

Les robots parmi nous

Pour une éthique des machines

En italique : mes ajouts ou les termes en anglais .

L'auteur, **Johan Rochel**, de nationalité suisse, est docteur en droit et philosophe. Il est membre de Centre d'éthique de l'Université de Zurich. Il est également fondateur d'Ethix, un laboratoire pour l'éthique et l'innovation.

Rencontrer Asimo (Introduction)

Asimo, développé en 2000 par Honda, est généralement considéré comme le premier robot humanoïde, car se déplaçant comme les humains et reconnaissant des obstacles.

Son nom fait aussi référence à Issac Asimov, célèbre auteur de science-fiction dans les années 1950-1980 où les robots sont dotés d'un « cerveau positronique ». Dans « Le cycle des robots », saga de 6 romans, l'auteur débute par l'énoncé des « trois lois de la robotique » : Première loi : Un robot ne peut porter atteinte à un être humain ni, en restant passif, laisser cet être humain exposé au danger. Deuxième loi : Un robot doit obéir aux ordres donnés par des êtres humains, sauf si de tels ordres entre en contradiction avec la première loi. Troisième loi : Un robot doit protéger son existence dans la mesure où cette protection n'entre pas en contradiction avec la première ou la deuxième loi. Asimov les complètera par la « loi zéro » : Un robot ne peut pas faire de mal à l'humanité, ni, par son inaction, permettre que l'humanité soit blessée.

Films, livres, ... nourrissent notre imaginaire où les robots sont utilisés comme serviteurs ou comme défouloirs pour assouvir nos pulsions. Au début dociles les robots peuvent-ils, tel Terminator, Blade Runner, mettre en péril notre espèce ? Les promesses de l'ingénierie robotique sont immenses : robots industriels, tondeuse ou aspirateur « smart », drone livreur et peut être véhicule autonome. Mais les désillusions nous guettent. Ces défis majeurs ne sont pas simples à aborder et comment éviter soit la croyance naïve d'une cohabitation dans un monde meilleur, soit à l'inverse, comment ne pas sombrer dans le catastrophisme ? L'éthique de la robotique ne se limite pas à l'éthique de la science-fiction. Les enjeux sont personnels, politiques, économiques et n'épargne aucun secteur.

Qu'est-ce qu'un robot ?

Naissance du mot robot. Le mot robot apparaît pour la première fois en 1921 dans une pièce de théâtre du tchèque Karel Čapek où il signifie un travail, une corvée. La pièce décrit une société où des robots prennent le contrôle des humains. En arrière-fond se joue la métaphore du nouveau prolétariat industriel où l'humain n'est plus appelé qu'à ressembler à une machine.

Définition d'un robot ? Autrefois on parlait d'automate. *Ce nom a une autre acception aujourd'hui.* On peut citer le « canard » de Vaucanson en 1728, les automates d'Henri Jacquet-Droz en 1774, etc... À partir du milieu du XXe siècle est venu s'ajouter l'informatique (traitement de l'information au sens large), avec notamment en 1954 le robot industriel Unimate sur les chaînes de montage de General

Motors et en 1973 le premier robot bipède Wabot-1 au Japon. L'une des définitions la plus connue est celle du roboticien George Bekey : **Une machine située dans un monde qui sent, pense et agit.** Dans le concept de « robot » il y a une « machine », un « outil », un « instrument » et un « programme ». Ce concept est trop riche et nous devons affiner et faire des choix ; autrement-dit notre manière d'aborder la question structure divers récits de société.

Une machine ... Premièrement, le robot est désigné comme une machine, un artefact ayant une fonction utilitaire (autre qu'artistique). Il ne fait aucun doute qu'un robot peut s'inspirer du vivant (déplacement, articulations, préemption, ...) ; à terme il n'est pas exclu que des matériaux organiques soient utilisés. Un robot aura alors l'apparence du vivant mais il lui manquera encore d'autres dimensions : sa propre création, son libre-arbitre, ...

Faut-il partir de l'humain pour se demander où commence le robot ? Imaginons qu'il soit possible de remplacer tous mes organes et tous mes membres par des éléments artificiels ; serai-je toujours un humain ? Faut-il privilégier certains organes y compris le cerveau ?

D'où les cyborgs et le transhumanisme (l'humain hybride et augmenté). Plus extrêmes encore les post-humanistes pensent que l'humain (l'homo sapiens) doit être dépassé, débarrassé de ses faiblesses et de ses limitations. Le débat n'est plus de la science-fiction et se matérialise dans le biohacking (*charcutage biologique*) visant à expérimenter sur son propre corps pour dépasser la condition humaine imposée par la nature. Plus raisonnable est la médecine réparatrice (greffes d'organes, robotique portative, exosquelette) qui travaillent de longue date au rétablissement de certaines capacités défaillantes. La robotique portative pourrait peu à peu soulever des questions éthiques. Les relations machine-humain, puis à terme les interfaces machine-cerveau, rapprochent le champ de l'éthique des robots et celui de l'éthique du transhumanisme.

Une machine qui sent, ... Le robot doit premièrement être capable ressentir, au sens de percevoir l'environnement qui l'entoure (température, pression, luminosité, odeur, ...), de s'y localiser et de s'y déplacer (ex : le véhicule autonome). Cela implique des senseurs qui captent des données.

Une machine qui sent, pense. Deuxièmement, le robot doit être capable de penser au sens de traiter les informations reçues et de diagnostiquer. Selon le roboticien Alan Winfield il faut y ajouter : formuler un objectif. Attention à ne pas tomber dans le piège de l'IA comme pouvoir presque magique. Pour simplifier, un système IA est défini comme un ou plusieurs algorithmes qui exécutent des opérations mathématiques, permettant ainsi de réaliser un objectif. A ses débuts l'IA a fonctionné sur un modèle symbolique prenant la forme d'arbre de décision « si ..., alors ... », sur la base de critères et d'options prédéterminés. Aujourd'hui l'IA fonctionne sur le paradigme de l'apprentissage automatique et de l'entraînement (*deep learning*). L'apprentissage peut être « supervisée » comme la reconnaissance faciale ou « non supervisée » comme jouer au jeu d'échec. Dans ce dernier cas la machine cherche entre les données des liens, les regroupe, les classe, les associe. Le grand défi est la pertinence des résultats. Il appartient à l'équipe de programmation de contrôler leur qualité, notamment de distinguer entre corrélation et causalité. L'Unesco fait la distinction entre les robots « déterministes » (qui suivent des instructions préprogrammées) et robots « cognitifs » (qui apprennent par eux-mêmes). Mais le robot doit apprendre à apprendre et de là pense-t-il vraiment ? De récentes avancées, comme le Chat GPT, tentent de l'affirmer.

Une machine qui sent, pense et agit. Troisièmement, après avoir récolté et traité des informations, le robot doit décider et enfin agir.

Une machine située dans le monde réel. Ces capacités exigent une existence physique. Est-il pertinent de faire le distinguo entre robot et entité numérique ? Cette question divise les spécialistes.

Cet ouvrage fait le pari qu'il faut maintenir une distinction entre outils numériques (logiciels, réseau internet, data centers) et robots (entité physique). Par leur capacité à agir directement sur le monde physique qui les entoure - et donc sur les humains - les robots soulèvent des défis éthiques avec toujours plus d'intensité. En plus de ces dangers directs, leur présence physique implique une réflexion sur leurs composants pour les construire, sur leur maintenance, leur durabilité, leur recyclage et leur consommation d'énergie.

Je complète par des définitions des mondes plus larges, ceux de la recherche et du numérique :

Numérique : Domaine qui intègre les ordinateurs, les robots, les automates, l'IA, etc... Il traite l'information sous sa forme numérique. On parle parfois de révolution numérique (la quatrième industrielle ou **4.0**), de monde numérique ou de **cyberespace** pour nommer l'époque/espace actuel. Dans le « virtuel » les « ordis », l'internet, les « applis » font de nous des « internautes connectés ».

Automate (d'après wikipedia) : Dispositif reproduisant en autonomie une séquence d'actions prédéterminées et sans intervention humaine.

Robot (d'après wikipedia) : Dispositif mécatronique, conçu pour accomplir automatiquement des tâches d'imitation ou reproduisant des actions humaines. La **mécatronique** est la combinaison synergique et systémique de la mécanique, de l'électronique, de l'automatique et de l'informatique (voir en annexe le diagramme de Venn). Le concept d'**automate programmable** est voisin de celui de robot. Leur conception est l'objet de la **robotique**, pour les tâches domestiques on parle de **domotique**.

Intelligence artificielle (IA) et intelligence artificielle générative (IAg) (d'après Le Robert) : Ensemble des théories et des techniques développant des programmes informatiques complexes capables de simuler certains traits de l'intelligence humaine comme le raisonnement, l'apprentissage, L'intelligence artificielle générative, est capable, à partir de grands volumes de données (textes, sons, images...), de dégager des modèles et d'en générer de nouveaux, ou d'améliorer les modèles existants. Les dispositifs IA comprennent une partie hardware (matériel informatique) et software (logiciels). Les automates et les robots sont de plus en plus dotés d'IA.

Cobotique (d'après wikipedia) : Domaine transdisciplinaire de collaboration, directe ou téléopérée, entre l'homme et les **cobots** (les robots non autonomes). On distingue ici les robots sériels et les robots parallèles.

Anthrobot (article de Science et Avenir) : robot biologique créé à partir d'un amas de cellules et capable d'induire la réparation de tissus notamment neuronaux. Ces systèmes ne sont ni construits ni programmés ; on ne peut donc pas vraiment parler de robots. **Xénobot** : robot biologique capable de s'auto-reproduire.

Interface neuronale (brain computer interface) : ce qui connecte directement le cerveau humain à des machines.

*Mais revenons sur le point principalement développé dans ce livre : les **robots humanoïdes**.*

Une définition au cœur du débat. Le robot sera donc compris ici comme une machine située dans le monde réel qui sent, pense et agit. Cela exclut les automates et les outils exclusivement numériques (IA). La norme ISO 8373 classe les robots en trois catégories : - industriels, - de service, - médicaux. Cette définition ne doit cependant pas être prise pour une vérité éternelle mais évoluant dans le temps et dans l'espace. Dans la littérature scientifique on parle de « systèmes sociotechniques » s'inscrivant dans un contexte marqué notamment par les interactions avec les humains.

Certains aspects ne sont pas ou peu évoqués dans ce livre :

- la **convergence NBIC** (Nanotechnologies, Biotechnologies, technologies de l'Information, sciences Cognitives) dont l'un des « techno prophètes » les plus actifs est Elon Musk,
- **l'homme augmenté ou le transhumanisme**. Voir notamment le livre de Guy Vallancien, chirurgien et membre de l'Académie de médecine : *Homo artificialis, plaidoyer pour un humanisme numérique*.
- la **révolution numérique** implique des disponibilités et la gestion en énergie et en matières premières qui sont incompatibles à un développement illimité des robots. Voir J.M. Jancovici, Guillaume Pitron, Philippe Bihouix, Maxime Nechtschein, Stéphane Sarrade,

Autonomie et responsabilité du robot.

Pour se frayer un chemin dans les débats autour de l'autonomie, il est utile de distinguer entre un usage d'ingénierie et un usage philosophique et éthique.

L'autonomie au sens de l'ingénieur. L'autonomie porte sur le « comment » le robot réalise une tâche, ou encore, selon la norme ISO, sur « la capacité d'exécuter des tâches à partir de l'état courant et de détections, sans intervention humaine ». L'autonomie du robot augmente lorsqu'il peut déterminer seul comment il réalise certaines tâches. En l'état actuel le robot n'a pas la capacité de remettre en cause les tâches qui lui sont confiées. Les synergies entre robotique et IA permettent d'améliorer cette capacité et les robots sont de plus en plus autonomes.

L'autonomie au sens du philosophe. Les philosophes font souvent une distinction entre « autonomie du premier degré », celle de déterminer les moyens pour atteindre des objectifs préfixés et une « autonomie du deuxième degré », celle de définir les objectifs. La faculté de se fixer des objectifs est-elle réservée exclusivement aux êtres humains ? L'extension de cette idée d'autonomie au monde animal a souvent été mal traitée comme de l'instinct (ce qui s'impose naturellement à eux). Chez l'Homme l'autonomie a un point commun avec la capacité à s'autogouverner, à savoir choisir ses objectifs de vie, de les hiérarchiser et de les réviser, sur la manière de les réaliser, etc. Cette autonomie est caractérisée par la « conscience de soi » et le « libre arbitre » ; mais est-elle véritable ou, au contraire, n'est-elle qu'un leurre dans un monde parfaitement déterminé ? Une majorité de philosophes pensent que la liberté est possible dès lors que l'on choisit les buts. Une minorité pense que la liberté requiert un choix pur et inconditionné et nous n'aurions qu'une « illusion » de liberté individuelle. Pour la réflexion sur les robots cette question doit être laissée ouverte.

Les robots sont-ils moralement responsables ? Il faut faire la distinction entre la responsabilité causale (ce qui en est la cause) et la responsabilité morale (ce qui en porte la responsabilité). La responsabilité morale sera attribuée à un individu à la double condition qu'il ait connaissance des faits (le savoir) et qu'il ait capacité d'agir (le contrôle). En reculant avec ma voiture j'ai écrasé un massif de fleurs. Suis-je responsable ? Oui. D'une part je l'ai vu, ou pouvais le voir, dans mon rétroviseur ; la condition « savoir » est remplie et d'autre part je pouvais freiner, changer de direction ; la condition « contrôle » est remplie. La question devient plus complexe si on complète ou modifie certains paramètres : 1) Je ne l'ai pas vu car je n'ai pas regardé dans mon rétroviseur, ou encore je savais mais je n'ai pas agi par négligence. Je suis responsable. 2) L'accélérateur de ma voiture s'est bloqué. Je porte la responsabilité d'un manque d'entretien de la voiture ou le producteur de la voiture doit répondre d'une malfaçon. Quand serait-il avec une voiture autonome ? Pour y répondre la *Society of Automotive Engineers* (SAE) a fixé cinq niveaux pour la conduite autonome, avec à chaque niveau les responsabilités respectives entre le conducteur et le véhicule.

Une longue chaîne d'acteurs et de responsabilités. Le mot anglais *accountability* (en français : *devoir rendre compte en transparence*) traduit la dimension de la responsabilité : une personne et/ou une

entreprise doit être identifiée pour répondre des actes d'un robot. Ces responsabilités doivent être égrenées tout au long des étapes de sa conception, production, utilisation, contrôle, maintenance et destruction. La robotique lance deux défis : celui de la transparence des décisions et celui de la répartition des responsabilités.

Un robot peut-il (malgré tout) agir moralement ? Aujourd'hui les robots, y compris les plus perfectionnés, n'échappe pas à une programmation humaine. Mais demain, qu'en sera-t-il avec leur auto-apprentissage ? Devrait-on accepter de programmer des robots menteurs ou, au contraire, un robot devrait-il toujours respecter certains principes moraux fondamentaux ? À savoir : inculquer des règles morales au robot (*exemple : les lois d'Asimov*), viser une moralité fonctionnelle ou développer et imposer des codes de comportement restreints (*très technique mais au cœur des préoccupations des GAFAMS*).

Design et interaction

Par sa présence matérielle, un robot s'impose à son environnement et son aspect induit de notre part certains comportements. Pour ceux qui conçoivent les robots l'éthique du design est donc primordial.

Donner au robot une forme humaine. Pour un robot industriel l'aspect physique est sans importance. En revanche l'apparence physique d'un humain ou celle d'un animal domestique, c'est-à-dire humanoïde, est essentiel lorsqu'un robot doit reproduire des tâches humaines comme porter, se déplacer, parler, donner un ordre et à l'avenir exprimer des sentiments. Dans les années 1970 le roboticien Masahira Mori l'a montré avec son graphique (*voir annexe*) qui montre le degré d'affinité ou de familiarité envers un robot en fonction du degré de son apparence humaine ; introduisant notamment la « vallée de l'étrange », une zone où le robot s'exprime par la parole, parfois complexe, mais sans être accompagnée d'une expression du visage qui traduirait un sentiment.

Interaction et manipulation. A terme les responsables du design d'un robot devront se demander s'ils doivent créer l'impression (*l'illusion ?*) d'une authentique interaction avec un humain. Autrement dit un humain devra être en mesure de déterminer facilement s'il traite avec un autre humain ou avec une entité artificielle. L'objectif n'est pas d'entraver l'interaction mais de clarifier sa nature. Dès aujourd'hui cette préoccupation existe lors des échanges téléphoniques ou par internet avec des centrales d'appel (*call centers*). Les robots doivent être créés pour servir des intérêts humains et on doit pouvoir les identifier en tant que tels. Un exemple bien connu est celui de Paro, ressemblant à un phoque blanc, qui réagit aux caresses. Pas de problème lorsqu'il est reconnu comme un objet animé ; mais qu'en est-il avec un enfant ou une personne démente ? (*6000 exemples, dont 350 en France sont commercialisés comme « robot émotionnel »*).

Je « saute » le cas des poupées sexuelles.

Représentation et diversité. En s'invitant dans notre quotidien et devant gagner notre confiance, les robots doivent avoir une représentation visuelle, mais également une voix conforme aux attentes des utilisateurs. Il existe donc des parallèles entre ce monde virtuel et celui du marketing et de la publicité. Les concepteurs doivent prendre des décisions de conception et de fabrication pas seulement technique ; ils amènent avec eux leurs biais cognitifs en termes de perception, d'opinions et de préoccupation. Ce risque de dérive peut toutefois être une opportunité si les concepteurs ont la préoccupation et parviennent à créer un cercle vertueux de confiance.

L'innovation n'est jamais neutre. Au-delà de la configuration matérielle il n'est ni possible ni même souhaitable de s'esquiver sur les missions et les limites éthiques des robots dans les jeux d'argent, le sexe, la violence, etc.

Je fais une parenthèse sur un livre de Serge Tisseron, psychiatre, psychanalyste et chercheur avec son livre : Le jour où mon robot m'aimera. Vers une empathie artificielle. Il écrit : - dos du livre : "Ils sont déjà parmi nous mais nous ne les voyons pas. Ce sont tous nos objets connectés. ... Bientôt certains d'entre eux auront une apparence humaine, déchiffreront nos émotions, nous parleront. Ils pourront même nous manifester de l'affection, voire de l'amour. Ce sera « pour de faux » ? Et, alors ! ... Avons-nous envie de ce bonheur-là ? Comment réagissons-nous au fait d'être surveillés en permanence ? Et que deviendra la relation de l'homme à ses semblables lorsque nous serons entourés de créatures humanoïdes (ou pas), programmés pour tout savoir de nous, anticiper nos désirs ? ..." et - page 65 : "Le problème est moins celui de savoir si le robot aura un jour une « vraie » intelligence ou de « vraies » émotions, que celui de comprendre pourquoi nous serions enclins à lui en attribuer."

Les robots dans le droit

En matière de nouvelles technologies le droit est souvent perçu comme étant perpétuellement en retard. Le droit et la technologie ressemblent aux courses poursuites entre Coyote et Bip Bip. Coyote (le droit) tente en vain de rattraper Bip Bip (la technologie), mais n'y parvient jamais, tant celle-ci est rapide et imprévisible. Ce récit donne une place de choix à l'éthique pour combler le vide entre eux. À cette lecture, inspirée de Bip Bip et Coyote, s'oppose une lecture interactive : l'innovation ne peut pas se dérouler en dehors d'un cadre juridique.

À quel catégorie juridique le robot appartient-il ? En l'état actuel, le droit considère le robot comme une chose. Comme toute chose il a une valeur marchande, peut être protégée par des brevets ou autres droits de propriété intellectuelle, mais n'a pas de statut juridique particulier. Comme il existe le statut de « personne morale » nous pourrions créer le statut de « personne artificielle » et le débat ressemblerait à celui des droits des animaux. Mais alors : Pourquoi ferions-nous la distinction en vivant et non-vivant ? Pour quels raisons le robot devrait-il être protégés ? Ou encore, faudrait-il utiliser diverses catégories d'humains pour distinguer les capacités de discernement ?

Le robot dans le droit privé. En matière de droit privé le robot est un objet qui a d'une part un concepteur/fabricant et d'autre part un utilisateur. Un robot simple ne pose pas de problème de responsabilité pour une cause identifiée. De même pour un robot doté d'un auto-développement, la responsabilité revient au concepteur ou à l'utilisateur qui a agi par négligence, détournement d'usage ou autres. Mais pour les robots encore plus perfectionnés (avec IA et plus demain), comment déterminer les responsabilités ? Le cas type du véhicule autonome représente un défi majeur pour la sécurité routière. Encore plus complexe, la collecte et le traitement de données privées (accès, conservation, cybersécurité, destruction) relèvent de la sécurité publique.

Le Parlement européen a adopté le 12 mars 2024 l' « AI act », la première réglementation au monde sur l'éthique et l'usage de l'IA ; cependant sa mise en application n'est pas attendue avant 2026.

Le robot dans le droit pénal. Aujourd'hui le robot peut être considéré comme un objet, ou comme le moyen d'un délit, et on juge l'humain ou la personne morale présente derrière le robot. Demain l'usage de systèmes autonomes posera le problème du contrôle humain. La légitimité de tels systèmes reposera (souvent pas toujours) sur la présence et le contrôle d'un humain capable de maîtriser le processus d'utilisation. Encore faut-il s'assurer de contrôler cette capacité. Il y a trois niveaux de contrôle : 1) L'humain est partie intégrante du processus décisionnel (*human in the loop*) et prend la décision finale. 2) L'humain est en situation de contrôler le système (*human on the loop*) sans intervenir dans le processus décisionnel ou uniquement en cas de problème. 3) L'humain est

complètement en dehors de processus décisionnel, sans capacité d'agir (*human out the loop*). La responsabilité des équipes ayant conçues et programmées le système est engagée.

Les **SALA** (*système d'arme létale autonome*), dont les **drones tueurs ou kamikazes**. Ils révolutionnent aujourd'hui les pratiques dans les conflits armés actuels. Les drones tueurs, isolés ou en groupe, peuvent de façon autonome : décoller, se positionner sur une cible ennemie, ne peuvent pas déclencher un tir létal sans accord humain (interdiction de la CCAC (Convention de Genève)), s'autoprotéger et enfin revenir à leur base. En date du 29 mars 2019, une majorité des gouvernements représentés à l'ONU se sont prononcés sur l'interdiction de tuer des SALA ; mais une minorité, notamment l'Australie, Israël, la Russie, le Royaume-Uni et les États-Unis, s'y sont opposés (wikipedia). En France, les leaders sont Nexter Robotics (groupe franco-allemand KNDS) et Thales. Citons également les robots poseurs de mines antipersonnels.

Les défis de justice sociale

Le robot, ou plus précisément la robotique, a des impacts sur la société, autrement dit induit des défis de justice sociale.

Une menace pour l'emploi ? Côté pile, les robots libèrent des tâches dangereuses, répétitives ou ingrates. Côté face, ces mêmes robots violent et volent des emplois. Les domaines d'activités les plus touchés ont d'abord été l'agriculture, puis l'industrie et la mobilité. Avec l'arrivée de l'IA et des cobots le tertiaire va encore être plus impacté ; dès aujourd'hui l'administration, le commerce et demain, le conseil, la santé, les services à la personne (tourisme), l'enseignement, la création artistique, Certaines régions, notamment celles où le secteur industriel est fortement implanté, sont plus sévèrement touchées. *Plusieurs études évoquent la disparition à moyen terme de la moitié des emplois actuels*. Par ailleurs les personnes qui collaboreront avec les cobots devront redéfinir leurs tâches.

Travail et société. En Europe occidentale, les quatre relations sociales indissociables sont : l'emploi, le travail productif, les moyens de subsistance et la reconnaissance sociale. Elles ne peuvent être dissociées l'une de l'autre ; or elles sont toutes remises en cause par les progrès de l'IA et de la robotique. Ce quatuor peut donc rapidement dérailler et au moins trois ensembles de mesures doivent nourrir la transition vers des futurs souhaitables :

- Adapter la formation obligatoire à l'école mais aussi l'apprentissage tout au long de la vie avec validation des expériences et des acquis.
- Préparer et mettre en place une réponse de politique sociale aux changements du marché du travail : transition responsable (anticipation des changements, plans sociaux accompagnés) ; transition sectorielle (public, privé) ; changements de la taxation du travail salarié et de celle des entreprises ; nouvelle taxation pour les robots (*très difficile à définir et à mettre en place*).
- Mener une discussion sur les domaines qui, comme l'aide à la personne, doivent échapper à la domotique et créer à ce titre le droit « au contact humain » .

Interaction ou relation ? Il est important de traiter les interactions entre les personnes humaines et les robots. Nous cherchons un coupable à notre colère quand nous parlons à notre voiture qui ne veut pas démarrer ou encore nous insultons notre ordinateur qui ne « mouline » pas assez vite. Le fait d'utiliser de tels objets a pour effet de nous hybrider à lui (Serge Tisseron). Par ce comportement, inconsciemment nous donnons à la machine un rôle dévolu jusqu'alors aux humains. Pour certains critiques l'attitude permettant aux robots d'acquérir des compétences semblables à celles des humains, est porteuse d'un grave risque de manipulation. Ce danger existe particulièrement avec

l'addiction aux réseaux sociaux, aux jeux interactifs en ligne et aux robots sexuels (le premier exemple, non informatique, a été le syndrome de la poupée Barbie). Dans le domaine du soin à la personne (médical ou esthétique) les spécialistes font la distinction entre soin « froid » (*cold care*) dispensé par la machine et soin « profond » (*deep care*) pratiqué par l'humain.

Les grands récits de société

Les grands récits de société représentent à la fois le périmètre général de nos débats et notre histoire collective. Ils peuvent être définis comme des cadrages (*framings*) ; certains aspects de la réalité font partie d'un cadre, d'autres sont en dehors. Pour appréhender le monde qui nous entoure nous sélectionnons notamment par des narrations sous forme de textes, d'images, de spectacles, Autrement dit : mon regard sur le réel dépend de la société dans laquelle j'évolue. Au niveau collectif on parle de « grands récits », comme les mythes, les contes, les légendes et maintenant la science-fiction du numérique. Trois problématiques structurent ces récits :

Les robots et les humains compétition ou coopération ? Dans un récit de compétition, les humains sont comparés aux robots sous l'angle de leur rapport de force où les robots menacent tôt ou tard de contrôler puis de soumettre les humains. Le terme d'IA porte cette idée. Dans un récit de coopération les robots sont au service des humains et le terme « cobots » les désigne. Les termes ne sont pas neutres. Le choix de ces récits n'est pas seulement un défi citoyen mais un enjeu économique et politique. Dans les cultures où l'humain est divinisé, les robots sont perçus avec suspicion et l'objet d'ingénieurs qui pensent pouvoir créer une vie artificielle allant jusqu'à dépasser l'humain.

Qu'est-ce qui fait la spécificité de l'homme ? Les jeux d'échecs et de go ont longtemps été présentés comme l'apanage de l'humain, mêlant ingéniosité et sens tactique. L'IA a remporté la partie contre Kasparov qui malgré tout considère qu'au contraire ce fut une victoire : Celle de l'homme plus la machine. Le robot fonctionne pour atteindre un objectif ; en revanche et en deuxième niveau, l'humain a la capacité de réviser ses valeurs et ses objectifs. Reste que l'immense défi de la liberté humaine s'articule, au niveau mondial, sur le couplage avec les données informatiques (*Big data*) et les processus de catégorisation. L'arrivée de la robotique nous offre là l'occasion de nous inscrire dans une forme de continuum semblable à un dégradé de couleurs.

Quel est le sens de la vie de l'homme ? Une société bâtie sur la performance ne peut que donner leur place à ces entités artificielles. Il nous faut cependant redéfinir en profondeur l'idée même de performance avec des critères qualitatifs (adaptation, créativité, inventivité), plutôt que quantitatifs (force, vitesse, puissance). La robotisation nous force à mieux imaginer, là où le travail humain rénuméré perd en importance et à nous confronter à un sens au-delà.

La vie éternelle ? Et si demain une enveloppe robotique devenait le lieu et l'outil de notre vie physique ? Cette question nous renvoie à la science-fiction (*ou divagations comme celle d'Elon Musk*). Par ailleurs elle entre en dialogue avec certaines traditions religieuses ou spirituelles où l'esprit se réincarne dans un nouveau corps physique.

Deux scénarios pour préparer l'avenir.

Deux grands scénarios émergent.

Le premier scénario se construit à partir des technologies disponibles aujourd'hui ou aisément concevables demain. Le robot se perfectionne, gagne en autonomie et en capacité d'action, mais reste ici un outil. Il s'invite dans de nombreux secteurs économiques et fait partie intégrante de notre quotidien ; cependant il garde sa pertinence, à savoir être au service des intérêts humains, sans révolution éthique ou juridique. Les principaux défis porteront ici sur la régulation et la

hiérarchisation des responsabilités, sur les politiques sociales de la transition dans les secteurs économiques touchés et sur les récits de société où le sens de la vie ne se résume pas à la seule valeur travail. En bref un scénario de continuité qui prend acte des progrès de la science.

Le second scénario s'inscrit dans un développement à plus long terme, encore largement porté par la science-fiction (*dystopies*). Le robot acquiert des capacités jusqu'alors typiquement humaines. Au-delà d'être autonome le robot développe la capacité de prendre conscience de son existence et de modifier ses objectifs en fonction de certaines de ses valeurs et de ses convictions propres. En d'autres mots le robot devient semblable à l'humain, un quasi-humain, *une forme modernisée de l'esclave*. Le robot devrait alors être traité avec respect et bienveillance. *Acceptera-t-on, demain, de lui confier une prise de décision, pouvant sauver ou mettre en péril des vies humaines ?* (Patrice Caine, PDG de Thales).

Afin de penser et nourrir un entre-deux scénario, on pourrait s'inspirer des débats autour de l'éthique des animaux. Mais la pertinence de ces scénarios est peut-être ailleurs : il nous tend un miroir où se reflètent nos prétendues spécificités, nos forces et nos faiblesses, nos ambitions et nos erreurs.

Pour aller plus loin, un livre que je n'ai pas lu : « Une éthique du numérique ». L'ouvrage rassemble des textes d'auteurs et autrices de spécialités académiques variées qui ont en commun de faire partie du Comité national pilote d'éthique du numérique (CNPEN) ou d'y être invités.

Et une réflexion de Raphaël Enthoven, philosophe et écrivain, dans son livre « L'esprit artificiel » (2024) : « Reste que le fait de réussir à mimer l'intelligence humaine ne suffit pas à prouver qu'on est intelligent. Dès lors, devons-nous considérer – ou pas – que nos cerveaux possèdent le monopole de ce qu'on appelle la pensée ? » et « Pourquoi continuons-nous de croire que nos créations seront un jour des créateurs ».... C'est un autre débat.

Pour terminer un dicton : « Ils ne savaient pas que c'était impossible alors ils l'ont fait » (Mark Twain) ».

Annexe : Diagramme de Venn

Graphe de Masahira Mori