

# Frontières du numérique et puissances du faux : le cas des hypertrucages (*deepfakes*)

Renée BOURASSA (1), Fabien RICHERT (2)

(1) CRILCQ, groupe de recherche ARCANES, Université Laval  
Renee.Bourassa@design.ulaval.ca

(2) CRILCQ, groupe de recherche ARCANES, Université du Québec à Montréal  
richert.fabien@uqam.ca



# Introduction

- Cette présentation s'inscrit dans le cadre des travaux menés au sein du projet de recherche ARCANES, financé par le Conseil de Recherches en Sciences humaines du Canada (CRSH) et composé d'un noyau principal de 4 chercheurs : Renée Bourassa (Université Laval, Canada), Jean-Marc Larrue (Université de Montréal, Canada), Samuel Szoniecky (Université Paris 8, France) et Fabien Richert (UQAM, Canada).
- Ce projet s'intéresse aux puissances du faux relatives à des stratégies d'illusion ou de tromperie telles qu'elles se manifestent dans les arts trompeurs ou dans les pratiques médiatiques de l'écosystème socionumérique actuel.
- Nous interrogeons les frontières du numérique en nous penchant sur les fabrications générées par les algorithmes génératifs de l'IA, soit les hypertrucages (deepfakes).



# Introduction

- D'une part, ces puissants dispositifs d'intelligence artificielle tracent de nouvelles frontières en devenant des moyens inédits de création algorithmique. Les *deepfakes* peuvent donc susciter des retombées positives lorsqu'elles sont mises au service de la création artistique ou de la médiation culturelle.
- D'autre part, ils contribuent à amplifier la crise de la vérité et de la confiance qui affecte les médiations dans nos sociétés. Ils peuvent déplacer les frontières poreuses entre le vrai et le faux en engageant des enjeux sociaux de taille.

## 1. Puissance du faux : le cas des hypertrucages (deepfakes)



The Chronicles of  
DeepTomCruise (Viral  
Tiktok Videos),  
VFXChris Ume

## 2. Multiplication des *deepfakes* dans la production cinématographique



The Irish Man, Martin Scorsese, 2019



Indiana Jones and the Dial of  
Destiny, James Mangold, 2023



### 3. Multiplication des *deepfakes* dans le secteur muséal



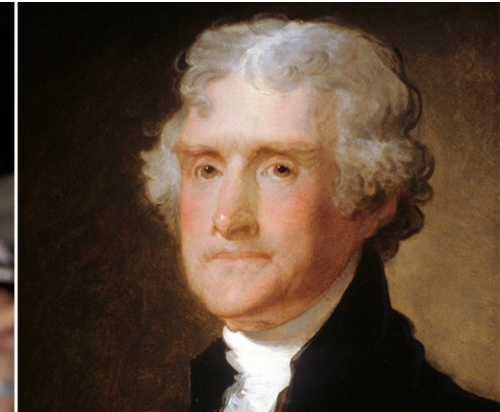
*La Jeune fille à la perle* (1665, Johannes Vermeer, collection permanente au Mauritshuis Museum de La Haye) fait scandale en mars 2023 lors d'un prêt temporaire au musée Rijksmuseum d'Amsterdam.

Outils de création d'image :  
MidJourney et Photoshop par  
Julian Van Dieken.

## 4. Puissance du faux, faiblesses du vrai : effets perturbateurs dans l'écosystème socionumérique

- La falsification associée aux médias ou leur dénonciation n'est pas un fait nouveau. En 1807, le président des États-Unis Thomas Jefferson écrivait :

« On ne peut plus croire ce que l'on voit dans un journal. La vérité elle-même devient suspecte en étant placée dans ce média pollué ».

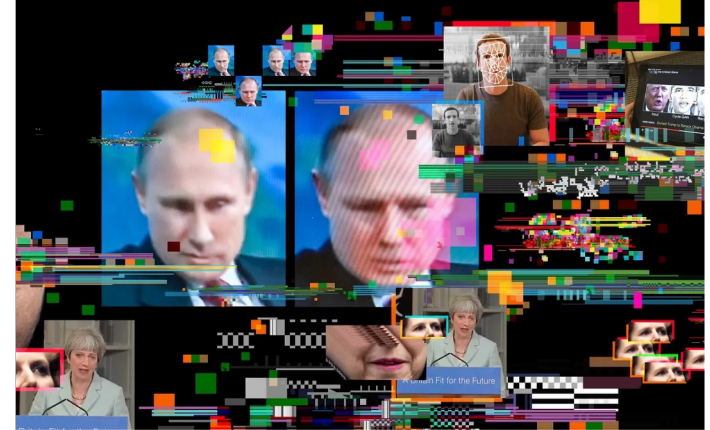


paper would find few subscribers. It is a melancholy truth, that a suppression of the press could not more completely deprive the nation of it's benefits, than is done by it's abandoned prostitution to falsehood. Nothing can now be believed which is seen in a newspaper. Truth itself becomes suspicious by being put into that polluted vehicle. The real extent of this state of misinformation is



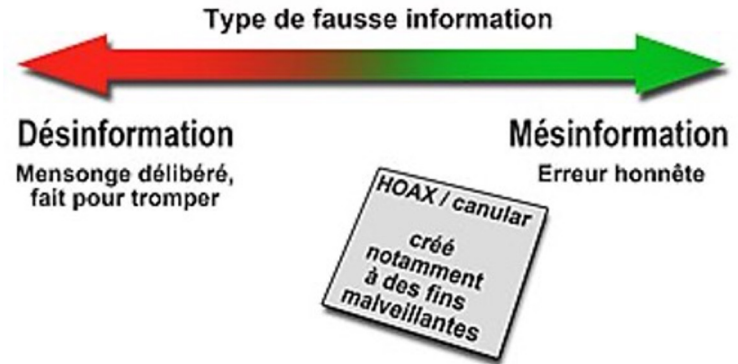
## 4. Puissance du faux, faiblesses du vrai : effets perturbateurs dans l'écosystème socionumérique

- Ces falsifications algorithmiques contribuent à nourrir les stratégies dites du «dividende du menteur » (Chesney et Citron, 2018) : toute information est fausse à moins de preuve du contraire ; puisque tout média (vidéo, audio, texte) peut être contrefait et que la tromperie est très difficile à détecter, cette situation jette le doute sur toutes les sources d'information et sur leur crédibilité, ce qui contribue à l'ère du soupçon généralisé.



## 4. Puissance du faux, faiblesses du vrai : effets perturbateurs dans l'écosystème socionumérique

- Désinformation : La désinformation est une information qui est fausse, et la personne qui la diffuse sait quelle est fausse.
- La mésinformation est une information fausse, mais que la personne qui la diffuse pense qu'elle est vraie.
- La malinformation est fondée sur la réalité, mais qui est utilisée afin de porter préjudice à une personne, une organisation, un pays.



## 5. Contrer les *deepfakes* : des solutions multi-dimensionnelles

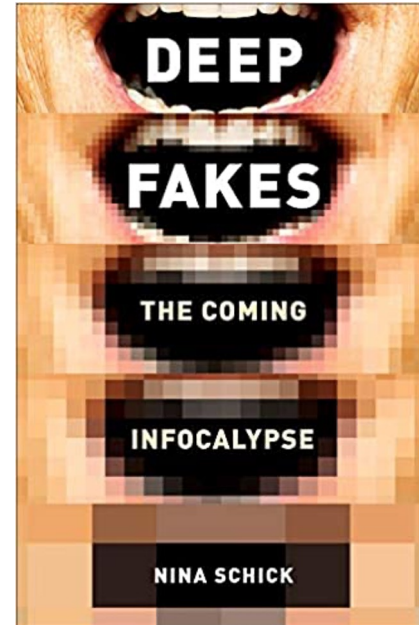
- Littéracie numérique et formation à l'esprit critique
- Action légale et politiques de gouvernance (plateformes, gouvernements)
- Intervention technologique



## 5. Contrer les *deepfakes* : des solutions multi-dimensionnelles

- Littératie numérique et formation à l'esprit critique :

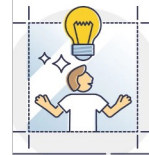
Développer la littératie informationnelle et numérique dans un contexte d'« infocalypse » (Schick, 2020) en tenant compte des biais cognitifs.



## 5. Contrer les *deepfakes* : des solutions multi-dimensionnelles

- Notre capacité de discriminer entre le vrai et le faux n'est pas neutre : elle repose largement sur des biais cognitifs, lesquels agissent comme des filtres la plupart du temps inconscients pour traiter l'information de toute nature qui nous parvient.
- Les biais cognitifs constituent des distorsions systématiques dans la pensée et le raisonnement, lesquels peuvent conduire à des erreurs de jugement et de prise de décision. Ils se manifestent par exemple lorsque la raison des individus est affectée par des émotions ou des préjugés qu'ils ne contrôlent pas.
- On peut même se demander s'il est possible d'échapper aux pièges à la fois perceptuels et sémiotiques tendus par notre propre cerveau.

### COGNITIVE BIASES



CONFIRMATION  
BIAS



LOSS  
AVERSION



GAMBLER'S  
FALLACY



AVAILABILITY  
CASCADE



FRAMING  
EFFECT



BANDWAGON  
EFFECT



DUNNING-KRUGER  
EFFECT

## 5. Contrer les *deepfakes* : des solutions multi-dimensionnelles

- Établir un cadre légal de régulation pour lutter contre l'escroquerie :

Aux États-Unis, le Digital Millenium Copyright Act (DCMA) permet à un plaignant de retirer un *deepfake* d'une plateforme.

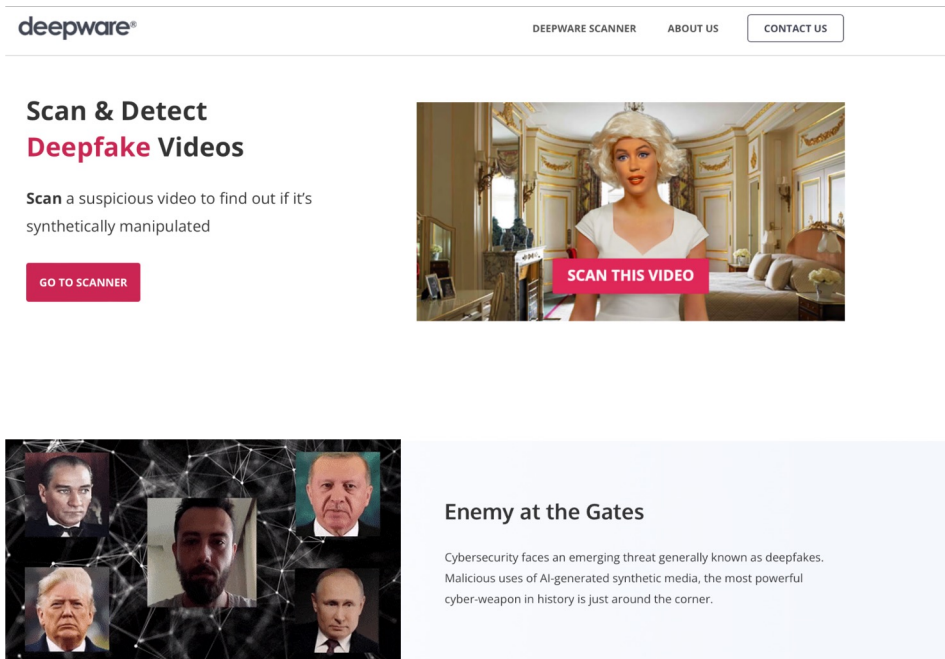
En Europe, le General Data Protection Regulation (GDPR) octroie le droit d'effacer ou d'être oublié. Cette législation s'applique dans le cas de *deepfakes* ou de représentations fictionnelles de toutes natures.



### 3. Contrer les *deepfakes* : des solutions multi-dimensionnelles

- Développer des solutions technologiques de détection

Utilisation de techniques et d'outils de détection comme les systèmes de vérification de l'authenticité du contenu, par l'identification de sa source et de son créateur, ou encore du média d'origine.



The screenshot shows the Deepware website. At the top, there's a navigation bar with the 'deepware' logo, links for 'DEEPWARE SCANNER', 'ABOUT US', and a 'CONTACT US' button. The main content area is titled 'Scan & Detect Deepfake Videos'. Below this, a text block says 'Scan a suspicious video to find out if it's synthetically manipulated' with a 'GO TO SCANNER' button. To the right is a video player showing a woman in a white dress in a luxurious room, with a 'SCAN THIS VIDEO' button overlaid. Below the main content, there's a section titled 'Enemy at the Gates' with a paragraph about deepfakes as a cybersecurity threat. At the bottom left, there's a collage of faces, including several of world leaders, overlaid with a network diagram.



## Conclusion

- Si les *deepfakes* ouvrent des potentialités créatives sans précédent pour l'art, l'éducation et la médiation culturelle, ce qui les légitime, ils ont également le pouvoir de bousculer nos rapports à la vérité en soulevant de nombreuses questions sociales, juridiques, éthiques et philosophiques. Ils mettent en jeu notre capacité à distinguer ce qui est faux de ce qui ne l'est pas.
- Afin de distinguer le réel de l'artifice et de statuer s'il s'agit d'une illusion créative ou d'une stratégie de tromperie dans le registre de la propagande, la question de l'intentionnalité est plus que jamais au cœur des questions éthiques et de la formation à l'esprit critique afin de déterminer si l'usage d'un média synthétique est malveillant ou non, entre usage culturel, désinformation, mésinformation et malinformation.





Rare photo de Mona Lisa et Leonardo da Vinci prise à Florence en 1504



# Bibliographie

- Bourassa, R. (2021). « Des arts trompeurs à la post-vérité : Puissances du faux et stratégies de tromperies », In H2ptm 2021. *Information : enjeux et nouveaux défis*, Saleh, I., Bouhaï, N., Leleu-Merviel, S., Roxin, I., Zacklad, M., Massou, L., (Eds.), Londres, ISTE éditions, 81-93.
- Bourassa, R. (2018). « Figures de l'être artificiel : du simulacre à la figure de synthèse ». In *Arts Trompeurs : Machines, Magie, médias*, Actes du Colloque Cerisy 2016, Kessler, F., Larrue, J.-M., Pisano, G. (Eds), Presses Universitaires du Septentrion, Paris.
- Bourassa, R. (2017). « Frictions de la fiction dans les imaginaires contemporains », In *L'ontologie du numérique. Entre mimesis et réalité*, Sens Public, Disponible à : <http://www.sens-public.org/article1279.html>
- Chadwick, J. (2022). See beyond the world's greatest works of art! AI artist DALLÉ can now imagine what's outside the frame of famous paintings including Girl with a Pearl Earring and Mona Lisa, Daily Mail, Disponible à : [www.dailymail.co.uk/sciencetech/article-11184387/AI-artist-imagine-whatsoutside-frame-famous-paintings-including-Girl-Pearl-Earring.html](http://www.dailymail.co.uk/sciencetech/article-11184387/AI-artist-imagine-whatsoutside-frame-famous-paintings-including-Girl-Pearl-Earring.html).
- Chesney, R. et Citron, D. (2019). « DeepFakes, A cooking Challenge for Privacy, Democracy, and National Security », *California Law Review*, vol. 107, num. 6.
- Gregory, S. (2022). Deepfakes, misinformation and disinformation and authenticity infrastructure responses: Impacts on frontline witnessing, distant witnessing, and civic journalism, In *Journalism*, vol. 22, num. 3, 708-729, Disponible à : [journals.sagepub.com/doi/pdf/10.1177/14648849211060644](https://journals.sagepub.com/doi/pdf/10.1177/14648849211060644).
- Holland,
- Holland, J.H. (1996). *Hidden Order: How Adaptation Builds Complexity*. Redwood City. Addison Wesley Longman, Boston.



# Bibliographie

- Jaiman, A. (2022). Deepfakes and Synthetic Media: Humanity at the Edge of an Uncanny Valley. Independently published, Illinois.
- Marreiros, G., Melo, F. S., Lau, N., Cardoso, H. L., Reis, L. P. (2021). Progress in artificial intelligence 20th epia conference on artificial, Springer, New-York.
- Mihailova, M. (2021). To Dally with Dali: Deepfake (Inter)faces on the Art Museum, In Convergence : The International Journal of Research into New Media Technologies, vol. 27, num. 4, 882-898.
- Naffi, N., et als (2021). Pour une éducation aux hypertrucages malveillants et un développement de l'agentivité dans les contextes numériques. In Éducation et francophonie, vol. 49, num. 2
- Schick, N. (2020). *Deepfakes: The Coming Infocalypse*. Grand Central Publishing, New-York.
- Tzeng, P. (2018). What can the Law do About Deepfakes? Disponible à <https://mcmillan.ca/insights/what-can-the-law-do-about-deepfake/>.