



Intelligence artificielle et perpétuation de l'espèce **Partie I – sens de la vie bonne et éthique de responsabilité**

Les humains se sont convertis en démiurges. Après avoir façonné **la nature** selon leurs besoins et parfois leurs caprices, après avoir commencé à modifier **le vivant**, ils dotent d'intelligence **des objets** inanimés. Ils ne programment plus ces machines. Elles apprennent d'elles-mêmes, à partir des millions de données qu'ils leur ingurgitent.

Le Dernier Homme ne sera pas un champion de la croissance économique, trader ou patron de start-up à succès, gagnant toujours plus pour dépenser tout autant. La fin de l'histoire et du politique, annoncée par Francis Fukuyama comme le triomphe de la démocratie libérale, pourrait être remise en cause. Le choc des civilisations, à supposer que la prédiction de Samuel Huntington se réalise, n'aura jamais l'importance du grand bouleversement technologique et sociétal en cours.

L'Intelligence Artificielle (IA) dite faible, qui a pris son essor au début de cette décennie, **interroge notre rapport au travail**. Elle menace d'ores et déjà de disparition ou, à tout le moins, de changements profonds des professions aussi nombreuses et variées, que les caissières, les routiers, les comptables, les traducteurs, les avocats et les médecins. Des millions de personnes, dont beaucoup sans qualification, devront retrouver un emploi, qui n'entre pas en concurrence avec ces machines.

L'homme pourrait aussi être confronté, dans quelques dizaines d'années à peine, à la **singularité**, dite aussi **intelligence artificielle forte**. Ces machines surpasseraient les plus brillants des humains dans toutes les tâches. La seule éventualité de ces entités semblables à nous, pose la question de **la spécificité de la nature humaine**. Elle interroge aussi **notre rapport à un autre** qui, tout en nous ressemblant, serait **capable de nous détruire**. Car cette supériorité pourrait s'amplifier au cours du temps pour atteindre rapidement une différence abyssale, compte tenu de l'augmentation constante de la puissance de calcul des ordinateurs.

D'aucuns envisagent le salut de notre espèce à travers **le transhumanisme**. L'homme serait **augmenté** par des technologies intégrées dans son corps, telles que des puces électroniques, des nanotechnologies, des membres bioniques, des connexions neuronales. Cette transformation physique interroge **notre relation au corps**. Elle nous pousse à nous positionner **par rapport aux objets et au reste du monde vivant**. Certains rêvent d'un **transhumanisme cognitif**, à savoir un cerveau humain interfacé directement avec la machine.

Dans des scénarios **posthumanistes**, la mémoire humaine deviendrait téléchargeable sur des supports informatiques. Sa conscience, sa mémoire et son intelligence ayant été numérisées, **l'homme**, affranchi de son corps, **ne connaîtrait plus la mort**. La copie informatique permettrait **une reproduction** à l'identique et à l'infini, **sans sexualité**. Cette nouvelle condition interroge **le sens de la vie**. Or « *la vie n'a de sens que si elle se sait finie et s'éprouve dans la finitude. Désirer une vie immortelle est un désir contradictoire, puisque c'est désirer la mort du désir.* » (L. Cournarie, *Monde, théâtre et transhumanisme*, NXU, 2018).



Ce qui relevait du mythe du golem ou des nouvelles de science-fiction d'Isaac Asimov prendrait-il forme ? La nouveauté et la profondeur des questions posées, la rapidité et l'ampleur des changements à l'œuvre requièrent de façon urgente des réponses philosophiques et éthiques, et les actes qui en découlent.

Les philosophes sont confrontés à une situation inédite. Jamais la formule de Socrate, « je ne sais qu'une chose, c'est que je ne sais rien », n'a été plus pertinente. Ainsi, une enquête réalisée en 2015 et 2016 par des chercheurs des universités d'Oxford et Yale auprès de 352 experts en IA et publiée mi-2017, souligne l'absence de consensus sur l'intelligence artificielle forte. Certains envisagent une arrivée de la singularité dans 20 ou 30 ans quand d'autres partent sur des hypothèses très conservatrices, atteignant les 200 ans, la moyenne étant de 120 ans. Il paraît difficile de conclure autre chose qu'**une probabilité significative (supérieure à 50%) de l'arrivée de la singularité... un jour.**

La mise au point d'un exocortex, un second cerveau cybernétique, utilisé afin d'y télécharger la mémoire biologique, semble autrement hypothétique. D'après le spécialiste des neurosciences Antonio Damasio, « *cette perspective demeure à ce jour peu plausible. Ces projets traduisent une vision limitée de ce qui constitue réellement la vie ; leurs promoteurs ne comprennent pas comment se produisent nos expériences mentales. [...] L'esprit est le produit d'interactions entre le corps et le cerveau. Il ne s'agit pas d'un phénomène purement cérébral.* »¹

Quant à l'intelligence artificielle faible, on découvre tous les mois de nouvelles innovations. Depuis quelque temps, des agents conversationnels répondent avec pertinence à la plupart des questions qui leur sont posées ; récemment, Amazon lançait son modèle de supermarché sans humain ; très bientôt, des camions autonomes rouleront sur les routes nord-américaines. Le rythme d'arrivée de ces trouvailles semble s'accélérer, sans pour autant que la fiabilité des prévisions s'améliore. Les conséquences économiques et sociales de cette mutation inédite des métiers et du travail font l'objet de larges supputations.

Le nombre d'emplois concernés s'annonce sans équivalent depuis la mutation de la mécanisation qu'a connue le monde agricole. En un siècle, les paysans étaient passés de plus de la moitié à moins de 3% de la population dans nos contrées ; ce changement, qui clôt une période remontant au néolithique, conduit Michel Serres à le considérer comme l'un des événements majeurs de l'histoire de l'humanité.² L'impact de l'IA faible, bien qu'inférieur, devrait être colossal. Foxconn, le principal sous-traitant d'Apple et de Samsung, a entamé en 2016 un large mouvement pour remplacer ses 1,3 millions d'ouvriers, en grande majorité chinois, par des robots³. Surtout, **les changements induits pourraient se réaliser en quelques années à peine.**

¹ A. Damasio, *L'ordre étrange des choses. La vie, les sentiments et la fabrique de la culture*, Odile Jacob, 2017.

² M. Serres, *Le temps des crises*, Le Pommier, coll. « Manifestes », Paris, 2009.

³ "[Foxconn replaces '60,000 factory workers with robots' - BBC News](#)". *BBC News*. Extrait le 06/03/2018.



Les conséquences psychologiques et sociologiques d'une cohabitation inédite avec ces entités anthropomorphes interroge et inquiète. Le comportement de nombre d'adolescents et d'adultes vis-à-vis des écrans pose déjà problème. Isolés dans un monde virtuel, nombre d'entre eux pourraient prolonger l'expérience dans le réel et favoriser la compagnie des machines à celle des hommes. Ces nouveaux rapports aux objets soulèvent également des interrogations métaphysiques, susceptibles de perturber les milliards de croyants à travers le monde. Des troubles sociaux ou géopolitiques ne sont pas à exclure.

Dès lors, comment penser notre relation à des entités, les intelligences artificielles, dont on ne sait pas jusqu'à quel point elles modifieront notre mode de vie ? si elles nous surpasseront ? Les philosophes n'ont pas le choix. Ce flou général doit s'intégrer à leur réflexion. Etant entendu que cette acception de philosophe englobe tout citoyen souhaitant participer au débat dans un esprit positif et bienveillant.

L'enjeu est considérable. Des personnalités respectées comme Elon Musk, Stephen Hawking ou Bill Gates, parlent de **menaces à brève échéance sur l'existence de l'humanité**. Plus que jamais, aucun scénario ne devra être laissé de côté, sans arguments objectifs. Il sera préférable de prêter attention aux hypothèses dites pessimistes, pour déterminer les axes de correction possibles, le cas échéant.

Aujourd'hui, les GAFAMI⁴ consacrent des budgets de recherche très importants à l'éthique de l'Intelligence Artificielle. Aucun de ces géants ne saurait renoncer aux formidables promesses de gains de l'IA, sans des pressions fortes du régulateur, de leurs clients ou de leurs actionnaires. Les avis d'entrepreneurs reconnus des NBIC, y compris lorsque certains alertent des dangers de l'Intelligence Artificielle, ne peuvent suffire. La régulation de l'IA ne doit pas reposer sur des seuls acteurs qui sont à la fois juge et partie. Des ONG, des think tank comme NXU, des intellectuels, des simples citoyens doivent se réapproprier le débat.

La gestion, la gouvernance et la régulation de l'intelligence artificielle sont **un enjeu démocratique**.

Interrogée sur sa propre nature, interpellée par la possibilité d'un autre à la fois si proche (son engeance) et si différent (non-vivant), susceptible de la surpasser et de la détruire, l'humanité ne peut répondre à ces questions et à beaucoup d'autres sans **savoir où elle va**.

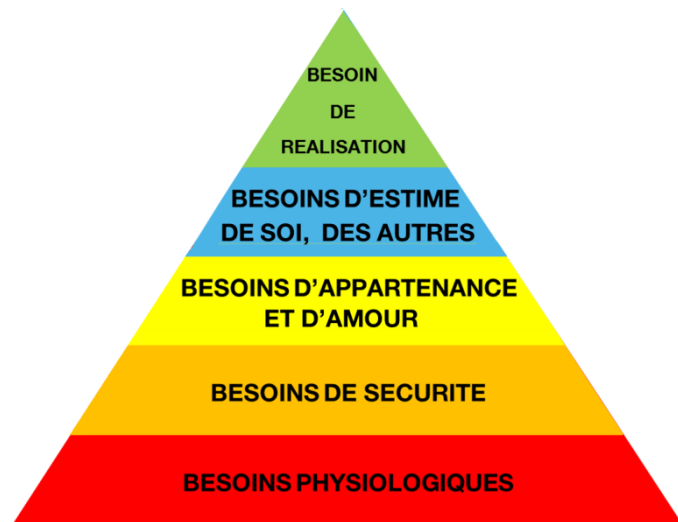
Or les philosophes ont de tout temps recherché **le sens de la vie bonne**. Les plus connues de ces interprétations sont le principe cosmologique (se réconcilier avec le Cosmos), le principe théologique (suivre les commandements divins), le principe humaniste (le « cogito » posé comme fondement de la pensée), le principe de la déconstruction (le soupçon vis-à-vis des illusions métaphysiques), auxquels Luc Ferry a ajouté le principe de l'amour (la période contemporaine du mariage d'amour). D'autres argueront que la quête du bonheur, théorisée dans *L'éthique à Nicomaque* d'Aristote ou élevée au rang de droit fondamental dans la Déclaration d'indépendance

⁴ Google, Apple, Facebook, Amazon, Microsoft, IBM. Microsoft et IBM sont aussi des acteurs très importants de l'IA.

américaine (« *tous les hommes sont créés égaux ; ils sont doués par leur Créateur de certains droits inaliénables ; parmi ces droits se trouvent la vie, la liberté et la recherche du bonheur⁵* »), est un but en soi. De plus, nous ne parlons ici que d'une vision occidentale. Imaginer un sens de la vie bonne commun à tous les hommes ressemble à une gageure. Personne n'empêchera quelqu'un de penser différemment de ses congénères. Le contraire reviendrait à nier la liberté d'opinion. Il y a, malgré tout, une condition : ce choix de vie ne doit pas menacer les autres humains dans leur liberté, encore moins dans leur intégrité.

Pourtant, **nous partageons, depuis notre naissance, un même sens de la vie, indépendant des modes et des cultures, inscrit dans nos gènes.**

Penchons-nous un instant sur la pyramide des besoins de l'homme réalisée par le psychologue Abraham Maslow. Construite à partir de l'étude de cas cliniques par le praticien et présentée en 1943 dans *A Theory of Human Motivation*, elle schématise et hiérarchise les motivations profondes de l'homme. Elle est basée sur l'idée qu'un besoin humain fondamental s'exprime



Pyramide des besoins de l'homme (A. Maslow)

derrière tout désir. Le marketing et les forces de vente, qui l'utilisent abondamment, ont démontré depuis longtemps sa pertinence et son efficacité.

Le premier niveau de la pyramide englobe les besoins physiologiques à satisfaire, tels que la respiration, la soif, la faim, les déjections, le sommeil, le maintien de la température corporelle, la reproduction. La sécurité occupe le deuxième niveau : l'homme doit se protéger de toute agression, y compris de ses congénères. Les trois niveaux suivants regroupent le besoin d'aimer et d'être aimé (troisième), le besoin de reconnaissance sociale (quatrième) et le besoin de donner un sens à sa vie (cinquième et dernier). Dès lors, on comprend que **le socle formé par les besoins des quatre premiers niveaux de la pyramide constitue un besoin commun à tous les humains.** Cette nécessité de survivre et de se reproduire, tout en interagissant avec nos semblables, porte un nom : **assurer la perpétuation de l'espèce.**

Le cinquième niveau en appelle à **un besoin de réalisation personnelle, un sens de la vie, propre à chaque individu,** marqué par la culture et l'histoire personnelle de chacun. Cependant, Maslow établit une hiérarchie des besoins selon les niveaux en commençant par le premier. Le besoin de perpétuation de l'espèce est donc prioritaire par rapport à la quête de sens individuel.

⁵ Wikisource. Traduction de Thomas Jefferson. Extrait le 17/03/2018.



Par le passé, la question de la perpétuation de l'espèce n'a jamais pris de véritable sens, car **elle ne se posait pas**⁶. Les technologies associées n'existaient pas encore (le nucléaire, les biotechnologies, l'IA) ou n'étaient pas suffisamment développées pour constituer un danger global (les émissions de gaz à effet de serre). En outre, l'homme ne disposait d'aucun moyen de prévenir des menaces naturelles, comme la chute d'une météorite géante ou une énorme éruption volcanique. Notre époque contemporaine a vu apparaître des risques d'origine humaine, mettant en danger la survie de notre espèce. Le *Centre for Study for Existencial Risk* de l'université de Cambridge a recensé quatre menaces majeures⁷ : la guerre nucléaire, l'usage de la biotechnologie par des terroristes ou des états-voyous, le changement climatique et l'intelligence artificielle.

Depuis la guerre froide, des mouvements pacifistes protestent contre les risques d'anéantissement nucléaire. Des biologistes renommés, comme Jacques Testart, se sont élevés contre les tentations d'eugénisme. Les conséquences délétères du dérèglement climatique sont anticipées et largement connues. Cependant, pour Nick Bostrom, fondateur du *Future Humanity Institute* de l'université Oxford, l'arrivée d'une intelligence artificielle forte, si elle devait se réaliser, représenterait une menace supérieure à toutes les technologies précédentes, y compris les armes nucléaires⁸. L'humanité est **responsable** de son propre avenir : si elle ne fait rien, elle risque de disparaître ; si elle agit, elle a beaucoup plus de chances de survivre. **Cette responsabilité de la possibilité d'une humanité future confère à la survie de l'espèce un sens éthique.**

Cette éthique de responsabilité de la perpétuation de l'espèce interpelle notre relation aux **machines** et aux **objets** à travers l'intelligence artificielle, mais aussi notre relation à **la nature** par le dérèglement climatique, **au vivant** par les excès de la biotechnologie, **aux autres humains** par la menace de guerre nucléaire. C'est notre rapport au monde et à notre évolution à travers le progrès technique qui est questionné.

Emmanuel Bertrand-Egrefeuille

⁶ Ce paragraphe doit beaucoup aux remarques de fond de L. Cournarie, que je remercie.

⁷ Fox News, ['Terminator center' to open at Cambridge University](#), 26 novembre 2012.

⁸ Cité par Raffi Khatchadourian, [The Doomsday Invention, Will artificial intelligence bring us utopia or destruction?](#), The New Yorker, 23 novembre 2015.