors de poursuivre les auteurs des vidéos, images ou paroles truquées.
- 10 La législation en matière de pénalisation des faits de *revenge porn* a subi des évolutions récentes pour permettre une meilleure appréhension du phénomène dans la loi pénale. Avant 2016, seuls les articles 226-1 et 226-2 du Code pénal étaient applicables à ce type de faits. Ils punissent d'un an d'emprisonnement et de 45 000€ d'amende « le fait, au moyen d'un procédé quelconque, volontairement, de porter atteinte à l'intimité de la vie privée d'autrui : 1° En captant, enregistrant ou transmettant, sans le consentement de leur auteur, des paroles prononcées à titre privé ou confidentiel ; 2° En fixant, enregistrant ou transmettant, sans le consentement de celle-ci, l'image d'une personne se trouvant dans un lieu privé »⁹. Ces articles, qui concernent les atteintes à la vie privée¹⁰, ont été interprétés de manière restrictive par la Cour de cassation¹¹. Pour la chambre criminelle, la notion de consentement contenue dans ces articles faisait obstacle à l'application des dispositions à un certain nombre de cas de *revenge porn*, notamment lorsque les images ou enregistrements utilisés par les auteurs de ce délit avaient initialement été captées ou enregistrées avec le consentement de la victime¹². Seule la diffusion était alors dénuée de consentement. Cette interprétation restrictive est venue justifier l'introduction d'un nouvel article 226-2-1 par la loi Lemaire¹³ en 2016.
- 11 Cet article dispose que « [lorsque] les délits prévus aux articles 226-1 et 226-2 portent sur des paroles ou des images présentant un caractère sexuel prises dans un lieu public ou privé, les peines sont portées à deux ans d'emprisonnement et à 60 000€ d'amende. Est puni des mêmes peines le fait, en l'absence d'accord de la personne pour la diffusion, de porter à la connaissance du public ou d'un tiers tout enregistrement ou tout document portant sur des paroles ou des images présentant un caractère sexuel, obtenu, avec le consentement exprès ou présumé de la personne ou par elle-même, à l'aide de l'un des actes prévus à l'article 226-1 »¹⁴. Cette nouvelle disposition corrige *a priori*¹⁵ certaines

des failles des articles 226-1 et 226-2 dans la pénalisation du *revenge porn* : d'une part, les images ou paroles enregistrées peuvent l'avoir été dans un lieu public, et pas uniquement dans un lieu privé ; d'autre part, une action est désormais possible y compris lorsque l'image a été captée avec le consentement de la victime¹⁶. En revanche, cette législation ne s'applique que partiellement aux vidéos hypertruquées diffusées sur des sites pornographiques. En effet, le code pénal ne traite pas de la création d'images ou de paroles, ou de la transformation d'images ou de paroles. On peut donc relever une inadéquation des dispositions légales actuelles visant les faits de *revenge porn*, d'autant que le droit pénal est d'interprétation stricte. Cette inadéquation constitue une limite manifeste à la lutte contre les vidéos et images hypertruquées par ce biais.

- 12 En résumé, les articles 226-1 et 226-2 du Code pénal, visant expressément les atteintes à la vie privée et utilisées pour lutter contre le *revenge porn* avant 2016, ne prenaient pas en compte les caractères sexuel et genré de ce type d'images visant dans leur très grande majorité des femmes¹⁷. Quant au nouvel article 226-2-1, visant expressément les vidéos à caractère sexuel et leur diffusion, il ne prend pas en compte la possibilité d'images trafiquées, modifiées, transformées ou créées de toute pièce.
- 13 En comparaison, l'État de Virginie, aux États-Unis, a adopté un amendement en 2019 visant explicitement les faits de vengeance pornographique *par l'utilisation de deepfakes*. L'article 5, §18.2-386.2, du Code de Virginie dispose désormais que « [quiconque] ayant l'intention de contraindre, harceler ou intimider diffuse malicieusement ou vend des images vidéo pornographiques ou des images fixes créées par tout moyen qui représentent une autre personne totalement nue ou en état de se déshabiller de manière à exposer les organes génitaux, la région pubienne, les fesses ou les seins d'une femme, lorsque cette personne sait ou a des raisons de savoir qu'elle n'est pas autorisée à diffuser ou vendre une telle image vidéo ou fixe est coupable d'un délit de Classe 1 »¹⁸. La précision d'une telle disposition permet davantage de pénaliser le trucage de vidéos à des fins de *revenge porn* et prend spécifiquement en compte le caractère genré de cette pratique par la mention explicite de parties des corps féminins.
- 14 Enfin, une dernière difficulté reste en suspens et interroge la capacité des instruments juridiques pour lutter contre ces images : celle de l'identification de l'auteur de l'infraction, qu'il s'agisse de l'auteur de la vidéo hypertruquée ou de sa diffusion. En effet, la diffusion sur des sites pornographiques de vidéos hypertruquées peut être réalisée de manière furtive¹⁹ et des VPN (*virtual private network* ou réseau privé virtuel)²⁰, peuvent être utilisés par les auteurs de ces vidéos.
- 15 Au-delà de la législation propre aux faits de *revenge porn*, les vidéos hypertruquées pornographiques peuvent également entrer dans le champ de la législation du droit à l'image et des atteintes à la vie privée. Là encore, les dispositions légales actuelles présentent des limites importantes pour appréhender ce phénomène nouveau.

B L'utilité relative de la législation du droit à l'image pour appréhender les *deepfakes*

- 16 Outre les dispositions du droit pénal régissant les cas de *revenge porn*, le droit à l'image pourrait présenter une utilité dans la lutte contre les vidéos hypertruquées, notamment pornographiques, tant pour les personnalités publiques que pour les victimes anonymes.

- 17 En droit français, le droit à l'image découle de l'article 9, alinéa 1 du Code civil, lequel dispose que « [chacun] a droit au respect de sa vie privée »²¹. Il permet à une personne, célèbre ou non, de s'opposer à la captation, la fixation ou à la diffusion de son image, sans son autorisation expresse et préalable. Ce droit s'applique donc de manière identique à toute personne et quel que soit le support de diffusion. Il a ainsi été jugé que « [toute] personne dispose sur son image, partie intégrante de sa personnalité, d'un droit exclusif qui lui permet de s'opposer à sa reproduction sans son autorisation expresse et spéciale ; de sorte que chacun a la possibilité de déterminer l'usage qui peut en être fait en choisissant notamment le support qu'il estime adapté à son éventuelle diffusion »²². Une autorisation expresse de la personne est nécessaire, quel que soit le lieu (public ou privé) dans lequel la personne a été prise en photographie ou filmée. Toute personne dont le droit à l'image n'a pas été respecté a la possibilité d'agir devant les tribunaux civils – y compris en référé²³ – afin d'obtenir le retrait des photographies ou images litigieuses et l'octroi de dommages et intérêts en réparation des préjudices subis. Par ailleurs, une action pénale est possible sur le fondement des articles 226-1 et 226-2 ci-dessus évoqués. Enfin, l'article 226-8 du Code pénal semble pouvoir être mobilisé puisque celui-ci punit d'un an d'emprisonnement et de 15 000€ 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 »²⁴.
- 18 Malgré ce panorama, la législation relative au droit à l'image ne convient pas totalement pour poursuivre les auteurs de vidéos hypertruquées pornographiques. Certes, le droit à l'image permet le retrait des vidéos et images sur décision des tribunaux civils, néanmoins la transformation d'images et le trucage algorithmique ne sont pas visés par les dispositions du code pénal qui ne sanctionnent que la captation, l'enregistrement et la transmission de paroles ou d'images sans le consentement de leur auteur. Or, l'algorithme n'utilise pas directement des images existantes pour les accoler dans un autre contexte mais transforme des images et vidéos grâce aux données récoltées dans d'autres images ou vidéos. L'enregistrement et la captation de paroles et d'images servent donc de support à la création de vidéos hypertruquées, ce qui peut rendre difficile l'application des articles 226-1 et 226-2 du Code pénal. Dit autrement, si le retrait des vidéos peut être prononcé sur décision des tribunaux civils, le principe d'interprétation stricte de la loi pénale pourrait faire obstacle à la poursuite devant les tribunaux correctionnels des auteurs de ces vidéos.
- 19 Seul l'article 226-8 semble s'appliquer en matière de vidéos hypertruquées, avec quelques réserves toutefois. Pour la Cour de cassation, « l'article 226-8 du code pénal ne réprime pas le montage en tant que tel, mais en ce qu'il tend à déformer de manière délibérée des images ou des paroles, soit par ajout, soit par retrait d'éléments qui sont étrangers à son objet »²⁵. Ce délit de manipulation du sens des images correspond alors à des faits de trucage de vidéos en ce que l'algorithme permet l'ajout d'images ou de paroles sur des images préexistantes dans le but de déformer délibérément les images initiales. Néanmoins, la notion même de montage ne recouvre que partiellement celle de trucage de vidéos, faisant plus référence au vidéo-montage (notamment par la technique de la coupe) qu'à la transformation d'images par algorithme.
- 20 A cela s'ajoute un dernier problème : aucune de ces dispositions relatives au droit à l'image ne punit de manière particulière le caractère pornographique ou à tout le moins sexuel du contenu diffusé. Or, on peut sans aucun doute analyser les vidéos

hypertruquées pornographiques non seulement comme une atteinte à la vie privée des individus, pourtant protégées tant par l'article 9 du Code civil que par l'article 8 de la Convention européenne des droits de l'homme, mais également comme une nouvelle forme de violence sexiste. La violence sexiste peut être définie comme « l'ensemble des violences, qu'elles soient verbales, physiques ou psychologiques, interpersonnelles ou institutionnelles, commises par les hommes en tant qu'hommes contre les femmes en tant que femmes, exercées tant dans les sphères publiques que privée »²⁶. Les actes sexistes renvoient ainsi, selon la secrétaire d'État chargée de l'Égalité entre les femmes et les hommes et de la lutte contre les discriminations Marlène Schiappa, au fait de « traiter différemment et négativement une femme par rapport à un homme [...] parce qu'elle est une femme »²⁷. Dans le cas des *deepfakes* pornographiques, il s'agit d'une pornographie non consentie rendue possible par la diffusion d'images sexualisées des femmes, visées dans 99% des vidéos hypertruquées pornographiques²⁸. Au-delà du débat sur les potentiels caractères violents et sexistes de la pornographie²⁹, c'est davantage le caractère non consenti de vidéos et images mettant en scène des femmes dans le cadre de pratiques sexuelles qui peut – et devrait – être assimilé ici à une violence sexiste.

- 21 En définitive, alors que les dispositions relatives au *revenge porn* prennent en compte le caractère sexuel des images diffusées avec ou sans le consentement de la personne représentée mais ne prennent pas en considération la possibilité de transformation et de trucage algorithmiques des images et paroles, les dispositions relatives au droit à l'image ne prennent quant à elles pas en compte les caractères sexuel et genré de l'atteinte à la vie privée. Une piste complémentaire envisageable reste la pénalisation de la diffusion de vidéos hypertruquées par le biais des infractions de presse, et notamment le délit de diffamation publique.

C L'imparfaite appréhension des *deepfakes* par la pénalisation des infractions de presse

- 22 La publication de vidéos hypertruquées pornographiques peut enfin être couverte par le délit de diffamation, prévu à l'article 29 al. 1 de la loi sur la presse de 1881. Cette loi est applicable à toutes les publications et pas seulement aux publications de presse, englobant *de facto* les publications réalisées par des personnes privées sur Internet. Deux cas de responsabilités doivent être distingués : l'engagement de la responsabilité du gestionnaire de la plateforme sur laquelle la vidéo hypertruquée est publiée d'une part, et l'engagement de la responsabilité de l'auteur de la publication d'autre part.
- 23 S'agissant du gestionnaire de la plateforme, il faut prendre en compte le fait que sur les plateformes de publication de vidéos pornographiques, le contenu peut être déterminé par les utilisateurs (ce sont eux qui publient les vidéos). Dans ce cas, le gestionnaire de la plateforme doit être considéré comme un hébergeur quant au contenu mis en ligne, il n'est donc pas responsable du contenu³⁰. La victime d'une *deepfake* pornographique ne pourra donc pas engager la responsabilité pénale du site diffusant la vidéo³¹. Elle pourra en revanche signaler le contenu sur la plateforme publique Pharos³² ou encore utiliser la procédure de la notification³³ à l'hébergeur de contenu en vue du retrait de la vidéo. Il lui faudra alors démontrer dans la notification le caractère illicite du contenu.
- 24 S'agissant de l'auteur de la publication, le contenu d'une *deepfake* peut être considéré comme diffamatoire en vertu de l'article 29 alinéa 1 de la loi de 1881. En effet, cet

article définit la diffamation comme « l'allégation ou l'imputation d'un fait qui porte atteinte à l'honneur ou à la considération ». Le fait imputé est dans le cas d'une vidéo hypertruquée précis, ce qui exclut donc l'infraction d'injure³⁴. Dans le cas d'une diffamation, il est nécessaire d'apporter la preuve de l'élément intentionnel. Toutefois, ce dernier est présumé dès lors qu'est apportée la preuve de l'intention de publication, ce qui est constitué dans le cas de vidéos hypertruquées publiées sur un site pornographique. Le problème de cette législation réside dans le délai de prescription d'une telle infraction. Le délai prévu par la loi de 1881 est de trois mois pour les diffamations publiques et celui-ci est porté à un an en cas de diffamation à raison du sexe³⁵. Cependant, l'identification du point de départ d'un tel délai soulève des difficultés. La Cour d'appel de Paris³⁶ avait fait des infractions sur Internet des infractions continues c'est-à-dire des infractions dont les éléments constitutifs se renouvellent chaque jour, ce qui était davantage protecteur des victimes. Cette solution a toutefois fait l'objet de critiques puisqu'en pratique l'infraction devenait *de facto* imprescriptible. La chambre criminelle de la Cour de cassation³⁷ a clarifié la situation en affirmant que les infractions commises sur Internet sont des infractions permanentes. Il en résulte que le point de départ du délai est le jour de la commission de l'infraction, et donc la date de la publication du contenu diffamatoire. L'insertion sur un autre site d'un lien menant au contenu litigieux constitue tout de même, selon la Cour de cassation, une nouvelle publication³⁸. L'infraction de diffamation n'est toutefois que peu protectrice des victimes dans le cas de vidéos hypertruquées à caractère pornographique du fait d'une diffusion souvent inconnue des victimes pendant une période relativement longue, remettant en cause le caractère instantané des délits de presse commis sur Internet.

- 25 En définitive, on aboutit au constat de l'imperfection des dispositions législatives actuelles à appréhender le phénomène des *deepfakes* et à lutter contre les conséquences de ces vidéos hypertruquées sur les victimes. Les dispositions relatives au *revenge porn* sont limitées par l'absence de prise en compte de la possibilité de transformation et de trucage algorithmiques des images et paroles ; les dispositions relatives au droit à l'image ne prennent quant à elles pas en compte les caractères sexuel et genré de l'atteinte à la vie privée ; enfin, les infractions de presse semblent finalement adaptées à la lutte contre les *deepfakes* à caractère sexuel ou pornographique mais sont limitées par les délais de recours.
- 26 Les insuffisances du cadre législatif français pour appréhender les *deepfakes* pornographiques se retrouvent également en matière de vidéos hypertruquées diffusées à des fins politiques.

II. La lutte contre les *deepfakes* politiques : les ambiguïtés du cadre législatif actuel

- 27 La seconde hypothèse de vidéos hypertruquées concerne les vidéos diffusées à des fins politiques. L'exemple type est la vidéo précédemment citée mettant en scène Barack Obama insultant son successeur. Si leur nombre reste relativement restreint³⁹, et si le phénomène ne semble pas encore très développé en Europe⁴⁰, leur fort pouvoir de déstabilisation politique ne doit pas être ignoré⁴¹ et devrait amener à les considérer avec le plus grand sérieux. Là encore, il s'agit de déterminer si l'arsenal législatif français, renforcé en 2018 par la loi sur la manipulation de l'information⁴², est suffisant,

comme le suggérait le Gouvernement dans sa réponse à la question écrite de la députée Caroline Janvier⁴³. A l'analyse de trois pans de la législation française actuelle, le constat de l'imperfection du cadre législatif français demeure : le droit au respect de la vie privée et le droit à l'image butent sur l'imprécision des dispositions, notamment du Code pénal, dans la lutte contre les *deepfakes* (A) ; la liberté artistique permet d'autoriser certaines vidéos hypertruquées (B) et la récente loi relative à la manipulation de l'information ne s'applique réellement que sur une durée précise et restreinte, faisant obstacle à la lutte contre les vidéos hypertruquées diffusées à des fins de déstabilisation politique (C). *In fine*, comme en matière de *deepfakes* à caractère sexuel ou pornographique, la législation française actuelle semble limitée pour appréhender ce phénomène.

A Le droit à l'image et le droit au respect de la vie privée des personnalités politiques : un cadre législatif trop imprécis pour être protecteur

- 28 Comme pour les *deepfakes* à caractère pornographique, les victimes de vidéos hypertruquées diffusées à des fins politiques peuvent se saisir du droit au respect de leur vie privée, tant dans son volet civil (article 9 du Code civil) que dans son volet pénal (article 226-1 et 226-2 du Code pénal). La portée de la protection offerte par le droit au respect de la vie privée est néanmoins à relativiser, puisqu'elle s'apprécie en fonction des circonstances dans lesquelles est invoqué ce droit⁴⁴. En principe, la notoriété de la personne n'a ainsi pas d'incidence sur l'interprétation du contenu de la vie privée⁴⁵. Dit autrement, les personnalités publiques, et par suite les personnalités du monde politique, ont, elles aussi, le droit de faire respecter leur vie privée et leur droit à l'image⁴⁶ mais une telle interprétation n'est toutefois pas automatique comme le révèle l'arrêt de Grande Chambre de la Cour européenne des droits de l'homme dans l'affaire *Von Hannover contre Allemagne n°247*. Dans sa décision, la Cour rappelle la nécessité de « [mettre] en balance [le] droit au respect de la vie privée et [le] droit à la liberté d'expression ». La liberté d'expression des auteurs de publications, y compris d'images, diffusées sans le consentement de la personne représentée, ne sera ainsi restreinte qu'à condition de prouver que la parution de ces éléments de la vie privée de la personne représentée n'apporte pas d'éléments à un débat d'intérêt général⁴⁸. Ainsi, la représentation de personnalités connues du public, y compris dans des vidéos hypertruquées, peut être autorisée à condition de prouver l'existence d'un débat d'intérêt général⁴⁹, la notion étant elle-même interprétée en fonction des circonstances de l'affaire. Dans le cas des vidéos hypertruquées de Barack Obama ou encore de Mark Zuckerberg, leur objet ayant été de dénoncer de nouvelles utilisations des nouvelles technologies⁵⁰, il est fort à parier que ce type de *deepfakes* pourraient être considérées comme ne violant pas le droit à l'image ni le droit au respect de la vie privée des personnes représentées au regard du droit français actuel⁵¹ ou au regard du droit européen des droits de l'homme.
- 29 Au-delà de ces considérations, l'absence de consentement de la personne représentée pourrait permettre de déclencher l'application des articles 226-1 et 226-2 du Code pénal, avec la limite déjà énoncée selon laquelle ces dispositions légales ne mentionnent pas la possibilité d'un usage de l'image à des fins de trucage et de manipulation. De même, les limites de l'article 226-8 du Code pénal sanctionnant la

manipulation d'images sont identiques en matière de vidéos hypertruquées à des fins politiques. Les victimes de *deepfakes* ne semblent donc pas pouvoir poursuivre les auteurs de ces vidéos ou de leur diffusion à l'aide de ces dispositions, tant s'agissant de vidéos pornographiques que de vidéos diffusées à des fins politiques, et cela en raison de l'interprétation stricte de la loi pénale et de l'imprécision de ces dispositions pour appréhender le phénomène des *deepfakes*.

- 30 Une autre voie consisterait à contester non pas la vidéo hypertruquée en elle-même mais l'utilisation d'images issues d'une banque d'images en ligne en vue de créer le *deepfake*. Sur ce point, le TGI de Paris a clairement énoncé « qu'on ne saurait déduire de l'éventuelle présence de clichés sur le réseau internet (...), la preuve que ceux-ci seraient de "libre parcours" et que cette situation priverait la requérante du droit d'agir sur le fondement de l'article 9 du Code civil »⁵², d'autant que la présence de telles images peut également être illicite à défaut d'autorisation de la personne représentée⁵³. *In fine*, les personnalités politiques ne semblent pas davantage protégées en matière de *deepfakes* politiques que ne le sont les victimes de vidéos hypertruquées pornographiques par les dispositions relatives au droit à l'image. Les dispositions sanctionnant les infractions de presse sont également limitées en partie s'agissant des *deepfakes* créées et diffusées à des fins politiques.

B Les infractions de presse limitées par la liberté d'expression et la liberté artistique : des *deepfakes* rendues juridiquement acceptables

- 31 Nous l'avons vu s'agissant des *deepfakes* à caractère pornographique, la publication de vidéos hypertruquées pourrait emporter la qualification de diffamation, puisque la vidéo impute un fait à autrui, lequel peut porter atteinte à son honneur ou à sa considération⁵⁴. Une telle vidéo peut également emporter la qualification d'injure si elle contient une « expression outrageante », des « termes de mépris » ou encore une « invective qui ne renferme l'imputation d'aucun fait »⁵⁵. Une fois la qualification établie, le juge judiciaire détermine si les faits sont couverts par la liberté d'expression de leur auteur. Ainsi, la liberté d'expression ainsi que la liberté de création artistique⁵⁶ viennent limiter le caractère diffamatoire de la vidéo diffusée et potentiellement truquée ou créée de toute pièce. L'appréciation du juge de l'application de la liberté d'expression repose notamment sur plusieurs conditions : la restriction à la liberté d'expression doit avoir été prévue par la loi (ce qui est le cas pour l'infraction de diffamation) ; elle doit poursuivre un objectif légitime (ici la protection des droits d'autrui) ; elle doit enfin être proportionnée (le caractère artistique de la vidéo hypertruquée permettant d'interpréter la proportionnalité de la restriction à la liberté d'expression avec plus de souplesse).
- 32 Ce cas de *deepfake* artistique n'est pas si rare : la vidéo de Mark Zuckerberg était, selon son auteur, un projet artistique visant à dénoncer la collecte de données par des entreprises devenues trop puissantes⁵⁷. De la même manière, la vidéo de Barack Obama insultant Donald Trump visait à dénoncer « la façon dont l'information se répand » selon Jonah Peretti, PDG de BuzzFeed à l'initiative de cette vidéo⁵⁸. Des vidéos similaires pourraient probablement émerger et viser des hommes ou des femmes politiques françaises. L'appréciation du juge est ainsi déterminante et elle se fait « au regard des valeurs sociales dominantes »⁵⁹. N'entrent pas dans le champ de la liberté d'expression les « agissements constitutifs d'infractions pénales (...), [les] comportements considérés

comme contraires aux valeurs morales communément admises au jour où le juge statue, étant précisé que ces notions doivent s'apprécier au regard de considérations objectives et non en fonction de la sensibilité personnelle et subjective de la personne visée »⁶⁰.

- 33 Il convient de préciser que les limitations à la liberté d'expression artistique admises tant en droit français qu'en droit européen prennent en compte la possibilité pour une législation de poursuivre au pénal un artiste du fait d'une œuvre⁶¹. Ainsi, la Cour européenne des droits de l'homme a affirmé que « l'artiste et ceux qui promeuvent ses œuvres n'échappent pas aux possibilités de limitation que ménage le paragraphe 2 de l'article 10 (art. 10-2). Quiconque se prévaut de sa liberté d'expression assume en effet, selon les propres termes de ce paragraphe, des "devoirs et responsabilités" ; leur étendue dépend de sa situation et du procédé utilisé (...) ; la Cour ne saurait le perdre de vue en contrôlant la nécessité de la sanction incriminée dans une société démocratique »⁶². Enfin, s'agissant des cas les plus complexes de vidéos hypertruquées, c'est-à-dire ceux où l'intention de leur auteur n'est pas rendue publique, leurs auteurs peuvent semble-t-il être poursuivis sur le fondement des infractions d'injure ou de diffamation, à condition toutefois de pouvoir identifier l'auteur de l'infraction.
- 34 La liberté d'expression et *a fortiori* la liberté artistique qui en découle permettent ainsi de rendre légitimes certaines vidéos hypertruquées, quand bien même leur diffusion poursuit des fins politiques. Toutes les *deepfakes* ne sont – et ne doivent – pas être interdites et un juste équilibre doit être trouvé, parfois, entre les droits et libertés de la personne dont l'image a été utilisée et ceux de l'auteur du trucage. Là encore toutefois, un encadrement spécifique du phénomène des *deepfakes* – précis et complet – pourrait prévoir avec une précision identique des dérogations pour aménager la liberté d'expression et la liberté artistique des créateurs de ces vidéos hypertruquées.
- 35 Ce constat de l'imprécision et l'incapacité à lutter finement contre le phénomène *deepfake* doit être étendu à la récente loi relative à la lutte contre la manipulation de l'information.

C La spécificité de la loi relative à la lutte contre la manipulation de l'information : un cadre législatif imparfait

- 36 Le gouvernement français s'est targué du caractère suffisant de la législation française actuelle pour prévenir et lutter contre la technique des *deepfakes*⁶³, ce dont nous pouvons douter au regard des éléments précédemment développés. Notamment, l'exécutif semblait considérer que la récente loi relative à la lutte contre la manipulation de l'information⁶⁴ est applicable aux cas de vidéos hypertruquées visant le personnel politique et par conséquent suffisante. Cette législation introduit notamment dans le code électoral de nouvelles infractions relatives à la manipulation de l'information, afin de « mieux protéger la démocratie contre les diverses formes de diffusion intentionnelle de fausses nouvelles »⁶⁵ selon le gouvernement. Elle distingue deux périodes : d'une part en période électorale, c'est-à-dire pendant la campagne et durant les élections, la loi crée une obligation de transparence pour les plateformes (publication des algorithmes, signalement des contenus sponsorisés, publication du nom de l'auteur de tels contenus...) et une action judiciaire en référé permettant de faire cesser la circulation de fausses nouvelles (la « fausse nouvelle » étant interprétée par le juge des référés autour de 3 critères définis par la loi de 1881 : elle doit être

manifestement inexacte ou trompeuse, elle doit être diffusée massivement et de manière artificielle, délibérée ou automatisée, elle doit constituer un trouble à la sincérité du scrutin) ; d'autre part, en dehors des périodes électorales, la législation crée un devoir de coopération des plateformes, dont le contrôle a été confié au Conseil Supérieur de l'Audiovisuel (CSA).

- 37 Si la loi relative à la lutte contre la manipulation de l'information a bien vocation à s'appliquer en matière de vidéos hypertruquées, plusieurs limites doivent cependant être relevées à l'application de cette législation. D'une part, la qualification de « nouvelle » ou de « fausse nouvelle » semble ne pas correspondre parfaitement à la réalité des *deepfakes*. En effet, toutes les vidéos hypertruquées ne diffusent pas nécessairement une (fausse) information ou une (fausse) nouvelle. A tout le moins, la législation sur la manipulation de l'information ne prend pas en compte le caractère proprement factice de la vidéo, indépendamment de l'information contenue par cette vidéo. D'autre part, la distinction opérée entre la période électorale et les périodes hors campagne électorale repose, s'agissant *a minima* des vidéos hypertruquées, sur une mécompréhension des effets réels de ces vidéos. En effet, en limitant la possibilité d'action en référé à la seule période électorale afin de protéger la sincérité des scrutins, le législateur écarte la possibilité d'appliquer cette législation aux actes de déstabilisation politique pouvant avoir cours durant les mandats politiques. Par ailleurs, la révélation de la nature factice d'une vidéo hypertruquée diffusée n'a pas nécessairement d'effet ensuite sur la déstabilisation politique que celle-ci a provoqué, puisqu'elle peut se poursuivre bien après cette révélation. Or, la finalité de la loi relative à la lutte contre la manipulation de l'information semble être justement de se prémunir contre ces nouvelles formes de déstabilisation politique puisqu'elle vise à « mieux protéger la démocratie »⁶⁶. Comme nous le démontrons à l'étranger la manipulation vidéo du discours de Nancy Pelosi⁶⁷, présidente démocrate de la Chambre des représentants, il est regrettable que l'action judiciaire en référé visant à faire cesser la circulation d'une information trompeuse ne soit donc restreinte qu'à la période électorale dans un objectif de protection de la sincérité des scrutins, puisque la manipulation de l'opinion peut avoir cours en dehors de cette période.
- 38 Pour conclure, on peut ainsi déplorer qu'à la suite de l'interpellation du gouvernement par la députée Caroline Janvier sur la nécessité de légiférer sur les *deepfakes*⁶⁸, l'exécutif n'ait concentré sa réponse que sur l'efficacité de la législation française actuelle à lutter contre la seule diffusion de fausses informations via des vidéos hypertruquées. Il ne semble en revanche pas avoir saisi la mesure du vide juridique qui persiste au sein de la législation française en raison de l'aspect parcellaire, incomplet ou imprécis de la protection que peuvent conférer les dispositions actuelles face aux *deepfakes* pornographiques comme politiques. Les *deepfakes* pornographiques se retrouvent ainsi à l'intersection de plusieurs législations qui ne sanctionnent jamais entièrement les faits en cause et ne permettent donc pas une appréhension fine du phénomène. Quant aux *deepfakes* diffusées à des fins politiques, si certaines doivent effectivement être autorisées en vertu du droit à la liberté d'expression et à la liberté artistique, on peut déplorer que le cadre législatif relatif au droit à la vie privée, au droit à l'image ou à la lutte contre les infos n'ait pas été adapté à la prise en compte du phénomène *deepfakes*, lequel dépasse assez largement la seule diffusion de fausses informations ou de fausses nouvelles. La déstabilisation politique qui peut résulter de ces vidéos

hypertruquées n'est pas – semble-t-il – pleinement prise en compte par le législateur à l'heure actuelle.

- 39 L'encadrement juridique spécifique de cette pratique pourrait, à un autre niveau, alimenter le débat relatif à l'opportunité d'une sanction des atteintes à la vérité. Alors que « la libre communication des pensées et des opinions est un des droits les plus précieux de l'homme »⁶⁹, certains discours portant atteinte à la vérité ont pu être interprétés comme constituant des abus de la liberté d'expression⁷⁰. Ce prisme de lecture, qui a déjà été décliné s'agissant de la négation de faits historiquement avérés⁷¹, se trouve renouvelé par certaines atteintes à la vérité rendues possibles par l'usage des nouvelles technologies de l'information et de la communication. L'atteinte à la vérité que constituent les *deepfakes* doit-elle être sanctionnée en tant que telle ou seulement en ce que ces vidéos portent atteinte à d'autres droits protégés ? Une préférence manifeste pour la première hypothèse impliquerait d'ériger les atteintes à la vérité au rang de problème d'intérêt général⁷², au risque de limiter grandement certains droits fondamentaux tels que la liberté d'expression et la liberté artistique. L'analyse du droit positif actuel et de sa capacité à lutter contre cette nouvelle pratique démontre bien au contraire que les *deepfakes* manipulent, certes, la vérité mais portent avant tout atteinte à certains droits fondamentaux ou intérêts publics protégés. Si le cadre juridique mérite d'être parfait pour appréhender ce phénomène, la démarche gagne selon nous à être conduite, non pas en défense de la vérité, mais bien en défense du droit à la vie privée, du droit à l'image, de la protection de la dignité, de l'honneur, de la considération des personnes, ou encore des intérêts démocratiques et de la sincérité des scrutins.

NOTES

1. THÉVENIN Lauren et MAZUIR Valérie, « Municipales à Paris : Benjamin Griveaux jette l'éponge », *LesÉchos.fr*, 14 février 2020, url :
2. La nature de ces vidéos entre *deepfake* ou *revenge porn* avait dans un premier temps soulevé quelques interrogations.
3. Sur des plateformes de vidéos pornographiques, dites « *hub* » c'est-à-dire une structure concentrant les flux (ici de données), les triant, les organisant et les diffusant.
4. Il s'agit de la technique dite « GAN » c'est-à-dire des « réseaux adversatifs générateurs ».
5. V. par exemple : SIX Nicolas, « Une vidéo truquée de Mark Zuckerberg façon « *deepfake* » éprouve les limites d'Instagram », *Le Monde*, 13 juin 2019 ; KOENIG Gaspard, « Les « *deep fakes* » ou la fin du débat démocratique », *Les Échos, Édits & Analyses*, 16 octobre 2019 ; PAUL Kari, « California makes 'deepfake' videos illegal, but law may be hard to enforce », *The Guardian*, 7 octobre 2019.
6. Deeptrace, 27 septembre 2019, Rapport « The State Of Deepfakes: Landscape, Threats and Impact ».
7. STUDER Bruno, député, Rapport n°990 fait au nom de la commission des affaires culturelles et de l'éducation sur la proposition de loi relative à la lutte contre la manipulation de l'information, 30 mai 2018 ; STUDER Bruno, député, Rapport d'information n°1296 fait au nom de la commission

des affaires culturelles et de l'éducation en conclusion des travaux de la mission d'information sur l'école dans la société du numérique, 10 octobre 2018.

8. Question écrite n°16587 de Mme Caroline JANVIER, publiée au JO le 5 février 2019, p. 1052.

9. Art. 226-1 Code pénal.

10. V. *infra*. (I/ B°/)

11. C. cass., crim., 16 mars 2016, n°780, ECLI:FR:CCASS:2016:CR00780.

12. *Ibid.* : « Attendu que, pour confirmer cette décision, l'arrêt énonce que le fait, pour la partie civile, d'avoir accepté d'être photographiée ne signifie pas, compte tenu du caractère intime de la photographie, qu'elle avait donné son accord pour que celle-ci soit diffusée ; Mais attendu qu'en se déterminant ainsi, alors que n'est pas pénalement réprimé le fait de diffuser, sans son accord, l'image d'une personne réalisée dans un lieu privé avec son consentement, la cour d'appel a méconnu les textes susvisés et le principe ci-dessus énoncé ».

13. Art. 67, loi n°2016-1321, 7 octobre 2016, Loi pour une République numérique, JORF n°0235 du 8 octobre 2016.

14. Art. 226-2-1 Code pénal. Nous soulignons.

15. Pour une critique de la formulation de l'article 226-2-1 du Code pénal, V. DETRAZ Stéphane, « Les nouvelles dispositions réprimant les atteintes à l'intimité sexuelle : faire compliqué quand on peut faire simple (Commentaire de l'article 226-2-2 du code pénal issu de la loi n°2016-1321 du 7 octobre 2016) », *Revue de sciences criminelles et de droit pénal comparé*, 2016/4, p. 741-753.

16. JOURDAA Laurent, « Les atteintes pénales à l'image des personnes face au droit du public à l'information légitime », *Journée d'étude: Variations sur l'effectivité du consentement*, Atelier droit pénal – Les atteintes au consentement en matière pénale, Juin 2018, Toulon, France, p. 7.

17. V. *supra* : 99% des *deepfakes* pornographiques concernent des femmes. V. Deeptrace, 27 septembre 2019, Rapport « The State Of Deepfakes: Landscape, Threats and Impact ».

18. Art. 5 § 18.2-386.2. A., Code de Virginie : « Any person who, with the intent to coerce, harass, or intimidate, maliciously disseminates or sells any videographic or still image created by any means whatsoever that depicts another person who is totally nude, or in a state of undress so as to expose the genitals, pubic area, buttocks, or female breast, where such person knows or has reason to know that he is not licensed or authorized to disseminate or sell such videographic or still image is guilty of a Class 1 misdemeanor ».

19. L'anonymat sur Internet relève davantage du mythe. En revanche, il est possible de rendre l'identification plus difficile en usant de certains outils informatiques, d'où l'idée de furtivité. V. par exemple : SYLVESTRE Guillaume. « Être furtif et anonyme sur Internet », *I2D - Information, données & documents*, vol. volume 54, n° 3, 2017, pp. 34-35.

20. Les VPN sont des extensions des réseaux locaux permettant la transmission de données de manière chiffrée, et offrant par suite un relatif anonymat dans l'utilisation d'Internet (relatif puisque l'identification est réalisée par le fournisseur du VPN, lequel renforce ensuite la confidentialité de l'utilisateur et peut notamment masquer l'identification de l'utilisateur).

21. C. cass., 1^{ère} civ., 16 juillet 1998, n° 96-15.610, Bull. 1998, I, n°259, p. 181.

22. TGI de Paris, 12 septembre 2000, *Charlotte R. épouse Jean-Michel J./SARL DF Presse*.

23. Art. 9, al.2, C. civ. : « Les juges peuvent, sans préjudice de la réparation du dommage subi, prescrire toutes mesures telles que séquestre, saisie, et autres, propres à empêcher ou faire cesser une atteinte à l'intimité de la vie privée ; ces mesures peuvent, s'il y a urgence, être ordonnées en référé ». V. également : LOUVET Marc-Noël. « Le juge des référés et la protection de la vie privée », *LEGICOM*, vol. 20, no. 4, 1999, pp. 27-33.

24. Art. 226-8 Code pénal.

25. C. cass., crim., 30 mars 2016, n°15-82039.

26. SIMONETTI Ilaria, « Violence et genre », in RENNES Juliette (dir.), *Encyclopédie critique du genre*, Paris, La Découverte, 2016, p. 681.

27. *JO Sénat*, Séance du 5 juillet 2018, p. 9220, url : <https://www.senat.fr/seances/s201807/s20180705/s20180705.pdf>. Ces propos ont été tenus lors des débats au Sénat relatifs à l'examen de la Loi n°2018-703 renforçant la lutte contre les violences sexuelles et sexistes, adoptée le 3 août 2018.
28. Données du rapport *Deeptrace* 2019.
29. Pour une synthèse de ces débats, V. TRACHMAN Mathieu et VÖRÖS Florian, « Pornographie », in RENNES Juliette (dir.), *Encyclopédie critique du genre, op. cit.*, pp. 479-485.
30. Loi n°2000-719, 1^{er} août 2000 relative à la liberté de communication.
31. Art. 6 §I-2 et §I-3, Loi n°2004-575, 21 juin 2004 pour la confiance dans l'économie numérique.
32. Il ne s'agit alors pas d'une procédure judiciaire.
33. Art. 6 §I-7, Loi n°2004-575, 21 juin 2004 pour la confiance dans l'économie numérique.
34. Sur la distinction entre injure et diffamation, V. ROMAN Diane et HENNETTE-VAUCHEZ Stéphanie, *Droits de l'Homme et libertés fondamentales*, Paris, Dalloz, 3^{ème} éd., 2017, p. 433.
35. Loi n°2004-204, 9 mars 2004 portant adaptation de la justice aux évolutions de la criminalité
36. Cour d'appel de Paris, 15 décembre 1999, *Légipresse*, mars 2000, n°169-III, p. 38, note Basile ADER.
37. C.cass., crim., 16 octobre 2001, D. 2001.
38. Cass. Crim 2 novembre 2016, pourvoi n°15-87.163 à propos de l'insertion d'un lien hypertexte publié sur un site personnel rendant à nouveau accessible un texte pouvant faire l'objet de poursuites en diffamation. LAVRIC Sabrina, « Presse : l'insertion d'un lien hypertexte constitue une nouvelle publication », *Dalloz actualités*, 17 novembre 2016.
39. 4% des *deepfakes* selon le rapport *Deeptrace* 2019.
40. La majorité des vidéos truquées politiques ou commerciales ayant été diffusées hors d'Europe.
41. Le rapport *Deeptrace* cite ainsi un scandale politique à caractère sexuel en Malaisie en juin 2019. V. rapport *Deeptrace*, *op. cit.*, p. 10. Dans le même ordre d'idée, la manipulation audio d'un discours de Nancy PELOSI, présidente de la Chambre des représentants aux États-Unis d'Amérique, s'apparente à un *deepfake*. V. rapport *Deeptrace*, *op. cit.*, p. 11.
42. Loi n°2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information, JORF n°0297 du 23 décembre 2018.
43. Réponse à la question écrite n°16587 de Mme Caroline JANVIER, publiée au JO le 15 octobre 2019, p. 9053
44. LEPAGE Agathe, *Libertés et droits fondamentaux à l'épreuve d'internet*, Paris, LexisNexis, 2002, p. 101.
45. *Ibidem*, p. 102.
46. V. également : CEDH, 24 juin 2004, 3^{ème} section, *Von Hannover c. Allemagne (n°1)*, req. n°59320/00, Rec. CEDH 2004-VI, notamment §63 : « la Cour considère qu'il convient d'opérer une distinction fondamentale entre un reportage relatant des faits – même controversés – susceptibles de contribuer à un débat dans une société démocratique, se rapportant à des personnalités politiques, dans l'exercice de leurs fonctions officielles par exemple, et un reportage sur les détails de la vie privée d'une personne qui, de surcroît, comme en l'espèce, ne remplit pas de telles fonctions » ; et § 69 : « Elle estime que toute personne, même connue du grand public, doit pouvoir bénéficier d'une « espérance légitime » de protection et de respect de sa vie privée ».
47. CEDH [GC], 7 février 2012, *Von Hannover c. Allemagne (n°2)*, req. n°40660/08 et 60641/08, Rec. CEDH 2012-I.
48. *Ibidem*, §109. La Cour en l'espèce pose les critères d'appréciation de la mise en balance entre la liberté d'expression et le droit au respect de la vie privée : 1) la contribution de la publication en cause à un débat d'intérêt général ; 2) la notoriété de la personne visée et l'objet de la publication ; 3) le comportement antérieur de la personne concernée ; 3) le contenu, la forme et

les répercussions de la publication ; 4) les circonstances de la prise des photos ou le mode d'obtention des informations et leur véracité, selon la nature de l'atteinte ; 5) la gravité de la sanction.

49. V. BIGOT Christophe. « La protection de la vie privée par la Cour européenne des droits de l'homme », *LEGICOM*, vol. 43, no. 2, 2009, pp. 43-49. Pour l'auteur, deux indications peuvent être observées dans la jurisprudence européenne quant à l'interprétation de la notion de « contribution à un débat d'intérêt général » : d'une part, la Cour européenne des droits de l'homme distingue le simple rattachement à l'intérêt général d'une contribution à un débat d'intérêt général ; d'autre part, « Pour la Cour, la volonté d'illustrer ne semble pas suffisante [...] pour caractériser une contribution à un débat d'intérêt général, même si l'article lui-même porte sur un sujet d'intérêt général ».

50. V. *infra*.

51. La critique et la satire du personnel politique est d'ailleurs protégée par la liberté d'expression. V. par exemple : CEDH, 15 mars 2011, *Otegi Monragon c. Espagne* : « les limites de la critique admissible sont plus larges à l'égard d'un homme politique, visé en cette qualité, que d'un simple particulier ».

52. TGI de Paris, ch. presse, 12 décembre 2000.

53. V. par exemple l'affaire *Estelle Hallyday* : CA Paris, 10 février 1999.

54. Art. 29 al. 1, loi du 29 juillet 1881.

55. Art. 29 al. 2, loi du 29 juillet 1881 ; art. 33, loi du 29 juillet 1881.

56. La liberté de création artistique est reconnue comme découlant de la liberté d'expression. V. à ce propos : CEDH, 24 mai 1988, *Müller et autres c. Suisse*, req. n°10737/84, notamment §33 : « Ceux qui créent, interprètent, diffusent, ou exposent une œuvre d'art contribuent à l'échange d'idées et d'opinions indispensables à une société démocratique. D'où l'obligation, pour l'État, de ne pas empiéter indûment sur leur liberté d'expression ».

57. Il s'agit ici du projet « Spectre » de Bill POSTERS et Daniel HOWE. V. POSTERS Bill, « Gallery : Spectre launches (Press release) », 29 mai 2019.

58. BRAUN Elisa, « La viralité d'une fausse vidéo d'Obama met en lumière le phénomène du « deep fake » », *Le Figaro*, 23 mai 2018.

59. ROMAN Diane et HENNETTE-VAUCHEZ Stéphanie, *op. cit.*, p. 434.

60. *Ibidem*.

61. CEDH, *Müller et autres c. Suisse*, préc.

62. *Ibidem*, §34.

63. Réponse à la question écrite n°16587 de Mme Caroline JANVIER, publiée au JO le 15 octobre 2019, p. 9053.

64. Loi n°2018-1202 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information, JORF n°0297 du 23 décembre 2018.

65. V. la communication gouvernementale autour de ce texte.

66. *Ibid.*

67. GRAND Harold, « 'Deepfake' : une vidéo trafiquée de Nancy Pelosi relayée par des proches de Donald Trump », *LeFigaro.fr*, 24 mai 2019.

68. Réponse à la question écrite n°16587 de Mme Caroline JANVIER, publiée au JO le 15 octobre 2019, p. 9053.

69. Art. 11, Déclaration des droits de l'homme et du citoyen, 26 août 1789.

70. Ce fût le cas pour les discours niant la Shoah. V. délit de négationnisme, art. 24bis, loi du 29 juillet 1881 sur la liberté de la presse, introduit par la loi n°90-615, 13 juillet 1990 tendant à réprimer tout acte raciste, antisémite ou xénophobe.

71. V. notamment : HOCHMANN Thomas, *Le négationnisme face aux limites de la liberté d'expression, Étude de droit comparé*, Paris, Pedone, 2013 ; HOCHMANN Thomas, « Négationnisme : le Conseil constitutionnel entre ange et démon », *Revue des droits et libertés fondamentaux*, 2016 chron. n°03.

72. En effet, le raisonnement relatif aux atteintes à la liberté d'expression repose sur la mise en balance entre cette liberté d'une part et les intérêts auxquels son exercice porte atteinte d'autre part. Mettre en balance la liberté d'expression et les atteintes à la vérité implique d'ériger la vérité comme intérêt général protégé. V. sur la mise en balance par exemple : HOCHMANN Thomas, « Négationnisme du génocide arménien : défauts et qualités de l'arrêt *Perinçek contre Suisse* », *Revue des droits et libertés fondamentaux*, 2015 chron. n°27, à propos l'affaire *Perinçek contre Suisse* relative à la condamnation d'un homme politique nationaliste turc du fait de ses déclarations niant l'existence du génocide arménien.

RÉSUMÉS

Face à un phénomène en pleine expansion qu'est la diffusion de vidéos hypertruquées sur internet – les *deepfakes* – le droit français actuel est-il en mesure de protéger les victimes de ces vidéos et de leur offrir des voies de droit efficaces ? Si le gouvernement semble convaincu de la capacité de la législation actuelle à lutter contre des vidéos *deepfakes*, on peut toutefois relativiser cette affirmation : tant dans le domaine des vidéos hypertruquées pornographiques que dans le domaine des vidéos *deepfakes* politiques, la législation actuelle reste incomplète pour lutter contre ces vidéos.

The broadcasting of deepfakes on the internet is a growing phenomenon that questions the capacity of the french Law to protect victims of these videos and to provide effective legal remedies. If the french government seems convinced that the current legislation can be used in order to fight effectively against these videos, this assertion is questionable : the current legislation is still incomplete to fight against both pornographic and political deepfakes.

INDEX

Mots-clés : NTIC - nouvelles technologies - revenge porn - violence sexiste - droit à l'image - diffamation - infox - liberté d'expression - vérité

Keywords : NICT - new technology - revenge porn - gender-based violence - image reproduction right - libel action - fake news - freedom of speech - trut

AUTEUR

CLAIRE LANGLAIS-FONTAINE

Claire Langlais-Fontaine est doctorante et ATER en droit public à l'Université Paris Nanterre – CTAD UMR7074 (équipe CREDOF)