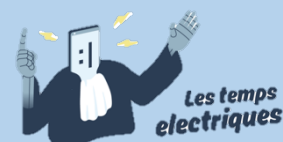


La critique de la technique : clé du développement de l'intelligence artificielle ?

Yannick Meneceur



une étude
publiée sur



« Il faut fonder le concept de progrès sur l'idée de catastrophe. Que les choses continuent comme avant, voilà la catastrophe. »

Walter Benjamin, Essai sur Charles Baudelaire

publié sur



Cette étude a été publiée sur le blog Les Temps Électriques en septembre 2020 – <https://lestempselectriques.net>, sous licence Creative Commons (BY-NC-ND)

Les vues et opinions exprimées dans ce document n'engagent que son seul auteur.

L'auteur tient à remercier Nicolas Régis, magistrat, pour son attentive relecture.

Sommaire

Introduction	7
Partie 1 Le mythe de la neutralité de la technique	11
#1.1. L'effet structurant des technologies numériques	11
#1.2. Une appropriation de la science et du progrès par le marché.....	12
#1.3. La technique comme seul horizon de progrès ?.....	13
#1.4. Les faiblesses structurelles de « l'intelligence artificielle » minimisées par les débats sur les usages.....	14
Partie 2 Les conséquences de décennies de gouvernance de la critique de la technique .	17
#2.1. La décrédibilisation des contre-discours sur la technique	17
#2.2. Un consensus sur les bienfaits de la technique limitant la construction de trajectoires originales.....	18
#2.3. La prolifération des discours éthiques en matière « d'intelligence artificielle »	19
#2.4. La mise à distance de la question de l'opportunité du recours à la technique	20
Partie 3 La difficile émergence d'un discours critique de l'informatique et de « l'intelligence artificielle »	23
#3.1. La critique précoce de l'informatique	23
#3.2. L'avènement d'internet et du web, moteurs d'une nouvelle économie... sans nouvelle réglementation	25
#3.3. La résurgence d'une critique profonde de l'informatique et de l'internet divertie par la résurrection du terme « intelligence artificielle »	26
Partie 4 Revitaliser le contenu d'une réglementation sur « l'intelligence artificielle »	29
#4.1. Des initiatives de réglementation prenant acte d'un lien entre innovation technologique et croissance économique	29
#4.2. Recentrer les travaux d'une réglementation de « l'intelligence artificielle » en connaissance du réel état de l'art	31
#4.3. Les défis et les opportunités d'une réglementation effective de « l'intelligence artificielle »	32
#4.4. Les fondements et les axes principaux d'une réglementation effective de « l'intelligence artificielle »	33

Introduction

Si l'année 2018 a été l'année de l'acmé des principes, chartes et déclarations sur l'intelligence artificielle (« IA^[1] »), l'année 2020 a bien démarré pour être un tournant vers des textes plus contraignants, enfin en capacité de protéger les individus et l'entière société de diverses dérives, déjà concrètes. À peu près toutes les organisations internationales se sont mises en ordre de marche dès 2019^[2], dans le cadre de leurs mandats respectifs, pour apporter leur contribution à ce qui pourrait constituer dans les années à venir une architecture globale de réglementation de « l'IA », appelée à concilier différents impératifs tels que l'innovation, la croissance économique et la protection des droits fondamentaux. Sur un plan local, nombre d'États ont adopté (ou sont sur le point d'adopter) des stratégies sur « l'IA » visant à s'assurer sur cette même scène mondiale leur leadership scientifique, technologique et moral, en y intégrant parfois un volet sur la régulation ou la réglementation^[3]. La Nouvelle-Zélande affirme ainsi être le premier État à avoir adopté des standards, sous forme de charte, en matière de régulation des algorithmes^[4]. Dans ce contexte de véritable course pour établir en premier des règles du jeu, un consensus s'exprime en ce qui concerne les potentiels bienfaits de cette « IA » pour l'humanité et la nécessité de créer de la confiance chez les milliards d'utilisateurs sur la planète pour en tirer les bénéfices attendus.

Toutefois, les lobbyistes de l'industrie numérique s'activent, notamment à Bruxelles, pour tenter de différer ces ambitions au prétexte de la nécessaire relance en pleine crise économique suite à la pandémie^[5]. Dans le même temps, une légère brise d'automne commence à se faire ressentir sur les différents emplois de cette technologie ; face aux promesses déçues, les investisseurs se rendent bien compte que les potentialités de nombre d'applications ont été surévaluées^[6].

Dans ce contexte, il semble tout à fait surprenant que ce début de prise de conscience ne s'accompagne pas de la part des régulateurs d'un sérieux inventaire permettant de distinguer avec précision les usages fondés de ceux relevant de la pure spéculation ni d'un examen encore plus rigoureux des implications sociales et politiques de cette technologie, autrement que sous le prisme de l'innovation et du progrès^[7]. La charte en Nouvelle-Zélande fait bien la promotion de la transparence des décisions algorithmiques, mais sans pour autant remettre en cause le principe même de certaines applications comme l'évaluation du risque de récidive en matière pénale^[8].

Ce constat s'explique si l'on replace cette « IA » dans une dynamique bien plus large et globale de conception du progrès par la technique^[9]. Dans les discours dominants, notamment relayés dans les médias généralistes, il n'est en effet procédé à aucune étude approfondie des capacités réelles de nouvelles technologies complexes afin de dépasser les discours commerciaux, tant la nouveauté est devenue synonyme de progrès dans nos

représentations. Tout ce qui semble techniquement faisable paraît même devenir souhaitable^[10], surtout si c'est rentable. La science, de plus, est trop souvent confondue avec croyance et de nombreux biais conduisent à donner crédit aux arguments d'autorité ou au bon sens plutôt qu'aux preuves^[11]. Le pouvoir de transformation de technologies, toujours plus invasives, n'est traité que sous l'angle d'une sempiternelle balance entre des bénéfices espérés et des risques redoutés, évacuant ainsi toute une série de « questions zéro » : sommes-nous sûrs avant tout que cela fonctionne comme nous l'imaginons ? N'existe-t-il pas d'autres solutions, moins sophistiquées, simples et peu coûteuses, pour délivrer le service recherché ? Est-ce qu'il n'est finalement pas créé des problèmes plus graves que ceux que l'on essaye de résoudre ?

L'on ne devrait donc pas totalement se satisfaire des postulats tenant pour acquis le bénéfice de cette « IA », même si elle est « centrée sur l'humain », et pour développer une « IA » de manière durable, les régulateurs devraient pouvoir prendre un pas de recul afin de conduire, par un débat contradictoire et éclairé, une réflexion sur ses exactes capacités. Il semblerait également opportun qu'ils puissent s'interroger sur les causes profondes de cette situation, qui résulte essentiellement d'une conception mercantile de la science.

Mais il faut dire que le mythe de la neutralité des technologies, et donc de « l'IA », est devenu extrêmement tenace (première partie) et que des décennies de gouvernance de la critique de la technique sont parvenues, assez substantiellement, à en réduire la portée et les effets comme dans bien d'autres domaines (deuxième partie). Ce qui s'est déjà joué avec la critique de l'informatique dans les années 1980 et de l'internet dans les années 2000 est train de se rejouer avec « l'IA », avec la même tonalité d'arguments, et il est à craindre que les atteintes aux droits de l'homme, à la démocratie et à l'État de droit finissent par se normaliser comme tribu d'un prétendu progrès (troisième partie). Les efforts de réglementation de « l'IA » risquent fort de se trouver fortement dévitalisés s'il est tenu pour acquis que seuls des dérives d'usages seront à encadrer, sans traiter les questions préalables de la qualité de la science et de la forme de société en train de se constituer avec l'emploi massif des outils numériques (quatrième partie).

Notes

[1] L'acronyme d'intelligence artificielle sera présenté entre guillemets par commodité éditoriale. L'ensemble des technologies recouvertes par ce terme ne constituent naturellement pas une personnalité autonome et, afin de se garder de tout anthropomorphisme, il a été choisi de résumer les termes plus appropriés « d'outils d'intelligence artificielle » ou « d'applications d'intelligence artificielle » par le seul terme « d'IA » entre guillemets.

[2] [Groupe de travail de haut niveau sur la coopération numérique, groupe international d'experts sur l'intelligence artificielle de l'UNESCO, réseau d'expert de l'OCDE sur l'intelligence artificielle, groupe d'experts de haut niveau sur l'intelligence artificielle de la Commission européenne, Comité ad hoc sur l'intelligence artificielle du Conseil de l'Europe](#)

[3] Pour une revue des initiatives sur l'IA, incluant ces stratégies, voir le site internet du Conseil de l'Europe dédié à l'intelligence artificielle : <https://www.coe.int/en/web/artificial-intelligence/national-initiatives> – Consulté le 14 août 2020

^[4] C. Graham-McLay, New Zealand claims world first in setting standards for government use of algorithms, The Guardian, 27 juillet 2020

^[5] J. Delcker, AI: Decoded: Cold winds are blowing around regulation, Politico, 5 juin 2020, accessible sur : <https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-cold-winds-are-blowing-around-regulation-the-ethics-of-contact-tracing-doubts-over-ai-to-treat-covid-19/> – Consulté le 14 août 2020

^[6] Voir le dossier Artificial intelligence and its limits dans la revue The Economist du 11 juin 2020, notamment An understanding of AI's limitations is starting to sink in, The Economist, 11 juin 2020 et C. Mims, AI Isn't Magical and Won't Help You Reopen Your Business, The Washington Post, 30 mai 2020

^[7] Voir les préambules et les introductions de documents produits par exemple par la Commission européenne, l'OCDE, l'UNESCO ou le Conseil de l'Europe, développés en #4.1. L'évolution de la notion de progrès et de son remplacement par le concept d'innovation ne sera pas discuté ici, bien qu'elle caractérise également la transition, à notre époque, d'un progrès social vers un progrès technique.

^[8] Véritable résurgence d'une conception déterministe de la matière pénale, les constats convergent pour en révéler les biais et les dangers. Voir Y. Meneceur, L'intelligence artificielle en procès, Bruylant, 2020, p. 99 et s. ou encore récemment : J. Burgess, Police built an AI to predict violent crime. It was seriously flawed, Wired, 6 août 2020

^[9] Le concept de « technique » renvoie tout d'abord à un ensemble large de moyens mis en œuvre par les humains pour satisfaire leurs besoins. La sophistication croissante de ces moyens et leur omniprésence dans notre vie quotidienne conduit toutefois à devoir considérer cette technique non plus comme un simple outillage, mais comme un système bien plus complexe, concrétisant également les rapports de force, sociaux et politiques, de notre société. C'est ainsi que sera entendu ce terme dans les présents développements.

^[10] D. Gabor, prix Nobel de physique en 1971, avait déclaré « ce qui peut être fait doit l'être, inéluctablement », caractérisant que tout ce qui était techniquement faisable par l'humain serait réalisé, même contre la morale ou l'éthique.

^[11] E. Klein, Le goût du vrai, Coll. Tracts, Gallimard, 2020, pp.4-5

Partie 1 | Le mythe de la neutralité de la technique

Il est bien connu que le lien entre les sociétés humaines et le système technique composé par l'ensemble de leurs artefacts est extrêmement étroit et chaque (r)évolution majeure a contribué à modeler substantiellement notre environnement, en se déployant parfois sur plusieurs siècles^[1]. Ainsi la portée de l'invention de l'imprimerie a dépassé la seule mécanisation de la reproduction d'ouvrages : la Réforme de l'Église, le siècle des Lumières, l'accès au savoir de manière générale ont été autant d'événements liés à cette invention. L'avènement des processus industriels au XIXe siècle a profondément recomposé les rapports entre individus ainsi que nos lieux de vie et nos modes de gouvernance^[2]. Aujourd'hui nous en serions à notre 4^e révolution industrielle avec la rencontre entre « le monde du physique, du numérique, du biologique et de l'innovation^[3] ». Cette révolution nous doterait de moyens nouveaux pour tenter de franchir de nouvelles frontières, même celle, absolue et imposée par la Nature, qui fait tendre tout chose de l'ordre au désordre : l'entropie^[4].

#1.1. L'effet structurant des technologies numériques

Ce lien entre les humains et leurs tout nouveaux artefacts rend plus que jamais nécessaire l'effort de décryptage de notre environnement sous un angle socio-technique, afin d'en saisir sa nouvelle composition et d'en découvrir ses modes de gouvernance et sa gouvernementalité^[5]. La transformation que nous sommes actuellement en train de vivre avec la traduction en données des moindres recoins de nos vies, aux fins de leur traitement algorithmique, nous dirige vers un tout autre modèle de société qui porte peut-être en lui quelques ferments d'amélioration de la condition humaine, mais aussi son lot de désenchantement, d'emprise, voire de totalitarisme. Et ceci non pas du seul fait de la manière dont nous utiliserions ces outils, mais du fait de la structure tissée par la généralisation de mécanismes informatiques et statistiques censés être en capacité de mieux apprécier que nous, et en toutes circonstances, un nombre toujours plus important de situations, des plus anecdotiques (le choix d'un restaurant) aux plus délicates (l'évaluation des chances de réussite d'un procès). Il devient pourtant évident que nous sommes en train de constituer une terrible intrication de cages logiques, dans lesquelles nos doubles statistiques s'engoncent de plus en plus même si l'on se pense plus que jamais libres de nos choix. L'exercice du pouvoir sur les individus, cette biopolitique théorisée par Michel Foucault, se complète donc d'un mécanisme original, de plus en plus autonome, dont le fonctionnement n'a rien de démocratique et qui congédie même la chose politique^[6]. Nous ne devrions donc pas nous laisser divertir par les pétitions de principe et les diverses assertions sur une prétendue neutralité de technologies, comme « l'IA », et où seuls les usages seraient susceptibles de générer des conséquences désastreuses.

Il pourrait être évoqué à titre de comparaison un autre domaine, comme le nucléaire. Si l'on ne disserte que sur les usages civils ou militaires, l'on manque ainsi d'appréhender une question préalable : souhaite-t-on mettre en œuvre une technologie dont on ne sait aujourd'hui traiter des déchets autrement que par le confinement ou qui nécessite une très haute technicité pour être maintenue sans créer de désastres ? Ce type d'arbitrage ne peut pas résulter de ressentis, d'impressions ou de simples opinions émanant d'une industrie ayant un intérêt économique au développement de cette technologie. C'est tout aussi vrai pour « l'IA », où la plupart des textes préparant une régulation tiennent pour acquis les bénéfices de l'emploi de cette technologie dans de très nombreux secteurs, en ne s'appuyant sur aucun autre fondement scientifique que les très nombreux arguments d'autorité délivrés par des « experts », parfois empêtrés dans de sévères conflits d'intérêts^[7]. La charge de cette preuve en arrive même à être étrangement renversée et ce serait aujourd'hui aux détracteurs de démontrer d'abord le juste fondement de leurs doutes, et non aux producteurs de ces services de prouver d'abord qu'ils délivrent le service attendu, sur la base de résultats rigoureux.

Pour prendre un autre exemple, l'usage de « l'IA » en appui au processus décisionnel dans les tribunaux pose la question du sens produit par le traitement statistique massif de la jurisprudence et la place de cette information – extrêmement incertaine – dans l'échelle normative^[8]. Mais les débats se perdent entre les promoteurs d'une justice plus prévisible et les détracteurs redoutant une robotisation des tâches, aboutissant aujourd'hui à une sorte de voie moyenne laissant intacts les problèmes de fond. Le développement en France d'un outil censé évaluer le montant de l'indemnisation pour la réparation du préjudice corporel atteste sans ambiguïté de cette incapacité de se doter d'une expertise de haut niveau pour évaluer, avec acuité, la viabilité d'un système avant même de se lancer dans sa conception^[9]. Une importante polémique est également née de l'usage d'un algorithme « prédictif » pour attribuer des notes à des élèves du baccalauréat international, au vu de la difficulté de les évaluer en pleine crise sanitaire. Alors que les administrateurs de ce diplôme espéraient adopter ainsi « l'approche la plus juste pour l'ensemble des étudiants », l'algorithme a, en réalité, produit des résultats contestés et pour lesquels aucune explication n'a été possible^[10]. La même situation s'est reproduite au Royaume-Uni, entraînant le retrait pur et simple de ces évaluations^[11]. Là encore, un recours hâtif à un dispositif d'apprentissage statistique a produit sans grande surprise des discriminations inacceptables alors que d'autres solutions, pas nécessairement numériques, étaient tout à fait envisageables pour produire une évaluation.

#1.2. Une appropriation de la science et du progrès par le marché

Il est vrai que les débats sur le numérique et « l'IA » semblent le plus souvent opposer les Modernes de notre époque à leurs Anciens, à savoir les techno-prophètes (accompagnés de

leurs techno-disciples) face aux techno-sceptiques (animés par une grande variété de motivations). Les premiers passeraient, aux yeux de leurs détracteurs, pour des innovateurs inconscients et les seconds pour des rétrogrades.

Cette dichotomie a le mérite de la simplicité et se retrouve employée dans de nombreux discours de ces dernières années, le plus souvent pour donner corps à une parole se déclarant sincère et équilibrée qui ne souhaite pas s'inscrire dans la tradition des poncifs habituels de l'un ou de l'autre de ces camps^[12]. Mais cette distinction présente un inconvénient, puisqu'elle ignore une autre ligne de fracture bien plus profonde et signifiante, celle séparant les tenants d'une totale neutralité des technologies et ceux essayant, au contraire, d'en analyser l'impact social profond. Elle ignore également le réel face-à-face s'opérant entre une science à la recherche d'objectivité et celle rongée par des conflits d'intérêts profonds, animée par l'influence financière et commerciale de puissantes entreprises cherchant, moins que la croissance d'un bien commun celle de leurs bénéficiaires et de leur influence sur le marché.

Il en résulte une appropriation de la science, et de l'idée même du progrès, par un projet de société qui ambitionne de transformer toute chose en marché, et où l'on ne sait plus réellement si telle ou telle invention nous est présentée avec engouement pour ses qualités intrinsèques ou pour que nous la consommons. La crise sanitaire a particulièrement révélé cette « épidémie de mauvaise science^[13] », et il semble que notre époque ait perdu le temps et le sens de la controverse scientifique. Controverse tout à fait salutaire, où il devrait être plus question d'échanger des preuves solides entre pairs^[14] que des croyances, des opinions préfabriquées ou des *punchlines* (phrases choc). La figure de l'ingénieur-entrepreneur en *hoodies* illustre bien la totale confusion entretenue aujourd'hui entre la démarche scientifique, qui s'était conçue jusqu'à un passé pas si lointain comme un projet d'amélioration de la société et des communs, et la démarche mercantile, dont la finalité est d'abord la recherche de profit individuel et la réalisation de soi.

#1.3. La technique comme seul horizon de progrès ?

Dans le même temps, cette prolifération d'artefacts a progressivement constitué un tout nouveau système, allant même jusqu'à l'interconnexion d'objets (*Internet of Things – IOT*) et la captation quasi permanente de flux d'informations et de constantes émanant d'individus. L'impact systémique et cumulé de tous ces développements sur notre société est donc à considérer avec sérieux, « l'IA » ajoutant une nouvelle brique décisionnelle qui embarque désormais dans cet environnement notre capacité d'action (*agency*).

Nous devrions pourtant être attentifs à la manière dont ces nouvelles technologies structurent déjà notre monde, même si nous nous pensons comme tout à fait conscients de leur pouvoir de transformation et libres de nos choix. Pour ne prendre que quelques exemples, l'on ne peut que constater que la forme de nos villes et de nos campagnes s'est

totalément adaptée à la généralisation de l'automobile ; l'avènement des mass media, puis aujourd'hui des réseaux sociaux a profondément transformé l'accès aux savoirs et l'exercice démocratique ; les échanges commerciaux se sont mondialisés avec l'accroissement des transports aériens et maritimes, créant des spécialisations et des interdépendances historiquement originales entre les différentes régions du globe. Et là encore, il peut être relevé que le trait commun de toutes ces évolutions est le succès du modèle économique de libre-échange, qui s'est consolidé à une échelle mondiale et s'est approprié au passage un mythe : la technique serait devenue notre seul destin et notre seul horizon de progrès et de développement.

Mais il semble que cette utopie scientiste suffise de moins en moins pour imposer dans le débat de nouvelles technologies, comme cette fameuse « IA » ou d'autres telles que les OGM (organismes génétiquement modifiés). Invariablement, les lobbyistes de ces industries et les partisans d'une croissance économique par l'innovation, dont certains se retrouvent aujourd'hui à de hauts postes de responsabilité politique, mobilisent une série de raisonnements instrumentalisant la crainte et entretenant un sentiment diffus de « retard » : si nous ne le faisons pas nous, d'autres le feront ; les cerveaux nationaux vont, à coup sûr, fuir vers des lieux plus cléments ; en réglementant, l'on offrirait un avantage compétitif à ceux n'ayant pas réglementé^[15] ; en s'arrêtant en si bon chemin, nous n'aurions pas accès à la nouvelle génération de technologie qui, bien entendu, réglerait tous les problèmes de la précédente. En d'autres termes, un discours nous enjoignant à nous adapter, de crainte de devenir obsolète^[16].

#1.4. Les faiblesses structurelles de « l'intelligence artificielle » minimisées par les débats sur les usages

Ce constat brut du fonctionnement du capitalisme industriel à notre ère numérique et des alliances nouées avec la science et la politique explique l'intérêt de promouvoir le concept d'une totale neutralité des technologies et les difficultés pour tenir un débat réellement approfondi sur « l'IA » et ses réelles capacités. Ajoutons également l'intérêt de décideurs publics ou d'universitaires soucieux de contribuer dans ces débats et qui, souvent par manque de culture technique, se recentrent sur des thématiques relevant de leur champ d'expertise^[17], et l'on obtient le consensus actuel centrant la réglementation de « l'IA » sur la seule question des usages.

Il en résulte concrètement un renvoi de tout examen sérieux des capacités de cette technologie à des controverses d'experts, ce qui en rend le contenu totalement inaccessible aux profanes. Même s'il n'était pas précisément question « d'IA », les débats sur les applications de suivi de proximité des contacts^[18] (*proximity tracing*) ont parfaitement illustré ce mécanisme, où les discours publics ont tenu pour acquise la capacité de ces applications à délivrer à minima le service espéré et ont traité l'ensemble des critiques de fond sur

l'inadaptation du Bluetooth comme un détail, alors même qu'il s'agissait là du cœur du problème^[19]. Les débats se sont étrangement centrés sur des questions de protection de la vie privée et de surveillance de masse alors qu'une condition nécessaire et préalable à toute discussion n'était pas remplie : le suivi de proximité de contact basé sur le Bluetooth, même adopté en masse, génère trop de faux positifs ou de faux négatifs pour apporter un réel appui aux brigades sanitaires^[20]. C'est un peu comme si un débat avait été provoqué sur les bénéfices et les dangers résultant de l'utilisation d'un marteau pour casser un œuf alors que la question semble pouvoir être assez vite évacuée en constatant que ce n'est pas l'outil adapté à cette tâche. Mais, même acculés face à cette réalité technique, nombre de décideurs publics n'ont pas rendu les armes et ont soutenu que même si une seule vie était sauvée par ces dispositifs numériques, cela mériterait les millions d'euros investis... mais parle-t-on des vies potentiellement mises en danger par un excès de confiance dans ces outils numériques ? En somme, toute une série d'arguments de pure circonstance, loin de toute rationalité et seulement culpabilisant, mobilisés à la seule fin de légitimer des choix pris dans l'urgence du moment. En alimentant les débats sur la seule question des usages de technologies, l'on contribue à faire accepter comme une réalité factuelle ce qui est loin d'en être une : non, « l'IA » n'est pas nécessairement en mesure de révolutionner en profondeur toutes les activités humaines et le traitement par des modèles statistiques de masses de données n'est pas à même de rendre tous les services espérés. Oui, en transformant le monde en données, on l'interprète et on le tord pour le faire rentrer de force dans des mécanismes algorithmiques. Oui, en considérant que les modèles produits par les « IA » sont dignes d'intérêt dans la plupart des circonstances, l'on confond la représentation de la réalité avec une certaine forme de son expression. Oui, « l'IA » et tous les types de traitements algorithmiques traduisent le monde au travers du prisme parfois déformant des mathématiques. Une lecture approfondie de la littérature académique sur le sujet montre bien que cette « IA » fait encore l'objet de bien trop de controverses, d'instabilité et de fragilités pour faire l'objet d'une diffusion de masse^[21]. Une réglementation efficace devrait donc arriver à se départir des discours dominants et ne pas considérer comme un détail ce qui constitue le cœur d'un vaste problème technique^[22].

Notes

[1] Même si certaines espèces animales parviennent bien à utiliser de manière simple des outils, l'espèce humaine s'est caractérisée par son appropriation de techniques et de combinaison de techniques ayant une influence sur sa propre évolution – voir F. Sigaut, *Comment Homo devint Faber. Comment l'outil fit l'homme*, CNRS éditions, 2013 cité par F. Jarrige, *Techno-critiques, Du refus des machines à la contestation des technosciences*, La Découverte/Poche, 2016, p. 24

[2] S'agissant de l'émergence de la statistique à des fins de gouvernance, voir O. Rey, *Quand le monde s'est fait nombre*, coll. Les Essais, Stock, 2016, p.95 et s.

[3] K. Schwab, *La quatrième révolution industrielle*, Dunod, 2017

[4] E. Sadin, *L'intelligence artificielle ou l'enjeu du siècle : anatomie d'un antihumanisme radical*, L'Echappée, 2018, p.15

^[5] A. Rouvroy et T. Berns, Gouvernamentalité algorithmique et perspectives d'émancipation, Réseaux 2013/1, n°177, 2013, pp163-196

^[6] E. Sadin, L'intelligence artificielle ou l'enjeu du siècle : anatomie d'un antihumanisme radical, op.cit., p.79 et s.

^[7] T. Metzinger, Ethics washing made in Europe, Der Tagesspiegel, 8 avril 2018

^[8] Pour illustrer avec quelques exemples ce qu'entendent réaliser les promoteurs d'un traitement mathématique, probabiliste ou statistique de la jurisprudence, voir notamment L. Godefroy, F. Lebaron et J. Lévy-Vehel, Comment le numérique transforme le droit et la justice vers de nouveaux usages et un bouleversement de la prise de décision, Rapport de recherche à la Mission Droit et Justice, Juillet 2019, O-M. Sulea, M. Zampieri, M. Vela, J. van Genabith, Predicting the Law Area and Decisions of French Supreme Court Cases, 2017 ou N. Aletras, D. Tsarapatsanis, D. Preotiuc-Pietro, V. Lamos, Predicting judicial decisions of the European Court of Human Rights: a natural language processing perspective. PeerJ CompSci, 2016

^[9] Y. Meneceur, DataJust face aux limites structurelles de l'intelligence artificielle, Les Temps Electriques, 19 juin 2020

^[10] Hye Jung Han, An algorithm shouldn't decide a student's future, Politico, 13 août 2020

^[11] C. Cohen, Au Royaume-Uni, un algorithme autour des notes d'étudiants crée la polémique, Le Figaro, 17 août 2020

^[12] Voir notamment Y. Meneceur, L'intelligence artificielle en procès – Plaidoyer pour une réglementation internationale et européenne, Bruylant, 2020, p.5

^[13] F. Goubet, Une épidémie de mauvaise science, Le Temps, 24 avril 2020 et H. Morin, « The Lancet » annonce le retrait de son étude sur l'hydroxychloroquine, Le Monde, 4 juin 2020

^[14] Selon le philosophe Karl Popper, les sciences procèderaient de « la coopération amicalement hostile des citoyens de la communauté du savoir » cité dans K. Boucaud-Victoire, Etienne Klein : « La vérité scientifique n'appartient nullement aux scientifiques », Marianne, 28 juillet 2020

^[15] Contre les idées reçues, la Chine est d'ailleurs en train de mettre en place une réglementation claire de « l'IA » : Voir les principes de gouvernance de l'IA publiés en 2019, accessible sur : <https://www.loc.gov/law/foreign-news/article/china-ai-governance-principles-released/> – Consulté le 21 août 2020

^[16] Ba. Stiegler, « Il faut s'adapter » : sur un nouvel impératif politique, Gallimard, coll. « NRF Essais », 2019

^[17] Il ne sera pas évoqué ici certains pseudos-experts intervenant de manière indifférente sur tout sujet à la mode, victimes de l'effet « Dunning-Kruger », biais cognitif où l'ignorance donne une assez grande assurance pour parler de sujets complexes avec très peu de prudence.

^[18] Notons que certaines de ces applications comportaient des algorithmes évaluant les risques de contamination, certains s'appuyant sur des modèles issus d'apprentissage automatique – Voir notamment Y. Bengio, Dépistage pair à pair de la COVID-19 basé sur l'IA, Blog personnel de Yoshua Bengio, 25 mars 2020

^[19] Voir à ce titre la publication particulièrement bien étayée : X. Bonnetain, A. Canteaut, V. Cortier, P. Gaudry, L. Hirschi, S. Kremer, S. Lacour, M. Lequesne, G. Leurent, L. Perrin, A. Schrottenloher, E. Thomé, S. Vaudenay, C. Vuillot, Le traçage anonyme, dangereux oxymore Analyse de risques à destination des non-spécialistes, 21 avril 2020, à mettre en parallèle avec les discours publics, notamment StopCovid: Cédric O refuse de reconnaître tout échec. "trop tôt pour faire le bilan", Challenges, 30 juillet 2020

^[20] Voir par exemple J-M. Manach, Covid-19 : pourquoi le « contact tracking » ne fonctionnera (probablement) pas, NextImpact, 10 avril 2020

^[21] Voir par exemple S. Ben-David, P. Hruběš, S. Moran, A. Shpilka, A. Yehudayoff, Learnability can be undecidable, Nature Machine Intelligence 1, 2019, pp.44-48 ou B. Georges, Le talon d'achille de l'intelligence artificielle, Les Échos, 15 mai 2017

^[22] Voir Y. Meneceur, L'intelligence artificielle en procès – Plaidoyer pour une réglementation internationale et européenne, op.cit., p.43 et s.

Partie 2 | Les conséquences de décennies de gouvernance de la critique de la technique

En première analyse, le consensus apparent sur les bénéfices incontestables de la technique pourrait paraître fragilisé avec la publication d'études étayées et convergentes dans de nombreux domaines, comme en ce qui concerne la menace environnementale^[1]. Le succès en France, et dans d'autres pays européens, de mouvements politiques écologistes à des élections témoigne de cette prise de conscience citoyenne et du relatif affaiblissement du discours imposé par le capitalisme industriel. Il est également intéressant de relever que les plus vives critiques actuelles de la technique ne proviennent d'ailleurs pas nécessairement d'idéologues ou de politiques, mais proviennent aussi des techniciens eux-mêmes, en pleine conscience des enjeux et des limites des divers moyens à notre disposition^[2]. De nombreuses publications parviennent aujourd'hui à dépasser les lieux communs en démontrant que derrière la prétendue neutralité des techniques, il y a surtout des enjeux de pouvoirs et que derrière les prophéties, il y a des artifices marketing maquillant une réalité bien plus modeste^[3]. Au final ce n'est d'ailleurs pas la technique en elle-même qui est remise en cause par certains de ces auteurs, mais l'asservissement résultant d'une certaine forme de transfert du sacré à la technique^[4]. Et c'est exactement cela qui se joue pour le numérique et « l'IA » qui, en derniers artefacts à la mode (et potentiellement rentables), sont instrumentalisés pour nourrir l'idée d'une révolution et d'une civilisation toujours en « progrès » continu par les sciences, alors que l'on se trouve en réalité en pleine confusion entre fins et moyens^[5].

#2.1. La décrédibilisation des contre-discours sur la technique

Durant ces dernières années, les contre-discours critiques des techniques ont eu bien des difficultés à émerger, à participer au débat public et à irriguer une pensée politique cohérente, notamment sous l'influence d'acteurs privés dont la puissance en est venue à être tout à fait comparable à celle des États^[6].

Il ne devient même plus concevable de remettre en cause les prétendues qualités ou l'utilité même d'une technologie potentiellement rentable. Les industries de la chimie, des médicaments ou du tabac se sont particulièrement illustrées en multipliant les controverses factices pour paralyser toute décision politique, aux moyens d'arguments simples et populaires et de l'instrumentalisation de la science^[7]. Les doutes ou les contestations sont souvent raillés, comparés aux doux fantaisistes pensant que la Terre est plate, ou renvoyés à des débats d'experts incompréhensibles pour les médias généralistes et l'opinion publique. Ralentir n'est même plus une option, la stase de la réflexion éclairée étant

systématiquement préférée au flux de l'action : « Move fast and break things », mantra de la Silicon Valley, révèle bien la stratégie de notre époque où la réalisation rapide de profits est préférée à la simple mesure des choses. Le sociologue Antonio Casilli affirmait même en août 2020 sur un réseau social que leur stratégie s'apparenterait plus à « tenter de casser des choses, s'excuser pour avoir tenté puis, quand tout le monde s'est calmé, les casser pour de bon ». Spécifiquement pour le numérique, la résolution de tout problème semble aujourd'hui devoir comporter un volet embarquant ce type de technologie, autant par solutionnisme que recherche de nouveaux marchés dans un contexte d'intrications profondes d'intérêts industriels et politiques^[8].

Pour parachever ce tableau, il faut reconnaître que la qualité et les motivations des discours techno-critiques sont d'une extrême diversité, sans réelle épine dorsale idéologique, et qu'il est aisé, en pure rhétorique, d'en fragiliser la teneur en les amalgamant tous aux déclarations les plus extrêmes. L'on y retrouve en effet pêle-mêle tant des décroissants, dont le seul objectif est de créer du dissensus, que des terroristes anarcho-primitivistes comme « Unabomber^[9] » ou de simples héritiers humanistes d'une pensée cherchant à se situer entre Jacques Ellul et Gilbert Simondon.

#2.2. Un consensus sur les bienfaits de la technique limitant la construction de trajectoires originales

La critique de la technique est donc le plus souvent ignorée, minimisée ou instrumentalisée dans les processus de décision politique et dans la plupart des travaux réglementaires. Remettre en cause le principe même du fonctionnement de certaines technologies, du fait des risques ou de leur manque de maturité, n'est le plus souvent pas entendu, d'autres intérêts, politiques, économiques et de croissance, semblant primer. Même les arguments de fond, les plus solides et les plus motivés, se retrouvent noyés dans des querelles d'experts et peinent à émerger.

Pourtant, c'est bien cette critique de la technique qui permettrait aux décideurs publics de disposer d'une grille de lecture pertinente pour interroger et analyser avec acuité les « innovations » et, ainsi, de pouvoir concevoir des trajectoires originales pour les politiques publiques. C'est aussi cette critique de la technique qui constitue un impératif démocratique et civique pour gouverner le changement vers une société soucieuse de son empreinte sur l'environnement. Mais c'est cette critique de la technique qui a été volontairement canalisée durant les quarante dernières années pour en réduire la portée et les effets, au point d'être aujourd'hui dévitalisée et presque moribonde. De la critique radicale des années 1970, qui concernait le recours à l'énergie nucléaire ou même déjà l'informatique^[10], les différentes formes contestations ont été progressivement dépolitisées au travers de réponses technocratiques. Qu'il s'agisse de divers comités d'experts, d'observatoires ou de stratégies participatives faisant émerger des consensus, tout a été entrepris pour marginaliser les

oppositions les plus virulentes et désamorcer tout contenu de nature à remettre en cause l'adoption des technologies.

Il en résulte un réel appauvrissement de la pensée, où les critiques, même étayées, en sont devenues à être considérées comme de simples « opinions » ou des « points de vue ». De manière tout à fait étonnante, il semble que notre époque accorde ainsi une crédibilité supérieure aux promoteurs d'une « IA » en capacité de traiter de manière tout aussi efficace des pixels que des décisions de justice qu'à ses détracteurs, dont le discours paraît naturellement bien moins séduisant si l'on se place dans une perspective de croissance économique par le numérique. Pourtant à se contenter de n'encourager et de ne nourrir que le discours dominant, l'on risque de réels accidents industriels, de nature à décrédibiliser durablement la technologie concernée ; cette attitude conduit aussi à se priver de nouvelles perspectives et de réelles découvertes. C'est exactement dans ce contexte de rejet de la critique que Yann LeCun a eu les plus grandes difficultés à faire émerger ses travaux sur les réseaux de neurones profonds (*deep neural networks*) dans les années 2000 alors que toute la communauté (ainsi que les investisseurs) se concentrait alors sur une autre forme d'apprentissage automatique (*machine learning*), les machines à vecteur de support (*support vector machines*^[11]). Le chercheur en plaisante maintenant, ayant conscience de son succès : « Maintenant il faut que je fasse plus attention à ne pas dire de bêtises parce que personne n'ose plus me dire que j'ai tort »^[12].

#2.3. La prolifération des discours éthiques en matière « d'intelligence artificielle »

La controverse scientifique est au cœur de la construction du savoir et des idées. En rendant reproductibles des observations, l'on donne évidemment les moyens de la critique et de la contradiction. Mais l'on rend aussi progressivement plus robuste ce qui n'est plus alors de l'ordre de l'intuition ou de l'assertion, mais de l'ordre de la démonstration. Ce travail est incontestablement en cours en ce qui concerne « l'IA », mais deux cercles, au moins, se sont constitués^[13]. Aux côtés de la recherche fondamentale et appliquée, qui explore encore la profondeur des réseaux des neurones et débat des avantages respectifs de l'apprentissage auto-supervisé ou de la formalisation de la causalité, coexiste une autre bulle, bien plus spéculative, débattant des effets de cette technologie et de la nécessité d'adopter des lignes de conduite éthiques centrées sur l'humain pour en prévenir les dérives d'usages.

L'observateur extérieur, peu averti, supposera une certaine porosité entre ces mondes et espérera que les progrès (et les doutes) des uns profitent aux autres. Comme avec la plupart des technologies complexes de notre temps, la réalité nous renvoie non seulement à une assez grande étanchéité entre ces deux espaces, mais aussi à une très forte emprise d'entrepreneurs instrumentalisant l'éthique pour « blanchir » une technologie loin d'être

encore aussi performante et construire un accord généralisé sur les machines, c'est-à-dire en l'excluant du champ politique et de l'espace démocratique.

Alors que les pionniers de « l'IA » comme Marvin Minsky semblaient mépriser l'éthique^[14], un discours a pris corps au milieu des années 2010 dans la communauté « spéculative » de « l'IA » afin d'importer des principes de conduite d'autres secteurs, notamment de la bioéthique, censés guider l'action des acteurs « opératifs ». Le résultat de cette intense production a été assez sévèrement critiqué par une partie de la communauté académique, en raison de son absence de sanctions en cas de manquements et de sa délicate opérationnalisation^[15]. Pour Rodrigo Ochigame, ancien étudiant chercheur au sein du Media Lab du MIT, le discours de « l'IA éthique » aurait même été stratégiquement aligné sur un effort de la Silicon Valley cherchant à éviter les restrictions légalement contraignantes de technologies controversées^[16]. La série télévisée Silicon Valley illustrera aussi dans l'un de ses épisodes l'hypocrisie de nombre de ces démarches, un des protagonistes, repentant de ses excès d'entrepreneur, lançant une nouvelle religion baptisée « Tethics » – combinaison de « Tech » et de « Ethics^[17] ». Qu'il s'agisse donc d'une éthique « pratique », à destination des développeurs, ou d'une éthique « mécanisme de gouvernance », à destination des institutions, il convient d'admettre que nombre des principes promus se révèlent ambigus, voire même contradictoires, qu'ils sont purement déclaratifs et qu'ils tiennent pour la plupart comme acquis les capacités de « l'IA » à fonctionner correctement s'il est fait preuve d'un certain nombre de précautions^[18].

#2.4. La mise à distance de la question de l'opportunité du recours à la technique

L'émergence d'un management par le risque de l'usage des technologies, qui part du postulat que ce ne sont pas les technologies étudiées qui sont susceptibles de présenter des défauts, mais seulement leur emploi, a clairement conduit à une dépolitisation de la question de l'emploi massif de technologies complexes. En se concentrant sur la manière de mettre en œuvre une technologie, au lieu de rechercher – avant toute autre réflexion – si nous devrions le faire, l'on soutient de fait la stratégie d'industries soucieuses de ne pas laisser d'espace à une quelconque critique portant sur l'intérêt et les qualités intrinsèques de leurs nouveaux produits.

Ce détournement des débats publics de l'opportunité vers l'usage, accompagné de pressions pour ne mettre en œuvre que de l'autorégulation, n'est toutefois pas spécifique à « l'IA » et il est bien connu que c'est quasiment l'entière démarche scientifique qui est aujourd'hui contaminée par ces méthodes du capitalisme industriel. À titre d'illustration, il pourrait être notamment cité l'Appel de Heidelberg, rendu public la veille du premier sommet de la Terre à Rio en 1992. Réunissant près de 70 prix Nobel et d'autres scientifiques de renom, cet appel proclame leurs inquiétudes face à l'émergence d'une idéologie

irrationnelle s'opposant au progrès scientifique et au développement économique et social. Peu de critiques se sont élevées sur l'instant face à ce texte dénonçant une « écologie irrationnelle », et seuls quelques intellectuels comme le philosophe Cornelius Castoriadis ont été en mesure de décrypter ce texte, en réalité orchestré par les industriels de l'amiante^[19] et omettant tout à fait volontairement d'évoquer les retombées négatives des faux besoins produits par tel « exploit » scientifique ou technique^[20].

Beaucoup de discours, pas nécessairement conscients ou coordonnées, participent donc aujourd'hui à écarter du champ de l'examen critique la question de l'opportunité pour ne mettre en avant qu'un certain nombre de questions centrées sur l'usage. En ayant réussi à imposer une telle narration, les acteurs industriels sont ainsi parvenus, avec le soutien des pouvoirs publics, à donner l'apparence qu'ils se préoccupent de manière sincère des conséquences sociétales de leurs actes alors qu'ils parviennent surtout à échapper à toute discussion de fond sur les prétendues qualités de certaines de leurs innovations. Il en va de même pour l'industrie numérique, puisque même si de plus en plus de personnes sont tout à fait conscientes que nous sommes passés d'un phénomène de salon à un outil de manipulation de masse^[21], questionner l'opportunité du recours à ces nouvelles technologies pour résoudre un problème reste parfois assimilé à des postures extrêmes, d'ultra-gauche anti-technologique ou de pessimistes. Or, c'est précisément en questionnant leur opportunité que l'on pourrait établir des clés de lecture permettant de dépasser les discours commerciaux et de nous doter de la capacité d'en évaluer le réel bénéfice sociétal, à même de prévenir tout « solutionnisme^[22] ». Questionner l'opportunité, c'est aussi crédibiliser les outils qui auront passé le filtre de cette interrogation et donner l'occasion d'une nouvelle politisation des trajectoires technologiques, consubstantielle à toute prétention démocratique.

Notes

^[1] Précisions en ce qui concerne « l'IA » que l'empreinte carbone résultant du calcul des modèles mathématiques avec de l'apprentissage automatique (machine learning) est considérable. En 2018, OpenAI a publié une étude montrant que les ressources de calcul nécessaires à la formation de grands modèles doublieraient tous les trois à quatre mois. En juin 2019, une autre étude a révélé que le développement de modèles de traitement de langage naturel (natural language processing) à grande échelle pourrait produire une empreinte carbone exponentielle. À ce rythme, « l'IA » pourrait représenter selon certains experts jusqu'à un dixième de la consommation mondiale d'électricité d'ici 2025.

^[2] En ce qui concerne « l'IA », voir par exemple M. David, C. Sauviat, Intelligence artificielle, la nouvelle barbarie, Éd. du Rocher, 2019

^[3] Voir les travaux de l'équipe du philosophe Daniel Leufer par exemple, qui a réuni les résultats de leur travail de recherche sur le site internet aimyths.org

^[4] Prolongeant ainsi le constat posé par J. Ellul, Les nouveaux possédés, Mille et une Nuits, 2003 (1973), p.316

^[5] Voir Y. Meneceur, L'intelligence artificielle en procès – Plaidoyer pour une réglementation internationale et européenne, Bruylant, 2020, p.397 et s.

^[6] [C. Kang, J. Nicas and D. McCabe, Amazon, Apple, Facebook and Google Prepare for Their 'Big Tobacco Moment', The New York Times, 28 juillet 2020](#)

^[7] Voir Y. Meneceur, L'intelligence artificielle en procès – Plaidoyer pour une réglementation internationale et européenne, op.cit., p.182 et s.

^[8] Voir à titre d'exemple [S. Fay, Coût et soupçon de favoritisme : l'appli StopCovid dans le collimateur d'Anticor, France Inter, 11 juin 2020](#)

^[9] Surnom donné par la presse au mathématicien Theodore John Kaczynski, auteur d'une série d'attentats aux colis piégés entre 1976 et 1982.

^[10] La loi informatique et libertés a été adoptée en France en 1978 à la suite d'un scandale majeur portant sur le croisement de fichier, voir #3.1 dans la troisième partie : La difficile émergence d'un discours critique de l'informatique et de "l'intelligence artificielle"

^[11] Pour un récit de ces difficultés, voir D. Cardon, J-P. Cointet, A. Mazières, La revanche des neurones. L'invention des machines inductives et la controverse de l'intelligence artificielle, Réseaux, 2018/5 (n° 211), p. 21

^[12] [C. Moulas, Yann LeCun: « Créer des machines qui ont autant de bon sens qu'un chat de gouttière », French Morning, 3 juin 2019](#)

^[13] Trois vagues éthiques sont distinguées, la première venant des philosophes qui ont posé des principes, la deuxième des techniciens qui ont voulu fixer techniquement les problèmes et la troisième, en cours, qui se manifesterait par des actions concrètes limitant le pouvoir des algorithmes, tels que le jugement d'un tribunal néerlandais relatif à un système de détection de fraude. Voir à ce sujet [C. Kind, The term 'ethical AI' is finally starting to mean something, VentureBeat, 23 août 2020](#)

^[14] Voir R. Ochigame, The invention of 'Ethical AI', The Intercept, 20 décembre 2019 où l'auteur cite un de ses anciens collègues du MIT Media Lab qui relatait le fait que Marvin Minsky avait l'habitude de dire que « un éthicien est quelqu'un qui a un problème avec ce que vous avez à l'esprit ».

^[15] B. Mittelstadt, Principles Alone Cannot Guarantee Ethical AI, Nature Machine Intelligence, novembre 2019

^[16] R. Ochigame, The invention of 'Ethical AI', The Intercept, 20 décembre 2019

^[17] Silicon Valley, saison 6 épisode 5 – Récapitulatif accessible sur : [O. Henderson, Silicon Valley Recap: We Are Experiencing Tethical Difficulties, Vulture, 24 novembre 2019](#)

^[18] De nombreuses méta-analyses ont été rédigées pour analyser la profusion des principes éthiques – Voir par exemple [A. Jobin, M. Ienca, et E. Vayena, The global landscape of AI ethics guidelines, Nature Machine Intelligence 1, septembre 2019, pp.389–399](#)

^[19] S. Foucart, L'appel d'Heidelberg, une initiative fumeuse, Le Monde, 16 juin 2012

^[20] C. Castoriadis, L'écologie contre les marchands. Sauvons les zappeurs abrutis !, Le Nouvel Observateur, 7-13 mai 1992, p.102 cité par F. Jarrige, Techno-critiques – Du refus des machines à la contestation des technosciences, op.cit., p.333

^[21] [F. Joignot, Sherry Turkle, la psy des nouvelles technologies. Hier technophile, aujourd'hui inquiète de l'assèchement des relations humaines, Le Monde, 25 février 2019](#)

^[22] E. Morozov, Pour tout résoudre, cliquez ici, Fyp éditions, 2014, p.18

Partie 3 | La difficile émergence d'un discours critique de l'informatique et de « l'intelligence artificielle »

Apparue dans le sillage de la Seconde Guerre mondiale et de la recherche militaire, l'informatique se présente comme la technologie de rupture de ces dernières décennies. Alors que le nucléaire a créé le désordre et le chaos, l'informatique, elle, se présente en outil d'ordre et de raison. Elle a d'abord accompagné les nouveaux besoins de calculs d'États cherchant à automatiser certaines tâches de calcul ou de stockage, puis a investi dans les années 1970 les salons des particuliers, dans un contexte de reconfiguration économique, politique et culturelle majeur. La miniaturisation toujours plus accrue des composants a conduit à la généralisation d'usages que nous connaissons aujourd'hui avec les téléphones intelligents (*smartphones*) et à la recomposition en profondeur des moyens de communication, des offres de service et de l'accès à l'information. Depuis début 2010, « l'IA » est présentée comme la dernière évolution majeure de ce courant technique, avec son lot de promesses, de disruptions et de perspectives d'un avenir meilleur. Mais les critiques de cette « IA » n'ont pas manqué et elles nous avertissent de dangers les plus divers, des plus fantaisistes aux plus substantiels. D'un scénario d'extermination de l'homme par les machines^[1] à un examen minutieux de l'impact concret de l'algorithmisation de notre monde sur l'humain^[2], aucune doctrine majoritaire ne semble se dégager de l'abondante littérature produite à ce sujet. Une constante, en revanche, paraît se dessiner : c'est le rejet de cette critique par les pouvoirs publics, soit en l'ignorant ou en la minimisant, soit en l'institutionnalisant. Le développement, dans certaines des organisations internationales, de textes pour créer une « IA digne de confiance », « éthique » ou « centrée sur l'humain » est certainement à lire en ce sens.

#3.1. La critique précoce de l'informatique

Reprenons brièvement le cours de l'histoire de l'informatique et la manière dont sa critique s'est opérée. Dès les premiers succès des ordinateurs, leur dualité d'usage a été assez rapidement perçue : outil des utopies libertariennes en Californie, offrant la possibilité d'un accès illimité à l'information et au monde pour s'émanciper d'un État oppressant, et outil de contrôle des populations, fichant, numérotant, surveillant de manière continue^[3]. Face au déferlement des micro-ordinateurs à la fin des années 1970 – début des années 1980, les analyses d'impact se concentreront sur trois axes principaux : le travail, notamment avec les risques de suppression d'emplois liés au traitement automatisé de données, la menace sur la qualité du lien social et la société de surveillance. En ce qui concerne ce dernier aspect, c'est exactement dans ce contexte que sont apparues en Europe les premières lois, sur la protection des données à caractère personnel, en conscience des souvenirs encore vifs du

fichage massif des populations durant la Seconde Guerre mondiale. Mais cette lame de fond technologique sur la société n'est pas seulement marquée par la prolifération d'objets de plus en plus miniaturisés et prétendument « intelligents », elle est aussi remarquable par la production d'un discours, devenu permanent, désignant l'informatique et ses avatars, tel que « l'IA », comme le principal facteur de progrès de notre temps.

Au Canada, Marshall McLuhan prophétisera son célèbre concept de « village global »^[4] dès les années 1960, pressentant que l'interdépendance résultant de ces outils électroniques allait rétrécir le monde. Jacques Ellul, pourtant critique sans concession de la technique, s'enthousiasmera même en 1982 en estimant que la micro-informatique offre une chance historique de réorienter le système technique dans un sens émancipateur^[5]. La même année, le magazine Times désignera l'ordinateur comme homme de l'année. Cependant, la mise en récit populaire s'écrira principalement depuis l'ouest des États-Unis, pétri de la culture de conquête d'individus croyant en leur destinée manifeste, avec la ferme intention de recommencer le monde avec de nouvelles valeurs. Des films comme *War Games*, *Tron*, ou des séries comme « Les petits génies » (*Whiz kids*), mettront en scène des adolescents contrôlant ces mystérieuses machines en frappant des lignes de codes cryptiques sur des écrans monochromes, en démontrant que cette technologie pouvait être tant l'instrument d'une nouvelle forme de totalitarisme qu'une forme d'instrument libérateur pour qui savait s'en servir.

Mais qu'il s'agisse des discours commerciaux des constructeurs ou des contes véhiculés par le soft power d'Hollywood, la méfiance dans la population restera assez vive face à ces instruments censés – déjà – apporter une solution dans un monde en crise après les chocs pétroliers. Dans les institutions européennes, l'idée de la vieille Europe frileuse devant l'innovation à l'heure de la mondialisation émergera déjà à cette époque, tout comme la présomption qu'en traitant la question sous l'angle des usages, l'on parviendrait à dépasser l'opposition stérile entre promoteurs de l'informatique et leurs critiques^[6]. Dans le même temps, les sciences sociales se concentreront progressivement sur l'examen approfondi du comportement des individus face à l'informatique, moins que sur son impact sur les structures. Cette « école des usages », pourtant consciente des conséquences de la généralisation de l'informatique, s'alignera progressivement sur le consensus selon lequel ce n'est pas la technologie en soi qui serait à mettre en cause, mais la manière dont les pouvoirs économiques et politiques la pervertissent. Voilà qui explique, au moins partiellement, les difficultés à construire un discours alternatif déconnectant l'évolution technique du progrès. Une victoire décisive pour une industrie parvenue à rendre incontournable sa production et qui, après avoir installé avec succès ses ordinateurs dans les bureaux de travail et les salons, allait parvenir à les interconnecter mondialement dans les années 1990 pour s'ancrer définitivement dans les moindres recoins de nos vies.

#3.2. L'avènement d'internet et du web, moteurs d'une nouvelle économie... sans nouvelle réglementation

La « bulle » née de l'émergence d'internet et de l'interconnexion massive de ces ordinateurs personnels durant les années 1990 a fait naître son lot d'espoirs et d'utopies. Aux critiques de l'informatique a ainsi succédé la perspective d'un accès généralisé au savoir et, surtout, une toute nouvelle économie. De quoi durablement tarir les méfiances. La puissance du simple lien « hypertexte », déjà imaginé par Douglas Engelbart dans les années 1960 et concrétisé par Tim Berners-Lee dans les années 1990 et qui parvient à créer des liens entre les savoirs, semble apporter une cinglante réponse aux inquiétudes imaginant une informatisation de la société digne de 1984^[7]. Les bénéfices du « web » et de son support technique, l'internet, paraissent tellement évidents, que bien peu de critiques de fond émergeront et les garages en Californie commenceront à nouveau à se peupler de jeunes entrepreneurs prêts à saisir leur chance. Techniquement, la logique du portail s'impose, comme celui créé par Yahoo! en 1995, qui devient la première grande entreprise du numérique. Le principe est alors d'intégrer sur une même page le maximum de services afin de retenir l'internaute le plus longtemps possible : des infos, de la météo, des annonces, un moteur de recherche, un service messagerie électronique... et de la publicité. Dans le même temps, les investissements dans l'informatique augmentent de manière substantielle, passant de 50,1 milliards de dollars par an dans le monde en 1980 à près de 154,6 milliards de dollars en 1990 ; au plus fort de la bulle de la « net » économie, ils atteignent même un pic de 412,8 milliards de dollars.

En Europe, la réglementation sur la protection des données à caractère personnel révèle encore dans ce contexte sa robustesse et sa pertinence. Elle parvient à rassurer sur les principales dérives et se trouve traduite au sein de l'Union européenne par une directive en 1995^[8], inspirée de la Convention 108 du Conseil de l'Europe^[9]. La construction juridique qui en résultera, jusqu'à l'adoption du RGPD^[10], sera la clé de voûte de la protection des individus dans ce tout nouvel environnement numérique qui commence, dès les années 2000, à agréger des quantités considérables de données. Mais, paradoxalement, c'est peut-être ce succès qui contribuera à négliger d'autres règles contraignantes, qu'il s'agisse d'encadrer le marché (le marché se corrigera d'ailleurs de lui-même dramatiquement le 14 avril 2000, suivi de la crise boursière post-attentats du 11 septembre) ou d'intervenir face à de nouveaux acteurs que l'on a du mal à catégoriser, avec une question centrale : quels sont leurs niveaux exacts de responsabilité ? Mais, là encore, cette question conduira à en négliger d'autres, bien plus profondes, comme les conséquences sur nos capacités de perception lorsque l'on emploie plusieurs heures par jour des applications redoublant d'ingéniosité et jouant de nos biais cognitifs pour retenir toujours plus notre attention^[11]... au point que la dopamine libérée à chaque gratification reçue pour une publication nous conduise à ressentir des effets de manque comparables à d'autres addictions lors de déconnexions^[12].

Le concept même d'une société de l'information et de la communication, promu avec la généralisation de l'internet et désignant l'ère post-industrielle, a pris acte de ce déterminisme technologique et a irrigué une action publique frileuse d'entraver l'innovation et de ralentir de très importantes perspectives de croissance^[13]. Il ne s'agit pas, bien entendu, de nier la place croissante et prépondérante du numérique à notre époque, mais bien plus d'interroger s'il paraît opportun de lier inconditionnellement l'amélioration de nos conditions de vie et le maintien du lien social à la sophistication d'outils dont les concepteurs captent une part de plus en plus importante de pouvoir sur l'organisation de nos vies en dehors de tout contrôle démocratique^[14]. Et, partant, d'interroger également comment les politiques de régulation ont bien plus servi à consolider de nouvelles positions dominantes qu'à protéger les individus et la société d'atteintes pourtant substantielles.

#3.3. La résurgence d'une critique profonde de l'informatique et de l'internet divertie par la résurrection du terme « intelligence artificielle »

C'est la justice qui interviendra en premier dans un contexte de frilosité des régulateurs de face aux géants du numérique. L'arrêt Google Spain de la CJUE du 13 mai 2014^[15] consacrera ainsi un droit à l'oubli (ou, tout du moins, un droit à l'effacement) et contribuera, dans un contexte où Edward Snowden venait de révéler l'année précédente un système de surveillance de masse dirigé depuis les États-Unis, à tempérer la perception publique d'un impact bénéfique du « web » et des outils numériques. Même si elles n'avaient pas totalement disparu depuis les années 1980, les approches critiques sur ces « nouvelles technologies » (qui ne le sont désormais plus) trouvent au milieu des années 2010 une nouvelle audience. Le refus d'utilisation de certains objets se développera, tout comme la manifestation publique de déconnexions volontaires. Ces attitudes témoigneront, moins que d'une technophobie, d'un choix idéologique et politique réfutant de contribuer à l'emprise sur nos vies de groupes industriels et d'une volonté de reprendre la main sur un quotidien de plus en plus synchronisé sur le rythme des notifications émanant des diverses applications peuplant nos téléphones mobiles. L'idée d'une nouvelle ère biopolitique est soutenue par des auteurs comme Lawrence Lessig^[16] (« code is law »), Antoinette Rouvroy^[17] (qui relève l'émergence d'une gouvernamentalité algorithmique) ou Adrien Basdevant^[18] (qui dénonce un « coup data »).

Mais, dans le même temps, un terme que l'on croyait durablement passé de mode est revenu dans le paysage à l'orée des années 2010. En véritable concept zombie, « l'IA » a ressurgi des tréfonds où elle avait été soigneusement enterrée à la fin des années 80 pour réenchanter toute l'industrie numérique. Trois mousquetaires (ou plutôt conspirateurs^[19]), Yann LeCun, Youshua Bengio et Geoffrey Hinton arriveront à démontrer que des algorithmes particuliers fondés sur une approche statistique des données, vieux de quelques décennies, sont en mesure de fonctionner remarquablement mieux que d'autres,

alors à la mode. L'apprentissage automatique (*machine learning*) et particulièrement l'apprentissage profond (*deep learning*) se révélera extrêmement efficace pour reconnaître des images ou des sons. Grisés par ces succès, la recherche et l'industrie numérique donneront une moindre priorité à la plupart de leurs autres travaux pour généraliser l'usage de ces algorithmes dans tous des domaines d'activités. Même en l'absence de preuves solides démontrant que les modèles établis restituent bien ce qu'ils sont censés établir, « l'IA » est commercialisée non seulement dans l'industrie ou les services, mais aussi dans des domaines comme la justice, ou la santé. Les perspectives de croissance vertigineuse^[20] de ce qui n'est qu'une application, pas tout à fait nouvelle, de l'informatique^[21] contamineront l'entière communauté, publique et privée. Les ambassadeurs ne manqueront pas pour convaincre les sceptiques des bénéfices majeurs d'instruments censés pouvoir décider de manière plus fiable que les humains, non pas à partir d'une connaissance empirique, mais d'une représentation statistique de n'importe quel environnement^[22].

Les mêmes causes produisant les mêmes effets, il est fourni aux détracteurs de « l'IA » les mêmes réponses qu'aux détracteurs d'autres innovations technologiques. À ceux qui dénoncent notamment la déshumanisation des activités transformées par cette technologie, l'aggravation des discriminations, la contribution à l'avènement d'une société de surveillance et l'amenuisement de l'autonomie d'action, il est répondu que résister à la généralisation de cette technologie, ce serait se mettre en retard par rapport à d'autres nations se souciant moins de régulation, ce serait aussi rater une occasion de développement économique majeur et ce serait – avant tout – résister au progrès. Les efforts d'encadrement se sont donc d'abord exprimés sous la forme de diverses formes de régulations très peu contraignantes, comme des lignes directrices, des principes éthiques ou des déclarations. L'emballement de la production de tels documents, à son acmé en 2018, a été également accompagné par une mobilisation progressive des organisations internationales à l'écoute des craintes du public, comme les Nations Unies, dont l'UNESCO, l'OCDE, l'Union européenne ou encore le Conseil de l'Europe. L'attention portée à la transformation numérique de la société s'est cependant focalisée sur les promesses et les opportunités de cette « IA », quitte à en adopter une définition relativement floue, et les débats se sont assez rapidement stabilisés sur la question des usages – qui seraient seuls à encadrer par des dispositions plus ou moins contraignantes. Sur la foi de succès dans des secteurs très spécialisés, il a été considéré sans grande rigueur scientifique qu'un emploi assez généralisé était possible. Encore une fois, les quelques résistants ont été relégués dans les rangs des néo-luddites du XXI^e siècle.

Notes

^[1] R. Cellan-Jones, [Stephen Hawking warns artificial intelligence could end mankind](#), BBC news, 2 décembre 2014 et E. Musk, entretien dans le documentaire [Do you trust that computer](#), 5 avril 2018

^[2] Voir, parmi d'autres, les développements réalisés par Bernard Stiegler, Dominique Cardon, Antoinette Rouvroy, Evgeny Morozov ou Eric Sadin.

^[3] Déjà, en 1964 à Berkeley, des manifestants ont arboré des cartes perforées autour du cou et ont dénoncé l'aliénation potentielle d'une transformation des humains en données.

^[4] G. Tremblay, De Marshall McLuhan à Harold Innis ou du village global à l'empire mondial, tic&société, Vol. 1, n°1, 2007

^[5] J. Ellul, *Changer de révolution – L'inéluctable prolétariat*, Coll. Empreintes, Seuil, 1982

^[6] Voir le discours du ministre des affaires étrangères de la RFA, Hans Dietrich Genscher, devant l'Assemblée Parlementaire du Conseil de l'Europe – 37^{ème} session, 1985, cité par F. Jarrige, *Techno-critiques, Du refus des machines à la contestation des technosciences*, op.cit., p. 293

^[7] L'industrie informatique s'était déjà appropriée cette image en... 1984, avec la publicité d'Apple, diffusée le 22 janvier 1984. Réalisée par Ridley Scott, cette publicité reprendra l'univers de George Orwell et promettra une année 1984 qui ne ressemblera pas au livre 1984 grâce à l'ordinateur Macintosh. Accessible sur : <https://youtu.be/axSnW-ygU5g> – Consulté le 17 août 2020

^[8] Directive 95/46/CE du Parlement européen et du Conseil, du 24 octobre 1995, relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données

^[9] Convention STE n°108, pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel

^[10] Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, abrogeant la directive 95/46/CE

^[11] B. Patino, *La civilisation du poisson rouge*, Grasset, 2019

^[12] T. Ong, Sean Parker on Facebook: 'God only knows what it's doing to our children's brains', The Verge, 9 novembre 2017

^[13] E. George, En finir avec la « société de l'information » ?, tic&société, Vol. 2, n° 2, 2008

^[14] E. George et F. Granjon (dir.), *Critiques de la société de l'information*, Paris, L'Harmattan, 2008, p.10

^[15] CJUE, gde ch., 13 mai 2014, aff. C-131/12, Google Spain SL et Google Inc. / Agencia Espanola de Proteccion de Datos et Gonzales

^[16] L. Lessig, *Code is Law*, Harvard magazine, 2000

^[17] A. Rouvroy et T. Berns, *Gouvernementalité algorithmique et perspectives d'émancipation*, Réseaux 2013/1, n°177, 2013, pp163-196

^[18] A. Basdevant, J-P. Mignard, *L'empire des données, Essai sur la société, les algorithmes et la loi*, Don Quichotte, 2018.

^[19] Terme employé par Yann LeCun lui-même, rapporté par J. Markoff, *Machines of Loving Grace – The Quest for Common Ground between Humans and Robots* : Harper Collins Publishers, 2015, p.150

^[20] Pour des projections financières relatives au développement de l'IA, voir la fiche d'information de la Commission européenne, marché unique numérique : http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51610 – Consulté le 17 août 2020

^[21] Il conviendrait en effet de distinguer le projet initial de l'informatique, calculatoire, de l'emploi de l'informatique aux fins de la reproduction de mécanismes de raisonnement et d'apprentissage, apte à qualifier le projet de « l'intelligence artificielle ». Certains courants ont toutefois assimilé pensée et calcul, au travers de la logique, notamment le courant « symboliste », pensant que la modélisation de la pensée ne pouvait être réalisée que par l'écriture d'un code informatique signifiant, enchaînant les règles logiques.

^[22] Pour découvrir un argumentaire courageux, transformant contre toutes les évidences les faiblesses de l'apprentissage automatique en force, afin de convaincre des prospectus de sa puissance à « prédire » les événements lire par exemple : *AI in manufacturing: Do umbrellas cause car accidents?*, accessible sur : <https://cosmotech.com/resources/do-umbrellas-cause-car-accidents/> – Consulté le 17 août 2020

Partie 4 | Revitaliser le contenu d'une réglementation sur « l'intelligence artificielle »

De manière assez surprenante, les bénéfices systémiques de la généralisation de l'informatique, notamment sur la productivité, restent à démontrer. En 1987, Robert Solow, prix Nobel d'économie, annonçait « on voit des ordinateurs partout, sauf dans les statistiques de productivité ». Ce paradoxe révélait l'absence de lien observable, au niveau micro-économique, entre les investissements informatiques et la productivité des entreprises. Le constat apparaît toujours d'actualité, puisqu'il s'agisse d'internet et de sa bulle, ou de la généralisation des téléphones intelligents (smartphones) et de « l'IA », l'économie peine à décoller^[1]. La crise économique résultant de la crise sanitaire due à la COVID-19 ne va vraisemblablement pas renverser la situation, puisque l'on a vu que cette « IA », pourtant promise à résoudre la plus vaste gamme de problèmes notamment en matière de santé, n'a eu en réalité qu'assez peu de résultats opérationnels^[2]. Il semble que nous héritons encore de l'influence d'analyses assez anciennes, comme celle de Daniel Bell qui estimait que la clé du passage de la société industrielle à la société post-industrielle serait dans l'augmentation de la productivité liée aux activités informationnelles^[3] – cette fameuse « société de l'information » – ou de Simon Nora et Alain Minc, selon laquelle il faudrait informatiser l'économie pour produire de la croissance^[4].

#4.1. Des initiatives de réglementation prenant acte d'un lien entre innovation technologique et croissance économique

Les préambules des travaux réalisés à Bruxelles, notamment par le groupe d'experts indépendants de haut niveau mis en place par la Commission Européenne, ou à Paris dans le cadre de l'OCDE ne démentent pas cette filiation de pensée. Les « Lignes directrices en matière d'éthique pour une IA digne de confiance^[5] » des experts mandatés par la Commission reconnaissent que les systèmes d'IA soulèvent des risques (paragraphe 10) mais déclarent avant tout que « Nous sommes convaincus que l'IA est susceptible de transformer la société de manière significative. L'IA n'est pas une fin en soi, mais plutôt un moyen prometteur d'accroître la prospérité humaine, en renforçant ainsi le bien-être individuel et de la société ainsi que le bien commun, et en étant porteur de progrès et d'innovation » (paragraphe 9). Le livre blanc de la Commission européenne sur l'IA^[6] identifie aussi un certain nombre de risques, mais ouvre sur un certain nombre d'affirmations : « [L'IA] va entraîner des changements dans nos vies en améliorant les soins de santé (précision accrue des diagnostics ou meilleure prévention des maladies, par exemple), en rendant l'agriculture plus efficace, en contribuant à l'adaptation au changement climatique et à

l'atténuation de ses effets, en augmentant l'efficacité des systèmes de production par la maintenance prédictive, en renforçant la sécurité des Européens et de bien d'autres façons que nous commençons à peine à entrevoir ». Elle précise : « La Commission prône donc une approche axée sur la régulation et l'investissement, qui poursuit le double objectif de promouvoir le recours à l'IA et de tenir compte des risques associés à certaines utilisations de cette nouvelle technologie ». De même, la « Recommandation du Conseil sur l'intelligence artificielle^[7] » de l'OCDE débute par ce constat : « L'intelligence artificielle (IA) est une technologie générique qui promet d'améliorer le bien-être des individus, de contribuer à une activité économique mondiale dynamique et durable, de stimuler l'innovation et la productivité, et d'aider à affronter les grands défis planétaires ». Pour les deux organisations, la question réside avant tout dans la nécessité de créer une « IA digne de confiance » (*trustworthy AI*), centrée sur l'humain (*human-centred*) et fondée sur les droits de l'homme afin de soutenir l'innovation et le progrès, ainsi qu'une croissance économique prometteuse. Dans son « Étude préliminaire sur les aspects techniques et juridiques liés à l'opportunité d'un instrument normatif sur l'éthique de l'intelligence artificielle^[8] », l'UNESCO souligne pour sa part que : « La nouvelle économie numérique qui voit le jour représente des défis et des possibilités immenses pour les sociétés d'Afrique et des autres pays en développement. Du point de vue éthique, l'IA devrait être intégrée dans les politiques et les stratégies nationales de développement, en prenant appui sur les cultures, les valeurs et les connaissances endogènes pour développer les économies africaines » (paragraphe 10). Le Conseil de l'Europe évoque également dans le préambule de sa Recommandation sur les impacts des systèmes algorithmiques sur les droits de l'homme : « Gardant à l'esprit le potentiel important d'innovation socialement bénéfique et de croissance économique que recèlent les technologies numériques^[9] ».

Or, en lisant ces présomptions sous le prisme du « paradoxe de Solow » et même si le taux de pauvreté absolu paraît reculer dans le monde durant ces dernières décennies^[10], il convient de constater que nous n'avons pas assisté à une résorption des inégalités et que les promesses d'une croissance revivifiée par l'investissement massif dans l'innovation numérique ne sont pas tout à fait au rendez-vous. De nombreuses applications de « l'IA », présentées comme une solution à des problèmes souvent bien complexes, se révèlent non seulement peu matures, mais surtout inadaptées à produire le service qu'elles prétendent rendre. Par exemple, reconnaître une image ou gagner au jeu de go à l'aide d'apprentissage automatique (*machine learning*) est une chose, mais penser pouvoir évaluer avec la même technologie la prétendue dangerosité d'un individu ou prédire l'indemnisation prononcée par un tribunal en est une autre. Non seulement parce qu'il s'agit d'une résurgence claire d'une conception déterministe des individus^[11] en matière pénale ou une méconnaissance absolue de ce qu'est un système juridique en matière civile^[12], mais aussi parce que, structurellement, manipuler des données quantifiables, dans des environnements fermés, n'est pas la même chose que de manipuler des données qualitatives, dans des environnements ouverts. L'on cherche également à amasser des quantités considérables de toutes sortes de données, pensant que l'on capitalise ainsi une ressource essentielle, alors

que, qualitativement, tout ne se vaut pas^[13]. Par ce manque de rigueur, l'on cautionne donc la production de résultats nécessairement décevants, voire dangereux, décrédibilisant potentiellement toute forme d'application et pesant ensuite sur la confiance des investisseurs et des usagers^[14].

Concevoir une réglementation de technologies comme « l'IA » devrait donc admettre, en préambule, que notre manière de quantifier et d'analyser l'innovation et le progrès n'est peut-être pas la bonne. L'innovation technologique doit-elle être nécessairement vue comme un progrès pour le genre humain ? La croissance économique est-elle un facteur inconditionnel d'amélioration des conditions de vie ? Ou faut-il définir et mesurer autrement le bien-être que par le seul prisme du PIB ? Par une réglementation effective et stricte de « l'IA », il s'agirait donc de parvenir à considérer ce que nous devrions en faire plutôt que de spéculer sur ce que nous pourrions en faire. Le progrès pourrait ainsi naître non pas de vagues principes éthiques, même transformés en textes juridiques, censés garantir des dérives, mais de l'exclusion pure et simple d'applications commerciales ou publiques bien trop hasardeuses.

#4.2. Recentrer les travaux d'une réglementation de « l'intelligence artificielle » en connaissance du réel état de l'art

Il paraît donc urgent d'arriver à se départir des éléments de langage et des mots-valises répétés à l'envi, qui ne sont pas singuliers à « l'IA », pour esquisser un regard objectif sur le numérique et concevoir une réglementation en capacité de traiter les réels enjeux de société. Il est naturellement indispensable d'arriver à prévenir des discriminations de tous ordres, à échapper à la constitution progressive d'une société de surveillance et à préserver la vie privée, à garantir notre autonomie d'action ou encore assurer notre liberté d'expression. Mais l'atteinte de ces objectifs doit se concevoir en réinvestissant une approche scientifique sérieuse, excluant en amont et d'emblée l'emploi de techniques et de technologies trop peu matures ou simplement inadaptées. Le scandale provoqué par la publication d'une étude extrêmement approximative sur l'effet d'un médicament sur le coronavirus par The Lancet est malheureusement représentatif de cette « épidémie de mauvaise science » évoquée en première partie de cette étude, le rédacteur en chef de cette revue reconnaissant que « cet épisode représente un échec complet pour la science^[15] ».

Les régulateurs, nationaux et internationaux, se doivent d'être informés de l'état de l'art précis et non simplement récepteurs de discours forgés par une industrie soucieuse d'asseoir toujours plus son hégémonie sur notre temps^[16]. La crise économique, qui nous touche déjà, ne doit pas nous donner l'occasion d'augmenter le maillage de l'emprise numérique au prétexte d'une recherche tout à fait hypothétique de croissance économique (elle n'était déjà pas au rendez-vous avant). Elle devrait, au contraire, nous inviter à nous

ressaisir pour évaluer exactement la qualité de ce qui nous est vendu et les conséquences d'un transfert de gouvernance, déjà bien avancé, entre des États abandonnant nombre de leurs prérogatives au bénéfice d'opérateurs privés, dont la vocation n'est ni de garantir, ni de renforcer l'intérêt général.

Une réglementation effective sur « l'IA » devrait donc s'appuyer sur une expertise de haut niveau dégagée des intérêts économiques ou de souveraineté des États, fondée sur des preuves solides et reproductibles. C'est l'ensemble de la démarche d'évaluation qui serait à dégager de ces influences pour constituer un réseau d'experts multidisciplinaires, représentant une conscience commune, dont la boussole serait solidement ancrée vers un progrès humain, si cher aux Lumières, et non vers la « technique ». Garantir, en amont, la qualité et la neutralité de la production scientifique, notamment de la recherche publique, c'est donc garantir que la production de normes, en aval, ait du sens.

#4.3. Les défis et les opportunités d'une réglementation effective de « l'intelligence artificielle »

Le contexte politique, économique et social n'est toutefois pas favorable à un tel ressaisissement, même si la crise sanitaire a confirmé de manière flagrante la faiblesse structurelle du complexe scientifico-industriel. Ajoutons à cela l'affaiblissement de la primauté d'une règle de droit au profit de mécanismes d'autorégulation en plus du transfert de gouvernance, déjà évoqué, vers des opérateurs privés et l'on comprend les difficultés pour les régulateurs pour imposer des vues strictes sur la régulation de « l'IA » en rupture avec le consensus ambiant sur l'irréductible conception du progrès par la technique. Il faut dire que le sentiment diffus de retard généralisé dans lequel nous sommes plongés durant ces dernières décennies^[17] nous impose une marche forcée vers un progrès technologique salvateur. Ce même sentiment nous enjoint à nous adapter à tout prix, le plus vite possible, sans se donner le temps de problématiser... et avec un regard en réalité fixé à court terme. C'est même à se demander si, après la longue ère de la primauté temporelle de la religion puis l'avènement, avec le temps des Lumières, de l'État de droit pour organiser au mieux les affaires humaines nous ne serions pas en train d'entrer dans une autre ère : celle d'une véritable primauté d'une interprétation du monde par les algorithmes (et d'un « État des algorithmes^[18] »), dont l'instrument opérationnel serait « l'IA ». Avec un résultat concret : la délégation, de fait, de l'essentiel de l'administration des affaires de notre société à des opérateurs privés, hors de tout contrôle démocratique.

Pourtant, au vu de l'impact de cette véritable doxa, il nous faudra bien parvenir à dépasser la dialectique nous étant imposée, liant irrémédiablement ces technologies numériques avec la croissance économique, pour échapper à l'ultime stade d'une évolution technocentrée, créant à la fois les conditions du mal et son remède terrible : le transhumanisme. Car ce qu'il y aurait à considérer dans les années à venir, c'est autant la sacralisation et les

dérives imposées par une technologie précise, comme « l'IA », que les dérives résultant de la conjonction de technologies telles que les nanotechnologies, les biotechnologies, les technologies de l'information et les sciences cognitives (NBIC). Là encore, la fameuse loi de Gabor n'a rien d'inéluctable, et ce sera la conjonction d'une approche scientifique rigoureuse et véritablement éthique, alliée à un cadre juridique fondé sur les droits de l'homme, la démocratie et l'État de droit qui donnera le cap d'une amélioration de la condition de l'humanité. D'une manière assez paradoxale, c'est même d'ailleurs la réelle prise en compte des discours critiques qui peuvent devenir leur meilleure garantie de la longévité d'un projet capitaliste et néolibéral.

Des réglementations strictes sont en effet déjà intervenues massivement dans des domaines industriels comme l'automobile (avec la prolifération de mesures de sécurité actives et passives dans les véhicules) ou les médicaments (avec la certification avant la mise sur le marché). Ces encadrements n'ont en rien ralenti l'innovation et, au contraire, ont été une garantie de confiance pour les consommateurs^[19]. Alors que cette même confiance est recherchée pour « l'IA », il semble surprenant qu'une partie de l'industrie numérique tente encore d'étouffer les critiques, pourtant substantielles, ou de ralentir l'adoption de normes strictes, à même de dépasser les seules incantations appelant à une « IA digne de confiance ». Reste encore à déterminer les fondements et les axes d'une réglementation qui serait effective et non cosmétique.

#4.4. Les fondements et les axes principaux d'une réglementation effective de « l'intelligence artificielle »

Parmi les fondements identifiés par la plupart des organisations internationales, les droits de l'homme ou les droits fondamentaux figurent de manière constante comme l'une des briques essentielles de leurs instruments juridiques. Bien que critiqués assez sévèrement, tant par les promoteurs d'un marché libre que les critiques d'un hyperindividualisme^[20], les droits de l'homme apparaissent en effet suffisamment généraux, transversaux et rassurants pour constituer la base d'une réglementation recherchant la confiance. C'est d'ailleurs sur ce fondement que le premier texte international de protection des données à caractère personnel, la Convention 108 du Conseil de l'Europe^[21], s'est appuyé en 1981 en rejoignant le corpus des traités de l'organisation, comprenant déjà la Convention de sauvegarde des Droits de l'Homme et des Libertés fondamentales^[22]. En ce qui concerne « l'IA », et toujours en référence à la Convention européenne des droits de l'homme, l'on y trouve des garanties substantielles en ce qui concerne le droit au respect de la vie privée et familiale (art. 8), le droit à la liberté d'expression (art. 10), le droit à un procès équitable (art. 6), l'interdiction des discriminations^[23] (art. 14) et, par construction jurisprudentielle, la dignité^[24]. Autant de garanties qui présentent un intérêt majeur pour prévenir des dérives bien documentées d'emploi de « l'IA » ou des algorithmes, telles que les prétentions d'évaluation des risques

de récidive d'un individu en matière pénale par exemple^[25] ou d'atteintes à la démocratie par des tentatives de manipulation de masse^[26].

Il sera tout d'abord réfuté ici toute une série d'arguments visant à écarter, à atténuer ou à instrumentaliser les droits de l'homme dans le seul objectif de « blanchir » « l'IA » ou de créer un sentiment de sécurité bien artificiel. Ainsi, il a pu être soutenu qu'il était inutile de créer de nouveaux textes spécifiques à cette technologie, du fait de son caractère novateur et mouvant, ou qu'il ne faudrait que des réglementations très spécialisées, ceci dans le seul objectif de laisser des interstices dans lesquels se faufiler pour échapper aux règles édictées. Il a pu aussi être fait référence à des concepts vagues (« centré sur l'humain ») ou des principes qui laissent une marge suffisante d'interprétation pour créer faussement un sentiment de confiance sans réelles garanties. Il s'agira donc de se concentrer sur des axes potentiels de réglementation ne considérant pas les droits de l'homme comme une clé universelle, mais comme une condition nécessaire et insuffisante à remplir seule cette fonction d'encadrement juridique de « l'IA ». C'est pourquoi il sera également fait appel à d'autres séries de principes qui, pour être totalement effectifs, devraient être également accompagnés de mécanismes de suivi.

Une telle réglementation devrait adopter tout d'abord une définition technologiquement neutre et suffisamment large de « l'IA » afin de construire un mécanisme juridique global et cohérent apte à appréhender l'essentiel des applications informatiques susceptibles d'avoir un impact significatif^[27] sur les individus ou la société. La notion de « systèmes algorithmiques », développée notamment dans une Recommandation du Conseil de l'Europe^[28], pourrait d'ailleurs au final être préférée au terme « d'IA ». La motivation sous-jacente à une telle approche est de parvenir à appréhender le plus grand nombre de situations spécifiques résultant de l'emploi de l'informatique, notamment celles tenant à la prise de décision ou à l'assistance à la prise de décision. Cette réglementation devrait ensuite poser un certain nombre de principes^[29], parmi lesquels on pourrait citer:

- **Un principe de précaution** : En cas de risque de dommages graves ou irréversibles résultant de l'emploi de systèmes algorithmiques, l'absence de certitude scientifique absolue ne devrait pas servir de prétexte pour remettre l'adoption de mesures effectives visant à prévenir tout dommage sur les individus, la société et l'environnement. La mise sur le marché ou l'emploi par le secteur public devraient être purement et simplement différés ou prohibés, cela n'empêchant pas la recherche de perfectionner des outils socialement acceptables avec un solide encadrement éthique^[30].
- **Un principe de proportionnalité de recours aux algorithmes** : Le recours à des systèmes algorithmiques, notamment les systèmes de prise de décision (ou d'aide à la prise de décision) ayant un impact significatif sur les individus et la société, serait à envisager que s'il en est démontré la nécessité, une particulière valeur ajoutée et s'il n'existe pas d'autres modalités aussi efficaces moins intrusives, et moins consommatrices d'énergie, pour exécuter une tâche ou offrir un service. Cette proportionnalité sera particulièrement

recherchée pour protéger les droits de l'homme et les libertés fondamentales, en prohibant l'emploi systématisé de technologies très intrusives dans les espaces publics, comme la reconnaissance faciale, sauf pour des circonstances exceptionnelles et sous un encadrement strict de l'autorité judiciaire. Cette proportionnalité serait également à envisager pour les services publics afin de limiter les effets de la fracture numérique entre les différentes parties de la population. Elle pourrait aussi être recherchée dans tout autre secteur d'activité afin de favoriser quand cela est nécessaire les interactions humaines directes, sans intermédiation numérique.

- **Le respect de la dignité humaine** : En convergence avec les nanotechnologies, les biotechnologies et les sciences cognitives, l'emploi de systèmes algorithmiques interfacés directement sur les êtres humains à des fins d'augmentation de leurs capacités ne devrait être envisagé que dans des situations particulières de compensation de handicap. La recherche et le développement de tels systèmes devraient être encadrés selon les mêmes principes directeurs que pour la biomédecine.
- **Le renforcement de la solidarité** : Les bénéfices résultant de l'emploi de systèmes algorithmiques devraient être redistribués équitablement afin que l'innovation participe au renforcement de la cohésion sociale, au bien-être des populations et des individus ainsi qu'à un développement humain harmonieux, comme cela est envisagé dans les objectifs de développement durable des Nations Unies.

Bien d'autres axes pourraient être évoqués pour compléter ce qui pourrait être intégré dans l'ordre juridique international sous la forme d'une convention cadre par exemple^[31], notamment en ce qui concerne les données (afin de trouver une articulation avec les dispositions relatives à la protection des données) ou la certification ex ante pour s'assurer de la conformité des systèmes^[32]. Même si les précédentes expériences de ce type d'instrument juridique de haut niveau démontrent une efficacité qui pourrait être discutée du fait de leur absence d'effet juridique direct^[33], la naissance d'un consensus international ambitieux en la matière ne pourra naître que progressivement.

L'intégration de principes reconnaissant que le progrès humain n'est pas inéluctablement lié à la technique et, en parallèle, un nouveau souffle donné à une recherche publique crédible, déconnectée d'intérêts financiers et marchands, sont vraisemblablement les clés d'un développement durable de nos sociétés, qui mériteraient d'être portées par un nouveau projet politique qui tirerait, de manière large, toutes les conséquences des enseignements de cette crise sanitaire.

Notes

^[1] De manière plus sectorielle, un rapport du Conseil de l'Europe (Commission européenne pour l'efficacité de la justice – CEPEJ) concluait en 2016 à la difficulté de démontrer un lien entre l'informatisation des tribunaux et l'amélioration de leur

efficacité – Rapport thématique : l'utilisation des technologies de l'information dans les tribunaux en Europe. Études de la CEPEJ n°24, 2016

^[2] J. Bullock, A. Luccioni, K. Hoffmann Pham, C. Sin Nga Lam, M. Luengo-Oroz, Mapping the landscape of Artificial Intelligence applications against COVID-19, UN Global Pulse, 24 avril 2020

^[3] D. Bell, *The Coming of Post-Industrial Society*. New York: Harper Colophon Books, 1974

^[4] S. Nora, A. Minc, L'informatisation de la société, Rapport au Président de la République, La Documentation Française, 1978

^[5] Groupe d'experts indépendants constitué par la Commission européenne en juin 2018, Lignes directrices en matière d'éthique pour une IA digne de confiance, Avril 2018

^[6] Livre blanc sur l'intelligence artificielle – Une approche européenne axée sur l'excellence et la confiance, Février 2020

^[7] OCDE, Recommandation du Conseil sur l'intelligence artificielle, OECD/LEGAL/0449, mai 2019

^[8] UNESCO, Étude préliminaire sur les aspects techniques et juridiques liés à l'opportunité d'un instrument normatif sur l'éthique de l'intelligence artificielle, 206 EX/42, Mars 2019

^[9] Conseil de l'Europe, Recommandation du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme, CM/Rec(2020)1, Avril 2020

^[10] La Banque mondiale a proposé, en 2018, deux nouveaux seuils de pauvreté : le premier, à 3,2 dollars par jour, qui conduit à compter 25 % de pauvres sur la planète, et le second à 5,5 dollars, qui implique près de 50 % de pauvres. Toutefois, ces comparaisons internationales ne portent désormais plus sur la « pauvreté absolue », avec différents seuils, mais sur la « pauvreté relative », qui recense les ménages vivant sous une certaine fraction du revenu médian. Selon cette grille de lecture, si la pauvreté absolue baisse dans les pays en développement (1,84 milliard d'individus en 1990 contre 766 millions en 2013), la pauvreté relative augmente (482 millions d'individus en 1990 contre 1,32 milliards en 2013). En d'autres termes, le dénuement total paraît diminuer alors que les inégalités progressent.

^[11] Au XIX^{ème} siècle, le professeur italien de médecine légale Cesare Lombroso avait cru pouvoir identifier les potentiels criminels sur la base d'études phrénologiques et physiognomoniques, c'est-à-dire en examinant la forme des crânes et les traits des visages. Bien que ses assertions aient été réfutées depuis lors, notamment pour faire émerger des politiques pénales centrées sur la réinsertion des individus, ce courant de pensée demeure encore actif et a été revivifié par les possibilités statistiques issues de « l'IA ». Voir notamment l'entretien avec Michal Kosinski, psychologue en sciences sociales computationnelles, qui pense avertir des dangers de « l'IA » et de sa capacité, par exemple, à identifier des homosexuels sur la base des traits de leur visage – UpNorth Production, iHuman, Accessible sur : <https://youtu.be/20kV0CcFSKE> (entretien à 35mn20) – Consulté le 21 août 2020

^[12] Les systèmes juridiques actuels sont très éloignés de l'idéal de rationalité que devait incarner en France le Code civil de 1804. Il existe une multitude de sources qui ne s'imbriquent pas de manière parfaitement cohérente et qui portent sur un ensemble de règles dont la signification demeure très indéterminée, ce que le théoricien du droit Herbert L. A. Hart qualifiait de « texture ouverte du droit » – Voir notamment H. L. A. Hart, *Le concept de droit* : Facultés universitaires Saint-Louis Bruxelles éd., coll. Droit, 1976

^[13] Sur les grandes ambitions de l'Union européenne sur cet amoncellement de données, voir D. Perrotte, Thierry Breton : « La guerre des données industrielles débute maintenant et l'Europe sera son principal champ de bataille », Les Echos, 17 février 2020

^[14] Voir le dossier Artificial intelligence and its limits dans la revue *The Economist* du 11 juin 2020, *op.cit.*

^[15] S. Delesalle-Stolper, «The Lancet» : «L'arrogance de l'Occident est responsable de dizaines de milliers de morts», Libération, 15 juin 2020

^[16] Sur la manière dont les médias relaient les discours de l'industrie numérique sur « l'IA », voir J. Scott Brennan, An Industry-Led Debate : How UK Media Cover Artificial Intelligence, University of Oxford, Reuters Institute for Study of Journalism, 13 décembre 2018

^[17] Ba. Stiegler, « Il faut s'adapter » : sur un nouvel impératif politique, *op.cit.*

^[18] R. Batko, J. Kreft, *The Sixth Estate – The Rule of Algorithms*, *Problemy Zarzadzania*, University of Warsaw, Faculty of Management, vol. 15(68), 2017, pages 190-209

^[19] Pour reprendre l'exemple automobile, le 1er juillet 1973, la ceinture de sécurité à l'avant des véhicules automobiles a été rendue obligatoire en France. Il s'agit alors du premier pays européen à imposer ce dispositif, après une année 1972 particulièrement meurtrière avec 18 034 personnes tuées. Mal acceptée au départ, cette mesure a permis pourtant de réduire dès le mois de décembre de 30 % le nombre de tués. Alors même que la population et le parc automobile se sont accrus depuis lors, 3 259 morts ont été à déplorer en 2018.

^[20] En ce qui concerne la critique des droits de l'homme, et sa réfutation, voir J. Lacroix et J-Y. Pranchère, *Les droits de l'homme rendent-ils idiots ?*, Seuil, 2019

^[21] Convention STE n°108, pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel

^[22] Convention STE n°5, de sauvegarde des Droits de l'Homme et des Libertés fondamentales dite « Convention européenne des droits de l'homme »

^[23] Il est à relever que l'article 14 de la Convention EDH n'est jamais applicable isolément, mais toujours conjointement à un autre article énonçant un droit ou une liberté. Voir à ce sujet F. Edel, *L'interdiction de la discrimination par la Convention européenne des droits de l'homme*, Dossiers sur les droits de l'homme n°22, Éditions du Conseil de l'Europe, 2010

^[24] Le protocole d'amendement à la Convention 108 (STCE n°223) a ajouté au préambule de ce texte une mention explicite à la dignité humaine : « Considérant qu'il est nécessaire de garantir la dignité humaine ainsi que la protection des droits de l'homme et des libertés fondamentales de toute personne »

^[25] Concernant l'algorithme COMPAS aux États-Unis, voir par exemple J. Angwin, J. Larson, S. Mattu, L. Kirchner, *Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks*, ProPublica, 23 mai 2016

^[26] Voir à ce titre le site internet du journal « The Guardian », *The Cambridge Analytica Files*, accessible sur : <https://www.theguardian.com/news/series/cambridge-analytica-files> – Consulté le 21 août 2020

^[27] Voir Y. Meneceur, *L'intelligence artificielle en procès – Plaidoyer pour une réglementation internationale et européenne*, *op.cit.*, p.385

^[28] Conseil de l'Europe, *Recommandation du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme*, CM/Rec(2020)1, Avril 2020 (Annexe, paragraphe 2) : « (...) on entend par « systèmes algorithmiques » des applications qui, au moyen souvent de techniques d'optimisation mathématique, effectuent une ou plusieurs tâches comme la collecte, le regroupement, le nettoyage, le tri, la classification et la déduction de données, ainsi que la sélection, la hiérarchisation, la formulation de recommandations et la prise de décision ». Il a retenu pour la présente étude cette définition : « système composé d'un ou de plusieurs algorithmes utilisés dans un logiciel pour recueillir et analyser des données ainsi que pour tirer des conclusions dans le cadre d'un processus conçu pour exécuter une tâche ».

^[29] Une liste plus exhaustive des éléments constitutifs d'une telle réglementation a été dressé dans : Y. Meneceur, *L'intelligence artificielle en procès – Plaidoyer pour une réglementation internationale et européenne*, *op.cit.*, p.383 et s.

^[30] A. Rességuier et R. Rodrigues, *AI ethics should not remain toothless! A call to bring back the teeth of ethics*, *Big Data & Society*, 2020

^[31] Voir Y. Meneceur, *L'intelligence artificielle en procès – Plaidoyer pour une réglementation internationale et européenne*, *op.cit.*, p.383

^[32] Y. Meneceur, *L'intelligence artificielle en procès – Plaidoyer pour une réglementation internationale et européenne*, *op.cit.*, p.351 et s. et Y. Meneceur, *La procédure de certification : une régulation intelligente de contrôle... de l'intelligence artificielle*, *Revue Experts* n°148, février 2020

^[33] Voir la convention cadre de l'ONU sur le changement climatique et les conventions cadres du Conseil de l'Europe sur la coopération transfrontalière ou celle sur les minorités.

Bibliographie (ordre chronologique décroissant)

E. Klein, *Le goût du vrai*, Coll. Tracts, Gallimard, 2020

A. Ressayeur et R. Rodrigues, *AI ethics should not remain toothless! A call to bring back the teeth of ethics*, *Big Data & Society*, 2020

C. Kang, J. Nicas and D. McCabe, *Amazon, Apple, Facebook and Google Prepare for Their 'Big Tobacco Moment'*, *The New York Times*, 28 juillet 2020

Artificial intelligence and its limits, *The Economist*, 11 juin 2020

Y. Meneceur, *L'intelligence artificielle en procès – Plaidoyer pour une réglementation internationale et européenne*, Bruylant, 2020

J. Bullock, A. Luccioni, K. Hoffmann Pham, C. Sin Nga Lam, M. Luengo-Oroz, *Mapping the landscape of Artificial Intelligence applications against COVID-19*, *UN Global Pulse*, 24 avril 2020

Conseil de l'Europe, *Recommandation du Comité des Ministres aux États membres sur les impacts des systèmes algorithmiques sur les droits de l'homme*, CM/Rec(2020)1, Avril 2020

Livre blanc sur l'intelligence artificielle – Une approche européenne axée sur l'excellence et la confiance, Février 2020

Y. Meneceur, *La procédure de certification : une régulation intelligente de contrôle... de l'intelligence artificielle*, *Revue Experts* n°148, février 2020

R. Ochigame, *The invention of 'Ethical AI'*, *The Intercept*, 20 décembre 2019

B. Patino, *La civilisation du poisson rouge* : Grasset, 2019

J. Lacroix et J-Y. Pranchère, *Les droits de l'homme rendent-ils idiots ?*, Seuil, 2019

Ba. Stiegler, « Il faut s'adapter » : sur un nouvel impératif politique, Gallimard, coll. « NRF Essais », 2019

M. David, C. Sauviat, *Intelligence artificielle, la nouvelle barbarie*, Éd. du Rocher, 2019

B. Mittelstadt, *Principles Alone Cannot Guarantee Ethical AI*, *Nature Machine Intelligence*, novembre 2019

OCDE, *Recommandation du Conseil sur l'intelligence artificielle*, OECD/LEGAL/0449, mai 2019

A. Jobin, M. Ienca, et E. Vayena, *The global landscape of AI ethics guidelines*, *Nature Machine Intelligence* 1, septembre 2019, pp.389–399

OCDE, *Recommandation du Conseil sur l'intelligence artificielle*, OECD/LEGAL/0449, mai 2019

UNESCO, *Étude préliminaire sur les aspects techniques et juridiques liés à l'opportunité d'un instrument normatif sur l'éthique de l'intelligence artificielle*, 206 EX/42, Mars 2019

J. Scott Brennen, An Industry-Led Debate : How UK Media Cover Artificial Intelligence, University of Oxford, Reuters Institute for Study of Journalism, 13 décembre 2018

Groupe d'experts indépendants constitué par la Commission européenne en juin 2018, Lignes directrices en matière d'éthique pour une IA digne de confiance, Avril 2018

E. Sadin, L'intelligence artificielle ou l'enjeu du siècle : anatomie d'un antihumanisme radical, L'Echappée, 2018

A. Basdevant, J-P. Mignard, L'empire des données, Essai sur la société, les algorithmes et la loi, Don Quichotte, 2018

D. Cardon, J-P. Cointet, A. Mazières, La revanche des neurones. L'invention des machines inductives et la controverse de l'intelligence artificielle, Réseaux, 2018/5 (n° 211), p. 21

K. Schwab, La quatrième révolution industrielle, Dunod, 2017

R. Batko, J. Kreft, The Sixth Estate – The Rule of Algorithms, Problemy Zarzadzania, University of Warsaw, Faculty of Management, vol. 15(68), 2017, pages 190-209

O. Rey, Quand le monde s'est fait nombre, coll. Les Essais, Stock, 2016

Rapport thématique : l'utilisation des technologies de l'information dans les tribunaux en Europe, Études de la CEPEJ n°24, 2016

J. Angwin, J. Larson, S. Mattu, L. Kirchner, Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks, ProPublica, 23 mai 2016

J. Markoff, Machines of Loving Grace – The Quest for Common Ground between Humans and Robots : Harper Collins Publishers, 2015

F. Jarrige, Techno-critiques, Du refus des machines à la contestation des technosciences, La Découverte/Poche, 2016 (2014)

E. Morozov, Pour tout résoudre, cliquez ici, Fyp éditions, 2014

E. George et F. Granjon (dir.), Critiques de la société de l'information, Paris, L'Harmattan, 2008

E. George, En finir avec la « société de l'information » ?, tic&société, Vol. 2, n° 2, 2008

G. Tremblay, De Marshall McLuhan à Harold Innis ou du village global à l'empire mondial, tic&société, Vol. 1, n°1, 2007

A. Rouvroy et T. Berns, Gouvernamentalité algorithmique et perspectives d'émancipation, Réseaux 2013/1, n°177, 2013, pp163-196

L. Lessig, Code is Law, Harvard magazine, 2000

J. Ellul, Changer de révolution – L'inéluctable prolétariat, Coll. Empreintes, Seuil, 1982

D. Bell, The Coming of Post-Industrial Society. New York: Harper Colophon Books, 1974

J. Ellul, Les nouveaux possédés, Mille et une Nuits, 2003 (1973)

La critique de la technique : clé du développement de l'intelligence artificielle ?

L'intelligence artificielle (IA) continue de susciter de grands espoirs pour les années à venir, en moteur espéré de la prospérité humaine et du bien-être. À ces généreuses promesses émanant directement ou indirectement de l'industrie numérique répondent des inquiétudes de plus en plus substantielles dans l'opinion publique, notamment face à des applications suscitant la polémique ou éternellement sur le point d'être pleinement fonctionnels. Après des années de discours éthiques, les régulateurs internationaux et nationaux commencent à se mettre en ordre de marche, mais ils risquent de passer à côté de la cible en proposant des textes parfois trop peu, ou parfois trop, ambitieux. Et si la juste mesure était susceptible d'émerger des discours critiques de la technique, dont la pertinence est bien trop souvent sous-estimée par l'entière communauté ?

Cette étude a été publiée sur le blog *Les Temps Électriques* en septembre 2020.



Yannick Meneceur est magistrat de l'ordre judiciaire, détaché au Conseil de l'Europe, affecté en tant que chef d'unité du développement numérique. Il est également chercheur associé à l'IHEJ et ancien élève de l'IRA de Lille (qualification analyste). En tant que magistrat, il a exercé des fonctions pénales, notamment en matière de lutte contre la cybercriminalité et la contrefaçon. Il a aussi dirigé au ministère de la Justice un important projet informatique national. Au Conseil de l'Europe, il a contribué à l'analyse comparative des systèmes judiciaires européens et a développé une expertise en matière de transformation numérique, notamment en ce qui concerne la régulation de l'intelligence artificielle. Intervenant occasionnel à l'ENM, à l'ENA et à l'Université de Strasbourg, il est aussi membre de conseils scientifiques (Club des Juristes, Institut PRESAJE) et de l'Observatoire éthique et IA de l'Institut Sapiens. Son premier ouvrage, « L'intelligence artificielle en procès », a été publié en mai 2020 dans la collection Macro Droit - Micro Droit aux éditions Bruylant.