Lex Electronica

Lex Electronica

L'hypertrucage : analyse du phénomène des « deepfakes » et recommandations

Simon Robichaud-Durand

Volume 28, numéro 4, 2023

Droit & Génération(s) numérique(s)

URI : https://id.erudit.org/iderudit/1108807ar
DOI : https://doi.org/10.7202/1108807ar

Aller au sommaire du numéro

Éditeur(s)

Centre de recherche en droit public Université de Montréal

ISSN

1480-1787 (numérique)

Découvrir la revue

Citer cet article

Robichaud-Durand, S. (2023). L'hypertrucage : analyse du phénomène des « deepfakes » et recommandations. *Lex Electronica*, *28*(4), 78–98. https://doi.org/10.7202/1108807ar

Résumé de l'article

Que ce soit l'hypertrucage, la technologie de permutation intelligente de visages, ou les « *deepfakes* », ces termes réfèrent aux montages ultra réalistes générés par l'intelligence artificielle ayant comme objectif de tromper son public. Aujourd'hui, cette technologie n'est plus exclusive aux studios hollywoodiens puisqu'elle est accessible à n'importe qui. Conséquemment, des montages générés, on dépeint des personnalités publiques, des chefs d'État et plusieurs victimes non consensuelles. L'hypertrucage à finalité pornographique et non consensuelle a rapidement émergé, ciblant majoritairement des victimes de sexe féminin causant de sérieux préjudices. Toutefois, les auteurs priorisent les dangers publics tels que l'ingérence étrangère, la perte de confiance dans les institutions publiques et la désinformation politique. Ainsi, nous tentons de comprendre pourquoi la majorité des écrits focalisent sur ces risques politiques alors qu'en réalité la grande majorité des montages d'hypertrucage est de finalité pornographique non consensuelle. De plus, considérant que la *Loi électorale du Canada* semble protéger contre les dangers des montages d'hypertrucages ayant une finalité politique, alors que les victimes de montages pornographiques et non consensuelles sont toujours sans recours concret, nous proposons une solution multiapproche regroupant quatre volets : la législation, la sensibilisation, l'innovation et la collaboration.

L'HYPERTRUCAGE : ANALYSE DU PHÉNOMÈNE DES « DEEPFAKES » ET RECOMMANDATIONS

Simon ROBICHAUD-DURAND[24]

24 L'auteur est diplômé du programme de maitrise LL.M avec concentration en droit et technologie (LL.M.) à l'Université d'Ottawa. Il est titulaire d'une licence en droit (LL.L) et d'un B.Sc.Soc. spécialisé en Développement international et mondialisation. Simon est associé à la Chaire Intelligence artificielle responsable à l'échelle mondiale de l'Université d'Ottawa (srobi134@uottawa.ca). **Note au lecteur** : les recherches menées pour l'écriture du présent article datent de la mi-2022 et ne font pas état des développements juridiques plus récents en la matière.

RÉSUMÉ

Que ce soit l'hypertrucage, la technologie de permutation intelligente de visages, ou les « *deepfakes* », ces termes réfèrent aux montages ultra réalistes générés par l'intelligence artificielle ayant comme objectif de tromper son public. Aujourd'hui, cette technologie n'est plus exclusive aux studios hollywoodiens puisqu'elle est accessible à n'importe qui. Conséquemment, des montages générés, on dépeint des personnalités publiques, des chefs d'État et plusieurs victimes non consensuelles. L'hypertrucage à finalité pornographique et non consensuelle a rapidement émergé, ciblant majoritairement des victimes de sexe féminin causant de sérieux préjudices. Toutefois, les auteurs priorisent les dangers publics tels que l'ingérence étrangère, la perte de confiance dans les institutions publiques et la désinformation politique. Ainsi, nous tentons de comprendre pourquoi la majorité des écrits focalisent sur ces risques politiques alors qu'en réalité la grande majorité des montages d'hypertrucage est de finalité pornographique non consensuelle. De plus, considérant que la *Loi électorale du Canada* semble protéger contre les dangers des montages d'hypertrucages ayant une finalité politique, alors que les victimes de montages pornographiques et non consensuelles sont toujours sans recours concret, nous proposons une solution multiapproche regroupant quatre volets : la législation, la sensibilisation, l'innovation et la collaboration.

INTRODUCTION

[179] La réputation d'un individu a toujours été source de conséquences pour une personne et aujourd'hui, à l'ère numérique, cette réalité est encore plus pertinente. Les licenciements motivés par des comportements inappropriés à l'extérieur du lieu de travail, mais rendus publics dus à l'Internet en attestent. Certains avancements technologiques tels que l'hypertrucage risque d'intensifier ce constat. Cette technologie permettant de générer des montages qui imitent des individus de manière si convaincante a permis aux réalisateurs cinématographiques de rajeunir un acteur, de garder l'anonymat pour éviter la persécution de régimes autoritaires (BERGERON et CARON, 2020, p. 214-215) et de faire jouer un acteur décédé depuis longtemps dans la série *Star Wars* (CITRON et CHESNEY, 2019, p. 1770). Cependant, cette technologie n'est pas exclusive aux studios hollywoodiens et elle peut maintenant être déployée par n'importe qui pour générer des montages qui dépeignent des personnes engagées dans des actes qu'ils n'ont jamais faits. Les algorithmes de l'hypertrucage sont mis au travail pour créer un contenu synthétique qui tente de tromper son public. Bien que les origines de l'intelligence artificielle remontent aux années 1950, notamment la publication d'Alan Turing intitulée *Can Machines Think* (TURING, 1950), l'avènement des montages hypertruqués est un phénomène récent qui a émergé en 2014 avec la création de réseaux adverses génératifs, mais rendus plus accessibles en 2017. Ainsi, l'hypertrucage est une problématique particulière aux générations numériques, qui sont habituées à consommer des informations par voie numérique. Ce texte tente de comprendre pourquoi la majorité des risques présentés sont axés sur la finalité politique, alors que la grande majorité des montages d'hypertrucage est pornographique et non consensuelle. Pour ce faire, nous aborderons premièrement la terminologie et la technique d'hypertrucage. Deuxièmement, nous survolerons certains des cas notoires quant à l'utilisation de ces montages. Troisièmement, nous aborderons les deux dimensions principales de l'hypertrucage, soit la dimension politique et la dimension pornographique afin de comprendre pourquoi les inquiétudes à l'égard des dangers politiques dominent. Quatrièmement, nous envisagerons certaines solutions qui ont été avancées par les auteurs canadiens. Finalement, nous proposerons une solution à quatre volets pour affronter l'hypertrucage pornographique non consensuel au pays.

1. TERMINOLOGIE ET TECHNIQUE D'HYPERTRUCAGE

1.1 TERMINOLOGIE

[180] Que ce soit la technologie d'hypertrucage, de « *permutation intelligente de visages* » (SIEKIERSKI, 2019, p. 1) ou son nom anglophone de « *deepfake* », ces termes sont souvent employés de façon interchangeable pour désigner des montages ultraréalistes, de nature trompeuse, produits par des systèmes d'IA. Pour ce faire, un dispositif d'IA permet de reproduire les détails intimes d'une personne en analysant des données d'entrées telles que des photos, des vidéos ou des enregistrements sonores

de cette dernière pour superposer ces détails sur un autre contenu numérique. Plus fréquemment, permuter le visage d'une personne sur le corps d'une autre. Aujourd'hui, ces montages comprennent tant la dimension vidéo qu'audio et sont même parfois exclusivement réalisés sous forme de montages audio pouvant causer d'énormes pertes pécuniaires[25]. Quoique la technologie pourrait être mobilisée pour générer des montages hypertruqués d'animaux, de personnages fictifs ou de célébrités décédées, de nos jours ces montages concernent majoritairement des personnes (JUDGE et KORHANI, 2021, p. 8) dans un objectif de tromper. Au Québec, le terme hypertrucage a été premièrement défini en 2019 par l'*Office québécois de la langue française* comme étant un « [p]rocédé de *manipulation audiovisuelle* qui recourt aux *algorithmes de l'apprentissage profond* pour créer des trucages ultraréalistes » (OFFICE QUÉBÉCOISE DE LA LANGUE FRANÇAISE, 2019). Certes, la tendance de manipuler le contenu multimédia n'est pas nouvelle. Bien avant l'arrivée de l'hypertrucage, des applications de multimédias existaient et ont été mobilisées pour créer du contenu trompeur (WESTERLUND, 2019, p. 40). Les termes « *cheapfakes* » ou « *shallow fakes* » sont employés pour faire référence aux montages plus élémentaires qui manipulent du contenu audiovisuel par le biais de ces logiciels multimédias (BERGERON et CARON, 2020, p. 210). Ce qui diffère les montages élémentaires des montages hypertruqués repose sur la technique utilisée pour générer le contenu réaliste.

1.2 TECHNIQUE

[181] Les montages d'hypertrucage sont créés à partir des dispositifs d'IA appartenant au domaine de l'apprentissage automatique, plus particulièrement le sous-domaine de l'apprentissage profond, qui tire son inspiration du fonctionnement du cerveau humain et tente de reproduire sa fonctionnalité à partir de processus algorithmiques (BERGERON et CARON, 2020, p. 210). Selon certains auteurs, la spécificité voire « *l'ingrédient magique* » (BERGERON et CARON, 2020, p. 210-211) de l'hypertrucage est l'utilisation d'algorithmes avancés nommés des réseaux adverses génératifs (generative adversial network) composés de deux réseaux de neurones artificiels, soit le générateur et le discriminateur, d'où le premier génère des montages et le second évalue le niveau de réalisme de ces montages afin de pouvoir améliorer, de façon autonome, sa capacité de générer des montages réalistes (Bergeron et Caron, 2020, p. 210-211). Ces dispositifs d'IA tentent de reconstituer synthétiquement les attributs d'une personne ciblée tels que ses mouvements, ses expressions faciales, ainsi que sa voix (BAILEY, BURKELL, DUNN, GOSSE et STEEVE, 2021, p. 252 ; JUDGE et KORHANI, 2021, p. 9 ; WESTERLUND, 2019, p. 40). Ce qui en résulte la création de montages ultra réalistes mettant en scène des individus entreprenant des actions ou disant des choses qui n'ont jamais réellement été faites ou dites (SIEKIERSKI, 2019 p. 1 ; WESTERLUND, 2019, p. 40). Étant donné que l'hypertrucage est le fruit des capacités technologiques informatiques, force est de constater que plus la performance de l'IA augmente, plus la qualité des montages croît, diminuant ainsi la capacité humaine de distinguer une vidéo réelle d'un montage hypertruqué (JUDGE et KORHANI, 2021, p. 7).

25 En 2019, il a été rapporté qu'un montage d'hypertrucage audio a été utilisé pour usurper la voix d'un haut dirigeant pour donner des instructions d'effectuer un transfert de fonds de 220 000 €. Voir Jesse Damiani, « A Voice Deepfake was Used to Scam a CEO Out of 243 000 $ » Forbes, en ligne : <https://www.forbes.com/sites/ jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/amp/>.

2. AVÈNEMENT ET CAS NOTOIRES

[182] Historiquement, l'hypertrucage est né d'une première publication vidéo sur la plateforme *Reddit* en novembre 2017, qui avait dépeint l'actrice Gal Gadot dans une vidéo engagée dans des actes sexuels avec son demi-frère (BERGERON et CARON, 2020, p. 209 ; GIESEKE, 2020, p. 1484). D'ailleurs, cette publication était en provenance d'un usager ayant comme pseudonyme le nom « *deepfakes* » d'où le nom anglais trouve son origine (GIESEKE, 2020, p. 1484-85 ; BERGERON et CARON 2020, p. 209). L'auteur de la publication originale soutenait avoir développé un algorithme pouvant superposer le visage de célébrités sur le corps d'actrices pornographiques dans des vidéos de nature sexuelle (BERGERON et CARON, 2020, p. 209-210 ; SEIKIERSKI, 2019, p. 1-2). Toutefois, quoique ce premier montage n'était pas sophistiqué (GIESEKE, 2020, p. 1484), une communauté d'usagers de la plateforme *Reddit* entièrement dédiée aux montages hypertruqués de célébrités féminines dépeintes dans des vidéos pornographiques s'est rapidement développée (GIESEKE, 2020, p.1484-85 ; CITRON et CHESNEY, 2019, p. 1763 ; BERGERON et CARON, 2020, p. 208-209). Ainsi, l'utilisateur au pseudonyme de deepfake avait rendu le code de son dispositif accessible sur le site Web d'hébergement de logiciels à source ouverte *GitHub* afin d'offrir ce code libre d'accès au public (BERGERON et CARON, 2020, p. 211). Conséquemment, le logiciel destiné à la création d'hypertrucage *Fakeapp* a émergé peu après (BERGERON et CARON, 2020, p. 211). En raison de l'accessibilité à la technologie, la communauté d'amateurs d'hypertrucage pornographique sur *Reddit* aurait atteint 90 000 membres quelques mois après la publication originale (WESTERLUND, 2019, p. 41-42 ; GIESEKE, 2020, p. 1485). Cependant, le grand public a commencé à prendre connaissance de l'hypertrucage en 2018, lorsqu'un montage de Barack Obama a émergé afin de mettre en lumière le phénomène de désinformation et les dangers de la technologie d'hypertrucage (WESTERLUND, 2019, p. 43). Ce montage était le fruit des chercheurs qui avaient créé un dispositif de réseaux de neurones artificiels permettant de faire parler des vidéos comme on le désire (CITRON et CHESNEY, 2019, p. 1760 ; SIEKIERSKI, 2019, p. 1). Le montage présentait Obama qui critiquait et insultait Donald Trump, le président des États-Unis de l'époque, afin d'avertir le grand public des dangers de l'hypertrucage (HALL, 2018, p. 60). Toujours en 2018, un autre montage hypertruqué a dépeint Donald Trump qui critiquait l'*Accord de Paris*, toutefois, ce montage était créé par le parti politique *Socialistische Partij* Anders pour des fins publicitaires afin de mobiliser le public belge à signer une pétition concernant les changements climatiques (HALL, 2018, p. 60 ; SCHWARTZ, 2018). D'ailleurs, le parti politique a tenté de justifier la création de cette campagne publicitaire en soutenant que la mauvaise qualité du montage aurait dû alerter le public quant à sa véracité (SCHWARTZ, 2018). En 2019, Nancy Pelosi, présidente de la Chambre des représentants des États-Unis, fut la victime d'un montage qui la présentait en ayant des difficultés à articuler ses mots. Malgré son niveau de sophistication élémentaire, ce montage a tout de même alimenté des courants de désinformation quant à des prétentions que Pelosi fût dans un état d'ivresse, ou dans un état de santé détériorée (DIXON JR, 2019, p. 35 ; CARON et BERGERON, 2020, p. 217-218 ; CBS NEWS, 2019). En 2020, la Première ministre

belge, Sophie Wilmès, fut la cible d'un montage. Le groupe Extinction Rebellion Belgium a publié un montage d'hypertrucage qui dépeint la Première ministre attribuant l'émergence du virus Covid-19 aux conséquences de la crise écologique (BERGERON et CARON, 2020, p. 217). Or, jusqu'à présent, la majorité des cas notoires entourant l'hypertrucage a dépeint des acteurs politiques. Ainsi, ces cas notoires à finalité politique sont-ils représentatifs de la majorité des montages d'hypertrucage retrouvés sur Internet ?

3. FINALITÉ PORNOGRAPHIQUE OU POLITIQUE DE L'HYPERTRUCAGE ?

[183] Les cas notoires d'hypertrucage abordés dans la section précédente créent un mirage quant à la finalité politique de ces montages, puisqu'en réalité les montages ont majoritairement une finalité pornographique. Heureusement, la doctrine canadienne a pris connaissance de cette réalité et certains auteurs ont proposé une approche axée sur les droits de la personnalité. Or deux types de préjudices pouvant découler de l'hypertrucage ont été identifiés, soit les préjudices privés et les préjudices publics (JUDGE et KORHANI, 2021, p. 12-15). La présente section abordera la finalité pornographique (3.1) et politique (3.2) des montages d'hypertrucage en s'appuyant sur ces deux types de préjudices potentiels afin de comprendre pourquoi les écrits entourant la technologie sont largement dominés par les dangers d'ingérence politique ou d'atteinte à la démocratie (3.3).

3.1 LA FINALITÉ PORNOGRAPHIQUE DE L'HYPERTRUCAGE

[184] Les préjudices privés sont subis sur le plan individuel, notamment par les victimes d'hypertrucage (JUDGE et KORHANI, 2021, p. 13). Ils portent atteinte à la réputation d'une victime, à la vie privée, à l'autonomie et à la dignité. Selon les auteurs les remèdes pour ce type de préjudice se réfèrent généralement à des dispositions législatives sur la diffamation, la vie privée ou la cyberintimidation, ainsi que les remèdes civils entourant la perte de revenu ou la perte d'emploi (JUDGE et KORHANI, 2021, p. 13 et 14). Toutefois, il est soutenu qu'en dépit de ces remèdes dits *potentiels*[26], il y a très peu de recours pour les victimes de préjudices privés d'un montage hypertruqué (JUDGE et KORHANI, 2021, p. 14). À l'heure actuelle, les victimes doivent donc mobiliser différents recours et tenter d'appliquer ces derniers à l'hypertrucage pornographique et non consensuel (BAILEY et al., 2021, p. 258). Quoique les préjudices privés qui découlent d'un tel montage varient, ils sont énormément préjudiciables. Sur le plan personnel, les victimes perdent le contrôle de leur image et de leur expression sexuelle sans aucune possibilité de consentir, et conséquemment, les victimes sont soumises à des sentiments tels que la peur, l'anxiété, la dépression, l'humiliation et la perte totale du droit à la vie privée (BAILEY et al., 2021, p. 253). En ce qui a trait à la sécurité des victimes, il est soutenu que les victimes risquent d'être soumises à du harcèlement subséquent découlant d'une publication d'hypertrucage à caractère sexuel (BAILEY et al., 2021, p. 253). Sur le plan financier, les montages d'hypertrucage pornographiques non consensuels sont tout aussi néfastes pour la

26 Puisqu'il n'y a pas officiellement de remèdes législatifs dédiés à l'hypertrucage et que les remèdes mis de l'avant dans la doctrine n'ont pas encore été confirmés par les tribunaux à l'égard des montages d'hypertrucage, nous employons l'adjectif potentiel.

santé financière des victimes, car à l'ère d'Internet, il est difficile d'échapper à un contenu numérique préjudiciable (DUNN et PETRICONE-WESTWOOD, 2018), et ce, peu importe son niveau de véracité. De plus, considérant qu'il est estimé que 90 % des employeurs effectuent des recherches sur Internet lors du processus d'embauche d'un candidat (CITRON et CHESNEY, 2019 p. 1774-75), il n'est pas difficile d'envisager comment un tel montage pourrait ruiner la carrière professionnelle d'une victime. Tenant compte de ces dangers et de l'apparence d'un vide juridique, est-il possible que les montages d'hypertrucage pornographiques non consensuels soient très peu fréquents ? Le rapport de 2019 intitulé *Deeptrace the State of Deepfakes* (AJDER et AL, 2019. [*Rapport Deeptrace*]) (ci-après rapport Deeptrace) démontre l'accroissement rapide des montages hypertruqués durant une période d'environ 9 mois qui s'étend de 2018 et 2019 (AJDER et AL, 2019, p. 1). De plus, le rapport confirme que la grande majorité des montages d'hypertrucage sur Internet sont pornographiques et non consensuels (AJDER et AL, 2019, p. 1) et que ce contenu cible presque exclusivement les femmes. Par exemple, selon ledit rapport de 2019, la mobilisation de cette technologie de permutation intelligente de visages sur des sites Web pornographiques cible des femmes puisque 100 % des victimes étaient de sexe féminin (AJDER et AL, 2019, p. 2). Plus particulièrement, les femmes américaines et les femmes sud-coréennes comptent pour 66 % des victimes de montages hypertruqués pornographiques non consensuels retrouvés sur des sites pornographiques (AJDER et AL, 2019, p. 2 ; BAILEY et AL, 2021, p. 253 ; JUDGE et KOHANI, 2021, p. 13). En outre, les personnes employées par l'industrie du divertissement comptent pour 99 % des victimes sur ces mêmes sites Web. Par ailleurs, le rapport mentionne que l'augmentation de sites Web dédiés à l'hypertrucage pornographique qui comporte des publicités semble indiquer qu'une industrie entourant le contenu hypertruqué pornographique est en plein essor (JUDGE et KOHANI, 2021, p. 6). Toutefois, les effets préjudiciables de cette industrie sont disproportionnés entre les sexes. À titre d'exemple, l'application *DeepNude*, associée à la création d'hypertrucage, qui génère du contenu synthétique de nature sexuelle à partir d'images d'entrées, n'arrive pas à générer des corps nus de sexe masculin et, conséquemment, cible exclusivement des victimes de sexe féminin (GIESEKE, 2020, p. 1482). Considérant que l'IA est largement dépendante des données d'entrée et que la quantité de données est primordiale, est-il possible que cette technologie soit uniquement mobilisée contre les personnalités publiques, d'où le contenu numérique de ces derniers est abondant ? Aujourd'hui, un montage hypertruqué peut être créé en utilisant une seule image d'entrée, tel qu'un égoportrait (« *selfie* ») (WESTERLUND, 2019, p. 41) et si jamais une grande quantité d'images d'entrée était nécessaire, il existe aussi des outils permettant de télécharger l'entièreté des images d'un abonné d'un réseau social donné (HARRIS, 2019 p. 101). De plus, des applications Web ont été conçues pour identifier des actrices pornographiques ayant des traits comparables à une victime potentielle, afin que les usagers puissent générer des montages hypertruqués encore plus réalistes (GIESEKE, 2020, p. 1488). Bien que lors des premières publications sur *Reddit* les montages ciblaient des célébrités, aujourd'hui l'utilisation de l'hypertrucage pornographique cible de simples citoyens dans un objectif de tromper et de donner l'apparence d'authenticité (BAILY et al., 2021, p. 252-53). En somme, considérant l'accessibilité des dispositifs pour la création d'hypertrucage, les outils et les applications développés pour assister à la création de montages, ainsi que la popularité des réseaux sociaux qui hébergent plusieurs photos de ces utilisateurs, ces montages peuvent effectivement cibler la grande majorité de la population.

3.2 LA FINALITÉ POLITIQUE DE L'HYPERTRUCAGE

[185] Le deuxième préjudice découlant des montages d'hypertrucages identifiés dans la doctrine est celui de nature publique. Les préjudices publics minent la confiance publique dans les institutions étatiques et ils sont très souvent examinés sous l'angle électoral. Selon les auteurs, les remèdes à ces préjudices se trouvent soit dans des dispositions pénales telles que celles à l'égard de la cybersécurité et l'inférence étrangère, ou soit dans la modération des plateformes numériques (JUDGE et KORHANI, 2021, p. 12-13). Du côté de la doctrine américaine, les auteurs parlent plutôt de préjudices sociétaux « *Harm to Society* » (CITRON et CHESNEY, 2019, p. 1776-1786), qui font référence aux risques tels que la distorsion des discours démocratiques, la manipulation électorale, la perte de confiance citoyenne dans les institutions publiques, l'amplification des divisions sociales, les dangers à la sécurité publique, les conséquences pouvant miner les relations diplomatiques ainsi que la sécurité nationale (CITRON et CHESNEY, 2019, p. 1776-1786). Ces préjudices sont tous intimement liés aux conséquences du phénomène de désinformation. Alors qu'en est-il des montages hypertruqués à finalité politique : sont-ils tout aussi présents que la désinformation dans l'écosystème numérique ? Selon le rapport *Deeptrace*, en 2019 les montages d'hypertrucages politiques sur YouTube représentaient 12 % de l'hypertrucage sur ce site Web (AJDER et al., 2019, p. 2). Toutefois, ledit rapport confirme que la grande majorité de la couverture médiatique a focalisé sur ces préjudices publics, notamment les dangers aux processus démocratiques et l'éventuelle combinaison de montages hypertruqués et les cyberattaques (AJDER et al., 2019, p. 9).

[186] Au Canada, les inquiétudes associées aux préjudices publics à l'égard de l'hypertrucage ont été soulevées dès 2019 par une publication parlementaire sur l'hypertrucage (SIEKIERSKI, 2019) ainsi que dans un rapport du *Centre de la sécurité et des télécommunications* (CENTRE DE LA SÉCURITÉ ET DES TÉLÉCOMMUNICATIONS, 2019). Ledit rapport met de l'avant que « les adversaires » peuvent mobiliser la technologie d'hypertrucage afin de menacer le processus démocratique au Canada, et ce avec souplesse (BERGERON et CARON, 2020 p. 250). D'ailleurs, le rapport soutient que 48 % des Canadiens se servent des médias sociaux comme une source d'actualité (CENTRE DE LA SÉCURITÉ ET DES TÉLÉCOMMUNICATIONS, 2019, p. 10), ouvrant la porte aux risques de désinformation par du contenu numérique trompeur. Plus récemment, un rapport de juillet 2021 énonce que 94 % des Canadiens ont, au moins, un compte sur une plateforme de réseaux sociaux (CENTRE DE LA SÉCURITÉ ET DES TÉLÉCOMMUNICATIONS, 2021, p. 10). Il n'est donc pas surprenant qu'une modification législative à la *Loi électorale du Canada* (*Loi électorale du Canada* (L.C. 2000, ch. 9)) soit survenue en 2018, prévoyant à l'article 91(1) (*Loi électorale du Canada* (L.C. 2000, ch. 9), art 91 (1)) une disposition pour les fausses déclarations dans le but d'influencer le processus électoral, ainsi qu'une disposition à l'article 481(1) de cette loi (*Loi électorale du Canada* (L.C. 2000, ch. 9), art 48 (1)) à l'égard de la distribution, la transmission ou la production de publication trompeuse sans autorisation ou dans l'intention de tromper. De plus, la Loi électorale du Canada prévoyait déjà une disposition qui a été ajoutée en 2014 à l'article 480.1 (*Loi électorale du Canada* (L.C. 2000, ch. 9), art 480.1) concernant l'usurpation de qualité avec l'intention de tromper. En ce qui a trait aux dangers

démocratiques, la publication parlementaire de 2019 soutient que le Canada est protégé contre les montages hypertruqués à finalité politique (SIEKIERSKI, 2019, p. 3-4). Cependant, les trois dispositions électorales mentionnées ci-dessus constituent l'entièreté du contenu législatif canadien à l'égard de l'hypertrucage politique (BERGERON et CARON, 2020, p. 251).

[187] Du côté de la doctrine américaine, les soucis politiques des montages d'hypertrucage sont nombreux et dominent la littérature américaine entourant cette technologie. Certains auteurs considèrent qu'en raison des biais de confirmation dans l'ère des réseaux sociaux, l'hypertrucage qui s'attaque au processus gouvernemental représente le préjudice le plus effrayant associé à l'utilisation de cette technologie (HALL, 2018, p. 58-59). De plus, il est soutenu que la présidence de Donald Trump n'a qu'intensifié cette tendance (HALL, 2018, p. 58-59 ; CITRON et CHESNY, 2019, p. 1786). D'ailleurs rappelons-nous que le montage de Pelosi avait été publié par ce dernier. Or, certains sénateurs américains ont présenté les montages d'hypertrucage comme une arme politique plausible capable de déstabiliser le processus électoral américain (HALL, 2018, p. 59). Ainsi, les dangers politiques ont mobilisé le gouvernement américain pour consacrer un budget considérable. Le *National Defense Authorization Act for Fiscal Year 2020*, la première loi fédérale américaine entourant l'hypertrucage, a réservé un budget de 738 millions de dollars afin de produire, entre autres, un rapport sur la militarisation étrangère de l'hypertrucage, une demande de notifier le Congrès des activités étrangères de désinformation mobilisant l'hypertrucage et un concours pour stimuler la recherche ou la commercialisation de technologies pour détecter les montages hypertruqués (BERGERON et CARON, 2020, p. 251-252).

3.3 POURQUOI LA PRIORISATION POLITIQUE ?

[188] Considérant que la majorité des montages d'hypertrucage sont à finalité pornographique et ne posent aucun danger pour la démocratie, pourquoi est-ce que les craintes à l'égard de ces montages sont majoritairement fixées sur les dangers politiques ? Certains auteurs soutiennent que les montages d'hypertrucage sont survenus en 2017, soit durant la présidence de Donald Trump qui a été marquée par des discours de fausses nouvelles et du scandale techno politique de *Facebook Cambridge Analytica* intensifiant les inquiétudes politiques à l'égard de la technologie (BERGERON et CARON, 2020, p. 217). Quoique nous observons ces constatations, nous souhaitons ajouter d'autres éléments servant à expliquer la fixation des préjudices publics causés par la technologie. Premièrement, les valeurs démocratiques sont centrales sur la scène internationale. À titre d'exemple, l'article 21 de la *Déclaration universelle des droits de l'Homme* prévoit l'expression d'élections honnêtes, ainsi que l'article 29(2) portant sur l'exercice des droits dans une « société démocratique » (Déclaration universelle des droits de l'homme, 1948). La notion de démocratie est subordonnée au processus électoral, donc le processus électoral mérite le plus haut niveau de protection à l'abri de toute forme d'ingérence. Deuxièmement, nonobstant le fait que les discours de Donald Trump étaient centrés sur cette idée de fausses nouvelles qui alimentaient les courants de désinformation, des questions entourant l'ingérence étrangère ont persisté tout au long de son mandat (JOURNAL DE MONTRÉAL, 2017 ; LE MONDE, 2017). Troisièmement, tel que nous l'avons soulevé dans une section précédente, la majorité des cas notoires usurpe des individus qui occupent des postes politiques de grande influence. De plus, la couverture médiatique

aborde principalement les préjudices publics découlant des montages politiques. D'ailleurs, il est soutenu que les montages d'hypertrucage les plus renommés sont ceux qui imitent des personnalités publiques. Conséquemment les inquiétudes quant à l'utilisation de la technologie sont plus souvent axées sur les montages trompeurs de ces personnalités (JUDGE et KORHANI, 2021, p. 3). Quatrièmement, l'hypertrucage est survenu à un moment où l'écosystème numérique est fertile pour la désinformation créant de véritables difficultés à modérer les contenus qui propagent de fausses informations. Considérant que la technologie de détection de montages hypertruqués est largement inefficace (CITRON et CHESNEY, 2019, p. 1788) puisqu'elle est en retard sur la technologie qui génère les montages (CITRON et CHESNEY, 2019, p. 1788 ; BERGERON et CARON, 2020, p. 213), les dangers démocratiques d'une population manipulée par un montage sont vraisemblablement envisageables. D'ailleurs, à titre d'exemple du danger lié à la désinformation, les adeptes de l'ancien président Trump ont récemment été mobilisés par une conspiration qui a mené à l'assaut du Capitole. Certes, si les auteurs de courants conspirationnistes étaient dotés de cette technologie de permutation intelligente de visages, l'étendue des conséquences pourrait réellement être infinie. Ainsi, les préjudices publics sont plus étudiés et conséquemment mieux encadrés. En revanche, l'utilisation de ces montages pour influencer l'opinion politique n'est pas autant répandue tel que soulevé dans les écrits (DUNN, 2020)[27]. Le climat politique américain expliquerait alors pourquoi les écrits doctrinaux américains sont fixés sur les implications politiques des montages d'hypertrucage, alors que les montages sont majoritairement pornographiques et non consensuels (DUNN, 2020 ; GIESEKE, 2020, p. 1482) et se trouvent sur des sites Web entièrement dédiés à ce type de contenu (AJDER et al., 2019 [*Rapport Deeptrace*]).

4. ÉTAT DU DROIT À L'ÉGARD DE L'HYPERTRUCAGE

[189] Jusqu'à présent, il n'existe aucun régime juridique au Canada qui traite, de manière expresse, les préjudices privés de l'hypertrucage pornographique non consensuel (JUDGE et KORHANI, 2022, p. 29), alors que pour les préjudices publics, le Canada a déjà adopté des dispositions législatives qui sembleraient être à la hauteur de l'hypertrucage ciblant le processus électoral (SIEKIERSKI, 2019, p. 3-4). Les auteurs proposent donc des remèdes déjà existants aux montages hypertruqués à finalité pornographique et non consensuelle qui varient grandement. La présente section survolera certaines solutions présentées pour remédier aux préjudices privés tels que les droits de la personnalité et la vie privée (4.1) ainsi que ceux du droit pénal (4.2).

4.1 LES DROITS DE LA PERSONNALITÉ ET DE LA VIE PRIVÉE

[190] Au Québec, les droits de la personnalité sont codifiés dans le *Code civil du Québec* (ci-après Code civil) (*Code civil du Québec*, RLRQ c CCQ-1991) et font partie de la *Charte des droits et libertés de la personn*e (ci-après Charte du Québec) (*Charte des droit et libertés des personnes*, RLRQ c-10). Pour le Code civil, le premier alinéa de l'article 3 énonce que « [t]oute personne est titulaire de droits de la personnalité, tel le droit à la vie, à l'inviolabilité et à l'intégrité de sa personne, au respect de son nom, de sa réputation et de sa vie privée ». De plus, l'article 35 du Code civil énonce que

27 De 00h:19m:35s à 00h :19m:55s.

« [t]oute personne a droit au respect de sa réputation et de sa vie privée » et l'article 36 du Code civil indique les situations qui « [p]euvent être notamment considérées comme des atteintes à la vie privée d'une personne […] » tel que le fait d'« [u]tiliser son nom, son image, sa ressemblance ou sa voix à toute autre fin que l'information légitime du public » (*Code civil du Québec*, RLRQ c CCQ-1991). *A priori*, les montages d'hypertrucage pornographique non consensuel semblent constituer une utilisation de l'image d'autrui qui est attentatoire *au respect de son nom, de sa réputation et de sa vie privée* des droits de la personnalité de l'article 3 du Code civil, cela demeure hypothétique puisque les tribunaux n'ont pas encore fait face aux montages d'hypertrucage (BERGERON et CARON, 2020, à la p. 224). Toutefois, ces derniers semblent avoir déjà été utilisés pour falsifier des éléments de preuves en litiges familiaux (ALEXANDER, 2021).

[191] En ce qui concerne la Charte québécoise, le droit à la vie privée est reconnu comme un droit fondamental de l'article 5 de la Charte du Québec qui énonce que « [t]oute personne a droit au respect de sa vie privée » (*Charte des droit et libertés des personnes*, RLRQ c-10, art 5). De plus, dans l'arrêt *Aubry c. Éditions Vice-Versa* (*Aubry c. Éditions Vice-Versa*, [1998] 1 R.C.S. 591), la Cour Suprême du Canada a expressément reconnu le droit à l'image comme une composante du droit à la vie privée de l'article 5 de la Charte du Québec (BENYEKHLEF et DÉZIEL, 2018, p. 547). Cependant, cet arrêt comportait une dimension économique, notamment l'utilisation non autorisée d'une image pour des fins commerciales. Par ailleurs, certains auteurs soulèvent qu'un montage hypertruqué pourrait ne pas constituer une atteinte à la vie privée en raison de son caractère faux qui ne révèle pas une information privée d'une victime au grand public (BERGERON et CARON, 2020, p. 225). L'utilisation sans consentement de photos, de vidéos ou autres types de renseignements personnels, de nature biométrique, afin d'entraîner un système d'IA à produire un montage, serait pourtant susceptible d'être considérée comme une violation à l'égard des lois de la protection des renseignements personnels (BERGERON et CARON, 2020, p. 226). Cependant, ces analyses demeurent hypothétiques car elles n'ont pas encore été produites devant les tribunaux.

[192] Pour la Common law canadienne, il est soutenu que le délit le plus approprié pour traiter l'hypertrucage pornographique est celui de l'appropriation de la personnalité (JUDGE et KORHANI, 2022, p. 29). Ce délit comprend aussi un élément commercial qui découle de l'utilisation d'une image d'un individu sans son consentement pour des fins commerciales (JUDGE et KORHANI, 2022, p. 38). Historiquement, le délit d'appropriation de la personnalité a été reconnu par la Cour d'appel de l'Ontario en 1973 dans l'affaire *Krouse c. Chrysler Canada Ltd* (*Krouse v. Chrysler Canada Ltd. et al.*, 1973 CanLII 574 (ON CA)). Dans cette affaire, la nature publique de la personnalité de Bob Krouse était centrale à l'instance puisque ce dernier était un ancien joueur de football américain, dont la personnalité avait été utilisée pour des fins publicitaires sans son consentement. Cependant, le délit d'appropriation de la personnalité s'applique tout de même à des personnes n'ayant pas de personnalité publique, et donc il serait possiblement applicable aux victimes d'un montage hypertruqué sous réserve que l'élément d'exploitation commerciale de la personnalité soit rempli (JUDGE et KORHANI, 2022, p. 38-39). En ce qui concerne l'aspect commercial du délit, les auteurs soutiennent qu'aujourd'hui le contenu généré par les utilisateurs représente une

valeur commerciale (JUDGE et KORHANI, 2022, p. 40) et que bien des plateformes numériques telles que *YouTube*, *Facebook*, *Instagram* et *Twitter* permettent de rentabiliser le contenu numérique de ses utilisateurs (JUDGE et KORHANI, 2022, p. 39) ce qui pourrait satisfaire au critère d'exploitation commerciale (JUDGE et KORHANI, 2022, p. 40). De plus, considérant les conclusions du rapport *Deeptrace* quant à l'industrie d'hypertrucage pornographique (AJDER et al., 2019, p. 6), notamment la publicité sur les sites dédiés à ce type de contenu, ce délit pourrait s'appliquer lorsqu'un montage pornographique et non consensuel est hébergé sur un site qui rentabilise ses contenus. Par ailleurs, l'affaire *Joseph v. Daniels* (Joseph v. Daniels, 1986 CanLII 1106 (BC SC)) de 1986 entourant une photo montrant le torse du demandeur est également mise de l'avant (JUDGE et KORHANI, 2022, p. 39). Cependant, dans cette affaire le tribunal a conclu qu'il n'y avait pas d'appropriation de la personnalité, car la photo en question ne permettait pas au public d'associer la photo à l'individualité du demandeur (JUDGE et KORHANI, 2022, p. 39). Considérant les montages d'hypertrucage qui s'approprient le visage d'une victime, le critère d'individualité identifiable au grand public du délit d'appropriation de la personnalité serait satisfait (JUDGE et KORHANI, 2022, p. 39), alors que les montages qui utilisent uniquement des images du corps d'une victime satisferaient difficilement ce critère. Nonobstant l'élément commercial nécessaire, ce délit serait applicable pour bien des victimes, mais pas l'entièreté puisque l'hypertrucage est une combinaison de deux différents contenus numériques. Par ailleurs, certains argumentent que le droit à la vie privée au Canada est mieux positionné pour traiter la question de montages d'hypertrucage pornographique non consensuel, car cinq provinces ont déjà légiféré des lois provinciales visant la pornodivulgation, soit l'Alberta, le Manitoba, la Nouvelle-Écosse, Terre-Neuve, Labrador et la Saskatchewan (BAILEY et al., 2021, p. 260-261). Ainsi, certains suggèrent d'amender ces lois provinciales pour qu'elles trouvent application à l'hypertrucage pornographique non consensuel (BAILEY et al., 2021, p. 260-261). De plus, s'agissant des autres provinces qui ne sont pas dotées de loi régissant la pornodivulgation, ces victimes pourraient intenter un recours en Common Law sous le nouveau délit de « *false light* » (BAILEY et al., 2021, p. 261, citant *Y. (V.M.) v. G. (S.H.)* (2019))[28] reconnu récemment en 2019 par la Cour supérieure de l'Ontario (BAILEY et al., 2021, p. 260-261).

4.2 LE CODE CRIMINEL DU CANADA[29]

[193] En ce qui concerne le Code criminel du Canada (ci-après C.cr.), plusieurs solutions sont aussi mises de l'avant. L'article 403 (*fraude à l'identité*) du C.cr.[30] pourrait s'appliquer s'il advenait qu'un juge confirme son application, notamment le volet identité des renseignements personnels à l'égard des montages d'hypertrucage (JUDGE et KORHANI, 2022, p. 18). Toutefois, les auteures soutiennent que la Couronne aura le défi de prouver que l'accusé a utilisé le montage pour gagner un avantage ou créer un

28 « The tort of false light protects people when content that would be highly offensive to a reasonable person is published about them and that content places them in a false light. Justice Kristjanson stated, unlike defamation, «[t]he wrong is in publicly representing someone, not as worse than they are, but as other than they are. The value at stake is respect for a person's privacy right to control the way they present themselves to the world ».

29 *Code criminel*, L.R.C. (1985), ch. C-46.

30 L'article 403 (1) du Code criminel énonce que « Commet une infraction quiconque, frauduleusement, se fait passer pour une autre personne, vivante ou morte : a) soit avec l'intention d'obtenir un avantage pour lui-même ou pour une autre personne; […] c) soit avec l'intention de causer un désavantage à la personne pour laquelle il se fait passer, ou à une autre personne […] ».

désavantage à la victime, ce qui se complique, lorsque nous considérons que bien des montages d'hypertrucage pornographiques et non consensuels servent à combler les désirs sexuels de leurs créateurs (BAILEY et al., 2021, p. 258). Toutefois, outre l'article 162.1 (*pornodivulgation*) que nous allons examiner ci-dessous, l'article 403 du C.cr., pourrait tout de même être pertinent dans une situation de pornodivulgation, où un motif de vengeance est établi puisque cela pourrait être considéré comme étant une action de *causer un désavantage à une personne* (*Code criminel*, L.R.C. (1985), ch. C-46, art 403(1)). Ailleurs, l'article 264 du C.cr. (*harcèlement criminel*), est aussi soulevé comme une disposition pouvant potentiellement s'appliquer aux montages hypertruqués pornographiques et non consensuels (BAILEY et al., 2021, p. 258-257). Toutefois, encore ici le destin de son application incombe au juge qui devra déterminer si un tel montage satisfait l'énoncé de l'article 264(1) du C.cr., qu'il « a pour effet de lui *faire raisonnablement craindre* – compte tenu du contexte – *pour sa sécurité* ou celle d'une de ses *connaissances* » (*Code criminel*, L.R.C. (1985), ch. C-46), art 264(1)) et l'énoncé du paragraphe 264(2) d) du C.cr. interdisant l'acte de « se comporter d'une manière menaçante à l'égard de cette personne » (*Code criminel*, L.R.C. (1985), ch. C-46), art 264(2)d) afin de considérer le montage comme un acte de harcèlement criminel. Toutefois, la doctrine affirme que l'application de ces deux dispositions à l'hypertrucage pornographique et non consensuel est incertaine et que le fardeau de la preuve reposant sur la Couronne dans une instance de harcèlement criminel est élevé (BAILEY et al., 2021, p. 257-258). Ailleurs, si un montage hypertruqué à caractère sexuel implique deux victimes mineures, tant le visage d'une victime superposée sur le corps de l'autre, l'article 163.1 du C.cr. (*Code criminel*, L.R.C. (1985), ch. C-46, art 163.1) portant sur la pornographie juvénile pourrait ainsi s'appliquer. Cependant, la doctrine est incertaine quant à son application s'il advenait que le visage d'une victime mineure est superposé sur un corps adulte ou l'inverse dans un montage à caractère sexuel (BAILEY et al., 2021, p. 259). Finalement, la grande majorité des auteurs canadiens pointent tous vers les modifications législatives récentes qui ont mené au nouvel article 162.1 du C.cr. (*Code criminel*, L.R.C. (1985), ch. C-46, art 162.1) (*pornodivulgation*), soit la distribution non consensuelle d'images intimes comme étant une disposition applicable à la problématique de l'hypertrucage pornographique non consensuel (BAILEY et al., 2021, p. 259 ; BERGERON et CARON, 2020, p. 246). Cependant, l'application de cet article semble faire l'objet d'un désaccord chez les auteurs. Certains auteurs, y compris l'auteur de la publication parlementaire, sembleraient conclure que l'article 162.1 du C.cr. est applicable aux montages d'hypertrucage pornographique non consensuel (SIEKIERSKI, 2019, p. 3 ; BERGERON et CARON, 2020, p. 246), alors qu'ailleurs la doctrine semble conclure que l'article ne s'appliquerait pas. Selon Bailey et al, puisque les montages d'hypertrucage pornographique non consensuel impliquent deux individus, soit les victimes dont le visage a été superposé sur un deuxième vidéo pornographique figurant une actrice pornographique consentante, la tâche de faire valoir que ces dernières portaient véritablement une atteinte raisonnable à la vie privée est quasi impossible et il incombera à la Couronne de le prouver (BAILEY et al., 2021, p. 259-260). En raison de ces problèmes potentiels d'application de l'article 162.1 du C.cr., un amendement pour inclure les victimes de montage d'hypertrucage à cet article, comme ce qui a été fait dans l'état américain de la Virginie et dans certains états Australiens, est suggéré (BAILEY et al., 2021, p. 260).

5. UNE SOLUTION À QUATRE VOLETS

[194] Jusqu'à présent, les législateurs du pays n'ont pas encore abordé la question de l'hypertrucage au Canada au détriment des victimes qui se trouvent sans aucun véritable recours conçu ou adapté pour affronter cette technologie émergente. Considérant la nature évolutive de la technologie et la difficulté de prévoir les problématiques pouvant émerger, nous proposons une solution à quatre volets, soit la législation (5.1), la sensibilisation (5.2), l'innovation (5.3) et la collaboration (5.4). Quoique chacun de ces volets puisse être considéré seul, le meilleur moyen de combattre l'hypertrucage est une combinaison de solutions telle que suggérée dans la présente section.

5.1 LÉGISLATION

[195] Premièrement, le volet législation fait appel aux législateurs fédéraux pour légiférer. Jusqu'à présent, l'avancement majeur concernant la réglementation de l'IA sur la scène internationale est la proposition de règlement par la Commission européenne pour le marché européen où nous pouvons constater que l'hypertrucage a été considéré. Figurant parmi les propositions législatives, l'article 52 (*Obligations de transparence pour certains systèmes d'IA*) cible les montages d'hypertrucage et exige aux utilisateurs de systèmes d'IA qui génèrent ces montages de prévenir l'audience de la nature modifiée du contenu (CE, *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*, COM/2021/206 final, art 52 (3)).

[196] Au Canada, l'élaboration d'un cadre législatif pour l'encadrement de l'IA a débuté, puisqu'en juin 2020 le projet de loi C-27 (Projet de loi C-27, 44e législature, 1re session, *Loi édictant la Loi sur la protection de la vie privée des consommateurs, la Loi sur le Tribunal de la protection des renseignements personnels et des données et la Loi sur l'intelligence artificielle et les données et apportant des modifications corrélatives et connexes à d'autres lois*) comportant une partie entièrement dédiée à l'IA (partie III) a été déposé et est maintenant à sa deuxième lecture (LEGISINFO, 2022). Cependant, le projet de loi demeure dans un état d'ébauche, car certaines notions primordiales telles que les critères pour qualifier un dispositif d'IA de *système à incidence élevée* de l'article 5 se réfèrent à un règlement qui n'a pas encore été publié (Projet de loi C-27, *Loi sur l'intelligence artificielle et les données* (Partie 3), art. 5). Ainsi, en ce moment, il est impossible d'analyser le contenu législatif pouvant protéger les Canadiens des effets néfastes de l'hypertrucage. Toutefois, nous croyons que les dispositifs qui servent à générer un contenu hypertruqué doivent être qualifiés de systèmes à incidence élevée selon les critères du règlement à venir.

[197] Cependant, pour l'instant la solution fréquemment mise de l'avant, malgré l'incertitude de son application, demeure l'article 162.1 du C.cr. portant sur la pornodivulgation. Force est de constater que cette disposition pénale est la plus analogue à la diffusion de montages hypertruqués pornographiques non consensuels. De ce fait, une modification législative à cette disposition nous semble la solution la plus logique et efficace. D'ailleurs, les propos de la juge Abella dans l'arrêt *A.B. c. Bragg*

Communications Inc (*A.B. c. Bragg Communications Inc.*, 2012 CSC 46, [2012] 2 R.C.S. 567) à l'égard des effets de « toxicité psychologique » de la cyberintimidation et la de propagation rapide dans l'univers numérique sont tout aussi pertinents dans le cadre des montages d'hypertrucages pornographiques non consensuels (*A.B. c. Bragg Communications Inc.*, 2012 CSC 46, [2012] 2 R.C.S. 567, para 20 à 22). D'ailleurs, l'ajout de l'article 162.1 au C.cr. était motivé par la volonté d'atténuer les préjudices privés qui découlent de la distribution d'images intimes, principalement ciblant les femmes et les filles (CED Criminal Law - Offences LXXXIII.1.(t); (ii) Criminal Law - Offences I LXXXIII - Sexual Offences 1 - Offences (t) - Publication of Intimate Image Without Consent (ii) - Elements of Offence, para §3013.1). Cette raison d'être semble faire écho aux préjudices privés causés par les montages d'hypertrucages pornographiques non consensuels et donc suit les motivations législatives. Par ailleurs, dans une affaire concernant la pornodivulgation, la Cour de justice de l'Ontario a rappelé les motivations parlementaires pour la création de cet article énonçant : « the growing trend to use technology as a tool of harassment and sexual exploitation, and the justice system's inability to respond because no offence existed at that time which addressed that type of conduct » (*R v. CP*, 2021 ONCJ 356, para 21). À notre humble avis, encore ici, ces soucis sont tout aussi pertinents à l'égard de l'hypertrucage pornographique non consensuel. Bien entendu, un des défis législatifs à un amendement à l'article 162.1 est d'assurer qu'il comprenne un langage qui assurera que son champ d'application ne devienne pas obsolète en raison de l'évolution rapide des systèmes d'IA.

[198] Finalement, le droit pénal comme solution aux montages d'hypertrucage n'est pas sans critiques. Selon certains, le droit pénal n'est pas une solution pour gérer les montages d'hypertrucage, car généralement les victimes sont plus préoccupées par le fait de faire retirer le contenu et de faire reconnaître le tort subi (DUNN, 2020, à 00h:37:m45s). Bien entendu, cette critique est certainement pertinente lorsque nous considérons une approche *ex post* à un préjudice découlant d'un montage hypertruqué. Toutefois, un des rôles essentiels du droit pénal est lié aux fonctions principales des peines telles que la dénonciation, la dissuasion et la réparation des torts causés qui sont énoncés à l'article 718 du C.cr. (*Code criminel*, L.R.C. (1985), ch. C-46, art 718). De plus, bien qu'il est soulevé que les réponses provenant du droit criminel à l'égard des crimes de nature sexuels peuvent laisser à désirer (DUNN, 2020, de 00h38:m:08s à 00h:38m:16s), plusieurs cybercrimes sont commis sur un territoire différent de celui du lieu des victimes, et de ce fait, une disposition dans le Code criminel canadien sur l'hypertrucage permettra l'application de la Convention sur la cybercriminalité (*Convention de Budapest*) ouvrant la porte à la possibilité d'engager le processus d'extradition de l'article 24 (*Convention sur la cybercriminalité*, Conseil de l'Europe, art 24), qui ne ferait que fortifier l'effet dissuasif à l'étranger. Quoique les auteurs proposent une diversité de remèdes civils pour combattre l'hypertrucage, nous ne partageons pas cette vision. Certes, les remèdes civils peuvent engendrer des dommages et intérêts pour les victimes qui intentent un recours, ces sommes doivent être accordées suivant un litige, mais cela impose des frais aux victimes pour faire valoir ces droits en justice. Par la suite, l'exécution d'un jugement civil dépend trop souvent de la volonté et de la situation financière du défendeur, de son emplacement et bien d'autres facteurs qui peuvent être générateurs d'obstacles additionnels aux victimes. Il est alors primordial que le Canada légifère afin de protéger les Canadiens

des préjudices qui découlent d'un montage d'hypertrucage pornographique non consensuel ce qui, selon nous, pourrait s'accomplir par un amendement à l'article 162.1 du C.cr. De plus, le cadre législatif du projet de loi C-27 devrait s'inspirer de la proposition européenne et consacrer une disposition à l'hypertrucage imposant une obligation de transparence analogue à l'article 52 de la proposition européenne.

5.2 SENSIBILISATION

[199] Deuxièmement, le volet sensibilisation fait appel tant aux chercheurs, aux universitaires, aux membres des médias qu'au gouvernement pour sensibiliser la population canadienne à l'égard des montages d'hypertrucage. Les générations numériques devraient être priorisées en raison de leurs interactions dans l'environnement numérique. Certes, une société mieux renseignée à l'égard de l'hypertrucage pourrait ainsi réduire sa vulnérabilité. Au Canada, les écrits ne font que commencer alors qu'aux États-Unis, les auteurs américains étudient déjà le phonème de désinformation de l'hypertrucage (WESTERLUND, 2020, p. 40). D'ailleurs, en raison d'une lacune doctrinale au Canada entourant la technologie, une publication parlementaire de 2019 intitulée *Hypertrucage - Que peut-on faire à propos du contenu audio et vidéo de synthèse ?* (SIEKIERSKI, 2019) s'est appuyée principalement sur trois courts articles publiés par des cabinets juridiques et un article doctrinal américain pour aborder la problématique de l'hypertrucage au Canada. Selon nous, cette constatation témoigne d'un vide doctrinal que représente l'hypertrucage au pays. Quoique la doctrine canadienne soit très mince, actuellement cette dernière est encore plus limitée en français. C'est ainsi que la communauté de chercheurs, d'universitaires et de journalistes devraient s'intéresser davantage aux dangers de cette technologie afin d'alerter le grand public canadien. De plus, les subventions gouvernementales à la recherche devraient donner priorité aux travaux sur l'hypertrucage pour stimuler les écrits canadiens.

5.3 INNOVATION

[200] Troisièmement, ce volet fait appel à la société civile pour trouver une solution au problème que pose l'identification des montages hypertruqués. Jusqu'à présent, certaines solutions ont été proposées telles que la fonction de hachage, servant à faciliter la tâche d'identification de données, utilisée avec la cryptomonnaie et le blockchain pour retracer les modifications de contenu (CHEIKOSMAN, HEWETT et GABRIEL, 2021), le blockchain pour héberger les actualités ou le contenu numérique (CHABAN, 2020), des caméras spécialisées qui téléchargent leur contenu sur le blockchain (PALMERO, 2020) et un site Web pouvant détecter les montages d'hypertrucage (GROH, s.d). Ailleurs, les chercheurs et concepteurs travaillent à développer des dispositifs pour identifier l'hypertrucage en se basant sur différentes composantes d'un montage tel que la réflexion de la lumière dans les yeux des sujets (RADIO-CANADA, 2021) ou les imperfections dans les empreintes digitales des images d'un montage (AGENCE FRANCE-PRESSE, 2021). Tout récemment, le dispositif *FakeCatcher*, développé par la firme américaine Intel, permettrait d'identifier les montages hypertruqués à un taux impressionnant de 96 % en se basant sur les flux sanguins des sujets (RADIO-CANADA, 2022). Bref, bien des solutions ont déjà été avancées, mais rappelons-nous que les dispositifs d'identification sont développés *ex post* et en mode rattrapage. Ainsi, le Canada devrait investir dans la recherche afin de

rattraper l'écart comme les États-Unis l'ont fait en consacrant un impressionnant budget pour stimuler la recherche ou la commercialisation de technologies pour détecter les montages hypertruqués en 2020.

5.4 COLLABORATION

[201] Finalement, la collaboration comme quatrième pilier d'une solution comprend deux types de collaboration, soit la collaboration entre les secteurs public et privé et la collaboration entre États. Pour la collaboration entre secteurs, il est primordial que le Gouvernement canadien collabore avec le secteur privé pour supporter et stimuler l'innovation pouvant mener à un système d'identification, tel que soulevé au volet précédent. Pour ce qui est de la collaboration étatique, outre la volonté étatique de collaborer pour affronter les montages hypertruqués, la collaboration entre États est tout aussi nécessaire afin d'être capable de réagir à la nature extraterritoriale de la cybercriminalité et, conséquemment, les montages d'hypertrucage pornographique non consensuel. Ainsi, les États devront collaborer entre eux à moins qu'ils se satisfassent uniquement des montages créés et distribués sur leur territoire, ce qui nous semble peu probable.

CONCLUSION

[202] Nous avons survolé la technologie de l'hypertrucage et la façon dont elle génère des montages qui dépeignent des personnes entreprenant des actions ou énonçant des paroles qui ne sont jamais réellement produites. Quoiqu'en théorie, l'hypertrucage peut être mobilisé pour déstabiliser le processus démocratique, la grande majorité des montages sont pornographiques et non consensuels. À l'heure actuelle, la législation canadienne serait à la hauteur pour protéger les Canadiens contre l'hypertrucage à finalité politique, alors qu'il n'existe aucun recours précis pour les montages pornographiques non consensuels. Depuis l'avènement de cette technologie, les écrits doctrinaux ont focalisé sur les risques politiques découlant de la technologie. Selon nous, la domination des écrits politiques s'explique en raison de plusieurs facteurs. L'importance associée à la démocratie sur la scène internationale, le climat politique de l'époque lorsque cette technologie a émergé, notamment les discours de fausses nouvelles et d'ingérence politique sont des facteurs qui expliquent la fixation sur les préjudices publics liés à la démocratie découlant de l'utilisation de la technologie. Puisque l'objectif de l'hypertrucage est de tromper son public, la technologie peut être considérée comme une nouvelle forme de désinformation cherchant à nuire à la réputation d'une personne ou des institutions publiques. Ainsi est venu le temps de se préoccuper des inquiétudes à l'égard des montages hypertruqués afin de considérer les victimes de préjudices privés qui subissent des montages d'hypertrucage pornographique non consensuel. Les préjudices privés qui découlent de ce type de montages sont hautement préjudiciables sur la vie privée des victimes, leur dignité et même leur sécurité financière. Plusieurs remèdes civils, tant du Code civil québécois que des délits de Common Law, sont suggérés, mais nous sommes d'avis que le droit criminel, plus particulièrement une modification législative à l'article 162.1 ciblant la pornodivulgation, permettrait mieux d'affronter les dangers de cette technologie. Finalement, le droit pénal pourrait créer un effet dissuasif et serait une solution à la hauteur de la problématique de l'extraterritorialité des auteurs de montages d'hypertrucage pornographique non consensuel.

[203] En conclusion, puisque les montages d'hypertrucage pornographique non consensuel sont hautement préjudiciables et qu'ils peuvent effectivement cibler la grande majorité des Canadiens, il est maintenant temps de passer à l'action, car à l'heure actuelle l'état du droit représente un vide juridique. De ce fait, nous proposons une solution multiapproche à quatre volets, soit la législation, la sensibilisions, l'innovation et la collaboration afin de mieux positionner le Canada vis-à-vis de l'hypertrucage.

BIBLIOGRAPHIE

Législation

Charte des droits et libertés des personnes, RLRQ c-10.

Code civil du Québec, RLRQ c CCQ-1991.

Code criminel (L.R.C. (1985), ch. C-46).

Loi électorale du Canada (L.C. 2000, ch. 9).

Projet de loi C-27, 44e législature, 1re session, *Loi édictant la Loi sur la protection de la vie privée des consommateurs, la Loi sur le Tribunal de la protection des renseignements personnels et des données et la Loi sur l'intelligence artificielle et les données et apportant des modifications corrélatives et connexes à d'autres lois*.

Législation étrangère

CE, *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*, COM/2021/206 final.

CE, *Règlement général sur la protection des données (UE)* 2016/676.

Nations Unies, *Déclaration universelle des droits de l'homme*, 1948.

Jurisprudence

A.B. c. Bragg Communications Inc., 2012 CSC 46, [2012] 2 R.C.S. 567.

Aubry c. Éditions Vice-Versa, [1998] 1 R.C.S. 591.

Joseph v. Daniels, 1986 CanLII 1106 (BC SC).

Krouse v. Chrysler Canada Ltd. et al., 1973 CanLII 574 (ON CA).

Doctrine : Monographies

BAILEY, J., BURKELL, J., DUNN, S., GOSSE, C. et STEEVE, V., « Chapter 10 AI and Technology-Facilitated Violence and Abuse », dans F. MARTIN-BARITEAU et T.

SCASSA, dir, *Artificial Intelligence an the Law in Canada*, Toronto, LexisNexis Canada, 2021.

BENYEKHLEF, K. et DÉZIEL, P-L, *Le droit à la vie privée en droit québécois et canadien*, Éditions Yvon Blais, Montréal, Thomson Reuters Canada Limitée, 2018.

CED Criminal Law — Offences LXXXIII.1.(t).(ii) Criminal Law — Offences I LXXXIII — Sexual Offences I 1 — Offences I (t) — Publication of Intimate Image Without Consent I (ii) — Elements of Offence.

Doctrine : Articles

BERGERON, V. et CARON, V., « Deepfake : distinguer le vrai du faux sur les implications juridiques d'une technologie "trompeuse" », dans BARREAU DU QUÉBEC, Service de la formation continue, *Développements récents en droit de la propriété intellectuelle*, 2020, V484, Montréal, Éditions Yvon Blais, p. 205-257.

CITRON, D. et CHESNEY, R., « Deep Fakes : A Looming Challenge for Privacy, Democracy, and National Security » (2019) *California Law Review*, vol. 107, p.1753-1820.

DIXON JR, H., « Deepfakes: More Frightening than Photoshop on Steroids », (2019) *The Judges Journal*, vol. 58, nº 3, p. 35-37.

DUNN, S. et PETRICONE-WESTWOOD, A., « More than "Revenge Porn" Civil Remedies for the Nonconsensual Distribution of Intimate Images », (2018) 38th *Annual Civil Litigation Conference*, vol. 16, p. 1-30.

GIESEKE, A.P., « The New Weapon of Choice : Law's Current Inability to Properly Address Deepfake Pornography », (2020) *Vanderbilt Law Review*, vol. 73, nº 5, p. 1479-1515.

HALL, H.K., « Deepfake Video : When seeing Isn't believing », (2018) *Catholic University Journal of Law and Technology*, vol. 27, nº 1, p. 51-76.

HARRIS, D., « Deepfakes: False Pornography Is Here and the Law Cannot Protect You », (2019) *Duke Law & Tecnology Review*, vol. 17, nº 1, p. 99-127.

JUDGE, E.F. et KORHANI, A.M., « Deepfakes, Counterfeits, and Personality » (2021) *Altaberta Law Review*, vol. 59.

TURING, A. M., « I.—Computing Machinery and Intelligence » *Mind*, vol. LIX, nº 236, October 1950, 433-460.

WESTERLUND, M., « The Emergence of Deepfake Technology: A Review », (2019) *Technology Innovation Management Review*, vol. 9, nº 11, p. 39-52.

YADLIN-SEGAL, A. et OPPENHEIM, Y., « Whose Dystopia is it Anyway ? Deepfakes and Social Media Regulation », (2021), *The International Journal of Research into New Media Technologies*, vol. 27, nº 1 p. 36-51.

Autres sources

AGENCE FRANCE-PRESSE, « *Facebook dit progresser dans la détection des hypertrucages* », 17 juin 2021, en ligne : (Radio-Canada), <https://ici.radio-canada.ca/nouvelle/1802382/facebook-hypertrucage-recherche-intelligence-artificielle-deepfake>.

AJDER, H., PATRINI, G., CAVALLI, F., CULLEN, L., « *The State of Deepfakes : Landscape, Threats, and Impact* », 2019, en ligne : <https://regmedia.co.uk/2019/10/08/deepfake_report.pdf>.

ALEXANDER, R., *What Family Judges Can Learn from Pornography*, 2021, en ligne (Canlii Connects) : <https://canliiconnects.org/en/commentaries/73749>.

CBS NEWS, « *Doctored Nancy Pelosi video highlights threat of "deepfake" tech* », 26 mai 2019, en ligne : <https://www.cbsnews.com/news/doctored-nancy-pelosi-video-highlights-threat-of-deepfake-tech-2019-05-25/>.

CENTRE CANADIEN POUR LA CYBERSÉCURITÉ, *Cybermenaces contre le processus démocratique du Canada : Mise à jour de juillet 2021*, en ligne : <https://www.cyber.gc.ca/fr/orientation/cybermenaces-contre-le-processus-democratique-du-canada-mise-jour-de-juillet-2021>.

CENTRE CANADIEN POUR LA CYBERSÉCURITÉ, *Le point sur les cybermenaces contre le processus démocratique du Canada en 2019*, à la p 10, en ligne : <https://www.cyber.gc.ca/fr/orientation/le-point-sur-les-cybermenaces-contre-le-processus-democratique-du-canada-en-2019>.

CHABAN, M. A.V., « Can Blockchain Block Fake News and Deep Fakes ? », le 30 novembre 2020, en ligne : *(IBM)*, <https://www.ibm.com/blogs/industries/blockchain-protection-fake-news-deep-fakes-safe-press/>

CHEIKOSMAN, E., HEWETT, N. et GABRIEL, K., « Blockchain Can Help Combat the Threat of deepfakes. Here's how », *We Forum*, 12 octobre 2021, en ligne : <https://www.weforum.org/agenda/2021/10/how-blockchain-can-help-combat-threat-of-deepfakes/>.

DAMIANI, J., « A Voice Deepfake Was Used to Scam a CEO out of $243,000 » : *(Forbes)*, en ligne : <https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/amp/>.

DUNN, S., « Suzie Dunn › Identity Manipulation : Responding to Advances in Artificial Intelligence and Robotics », 2020, *CDTS/CLTS UOTTAWA*, (à 00h:19m:35s de 00h:19m:55s, en ligne (vidéo) : <https://www.youtube.com/watch?v=wl13HRwrNNk>.

GROH, M., « Detect DeepFakes: How to counteract misinformation created by AI », s.d., en ligne : (MIT media Lab) <https://www.media.mit.edu/projects/detect-fakes/overview/>.

JOURNAL DE MONTRÉAL, « Trump suggère au Sénat d'enquêter sur les médias américains », 5 octobre 2017, en ligne : <https://www.journaldemontreal.com/2017/10/05/trump-suggere-au-senat-denqueter-sur-les-medias-americains-1>.

LEGISINFO, Sommaire C-27, 2022, en ligne : (Parlement du Canada), <https://www.parl.ca/LegisInfo/fr/projet-de-loi/44-1/c-27>.

LE MONDE, « Obama aurait été averti de l'ingérence russe dans la campagne présidentielle dès l'été 2016 », 23 juin 2017, en ligne : <https://www.lemonde.fr/ameriques/article/2017/06/23/obama-averti-de-l-ingerence-russe-des-l-ete-2016_5150298_3222.html>.

OFFICE QUÉBÉCOISE DE LA LANGUE FRANÇAISE, 2019, Fiche terminologique, hypertrucage, en ligne : <https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/26552557/hypertrucage>.

PALMERO, J., « Deepfakes Why you can't Believe Everything you See », *Ted X Talks*, 12 février 2020, en ligne : <https://www.youtube.com/watch?v=JXBBalHl_cI>.

RADIO-CANADA, « Des scientifiques ont trouvé un moyen pour déjouer les hypertrucages », 15 mars 2021, en ligne : <https://ici.radio-canada.ca/nouvelle/1777459/scientifiques-hypertrucages-web-systeme>.

RADIO-CANADA, « Intel dévoile un outil de détection des hypertrucages efficace à 96 % », 17 novembre 2022, en ligne : <https://ici.radio-canada.ca/nouvelle/1933488/intel-hypertrucages-outil-deepfake-fakecatcher-intelligence-artificielle>.

SCHWARTZ, O., « You Thought Fake News was Bad ? Deep Fakes are where Truth Goes to Die » *The Guardian*, 12 novembre 2018, en ligne : <https://www.theguardian.com/technology/2018/nov/12/deep-fakes-fake-news-truth>.

SIEKIERSKI, B.J., « Hypertrucage : Que peut-on faire à propos du contenu audio et vidéo de synthèse ? (en bref) », Bibliothèque du Parlement du Canada, publication nº 2019-11-F, Ottawa, 8 avril 2019.