L'IA aux frontières de l'esprit

Introduction

Nicolas Léger et Adrien Tallent

Nous y sommes. Notre quotidien a rejoint ce qui, il y a encore peu, relevait de la science-fiction. L'intelligence artificielle (IA) est dans nos smartphones, au cœur de l'actualité technologique et même géopolitique. Les prophéties fusent : la fin de l'humanité est-elle en marche ? Serons-nous remplacés dans notre travail, réduits à des prestataires algorithmés ? Ce bouleversement technologique est à n'en pas douter un bouleversement anthropologique. Cerner sa singularité et la rupture que l'IA marque est la première exigence faite à la pensée. Pourtant, quand on suspend fascinations et angoisses, il est évident que l'intelligence artificielle ne surgit pas *ex nihilo* et s'inscrit dans un projet technologique et des imaginaires déjà anciens. *« Nous pouvons espérer que les machines pourront égaler les hommes dans tous les domaines purement intellectuels. »* Voilà comment Alan Turing, père fondateur de l'« intelligence artificielle » (cette expression est néanmoins forgée en 1956 par l'informaticien John McCarthy), concluait un article publié en 1950 et intitulé "Computing machinery and intelligence". Nous étions alors dans les premières années du rêve de l'IA. Sept décennies plus tard, ces systèmes automatisés sont massivement utilisés dans tous les domaines – social, économique, politique, technique – de la vie humaine. Nous parons l'IA de fonctions prédictives, prescriptives, voire injonctives. Nous lui déléguons tout ou une partie de nos processus de prise de décision. Comme d'autres ruptures technologiques avant elle, elle nous somme de réinterroger notre devenir à l'aune des possibles qu'elle ouvre, autant que notre présent, qu'elle transforme déjà. Nous faisons ici l'hypothèse que se joue cette fois une mise en concurrence de la technologie avec ce que nous avions jusqu'ici coutume de définir comme propre à l'humain :

son esprit. Mise en concurrence qui confine à la guerre, tant les fronts ouverts sont multiples. Le front pionnier de la Silicon Valley est le lieu contemporain d'où se déploie une conquête agressive et transgressive. Mais comme toutes les ruptures technologiques, l'IA ouvre des horizons nouveaux : une *terra incognita* se profile et exige d'avoir conscience d'où nous partons, et avec quels bagages.

Une aliénation parachevée ?

Une telle affirmation nécessite un double travail de démythification et d'appréciation du contexte contemporain troublé dans lequel l'IA se déploie à grande échelle. Plutôt que des réponses, toujours précaires et hypothétiques, il s'agit ici de formuler des questions probes à l'aune d'une inquiétude et d'une vigilance nécessaires. Le projet cybernétique, repris par les grandes entreprises de la Silicon Valley, porte des représentations de l'homme et des sociétés qui, aussi fantasques paraissent-elles, n'en demeurent pas moins efficientes et productives. Nous assistons au triomphe hégémonique du paradigme technique appliqué au corps et à l'esprit humain. Efficacité, performance, simplification sont les horizons d'un modèle qui nous promet une automatisation toujours accrue des processus. Processus de production, de création, d'organisation sont évidemment en première ligne, mais les ambitions ne s'y cantonnent pas : justice, santé, éducation sont envisagées comme des domaines d'extension de la rationalité algorithmique. Le pseudo-philanthropisme affiché ne saurait masquer une autre hégémonie tout aussi inquiétante : le monopole économique et idéologique de ces grandes entreprises. Celui-ci réduit drastiquement toute possibilité d'appropriation des IA par les divers acteurs sociaux et de souveraineté de leurs usagers. Que les finalités soient vantées comme émancipatrices ou régulatrices, elles s'appuient avant tout sur une modélisation réductrice et dévitalisante de l'individu. Les processus d'aliénation et de désindividuation déjà à l'œuvre avec l'avènement du numérique s'accélèrent. L'extension du domaine de l'algorithmique est sans précédent, conjuguant l'horizontalité et la désintermédiation des usages avec une verticalité d'une intensité inédite, dans notre rapport à une technologie qui nous commande. L'IA, artefact humain, s'avance comme un sphinx qui nous interroge sur ce

qui nous définit comme personne, interlocuteur, créateur ou même citoyen, c'est-à-dire dans notre propre singularité. La crainte ou le fantasme d'une « singularité », d'une IA devenue démiurge, autonome et dominante occulte pourtant ce qui est profondément en jeu : l'acceptation d'une réduction de l'humain à un ensemble de compétences et d'usages qui peuvent dès lors être copiés, améliorés, dépassés. Depuis plusieurs années fleurissent des tests en tous genres censés démontrer que l'IA est désormais plus performante que l'humain dans presque tous les domaines. L'IA comprend mieux un texte, l'écrit mieux, résout mieux des problèmes, soigne mieux, décide mieux, est plus empathique… Ces tests participent d'un discours de justification de l'automatisation généralisée – automatisation non seulement des capacités humaines, mais aussi de l'humain lui-même.

L'individu est ainsi vu comme un objet de gouvernance plutôt qu'un sujet citoyen. En 1948, le père Dominique Dubarle, très au fait des développements techniques, dans un article paru dans *Le Monde*, faisait déjà le parallèle entre le Léviathan hobbesien et la *« machine à gouverner*[1] *»*. Dans un article pour la revue *Esprit*, deux années plus tard, il entrevoyait avec inquiétude la gestion cybernétique des affaires humaines[2]. Le Léviathan de la Silicon Valley a aujourd'hui émergé et affirme son emprise. Ainsi, Elon Musk, dans le bureau ovale, cœur de la démocratie américaine, se fait le chantre de la dérégulation et de l'optimisation budgétaire et institutionnelle. Nous ne pouvons plus l'ignorer, l'IA est élaborée dans un creuset idéologique et philosophique qui n'a rien de neutre. C'est une certaine vision de l'homme et du monde qui conditionne ses élaborations actuelles et ses usages, où le techno-solutionnisme fonctionne comme un nouvel horizon eschatologique. Il se veut la promesse d'un monde régi par l'IA, sauvé de l'effondrement climatique et délivré de l'erreur humaine et de ses failles. La figure du progrès technique se substitue dès lors à celle du salut, et les prophètes de la Silicon Valley n'hésitent pas à endosser ce rôle messianique. L'IA n'est pas une entité qui flotterait au-dessus de la société, mais le produit d'une chaîne de décisions humaines. Elle est inscrite dans des structures de pouvoir et de valeurs qui lui préexistent, mais qu'elle renforce. Elle sélectionne, classe,

1 - Dominique Dubarle, « Vers la machine à gouverner… », *Le Monde*, 28 décembre 1948.
2 - D. Dubarle, « Idées scientifiques actuelles et domination des faits humains », *Esprit*, septembre 1950.

hiérarchise, optimise, automatise, consomme. Elle reconduit certaines visions du monde plutôt que d'autres. Elle matérialise des rapports de force sociaux, économiques et politiques. La neutralité devient alors une justification pour la mise en œuvre de ces technologies, auxquelles il n'y a plus de raison *a priori* de s'opposer dès lors que la neutralité a la force de l'évidence. Nous ne sommes pas loin des discours de darwinisme social qui font aujourd'hui recette outre-Atlantique.

Les défenseurs de la supposée neutralité de l'IA et les alarmistes représentent deux faces d'une même pièce. Tous empêchent de considérer que derrière la machine artificielle se trouvent des personnes, physiques et morales, qui mettent en marche une certaine vision du monde. Un monde qui serait parfaitement optimisé, prévisible, contre l'imprévisibilité démocratique et le temps de la délibération. Les promoteurs d'une version technologique de l'autoritarisme et du *management* politique sont ainsi parvenus jusqu'au pouvoir de la première économie du monde. Ce sont Peter Thiel (milliardaire, investisseur des premières heures de Facebook et admirateur de la pensée de René Girard[3] ; il a notamment financé la campagne de J. D. Vance dans l'Ohio en 2022) ou Elon Musk. Il est donc évident que la réduction au *process*, à l'automatisation, est bien une menace démocratique. La génération massive et infinie de discours et d'images est désormais possible dans une ère où le vrai et le faux indiffèrent, où les passions politiques et le discrédit prévalent. L'IA conversationnelle débarque dans un monde où le désarroi affectif et la désagrégation des rapports humains sont un sujet d'inquiétude pour nos sociétés. L'IA pourrait bien vite être considérée comme un compagnon, un interlocuteur digne de notre empathie ou de notre attachement : un *alter ego*. Elle serait un réconfort, une bulle hyperadaptative dans des temps désenchantés. Loin d'être de simples instruments de gestion, les technologies prédictives promettent des *« décisions prises à l'avance*[4] *»*. Elles ne se contentent pas d'informer ou d'accompagner nos choix : elles les précèdent, orientent et modifient nos comportements avant même que nous ayons conscience d'avoir pris une décision. Dès lors, *« le pouvoir s'exerce par régulation anticipative et non par réglementation effective*[5] *»*. La pratique

3 - Adrien Tallent, « Facebook, antichambre du trumpisme », *Esprit*, mai 2022.
4 - Günther Anders, *L'Obsolescence de l'homme*, t. I, *Sur l'âme à l'époque de la deuxième révolution industrielle (1956)*, Paris, L'Encyclopédie des nuisances/Ivrea, 2002, p. 16.
5 - Anne Alombert, *Schizophrénie numérique*, Paris, Allia, 2023, p. 59.

du pouvoir change. Le nouveau monde de l'IA acquis au paradigme de l'efficacité ne peut se satisfaire des incertitudes et de l'imprévisibilité de la délibération publique et citoyenne. Sur ce point, l'exemple chinois de *social credit scoring* – quoique parfois soumis à la surinterprétation d'une surveillance panoptique centralisée[6] – démontre le risque d'un usage autoritaire de la possibilité du contrôle technique. Il s'agit d'affirmer et de mettre en avant l'irréductibilité de l'humain face à l'IA, en réfutant l'inéluctabilité de l'automatisation. Cela suppose un changement de perspective : ne plus se contenter d'encadrer les Gafam, mais s'interroger sur leur existence même ; ne pas simplement réguler l'IA, mais en contester les prémisses. La technique est toujours porteuse d'un projet politique. L'enjeu est de réinvestir cet espace de choix et de redonner sens à une pensée critique capable de poser les bonnes questions.

La candeur de l'idée selon laquelle il existe de bons ou mauvais usages de l'IA au gré des choix des individus n'est plus de mise. L'IA n'est pas un simple outil et, si c'est une machine, ses interfaces langagières, anthropomorphiques, engagent de nouveaux rapports. Il nous faudra apprendre à vivre avec, dans la convivialité[7] et non dans une relation de fascination ou de soumission. Face à ce miroir trompeur, le risque est de survaloriser l'intelligence comme simple capacité de calcul et d'ignorer les ressources irremplaçables de l'erreur, de l'affect et de la bifurcation. Nous sommes à la croisée des chemins : l'inquiétude, le refus, le choix et la pratique critique se doivent d'être les points cardinaux précieux de notre boussole.

Face à ce miroir trompeur, le risque est de survaloriser l'intelligence comme simple capacité de calcul.

Accompagner, comprendre, résister ?

Il serait cependant vain et paresseux de se laisser définitivement emporter par l'angoisse dystopique. Ce dossier s'attache ainsi à montrer qu'il nous faut puiser dans notre héritage et notre patrimoine pour permettre une tension et une résistance devant ce qui se déploie comme une lame de fond. En rappelant ce que signifient « penser », « créer », « dialoguer » ou

6 - Entretien avec Séverine Arsène, « Le pouvoir numérique chinois », *Esprit*, décembre 2020.
7 - Sur ce point, se reporter à Alain Damasio, *Vallée du Silicium*, Paris, Seuil, 2024.

même « aimer », il est possible de se prémunir contre cette agressivité du béhaviorisme numérique. Une IA ne « parle » que dans la mesure où l'on réduirait le langage à une simple probabilité d'agencements d'informations. Elle ne sera pas non plus « conscience », dès lors qu'une conscience sera entendue comme incarnation, corps pensant. La pensée de la Silicon Valley, faisant la part belle au cognitif, se heurte à ces irréductibles que sont le corps et la chair. Sans entrer dans la multiplicité des approches du corps promues par les nouvelles technologies, ce dernier y est avant tout considéré comme un objet de performance à améliorer ou à réparer et un support affectif à solliciter. Mais il sera peut-être l'ultime bastion de résistance, garant de notre rapport au monde.

Dans le même temps, les interrogations soulevées par les IA promues par la Silicon Valley ne doivent pas nous faire oublier la myriade d'initiatives et de réflexions en cours et déjà anciennes. L'IA peut ne pas être ce monstre technique qui engloutit tout principe démocratique sur son passage. Des expérimentations démocratiques réussies fondées sur le numérique ont ainsi vu le jour un peu partout à travers le globe – comme à Taïwan à la suite de la révolution des Tournesols de 2014 et la mise en place de la plateforme vTaiwan ou Decidim à Barcelone, *open source* et sous le contrôle des citoyens –, qui laissent espérer un autre virage technique, cette fois démocratique. Mais la création d'IA plus responsables et conscientes des enjeux contemporains, si elle est souhaitable et indispensable, ne sera pas pour autant suffisante. Dans ce que Bernard Stiegler nomme la société de la *« disruption*[8] *»*, où tout doit toujours aller plus vite et où la technique doit en permanence disrupter la société, la régulation et l'éthique apparaissent comme des « après coup » bien fragiles. Évidemment, une régulation démocratique est nécessaire et constitue à certains égards notre force et la singularité européenne. Elle devient un moyen de regagner des marges de manœuvre. Mais il ne faut aucunement sous-estimer la dynamique libidinale qui porte le désir d'IA : la rapidité, l'attestation de soi, la fluidité et la délégation des tâches sont des pentes périlleuses, renforcées par l'anthropomorphisme en jeu. Autant de promesses relayées par un marketing massif. C'est pour cela que l'appropriation doit être rendue possible dès la conception de ces

8 - Bernard Stiegler, *Dans la disruption. Comment ne pas devenir fou ?*, Paris, Les Liens qui libèrent, 2016.

technologies et leurs applications. Un « faire » engagé dans un sens et des finalités alternatives, valorisées et concrètement soutenues. L'IA est porteuse de grandes promesses et de grandes avancées techniques et technologiques dans des domaines aussi divers que la prévision météorologique, la défense des langues, la santé, la traduction, la sécurité, la lutte contre la crise climatique ou les risques cyber. Mentionnons, à titre d'exemple, le projet « Delacroix numérique » lancé par Sorbonne Université qui propose une plongée dans l'œuvre du peintre[9] (numérisation des 2 500 lettres de sa correspondance, analyse fine et croisée de ses œuvres, reconstitution d'œuvres détruites…).

À l'imaginaire de l'efficacité, de l'hyper-rapidité et d'un prétendu triomphe de la rationalité, il nous faut opposer, si ce n'est imposer, d'autres représentations. Sur la ligne de front que notre dossier définit, la littérature et la fiction sont une précieuse ressource. L'écrivain de science-fiction Philip K. Dick, dans plusieurs de ses récits, met à l'honneur la figure du bricoleur, variante particulière du grain de sable dans la machine[10]. Amateur, curieux, habile, il se distingue par sa profonde humanité, jusque dans la maladresse. Hors des rets de la technicité ou de la rentabilité, il parvient grâce à la somme de ses expériences et au gré de ses curiosités à réparer voire – parfois malgré lui – à défaire les systèmes les plus oppressants ou les plus étouffants. Autrement dit, il se donne comme un *échappement* et une échappatoire. Ce détour par la fiction est le rappel simple d'une irréductibilité de l'homme. Gilles Deleuze, dans son analyse des *« sociétés de contrôle »*, proposait quant à lui la figure du *hacker*, à rapprocher de celle du saboteur. Le *hacker* ne se contente pas d'optimiser le système, il en perturbe les logiques, créant des brèches pour développer d'autres usages et d'autres imaginaires. La situation, sur ce plan, n'est pas inédite. Le développement des technologies de l'Internet, souvent présenté comme une promesse de décentralisation et d'émancipation, a pourtant conduit à un renforcement inédit des asymétries de pouvoir et de l'accaparement des dispositifs. L'utopie initiale d'un Internet horizontal et démocratique, qui aurait permis aux citoyens de prendre le contrôle de leur environnement informationnel dans le

9 - « "Delacroix numérique" : l'IA au service de l'art et des humanités », Sorbonne Université, en ligne le 13 février 2025.

10 - Entretien avec David Lapoujade, « La recomposition des mondes. Dans l'atelier de Philip K. Dick », *Esprit*, décembre 2021.

cadre d'un contrat social plus participatif, s'est heurtée à la concentration des infrastructures et des ressources dans les mains d'un petit nombre d'acteurs privés et publics. À l'ère de l'IA, la question des communs numériques[11] s'impose plus que jamais et avec force. Nombreux sont les projets et les tissus associatifs s'attachant à la transparence des données et soucieux des finalités des technologies élaborées.

L'IA comme technologie renouvelle et trouble l'acception du terme « machine » tel qu'admis jusqu'à maintenant. Elle constitue, en quelque façon, une forme d'altérité. Comme dans le film *Her* de Spike Jonze, elle se fond dans l'intime jusqu'à redéfinir les contours du lien à l'humain et au corps. L'élément le plus saillant du récit n'est pas tant l'emprise que l'IA acquiert sur son compagnon humain que son maintien dans une altérité radicale : elle s'émancipe de son propriétaire, le laissant esseulé. Toute fantasmatique qu'apparaît cette figure de l'IA, elle permet cependant d'envisager un rapport intrigant et curieux à ce qui, déjà, se donne comme un redoublement du réel avec son étoffe propre et étrange. Le maintien de l'IA dans son inquiétante étrangeté ouvre des possibilités bien plus prometteuses que l'asservissement auquel nous paraissons condamnés. Plutôt que de penser l'IA comme une dévoration de notre humanité, il serait possible de l'envisager comme un point *contre* et *avec* lequel il nous faudra vivre, sans renoncer à définir ce qui doit nous rester essentiel, commun et inaliénable. À nous, collectivement et individuellement, d'en prendre la mesure et de ne pas tomber sous l'emprise d'une optimisation factice de soi et de nos vies. C'est à cette condition que l'exploration de nouveaux horizons de l'IA pourrait se faire, avec cette technologie non pas pour guide, mais comme compagnon de périple.

11 - Bernard Jarry-Lacombe, Jean-Marie Bergère, François Euvé et Hubert Tardieu, *Pour un numérique au service du bien commun. Questions anthropologiques et éthiques*, Paris, Odile Jacob, 2022.