

# (I)<sup>2</sup>A

## L'ingénieur et l'IA

### 1. Introduction : un sujet vaste

Une certaine dose d'inconscience, peut-être même de folie, est sans doute nécessaire pour oser accepter la demande de rédaction d'un éditorial qui devrait aborder les questions éthiques soulevées par le développement de l'intelligence artificielle. Folie et inconscience, tant le sujet à traiter est vaste, d'autant plus que le développement de l'intelligence artificielle est quasi exponentiel.

Ajoutons que je ne suis pas ingénieur, ce qui ne facilite pas les choses pour repérer les enjeux éthiques d'un sujet... qui me dépasse. Je remercie d'autant plus les membres du Comité de rédaction de la revue dans laquelle je fais mon entrée pour la confiance qu'ils veulent bien m'accorder.

Néanmoins, à force de côtoyer des ingénieurs, qu'ils soient mes étudiants en formation initiale ou acteurs dans la vie professionnelle à l'occasion de jurys ou de visites d'entreprises, j'ai fini par comprendre les grands principes de sujets techniques divers plus passionnants les uns que les autres et j'ai progressivement partagé la logique de l'ingénieur.

C'est l'un d'eux qui m'amène à relever le défi de ce court texte. En décembre 2017, Amaury présentait, dans le cadre d'un cours d'éthique, un travail qu'il avait intitulé : « *L'homme bionique* ». J'avais été interpellé par les questions soulevées par « mon » étudiant, montrant la transformation annoncée de l'être humain vers un état hybride alliant anatomie et technique. Amaury concluait sur l'hypothèse d'une nouvelle étape dans l'évolution du vivant avec l'avènement d'une nouvelle espèce humaine qu'il appelait dans la ligne d'autres scientifiques : « *L'homme bionique* ».

Je m'étais promis de revenir sur le sujet pour l'approfondir quelque peu. Je vois dans la proposition d'aujourd'hui, une belle occasion d'honorer la promesse que je m'étais faite à l'époque.

L'ouvrage est sur le métier, il faut donc y aller.

Le lecteur qui attendrait un condensé des fondements, des axes et des contenus d'une éthique de l'intelligence artificielle sera nécessairement déçu.

Je choisis au contraire de limiter mon propos pour des développements courts qui sont davantage à considérer comme des ouvertures vers un processus de construction d'une éthique pour l'intelligence artificielle qui devrait se réaliser au fur et à mesure des avancées de celle-ci. Je considère en effet l'éthique comme le catalyseur de l'humain, la limite de l'inter - dit qui permet un dialogue en profondeur à la recherche de ce qui fonde l'humain et qui me fait être. L'éthique est ainsi toujours en actes. Une fois les fondamentaux posés, l'éthique doit rester en phase avec tout ce qui se passe dans l'existence humaine et poser les jalons nécessaires pour garantir l'humain. Elle se fait ainsi en partie en accompagnant les avancées humaines, un peu comme la locomotive avançait à la conquête de l'ouest à mesure que les rails étaient posés et fixés pour construire la voie ferrée.

Dans ce qui suit, je limiterai mon propos autour de trois axes : tout d'abord, le développement des sciences et des techniques, celui de l'homme bionique ensuite, et enfin la question de la morale de l'intelligence artificielle.

## **2. Pistes pour une démarche éthique**

### **2.1. Le développement des sciences et des techniques**

D'un point de vue éthique, le développement de l'intelligence artificielle s'inscrit pleinement dans celui des sciences et des techniques, et l'on ne peut que s'en réjouir.

Le développement des sciences et des techniques est éthique dès lors qu'il participe au développement du bien commun. Le bien commun se distingue de l'intérêt général et vise la mise en place de toutes les conditions matérielles, sociales, morales, culturelles, économiques, de formation et d'éducation qui permettent à chaque personne, considérée individuellement ou en tant que membre d'un groupe, d'une communauté ou d'une collectivité de devenir pleinement elle-même, de développer ses talents, de vivre ses vocations conformément à son identité spécifique d'être humain. L'intérêt général vise le maintien du système, le souci du bien commun se traduit par une attention personnalisée aux membres du système et à la qualité des relations qu'ils entretiennent entre eux.

Tous les humanistes devraient pouvoir se retrouver dans cette définition du bien commun et la colorer sur base des conceptions et des convictions de chacune et de chacun, dans le respect de la liberté de chacun, comprise comme cette capacité de devenir pleinement lui-même selon la spécificité de son être. Je vise ici la liberté intérieure de l'homme en lien étroit avec la libération de tout ce qui entrave cette liberté et enferme l'homme dans une dépendance de ses passions, de ses addictions, de ses pulsions. L'individu est invité, s'il le désire et s'il le peut, à se libérer de tout ce qui est brut en lui pour devenir un être raffiné, qui sait se tenir et se maîtriser.

Le bien commun favorise cette rencontre de chaque homme avec lui-même. La technique en est un instrument privilégié. Non seulement par le confort matériel qu'elle permet, mais aussi par le questionnement éthique qu'elle suscite.

La part de la technique dans la maîtrise de la nature n'est plus à démontrer. Avec la conséquence du passage progressif d'un état subi à un état construit : au rythme des avancées techniques, l'homme s'est de plus en plus et de mieux en mieux libéré des nécessités et des contraintes naturelles pour accéder à une vie plus confortable et plus agréable.

Habitation et vêtement pour échapper aux rudesses du climat par les techniques liées à l'architecture, au tannage, au travail des cuirs et des peaux. Poterie, cuisson, épices, boissons fermentées et alcoolisées, natte, tressage, tissage pour des litières, pour répondre aux besoins naturels de la faim, de la soif et du sommeil.

Par la technique encore, l'homme peut quitter son milieu naturel de la terre ferme pour évoluer dans le milieu aquatique et marin par la natation ou la navigation qui s'est développée des barques creusées dans un tronc d'arbre aux bateaux à voile et à moteurs et aux sous-marins à différentes propulsions. Évolutions similaires pour le déplacement dans les airs : de la montgolfière aux navettes spatiales, en passant par les parachutes, les avions à hélices ou à réaction.

Même évolution et même affranchissement des contraintes par la technique dans le domaine de la communication : du sémaphore à feux jusqu'aux « Iphones » et « smartphones », en passant par les téléphones classiques à cadran ou à touches, sans oublier les cabines téléphoniques que les plus anciens d'entre nous ont encore connues.

Dans chaque situation, les avancées vont dans le même sens d'une maîtrise de l'homme sur la nature grâce à la technique et d'un développement de celle-ci pour apporter une solution aux problèmes posés à l'homme par la nature.

Michel Onfray évoque d'une manière juste l'impact des sciences et des techniques sur la civilisation : « *L'histoire de l'humanité coïncide avec l'histoire des techniques. Des inventions suffisent ainsi parfois à déclencher de véritables révolutions de civilisation : le feu par exemple, et les techniques associées, la métallurgie, la fonderie, l'usage des métaux, donc les outils pour l'agriculture ou les armes pour la guerre ; la roue également, et la modification des distances avec l'invention de moyens de transports d'hommes, d'animaux, de biens, de richesses, de marchandises, d'alimentation, d'où le commerce ; puis le moteur, dont l'énergie rend possible les machines, donc l'industrie, les manufactures et le capitalisme, mais aussi les voitures, les camions, les trains, les avions ; l'électricité transforme également*

*la civilisation en permettant l'évolution des moteurs, certes, mais aussi en transfigurant le quotidien domestique : chauffage, éclairage, électroménager, radio et télévision. » [1].*

Les sciences et les techniques sont à la couture entre nature et culture. Elles font passer l'homme de l'une à l'autre, la culture étant la nature transformée, maîtrisée et signifiée. Le bien commun s'inscrit dans ce passage.

Il y a cependant des revers à la médaille : les trains déraillent, les avions se crashent, il n'y a pas de voitures sans accidents, ni de bateaux sans naufrages. Sans compter que la maîtrise de la nature s'est emballée sans maîtrise jusqu'à la maltraiter, la défigurer et la détruire. La planète est en péril et l'heure en est, de manière urgente, à la maîtrise de la maîtrise. La question de notre survie est à l'ordre du jour et à nouveau, sciences et techniques sont appelées à relever de nouveaux défis pour produire et consommer d'une manière éco-responsable. Sans compter les questions liées à la juste répartition des technologies issues du développement des sciences et des techniques.

L'intelligence artificielle s'inscrit pleinement dans l'évolution qui vient d'être rapidement retracée.

## **2.2. L'avènement de l'homme bionique**

Dans ce qui suit, nous nous inspirons en tout ou en partie de propos glanés sur le net ou dans l'actualité, et dont les références sont reprises plus loin.

L'informatique, le virtuel, l'intelligence artificielle sont l'aboutissement de l'évolution des sciences et des techniques qui vient d'être évoquée. Leur développement explose à des vitesses vertigineuses et nous fait entrer dans un changement radical de nos manières humaines de vivre dont nous sommes loin de mesurer les conséquences. Certains n'hésitent pas à parler de mutation, de changement de paradigme.

Les progrès dans le domaine sont fulgurants [2].

Tout débute avec l'informatique qui traite l'information de manière automatique au moyen de machines capables de traiter l'information dont il est question grâce à des algorithmes qui livrent une série d'instructions à suivre dans un certain ordre pour effectuer une tâche de manière automatique sans avoir besoin de comprendre ou de réfléchir. Ils agissent un peu comme une recette de cuisine. Cette série d'instructions est alors traduite dans un langage compréhensible par un ordinateur et constitue le programme informatique. Tout programme qui décrit un algorithme peut être exécuté parfaitement et très rapidement par un ordinateur.

Dès les années 50, les chercheurs se sont mis en quête d'algorithmes qui rendraient les ordinateurs aussi intelligents que l'être humain. Il y eut de belles avancées jusque dans les années 70, mais les résultats escomptés n'étaient pas au rendez-vous. Il fallut déchanter, ce qui provoqua la diminution des fonds disponibles pour la recherche dans ce domaine. Ce fut le premier hiver de l'intelligence artificielle.

Il y eut un nouvel engouement avec l'avènement des systèmes experts qui tentent de reproduire les raisonnements d'un expert comme un médecin par exemple. Le problème était qu'il fallait entrer les connaissances des hommes dans les machines, alors que le but était que les machines puissent apprendre toutes seules. Ce fut un second hiver pour l'intelligence artificielle.

Coup de tonnerre à la fin des années 90 : Deep Blue, le superordinateur d'IBM bat le champion du monde d'échec, Gary Kasparov. Et un coup de tonnerre pouvant en cacher un autre, en 2016, l'algorithme de Google bat un des meilleurs joueurs de Go, alors que cela n'était pas envisagé comme possible avant plusieurs années.

Il devient dès lors légitime de se poser la question de savoir si les ordinateurs sont intelligents ou pas, ce qui suppose de définir au préalable ce qu'est l'intelligence biologique pour définir en regard, celle de l'intelligence artificielle.

Les sciences cognitives étudient l'intelligence humaine et ont permis d'identifier les différents mécanismes qui sont à l'œuvre dans le fonctionnement de la pensée qui nous permet de percevoir, de parler, de bouger, de mémoriser, de raisonner, de planifier, de faire preuve d'abstraction, de créativité. Il s'agit des fonctions cognitives.

Pour simuler l'intelligence humaine, l'idée était de reproduire les fonctions cognitives. Les chercheurs ont privilégié deux approches qui sont complémentaires et qui sont à l'image de notre façon d'apprendre : soit par transmission de connaissances, soit par tâtonnement à partir de l'observation du monde et de l'expérimentation.

D'un côté, l'approche symbolique essaye de reproduire le processus cognitif de l'être humain en tentant d'imiter la manière dont il raisonne et utilise ses connaissances. La transmission des savoirs reposant sur un ensemble de savoirs, de règles et de procédures, il est possible d'expliquer la démarche suivie, étape par étape pour accéder au savoir. Sur ce modèle, un programme expliquant à la machine comment procéder étape par étape peut ainsi être écrit. La machine pourra ensuite l'exécuter beaucoup plus rapidement que n'importe qui et sans étourderie. Cette approche a eu le vent en poupe dans les années 80 et a donné naissance aux systèmes experts dont nous avons déjà parlé et qui, pour rappel, essayent de reproduire le raisonnement d'un expert comme un médecin ou un juriste.

D'un autre côté, l'approche numérique tente de reproduire comment l'être humain apprend par essai et erreur. Cette approche occupe le devant de la scène aujourd'hui et est plus complexe que l'approche symbolique : l'explication de comment nager ne remplace en rien l'expérience concrète de la nage. C'est en nageant que l'on sait nager !

Et là réside le défi technique... Car il s'agit de rendre l'ordinateur intelligent à partir de ce que le technicien lui injecte, ce qui n'est pas une mince affaire quand il s'agit d'expérience et de savoir-faire !

Il convient de garder à l'esprit la distinction entre l'intelligence biologique et l'intelligence artificielle.

Être intelligent pour un être vivant, c'est savoir répondre rapidement à une situation complexe, c'est trouver la réponse satisfaisante plutôt que la réponse optimale, c'est savoir trouver ce qui est important dans une information, c'est savoir s'adapter à une situation si les conditions changent ou réagir à une situation imprévue. Et tout cela en prenant appui sur une stratégie caractéristique du vivant qui mélange émotion, motivation, créativité. Sur ces terrains, l'intelligence du vivant dépasse l'intelligence artificielle. Sa force est la complémentarité de ses aspects émotionnels et logiques. Les intelligences artificielles au contraire sont très techniques et très spécifiques : elles ne peuvent accomplir que la tâche pour laquelle elles ont été programmées ou paramétrées : ainsi le programme qui bat le champion du monde d'échec, est mis en échec pour reconnaître un lapin ! En revanche, les intelligences artificielles sont ultraperformantes et souvent meilleures que les humains dans l'accomplissement d'une tâche précise. Sans doute est-il plus indiqué de parler d'intelligences artificielles en visant ainsi leur diversité en fonction des tâches à accomplir plutôt que d'intelligence artificielle qui serait mise en concurrence à l'intelligence du vivant. Convenons donc que parler d'intelligence artificielle, c'est considérer les multiples intelligences programmées pour accomplir des tâches spécifiques.

Jeu de mots assurément. Au singulier ou au pluriel, l'intelligence artificielle est répartie dans un développement aux allures exponentielles grâce à la puissance de calcul des ordinateurs ainsi qu'à l'optimisation des algorithmes. Les programmes de reconnaissance vocale, de reconnaissance d'image, ou ceux de traduction sont particulièrement performants grâce à l'apprentissage automatique. Dans les maisons ou dans la société, dans tous les domaines allant de la finance à l'agriculture, en passant par la médecine et l'éducation, l'intelligence artificielle est sortie des laboratoires et est présente. L'intelligence artificielle impacte notre quotidien au travers de l'utilisation des objets de notre quotidien : appareil photo, enceintes connectées, mécanisme de recommandation sur les réseaux sociaux. Nous baignons dans un monde d'algorithmes.

D'un point de vue éthique, le développement de l'intelligence artificielle s'inscrit dans le juste service du bien commun dès lors qu'elle participe à soulager la personne humaine de ce qui pourrait entraver sa liberté.

À titre d'exemple, le témoignage émouvant de cette artiste-peintre, amputée de trois membres à la suite d'un accident de la vie, et qui apprend de nouveau à dessiner grâce à un bras bionique. Les équipes de France Télévisions suivent Priscille Déborah depuis deux ans dans ses avancées. Le 23 décembre dernier, elle expliquait dans le Journal de 20 heures sur France 2 [3], comment elle s'exerce avec ténacité pour s'approprier son nouveau bras en le commandant grâce à la pensée. Pour cela, les chirurgiens ont dû réveiller plusieurs nerfs de son coude et les ont reconnectés à plusieurs muscles du bras. Ce sont leurs mouvements captés par des électrodes qui activent la prothèse. Il s'agit d'être extrêmement précis. La prothèse doit en effet coller le plus possible au muscle de la patiente. Le prothésiste qui suit Priscille Déborah depuis plusieurs années fabrique chaque mois de nouvelles emboîtures pour atteindre la précision nécessaire. Le prothésiste explique qu'un trop fort serrage avait provoqué un écart d'à peine 3 millimètres entre le bras et la prothèse. Le muscle ne pouvait pas travailler correctement lors de la contraction, ce qui provoquait des mouvements contradictoires. C'est dire la prouesse des équipes médicales pour réussir ce qu'elles ont accompli. Depuis mars 2020, Priscille Déborah porte réellement son bras bionique. Il lui reste désormais à apprendre à coordonner toujours mieux ses gestes et à maîtriser sa force, par exemple quand elle doit casser un œuf. L'artiste est dotée d'une force et d'un courage exemplaires. Elle annonce avec plein d'optimisme et avec un large sourire aux lèvres que si elle s'exerce chaque jour à dessiner, dans un an on peut se revoir pour sa prochaine exposition.

À titre d'exemple encore, le propos de Thierry Geerts, patron de Google Belgique, lors de la promotion de son dernier livre « *Homo digitalis* » indiquant que l'exactitude d'un diagnostic à la suite d'une mammographie est de quelque 33 % et monte à 98% dès lors qu'il est posé via l'assistance de l'intelligence artificielle : les algorithmes permettent en effet de regarder les clichés pixel par pixel et de détecter ainsi toute anomalie [4-5].

À titre d'exemple enfin, l'intelligence artificielle au service de l'embauche : l'intelligence artificielle se met à recruter. Il ne suffit donc plus de se faire remarquer par un recruteur ni de convaincre un employeur potentiel, il faut encore viser la validation de l'algorithme qui décortique tout ce qu'un candidat produit. Les uns diront le bénéfice de n'être jugé que sur les compétences, les autres objecteront que l'algorithme ne permet de trouver que ce que l'on souhaite trouver et écarte ainsi toute occasion pour une diversité de profils. Être recruté par des robots signifie la fin des entretiens classiques : ce sont des algorithmes qui classent et qui évaluent les candidats avec 25% de chances seulement d'être pris si l'on fait partie des mauvais can-

didats et 90 % si l'on fait partie de la bonne catégorie. L'intelligence artificielle analyse la voix, le discours, mais aussi les expressions du visage, un froncement de sourcils, un sourire ou un petit geste. Tout cela fait augmenter la note ou la baisser.

Les avis sont partagés sur la confiance à réserver aux machines ou au contraire à l'inquiétude à leur manifester. Les uns soulignent le caractère impersonnel d'un recrutement par la machine, indiquant l'absence du côté émotionnel et le manque de ressenti : « Il faut arrêter de vouloir remplacer les humains coûte que coûte ». Les autres s'accrochent au robot : « S'il est formaté et s'il pose les bonnes questions par rapport à l'emploi visé, pourquoi pas ! ».

L'algorithme scanne le CV, l'analyse, en extrait les données utiles et attribue une note au candidat en fonction des 6 millions de CV déjà analysés par la machine. Plus les compétences du candidat correspondent au profil des employés précédemment recrutés, plus le pourcentage attribué sera élevé. Une manière plus impartiale de traiter les candidatures selon certains, l'algorithme se consacrant uniquement sur les compétences, les qualités, les expériences, ce que le candidat sait faire ou non, indépendamment de la présentation formelle du CV. De facto, selon les partisans de l'intervention de la machine, le traitement des candidatures sera plus juste et plus efficace, avec un gain de temps considérable par la focalisation sur les seuls CV proches des besoins. Ces mêmes convaincus de l'efficacité de la machine attestent ne pas passer à côté de certains CV atypiques mais au demeurant intéressants, dès lors qu'ils gardent la main sur le choix des CV et qu'ils peuvent rencontrer physiquement les candidats sélectionnés. Reste que la première sélection est opérée par la machine.

Certains algorithmes vont plus loin : aux USA, les expressions du visage sont analysées à partir de courtes vidéos. Un signe peut trahir la personnalité, la voix et le discours sont scrutés par l'intelligence artificielle à partir de 200 critères allant de la richesse du langage à l'élocution. Ainsi, sur cinq minutes de vidéo, le candidat utilisant 800 mots différents est considéré comme maîtrisant mieux son sujet que celui qui en utilise seulement 150. Le but recherché est de mettre en avant des éléments factuels dans le processus de recrutement.

L'intelligence artificielle apparaît moins discriminante que les humains. C'est oublier que l'algorithme n'est jamais neutre. Il apprend sur des données qui reposent sur les choix de ceux qui l'ont construit. Si un algorithme est entraîné sur base de recrutements passés, et si ces recrutements ne sélectionnaient que des hommes, ce que l'algorithme va pouvoir observer, il va reproduire le biais existant et ne présenter que des CV rédigés par des hommes.

De plus, face à l'efficacité pseudo-neutre de la machine, certains candidats auraient trouvé la parade d'adopter volontairement le discours et les expressions préférées de la machine, de quoi pirater finalement les robots.



Les exemples qui viennent d'être présentés s'inscrivent dans la prodigieuse aventure de l'humanité qui est marquée par des progrès considérables. Ceux-ci sont loin de s'arrêter et l'on ne peut que s'en réjouir.

L'intelligence artificielle permet des avancées spectaculaires qui participent au bien-être et au vivre-mieux de l'humain. D'un point de vue éthique, il s'agit d'acquis indéniables et hautement appréciables. Mais il importe de pousser le questionnement plus loin et d'envisager quelle devrait être la morale de l'intelligence artificielle. Et là, je resterai fidèle à la posture d'un regardeur furtif qui pourra voir l'étendue de la tâche à accomplir mais qui n'en dressera que des contours qui devraient être affinés par une étude systématique qui dépasserait le cadre de cet éditorial.

### **2.3. La morale de l'intelligence artificielle**

Le développement des sciences et des techniques a de quoi donner le vertige.

Et il est loin d'arriver à sa limite. Le 21 janvier dernier, le Président Macron annonçait qu'une enveloppe de 1,8 milliards d'euros sur 5 ans serait réservée au quantique, un secteur qui devrait révolutionner l'informatique. L'ordinateur quantique sera une machine 100 millions de fois plus rapide que les ordinateurs actuels les plus puissants, ce qui révolutionnera les pratiques de recherche, de développement et de production. Les ordinateurs actuels reposent sur des micro-processeurs capables d'effectuer des milliards d'opérations par seconde à l'aide du langage binaire : mais la course à la performance se heurte aux limites de la miniaturisation. Alors que l'ordinateur quantique utilise la physique quantique en exploitant notamment la particularité qui veut qu'un atome peut être à différents endroits en même temps : les calculs peuvent ainsi se faire en parallèle, être superposés, ce qui multiplie la puissance de calcul de l'ordinateur... Le lecteur comprendra que j'apprécie n'être qu'un regardeur furtif !

La crainte est de voir science et technique dépasser l'homme au point de prendre le pas sur lui. L'homme perdrait la maîtrise. La machine deviendrait autonome, se développerait par elle-même et supplanterait l'homme en devenant plus intelligente que lui. Il y a débat dans le monde scientifique sur les options à prendre : les uns mettant en garde sur le péril qui menace l'humanité – la singularité technologique – par lequel le statut de l'homme dans la nature sera bouleversé ; les autres considèrent que la loi de Moore – les scientifiques la connaissent bien – ne se poursuit pas indéfiniment au point de traduire dans le registre technologique, la loi universelle de l'évolution conduisant du néant primitif à l'épanouissement de la spiritualité universelle. Ils prennent distance face à l'affirmation selon laquelle l'esprit humain s'affranchirait de son enveloppe charnelle et cérébrale et se téléchargerait « *sur une machine pour devenir souffle pur et simple 'pneuma'* » [6].

Est-il possible de trancher ? La machine restera-t-elle toujours en arrière de l'être humain ? L'intelligence artificielle ne va-t-elle pas donner à la machine une connaissance de son état et déclencher en conséquence des réactions adaptées ? Ce qui constituerait une forme spécifique de conscience pour la machine ?

En est-on si loin ? « *L'apprentissage automatique dote les machines d'une capacité à construire d'elles-mêmes des connaissances et à les utiliser pour se reconfigurer en réécrivant leurs propres programmes* » [6]. Les comportements des ordinateurs deviennent ainsi de plus en plus difficiles à anticiper. En outre, sur base des techniques d'apprentissages qui les rendent autonomes techniquement, les machines sont maintenant capables de rédiger des articles clairs et argumentés. En septembre dernier, « *The Guardian* » publiait une chronique écrite par un robot : « *Je ne veux pas détruire l'humanité* ».

Jean-Gabriel Ganascia rappelle que les résultats des machines tiennent essentiellement aux algorithmes d'apprentissage supervisé, ce qui suppose qu'un professeur instruit la machine en indiquant pour chaque instance, sa classe ou sa catégorie. Les machines « *restent ainsi soumises aux catégories et aux finalités imposées par ceux qui auront annoté les exemples utilisés durant la phase d'apprentissage* » [6]. Les machines ne sont davantage pas capables d'invention dans la mesure où elles « *ne modifient pas d'elles-mêmes le langage dans lequel s'expriment les observations qui alimentent leurs mécanismes d'apprentissage et les connaissances qu'elles construisent* » [6]. L'auteur de conclure que dans l'état actuel des avancées technologiques, rien ne permet d'affirmer que les ordinateurs pourraient se passer des hommes pour se développer, ni qu'ils puissent les dépasser et acquérir leur autonomie.

Jusqu'à preuve du contraire ! Mais serait-ce si grave ?

Je passe en revue rapidement les questions soulevées par ce qui vient d'être présenté.

***Tout d'abord***, l'homme bionique. Quelle merveille ! L'exemple du bras bionique est une prouesse qui permet à une personne de redevenir elle-même en retrouvant les chemins d'expression d'elle-même. La personne est un dedans qui a besoin d'un dehors disait Mounier. Et même si la nature de l'homme dépasse l'accidentel liés aux aléas de la vie, le corps reste le lieu et le moment de l'être que je suis. Il est pleinement éthique de développer la technique à son service.

De même, il s'avère que l'homme est loin d'utiliser toutes les potentialités de son cerveau. Si une modalité de l'intelligence artificielle – la fameuse puce dont parlent certains – permettait d'élargir l'utilisation du cerveau humain, ce serait une belle opportunité, à la réserve près de canaliser les aspirations de Prométhée. Je suis favorable à toute technologie qui reste au service de l'homme dans le respect de son identité profonde et avec la finalité de mener celle-ci jusqu'où elle peut aller. Épanouir l'homme reste une priorité éthique dès lors qu'il s'agit de le faire être selon sa

nature. Tant mieux si l'intelligence artificielle y participe dans les limites largement explicitées en bioéthique. Le corps de l'homme ne doit jamais devenir un prétexte à des expérimentations pour des fins qui ne seraient pas nobles et dont les horreurs liées aux années 40 résonnent toujours dans de nombreuses mémoires.

*Ensuite*, les fabuleuses performances de la machine. Toujours plus et pleinement acceptées de mon point de vue, si elles restent au service de l'homme. Et si, dans les conditions qui viennent d'être rappelées, la réalité dépassait la fiction – la révolution quantique est à nos portes – et que le monde des machines devenait une sphère autonome, l'homme devrait se situer par rapport à ce qu'il aurait participé à mettre en place. Et ce ne serait pas la première fois dans son histoire qu'il serait confronté à des forces qui le dépassent. Les forces de la nature en sont un bel exemple.

Reste que le questionnement éthique doit rester vigoureux dans la manière d'enseigner les machines. Les algorithmes reflètent les heurs et les malheurs de nos sociétés, les bonnes intentions comme les plus perverses. Les algorithmes sont porteurs des idéologies de ceux qui les alimentent. Ainsi, les machines peuvent être racistes, homophobes, sélectionner des personnes blanches ou de sexe masculin ou prôner des idéologies extrêmes. Pour rappel, les machines peuvent recruter, ou écrire des articles d'opinion sur base des contenus que les concepteurs leur font apprendre !

Il importe de rester très vigilants. Des Comités d'éthique doivent cadrer et recadrer. La course à la performance ne doit pas occulter les dérives. Le récent licenciement de Timnit Gebru par Google pose question. Il ne s'agit pas seulement d'alimenter la machine de données en quantité tous azimuts, il faut dans le même temps veiller à leur qualité. Qu'un(e) chercheur(se) fasse correctement son travail et signale les faiblesses d'une pratique – l'article qui ne cadrerait pas avec les critères de sélection établis par Google et faisait donc problème présentait une histoire des programmes automatique du langage naturel, un résumé des quatre principales faiblesses des grands modèles statistiques des langages ainsi que des pistes pour des travaux de recherches complémentaires – rappelle un cadre éthique et se fasse licencier... n'est pas éthique. Toute œuvre humaine se construit et progresse avec un mélange de combustible noble et de combustible ignoble. L'humain a toujours à se frayer un chemin dans la séparation entre l'ivraie et le bon grain. Des choses commencent à bouger timidement, puisque, suite au licenciement de Timnit Gebru, des employés de Google viennent de publier une lettre ouverte pour protester contre le licenciement de leur collègue en réclamant trois choses à leur Direction : une explication sur la censure de la publication de leur collègue ; plus de transparence vis-à-vis du grand public ; un engagement ferme de la Direction pour plus d'intégrité dans la recherche et plus de liberté académique afin que les principes de la philosophie de recherche chez Google ainsi que ceux de Google pour ce qui concerne les intelligences artificielles, soient respectés.

Cette lettre ouverte a été signée par 1604 employés sur un total de 100.000 que compte l'entreprise Google. C'est un petit pas, mais qui a le mérite d'exister. Il traduit la prise de conscience et la volonté qui lui est associée de s'opposer à toute forme de censure sur la recherche de l'éthique des algorithmes.

*Enfin*, la morale de la machine.

L'intelligence artificielle peut être considérée comme un super logiciel qui a appris à apprendre en vue de résoudre des tâches complexes à la place des humains.

Laura Wojcik [7], journaliste au Monde, se demande pourquoi l'intelligence artificielle a besoin d'éthique. Elle liste des situations montrant l'urgence d'un tel besoin.

Les cas où l'intelligence artificielle raisonne de manière trop simpliste et fonctionne de manière imprécise.

Les voitures autonomes doivent décrypter et analyser ce qui se passe sur l'autoroute. La réalité y est complexe et pleine d'imprévus. Des expériences ont montré que la machine était en difficulté pour repérer des stickers collés sur un panneau « Stop » et ne pas le confondre avec un panneau de limitation de la vitesse à 50 km/h. Qu'en sera-t-il quand ces voitures devront gérer des piétons imprévisibles, des cyclistes peu identifiables ou des dilemmes moraux épineux comme éviter un mur mais faucher des piétons sur le côté de la route, ou foncer sur le mur, épargner les piétons mais mettre en péril les occupants de la voiture ?

D'autres problèmes se posent quand l'intelligence artificielle apprend vite et obéit. En 2017, des chercheurs de l'Université de Stanford ont affirmé avoir appris à une intelligence artificielle à déchiffrer l'orientation sexuelle de personnes, et ce juste en scannant les traits de leur visage à partir de photos. Cela a provoqué beaucoup de remous, au point que les chercheurs ont indiqué que leur travail était une provocation pour montrer que la technologie allait beaucoup trop loin. Autre exemple angoissant : Tay, un robot intelligent de discussion de Microsoft a dû être désactivé en catastrophe. L'intelligence artificielle était censée devenir humaine en discutant avec d'autres internautes, mais certains lui ont montré comment détester l'humanité et même devenir nazi.

Ainsi, c'est au cœur même du fonctionnement de l'intelligence artificielle que se posent certaines limites éthiques : l'intelligence artificielle est aussi intelligente que les données qu'elle dévore pour réfléchir. Si la base de données de départ est biaisée et discriminante, le résultat donné par la machine sera biaisé et discriminant. Le fait que le logiciel fonctionne de manière opaque ne facilite pas les choses.

Laura Wojcik propose trois pistes pour réinjecter un peu de morale dans l'intelligence artificielle. *Tout d'abord*, en finir avec les boîtes noires des machines et créer des technologies qui soient lisibles et traçables en vue de démêler les bugs ou les résultats discriminants. Une science existe déjà pour déconstruire les raisonnements de la machine : il s'agit de la rétro-ingénierie. *Ensuite*, impliquer des sociologues, des anthropologues et des philosophes sur des projets scientifiques, mais aussi former ingénieurs et mathématiciens aux questions éthiques directement à l'université. C'est le cas depuis 2014 à l'Université de Stanford : « Initier nos étudiants aux questionnements sur l'éthique, les amener à questionner les fondements de leurs prises de décisions et de leurs actions avant qu'ils doivent être dans cette position en situation réelle où ils auront à prendre ces décisions en temps réel et peut-être sans le recul intellectuel dont ils pourraient bénéficier en classe, c'est ça notre ambition » (John Hennessy, Ancien Président de l'Université de Stanford). Plus récemment, en mars 2019, l'Université de Stanford a ouvert un centre de recherche dédié à l'intelligence artificielle humaine. Cela va dans le même sens que des dizaines d'autres initiatives d'autres Universités ou d'entreprises. *Enfin*, plancher sur le cadre légal pour encadrer les pratiques à l'échelle d'une entreprise, d'un pays ou d'un continent. Un panel de 52 experts travaillent actuellement à la Commission européenne sur un « Guide d'éthique en intelligence artificielle ». L'ambition est de faire des acteurs européens, des leaders en intelligence artificielle éthique et transparente.

### **3. Conclusion : Vers plus d'intelligence**

L'humain n'en est pas encore à vouloir spontanément le bien. Du moins pas tous les humains et pas tout le temps.

L'humain est en marche depuis longtemps, dans une histoire toujours fluctuante, toujours en train de se réinventer, de se refaire. Avec la volonté jamais éteinte de raconter ses grands fondements.

L'humain vit un moment historique : d'un côté et pour la première fois, l'humain se rend compte qu'il est allé trop loin. Il se rend compte qu'il a transformé la planète en sa scène, qu'il a supprimé des espèces, qu'il a transformé le paysage, qu'il est en train de modifier le climat. Cette conscience est totalement neuve. D'un autre côté, l'humain est appelé à prendre ses responsabilités et il dispose de moyens formidables au service de ses choix. L'intelligence artificielle en est un prodigieux. Il doit donc agir vite et bien. Il doit aller de l'avant.

Face aux enjeux, je suis pessimiste à la façon de Kant sur le destin de l'humain. L'humanité ne progresse qu'à travers des catastrophes. C'est le mal qui est le moteur du progrès et non la volonté de faire le bien. Dans sa « Philosophie politique », Kant explique que c'est parce que les guerres vont devenir mondiales, qu'il faudra un jour

créer une Société des Nations qui puisse tenter de régir les conflits. L'humain progresse pour éviter le pire, après qu'une catastrophe soit advenue.

C'est ce que Camus appelle « la ruse de la raison ».

Ainsi les catastrophes liées à la pandémie de la Covid, celles liées au climat vont sans doute nous apprendre des choses sur nous et nous faire réagir.

Je ne doute pas que les biais pris par l'intelligence artificielle vont nous pousser à négocier les tournants nécessaires pour garder le cap de notre histoire humaine qui ne peut devenir que toujours plus humaine.

Dr Phil. Roland URBAIN  
Enseignant d'HENALLUX-Pierrard

#### 4. Sources

Nous nous sommes inspirés en tout ou en partie des propos développés dans les références suivantes :

- [1] ONFRAY, M., *Antimanuel de Philosophie*, Rosny, Bréal, 2001, p.96.
- [2] PIXEES SCIENCEPARTICIPATIVE, (mis en ligne le 6 avril 2020), *C'est quoi l'intelligence artificielle ? Découvrir #1*.  
Adresse URL : <https://www.youtube.com/watch?v=yQLmgw3rCIM>
- [3] FRANCEINFO, (mis en ligne le 23 décembre 2020), *Le 20h de France 2 : journal télévisé du 23 décembre 2020 en replay*.  
Adresse URL : [https://www.francetvinfo.fr/replay-jt/france-2/20-heures/jt-de-20h-du-mercredi-23-decembre-2020\\_4205083.html](https://www.francetvinfo.fr/replay-jt/france-2/20-heures/jt-de-20h-du-mercredi-23-decembre-2020_4205083.html)
- [4] GEERTS, T., *Homo digitalis*, Bruxelles, Racine, 2021.
- [5] TAMIGNIAU, M., (mis en ligne le 26 février 2021), *Comment la digitalisation va changer le monde: "L'école est devenue un concept artificiel" – RTL Info*.  
Adresse URL : <https://www.rtl.be/info/belgique/societe/comment-la-digitalisation-va-changer-le-monde-l-ecole-est-devenue-un-concept-artificiel--1282366.aspx>
- [6] GANASCIA, J.-G., *Le mythe de la singularité : Faut-il craindre l'intelligence artificielle ?*, Paris, Seuil, 2017, pp. 40, 62-63, 67, 69.

- [7] WOJCIK, L., *Paris streets – Laura Wojcik, journalist and photojournalist based in Paris.*  
Adresse URL: <https://www.laura-wojcik.com>
- [8] FRANCE TELEVISIONS, (mis en ligne le 17 février 2021), *La Grande Librairie – Eric-Emmanuel Schmitt, Hédi Kaddour, Jennifer Richard et Eugène Ebodé en streaming – Replay France 5 | France tv.*  
Adresse URL : <https://www.france.tv/france-5/la-grande-librairie/la-grande-librairie-saison-13/2254567-emission-du-mercredi-17-fevrier-2021.html>