

Executive Master

Management des médias et du numérique

Promotion 2017/2018

Mémoire professionnel

« Deepfakes, une nouvelle arme de déception massive »

GÉRALD HOLUBOWICZ

Paris – le 19 Mars 2019

Remerciements	4
Introduction	6
1 Voir pour croire ou croire pour voir. L'image a-t-elle toujours représenté la réalité ?	9
1.1 Le Réalisme, le vraisemblable et l'objectivité de l'image	10
1.1.1 Qu'est-ce que le Réalisme ? L'exemple de Johannes Vermeer.....	10
1.1.2 Le vraisemblable, un début d'éloignement de la représentation réaliste	13
1.1.3 L'objectivité, une illusion cherchée par tous.....	15
1.2 La construction d'un réel collectif à travers l'image.....	21
1.2.1 La photographie, premier artisan de la représentation	22
1.2.2 L'image se met en mouvement.....	23
1.2.3 La manipulation par l'image devient industrielle	24
1.3 La propagande par l'image, le réel disparaît au profit du message	26
1.3.1 Deux camps, une même approche manipulateur	26
1.3.2 L'après guerre marque un tournant dans l'histoire du réel représenté	28
1.3.3 Kennedy ouvre les portes de l'ère moderne	30
2 Les médias synthétiques, les deepfakes et les internets	34
2.1 Naissance et développement du « faux authentique »	36
2.1.1 Fake news, deepfakes, médias synthétiques	37
2.1.2 La nature profonde des « fakes »	37

2.1.3 Histoire de l'intelligence artificielle, des réseaux neuronaux et de l'apparition des <i>deepfakes</i>	41
2.2 L'ère des armes de déception massive	48
2.2.1 Les fondements d'une ère de tromperie numérique	49
2.2.2 L'algorithme fabricant d'illusions.....	50
2.2.3 Quand il se sent attaqué, le corps réagit.....	52
3 Les scénarios d'une réponse	57
3.1 Solutions technologiques	58
3.1.1 Le framework Lessig comme cadre d'action	59
3.1.2 Cryptage et sécurisation, éléments de la lutte anti-deepfake.....	60
3.1.3 Les outils sont-ils les seuls levier d'influence ?.....	63
3.2 Solutions légales, politiques et éthiques.....	64
3.2.1 La nécessité de l'éducation aux médias.....	64
3.2.2 La formation des professionnels.....	65
3.3 Solutions médiatiques	66
3.3.1 Les journalistes jouent à un jeu « fini ».	67
3.3.2 Les auteurs de fake news et de deepfakes jouent une partie infinie	68
3.3.3 L'ébauche d'une solution, changer de partie	68
Conclusion	71
Bibliographie.....	72

Remerciements

« ... Ne te fie pas à tes yeux. Tout ce qu'ils te montrent, ce sont des limites, les tiennes. Regarde avec ton esprit, découvre ce dont d'ores et déjà tu as la conviction et tu trouveras la voie de l'envol... »

Richard Bach, *Jonathan Livingston le Goéland*, 1970

À Giulia, merveilleux goéland qui accompagne mon vol, qui a su m'encourager chaque jour à rejoindre mon horizon.

Merci également à tout ceux qui ont répondu à mes interrogations en me livrant leur réflexion et plus particulièrement à :

- Mon tuteur, Frédéric Filloux, journaliste, spécialiste des médias et de l'économie numérique, fondateur de *Monday Note*
- Jeff Jarvis, journaliste, auteur et professeur à la CUNY's Newmark J-school
- Claire Wardle, fondatrice de *Firstdraft News* et ancienne directrice du Tow Center for Digital Journalism à l'université de Columbia.
- Guillaume Brossard, fondateur du site *Hoaxbuster*
- Cédric Mathiot, journaliste à *Libération* et responsable du service [Checknews.fr](https://www.checknews.fr)
- Samuel Laurent, journaliste au *Monde* et responsable du service *Les Décodeurs*
- Vincent Nozick, chercheur et maître de conférences à l'université Paris-Est Marne-la-Vallée, spécialiste de la détection et de la falsification d'images.

Introduction

« L'universalité des hommes se repaît de l'apparence comme de la réalité ; souvent même l'apparence les frappe plus que la réalité même »

Machiavel, *Discours sur la première Décade de Tite-Live*, 1513-1520¹

La place de l'image dans notre société « iconocentrée » est envahissante, totalisante, omniprésente. Jusque-là, pas de révélation poignante. Grâce à une manipulation experte et quotidienne des images, nous nous inventons de « fausses vies » sur les réseaux sociaux ; nous laissons le web s'emparer quotidiennement de notre intime et valider une réalité complètement distordue. Si le commun des mortels peut arriver à nous faire avaler à travers les photos qu'il déballe sur Instagram une version édulcorée de sa vie, que se passe-t-il quand des professionnels de l'image s'affairent à fabriquer de fausses informations ?

Nous nous attaquons à un sujet que tout journaliste, tout enquêteur, tout philosophe a essayé de percer à jour : quelle sont les limites de la représentation du « vrai » ? Nous sommes pleinement conscients que ceci mériterait une étude beaucoup plus approfondie et c'est pour cela que nous allons nous concentrer sur les dérives modernes de la question, en essayant d'explicitier le phénomène des *deepfakes*, ces nouvelles vidéos qui détournent la réalité, et la nouvelle pierre qu'elles apportent à l'édifice du « vrai faux » ou plutôt du « faux vrai ». Cependant, pour ce faire, nous ne pouvons pas nous épargner

¹Discours sur la première décade de Tite-Live

Trad. de l'italien par Alessandro Fontana et Xavier Tabet, Collection Bibliothèque de Philosophie, Gallimard

une contextualisation des moments marquants de la représentation du vrai dans l'histoire iconographique moderne. En passant par Vermeer et le réalisme de ses peintures jusqu'aux premiers photographes et les premiers cinéastes nous explorerons ensemble le lien qui se tisse tout au long du XIXème et du XXème siècle entre le public occidental et l'image. Nous verrons en effet avec le présent travail comment l'image, qu'elle soit statique ou en mouvement, a toujours posé un souci de véridicité. Le but de ce mémoire est donc de proposer une réflexion transversale sur l'influence des *deepfakes* ou « médias synthétiques »² sur la société dans un contexte désormais qualifié – à tort ou à raison – d' « ère de la post-vérité »³

La nature profondément technologique des médias « synthétiques » questionne tout d'abord les circonstances de leur émergence. Quels sont les événements qui ont mené à leur développement ? Sont-ils des objets résolument innovants ou les dignes héritiers des médias de propagande ? Et, surtout, sont-ils technologiquement mûrs ou font-ils partie des premiers représentants d'une nouvelle forme de médias ?

² Les médias synthétiques sont des médias générés par l'intelligence artificielle et notamment par les GANs (réseaux de neurones artificiels) quand les deepfakes font référence essentiellement aux vidéos apparues d'abord sur Reddit puis ailleurs sur le net, mettant en scène des actrices porno sur le corps desquelles le visage d'actrices célèbres avait été "collé".

³ Concept que l'auteur américain Ralph Keyes introduit dès 2004 dans son livre "The Post-Truth Era: Dishonesty and Deception in Contemporary life" et que le dictionnaire d'Oxford, puis le Larousse et le Petit Robert ont consacré en 2016 à la suite de la campagne précédant le référendum du Brexit. Une des définitions pourrait être: "La post-vérité résulte de tous les flux d'information très et trop approximatifs, tronqués, mensongers, propagandistes, propres aux réseaux sociaux et aux données numériques."(De l'utilité de la notion de post-vérité Gil Delannoï - Dans Le Débat 2017/5 (n° 197), pages 4 à 12)

Pour saisir l'incroyable complexité de ce phénomène, on ne saurait faire l'impasse sur notre rapport aux médias plus conventionnels qui, au fil des décennies sont devenus vecteurs de culture et piliers de nos raisonnements contemporains. En observant l'influence qu'ont ces formes médiatiques sur notre société, il sera plus aisé de considérer les conséquences potentielles qu'auront les médias synthétiques. De la même façon, nous observerons notre rapport à la « vérité » qu'est supposé véhiculer le médium.

1 Voir pour croire ou croire pour voir. L'image a-t-elle toujours représenté la réalité ?

Comment s'est construite la représentation du réel que nous partageons ? Comme nous l'avons annoncé lors de notre introduction, avant d'essayer de comprendre la place qu'occupent les *deepfakes* et autres contenus synthétiques depuis deux ans dans nos sociétés, il est légitime de s'interroger sur la place de l'image en Occident et d'en interroger l'histoire. Nous étudierons l'image de Vermeer, du Réalisme en peinture, de la photographie et du cinéma, pour dégager la relation parfois ambiguë que nous entretenons avec cette représentation de la réalité. Car, en somme, c'est de cela qu'il s'agit : comprendre cette croyance que nous plaçons dans les images pour mieux analyser dans un second temps la portée trompeuse des deepfakes.

Il s'agit de saisir le sens de cette volonté permanente de tirer un trait direct entre le sujet de référence et sa représentation, entre la réalité et son image. La société moderne, que Debord décrit comme une société du « spectacle » où « l'image construite et choisie par quelqu'un d'autre est devenue le principal rapport de l'individu au monde qu'auparavant il regardait par lui-même »⁴, fait face à une alternative sans précédent. Elle peut se résigner à perdre le lien avec la réalité et s'enfermer dans un scepticisme dogmatique, ou se battre pour conserver une vision aussi fidèle que possible de la réalité à travers ses représentations. De notre compréhension des trajectoires passées dépendra probablement la pertinence des réponses à venir.

⁴ Guy Debord, *La société du spectacle*, Paris, Buchet-Chastel, 1967.

1.1 Le Réalisme, le vraisemblable et l'objectivité de l'image

Il convient dans cette première partie de définir un point de départ dans notre analyse. C'est au XVI^{ème} siècle que notre investigation commence pour ensuite se prolonger à travers les époques jusqu'à l'ère industrielle et les prémices de la Seconde Guerre mondiale. Pourquoi remonter si loin ? L'hypothèse défendue ici tient dans le fait que notre rapport à la réalité s'est construit à travers le temps notamment par le biais des images qui ont émaillé notre histoire. La société visuelle dans laquelle nous nous trouvons et où apparaissent les deepfake hérite de cette histoire et de ce rapport complexe à la réalité et à son objective représentation.

1.1.1 Qu'est-ce que le Réalisme⁵ ? L'exemple de Johannes Vermeer

Dès le XVI^e siècle et le « siècle d'or néerlandais »⁶, les peintres, Vermeer et Rembrandt en première ligne, poussent la vraisemblance de leurs œuvres à leur paroxysme. Bien que la tentative de représenter la nature, les hommes et leur condition avec fidélité ne soit pas spécifique à cette époque, il est intéressant de se pencher sur les peintures de Johannes Vermeer⁷ qui, après une longue période d'oubli, fut redécouvert par Théophile Thoré-Bürger, journaliste français, critique d'art du milieu du XIX^e siècle et défenseur du Réalisme.

⁵ Nous écrivons dans ces pages Réalisme pour désigner le mouvement du XIX^{ème} siècle dans lequel les écrivains peignent la réalité de leur temps, explorent la vie quotidienne sous toutes ses formes et analysent la société, par extension, les artistes embrassant cette démarche font également partie de ce courant.

⁶ Le siècle d'or néerlandais désigne une période du XVI^{ème} au XVII^{ème} siècle aux Pays Bas pendant laquelle un épanouissement culturel intense va bouleverser les arts, la littérature et la science en produisant les œuvres et les travaux les plus remarquables.

⁷ Johannes Vermeer est un peintre baroque Néerlandais originaire de Delft né en 1632 et mort en 1675 dont le tableau le plus connu, La Laitière, est considérée comme la Joconde. Il fut redécouvert au milieu du XIX^{ème} siècle au moment de l'essor du Réalisme en France par le critique journaliste Théophile Thoré

La peinture de Vermeer donne une « impression de vie » que nous ne retrouvons guère dans d'autres peintures. Cette impression provient de la lumière, vivante, superbe, extrêmement réaliste. Le traitement de cette lumière trahit l'usage par le peintre d'une technique particulière, connue des grands maîtres avant lui, mais que Vermeer semble affectionner particulièrement : la *camera obscura*⁸. C'est le principe au cœur de tout appareil photo moderne.

Bien que la question ne soit pas tranchée officiellement, il semble que Vermeer n'utilisait pas la *camera obscura* pour tracer le dessin de ses tableaux, mais plutôt pour en sublimer le rendu. Des travaux publiés en 2001 par le peintre anglais David Hockney et l'américain Charles Falco soutiennent cette hypothèse en explorant le travail d'autres artistes comme Jan van Eyck ou Lorenzo Lotto, Caspar van Wittel, Rembrandt ou Caravaggio. Ils suggèrent dans leur livre, *Secret Knowledge*⁹, le recours probable à une variété de techniques similaires à celles qu'aurait employé le peintre de Delft. Sur ce point, la thèse Hockney-Falco avance notamment l'usage de la *camera obscura* comme une des techniques révolutionnaires qui aurait permis l'émergence des peintures réalistes.

Plusieurs voix, dans le milieu de l'histoire de l'art, s'opposent à cette hypothèse mais la nature quasi sacrée de ces maîtres n'aide certainement pas à y voir plus clair, tant les tabous autour de leur génie créatif sont nombreux. Cependant, nous considérons ici que l'usage de procédés techniques avancés pour l'époque n'est pas à exclure et ne

⁸ Invention du scientifique arabe Ibn Al-Haytham, la *camera obscura* se compose d'une pièce sombre dont l'un des murs est percé d'un trou muni d'un dispositif optique permettant de capturer la lumière (une lentille). Sur le mur opposé au trou, se forme naturellement l'image de ce qui se trouve à l'extérieur de la chambre et dont la lumière émise traverse le trou.

⁹ *Secret Knowledge: Rediscovering the Lost Techniques of the Old Masters*, David Hockney, Thames & Hudson, 2001

remet en aucun cas la maîtrise et l'expertise de ces peintres fabuleux. Comme tous les artistes, les maîtres précurseurs du Réalisme s'employaient à vivre de leur art en menant une activité de recherche et développement qui leur assurerait un avantage face à une concurrence rude. Léonard de Vinci était d'ailleurs en son temps l'un des plus prolifiques à ce sujet ; ses nombreux sujets d'études sur la nature des pigments, les techniques de peinture et la *camera obscura* montrent sa préoccupation à faire évoluer son art à l'aide de nouvelles techniques.

Vermeer semblait donc employer la *camera obscura* pour apporter des touches de lumière singulières en utilisant un flou de lumière ponctuelle pour suggérer la profondeur des plans (effet de bokeh¹⁰). C'est cette technique qui donne aux tableaux de Vermeer l'apparence du réel. Le pointillisme¹¹, les touches d'éclats disposées çà et là sur la toile, bien que n'étant pas réalistes, elles peuvent en effet être reproduites par un dispositif optique, apportent une touche de réalisme : l'œil interprète ces formes de couleurs comme étant probablement tirées du réel. Dans ce cas, le mélange entre réel et « artificiel » sur la toile arrive à créer chez le spectateur la nécessaire conviction que ce qu'il est en train de regarder provient du réel. Vermeer tente d'insuffler la vie dans ses toiles et, par une curieuse alchimie, parvient par exemple à restituer la paisible intensité d'un après-midi passé chez une crémère de Delft. Ce faisant, il plante les bases du réalisme et par la même occasion invite le spectateur à interroger son rapport au réel en offrant luxe de détails à ses tableaux désormais devenus célèbres.

¹⁰ Le bokeh est l'effet de flou qu'on aperçoit en arrière plan sur certaines photos de portrait réalisées à l'aide d'un petit téléobjectif.

¹¹ Le style de Vermeer se caractérise par des couleurs transparentes obtenues par l'application sur la toile de peinture en couches légèrement granuleuses, une technique appelée **pointillée**, très différent du pointillisme popularisé par Georges Seurat.

1.1.2 Le vraisemblable, un début d'éloignement de la représentation réaliste

L'exhumation de Vermeer par Thoré-Bürger en 1842 crée un lien unique entre une société Néerlandaise du XVII^e siècle en plein renouveau, portée par une prospérité éclatante, et une société française du XIX^e en pleine révolution industrielle se réinventant complètement autour d'une approche rationnelle du monde qui l'entoure. Par nature, la peinture de Vermeer, de Rembrandt et de leurs contemporains néerlandais s'attache à la vraisemblance des scènes, à mettre le spectateur en contact direct avec le réel, sans artifice. En héritant plus tard des influences de ces peintres néerlandais, le Réalisme rejette tout ce qui relève de l'imaginaire, tente d'apporter une vision objective du monde en représentant ce que la majorité de l'humanité admet comme étant le « vrai ». C'est dans la nature même de cette recherche de vérité que le Réalisme plante les bases de la relation complexe, qui se développe dans les siècles suivants et jusqu'à nos jours, entre cette vérité et l'objective présence du sujet représenté, ce que Barthes a décrit comme « ayant été là »¹². En observant ces œuvres, le public s'interroge sur la nature même de ce qui est représenté. Est-ce le réel ? Est-ce une fiction ? Est-ce un mélange des deux ? Ou bien la projection d'une vision du monde personnelle et soumise à interprétation ?

Le réalisme qui émane de ces toiles trouble les observateurs en cela qu'ils savent que la représentation de la réalité qu'ils ont en face d'eux provient du talent d'un homme qui, en toute subjectivité, a tracé les lignes, appliqué les pigments et passé le vernis qui servira à créer l'œuvre. La main, prolongement de l'esprit, trahit dans une certaine mesure cet œil créatif, qui ne peut qu'embellir une réalité parfois morne. Mais cette

¹² « Ça-a-été » (...) cela que je vois s'est trouvé là, dans ce lieu qui s'étend entre l'infini et le sujet (operator ou spectator) ; il a été là, et cependant tout de suite séparé ; il a été absolument, irrécusablement présent, et cependant déjà différé. » Roland Barthes, *La chambre claire*, 1980.

subjectivité de l'artiste est compensée à la fois par une technique accrue, permettant de singer la réalité par tout un tas de stratagèmes, et par la propre volonté des spectateurs de croire en la réalité de l'œuvre elle-même.

En plongeant son regard dans les plis de peinture déposés sur la toile, il est difficile de s'extirper de la représentation du sujet. Mais dès lors qu'un pas ou deux amènent le regard à embrasser l'ensemble de la toile, les détails s'estompent et l'illusion de réalité advient. La frontière entre signifiant et signifié s'efface, pour finalement les faire se rejoindre et donner la parfaite illusion que ce qui se trouve face à soi n'est autre qu'une parfaite représentation de la réalité. Ce réel, ce que le public tient pour vrai, se confond donc avec sa représentation parce que, d'une part, l'artiste a tenté avec maestria de représenter le plus fidèlement possible ce qu'il a vu et que, d'autre part, l'observateur a cru en sa vraisemblable authenticité. Cette ténacité du spectateur à penser, dans le contexte d'observation de l'œuvre, que celle-ci représente la vérité d'une situation, joue un rôle structurant dans le sentiment de vérité qui se dégage de l'œuvre.

Vermeer et les peintres Réalistes qui lui succèdent tentent donc d'imprimer sur la toile toutes les apparences de la réalité, plus que la réalité elle-même. « Faire vrai consiste donc à donner l'illusion complète du vrai, suivant la logique ordinaire des faits, et non à les transcrire servilement dans le pêle-mêle de leur succession. J'en conclus que les Réalistes de talent devraient s'appeler plutôt des Illusionnistes. »¹³ nous dit Maupassant dans sa préface de *Pierre et Jean* en 1888. C'est bien vers l'illusion que

¹³ Pierre et Jean, Guy de Maupassant, Édition Paul Ollendorff, 1888

nous nous dirigeons progressivement à travers la photographie et l'arrivée imminente du cinéma quelques décennies plus tard.

1.1.3 L'objectivité, une illusion cherchée par tous

La société moderne se développant, la relation aux œuvres picturales se démocratise. Les publics des sociétés occidentales vont peu à peu développer une relation complexe aux images qui les accompagnent, oscillant entre crédulité et scepticisme. Le Réalisme n'est qu'une des premières étapes d'un long processus qui va de plus en plus entrelacer la notion de réalité et les représentations de cette réalité. En peinture, la quête du vrai dérivera entre autres dans la volonté d'immortalisation de l'instant avec l'Impressionnisme ou dans la déconstruction de la réalité avec le Cubisme. Bien que ces courants artistiques auraient beaucoup à apporter à l'édifice de la recherche de la réalité, nous décidons maintenant de faire un petit bond en avant pour nous concentrer sur les deux événements qui conditionneront à jamais le rapport moderne et contemporain avec la vérité de l'image : l'avènement de la photographie et des images en mouvement.

C'est avec la première photographie prise par Nicéphore Niepce¹⁴ en 1826 que la peinture perd définitivement son pouvoir de représentation du réel. *Le Point de vue du Gras*, première image capturée par le biais d'un dispositif optique et fixé sur une plaque d'étain enduite de bitume de Judée, sonne le glas de cette époque de représentation subjective, pour lui substituer une objectivité tout apparente. Qu'est-ce donc la photographie sinon l'apparente objectivité d'un enregistrement ?

¹⁴ Nicéphore Niepce est un ingénieur français né à Chalon-sur-Saône. Inventeur de la photographie, il a réalisé le célèbre cliché « Point de Vue du Gras », une héliographie produite sur une plaque d'étain recouverte de bitume de judée. La photographie est aujourd'hui conservée au Harry Ransom Center à l'université du Texas à Austin, Etats-Unis.

Se mêlent dans ce nouveau médium toutes les idées qui traversent le XIX^e siècle, un siècle résolument tourné vers la technique et sa suprématie. Nous sommes à l'heure du « Scientisme »¹⁵, né des idées de Descartes et Condorcet, qu'Ernest Renan décrit comme « l'organisation scientifique de l'humanité ». C'est le temps de la construction des chemins de fer en Europe et aux États-Unis, de la création du télégraphe et des longs courriers qui traversent l'atlantique grâce à la vapeur, du développement de l'électricité et de l'électromagnétisme et, plus tard, du téléphone, de la radio et du cinéma. On peut désormais se déplacer rapidement sur de très longues distances et communiquer avec l'autre bout du monde en un temps record.

La photographie accompagne le mouvement. Elle est d'abord popularisée par Daguerre après qu'il a enfin découvert les méthodes de fixation de l'image en 1837, et devient rapidement l'instrument des portraitistes. Près de 80% des images réalisées au cours du XIX^e siècle sont des portraits : les bourgeois et les classes moyennes de l'époque, auxquels les portraits peints étaient inaccessibles, se jettent sur l'opportunité de se faire tirer le portrait pour rejoindre la bonne société. La photographie de soi devient à la fois un acte de mémoire, mais aussi un acte social qui classe l'individu.

Nadar¹⁶, ancien caricaturiste, réalise des centaines de portraits, dont ceux de Baudelaire, Courbet, Hugo, Zola, Mallarmé, Maupassant et Jules Vernes. Son travail, comme celui des autres photographes portraitistes de l'époque, comme Disdéri (qui

¹⁵ Le scientisme est une vision du monde, apparue au XIX^e siècle, qui avance que la science expérimentale a priorité sur les formes plus anciennes de référence sociale et religieuse.

¹⁶ Félix Tournachon, dit Nadar, est un portraitiste français du XIX^e siècle connu pour ses photographies d'hommes célèbres. Jules Vernes, Alexandre Dumas, Charles Baudelaire ou Georges Sand ont été parmi ses clients.

invente la carte de visite photographique), se confond avec la peinture. Les mêmes techniques de composition, les mêmes décors, le même soin apporté à la lumière sont employés pour donner l'impression du vivant. Les portraits mis en scène sont parfois un peu figés et convenus mais, petit à petit, on recadre vers le visage et l'expression des traits se détend, apportant ce sentiment saisissant d'être en présence du sujet.

Au fil du temps, le studio où se trouve confinée la lourde chambre photographique ne convient pas à Monsieur Tournachon, le vrai nom de Nadar. Il commence alors à transporter sa chambre photographique comme on transporte le chevalet du peintre dans toutes sortes de lieux insolites. L'une de ses expérimentations porte sur les catacombes, où il entreprend de réaliser les premières photographies éclairées à l'aide d'une source de lumière artificielle. On découvre dans ces images de longues silhouettes habillées d'une blouse de travail et coiffées d'un chapeau à larges bords triant des monceaux de squelettes dans un souterrain parisien. La lumière crue, visiblement éblouissante, paralyse les sujets, les faisant ressembler dans leur rigidité aux morts qui les entourent. (Cf. Nadar catacombes) Le dispositif d'éclairage, inédit à l'époque, repose sur un flash de poudre de magnésium, qui contraint le photographe à utiliser des temps de pose relativement longs. Pour éviter de trop nombreuses erreurs, Nadar préfère donc utiliser des mannequins comme sujets plutôt que des hommes, qui à l'évidence bougeraient trop. En tentant de recréer le vrai, de capturer objectivement la réalité d'une situation, Nadar n'a pas eu d'autre choix que de la travestir et de fabriquer ces illusions vraisemblables du travail des ouvriers des catacombes. Cette démarche, dont la portée n'a certainement pas été perçue par Nadar lui-même, inscrit dans les fondamentaux de la photographie cette bataille pour l'objectivité.

La photographie remplace la peinture dans la quête d'un certain réalisme parce que son dispositif repose sur une technologie qui ne laisse que très peu de place à la main de l'homme. L'optique, qui sert à faire pénétrer la lumière dans la chambre noire, n'est certes pas neutre, mais, à l'époque de Nadar, son utilisation ne sert pas encore consciemment une narration spécifique. Le processus de fixation de l'image (plaque puis procédé de révélation de l'image en laboratoire) ne constitue qu'une étape purement technique à l'époque. L'ensemble du processus photographique témoigne apparemment de son innocuité à transformer le réel ainsi saisi et donc de son apparente objectivité. La nature des sujets photographiés contribue aussi grandement à cette réputation d'objectivité. Quoi de plus facile en effet que de comparer celui qui est portraituré à son portrait et constater que hormis la couleur, tout y est ?

Pourtant, très vite, le public va pouvoir se rendre compte que l'objectivité du médium n'est pas si franche et que bien souvent, les opérateurs photo se plaisent, à l'instar de Nadar et de ses mannequins, à modifier des éléments de composition pour donner l'illusion de l'objectivité.

De l'autre côté de l'atlantique, la situation se durcit dans la seconde moitié du XIX^e siècle. En 1861, la guerre de Sécession éclate entre les États-Unis d'Amérique dirigés par Abraham Lincoln et les États confédérés d'Amérique dirigés par Jefferson Davis. Les soldats ne sont pas les seuls à partir à la guerre. Quelques photographes leur emboîtent le pas pour témoigner de la dureté des affrontements et documenter ce moment d'histoire. Parmi eux, un Écossais d'une quarantaine d'années au regard franc, le cheveu noir et souple, la barbe longue et bien taillée. Il s'appelle Alexander Gardner¹⁷.

¹⁷ Alexander Gardner est un photographe américain d'origine écossaise de la fin du XIX^e siècle connu pour ses photographies de la guerre de Secession et son portrait d'Abraham Lincoln. Mathew Brady est également un

Il faut le voir poser dans son studio new-yorkais, le regard porté au loin, pour comprendre l'intensité du personnage.

Lorsque la guerre éclate, en 1861, Gardner est assigné à prendre le portrait des soldats qui partent à la guerre. Très vite, il apprend que son mentor, Matthew Brady, a réussi à assister à la bataille de Manassas et à en rapporter des images. Avec une vingtaine d'assistants, il part alors couvrir le conflit à l'aide de chambres portatives. Son rang de capitaine lui permet d'accéder au champ de bataille de Antietam où il saisit l'horreur du massacre (près de 22 000 morts en un jour), les corps jonchant les champs à perte de vue. Ses photos sont exposées en 1862 par Brady pour montrer au grand public les réalités du conflit et le *New York Times* dira de lui que « s'il n'a pas ramené les corps en personne dans nos jardins, il a fait avec ses images quelque chose d'assez proche ».

Gardner a photographié la guerre civile plus que quiconque et son approche nous paraît aujourd'hui celle d'un photojournaliste aguerri. Il semble montrer avec objectivité les conséquences de la guerre qui a déchiré les États-Unis pendant quatre ans. Ce que découvre l'historien William Frassanito¹⁸, près d'un siècle plus tard, en 1975, contredit pourtant cette idée. Après un examen approfondi de certains clichés pris par Alexander Gardner sur le site de la bataille de Gettysburg, il observe que le corps d'un même soldat a été déplacé et réutilisé plusieurs fois dans des postures et des lieux différents. L'objective réalité des photos de Gardner n'est donc qu'une illusion.

photographe américain considéré par beaucoup comme le premier photographe de guerre. Il fut le mentor d'Alexander Gardner avant qu'une brouille ne les éloigne.

¹⁸ Gettysburg: A Journey in Time, William Frassanito, 1975

Ce qu'il est important de remarquer, c'est que Gardner ne s'est jamais prétendu autre chose qu'un artiste. Nous sommes très loin de la figure contemporaine du photoreporter. Ce qui choque aujourd'hui ne choquait personne à une époque où l'image photographique était encore jeune. La volonté de Gardner était avant tout de témoigner d'une situation de la meilleure façon possible en captant la réalité d'une bataille. L'objectivité du regard ne consistait pas une fin en soi, en revanche il fallait que ces images transpirent la mort qu'elles montraient, il leur fallait exsuder une certaine vérité, un parfum de vraisemblance que le public d'alors accepterait.

La représentation de la réalité par l'image revêt différents aspects, dont la complexité ne peut être pleinement abordée dans ces pages. Mais à travers les exemples abordés, nous constatons que, depuis les peintres néerlandais du XVII^e siècle jusqu'aux photographes de la fin du XIX^e siècle, la tentative de représenter le réel se heurte aux besoins de vraisemblance. La proposition d'une image à un public suppose de ne pas seulement user d'un jeu de techniques artisanales pour représenter mécaniquement la chose réelle, mais bien d'instiller dans ces images produites une dimension supérieure, une émotion intrinsèque, pour lui donner une forme de vraisemblance, de sorte que la réalité et sa représentation ne fassent qu'un aux yeux de ce public. Il s'agit donc de créer l'illusion du vrai et de convaincre à travers différents stratagèmes les observateurs que cette illusion n'est que la stricte et objective transposition de ce qui a été vu par le créateur de l'image.

Cette croyance du public dans la représentation comme étant objectivement la reproduction du réel va se construire au fil du temps par le biais de différents mécanismes d'acceptation sociale. Elle débouchera sur la construction collective d'une

certaine vision du monde par le biais de sa représentation iconographique. La société, en tant qu'entité propre, commence à croire en sa propre existence par le biais de ces images « plus vraies que nature » et les ajoute au récit écrit pour en définir les contours, comprendre ce qui s'y passe et agir en conséquence.

En somme, l'industrialisation progressive des sociétés occidentales du XIX^e siècle et l'apparition des premières technologies de communication longue distance permettent de relier les hommes comme jamais auparavant. La production d'images s'accélère du fait de sa mécanisation et des techniques de reproduction qui permettent les copies rapides sans dégradation aucune. L'ensemble combiné permet à des images à forte connotation émotionnelle de circuler à travers le monde très rapidement. Cette accélération soudaine et fondamentale pour la suite marque le passage d'une société de l'écriture, fondée sur la logique et la raison qu'ont portées les penseurs tels que Descartes ou Kant, à une société de l'image fondée sur la dialectique et l'émotion, dont les fondamentaux reposent sur l'iconographie.

1.2 La construction d'un réel collectif à travers l'image

Nous basons ici notre réflexion, sans l'évoquer faute d'espace, sur la thèse de Bernard Delforce¹⁹ qui affirme que les conditions de production de l'information et que la construction sociale du sens peut être examinés sous une approche constructiviste. « Pour parler du réel il faut en construire une représentation » nous dit Delforce, c'est donc en étudiant la façon dont l'histoire est rapportée par l'image que nous comprendrons la représentation qui en est faite à travers le XX^e siècle.

¹⁹ Le constructivisme : une approche pertinente du journalisme, Bernard Delforce, 2004

1.2.1 La photographie, premier artisan de la représentation

Si le XIX^e siècle a été un siècle d'industrialisation, le XX^e est, sans conteste, un siècle de massification. Les innovations qui voient le jour dans les dernières décennies du XIX^e se répandent sur la surface du globe à la vitesse grand V.

Il faut parler du cinéma, que les frères Lumière²⁰ « inventent » en 1895 et des premières tentatives, qui relèvent à la fois du test technique et du témoignage du réel. Près de 108 films sont tournés par Auguste et Louis. Parmi tant d'autres : *La Sortie des usines Lumières à Lyon*, *L'Arroseur arrosé*, *Baignade en mer* et, le plus célèbre d'entre eux, *L'Arrivée du train en gare de la Ciotat*, dont la légende raconte que les spectateurs, pris de panique face au réalisme de la scène, auraient quitté la salle de projection de peur que le train ne leur fonce dessus. Cette « invention », qui survient à l'apogée de la révolution industrielle du XIX^e siècle, ne donne pas le mouvement aux images à proprement parler – le kinétoscope d'Edison²¹ l'avait fait dès 1891 –, mais offre au grand public l'accès à ces films. De ces premières projections naît un engouement du public pour les salles obscures où toute une génération de créateurs vont pouvoir montrer leurs œuvres. Georges Méliès, le plus connu des réalisateurs français de l'époque, investit le monde de l'imaginaire jusqu'à ce que la guerre de 1914-1918 éclate et que la soif d'informations ne soit plus forte.

²⁰ Auguste et Louis Lumière comptabilisent près de 170 brevets à leur actif, dont celui d'un nouveau type de kinétoscope permettant l'enregistrement mais aussi la projection de films. C'est l'invention du cinéma telle qu'on le connaît aujourd'hui.

²¹ Le kinétoscope est le plus ancien dispositif de l'histoire du cinéma, destiné à visualiser une œuvre photographique donnant l'illusion du mouvement. Il va de pair avec le kinétophore, première caméra de cinéma.

Pendant la Première Guerre mondiale, d'autres images, des images d'information, s'invitent dans la presse. *Le Monde illustré* et *Le Petit journal* dépeignent la vie des poilus dans les tranchées à travers des instantanés, photographies prises sur le vif. *L'Illustration*, précurseur « picture magazines », couvre l'arrivée des Américains dans le saillant de Saint-Michel. Le cinéma devient un lieu où l'on vient s'enquérir de l'effort des troupes. Le cinéma de création est décimé, les artistes sont envoyés au front, rares sont ceux qui restent en arrière. En France, la SCA (Section cinématographique de l'Armée) et la SPA (Section photographique de l'Armée)²² sont créées pour relayer la propagande officielle et montrer les avancées victorieuses de l'armée.

1.2.2 L'image se mets en mouvement

Après la guerre, à la fin des années 1920, le *New York Times* commence à publier des photographies en première page et, dans le même temps, d'autres publications lui emboîtent le pas, avec la création des légendaires *VU* par Lucien Vogel, en 1928, et *Life Magazine*²³ par Henry Luce, en 1936. Ce sont aussi les années de la démocratisation du cinéma. Elles voient la sortie des deux grands documentaires fondateurs : *Nanouk of the North* de l'américain Robert Flaherty et *L'Homme à la caméra* du russe Dziga Vertov, qui accompagnent les années folles et la soif de vivre et de découvrir le monde qui les occupe. À ce moment, l'image change de signification avec l'arrivée des premiers films de propagande montrant les poilus au front. De l'illusion imaginaire, on passe à l'illusion

²² « Au printemps 1915, un accord entre le ministère de la Guerre et les maisons cinématographiques aboutit à la création de la Section cinématographique de l'armée (SCA). Peu après, les ministères de l'Instruction publique et des Beaux-Arts, de la Guerre et des Affaires étrangères créent la Section photographique de l'armée (SPA). Dans les deux cas, l'objectif est de contrer la propagande allemande. » Sené Xavier, France Archives

²³ *Vu* et *Life* sont ce qu'on appelle des *Pictures Magazines* dont la production est entièrement tournée vers la photographie. Le premier est français et le second américain. Tous deux se sont fait concurrence avant que *Vu* ne ferme ses portes. Les plus grands photographes du monde ont été publiés dans leur pages.

de la propagande. L'image, on le sait, ne représente plus le réel, ni même la version subjective de l'opérateur de cette réalité, mais la version politique d'une réalité.

La photographie quant à elle n'est pas en reste avec la Révolution russe. Elle devient l'instrument de propagande post-révolution d'un Staline en quête d'absolu. Toutes les figures de la révolution seront peu à peu gommées de l'iconographie officielle au fur et à mesure des limogeages, des disparitions et des exécutions. Dès 1920, Trosky, Kamenev, Malchenko ou le chef du NKVD²⁴ lejev sont les premières figures à être littéralement effacées des clichés officiels. Le régime refait l'histoire, il la manipule en changeant le visible, de sorte que les yeux ne voient plus ce qu'ils savent avoir vu exister. Dans la Russie soviétique, l'image est soupçonnée, on sait qu'elle peut désormais témoigner à charge, montrer ce qui n'a pas été, placer les individus dans des situations compromettantes. En Occident, on se méfie des rouges et de leur propagande, et pourtant...

1.2.3 La manipulation par l'image devient industrielle

Pourtant, la propagande s'invite chez l'Oncle Sam comme instrument de manipulation à la solde des industries américaines. Il s'agit de persuader les populations du bien-fondé de l'économie de marché et du capitalisme face à la montée du communisme. À New York, le neveu de Freud, Edward Bernays²⁵, devient en quelques coups d'éclat l'un des pères de la propagande politique et industrielle. Dès 1917, il œuvre à la commission Creel pour inciter les Américains à rejoindre la guerre qui fait rage en

²⁴ Le NKVD, commissariat du peuple aux Affaires intérieures, était l'organisme d'État en charge de combattre le crime et de maintenir l'ordre public de 1934 à 1946 en Russie soviétique.

²⁵ Edward Bernays est considéré comme le père de la propagande politique institutionnelle et de l'industrie des relations publiques, ainsi que du consumérisme américain. Neveu de Sigmund Freud, il a contribué à populariser le tabac chez les femmes.

Europe, puis, en 1920, il aide la Beech-Nut Packing Company, grosse productrice de produits issu du cochon, à populariser le bacon à travers une campagne de pub nationale vantant les mérites du petit déjeuner « Bacon and eggs ». En 1929, commandité par la Lucky Strike Company, il parviendra à imposer la cigarette chez les femmes américaines en profitant du mouvement des suffragettes américaines et de leur désir de liberté, faisant de la cigarette un symbole d'émancipation.

Le livre de Bernays, *Crystallizing Public Opinion*²⁶, publié en 1923, décrit par le menu les grands principes de la « fabrique du consentement » que dénoncera plus tard le linguiste américain Noam Chomsky, entre autres. Un ouvrage qu'un journaliste rapporte avoir vu en 1933 dans la bibliothèque de Joseph Goebbels, le ministre de la propagande du III^e Reich, et qui semble avoir inspiré la stratégie nazie. Bernays joue sur les images mentales que nous avons du monde et tente, avec succès, d'y imposer celles de ses clients en usant du cinéma, de la photo, de l'illustration pour véhiculer les principes fondateurs d'une idée. Avec lui, l'image devient instrument, vecteur d'influence, agent de propagation. Le réel s'efface devant le message, l'illusion devient reine et les mythes modernes s'établissent fermement dans l'iconographie. Le cow-boy de Marlboro, la pomme d'Apple sont les plus beaux exemples contemporains de la pertinence des travaux de Bernays. Dans son monde, le réel doit être mis en scène, il se fabrique pour livrer une narration parfaite dans laquelle le produit et l'idéologie viennent parfaitement s'insérer. Difficile pour le public de discerner le vrai du faux. Tout devient vraisemblable ou acceptable, les relais d'opinion crédibles.

²⁶ Crystallizing Public Opinion, Edward Bernays, 1923

Cependant, une ombre plane. Certains leaders politiques de l'époque ont compris mieux que d'autres le rôle de l'image sur l'opinion. Il faut voir l'iconographie fasciste de Mussolini et les essais gestuels d'Hitler en 1925, capturés par son photographe personnel Heinrich Hoffman, pour comprendre que les futurs dictateurs entendaient faire de l'image un axe central de leur communication.

1.3 La propagande par l'image, le réel disparaît au profit du message

L'image animée ou fixe devient dans les années 30, un outil de propagande majeur au service des industries et de la politique. Les regards se construisent dans cette lumière qui délivrent les informations avec une complaisance coupable pour l'approximation et la manipulation.

1.3.1 Deux camps, une même approche manipulateur

Le rapport à l'image se construit dans une ambivalence complexe au fil des années 1930 et 1940. La photographie et le cinéma tentent de rapporter la réalité des événements en informant le public des soubresauts du monde. Les travaux de Robert Capa notamment, pendant la guerre civile espagnole en 1936, marquent le début de cette relation de confiance avec l'image de presse et les photographes de guerre, qui risquent leur vie pour rapporter des images. Après lui suivront David Seymour, Georges Rodger et Cartier-Bresson, fondateurs de l'agence Magnum²⁷, première agence de photographie d'information indépendante à voir le jour après la guerre. Pour ces

²⁷ L'Agence Magnum a été fondée après la libération à Paris par cinq photographes, Robert Capa, David Seymour, Georges Rodger et Cartier-Bresson. Elle tire son nom au Magnum de champagne utilisé pour fêter l'ouverture de l'agence. C'est depuis une des agences les plus réputées au monde.

photographes, le « filet noir » autour de leurs images sera le signe distinctif de l'authenticité des cadrages, de l'absence de manipulation des négatifs.

Du côté des documentaires, les films propagandistes se popularisent, comme ceux de Leni Riefenstahl, dont le plus marquant, *Olympia (Les Dieux du Stade)*, et ses 400 000 mètres de films donneront aux Jeux olympiques de 1936 à Berlin l'exposition rêvée par Goebbels, le chef de la propagande nazie. L'ultime réalisation de Riefenstahl, *Triumph des Willens (Le Triomphe de la volonté)* reste un chef-d'œuvre pictural à la solde d'un régime du mal, un outil au service d'une illusion sordide cherchant à maquiller l'horreur d'une idéologie totalitaire et antisémite.

Aux États-Unis, on fait appel à la photographe Dorothea Lange, dont le travail, notamment son portrait de Florence Owens Thompson, avait été remarqué à l'occasion d'une mission de la Farm Security Administration en 1936 aux côtés de Walker Evans et Russel Lee. En 1942, elle documente, pour le compte de la War Relocation Authority, l'internement des Nippons-Américains qui fait suite à un décret-loi de Roosevelt. Une commande publique chargée de montrer au public les bonnes conditions de traitement de ces citoyens accusés de former une cinquième colonne après l'attaque de Pearl Harbor que Lange, contrairement à Riefenstahl, traite sans aucune compromission.

Dans une candeur presque innocente, les films d'information de l'époque montrent les internés, des citoyens américains de première génération, arriver par bus dans les War Relocation center, qui sont en réalité des camps de concentration aménagés au milieu du désert du Nevada, accompagnés par une musique guillerette et un commentaire rassurant. La réalité des camps est montrée, mais le gouvernement

maquille leur véritable nature. De chaque côté, l'image devient l'outil de ce que Noam Chomsky appelle « la fabrique du consentement »²⁸. On modèle l'opinion publique comme de la glaise en utilisant l'image non plus comme une représentation du réel, mais comme l'incarnation de la réalité souhaitée. « La première victime d'une guerre, c'est la vérité » nous dit Eschyle²⁹, et dans les guerres qui vont se succéder, qu'elles soient armées ou politiques, la vérité va souffrir toutes les tortures.

1.3.2 L'après guerre marque un tournant dans l'histoire du réel représenté

La Seconde Guerre mondiale a pris fin. Dans les travées du tribunal de Nuremberg, plus une âme ne respire. Les soldats qui en surveillent l'entrée, planqués sous leurs grands casques blancs, gardent les yeux rivés droit devant sans oser détourner le regard. Un grand écran de toile, tendu à la hâte par les assistants du procureur en chef Robert Jackson, se dresse contre l'un des murs de la salle étouffante. Il attend là de terribles images, tournées par les opérateurs de la section de l'OSS affectée aux activités cinématographiques qu'un certain John Ford a dirigées quelques mois plus tôt³⁰. Quand l'obscurité emplit la salle, le témoignage noir et blanc, méticuleux, gravé sur pellicule, révèle enfin l'atrocité du crime nazi aux yeux du monde moderne.

Le procès de Nuremberg marque sans aucun doute un tournant dans la conscience universelle. C'est aussi le moment où l'image montre la réalité de la solution finale à une

²⁸ « La Fabrique du Consentement » est un essai co-écrit par Edward S. Herman et Noam Chomsky sur l'industrie médiatique aux États-Unis. Les deux auteurs proposent un « modèle de propagande » pour comprendre la mesure dans laquelle « les médias constituent un système qui sert à communiquer des messages et des symboles à la population ». *La Fabrication du consentement : De la propagande médiatique en démocratie*, Agone, 16 octobre 2008.

²⁹ Les Suppliants, Eschyle

³⁰ En 1945, le réalisateur John Ford encore tout auréolé du prestige de son Oscar reçu pour "Les raisins de la colère" en 1940 et qui a fondé dès 1939 la Naval Field Photographic Unit dans le but de mettre les talents d'Hollywood au service de l'armée, suit l'armée de Patton en Allemagne. Il participe alors à la préparation du Procès de Nuremberg en rassemblant des documents filmés pour l'accusation. Il filmera également le procès.

audience pétrifiée d'horreur et, au-delà, à des nations entières. Les films projetés au cours du procès, dont la production a suivi de strictes procédures de documentation, apportent les éléments de véracité indéniables du crime contre l'humanité orchestré par le III^e Reich. Nous en revenons plus que jamais au « ça a été » de Barthes : une irréfutable réalité qu'ont découvert les cameramans dès la libération des camps. Dans les mains des procureurs, ces films deviennent documents de preuve et incarnent, aux yeux de tous, ce qui s'est réellement produit, venant mettre un terme à un débat de près de trois ans sur l'existence de la « solution finale ».

Bien sûr, tous les films collectés par les équipes de Jackson et ses équipes n'ont pas été utilisés. Certains documents tournés par l'Armée rouge, qui avaient été jugés peu crédibles, furent exclus par peur de manipulations ou de discrédit. Mais, dans cet effort d'exigence et de transparence, la vérité par l'image fut débattue et pour un temps sanctuarisée, au bénéfice d'un procès historique. Certains ont évidemment mis en doute l'authenticité des documents présentés au cours de cette année de procès. Comme Goehring le fit sur le banc des accusés en les qualifiant de « faux grossiers », certaines voix s'élèvent alors pour disqualifier le travail de documentation effectué par le procureur Jackson. Des contestations, des doutes, des remises en question, des mensonges perpétrés pour d'obscures raisons, mais dont les effets ne sont pas dénués de portée.

À partir de ce moment, l'opinion va à jamais se méfier de l'image comme témoignage objectif de la réalité. Le questionnement prédomine, le scepticisme prend le pas définitivement, le regard critique se renforce et le doute devient la norme, quand ce n'est pas de l'indifférence qui s'installe face au flot continu des images qui vont peupler les pages des magazines et le poste de télévision à partir des années 1950.

Sur ce point, Hannah Arendt commente cette indifférence à la vérité de la façon suivante : « le résultat d'une substitution cohérente et totale de mensonges à la vérité des faits, n'est pas que les mensonges seront maintenant acceptés comme vérité, ni que la vérité sera diffamée comme mensonge, mais que le sens par lequel nous nous orientons dans le monde réel – et la catégorie de la vérité relativement à la fausseté compte parmi les moyens mentaux de cette fin – se trouve détruit. »³¹

1.3.3 Kennedy ouvre les portes de l'ère moderne

C'est avec le président John Fitzgerald Kennedy que le rapport de l'opinion à l'image va encore profondément s'éroder. Avec son ascension, l'image est employée à une nouvelle sorte de propagande, allant de la manipulation politique à la documentation historique. De l'épisode de la Baie des Cochons à la crise des missiles russes, des premiers pas sur la Lune à la guerre du Vietnam, chacun de ces épisodes historiques a été l'objet d'une documentation iconographique controversée, où manipulation, illusion et objectivité se sont côtoyées sous les yeux méfiants du public occidental. L'alunissage de Neil Armstrong et Buzz Aldrin reste d'ailleurs encore à ce jour un sujet de controverse incroyable et passe pour une supercherie grotesque dans les sphères complotistes.

Le 22 novembre 1962 met fin à cette longue séquence dans une sorte d'apothéose morbide. Sous les yeux d'une nation entière, le président Kennedy est assassiné en direct. Les images entourent l'événement, le documentent sous tous les angles possibles sans pour autant lever le voile de suspicion qui plane dans l'esprit du public. Le film

³¹ Hannah Arendt « Vérité et politique », dans La crise de la culture, folio poche p. 327-328

d'Abraham Zapruder tourné à l'aide d'une Bell & Howell, modèle 414 PD 8 mm Zoomatic série « Director » achetée quelque mois auparavant, montre dans le détail les 26,6 secondes qui ont conduit à la mort du Président. Objectif témoignage de ce qui a été, pris dans l'innocence du moment par un banal témoin de l'histoire, le film est devenu l'objet de tous les soupçons de manipulation. Rapidement, il part entre les mains des services secrets pour être développé puis copié, avant de revenir dans les mains de Zapruder, qui le vend quelques semaines après à *Life Magazine*. Seules les images fixes seront publiées (hormis l'image 313, celle montrant le crâne de Kennedy éclater sous l'impact d'une balle). Il faudra attendre 1970 pour qu'il soit diffusé à la télévision, dans l'émission *Underground News* de Chuck Collins sur WSNS-TV Chicago et 1975 pour que ABC le diffuse à une audience nationale dans *Good Night America*, présenté par Geraldo Rivera. Rien d'anormal en réalité quand on considère le choc que cet événement a représenté dans la psyché américaine, mais une porte ouverte aux doutes et mises en cause pour tous les tenants du complot.

Des années 1960 à la fin des années 1990, le rapport que la société occidentale entretient avec l'image évolue vers un scepticisme chronique, que nombre de documents manipulés ne font que renforcer. Avec l'arrivée de l'informatique et des ordinateurs personnels, le développement des logiciels de retouche d'image et la création du Web en 1989, le tangible, le support analogique se dématérialise peu à peu pour ne devenir plus qu'une suite de 0 et de 1. La technologie rend l'image manipulée de plus en plus vraisemblable.

La trajectoire décrite dans ces pages montre que le rapport entre le public et l'image se complexifie avec les époques. La recherche de réalisme qui pousse l'artiste à l'observation minutieuse de son environnement dans un effort de reproduction millimétrique cède rapidement la place à une subtile magnification de ce réel, qui devient un objet pétrissable dont on peut à loisir s'accommoder pour en changer les contours. La photographie comme le cinéma appartiennent à des genres qui prétendent reproduire scientifiquement le réel, mais dont la perfection, relativement ennuyeuse, invite à l'amélioration sinon à sa mise en scène.

C'est l'usage intentionnel de l'image à des fins d'information ou de propagande qui change la nature du rapport qu'a tissé le public avec l'image au fil des siècles. Cet usage ajoute des échelons de subjectivité à l'objectivité mécanique du dispositif de prise de vue et vient en brouiller la valeur. Entre le regard du preneur de vue et celui du public, qui jadis observaient la même chose, d'autres regards viennent s'intercaler. Ils donnent un sens nouveau à l'image et enrichissent son contexte de compréhension en lui apposant un message adapté aux intentions du diffuseur. Dans son essai *L'Image, ou ce qu'il advint du rêve américain* Daniel Boorstin³², souligne les dérives d'une société de l'image (pour lui, la société américaine, mais que la mondialisation a globalisée) où « la reproduction ou la simulation d'un événement est perçue comme plus réelle que l'événement lui-même ».

Qu'il s'agisse d'information ou de propagande, ces images-messages, dénaturées de leur mission première, suscitent progressivement la confusion des publics et leur

³² L'image : ou ce qu'il advint du rêve américain, Boorstin, Daniel, Paris : R. Julliard , c1963

questionnement. Les motivations, jadis assez simples, deviennent de plus en plus illisibles à mesure que la communication emprunte les codes de l'information. C'est cet entrelacement des genres qui amène le public à perdre la notion de ce qui est réel ou pas et à douter. Ainsi, quand l'époque permet non seulement de brouiller le sens des images, mais également leur fabrique même, grâce aux procédés de manipulation numérique, c'est tout le lien de confiance avec ce que l'on voit qui est remis à plat. La valeur du réel se délite au fur et à mesure dans ce qu'Umberto Eco qualifie de « faux authentique » c'est-à-dire un monde où la valeur du signe prend plus d'importance que la réalité à laquelle il est rattaché et où son échange prend l'ascendant, rendant le réel encore plus difficile à discerner. C'est sur ce nouveau terrain du réel que vont prospérer les *fake news* et les *deepfakes*.

La science et l'information sont particulièrement mises à mal sur ce point, mais si les deux partagent un certain nombre de challenges, nous nous intéresserons ici davantage à l'information et notamment à l'influence sur l'activité des médias et des journalistes chargés de produire cette information pour le bien commun.

2 Les médias synthétiques, les deepfakes et les internets

La vidéo est apparue sur les réseaux sociaux dans la nuit du 3 novembre, deux jours à peine avant le scrutin de 2024. Très vite, elle se répand comme une trainée de poudre sur Whatsapp Messenger, la nouvelle plateforme cryptée qui a pris le pas sur le bon vieux Facebook, sans que personne au sein de la campagne n'en soit alerté. Depuis quelques mois déjà, le réseau de Mark Zuckerberg joue un rôle crucial dans la campagne en propageant de fausses nouvelles au sein de ses groupes privés et sa place dans le débat démocratique ne cesse d'inquiéter. Les images de mauvaise qualité montrent la rencontre entre une jeune femme aux longs cheveux noirs avec deux individus en costume sombre. On peine encore à comprendre qui sont les protagonistes. La discussion semble déjà engagée et, à en juger par ce qui est dit, on comprend clairement qu'il s'agit d'influencer les résultats de l'élection à venir en opérant un virus informatique capable de changer le comptage de voix des machines de vote électronique. Le document fait l'ouverture des Live News sur YouTube.

Sur YCNN, YMSNBC ou YFOX, les commentatrices s'insurgent : comment Alexandria Ocasio Cortez, le renouveau démocrate, l'espoir de toute une nation face à la présidente Ivanka Trump, a-t-elle pu commettre un acte aussi grave ? Malgré les plus fortes dénégations de la candidate socialiste, les plus fidèles des fidèles lui tournent le dos, déçus par de tels agissements. Le *New York Times* et le *Washington Post* lancent aussitôt leur cellule d'investigation sur l'affaire. Tout le monde est pris de court, le temps manque pour vérifier l'authenticité du document. Il ne faudra que quelques jours pour désamorcer la supercherie, mais trop tard, le premier deepfake politique de l'histoire a bouleversé l'avenir de toute une nation. Une deuxième Trump va siéger dans le bureau ovale et plus rien ne peut renverser la situation. Une partie de l'électorat d'Ocasio-Cortez,

bouleversée par les révélations du *NYT*, tente encore de faire annuler l'élection, mais sans preuve formelle d'une manipulation téléguidée par le camp adverse, la tricherie n'est pas avérée. L'autre partie doute encore de la véracité des révélations post-scrutin et tourne définitivement le dos à la politique.

Ce scénario, bien qu'évidemment fictionnel, n'a rien d'improbable quand on considère les efforts déjà entrepris pour discréditer la représentante du 14^e district de New York, Alexandria Ocasio Cortez³³. En janvier dernier, un faux nu de la politicienne a circulé sur les réseaux sociaux avant qu'un petit groupe de vérificateurs ne mette fin à la rumeur. Ce n'est donc pas la nature du contenu qui pose ici problème, bien qu'elle suppose un certain nombre d'enjeux techniques, mais bien la temporalité dans laquelle s'inscrit cette fausse vidéo.

La survenance d'un *deepfake*³⁴ à un moment déterminant, choisi à dessein pour contrecarrer ici une élection, mais aussi peut-être le rachat d'entreprise, le vote d'une loi, la popularité d'une idée ou d'un courant de pensée, l'émergence d'un leader, questionne la réactivité des anticorps naturels de nos sociétés que sont les journalistes. Dans quelle mesure les rédactions auront-elles le temps et les ressources pour vérifier et recouper les informations contenues dans ce *deepfake* ?

Si l'on considère d'ailleurs les dernières déclarations du patron de Facebook Mark Zuckerberg³⁵, qui affiche sa volonté de fermer un peu plus sa plateforme et d'en faire un endroit privé, les journalistes seront-ils à même de détecter ces *deepfakes* avant qu'ils ne

³³ Alexandria Ocasio-Cortez une femme politique américaine. Elle est élue le 6 novembre 2018 représentante du 14^e district de New York à la Chambre des représentants des États-Unis, devenant la plus jeune représentante jamais élue au Congrès américain.

³⁴ Voir définition plus bas

³⁵ Dans un post du 6 mars 2019, Mark Zuckerberg, le patron de Facebook annonce un changement de politique majeur pour la plateforme qui va désormais parier sur les échanges privés et donc refermer les publications à un cercle très restreint.

provoquent l'irréparable ? L'utilisation de Whatsapp pendant la campagne présidentielle brésilienne de 2018 par Jair Bolsonaro³⁶ montre bien les dangers potentiels de cette combinaison *deepfake*/réseaux sociaux et suggère que l'exercice sera compliqué même pour les titres disposant de moyens importants. Après une définition des *fake news* et des *deepfakes*, nous reviendront sur les conditions de leur émergence pour terminer en examinant les enjeux qu'ils soulèvent.

2.1 Naissance et développement du « faux authentique »

Nous vivons la fin d'un monde purement analogique, sans comprendre réellement le monde numérique dans lequel nous entrons. Dans cette partie c'est donc le changement profond que provoque le numérique dans notre rapport à l'image que nous tentons d'appréhender. Pas dans l'usage que nous en faisons, mais bien dans la façon dont nous comprenons le médium. Très vite, la boîte de pandore a été ouverte. Photoshop prend son envol et devient un verbe.

La possibilité de la modification d'une image par un particulier ne passe plus par l'expertise photographique acquise en chambre noire. Les effets spéciaux renforcent leur présence au cinéma, les dinosaures réalistes débarquent sur grand écran. L'image visiblement manipulée devient peu à peu crédible avant de, plus tard, devenir probable aux yeux des plus crédules. Internet commence à viraliser les premiers *memes* et, sans s'en rendre compte, les sociétés occidentales approchent à grands pas du phénomène des *fake news*.

³⁶ Pendant la campagne présidentielle Brésilienne de 2017, le candidat Jair Bolsonaro a fait campagne notamment en utilisant l'application Whatsapp où il a diffusé nombre de fake news.

2.1.1 Fake news, deepfakes, médias synthétiques

Une étude rapide de la littérature professionnelle, technique et scientifique abordant le problème des *fake news* et des *deepfakes* conduit à une observation : les termes pour désigner ce nouveau type de média sont vagues et non explicites. Il s'agit donc dans un premier temps de définir les contours de ces notions encore vagues, héritées pour la plupart de l'anglais. Nous le verrons plus tard, la notion même de *deepfake* comporte une dimension à la fois sémantique et technique tout à fait différente de ce qu'on désigne aujourd'hui par le terme *fake news* (ou *infox*³⁷, fausses informations en français). Revenons donc à l'essentiel des définitions.

2.1.2 La nature profonde des « fakes »

Le terme *fake news* constitue un point de départ intéressant, puisqu'il partage avec les *deepfakes* le même terme : « fake ». Celui-ci désigne quelque chose de faux en anglais, quelque chose de truqué ou qui se substitue à un « original » dans un but trompeur. Il s'agit ici déjà de manipuler le réel, de le modifier et d'en donner l'illusion de véracité. Il s'agit de tromper en s'attribuant tous les codes du réel pour rendre indiscernable le *fake* de l'original. Sur les forums, on parle volontiers de *fake* identité lorsqu'un utilisateur prétend être quelqu'un d'autre. Dans le contexte qui nous occupe, à savoir celui de l'information, le terme *fake news* désigne évidemment une fausse information.

³⁷ Lorsqu'il s'agit de désigner une information mensongère ou délibérément biaisée, répandue par exemple pour favoriser un parti politique au détriment d'un autre, pour entacher la réputation d'une personnalité ou d'une entreprise, ou encore pour contredire une vérité scientifique établie, on pourra recourir au terme « information fallacieuse », ou au néologisme « infox », forgé à partir des mots « information » et « intoxication ». JORF n°0229 du 4 octobre 2018 texte n° 113

L'expression apparaît dans la sphère publique en 2016 à l'occasion de la campagne référendaire du Brexit en Grande-Bretagne et pendant la campagne présidentielle américaine qui verra l'élection de Donald Trump en novembre de la même année. Nombreux sont les observateurs à s'accorder sur le fait que les *fake news* ont fait basculer ces scrutins et que le rôle des réseaux sociaux, notamment Facebook et Twitter, fut prépondérant dans la dissémination et l'entretien de ces fausses nouvelles. Des révélations par la presse ont par ailleurs montré ultérieurement la responsabilité de ces réseaux et les enjeux de pouvoir qui animaient ces campagnes de propagation de *fake news*. Il faut donc comprendre que le terme *fake news* ne se limite pas à désigner une fausse nouvelle : sa dimension excède ce périmètre et englobe de facto l'intentionnalité du message et son moyen de propagation.

Les *fake news* participent d'une seule et même intention de désinformation volontaire et, si elles sont parfois relayées sur des canaux de diffusion traditionnels (presse écrite, radio, télévision), ce sont davantage les réseaux sociaux et les messageries privées qui en sont le véhicule. Le moteur de cette propagation reste bien entendu le public lui-même, qui répond au caractère souvent polémique des *fake news*. L'émotion joue à plein dans cette dynamique de partage où deux facteurs entrent en jeu : la surpondération de la valeur informative des *fake news* et le degré de méfiance envers l'autorité morale chargée de contrôler ou contredire les fausses informations. Mais toutes les *fake news* sont-elles égales ?

Il est possible de catégoriser l'ensemble de ces fausses nouvelles en sept grands groupes, comme l'a fait l'organisation First Draft³⁸. On distingue alors la parodie, les

³⁸ First Draft est une organisation fondée par Claire Wardle qui se consacre à soutenir les journalistes, les universitaires et les technologues qui s'efforcent de relever les défis liés à la confiance et à la vérité à l'ère numérique.

contenus délibérément trompeurs destinés à discréditer une information ou un individu, l'imposture, le contenu fabriqué à 100%, les connexions douteuses entre deux informations destinées à faire émerger un sens nouveau, la contextualisation trompeuse et la manipulation du contenu. Cette typologie répond à des motivations particulières qui englobent de mauvaises pratiques journalistiques, la propagande, la recherche de profit ou d'influence politique, la provocation, la passion, la défense de causes spécifique ou simplement l'humour.

Les *fakes news* ne sont pas nécessairement complexes. Il peut s'agir d'un chiffre modifié pour donner l'impression d'un rapport de force différent, ou d'une citation inventée pour discréditer un individu, ou bien encore d'une légende inventée pour donner un sens nouveau à l'image qu'elle accompagne. En revanche les *fake news* se doivent d'être virales et sont pensées dans ce but. La typologie des groupes véhiculant les infox montre d'ailleurs l'intérêt de cette viralité : groupes extrémistes, terroristes ou radicaux, complotistes, puissances étrangères, gouvernements étrangers, entreprises concurrentes... il s'agit d'influencer le cours des événements pour en tirer un bénéfice direct ou indirect.

Considérons à présent ce qu'on appelle communément les *deepfakes*. Nous le verrons plus bas, l'origine du nom ne repose absolument pas sur les mêmes bases. Bien qu'il s'agisse de faux, le *deep* fait davantage référence à la technique employée pour fabriquer ce type de contenus qu'à la profondeur supposée de la manipulation (même si celle-ci est souvent corrélée).

Son action se porte principalement sur le fact-checking en collaboration avec les médias et les plateformes sociales les plus importantes comme Facebook ou Google.

Si, pour des raisons de commodité, nous continuons de parler de *deepfakes* dans le reste de ces pages, nous aurons en tête en réalité ce que certains qualifient désormais de « média synthétique » et dont la signification a été notamment établie par le groupe de travail réuni par les organisations Witness et First Draft, dont les considérations ont fait l'objet d'un rapport intitulé *Mal-uses of AI-generated synthetic media and deepfakes: pragmatic solutions discovery convening*, publié en juin 2018³⁹.

Cette notion de « médias synthétiques » recouvre un ensemble de réalités plus large et permet d'embrasser la diversité des scénarios possibles concernant la montée en puissance des *deepfakes* dans l'écosystème du Web et des médias. Ceci étant posé, on peut donc considérer comme « média synthétique », tout média, vidéo ou audio, généré ou modifié par un algorithme. Cette définition permet de développer une typologie des cas de *deepfakes* les plus fréquemment rencontrés sur le Web. Trois possibilités s'offrent à nous pour réaliser cette typologie.

- La première consiste à ne considérer que les usages pour déterminer dans quelle situation un *deepfake* peut-être employé (classification proposée par Tormod Dag Fikse dans sa thèse⁴⁰).
- La seconde consiste à ne considérer que leur fondement technique pour envisager comment le *deepfake* peut être produit et à l'aide de quel outil (classification proposée par le rapport Witness/First Draft)
- La troisième consiste à croiser les deux premières pour tenter d'obtenir un scénario crédible d'utilisation.

³⁹ Mal-uses of AI-generated synthetic media and deepfakes: pragmatic solutions discovery convening, par Firstdraft, 2018

⁴⁰ Imagining Deceptive Deepfakes, an ethnographic exploration of fake videos par Tormod Dag Fikse, 2018

Les *deepfakes*, dans leur diversité et leur originalité technologique, posent non seulement des problèmes éthiques, mais questionnent notre rapport à l'information et à la réalité. Une entreprise peut les utiliser afin de déstabiliser une entreprise rivale et lui ravir des parts de marché, un gouvernement peut les employer dans l'optique de déstabiliser une organisation politique interne ou étrangère, la réputation de célébrités ou de citoyens ordinaires peut être engagée sur la foi de vidéos compromettantes, des idéologies peuvent se propager en véhiculant toutes sortes de stéréotypes.

2.1.3 Histoire de l'intelligence artificielle, des réseaux neuronaux et de l'apparition des *deepfakes*

Il s'agit à ce moment de retracer avec exactitude la chaîne d'événements qui a conduit à l'émergence des médias synthétiques, plus connus sous le nom de *deepfakes*. Une chaîne d'événements qui nous conduit tout droit vers le milieu de la Seconde Guerre mondiale quand, en 1943, Walter Pitts, un logicien américain et Warren McCulloch, un chercheur en neurologie à l'origine de la cybernétique, publie *A Logical Calculus of Ideas Immanent in Nervous Activity*, dans lequel ils proposent un modèle mathématique permettant de combiner algorithmes et neurosciences pour répliquer le processus de pensée humaine. Un modèle révolutionnaire pour l'époque, qui va mettre Alan Turing⁴¹, célèbre pour avoir décrypté le code derrière la machine Enigma servant de base de communication aux nazis pendant la guerre, sur la voie de l'apprentissage des machines et de l'intelligence artificielle.

⁴¹ Alan Turing est un mathématicien et cryptologue britannique, auteur de travaux qui fondent scientifiquement l'informatique. Il a notamment aidé à décrypter la fameuse machine nazie « Enigma » qui servait à encrypter les communications allemandes pendant la seconde guerre mondiale.

Ce n'est qu'en 1965, après les travaux de Turing et Frank Rosenblatt, qui posent les bases techniques du réseau de neurones, qu'Alexey Ivakhnenko crée le premier réseau d'apprentissage profond (*Deep Learning Network*⁴²) et invente le GMDH (Group Method of Data Handling), une famille d'instructions algorithmique permettant d'analyser des données sur le modèle des réseaux de neurones biologiques. Ivakhnenko et son équipe parviennent à créer un réseau d'une profondeur de huit couches (en 2017, la profondeur était de 1 000 couches) et démontrent dans le même temps la capacité d'un ordinateur à « apprendre » automatiquement. Quelques progrès seront faits entre les années 1970 et 1990, dont la reconnaissance de récurrences visuelles en 1980 par le Neocognitron de Kunihilo Fukushima, mais c'est avec la montée en puissance des capacités de calcul des ordinateurs personnels au début des années 1990 que les choses commencent à rapidement évoluer.

Yann LeCun⁴³, chercheur français auteur d'une thèse à la base de l'apprentissage automatique et de l'apprentissage profond, Christopher Watkins et d'autres vont alors grandement contribuer au développement de ce qu'on appelle communément l'intelligence artificielle. En 2012, le concours de reconnaissance d'images organisé par ImageNet (imagenet est une base de données regroupant des millions d'images à classer par catégories destinées à l'entraînement des réseaux neuronaux profonds) est remporté pour la première fois avec un réseau neuronal profond. C'est la renaissance du Deep Learning.

⁴² Le « deep learning » fait partie d'une famille de méthodes d'apprentissage automatique fondées sur l'apprentissage de modèles de données.

⁴³ Yann Le Cun est un chercheur français en intelligence artificielle et vision artificielle. Il est considéré comme l'un des inventeurs de l'apprentissage profond et travaille désormais pour Facebook.

En 2014, deux innovations dans le domaine de l'intelligence artificielle et des réseaux de neurones permettant l'avènement des *deepfakes* bouleversent le monde de l'apprentissage profond. La première se nomme Deepface. Développé par Facebook sous la supervision du français Yann Lecun, le système utilise les réseaux neuronaux pour identifier les visages avec une précision de 97,35 %. Le saut qualitatif est gigantesque puisque cela représente une amélioration de performances de 27 % par rapport à la précédente solution et rapproche la machine de l'humain dont la précision est de 97,5 %. La seconde innovation, les Réseaux antagonistes génératifs (en anglais *Generative Adversarial Networks* ou GAN), nous vient de Ian Goodfellow, un chercheur en intelligence artificielle (I.A.) américain employé par Google Brain, la branche d'Alphabet (maison mère de Google) en charge des recherches sur l'intelligence artificielle. Ces GANs sont des algorithmes d'apprentissage « non-supervisés » qui permettent de créer des images extrêmement réalistes. C'est leur autonomie de fonctionnement qui les distingue des autres types d'algorithmes. Les GANs reposent sur deux réseaux qui sont placés en compétition l'un avec l'autre. Le premier va générer un certain nombre d'échantillons à partir d'une base de données, c'est le « générateur ». Le second reçoit des échantillons réels et fabriqués et tente de déterminer s'ils sont issus du générateur ou s'il s'agit de véritables échantillons, c'est le discriminateur. Sans eux, pas de *deepfake*. Les résultats sont prometteurs et peu à peu suscitent l'intérêt d'autres chercheurs désireux de réaliser des applications pour l'industrie cinématographique.

En 2016, un groupe de cinq chercheurs appartenant à l'université Erlangen-Nuremberg, au Max-Planck-Institute for Informatics et à l'université de Stanford, publient *Face2Face: Real-Time Face Capture and Reenactment of RGB Videos*, dans lequel ils détaillent leurs travaux sur la manipulation d'images vidéo en temps réel. Au cœur de leur travail, la possibilité de transférer les expressions du visage d'une personne, un

acteur par exemple, sur un avatar numérique créé en temps réel. Les exemples associés à la publication sont saisissants. On y voit des vidéos de Daniel Craig et d'Arnold Schwarzenegger, de Georges W. Bush et de Donald Trump dont les expressions du visage sont calquées artificiellement sur les mouvements du visage des chercheurs testant leur solution. La même année, les studios Disney sortent *Rogue One*, épisode *prequel* de la franchise *Star Wars*, dans lequel apparaissent une jeune Carrie Fisher et un Peter Cushing ressuscité (l'acteur est mort en 1994), dont les visages ont été entièrement recréés par ordinateur à l'aide d'un procédé de capture du mouvement très perfectionné. Si le résultat a suscité de vives réactions de la part de la communauté, ce n'est pas à cause du réalisme des captures qui, trop rigides, ne transmettaient pas l'essence du jeu des acteurs. En comparaison, le résultat obtenu par le groupe de cinq chercheurs avec un investissement considérablement moins important a provoqué l'incrédulité de nombre d'observateurs. La même année, Adobe dévoile lors de son Adobe Max Creativity Conference le projet « VoCo » capable de reproduire fidèlement à partir d'un échantillon de 20 minutes la voix de n'importe quel individu, et Adobe Sensei, l'intelligence artificielle associée au logiciel de retouche d'images Photoshop. La puissance de l'intelligence des réseaux de neurones montre ici tout son potentiel pour l'industrie créative et préfigure déjà des évolutions à venir dans ce secteur.

Quelques mois plus tard, en 2017, trois chercheurs de l'université de Washington publient *Synthesizing Obama: Learning Lip Sync from Audio*, un papier dans lequel ils décrivent leur méthodologie pour synchroniser une piste audio A à une vidéo B tout en faisant en sorte que les mouvements labiaux de la vidéo B collent aux mots prononcés dans la piste audio A. Une fois encore le résultat donne matière à réflexion tant il est impressionnant. Le président Obama tient un discours qu'il n'a jamais tenu dans ces circonstances ni dans le lieu où il se trouve. L'illusion est parfaite et la production d'une

telle vidéo n'a finalement coûté que peu d'efforts quand, une dizaine d'années auparavant, il aurait fallu une armée de techniciens de l'image pour aboutir à un tel résultat.

En novembre de la même année, tout bascule. Sur le réseau social Reddit, où se retrouvent les communautés de gamers, les geeks et autres dépositaires de la sous-culture pop véhiculée par le Web, un utilisateur anonyme agissant sous le pseudonyme « u/deepfake » crée le sous-groupe (*subreddit*) r/deepfake. Dans ce groupe, il partage des vidéos pornographiques dans lesquelles le visage des actrices X a été remplacé par celui de célébrités issues du cinéma hollywoodien. La méthode employée pour réaliser ces manipulations repose très clairement sur le *deep learning* ainsi que l'indiquent certaines archives retrouvées sur Github et anonfile. Celles-ci montrent que u/deepfake a téléversé sur le réseau un fichier compressé relativement modeste comprenant 17,6 Mo d'images de référence de Donald Trump et Nicolas Cage (déjà connu dans cette communauté pour être un sujet favori de création de *mèmes*⁴⁴) ainsi que 398 lignes de code écrites en python réparties dans sept fichiers différents, dont presque la moitié (170) font appel à des bibliothèques d'apprentissage profond (*deep learning*) en libre accès telles que Keras et TensorFlow, mises à disposition par Google. L'ensemble repose sur des algorithmes proches de ceux présentés par le fabricant de cartes graphiques Nvidia à la « Conference on Neural Information Processing Systems » en 2017.

Rapidement, le phénomène prend de l'ampleur et dès le mois de décembre 2017, les premiers articles apparaissent dans la presse spécialisée et grand public, notamment

⁴⁴ Un mème est un terme inventé par Richard Dawkins dans son « Gène égoïste » (1976). Il s'agit d'un bout de culture pop qui se propage sur internet principalement sous la forme d'une image ou d'une vidéo. Le mème utilise des visages populaires pour les mettre en scène de façon extravagante. Nicolas Cage, Rick Astley ou Donald Trump sont des figures très utilisées.

dans les colonnes du *New York Times*, du *Washington Post*, du *Guardian*, de *Motherboard*, de *Wired*, du *Times* et sur la BBC. Les articles ainsi recensés s'inquiètent du phénomène encore presque exclusivement centré sur la fabrication de courtes séquences pornographiques. Les auteurs y évoquent que les actrices Gal Gadot, Maisie Williams, Aubrey Plaza, Scarlett Johansson, Taylor Swift et Daisy Ridley, victimes de cette première vague. Au fil des semaines, le phénomène semble gagner en popularité.

Le 2 février 2018, quelques jours avant d'être définitivement banni de Reddit, le groupe *r/deepfake* compte 88 158 lecteurs, mais l'engouement semble fort puisqu'en suivant un lien depuis la page archivée par *archive.is* on découvre une page datant du 7 février 2018 où le nombre de lecteurs passe à 91 313 lecteurs soit une augmentation de 3155 lecteurs en cinq jours (une moyenne de croissance de près de 600 lecteurs/jours). D'autres communautés se forment ailleurs sur le Web. *Deepfake.cc* compte au 12 février 2018 un total de 6391 membres avant sa fermeture quelques semaines plus tard. *Mr Deepfake*, communauté encore active, compte au 16 mars 20 192 234 membres (soit 1830 membres gagnés depuis le dernier relevé en date du 3 mars 2019, moyenne de croissance de 140 membres/jour). *GifFakes* sur Reddit compte encore un total de 10 252 membres inscrits. Difficile d'évaluer avec certitude la taille réelle de la communauté, les doublons sur différentes communautés, les groupes bannis et les reports sur des groupes plus larges n'aident pas à une juste estimation, cependant on peut estimer qu'un minimum de 90 à 100 000 personnes a été exposé et/ou on prit part à la fabrication de deepfakes à travers ces groupes, ce qui constitue un nombre conséquent d'individus en état de comprendre les techniques employées et capables de les utiliser.

Dans la même période, entre janvier et février 2018, les plateformes Discord, Gfycat, et les sites Pornhub⁴⁵ et Twitter bannissent les *deepfakes*, quand d'autres tentent de prospérer sur cette nouvelle mode et ouvrent des sites dédiés aux *deepfakes* pornographiques.

Loin d'avoir éradiqué le phénomène, ces actions de contingence ont éveillé la curiosité de nombreux médias, exposant du même coup le grand public à ce phénomène nouveau. Un rapport de la société Deepttrace publié fin 2018 indique que le nombre de recherches pour le terme « *deepfake* » sur Google a connu une forte croissance entre novembre 2017 et novembre 2018. Le nombre de requêtes a explosé de moins de 200 000 à près de 500 000 entre mai 2018 et novembre 2018. Un dispositif de recherche développé par un développeur du journal *Libération* pour les besoins de ce mémoire a permis de récolter des données entre janvier et mars 2019 et elles confirment la tendance à la hausse du nombre d'articles écrits sur le sujet en anglais et en français.

Le même rapport publié par Deepttrace se penche sur la taille de la communauté de développeurs travaillant sur les projets en libre accès permettant de créer des *deepfakes*. Là aussi, le constat est sans appel puisqu'on constate une croissance constante du nombre de « commits » (une modification du code validé par la communauté en vue d'en améliorer la performance ou d'y ajouter des fonctionnalités). En décembre 2018, le rapport indique pour le répertoire git *deepfakes/faceswap* un nombre de « stars » (personnes suivant le git) à 7728 et un nombre de « forks » (versions créées) de 2307. À la mi-mars 2019, le nombre de « star » pour le même répertoire monte à près

⁴⁵ Pornhub est un site pornographique canadien qui diffuse des vidéos pornographiques gratuitement. En 2019, le site se place à la 28^{ème} place des sites les plus visités au monde, à 4 places de Netflix.

de 16 103 et le nombre de « fork » s'élève à 5373. La seule période de février 2019 totalise 69 commits contre 22 en janvier, 23 en décembre 2018 et 20 en novembre 2018.

La communauté autour des deepfakes est désormais comparable à celle de facebookresearch/detectron (19 502 stars et 4081 forks, chiffre relevé au 16 mars 2019). Les tendances sont les mêmes lorsqu'on observe le nombre de publication sur les Réseaux antagonistes génératifs ou « GAN ». En 2016, on en comptait à peine une cinquantaine, en 2017 presque 500 et en 2018 près de 900. Par contraste, seules 25 publications dans le monde portaient en 2018 sur les contre-mesures anti-deepfakes.

Rien n'indique donc à ce stade que l'engouement pour ces nouvelles techniques retombe au contraire, tous les indicateurs suggèrent qu'une accélération de la croissance et de l'activité des communautés de développeurs.

2.2 L'ère des armes de « deception ⁴⁶ » massive

Les élections américaines de 2016 furent donc le théâtre des pires manipulations, les fameuses « fake news » font leur apparition. La nation Facebook a découvert à cette occasion les pires vidéos sur Hilary Clinton et Bernie Sanders sans se douter un instant que derrière elles se cachaient des fermes à troll commandées par les services secrets russes. Twitter est devenu un terrain de jeu pour les extrêmes de tout poil, les montages photo grossiers se sont multipliés pour discréditer soit une actrice, soit un politicien, une activiste ou un scientifique. Des groupes de centaines de milliers de membres actifs se sont formés spontanément sur ces réseaux pour échapper dans un mouvement de défiance inédit aux circuits de l'information vérifiée. Ils sont rapidement devenus les

⁴⁶ Le terme « deception » est ici utilisé pour ce qu'il est, un mot anglais signifiant tromperie ou supercherie. Le terme français n'a pas été retenu pour de simples raisons esthétiques.

cibles idéales pour faire grandir toutes sortes de théories conspirationnistes qu'un contexte plus que troublé favorise grandement. En 2018, l'heure des Deepfakes a sonné.

2.2.1 Les fondements d'une ère de tromperie numérique

Les armes de « deception » massive que toute personne attachée à un processus de vérification rigoureuse de l'information et à son rôle dans l'établissement d'un consensus collectif redoute sont là. Elles sont filles de l'improbable mariage entre intelligences artificielles, Photoshop et les « trolls⁴⁷ » des fermes à clics. Les *deepfakes* inquiètent. Les médias commencent à en parler. Les services de fact-checking, trop occupés à vérifier les *fake news* existantes, négligent le phénomène et pour cause, pour le moment, rien de bien méchant n'est encore à mettre au compte des *deepfakes*. Deux ou trois célébrités se plaignent de voir leur visage collé au corps d'actrices porno, Nicolas Cage devient Indiana Jones, un fake Barack Obama nous prévient des dangers de ces médias synthétiques et c'est à peu près tout. Mais décrire ce que sont les *deepfakes* ne permet pas de comprendre ce qui passe en réalité. À ce stade, c'est la fabrique du consensus de « ce qui est » et de « ce qui n'est pas vrai » qui se délite petit à petit sous nos yeux. Le fil qui nous lie et nous permet de faire société s'effiloche.

Un fossé irrémédiable se crée entre, d'une part, la capacité de la technologie à créer des images et des sons synthétiques crédibles aux yeux et aux oreilles de tous et d'autre part, ce que le public connaît de ces avancées technologiques et de leur capacité à créer de parfaites illusions. Ni le public, ni les professionnels de l'information ne sont désormais en capacité de reconnaître une image manipulée d'un simple coup d'œil. Aucun d'entre eux n'a reçu l'éducation nécessaire à ce type d'analyse. Plus grave, la

⁴⁷ Le terme troll désigne, dans le jargon de l'internet, un personnage malfaisant dont le but premier est de perturber le fonctionnement des forums

délégation de la vérification des contenus à un « agent d'authentification » incarné aujourd'hui par la presse combinée à l'accélération des technologies reposant sur des algorithmes puissants que seuls les spécialistes du domaine peuvent appréhender, prive l'ensemble de la collectivité des outils de décryptage essentiels de leur environnement médiatique immédiat.

Ainsi lorsque ces « agents d'authentification » que sont les journalistes sont bernés par l'ultra-réalisme de certaines créations visuelles, ou que, leur scepticisme est mis à mal par des vérifications dont ils ne maîtrisent pas immédiatement les processus, ce n'est pas leur propre réalité qui est altérée, mais bien, par un phénomène d'amplification massif lié à la nature même des médias, celle de millions d'individus. Quand par exemple, pendant la campagne présidentielle américaine de 2004, une photo place John Kerry aux côtés de Jane Fonda dans un rassemblement contre la guerre du Vietnam, personne ne détecte la supercherie. Photoshop n'existe que depuis une petite dizaine d'années et personne n'imagine que ce type d'image puisse être issu d'une manipulation. Personne n'imagine cela faisable. Personne n'a suffisamment encore été exposé à ce type de canular pour, dans un réflexe pavlovien salvateur, associer immédiatement le document à une manipulation.

2.2.2 L'algorithme fabricant d'illusions

Comment imaginer dès lors que l'on puisse détecter de plus subtiles manipulations ? Comment imaginer que nous puissions d'une simple écoute détecter les voix synthétiques développées par Lyrebird, Adobe⁴⁸ Voco ou Baidu Deepvoice qui sont capables de reproduire le ton, la sonorité et la tessiture de la voix de n'importe quel

⁴⁸ Adobe est une entreprise américaine leader dans le domaine de la retouche d'image. Son logiciel phare est photoshop, un des logiciels les plus piraté de l'histoire du web.

individu ? Comment imaginer que nous puissions discerner à l'œil nu les changements appliqués à une image ou une vidéo par Abodbe Cloak ou Pixelmator ? Comment penser que nous soyons suffisamment éduqués à ce type de techniques pour remarquer le « faceswap⁴⁹ » effectué sur une vidéo à l'aide de Face2face ou Deepvideo portrait ? Peut-on simplement comprendre à la vue d'un Obama plus vrai que nature que les mots qu'il prononce dans une vidéo ne sont pas les siens, mais ceux reproduits par un algorithme spécialisé ? Dans certains laboratoires, les algorithmes parviennent même à recréer des visages entiers en 3D à partir de photographies et à les animer de façon réaliste en leur prêtant des expressions faciales réalistes.

Bien sûr pour autant d'exemples de supercherries non détectées on peut trouver dix autres exemples montrant que les vérifications sont efficaces, mais la popularité de plus en plus forte de certaines fake news fait craindre que nous soyons arrivés à une situation de plateau, un moment où médias et journalistes ont atteint leur pic d'efficacité maximum. Si leur popularité et leurs audiences continuent de côtoyer les fonds de sondage, les médias n'auront à leur disposition que peu de moyens pour se battre contre des deepfake très élaborés.

Certainement qu'une course à l'armement va s'engager entre les créateurs de médias synthétiques malintentionnés et la communauté chargée d'authentifier les contenus. Certains chercheurs comme le français Vincent Nozick ou les chercheurs américains de la DARPA⁵⁰ développent des outils de détection de ces contrefaçons high-tech. Certaines vérifications physiques permettent de tirer rapidement les choses au clair, l'angle de la lumière, le déplacement des ombres, certains comportements physiques de

⁴⁹ Le faceswap est une technique de photomontage permettant d'échanger deux visages entre eux.

⁵⁰ La Defense Advanced Research Projects Agency, est une agence du département de la Défense des États-Unis chargée de la recherche et développement des nouvelles technologies destinées à un usage militaire.

l'image. D'autres reposent sur l'utilisation d'intelligences artificielles, notamment les techniques d'amplification vidéo Eulerienne⁵¹ qui permettent de voir les distorsions physiques imperceptibles à l'œil nu. À mesure que le corpus de médias synthétique va augmenter, les algorithmes de détection auront la capacité de travailler sur des bases de données de plus en plus larges, améliorant ainsi leur capacité d'identification de contenus détournés. Mais à mesure que ces techniques de détection évolueront, les algorithmes de création de contenus manipulés seront capables de produire des résultats intégrant les dernières méthodes de détection pour rendre leurs productions encore plus crédibles. Le jeu n'a pas de fin.

2.2.3 Quand il se sent attaqué, le corps réagit

Nous ne verrons probablement pas arriver le premier deepfake d'information et nous réaliserons – certainement trop tard – que nous avons été manipulés. Dans un contexte où la relation à l'information devient de plus en plus troublée, la survenance des médias synthétiques constitue une menace supplémentaire sur les équilibres des démocraties occidentales. L'effet de propagation exponentielle des réseaux sociaux, combiné à un problème global de désinformation et d'éducation aux médias et au numérique, à la surpersonnalisation des plateformes, aux techniques addictives de captologie⁵², à l'automatisation des bots⁵³ politiques, à la polarisation des extrêmes et à l'éloignement des audiences des médias historiques agissent comme catalyseurs. Le problème des *deepfakes*, encore vécu par certains comme un terme à la mode « qui ne

⁵¹ Développé par deux chercheurs du MIT, l'algorithme d'amplification Eulérienne permet de révéler des variations subtiles qui sont invisibles à l'oeil nu et notamment celles des pulsations cardiaques sur un visage ou la vibration de matériaux.

⁵² Le terme « captologie » a été inventé par B.J. Fogg de l'université de Stanford. Celui-ci a publié en 2003 *Persuasive Technology: Using Computers to Change What We Think and Do*. La captologie se sert de l'étude des comportements pour permettre à la technologie d'influencer les utilisateurs. Typiquement, le défilement infini du mur de Facebook, les systèmes de récompenses etc font partie de la captologie.

⁵³ Un bot est un mini-robot, un programme qui agit automatiquement. Ils sont principalement utilisés pour effectuer des tâches pénibles sur le web comme « liker » des pages automatiquement ou diffuser des messages massivement.

pose pas de réelle menace » semble dépasser la simple mode si on en juge par les mesures législatives qui émergent ici où là.

En quelques mois les États-Unis, la Commission européenne, la Grande-Bretagne, la France et d'autres ont commandité des rapports sur la situation des fake news, des deepfakes et de leurs effets sur le débat public et l'équilibre démocratique des pays concernés. Au États-Unis, la proposition de loi S3805, connue sous le nom « Malicious Deep Fake prohibition act of 2018 » a été déposée par le sénateur républicain Ben Sasse en décembre 2018 pour pénaliser la création et la diffusion de deepfakes. En France, c'est la « proposition de loi relative à la lutte contre la manipulation de l'information » adoptée le 20 novembre 2018 qui encadre la production et la diffusion de fausses informations. En chine et en Biélorussie des lois ont également été votées pour restreindre l'utilisation de deepfakes et la diffusion de fake news. Ce sont au total près d'une cinquantaine de pays qui ont pris des mesures contre la propagation de fausses informations, mais les motivations sont diverses et certaines d'entre eux des plus autoritaires comme la Chine, la Biélorussie, la Russie, les pays du Golfe, certains états africains et d'Asie en ont profité pour porter atteinte à la liberté d'expression et museler des oppositions trop bruyantes. Les autres en revanche cherchent à contre-carrer les potentielles déstabilisations et les usages malintentionnés.

La terrible promesse des deepfakes en ce domaine semble d'ailleurs particulièrement sérieuse. Préjudices contre les personnes ou les associations, les entreprises et les organisations publiques, sabotage économique, manipulation du débat public, érosion téléguidée des institutions, exacerbation des polarisations politiques, sabotage diplomatique et mis en péril de la sécurité des états, sape du rôle des

journalistes la liste des menaces sur la démocratie et la sécurité nationale américaine que dressent Bobby Chesney et Danielle Citron⁵⁴ ne semble plus finir. Le point le plus terrible semble pourtant celui-ci. Dans une sorte de retournement de situation improbable, il n'est pas impossible de voir dans un débat l'un des intervenants, poussé dans ses retranchements par un document authentique, s'exclamer qu'il s'agit d'un deepfake authentique. La confusion dans l'esprit du public serait terrible puisque le doute à l'égard de l'information et du travail des journalistes serait alors porté à son paroxysme.

Il suffit de constater l'accumulation dans l'espace public et notamment sur le web des articles mettant en garde contre les deepfakes et les titres alarmistes qui sont employés : « Le fake porn n'est que la partie émergée de l'iceberg. », « Voices : faites dire ce que vous voulez à des célébrités avec ce simulateur vocal », « L'un de ces visages a été créé aléatoirement par une intelligence artificielle. », « Que faut-il craindre des "deepfakes", ces redoutables vidéos falsifiées ? », « Deepfake : Est-il possible de détecter ces vidéos truquées ? »⁵⁵. La presse prépare son public à faire face aux deepfakes, mais dans le même temps, par le choix sémantique qu'elle opère en rédigeant ses titres et ses accroches, elle instille un sentiment de danger imminent dans l'esprit de chacun.

La nature profondément complexe des médias synthétiques soulève bien entendu des questions qui s'étire au-delà du simple champ technique. Ce n'est pas l'existence des GANs eux-mêmes qui pose les questions éthiques, mais bien les raisons qui ont conduit les chercheurs, ingénieurs et autres mathématiciens à créer ces nouvelles machines qui à leur tour permettent de créer des médias synthétiques trompeurs. Si

⁵⁴ Deep fakes: A looming challenge for privacy, democracy and national security par Bobby Chesney et Danielle Citron, 2018

⁵⁵ Articles du Journal du Geek et de Télérama en date de 2019

« *Code is Law* » (le code est la loi), alors on ne peut s'empêcher de considérer que la loi qui est ici créée devant nos yeux et qui commence à envahir tous les aspects de nos vies jusqu'à en menacer certains aspects, ne répond à aucun des critères de la loi des hommes. Si l'algorithme est capable de modeler notre perception du réel à travers les ans, s'il est capable en toute autonomie de choisir et d'interpréter nos désirs ou de faire dire ou faire des choses qui n'ont jamais existé à un individu. Si l'algorithme permet à des individus de tromper des populations entières alors cet algorithme comporte un risque que la société dans son ensemble doit observer de près. Les laboratoires d'étude ne sont pas des espaces neutres dans lesquels ceux qui y travaillent oublient leurs biais cognitifs, culturels et leurs inclinaisons idéologiques. L'expression du travail de ces femmes et de ces hommes dans ce cas ne reste pas confinée dans les murs blancs des salles de recherches, mais se répandent sur le web à la portée de tous pour le meilleur et pour le pire. Bien sûr, si nous considérons le meilleur, il est évident que l'industrie du divertissement, cinéma, jeu vidéo, télévision et d'autre encore sera le principal client des ces solutions qui sauront nous conduire dans des univers jamais explorés. Mais enfin, pour quel objectif ? Celui de nous amuser ? De nous permettre d'échapper aux réalités morbides que certains annoncent déjà ? Le jeu en vaut-il la chandelle ? En revanche si l'on considère le pire, c'est là tout un univers de manipulation, de déception et d'illusions qui s'offre à nous. Quelles que soient les motivations, économiques ou politiques, cet univers forgé par les médias synthétiques où le doute permanent, le scepticisme, la polarisation sera de mise ne peut être en aucun cas souhaitable. Il relève de la responsabilité des hommes et des femmes qui s'emparent de ces sujets dans les laboratoires ou sur le Web de considérer ces questions avec tout le sérieux qu'elles méritent, mais pour cela il faut qu'elles soient armées des bons outils. La formation individuelle, la culture numérique et une collection d'outils assurant un contrôle minimal

de l'influence de ces technologies semblent nécessaires et c'est ce que nous allons voir dans cette troisième partie.

3 Les scénarios d'une réponse

Nous pourrions facilement dresser dans ces pages un portrait particulièrement sombre du futur qui nous attend. Rien n'invite au premier coup d'œil à faire montre d'un optimisme débridé quant à la place que prendront les *deepfakes* dans l'écosystème médiatique de demain. Au vu de la vivacité de la communauté qui est en train de se former autour des outils de création de médias synthétiques, le problème ne va pas aller en s'amenuisant. La disponibilité des technologies, la tendance à la fermeture des réseaux sociaux et les opportunités de déstabilisation sont autant de facteurs qui promettent aux *deepfakes* un avenir supposément radieux.

Pour autant, il n'est pas question de céder à la panique. La première vague de *fake news* nous enseigne un certain nombre de choses sur notre comportement et notre niveau de préparation, que nous aurions tort de balayer d'un revers de la main. Par ailleurs, si nous approchons la question avec un angle systémique, une variété de solutions s'offre d'ores et déjà à nous. Enfin, considérons la société comme un corps biologique, une entité vivante. Lorsqu'un individu en pleine santé, fort, solide, bien alimenté, croise le chemin d'un virus de la grippe, il passe quelques jours à greloter, ses défenses s'activent avec force, puis il se rétablit rapidement et oublie même jusqu'à avoir été malade. En revanche, lorsqu'un individu affaiblit par une maladie ou une mauvaise alimentation entre en contact avec le virus de la grippe, ses défenses n'agissent pas, dépassées par les événements, elles laissent le virus s'établir, alors les conséquences peuvent être catastrophiques voire, dans certains cas, fatales. Notre corps social est malade, les sociétés occidentales sont affaiblies et ne fonctionnent plus suffisamment correctement pour donner à ces anticorps naturels que sont les médias et les journalistes la force de se battre contre les fausses

informations et les *deepfakes*. Dans cette dernière partie, nous allons dresser le portrait de la menace et suggérer les voies d'actions qui nous semblent aujourd'hui opportunes.

3.1 Solutions technologiques

« Si vous connaissez vos ennemis et que vous vous connaissez vous-même, mille batailles ne pourront venir à bout de vous. Si vous ne connaissez pas vos ennemis, mais que vous vous connaissez vous-même, vous en perdrez une sur deux. Si vous ne connaissez ni votre ennemi ni vous-même, chacune sera un grand danger. »

Sun Tzu, *L'Art de la Guerre*.

Le tableau dépeint jusque là n'invite guerre à l'optimisme le plus béat, mais la technologie n'est en soit ni bonne ni mauvaise et c'est bien son usage malveillant que les journalistes, médias et acteurs publics doivent s'attacher à attaquer. L'essor des outils permettant la création de médias synthétiques participe en réalité plutôt d'un rêve formulé par la Silicon Valley, influencé par les expériences psychédéliques des années 1960 – et notamment celle de l'International Foundation for Advanced Study qui testa le LSD pour le compte de l'armée américaine. Cinquante après le Summer of Love, le LSD fait son grand retour en microdose chez tous les techies et geeks de la zone qui, dans l'espoir de marcher dans les pas de Steve Jobs, le fondateur d'Apple, cherchent à nouveau à changer le monde et les consciences. Les *deepfakes* et le travail sur les GAN, dans leur forme la plus hétérodoxe, s'approchent des premiers travaux sur l'intelligence artificielle et de cette vision naïve de l'innocuité technologique. Il suffit de voir la présentation par NVIDIA, en mars 2019, de son logiciel de création de paysages pour comprendre que le présentateur, emporté par l'enthousiasme, oublie la portée de l'invention qu'il rend publique.

3.1.1 Le framework Lessig comme cadre d'action

Que peut-on donc faire pour contrer les utilisations malveillantes de *deepfakes* ? La référence en matière de réflexion légale sur le Web reste celle du framework imaginé par Lawrence Lessig⁵⁶, professeur de droit et juriste américain, qui fonctionne de la façon suivante. Considérons une cible que l'on doit réguler. Cette cible est contrôlée par quatre composantes différentes. La première est la loi. Dans le cadre des *deepfakes*, on l'a vu plus haut, les États ont déjà fait ce premier pas de régulation et renforceront probablement leur arsenal à mesure que la menace se fera de plus en plus concrète. La seconde est la norme. Si pour le moment l'usage des *deepfakes*, bien que marginal, reste essentiellement cantonné au porno et à la culture pop, rien n'interdit de penser que la norme sociale, l'étiquette sur le Web, poussera à une utilisation plus grand public de ces outils et que les derniers à s'en servir pour des usages moralement répréhensibles seront marginalisés. La troisième composante est l'architecture, le code dans ce cas. Il est essentiellement libre, repose sur la culture open source, chacun est libre de s'en saisir sans entrave, de le manipuler, d'en créer des versions différentes. La dernière composante est le marché. Y a-t-il un marché pour les *deepfakes* ? Probablement. Qu'il soit légal (industrie du cinéma, de la création visuelle) ou non (officine de déstabilisation, farmtroll, etc.), la fabrication de médias synthétiques sera à coup sûr au centre d'un marché économique structuré. Maintenant, considérons l'un de ces éléments en particulier, la loi. Celle-ci peut influencer la norme en condamnant comme elle le fait les comportements déviants considérés comme dangereux. Elle peut également réguler le marché, soit en créant les conditions d'émergence d'un secteur industriel favorable au développement des *deepfakes* ou en régulant le marché du logiciel libre. Enfin, elle peut

⁵⁶ Lawrence Lessig est un juriste américain à l'origine des Creative Commons. Il a développé une théorie établissant que le Code fait loi et qui implique que ceux à l'origine du code ont une action concrète sur la sphère publique.

influer sur l'architecture, c'est-à-dire le code, en régulant la façon dont on s'échange les informations sur les *deepfakes*, en demandant l'installation de codes permettant l'identification de chacune des productions ou en limitant les possibilités des GANs et donc leur intelligence.

3.1.2 Cryptage et sécurisation, éléments de la lutte anti-deepfake

Ce dernier point nous semble crucial puisque c'est certainement la part la moins maîtrisée par l'homme dans le réseau neuronal : l'intelligence. L'architecture des réseaux de neurones compétitifs est bien entendu connue et maîtrisée, mais ce sont les interactions entre ces réseaux qui sont inconnus, particulièrement les signes que l'un reconnaît comme étant un signe distinctif d'une action à réaliser. Par exemple donc, la loi pourra réguler le nombre d'exemples contenus dans un échantillon destiné un réseau de neurones, mais elle ne pourra jamais réguler ce que le réseau de neurones reconnaît, car personne ne sait – pas même leurs créateurs – ce que « voient » ces réseaux. Une anecdote illustre parfaitement cela : on peut réguler le fait que les échantillons ne doivent pas présenter de biais racistes, mais il est impossible de réguler le fait que pour déterminer la différence entre le ping-pong et le tennis, les réseaux observent en particulier la couleur des cheveux, la taille et la couleur de peau des joueurs (le ping-pong étant majoritairement joué par des joueurs asiatiques, du fait de son essor en Chine puis au Japon à partir des années 1950).

En extrapolant ce concept, on peut dès lors imaginer les actions possibles suivantes : la première consiste à empêcher la contrefaçon des documents. Il s'agit là d'appliquer un système inviolable de chiffrement permettant l'identification des contenus créés par un appareil d'enregistrement connu et édités sur une plateforme disposant du

dispositif de cryptage. La voie naturelle ici repose sur la blockchain⁵⁷, qui fonctionne sur un registre de données infalsifiable permettant de transmettre des données sans avoir recours à un tiers de confiance. Cette solution n'interdit pas la modification des contenus, mais en revanche garantit l'authenticité du contenu sur l'ensemble de la chaîne de production.

L'algorithme de la blockchain nécessite de résoudre un problème mathématique complexe pour décrypter les transactions (c'est-à-dire les échanges de contenus) qui nécessitent une puissance de calcul importante par ceux qui tentent de décrypter un block (les mineurs) afin d'obtenir une preuve de travail. Sans trop entrer dans le détail du fonctionnement de la blockchain ici, on notera simplement que l'usage de cette blockchain soulève quelques questions d'ordre éthique (consommation énergétique importante des unités de calcul, appât du gain par l'acquisition de cryptomonnaies, faible nombre de transactions possibles). Deux problèmes se posent pourtant lorsqu'on envisage le déploiement d'une telle solution : il faut que l'ensemble des constructeurs de matériels permettant d'enregistrer du contenu adopte un standard d'encryption pour être certain que tous les contenus produits soient identifiables et traçable, ce qui requiert un consensus international autour d'une norme et d'un protocole technique compatible et accepté par tous. Pas impossible, mais compliqué. Second problème, le traçage et l'identification de tous les contenus produits posent des problèmes fondamentaux de respect de la vie privée et ne sont guère compatibles avec une société en prise directe avec ces questions. Reste que deux sociétés, Civil et Terciv, proposent de combattre les *fakes news* et les *deepfakes* grâce à la technologie blockchain.

⁵⁷ La blockchain est une technologie de stockage et de transmission d'informations sans organe de contrôle centralisé. Il s'agit d'une base de données distribuée dont les informations envoyées par les utilisateurs et les liens internes à la base sont vérifiés et groupés à intervalles de temps réguliers en blocs, l'ensemble étant sécurisé par cryptographie, et formant ainsi une chaîne.

Le Hashgraph, nouveau venu dans le monde des technologies de cryptage, suscite ainsi un intérêt grandissant depuis la publication de son manifeste en 2016. Plutôt mieux sécurisé, il tolère davantage les fautes byzantines naturelles⁵⁸ (les erreurs dues au réseau lui-même) et volontaires (les attaques). Il est également moins gourmand en énergie, son système de calcul ne repose pas sur le consensus général et ne peut être bloqué par défaut de consensus. Enfin, il présente un nombre de transactions largement supérieur à la blockchain (250 000 par seconde) et une capacité à grossir bien supérieure. En revanche, c'est un système encore confidentiel, qui peine à s'imposer devant la blockchain, dans laquelle beaucoup ont d'ores et déjà investi.

Autre solution de cryptage, le taggage des contenus à l'aide d'une fonction de hachage, c'est-à-dire le calcul d'une empreinte unique à l'image d'entrée permettant d'identifier rapidement le contenu avec lequel on est en présence. On peut également dans ces hypothèses penser aux systèmes compétitifs évaluant la sécurité des contenus par l'opposition de deux algorithmes jouant dans deux camps opposés et cherchant à tester les différentes options coût/gains disponibles pour éviter l'apparition de failles dans le réseau.

Un système de vérification en peer-to-peer peut également être mis en place pour réintégrer un élément humain dans la boucle. Plus proches des solutions conventionnelles proposées notamment par les journalistes, ces systèmes de vérification fonctionnent sur des niveaux de faible sécurité qui tolère la correction a posteriori et dans un délai obtenu plus long que ceux obtenus avec un algorithme.

⁵⁸ En informatique, on appelle panne byzantine ou comportement byzantin tout comportement d'un système ne respectant pas ses spécifications,

3.1.3 Les outils sont-ils les seuls levier d'influence ?

Les solutions logicielles destinées aux particuliers et aux professionnels doivent être également proposées afin de développer une mise en pratique des préceptes appris à l'école ou en formation continue. Il existe une collection d'application permettant de vérifier l'authenticité de contenus comme Proofmode, Invid ou Truepic et il est important de développer le soutien à ces solutions et la généralisation de leur usage dans la sphère privée.

Mais, plus généralement, c'est une question philosophique et éthique profonde qui s'impose ici. Peut-on se permettre, sur la question de la vérification des informations, de nous reposer sur des plateformes ou des services dont le code ne serait pas en libre accès, disponible sous licence creative common, et financé par la communauté directement ? En remettant les clefs du Web aux GAFa et BATX, nous abdiquons toute souveraineté sur le Web et refusons de participer à son gouvernement alors même que, comme technologie singulière, le Web ne constitue pas un repli de notre réalité, mais bien un de ses prolongements naturels. Nous ne pourrions pas nous insurger de la prolifération de fausses nouvelles et des médias synthétiques ni de leur conséquence si nous évitons systématiquement et soigneusement de nous intéresser à la fabrique intime du Web et à son animation politique, économique ou philosophique. La privatisation des moyens de communication ne peut que signifier une chose et nous rappelle les tentatives propagandistes connues sous d'autres régimes dans d'autres époques : la perte de contrôle sur nos vies.

Se saisir des questions technologiques et activement militer pour la mise en place de solutions – dont certaines existent déjà – capables d'endiguer le flot de fausses nouvelles et d'authentifier la nature des documents partagés sur le Web relève donc d'une activité citoyenne normale et saine, et repose sur un socle éducatif fort.

3.2 Solutions légales, politiques et éthiques

Il s'agit enfin d'explorer d'autres solutions qui ne relèvent pas de l'usage direct de la technologie et n'entrent pas encore dans le champ que nous pensons réserver exclusivement aux médias et aux journalistes.

3.2.1 La nécessité de l'éducation aux médias

Nous l'avons évoqué, une des plus grandes menaces des *deepfakes* reste la confusion que de tels contenus vont susciter dans l'esprit du public. Il est donc nécessaire, pour les médias notamment, mais plus largement pour la société dans son ensemble, de reconnaître l'éducation aux médias et au numérique comme une priorité tout aussi importante que l'apprentissage de la lecture et des mathématiques élémentaires. Qu'on le veuille ou non, les sociétés occidentales sont entrées dans une ère profondément marquée par la capacité à maîtriser les concepts informatiques les plus complexes. Si sa subsistance à long terme n'est pas assurée du fait de la raréfaction des ressources premières nécessaires à son maintien et à la production de ses supports, il n'en reste pas moins que, sauf preuve contraire, la tendance à l'expansion du numérique dans nos vies n'est pas immédiatement remise en cause. On ne peut donc espérer survivre dans cet environnement profondément marqué par cette pénétration du digital sans en connaître les codes et en maîtriser un socle de connaissance élémentaire, socle qui ne se résume pas à la simple utilisation des réseaux sociaux ou à celle de terminaux particuliers. L'éducation aux médias revêt donc une importance capitale. Il s'agit de faire comprendre ici les dynamiques en jeu lors de la création d'une information, le processus de détection, de collecte, de vérification, d'ajout de perspective et d'analyse, de mise en forme et de diffusion complètement inconnu du grand public. Il s'agit d'ouvrir les portes et de décortiquer en toute transparence la véritable mécanique qui anime encore toutes les rédactions, les enjeux éthiques et moraux, les débats politiques et philosophiques qui les

animent. La question de la crédibilité et de la confiance doit être également centrale dans cette recherche d'éducation aux médias, afin de faire comprendre la nature de la responsabilité qui pèse sur les médias.

3.2.2 La formation des professionnels

Enfin, il s'agit de ne pas cacher les errements multiples, les excès de confiance, les erreurs, les arrogances et les gestions hasardeuses qui expliquent la défiance à l'encontre du quatrième pouvoir. Le tableau ne serait évidemment pas complet sans une étude approfondie de la nature et des mécaniques qui animent les réseaux sociaux. L'analyse des phénomènes de propagation de l'information, les questions éthiques et les bouleversements fondamentaux sur les notions de vie et de sphère privée chéries des sociétés occidentales, doivent être centrale dans nos préoccupations. Tout comme le rôle particulier des réseaux dans la convection démocratique, la polarisation des extrêmes politiques et intellectuels, le phénomène de bulles, l'isolement des individus au sein du collectif, les questions de narcissisme comme celle des comportements tribaux identifiables sur le Web. Voilà en somme un aperçu des connaissances à acquérir afin de former des citoyens complets capable de réfléchir à leur place dans la société et aux enjeux, capables de résister aux manipulations et notamment aux fausses informations comme à tout autre média synthétique.

Pour cela, il s'agit d'effectuer un saut qualitatif majeur dans la formation des adultes, des professionnels. On rencontre là aussi un déficit d'expertise particulièrement marquant qui ne présage rien de bon lorsqu'on se penche un peu sur la question de la survie des médias dits « historiques », c'est à dire pré-Internet. Nombre d'outils sont d'ores et déjà fournis par des organisations telles que Firstdraft, le Google News Initiative ou l'European Journalism Center en Europe et la Knight Foundation aux États-Unis. Mais nombre de ces outils sont issus d'activités directement ou indirectement financées par les

géants américains du Web et ce ne peut être en aucune une solution pérenne. Il s'agit donc là aussi de développer des outils libres destinés aux professionnels sinon supérieurs, au moins égaux en qualité, afin d'assurer une diversité des sources fondamentale. Déjà, Google s'invite dans les écoles françaises pour assurer des cours d'éducation au numérique, palliant ainsi à l'incapacité des enseignants mal formés. Il est impensable de laisser au secteur privé, qui envahit déjà l'espace de l'informatique domestique avec ses solutions – certes excellentes – de bureautique, le monopole de l'éducation numérique.

3.3 Solutions médiatiques

Considérons un instant la situation sous le prisme de la théorie des jeux et notamment à travers le prisme des jeux finis et infinis apporté par James Carse dans son livre *Finite and Infinite Games*⁵⁹, récemment remis au goût du jour par l'entrepreneur américain Simon Sinek. Carse considère qu'il y a deux types de jeux, les jeux finis et les jeux infinis. Les jeux finis sont ceux dans lesquels les participants suivent une règle du jeu, obéissent à ces règles et définissent qui a gagné et qui a perdu à la fin du jeu. Traditionnellement, les activités comme la guerre, la politique, le sport font partie des jeux finis. Dans le jeu infini, il n'y en a qu'un par définition, les participants n'ont qu'un seul but, celui de perpétuer le jeu. Les règles peuvent changer, les limites du jeu également. Les participants aux jeux finis mobilisent leurs forces pour emporter la partie quand les participants au jeu infini parient sur la patience et sur le renoncement des autres joueurs. Un joueur jouant un jeu fini contre un joueur jouant un jeu infini n'a aucune chance de gagner.

⁵⁹ *Finite and Infinite Games*, James P. Carse, 1986

3.3.1 Les journalistes jouent à un jeu « fini ».

Maintenant, revenons à notre sujet, la lutte contre les fausses informations et la propagation des *deepfakes*. En 2016, la réaction de la presse fut unanime et cohérente avec l'ADN de la profession. À travers le monde, les rédactions ont mobilisé leurs forces pour constituer des services entièrement dédiés à la vérification des infox. Les services, en fonction de la taille des journaux concernés, regroupent une dizaine de journalistes pour les plus importants. Bien entendu, considérant la santé actuelle des journaux, tous n'ont pas le loisir de mobiliser ce type de force et s'en remettent aux efforts de leurs collègues dotés de plus de moyens. En considérant qu'un journaliste peut traiter une fausse information et vérifier l'ensemble des éléments d'une histoire par jour, un service de fact-checking d'une dizaine de journalistes peut traiter 200 informations par mois. Ces services sont bien évidemment soutenus par les rédactions et les directions de journaux qui acceptent d'y allouer des ressources financières dans l'espoir, parfois cynique, de gagner des parts de marchés en redorant le blason quelque peu défraîchi par des années de laisser-aller. Les services de fact-checking ne fournissent évidemment pas un travail complémentaire entre les uns et les autres, mais souvent relativement redondant, de sorte que les chances de traiter un grand nombre de fausses nouvelles s'en trouvent réduit.

Les règles du jeu sont donc simples : débusquer les fausses informations où elles se trouvent, les vérifier et publier le compte-rendu de l'opération, puis recommencer encore et encore. À mesure que le nombre de fausses informations augmente, les rédactions doivent augmenter le nombre de journalistes assignés au fact-checking. Lorsque la complexité de la vérification augmente, c'est le nombre de journalistes et les moyens engagés en technologies, logiciels et autres ressources qui croît de la même manière. À un certain point, soit le résultat de ces efforts paie et résulte en une augmentation des abonnements, ce qui permet, éventuellement, d'assurer de nouvelles

ressources et/ou de nouvelles recrues, ou bien rien ne se passe et les moyens comme le nombre de journalistes dédiés au fact-checking stagnent.

3.3.2 Les auteurs de fake news et de deepfakes jouent une partie infinie

Le grand cycle de l'information se range du côté du jeu infini. La vie se charge de fournir chaque jour son lot d'événements propices ou non à la fabrication de fausses informations par l'armée de trolls disséminés à travers le monde. De leur côté, ni deadline, ni fiche de paie, pas toujours d'objectif précis hormis celui de rire un bon coup ou au contraire, un objectif extrêmement précis (déstabilisation, discrédit, intervention politique) et des moyens proportionnels aux enjeux, c'est-à-dire colossaux.

Voilà où se trouve aujourd'hui la lutte contre les *fake news* et les *deepfakes*. Un jeu fini est engagé contre des participants à un jeu infini. Peu importe le volume des moyens engagés dans la vérification des infox, si l'on en croit la théorie de James Carse, médias et journalistes ne seront jamais en mesure d'emporter cette bataille tant qu'ils continueront de jouer à ce jeu fini.

3.3.3 L'ébauche d'une solution, changer de partie

Pour avoir une chance de l'emporter, il s'agit donc de se placer sur le même plan que l'adversaire, c'est-à-dire dans un jeu où le seul objectif est de perpétrer le jeu lui-même. Concrètement, il s'agit donc d'investir ses efforts sur une mission de long terme plutôt que de se battre pied à pied, comme de bons petits soldats, contre chaque danger. Le choix de la mission revient bien entendu à chaque rédaction, mais la tendance qui se dessine semble pointer vers un rapprochement des journalistes de leur public. Le *Guardian*, *Mediapart*, *De Correspondent*, le *New York Times* ou le *Washington Post*, parmi

les 158 médias recensés par le site MembershipPuzzle.org⁶⁰, ont choisi de développer une stratégie résolument tournée vers une mission : retisser le lien ténu entre un titre et son public. Bien entendu, il ne s'agit pas ici de nier l'importance du travail fourni chaque jour par les services de fact-checking, ni même de penser qu'il s'agit du jour au lendemain de supprimer l'ensemble de ces services pour retrouver la confiance du public, ils y contribuent largement. Non, il s'agit pour les rédactions de comprendre que le fait de concentrer la question du fact-checking sur une dizaine de journalistes ne les dispense pas des efforts nécessaires pour voir plus loin que le bilan financier trimestriel.

La réponse au phénomène des *fake news* et des médias synthétique ne peut pas être exclusivement technologique sans que cela entraîne une certaine forme de négation de ce qu'est l'information, une matière organique, mouvante, principalement basée sur la confiance entre les individus qui produisent cette information et ceux qui la reçoivent. Les algorithmes, s'ils apportent une assistance non négligeable dans le traitement des données et leur restitution, ne sont pas l'alpha et l'oméga des relations humaines. Dans leur recherche perpétuelle d'innovations capables de les sortir de l'ornière financière dans laquelle ils se trouvent, les médias du monde entier se sont précipités et continuent encore de le faire, vers la technologie. Comme au XIX^e siècle, une forme de scientisme s'est emparée de certaines salles de rédaction, qui ne jurent plus que par le développement produit, la data, son analyse et son exploitation.

Cet effort dépend grandement de la santé économique des titres qui doivent, pour se mettre à jour, embaucher un nombre conséquent de « fonctions support », comme on les appelle pudiquement, à savoir développeurs, chefs de projets, analystes data, etc. Un

⁶⁰ The Membership Puzzle est une plateforme créée par Jay Rosen, professeur de journalisme à la NYU et le journal néerlandais créé en 2013 à la suite d'une campagne de crowdfunding de plus d'un million d'euros. La plateforme se focalise sur l'innovation des modèles d'affaires dans la presse.

investissement que peu de titres sont à même de réaliser et qui sera d'autant plus difficile à entreprendre si le compas reste pointé sur une route où l'accélération est exponentielle.

Pour jouer un jeu infini, il s'agit donc de parier sur l'humain et de renvoyer les journalistes au contact de leur public, il s'agit de les inciter – voire de les forcer – à reprendre pied dans une réalité qui leur échappe. Tout cela est une question de culture de la proximité, de l'écoute, de l'ouverture et de l'échange. Les rédactions doivent reprendre leur place dans le tissu social et apporter un service unique. C'est à ce prix qu'elles retrouveront l'écoute de leur public. Il s'agit de quitter cette course à l'innovation pour embrasser une véritable dynamique de progrès.

Conclusion

L'intention initiale de ce travail visait à analyser l'effet des deepfakes sur les rédactions et d'observer quels étaient les tactiques mises en place en amont de leur arrivée massive dans la sphère publique pour les contrer. L'étude des médias synthétiques nous a progressivement poussé à reconsidérer notre angle et à embrasser une analyse plus large du phénomène. En convoquant l'histoire de la peinture Néerlandaise du XVI^{ème} siècle, puis celle du mouvement Réaliste, puis celle de la photographie et du cinéma nous avons confirmé que l'image, en tant que représentation du réel, avait profondément marqué notre rapport au monde et bouleversé le consensus collectif créé au fil des siècles sur ce qui est, qui n'est pas et qui est peut-être. Il en résulte que ce n'est pas l'exotique technologie des deepfakes et le challenge qu'elle pose pour le petit monde des journalistes qui représentent une menace, mais plutôt l'incidieuse détérioration de notre rapport au réel. La technologie des GANs, neutre par essence, est aujourd'hui employée à confectionner de petites vidéos plus ou moins amusantes prêtes à être diffusées sur le web. Demain, cette même technologie, servira sans aucun doute à changer la réalité autour de nous. Elle sera l'instrument de manipulations politiques ou idéologiques comme l'ont été avant la peinture, la photographie et le cinéma et comme le sera – si son usage se popularise – la réalité virtuelle ou d'autres formes de représentations visuelle ou sonore. La technologie évolue mais les intentions humaines qui motive les usages restent les mêmes. Ce n'est donc pas dans la technologie qu'il faut chercher une solution, mais dans notre approche et notre rapport à celle-ci. Sans éveil des consciences, sans éducation à l'image, à l'information, aux nouveaux codes de communication, l'impact de ces médias synthétiques ne sera que plus grand et devastateur.

Bibliographie

Livres et publications universitaires :

- Eye of the Beholder: Johannes Vermeer, Antoni van Leeuwenhoek, and the Reinvention of Seeing par Laura Snyder, 2015
- Secret Knowledge: Rediscovering the Lost Techniques of the Old Masters par David Hockney 2001
- On Alexander Gardner's Photographic Sketch Book of the Civil War par Anthony W. Lee, Elizabeth Young, 2008
- Image dialectique, anachronisme et vérité, par Fabienne Ankoua
- L'image, entre simulacre et vérité, Vassiliki-Piyi Christopoulou, L'Esprit du temps, « Imaginaire & Inconscient » 2014/2 n° 34, pages 35 à 46
- Les Paradoxes de L'image, Monique Sicard, Hermès 21, 1997
- The power of Images in the Age of Mussolini par Valentina Follo, 2013
- Le monde de l'image, cahier hors série de la Revue n°43 Sciences Humaines, 2003-2004
- Les Mécanismes de la transformation du réel dans la propagande photographique nazie : analyse d'un album illustré du congrès du parti nazi à nuremberg en 1933 par Samuel Gaudreau-Lalande, 2011
- La Faiblesse du Vrai, Myriam Revault d'Allonnes, 2019
- Journalisme et réalité : l'argument constructiviste par Gilles Gauthier dans Communication et Langues n°139 pp. 17-25, 2004
- Data and society, Media Manipulation and disinformation online par Alice Marwick et Rebecca Lewis, 2018
- The Oxygen of amplification par Whitney Phillips, 2018

- Social Media & democracy par Kris-Stella Trump, 2018
- Information disorder: Toward an interdisciplinary framework for research and policy making par Claire Wardle et Hossein Derakhshan
- The Science of Fake News dans Insights, Science par David Lazer, 2018
- Mal-uses of AI-generated synthetic media and deepfakes: pragmatic solutions discovery convening, par Firstdraft, 2018
- Fake news: l'arbre qui cache la forêt par David Chavalarias, AOC 2018
- Media Litteracy, Democracy and the challenge of Fake News par Lance E. Mason, Daniel G. Krutka et Jeremy Stoddard, Journal of Media Litteracy, 2018
- Deep fakes: A looming challenge for privacy, democracy and national security par Bobby Chesney et Danielle Citron, 2018
- The spread of true and false news online par Soroush Vosoughi, Deb Roy et Sinan Aral, Science, 2018.
- Imagining Deceptive Deepfakes, an ethnographic exploration of fake videos par Tormod Dag Fikse, 2018
- MesoNet: a compact facial video forgery detection network par Darius Afchar, Vincent Nozick, Junichi Yamagishi et Isao Echizen, 2017
- Total moving face reconstruction, par Supasorn Suwajanakorn, Ira Kemelmacher-Shlizerman et Steven Seitz, 2016
- What Makes Tom Hanks look like Tom Hanks, par Supasorn Suwajanakorn, Ira Kemelmacher-Shlizerman et Steven Seitz, 2017
- Synthesizing Obama: learning lip sync from audio par Supasorn Suwajanakorn, Ira Kemelmacher-Shlizerman et Steven Seitz, 2016

- High-resolution image synthesis and semantic manipulation with conditional GAN's par Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz et Bryan Catanzaro, NVIDIA corporation/UC berkeley 2018.
- Progressive growing of GANs for improved quality, stability and variation par Terro Karras, Timo Aila, Samuli Laine et Jaako Lehtinen, ICLR 2018
- A multidimensional approach to disinformation, par Commission Européenne, 2018
- Disinformation and “fake news”: `Final report par The British House of Common, 2019
- State-Sponsored Trolling par Carly Nyst et Nick Monaco, the Institute for the Future, 2018
- Rapport sur la proposition de loi relative à la lutte contre la manipulation de l'information par Mr le député Bruno Studer, 2018
- The Malicious use of Artificial intelligence: forecasting, prevention and mitigation, 2018
- The Future of political warfare: Russia, the west, and the coming age of global digital competition par Alina Polyakova et Spencer Boyer, 2018
- Worldwide threat assesment of the US intelligence community par Daniel Coats, Director of National Intelligence, 2019.

Film:

- Tim's Vermeer, produit par Penn Jillette, 2013

Sources Internet :

- L'(in)visible de Vermeer en six tableaux par Alison Sinard, 2017 (France Culture)
- Bibliographie des sources internet disponible ici

(<https://www.notion.so/journalismdesign/6464bff637ca473eaaf498fd14977dc4?v=1de8c626435140c88a5d010e144c48f1>) en accès libre