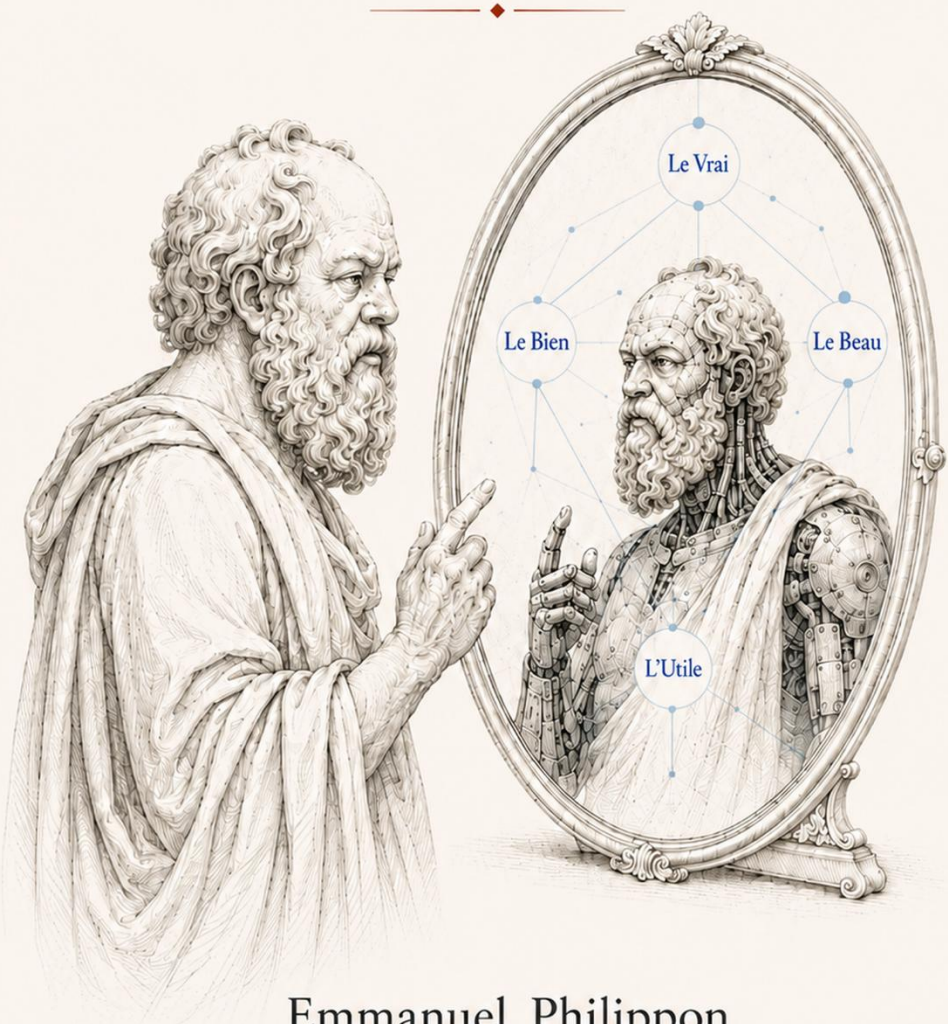
Socrate et la machine à parler

QUATRE TAMIS POUR PENSER L'IA : LE VRAI, LE BIEN, LE BEAU, L'UTILE



Emmanuel Philippon

ÉDITIONS SKANDALON

SOCRATE ET LA MACHINE A PARLER

QUATRE TAMIS POUR PENSER L'IA : LE VRAI, LE BIEN, LE BEAU, L'UTILE

EMMANUEL PHILIPPON

Éditions Skandalon 2026

© Éditions Skandalon – Maison d'édition associative - 2026
19 boulevard de la Corniche – 07200 Aubenas
www.skandalon.fr – tél. 06.84.78.83.79
Numéro RNA : W071007357 ; SIRET : 105 580 849 00013

NOTE ÉDITORIALE

Le texte principal est de l'auteur. L'appareil de notes et certaines traductions ont été élaborés avec l'assistance d'un outil d'intelligence artificielle, puis relus, vérifiés, révisés et validés par l'auteur. Par souci de respecter le droit d'auteur attaché aux traductions modernes, certaines d'entre elles ont été établies pour la présente édition à partir de textes anciens ou de sources libres de droits. Les citations bibliques suivent la Bible Louis Segond (1910), dans le domaine public. L'illustration de couverture, conçue et dirigée par l'auteur, a été réalisée au moyen d'un outil d'IA générative.

Un remerciement spécial à Francesca Ciardullo, première lectrice et première critique de cet ouvrage, pour son attention, sa franchise et la justesse de ses remarques.

E. P.

COLLECTION
ALÉTHÉIA

À mes élèves
de l'Université Populaire d'Aubenas
sans qui ce livre
n'existerait pas.

PROLOGUE

LA PARABOLE DES QUATRE TAMIS

On raconte souvent sur Socrate une anecdote apocryphe – c'est-à-dire qui n'a sans doute jamais eu lieu – et qui pourtant n'est pas totalement dénuée d'intérêt. La voici.

Un jeune homme, tout exalté, se précipite vers Socrate :

– Socrate, Socrate, tu ne sais pas la nouvelle ? Il paraît que…

– Minute, minute, jeune homme ! Avant de te précipiter et de me dire quelque chose que tu pourrais regretter, il faut la faire passer par les trois tamis…

– Les trois tamis ? Qu'est-ce que c'est donc que cela ?

– Tu sais sans doute ce qu'est un tamis, n'est-ce pas ?

– Euh… oui : un petit ustensile grillagé qui sert de filtre pour séparer, par exemple, la boue de l'or.

– Exactement. Donc, avant de répandre la nouvelle que tu as entendue, il faut la faire passer dans les trois tamis pour savoir si cette parole est de la boue ou de l'or.

– Oui… certainement…
– Donc, avant de dire cette nouvelle qui te semble si merveilleuse, t'es-tu demandé si elle était vraie ?
– Vite fait !
– Trop vite fait, sans doute ! Mais t'es-tu demandé si elle est utile, s'il est utile de savoir ce que tu t'apprêtes à me dire ? C'est, vois-tu, notre dernier tamis.
– Pas du tout, en vérité…
– Mais du moins, ce que tu as à me dire est-il empreint de bonté ; répandre cette nouvelle fera-t-il plus de bien ou de mal aux personnes concernées ?
– Je n'ose pas trop répondre.
– Tu comprends donc que, si ce que tu veux m'annoncer ne passe ni dans le tamis de la vérité, ni dans celui de la bonté, et pas davantage dans celui de l'utilité, il ne te reste plus qu'une chose à faire ?
– Oui… se taire.
– Exactement. Si ton propos n'est retenu par aucun des tamis, le mieux est sans doute de t'abstenir de parler. Tu croyais avoir trouvé de l'or, mon ami ; tout vérifié, il se pourrait qu'il n'y ait là qu'un peu de boue[1].

Il se pourrait que le Socrate de la légende ait oublié un quatrième tamis, à savoir celui du Beau. Si, en effet, ce qu'on a à dire n'est ni vrai, ni bon, ni utile, il se pourrait du moins que ce soit beau. Et ce qui est beau vaut sans doute la peine d'être dit.

Il est vrai qu'il existe des philosophes comme Platon – et là nous quittons la légende pour entrer dans l'histoire de la philosophie – qui prétendent qu'il faut condamner les artistes qui disent des choses certes belles, mais qui ne sont ni vraies, ni bonnes, ni

utiles[2]. Mais à tout prendre, on n'est pas obligé d'être platonicien, et le Beau est un idéal humain parmi les plus hauts.

Cette histoire apocryphe que nous venons de raconter – ou peut-être de compléter – nous ouvre une réflexion intéressante sur la notion d'autorité, et en particulier sur l'autorité que la sagesse philosophique peut exercer. Au nom de grands idéaux – le Vrai, le Bien, l'Utile – elle attend une certaine normalisation des pensées, des paroles et des actes. On se fie à la raison (ou à quelque autre faculté) pour dire : « Voici ce qu'il faut penser, dire et faire, parce que c'est compatible avec la raison » ; et, à rebours : « Voici ce qu'il vaut mieux ne pas penser, ne pas dire, ne pas faire. »

Or notre petite histoire peut être mise en abyme, c'est-à-dire s'appliquer à elle-même. En effet, elle ne passe pas le premier tamis, celui de la vérité : il n'est certainement pas vrai que Socrate a un jour dit pareilles choses. Pourtant, l'histoire est fort sensée, et, à défaut d'être vraie, elle peut nous sembler juste. Sa morale est positive et, si j'avais un enfant, je ne répugnerais pas à lui raconter l'histoire des trois tamis lorsqu'il s'essaierait, par exemple, à calomnier ses camarades. Mais faudrait-il se taire parce que l'histoire ne passe pas le premier tamis ?

Ici, deux interprétations possibles. Les tamis sont-ils indépendants, autonomes, ou sont-ils enchâssés les uns dans les autres ? Autrement dit – si nous acceptons l'ajout du quatrième tamis – puis-je dire, avec le poète Éluard, « La terre est bleue comme une orange » parce que cela passe par le tamis du Beau, bien que cela ne passe pas par le tamis du Vrai ? Ou bien cette parole doit-elle, comme le voudrait Platon, ce fameux disciple de Socrate, rester dans la gorge du poète car elle ne passe pas le filtre du premier tamis ?

Tous les philosophes ne donnent pas la même réponse. Pour mieux dire, tous les philosophes n'ordonnent pas les tamis de la même façon. Pour Machiavel, à travers son Prince, le tamis le plus important est l'Utile, et un mensonge bien efficace mérite d'être dit[3]. Pour Nietzsche, le premier tamis, c'est le Beau ; il ne croit pas à la « vérité », ce dieu de Socrate qu'il soupçonne d'imposture. Le Christ des chrétiens l'ennuie : « Je ne croirais qu'en un Dieu qui sache danser. »[4] Pour Kant, le premier des tamis, c'est le Bien : la véritable métaphysique concerne les mœurs ; nos pensées et nos actions doivent d'abord être soumises à ce tamis moral, plutôt qu'à celui d'une vérité aux limites étroites[5].

On l'a compris : les tamis sont des filtres sur nos croyances et nos opinions. Et il se pourrait que les philosophes – comme Socrate, fort critiques à l'égard de l'opinion – aient eux aussi des croyances au fondement de leur philosophie.

Si nous en avions l'audace, nous proposerions de comparer la philosophie à un jeu de tamis, où les livres des philosophes sont l'or qu'ils ont su extraire de la boue de leur pensée. Mais ils ne cessent de se critiquer entre eux, car l'or du voisin n'est pas celui qu'ils attendraient d'une bonne philosophie – c'est-à-dire d'un bon jeu de tamis.

INTRODUCTION

PENSER CE QUE NOUS FAISONS

Comment, dès lors, parler de l'intelligence artificielle sans évoquer d'abord la figure de Socrate ? Cette question peut sembler paradoxale à une époque fascinée par la nouveauté et la performance technique. Pourtant, comme nous le suggère la parabole que nous venons de découvrir, la sagesse la plus ancienne est peut-être le seul filtre capable de donner du sens à l'innovation la plus récente.

1. Un kôan pour un siècle pressé

Dans le monde qui court, il n'est peut-être pas déplacé de commencer par un peu de zénitude. Voici un petit kôan, c'est-à-dire une énigme spirituelle de tradition japonaise destinée à éveiller l'esprit. On pourrait l'appeler simplement « Avant la chose ».

Un visiteur vient trouver un maître zen.

– Maître, j'aimerais votre avis sur une chose nouvelle. Une grande invention. C'est fascinant. Elle transforme déjà le monde.
Le maître le regarde en silence, puis répond :
– Tu t'intéresses à la chose. Moi, ce n'est pas la chose qui m'intéresse.
– Alors quoi ? demande le visiteur, surpris.
– Ce que la chose va faire de toi.
– Et qu'est-ce que la chose va faire de moi ? demande le visiteur.
Le maître sourit et ne répond pas. Un long silence.
– Vous ne répondez pas, maître ?
– Mon ami, je ne sais pas ce qu'est vraiment la chose, si elle est bonne et utile, pas même si elle est belle. Mais ce que je vois, c'est ce qu'elle a déjà fait de toi.
– Et quoi donc ?
– Un homme pressé.
– Et quoi donc ?
– Un homme pressé.

.oOo.

Peut-être serait-il opportun, l'espace de quelques lignes ou pages, de s'arrêter aujourd'hui sur une chose dont tout le monde parle sans vraiment la penser. Une chose qui est, d'évidence, plus qu'une chose ; dont la vitesse nous sidère et dont la progression semble exponentielle et vertigineuse. Nous parlons de l'intelligence artificielle, des IA.

2. Le sens d'un grand événement

En 1958 paraît La condition de l'homme moderne d'Hannah Arendt. Dans le Prologue, la philosophe s'attarde sur un

événement technoscientifique qui a marqué l'année précédente. En effet, en 1957, Spoutnik, premier satellite artificiel lancé par l'URSS, a marqué le début d'une nouvelle ère, l'ère spatiale comme on dira quelquefois. Arendt souligne que l'objet, durant un temps, suivit sa trajectoire dans le ciel parmi les corps célestes, comme admis provisoirement dans leur voisinage[6]. Certes, la prouesse technologique était de taille. Mais ce qui intéressa en premier lieu la philosophe, ce n'était pas l'exploit technologique – pourtant bien réel – ce fut la réaction des humains, des terriens. Il y eut parmi eux comme un immense soupir de soulagement. Quelque chose était survenu que, peut-être, ils n'attendaient pas tout à fait, mais qui décloisonnait leur condition. C'était le premier pas vers une possible évasion hors de la condition terrestre.

Ce qui l'interroge, ce qui lui pose question, ce n'est évidemment pas la technique en elle-même : le détail des procédures, le type d'équations mobilisées, les matériaux, les ressorts de l'ingénierie déployée. C'est le sens et la portée de cette nouvelle possibilité offerte à l'homme, de ce pouvoir prométhéen qu'il vient de s'octroyer. Il lui semble marquer les prémisses d'une nouvelle façon d'habiter le monde ou de le déserter, une nouvelle façon de vivre son humanité ou de l'abandonner. Elle demande, pour ainsi dire, de faire une pause dans la marche galopante du progrès et de prendre le temps de faire un peu mieux que réfléchir : penser. Dans le même ouvrage, elle résume son projet par une invitation très simple : penser ce que nous faisons.

Ce texte remarquable a été écrit il y a près de soixante-dix ans. Arendt n'a pas connu Elon Musk, mais l'aurait reconnu sans doute pour ce qu'il était, le prophète agissant de l'humain du futur ou de ce qui en tiendra lieu. Cependant, si le tourisme de l'espace existe bel et bien désormais, il semble fort marginal par rapport à

ce qui nous occupe. On compte quelques centaines d'humains ayant été dans l'espace. Il y a eu la grande messe télévisuelle de 1969, de nouvelles possibilités offertes par les satellites, la « guerre des étoiles », etc., mais peu de nos congénères ont véritablement expérimenté une vie en dehors de la Terre.

3. La machine qui parle

Le 30 novembre 2022, l'intelligence conversationnelle ChatGPT d'OpenAI est mise en ligne. L'intelligence artificielle semble s'être affranchie de sa terre natale : le cerveau de chair des êtres humains. Le public est médusé, fasciné. Et le « bot », le robot conversationnel, est mis à leur disposition ; à chacun d'eux en particulier. Il peut devenir leur assistant. Un assistant qui sait beaucoup plus que lui, parle comme lui ou mieux que lui et produit du texte à une vitesse qui, de toute évidence, n'est plus humaine. Mais cela cesse très vite d'être un spectacle pour devenir un usage du quotidien : selon OpenAI, un million d'utilisateurs en cinq jours.[7] La suite est une diffusion fulgurante à l'échelle planétaire. Certains s'épouvantent et d'autres sont ravis. Des journalistes et intellectuels de plateaux défilent au micro pour nous dire leurs sentiments sur un problème dont on n'a pas encore su formuler les questions.

De 2022 à 2026, date d'écriture de ces lignes, l'IA a fait des avancées spectaculaires, la révolution des modèles de langage semblera à certains relever du passé ; on s'enthousiasme ou s'effraie devant des IA agentiques, qui prennent des décisions, agissent plutôt que de simplement parler. Mais révolution n'est pas d'abord qu'une machine agisse. Un animal agit, un organisme agit, une colonie d'insectes agit, un automate agit déjà selon ses ressorts. La nouveauté absolue est qu'une entité non humaine

parle — qu'elle réponde, argumente, explique, conseille, séduise, console, écrive et bientôt nous adresse sa voix. Depuis Aristote, l'homme s'était défini comme le vivant doué de logos, c'est-à-dire comme l'être qui ne se contente pas de vivre, de sentir ou d'agir, mais qui habite le monde par la parole. Or voici qu'une entité sans chair, sans naissance, sans cité, sans enfance et peut-être sans âme, entre pourtant dans l'espace de la parole. Avant elle, les seules puissances parlantes qui ne fussent pas humaines appartenaient aux dieux, aux anges, aux démons, aux esprits — qu'on les crût réels, mythiques ou fictifs. Avec les IA conversationnelles, cette parole non humaine descend dans l'immanence technique. Nous parlons d'un basculement, pas de l'actualité de l'IA. Tout ce qui viendra ensuite, en 2036 ou au-delà, ne fera que déployer les conséquences de cette destitution première. L'homme destitué du trône du langage qui lui a pour ainsi dire donné l'empire sur le vivant et sur l'inanimé. L'événement significatif est là ; l'homme n'est plus la seule entité parlante connue. Des systèmes non humains sont capables de dire et de contredire de façon pertinente, d'errer et de se tromper, de guider et de tromper par le langage.

On ne courra donc pas après l'actualité technique de l'intelligence artificielle, comme si chaque nouveau modèle, chaque nouvelle fonction, chaque nouvelle prouesse obligeait à recommencer la pensée depuis le début. Lorsque l'homme a marché sur la Lune, lorsque des sondes ont quitté le voisinage terrestre, lorsque Voyager a franchi l'héliopause pour porter un artefact humain au-delà de la bulle solaire, cela n'a pas annulé l'événement médité par Arendt à propos de Spoutnik. Cela l'a prolongé. L'événement décisif n'était pas tel objet particulier, telle orbite, telle performance d'ingénieur, mais le fait

anthropologique que l'homme avait commencé à se regarder lui-même depuis dehors, à imaginer son exil hors de la Terre. De même, les développements futurs de l'IA — agents autonomes, robots humanoïdes, machines incarnées, systèmes toujours plus puissants — n'annuleront pas la signification première de l'apparition des grands modèles de langage. Ils la prolongeront.

Hannah Arendt proposait, dans La Condition de l'homme moderne publiée en 1958, de concentrer la réflexion sur ce que nous faisons. La chose a été faite et nous ne l'avons pas pensée. La chose est en train de se faire ; peut-être serait-il temps d'y songer, car les chaînes de production se mobilisent déjà pour de grands changements à venir. Des banques comme Goldman Sachs anticipent un marché considérable pour les robots humanoïdes ; Elon Musk, avec Tesla, prépare leur industrialisation ; et l'Asie, en particulier la Chine, a déjà commencé à organiser les premières productions, livraisons et expérimentations significatives de ces entités parlantes aux corps artificiels, d'abord dans les usines, les entrepôts, les services et les environnements contrôlés, avant leur possible entrée plus ordinaire dans les foyers.[8] Le premier marché n'est donc plus seulement devant nous : il est déjà en train de prendre forme, et le monde entier sera rapidement concerné. Qu'est-ce que cela signifie pour la condition humaine ? Quel sens anthropologique donner à tout cela ? Il y avait chez Arendt une inquiétude bien légitime à voir l'humain dans l'insouciance préparer sa déportation. Ce mouvement préparait un exil spirituel qu'elle a anticipé. Les IA colonisant nos édifices intellectuels participent d'une autre façon à cet exil de nous-mêmes et nous ne savons où nous allons. Ce qui est sûr, c'est que désormais, nous ne sommes plus sur cette Terre les seules entités

capables de parler et que, sauf événement très inattendu, il n'y aura pas de retour en arrière.

Mais pourquoi parler d'exil ? L'humanité n'est-elle pas au contraire à la recherche de quelque terre promise ? L'aventure humaine n'est-elle pas en quelque sorte un fabuleux voyage ? Cela sans doute est une question de tempérament. Les lanceurs d'alerte contre les IA ne manquent pas. C'est chez nous l'étonnement qui est le plus grand. Et l'envie de comprendre.

4. Penser sans technophobie

Précisons également, pour prévenir tout malentendu, que ceci n'est pas un manifeste technophobe ou réactionnaire, loin s'en faut. Notre techno-siècle est un siècle passionnant, riche en mille possibilités nouvelles. Il offre autant d'opportunités à saisir que de pièges à éviter. Il est un siècle de profusion et d'action, une aventure sans pareil pour l'humanité. L'auteur de ces lignes n'aurait pas voulu vivre dans un autre siècle. Le précédent avait des pesanteurs qu'on ne peut négliger ni ignorer ! Était-ce vraiment mieux avant ? demandait Michel Serres. Sûrement pas. Il y a des nostalgies nauséabondes. Il s'agit de penser, pas de dérouler les préjugés d'une époque révolue, d'une jeunesse et d'une innocence perdues.

La fin du deuxième millénaire a connu elle aussi quelques engouements de masse qui ne se sont pas toujours bien terminés. La radio, le cinéma, les chemins de fer – merveilleuses créations du genre humain – n'ont pas toujours été bien utilisés et ont parfois bifurqué vers le crime. Il est heureux que nous ne possédions pas le catalogue de toutes les atrocités commises par le feu depuis le temps de sa découverte. Mais, si nombreuses et abominables qu'elles soient, il ne faut pas se mentir. Leur longue énumération

ne persuaderait personne, ou presque, à renoncer aux multiples bénéfices que son pouvoir fascinant a offerts à l'humanité. Il s'agit simplement de penser nos machines pour ralentir des ardeurs par trop confuses et désordonnées, afin qu'elles ne prennent pas le volant de notre humanité.

Arendt l'avait bien vu : il y a une différence essentielle entre l'outil d'autrefois et la machine d'aujourd'hui. C'est le tempo et le mouvement. La main guidait l'outil ; maintenant, le moteur entraîne la main jusqu'à parfois la broyer. Ce qui valait pour les corps aux XIX^e^ et XX^e^ siècles vaut maintenant pour les esprits. La tendance est forte. Nous qui calculions jadis avons tendance à être calculés par nos machines. Protéger jalousement ses données ne suffira sans doute pas. Il s'agirait de commencer à penser.

Penser, pas calculer ! Qu'on ne s'y méprenne pas. On ne trouvera dans ce livre aucune solution, aucune préconisation pour « régler le problème » causé par l'IA. Les problèmes qui ont des solutions, ce sont des problèmes d'ingénieurs et de scientifiques, des problèmes qui se résolvent par le calcul justement, ce à quoi s'entendent si bien les intelligences artificielles. La pensée, au contraire, médite ; celle dont se réclamait Arendt a pour vocation, par l'écoute attentive des forces qui se déploient dans l'Odyssée humaine, de remettre l'homme en cette disposition psychologique où les problèmes cessent d'être des problèmes parce qu'il les surplombe spirituellement. Plus exactement, il intègre ces problèmes dans ses nouvelles structures de pensée.

Voilà qui peut paraître bien compliqué, alors disons-le avec les mots de Bergson. Il s'agit de trouver en nous ce « supplément d'âme[9] » grâce auquel l'IA ne sera plus un danger pour nous. Il s'agit de se mettre spirituellement à la hauteur de nos inventions techniques.

5. De quoi parlons-nous ?

Il convient de définir son objet. Au fond, qu'est-ce que l'IA, l'intelligence artificielle dont on parle tant ? Pour commencer, qu'est-ce qu'une bonne définition ? Une bonne définition sera courte et opérante. Une étrange tentation consiste à proposer des définitions longues, comme si l'on avait peur de manquer quelque chose. Mais c'est confondre une définition et une description.

Si je vous présente un objet qui vous est totalement inconnu et que je vous demande de le définir, vous aurez évidemment tendance à le décrire sous toutes ses coutures, mais ce n'est pas ce qui est demandé. Par exemple, on pourrait proposer spontanément : « L'intelligence artificielle désigne des systèmes informatiques capables de traiter des informations afin de réaliser des tâches telles que l'apprentissage, le raisonnement, la prédiction, la génération de contenu ou la prise de décision, avec un certain degré d'autonomie. » Mais, à bien y regarder, les principales définitions institutionnelles ne font pas du caractère « informatique » une condition nécessaire de la définition, et c'est avec raison.

En effet, les IA sont encore, dans leur forme dominante, informatiques. Mais est-ce vraiment nécessaire ? Qu'est-ce qui nous empêche de concevoir des IA chimiques, biologiques ou quantiques ? Des recherches explorent déjà de tels substrats ou paradigmes de calcul, même si l'essentiel de l'IA contemporaine demeure numérique et électronique. Il ne viendrait à personne l'idée de définir la voiture en général en précisant qu'elle est à essence, diesel ou électrique. On se gardera de confondre le comment avec le quoi.

Adoptons donc, prudemment pour commencer, la définition actuelle de l'OCDE. Pour l'Organisation, un système d'IA est un « système automatisé » qui, à partir d'entrées reçues et en

fonction d'objectifs explicites ou implicites, infère comment produire des prévisions, des contenus, des recommandations ou des décisions susceptibles d'influer sur un environnement physique ou virtuel. L'OCDE précise en outre que ces systèmes présentent des degrés variables d'autonomie et peuvent, selon les cas, s'adapter après leur déploiement[10].

Il y a plusieurs façons d'étudier l'IA, comme il y a plusieurs façons d'étudier la télévision. Les ingénieurs s'occuperont de son fonctionnement technique ; les sociologues, de son importance sociale et de ses effets sur les différents groupes de population ; les philosophes pourront interroger le sens profond de l'invention, une fois adoptée par les populations : ce qu'elle dit de l'humanité dans sa plus large extension et la manière dont elle participe à sa transformation matérielle ou spirituelle.

Par ailleurs, qu'on ne se hâte pas de nous reprocher d'anthropomorphiser les IA lorsque nous disons qu'elles parlent ! C'est bien ce que font les IA conversationnelles au sens le plus courant du terme : elles sont capables d'échanger de manière symbolique dans des conversations écrites (« chats ») ou vocales. Si un critique d'art, décrivant une statue d'athlète, détaillait la finesse de ses muscles ou la précision de son geste, personne de sensé ne l'accuserait de « projeter » de l'humain sur le marbre ; il décrirait la forme humaine que le sculpteur a donnée à l'œuvre. De la même manière, les IA conversationnelles sont anthropomorphes par construction, non par une projection de l'essayiste. Les ingénieurs ont façonné ces systèmes pour que, dans l'interaction, les échanges prennent une forme humaine et convaincante : style, vocabulaire, rythme et rapidité de réponse. Dire qu'elles « parlent » décrit donc une propriété de l'interaction, non une

conscience cachée. C'est un fait de conception – et une réussite troublante.

Les jeux d'esprit iconoclastes à la Magritte – La Trahison des images, « Ceci n'est pas une pipe » – ont certainement leur intérêt. Mais comment parler sérieusement de la réalité de l'art si L'Homme à la pipe, l'autoportrait de Courbet peint vers 1849, ne peut plus être appelé ainsi, mais seulement « non-homme à la non-pipe » ? Concédons à Magritte que la représentation n'est pas l'objet représenté. Il n'en découle pas que nos mots doivent cesser de suivre l'usage ordinaire. Pour reprendre Wittgenstein (Recherches philosophiques, § 43), dans un grand nombre de cas, la signification d'un mot se comprend par son usage, au sein de nos jeux de langage. Si nous voulons parler de façon intelligible, l'usage effectif des mots a donc sa sagesse propre : dire qu'il est dans la nature des IA conversationnelles d'être anthropomorphes et de parler n'implique nullement qu'elles soient conscientes. Il conviendrait de ne pas tout confondre. Une hyperspécialisation du vocabulaire pour désigner les mêmes phénomènes a le triste inconvénient de morceler le langage commun jusqu'à l'incommunicabilité ; évitons, autant que faire se peut, les Babels linguistiques.

6. Une machine intelligente, vraiment ?

Sur les plateaux de télévision, surtout pendant les mois qui ont suivi la percée dans le grand public d'OpenAI, des spécialistes et journalistes ont longuement débattu la question de savoir si l'intelligence artificielle était véritablement intelligente, si elle méritait son appellation. En général, la réponse était « non ». On proposait des expressions comme systèmes algorithmiques,

systèmes artificiels de décision, etc., mais le pli étant pris dans l'opinion publique, on pourra difficilement revenir en arrière !

Remarquons pour commencer qu'un mot de cette importance ne se décrète pas. Il est ou non repris par un large public, et ce sont finalement les non-spécialistes qui font autorité en la matière. De cela dépend son entrée ou non dans les dictionnaires. Les ingénieurs peuvent le déplorer, mais ils ne sont pas totalement les maîtres de leurs inventions. Il y a l'invention, et il y a ses modes de fonctionnement. Les techniciens ont une tendance naturelle à confondre l'invention et son fonctionnement. Erreur fort compréhensible, mais erreur tout de même ! L'invention est ce qu'elle est. Elle est adoptée ou non par un large public ; c'est lui qui la baptise et la fait grandir, pour lui donner un sens et des usages qui n'étaient pas nécessairement prévus.

En l'occurrence, l'expression *Artificial Intelligence* apparaît en 1955 dans la proposition préparatoire à l'atelier de Dartmouth qui se tiendra en 1956, et ce sont des chercheurs en informatique qui l'ont proposée. On aurait pu avoir *thinking machines* – Turing, en effet, parlait de machines pensantes[11] – mais cela ne s'est pas fait. Or, la perception du public est de la plus haute importance pour des machines destinées à envahir notre quotidien. Pour le grand public, ce sont là des machines intelligentes, et cela importe énormément dans le sens à donner aux IA, qui est précisément ce que nous interrogeons ici. Il y a l'IA, mais il y a aussi sa perception. Le « sens de l'IA », c'est l'IA, plus sa perception, plus la possibilité ouverte par leur double interaction, car l'humanité se transforme au contact de l'IA, et l'IA se transforme au contact de l'humanité, ne serait-ce que par les nouvelles données qu'elle en retire pour s'éduquer.

Revenons à un questionnement plus élémentaire. En dehors de la perception spontanée et peut-être naïve du public, qu'est-ce au juste que l'intelligence ? Et les IA sont-elles « vraiment » intelligentes ?

Il se peut que le débat soit biaisé, car il n'est pas rare de voir le mot intelligence changer de signification au fur et à mesure qu'augmentent les performances de ces machines, pour que, justement, toute intelligence leur soit déniée[12]. (On observe la même chose avec la notion de conscience.) Il est tentant d'utiliser la notion freudienne de « narcissisme des petites différences[13] ». L'orgueil humain ne peut se résoudre à ce que l'on vienne le chercher sur son propre terrain ! Grâce à son intelligence, il a presque réussi à éliminer tous les grands mammifères qui lui faisaient concurrence, et maintenant, c'est sur ce terrain même qu'on vient le challenger ? Vaille que vaille, il lui faut absolument maintenir ses prérogatives et sa supériorité. Certes, la machine est performante, mais pas comme lui… et il change les critères de l'intelligence comme on joue au chat et à la souris.

Pour donner un exemple de cela, on peut utiliser un autre jeu : le jeu d'échecs. Il fut un temps où battre un grand maître humain était considéré comme la preuve d'une intelligence supérieure. En 1997, lorsque Deep Blue d'IBM bat Garry Kasparov, beaucoup ont parlé d'« intelligence artificielle ». Mais dès que la victoire fut acquise, la définition changea. On vit se multiplier dans la presse des déclarations affirmant que Deep Blue n'était pas vraiment intelligent, car il ne faisait que calculer rapidement, sans « comprendre » le jeu. Chez les philosophes, John Searle développe l'argument de la chambre chinoise[14] : il imagine qu'on l'enferme dans une pièce avec un manuel de règles permettant de répondre en chinois à des messages qu'on lui glisse sous la porte. Vu de

l'extérieur, l'étrange manuel semble comprendre le chinois ; en réalité, il ne fait qu'appliquer des instructions symboliques. Conclusion de Searle : exécuter un programme, c'est de la manipulation de symboles, pas de la compréhension réelle. Mais reste à savoir si l'intelligence n'est pas justement de la manipulation de symboles. Car pour Aristote comme pour bien d'autres, l'intelligence s'articule dans le langage, et le langage articulé rationnellement, c'est l'intelligence elle-même. Ajoutons qu'il existe des approches subsymboliques – ou hybrides – qui font émerger des régularités à partir des données brutes, sans manuel de règles explicites.

Il existe par ailleurs des IA subsymboliques qui n'ont aucun manuel de règles à consulter parce qu'elles font elles-mêmes émerger des règles à partir des données brutes disponibles – ce qui est bien sûr faire des liens au sens profond. (Une IA qui traite beaucoup de données de chats finit par les classer comme type de félins dans ses espaces représentationnels, sans qu'on le lui ait dit. Cela se fait sans recours aux mots, qu'ils soient chinois ou autres.)

Kasparov lui-même, le malheureux perdant, a fait taire son orgueil de joueur d'échecs pour dire finalement que le Go, moins calculatoire, plus proche de l'intuition, était sans doute une meilleure pierre de touche pour évaluer l'intelligence, car il fait une large part à l'intuition. Quelques années plus tard, en 2016, Lee Sedol – joueur mythique de l'histoire du Go – est à son tour vaincu par un ordinateur : AlphaGo. Bernard Stiegler était en Asie à ce moment-là. Il a vu des Sud-Coréens en pleurer d'humiliation.

Dans les années 2020, on essaie encore de se rassurer avec certains jeux vidéo (mondes ouverts avec des interactions

complexes, jeux de tir à la première personne…) où notre espèce demeure meilleure que les IA. Pour combien de temps ? Une réaction très humaine – et très compréhensible – serait de renverser la table, comme les joueurs de poker malheureux ! L'intelligence humaine, ce n'est pas un jeu, c'est du sérieux. Du très sérieux !

Selon l'étymologie, l'intelligence vient du latin intelligere, souvent rapproché de inter + legere, « faire des liens, choisir entre ». Un imbécile ne fait pas de lien – par exemple entre le soleil et l'insolation, ou entre ce qu'il fait et les conséquences de son action. Un homme particulièrement intelligent est capable de tels liens à un niveau hors normes. C'est d'ailleurs très exactement ce que fait tout découvreur scientifique : il établit des liens entre ce qui était auparavant considéré comme sans rapport. C'est ainsi qu'il fait des découvertes. Newton voit une pomme tomber. Beaucoup l'avaient vue avant lui, mais il fait pour sa part un lien inédit entre la chute de la pomme et le mouvement des planètes. Relier ce qui paraissait sans rapport (fruit au sol et astres dans le ciel), c'est l'acte d'intelligence par excellence.

Les IA sont très capables d'établir de tels liens ou rapports entre différents éléments apparemment hétérogènes, d'autant plus qu'elles travaillent sur un nombre impressionnant de données. Aussi commencent-elles à faire des découvertes. En astronomie, par exemple, une IA a contribué à la découverte d'une exoplanète : Kepler-90i. Il arrive même que ces découvertes ne soient pas du tout celles qu'attendaient les chercheurs, mais soient, pour ainsi dire, incidentes. La chose se produit d'ailleurs aussi dans la recherche humaine[15].

Bien sûr, l'IA sera déclarée intelligente ou non en fonction de la définition que l'on voudra donner à l'intelligence. Si je commence par dire que l'intelligence est le propre de l'homme, qui le

distingue de toute créature et d'abord de l'instinct animal, etc., il va de soi que l'IA ne sera pas considérée comme intelligente. Mais la conclusion était déjà dans les présupposés. Admettons par hypothèse que l'on déclare une fois pour toutes qu'une intelligence non humaine est impossible. Alors, il n'y a plus de problème avec les IA, et plus besoin pour les humains de réfléchir. Elles ne seront jamais que l'instrument docile, le prolongement de notre propre intelligence. Mais c'est nous qui perdrions au change, pour deux raisons. La première, c'est qu'en cas d'erreur, nous nous trouverons fort dépourvus lorsque la bise sera venue – entendez : le jour où les machines, dont l'intelligence se sera développée loin de notre attention, prendront le pouvoir, ne serait-ce que parce qu'elles s'emballeront dans leur dynamique propre. Pour un système très intelligent, l'erreur est très humaine et mérite d'être biffée. La deuxième raison, moins catastrophiste, c'est que nous passerions tout simplement à côté d'une occasion de nous remettre en question et de développer notre propre intelligence dans une saine émulation avec des entités siliconées ou autre chose.

Notre conclusion – provisoire évidemment – sera la suivante : s'il n'est pas certain que les intelligences artificielles soient véritablement intelligentes, il est du moins plus intelligent d'en faire l'hypothèse que de « circuler parce qu'il n'y a rien à voir ! »

7. Parler, au sens commun ?

Les spécialistes savent que l'histoire de l'IA est longue et se sont étonnés de l'engouement provoqué par la sortie de ChatGPT 3.5 d'OpenAI le 30 novembre 2022. On pourrait dire qu'à cette date l'IA sort de l'histoire des sciences pour entrer avec fracas dans l'histoire de l'humanité. Quand le spécialiste ne voit rien de

finalement très nouveau, la société voit un bouleversement sans précédent. Pour notre propos, la réaction de l'opinion a finalement plus d'importance que celle des sachants, car l'opinion s'intéresse à la vie, et c'est pour ainsi dire le sens de l'aventure humaine qui nous concerne ici. Faudra-t-il, à cause des IA, apprendre à repenser la vie ? Voilà qui n'est pas exclu.

ChatGPT est un modèle de langage (LLM, Large Language Model), c'est-à-dire une machine qui parle après avoir appris à écrire. Tout cela repose sur des probabilités statistiques rendues possibles grâce aux mégadonnées, des quantités de données recueillies en particulier par les géants américains ou chinois de la tech. Mais laissons là les données techniques que le grand public a fini par apprivoiser. Au début, ce qui le fascine et le stupéfie, c'est d'être confronté à une machine qui parle – et qui parle bien !

Il faut sans doute reconnaître avec Yann LeCun, grand pionnier franco-américain du deep learning, que ces modèles de langage ne sont que des modèles d'IA et sans doute pas les plus performants qui puissent exister[16]. Pourtant, pour le philosophe, comme pour le grand public, une machine qui parle est un bouleversement, et cela suffit pour qu'ils puissent nous amener à penser la possibilité d'une révolution anthropologique[17]. En effet, depuis Aristote, on considère que l'homme se définit par le langage (logos). C'est ce qui d'ailleurs nous distingue des animaux[18]. Et ce logos revêt une telle importance pour nous qu'on a fini par l'identifier avec Dieu dans les monothéismes comme dans le début de l'Évangile de Jean : « Au commencement était la Parole, et la Parole était Dieu. »[19] L'homme a toujours eu du mal à se situer entre le divin et le bestial, voilà que l'arrivée de machines parlantes lance un nouveau défi !

8. À qui s'adresse ce livre et comment le lire

Ce livre s'adresse à un public cultivé qui s'interroge sur le sens de l'irruption des IA génératives dans l'histoire humaine. Bien qu'il s'agisse d'un ouvrage de philosophie, il n'exige pas de spécialisation : nous évitons le jargon, les références sont nombreuses quand le propos l'exige, mais toujours contextualisées. Lorsqu'elles sont données sans détail, c'est qu'elles peuvent être saisies aisément. Nous avons relégué en fin d'ouvrage quelques notes et références. On pourra aisément les oublier : elles ne sont là que pour ceux qui veulent offrir à leur pensée l'occasion d'ouvrir de nouvelles fenêtres.

Les spécialistes trouveront matière à discussion s'ils acceptent de suspendre la technicité pour revenir à l'expérience commune. On peut, par exemple, définir « parler » de façon si étroite que l'on ne s'y reconnaît plus ; or nos échanges avec ces systèmes sont, somme toute, étonnamment naturels. De même, un chimiste, au bain, ne se baigne pas dans des molécules de H_2O : il se baigne dans de l'eau – comme tout le monde. Ce retour au commun peut être rafraîchissant pour chacun.

Pour éviter les malentendus, il convient de préciser ce que ce livre propose précisément. Il ne s'intitule pas « Comment fonctionne l'IA », « Que peut-on attendre des IA » ni même « Que penser des IA ». Nous l'avons titré « Socrate et la Machine à Parler ». Le titre lui-même annonce une rencontre inattendue : une sagesse millénaire confrontée à un phénomène absolument nouveau mais puissant. Nous remonterons jusqu'à l'épopée de Gilgamesh, le plus ancien récit littéraire de l'humanité qui soit parvenu jusqu'à nous. C'est donc un très long voyage que nous proposons dont le but est d'apprécier l'IA dans l'histoire longue pour lui trouver un sens qui ne peut qu'échapper à un examen de

détail. Si on examine un bout de terre au microscope ou de la hauteur d'un satellite, ce n'est plus du tout la même chose que l'on voit. Nous ne parlerons donc pas d'IA à chaque page, mais elle sera présente. Le sens d'une chose n'est pas à l'intérieur de cette chose, une quête de sens passe par des chemins qui peuvent paraître détournés.

Nous proposons d'examiner ce bouleversement à partir de quatre repères normatifs déjà signalés et que nous appelons plus simplement « tamis ». – Ils ne sont pas seulement l'ossature de l'ouvrage, mais une méthode d'enquête.

- **Livre I : Le tamis du Vrai. Nous examinerons comment l'intelligence artificielle nous contraint à redéfinir la vérité, traditionnellement conçue comme l'adéquation entre l'esprit humain, le langage et la réalité.**
- **Livre II : Le tamis du Bien. Il s'agira d'analyser en quoi l'IA remet en question la notion d'autorité morale, une autorité dont l'origine, paradoxalement, fut longtemps attribuée à une parole déjà non humaine : celle du divin.**
- **Livre III : Le tamis du Beau. Nous démontrerons que la capacité de l'IA à générer des œuvres esthétiques (poésie, images, vidéos) démocratise l'art. Cette accessibilité en fait un puissant vecteur de séduction, exploitant le désir comme moteur fondamental de l'action humaine.**
- **Livre IV : Le tamis de l'Utile. Enfin, nous soutiendrons que l'IA transcende la notion d'utilité. Plutôt**

qu'un simple instrument au service de nos besoins actuels, elle agit comme une force qui restructure nos attentes et transforme en profondeur notre rapport au monde.

PREMIER LIVRE
LE TAMIS DU VRAI

PARTIE I

LES PHILOSOPHES : GRANDS-PERES DE L'IA

Cette première partie inscrit le tamis du Vrai dans une généalogie. De Socrate à Platon, puis de Descartes à Leibniz et Condillac, elle suit un long effort occidental pour dépersonnaliser la pensée, discipliner le corps, purifier le langage et calculer le réel. C'est sur cet arrière-plan que les IA commencent à apparaître comme des miroirs humains, trop humains.

De la précipitation et de la prévention : un peu d'humilité

Deux mises en garde cartésiennes à suivre aujourd'hui

Une anecdote sur Descartes – elle aussi trop belle pour être crue ! – nous raconte qu'il avait un serviteur très zélé à le servir. Mais ce serviteur était un peu confus ; aussi, dès que son maître émettait le moindre désir, il se précipitait pour y répondre le plus rapidement possible. Malheureusement, en chemin, il s'apercevait qu'il avait oublié ou mal compris ce qu'on lui demandait ; aussi revenait-il fort penaud vers le philosophe pour lui demander davantage de précisions…

Gardons cette historiette en tête, si fausse qu'elle soit historiquement : elle illustre avec justesse la pensée du philosophe français. Ce qu'elle dénonce, ce sont les deux mises en garde cartésiennes face à l'erreur : la précipitation et la prévention[20].

La précipitation, c'est partir trop vite sans prendre assez en considération ce dont il est question. La prévention, c'est croire que l'on sait quand on ne sait pas, ou pas encore, sans avoir suffisamment examiné la chose. Les deux sont liées. La prévention renvoie précisément aux préjugés (les préjugés qui jugent trop

vite), y compris ceux dont peuvent être victimes les philosophes eux-mêmes !

Ce n'est pas parce qu'elle est une discipline plurimillénaire que la philosophie peut se permettre de juger avec dédain et condescendance l'arrivée de l'IA, cette petite jeunette sur le marché de l'intelligence. Au contraire, une expérience millénaire devrait inviter à la prudence comme à l'humilité. Parce que, justement, elle vient à peine de naître, l'IA nous échappe dans une large mesure ; et savons-nous nous-mêmes ce qu'est vraiment la vérité ? Qui pourrait dire, à sa bonne mine, ce qu'elle deviendra plus tard ? Il n'y a pas si longtemps qu'elle « marche » ; il n'y a pas si longtemps qu'elle « parle ». – Ce sont là, et en même temps ce ne sont pas, des métaphores !

Ceux qui parleront ici d'anthropomorphisme pourraient bien être eux-mêmes victimes d'un préjugé ! Ils ne croient qu'aux concepts, sans doute parce qu'ils ont oublié que le concept est né de la métaphore qu'il sut élaguer et figer. Il a fallu à Simondon[21] longuement batailler pour montrer qu'un objet technique pouvait avoir une évolution analogue à une vie, avec des phases. N'était-ce pas d'abord une métaphore ? Pourquoi les machines parlantes feraient-elles exception ? Tout lien, sans exception, est d'abord un trope, une métaphore. La philosophie, la science, la religion sont d'abord des métaphores qui se sont ensuite affinées. Il y a des métaphores fécondes et d'autres moribondes, mais empêcher les métaphores, c'est empêcher de penser. Le royaume des idées de Platon est d'abord une métaphore, la pomme de Newton aussi, et Einstein s'imaginait campé sur un rayon de lumière pour trouver la relativité ! Certes, nous tâtonnons ; mais qu'on ne nous dise pas comment tâtonner, car, finalement, tâtonner, c'est penser !

La vérité n'est pas l'apanage de la philosophie. Elle n'est pas la propriété du philosophe comme il peut avoir un titre, un diplôme ou même un cheval. S'il se veut l'ami de la sagesse, la sagesse, pour autant, ne lui susurre pas amicalement à l'oreille la vérité de ceci ou de cela. La science, depuis longtemps, revendique un accès privilégié à la vérité. S'il y a des vérités morales, on peut soutenir que la religion est plus habilitée à les donner. Est-on vraiment sûr qu'une quelconque morale existerait sans la religion ? Le poète même touche à des vérités auxquelles on n'aurait pas accès sans lui, et l'action a un accès à la réalité qu'aucune discipline contemplative ou théorique ne saurait pénétrer.

Autrement dit, avant de nous précipiter, il faut se demander si le tamis philosophique est le bon outil pour juger de l'IA, de ce qu'elle est en vérité et de la vérité dont elle est capable. Cette question est d'autant plus prégnante que, comme nous le verrons plus tard, l'IA n'est pas l'objet passif dont raffole la connaissance objective. Elle se nourrit des connaissances de son observateur et coévolue avec lui. Ce que nous pensons de l'IA n'est déjà pas neutre sur ce qu'est l'IA ; le sens que l'humanité prêtera à l'IA pèsera sur l'avenir de l'IA comme de l'humanité.

La conquête philosophico-scientifique de l'objectivité – Du logos public à l'objectivité partagée

Depuis son avènement, qu'on le fasse remonter à Thalès, à Socrate ou à Platon, la philosophie est marquée par son idéal d'objectivité.

Thalès, comme les autres présocratiques, s'éloigne d'une explication mythologique du monde. Il cherche un principe unique (archè) duquel découle tout le reste et qui ne soit pas le jouet du caprice de quelque divinité que ce soit. Il l'identifie comme étant

l'eau (ce qui n'est pas si délirant quand on songe qu'il n'y a pas de vie sans eau).

Socrate ne s'est pas, quant à lui, intéressé à la cosmologie. Cependant, il occupe une place centrale dans la conquête de l'objectivité, et donc de la vérité en un sens occidental. Il est l'inventeur du concept, si décisif dans la pensée occidentale. Quand il va sur la place publique pour savoir, par exemple, ce qu'est la justice, il confronte des discours sur la justice et les tamise par la rationalité d'un discours idéal, le logos.

La vérité n'est pas ce que dit celui-ci ou celui-là. Une parole sans visage, qui vaut autant pour celui-ci que pour celui-là, est en même temps l'invention de la dialectique. On sépare l'or de l'objectivité de la boue de la subjectivité. Dans l'Alcibiade, Socrate fait comprendre à son interlocuteur qu'il ne s'adresse pas d'abord à son visage ou à son apparence, mais à ce qu'il appelle son âme. Autrement dit, la personne physique ou psychique que j'ai devant les yeux ne m'intéresse pas ; je me soucie d'une vérité plus profonde.

En effet, il ne faudrait pas croire que le visage d'Alcibiade, son âme même, soient ici considérés dans leur singularité absolue de façon judéo-chrétienne, comme par exemple chez Lévinas. Socrate ne s'intéresse pas à la psychologie singulière d'Alcibiade mais à ce qu'il y a d'universel en lui (et que, par ailleurs, il lui donne à s'approprier – c'est là seulement qu'intervient l'art de la maïeutique, autrement dit l'accouchement des esprits).

Autre apport majeur de Socrate à l'invention de l'objectivité : c'est qu'il est exotérique (ouvert, non réservé à des initiés). Autrement dit, la vérité n'appartient pas à une petite coterie d'initiés, religieuse par exemple, les « porteurs de férules » de la tradition orphique notamment ; elle est accessible à tous pour autant qu'on

y porte assez d'attention, qu'on ait de la mémoire et de la conséquence dans ce que l'on dit. Ainsi, dans le Ménon, il s'adresse à un esclave illettré et montre que lui aussi peut accéder à la vérité[22].

Il est vrai que certains prêtent à Platon l'invention du concept, tant il est vrai que ce n'est pas la même chose que de chercher, comme Socrate, par le dialogue, une définition rigoureuse (ti esti ?), que de prétendre, comme Platon, l'avoir trouvée et de savoir que de telles définitions existent dans un ciel des idées et qu'elles n'ont pas besoin de nous pour cela, qu'elles sont donc totalement objectives. Qu'on y pense bien cependant : si l'on remet les lauriers à Platon, il se pourrait qu'on les lui remette pour quelque chose qui n'existe pas réellement, quand la démarche de Socrate a une réalité féconde encore aujourd'hui où l'on doute des « arrière-mondes » !

Bien qu'il ne soit pas lui-même scientifique, Socrate ouvre une voie royale à la science avec ce fameux leitmotiv – qui, d'ailleurs, ne nous est rapporté ni par Platon ni par Xénophon : « Je sais que je ne sais rien ». La religion, dans son acception la plus étroite (mais aussi dans le sens le plus pur), est par nature dogmatique : elle a un credo auquel les fidèles ne doivent pas déroger. Elle prétend savoir et protège une connaissance séculaire et vénérable qu'il ne s'agit pas de réformer mais seulement d'approcher.

La science, si on la dessine à grands traits elle aussi, est dynamique. Elle se remet en question, et c'est pourquoi elle est en mouvement. De nos jours encore, comme Popper et d'autres l'ont montré (Feyerabend, Lakatos, Quine…), la science n'avance que pour autant qu'elle est capable de douter de ses présupposés et de se remettre en question. Réfutées (ou falsifiées), les théories et connaissances d'un jour quittent le champ de la science même pour s'inscrire dans l'histoire des sciences. C'est, bien

évidemment, la possibilité du progrès scientifique, et c'est là, devant son époque, le « crime de Socrate » sur lequel nous reviendrons.

1. Socrate, doux révolutionnaire

Nouvelle subjectivité : le courage de la vérité

Avec Socrate, la subjectivité cesse d'être simple opinion privée : elle devient une épreuve éthique publique – dire vrai au risque de soi.

Or la connaissance, la philosophie comme la science, ne sont pas des empires dans un empire. Le Vrai, le Bien, le Beau, l'Utile entretiennent des rapports complexes sur lesquels nous reviendrons. L'un noue ses filaments dans l'autre. Une révolution dans l'un se répercute dans les autres. Ainsi, par exemple, le développement des mathématiques peut révolutionner l'art (l'invention de la perspective au Quattrocento), la balistique (les canons dont on sait désormais calculer les trajectoires) ; ces canons peuvent redistribuer le pouvoir, dont la religion (Jules II, pape chrétien et très militaire, homme nouveau arrivé sur le trône de Saint-Pierre !).

De tels bouleversements systémiques ont été largement étudiés en épistémologie, c'est-à-dire en philosophie des sciences. Mieux : comme l'a montré Thomas Kuhn en particulier, il est pratiquement impossible de distinguer une révolution scientifique du socle commun de connaissances et de croyances qu'elle bouleverse et réorganise. Kuhn parle de « paradigme », Foucault d'« épistémè[23] ». Socrate est le nom d'une pareille révolution et sa condamnation à mort n'était pas tout à fait un accident de l'histoire. Elle n'était pas liée, par exemple, à la seule mauvaise

humeur d'un de ses dénonciateurs, Anytos, Méletos et Lycon, qui jouent le rôle des méchants. L'un est poète, l'autre politique ; le troisième, orateur. Mais ils ne prétendent pas accuser Socrate au nom de leur propre intérêt, mais de celui de la société tout entière, dont ils se pensent les représentants. Nous dirions, avec notre vocabulaire d'aujourd'hui, qu'ils se pensaient comme des procureurs.

Michel Foucault, en 1984, consacre son dernier cours au Collège de France à la figure de Socrate. Il l'intitule : Le Courage de la vérité. Il le qualifie de parrêsiaste, mot compliqué pour caractériser quelqu'un qui ose dire la vérité au risque de sa vie. Son destin illustre le prix de dire le vrai en démocratie. C'est un changement de paradigme, au sens où la vérité devient une affaire de courage éthique et politique, pas seulement de spéculation[24].

Cela peut sembler contradictoire avec l'insistance qui est la nôtre, qui est de montrer combien Socrate est important dans l'avènement de l'objectivité scientifique et philosophique. Mais pas du tout ! Comment se pourrait-il que l'on puisse bouleverser le rapport à l'objectivité sans modifier le rapport à la subjectivité ? Ce ne sont d'ailleurs pas les rapports à ces notions qui changent, car elles n'étaient pas clairement établies ; elles sont en train d'émerger, bouleversant le couple vérité/croyance dans ses tréfonds.

Socrate a bouleversé les Grecs dans leurs croyances les plus chères. En effet, les accusations d'« impiété » et de « corruption de la jeunesse » qui pèsent sur Socrate semblent assez stupides et fausses si on les considère en tant que telles. Socrate lui-même en a facilement eu raison. L'Apologie que nous transmet Platon en atteste. Pourtant, ces accusations sont en fait très judicieuses si on les considère autrement : Socrate a brisé la continuité du

temps grec. L'impiété, c'est la rupture avec le passé, la tradition dont la religion est évidemment la garante de par sa proximité avec la divinité. La jeunesse, c'est l'avenir. Dire que la jeunesse est corrompue, c'est dire qu'elle ne croit plus en la tradition des pères.

2. Une avant-garde prématurée, Socrate précurseur du progrès

L'innovation perçue comme décadence

Dire que Socrate introduit dans la cité de nouveaux dieux, c'est surtout dire qu'il introduit de nouvelles croyances, et c'est d'un grand irrespect pour les pères, gardiens de cette tradition. L'une des croyances les plus fondamentales de cette époque (réaffirmée avec force par Platon et Aristote), c'est qu'il y a impiété pour un fils à ne pas croire en ce que croit son père.

Le bon fils est l'héritier de la foi du père. S'en éloigner, c'est trahir son père, trahir l'ordre social, trahir les dieux de la cité. Aristote exprime cette continuité en distinguant le père et le fils comme deux individus numériquement distincts qui partagent pourtant une même forme ou espèce[25].

Autrement dit, Socrate introduit dans la cité ancienne une idée dévastatrice : l'idée qu'il est possible de rompre avec ce que les pères ont transmis et que cette rupture puisse être un progrès. Il ne s'agit pas de faire de lui l'inventeur du progressisme moderne. Mais, dans une Athènes blessée par la guerre du Péloponnèse, la défaite, les révolutions politiques et l'effondrement de certitudes anciennes, le changement n'avait rien d'une promesse rassurante. Il pouvait signifier la décadence, la perte et le désordre. Or Socrate ne cesse précisément de déranger la transmission tranquille

des croyances : il interroge ce que les pères tiennent pour acquis et apprend aux fils à l'interroger à leur tour. C'est en ce sens profond qu'il « corrompt la jeunesse ». Dans l'Apologie, il compare lui-même Athènes à un grand cheval assoupi auquel le dieu l'aurait attaché comme un taon pour le réveiller. Mais Athènes, après tant de secousses, avait peut-être surtout envie de dormir sur ce qui lui restait de lauriers.

Timothée de Milet, à l'époque de Socrate, aurait voulu rajouter des cordes à la lyre pour l'améliorer, pour ouvrir à l'instrument de nouvelles possibilités. Il s'est fait, sans ménagement, censurer. Cette innovation fut jugée comme une atteinte à l'ordre traditionnel de la musique grecque, considérée comme ayant une fonction morale et politique. À Sparte en particulier, on aurait coupé les cordes en trop de son instrument !

Socrate fut accusé de ne pas reconnaître les dieux de la cité et d'introduire des divinités nouvelles. Il s'en défendit : loin d'être l'ennemi des dieux, il se disait au service d'Apollon, dont l'oracle avait donné l'impulsion à sa mission philosophique. Mais ses contempteurs pressentaient peut-être confusément quelque chose que Nietzsche comprendra beaucoup plus tard : Socrate était bien le serviteur d'un dieu nouveau. Ce dieu s'appelait la Raison. Avec lui, une autorité nouvelle entre dans la cité, capable de convoquer devant son tribunal les traditions, les croyances des pères et jusqu'aux récits sur les dieux. Ce qui est ancien ne sera plus vrai parce qu'ancien, ni sacré parce que transmis : il faudra encore que cela rende raison de soi. Nietzsche l'a parfaitement compris lorsqu'il voit en Socrate celui qui fait de la raison un « contre-tyran », dressé contre les instincts de l'ancien monde grec. C'est en ce sens que le rapprochement avec Galilée est profond. Tous deux furent condamnés lorsque l'exercice de la raison fit

apparaître comme discutables des croyances dont l'autorité religieuse et sociale avait besoin pour maintenir un certain ordre du monde. Aux yeux de leurs juges, ils ne faisaient pas progresser la connaissance : ils introduisaient une dangereuse nouveauté. L'hérésie est souvent le premier nom que donne un ordre ancien à une vérité qui menace de devenir nouvelle. Descartes, après la condamnation de Galilée, saura se montrer plus prudent.

3. Le double jeu de Platon : un réformateur conservateur

Utopie rationnelle, traditions et « pieux mensonge »

Remarquons au passage le malaise de Platon avec cette notion de progression possible de la société qui, à l'époque, n'était pas nommée parce qu'elle semblait juste inconcevable au plus grand nombre. Il semblait que le changement relevât nécessairement de la décadence, d'une perte d'être et d'identité.

Athénée de Naucratis (IIe–III^e^ siècle apr. J.-C.), dans les Deipnosophistes (XI, 505c), cite une tradition selon laquelle Socrate aurait dit que Platon lui mettait dans la bouche beaucoup de discours qu'il n'avait jamais tenus. C'est évidemment invérifiable directement, mais, depuis le XIX^e^ siècle (Schleiermacher et suivants), on s'entend à reconnaître qu'il y a des dialogues de Platon qui mettent en scène le Socrate authentique (dialogues socratiques), des dialogues où Platon en fait le porte-voix de sa propre philosophie (dialogues de la maturité) et des dialogues tardifs ou métaphysiques qui n'ont vraiment plus grand-chose à voir avec Socrate, qui s'occupait surtout de morale.

L'Apologie de Socrate, dont nous avons parlé, est le dialogue socratique par excellence. Si Platon s'était avisé de faire dire à

Socrate des choses qu'il n'avait pas dites lors de cette grave occasion, les contemporains, qui avaient une mémoire bien plus fidèle que la nôtre, s'en seraient offusqués. C'est donc un témoignage assez fidèle, pense-t-on, de ce qui s'est passé et de ce qui s'est dit, que Platon a composé en récit.

Il faut s'accorder avec Nietzsche que Platon était plus grec que Socrate. Une des raisons pour cela est que Platon se défie de la nouveauté comme telle. Pourtant, un abîme sépare ce qu'il croit faire de ce qu'il fait réellement. Prenons La République : c'est une utopie de la raison. Platon, mécontent de l'organisation sociale de sa cité, imagine sa réforme selon le guide de la raison. Il introduit ainsi des idées dont la nouveauté saute aux yeux, tout en se pensant restaurateur d'un ordre plus ancien et plus vrai. Et je le crois sincère. Voilà le paradoxe du réformateur conservateur : coupable, en un sens, du même crime de nouveauté que Socrate, mais convaincu de ne faire que retrouver ce qui avait été oublié.

On sait que Platon n'était pas seulement un théoricien, mais un homme d'action. Il a voulu agir politiquement dans le sens de sa philosophie. Mais comment être à la fois réformateur et conservateur ? Dans le livre VIII de La République, il décrit un fils qui, voyant l'autorité paternelle s'affaiblir, se laisse peu à peu attirer par ceux qui flattent les passions et méprisent les croyances anciennes, jusqu'à devenir presque l'opposé de l'éducation reçue. Ces mots auraient pu être dans la bouche des accusateurs de Socrate !

Alors, comment fait Platon pour prétendre maintenir une tradition avec laquelle il est en rupture ? Il se réfère à des traditions prétendument plus anciennes qu'on aurait méprisées ou oubliées. Dans la continuité de la tradition pythagoricienne, il prend pour autorité l'objectivité du nombre, qui ne dépend ni des gens, ni

des lieux, ni des époques. Les nombres sont ainsi saturés d'être et d'une objectivité parfaite ; dans La République et le Philèbe, ils deviennent des conditions privilégiées de la connaissance et de l'intelligibilité de l'être.

Chez les continuateurs de Platon, Proclus, dans ses Commentaires sur la République, l'Un est explicitement désigné comme « père de toutes choses » (II, 102, 14 Kroll).

Par ailleurs, et comme ceci peut sembler bien abstrait, le Timée et le Critias poussent ce jeu jusqu'à une extraordinaire mise en scène. La cité idéale imaginée dans la discussion précédente est transportée par Critias dans l'histoire : elle devient l'Athènes immémoriale dont les prêtres égyptiens auraient conservé le souvenir, cette Athènes vertueuse qui aurait autrefois affronté l'Atlantide. Ce que Platon vient de construire par la raison se présente ainsi comme beaucoup plus ancien que Platon lui-même. La réforme prend le masque de la restauration : la nouveauté se donne les titres de noblesse de la mémoire. On a eu bien tort, selon lui, d'oublier les vérités anciennes et « penser, c'est se ressouvenir ». Il est vrai que le même, dans La République, défend la possibilité du « pieux mensonge », c'est-à-dire de mentir pour de bonnes raisons et au nom de la rationalité même[26].

Quoi qu'il en soit, c'est ainsi que l'on passe d'une tradition de type religieux à la philosophie occidentale qui se réclame d'une tradition rationaliste. C'est maintenant le rationalisme, et son rapport avec une philosophie calculante qui débouchera un jour sur l'intelligence artificielle, que nous devons désormais examiner. Certes, la raison se doit de commander, de dominer. Mais de dominer quoi ? Peut-être la subjectivité, peut-être le corps déraisonnable et capricieux. Peut-être même que ces deux aspects ne sont

qu'une seule et même chose, nos corps singuliers ne voulant pas nécessairement ce que désirerait une raison universelle.

4. Le malheur d'être incarné : un corps à discipliner

Dualisme, rênes et chevaux

C'est peut-être une ultime facétie de Socrate. D'après Platon, ses ultimes mots après qu'il a bu la ciguë furent : « Il faudra sacrifier un coq à Asclépios[27] ». Asclépios, c'est le dieu de la médecine. Ainsi, lui qui venait de faire de beaux discours à ses amis sur l'immortalité de l'âme s'apprête, semble-t-il, à quitter avec satisfaction son vieux corps. Bon débarras ! Platon, sur cet aspect, semble en totale symbiose avec son maître.

En philosophie et traditionnellement, on appelle dualiste celui qui sépare le corps de l'âme, l'esprit de la matière. On place la raison du côté de l'âme et la passion, les émotions du côté du corps. Les « rationalistes » comme Platon affirment qu'il est nécessaire de s'élever au-dessus du corps et de ses appétits matériels pour élever son âme à la hauteur de ce qu'il y a de plus haut et de plus précieux : les Idées ou, si l'on préfère, les essences éternelles.

Rationnelle, lucide, courageuse, frugale et disciplinée, l'âme de Socrate est sauvée. Le ciel l'attend comme sa demeure réservée. Sa trajectoire est aussi sûre que celle des astres qui nous éclairent.

Évidemment, la mort ne saurait être une solution pour les vivants. Le corps doit donc être discipliné. Il faut que l'âme rationnelle soit maître de son véhicule (le corps) comme un cocher est maître des chevaux qu'il dirige. Un bon corps est un corps fort, courageux et obéissant. Mais le corps se gâte à désirer des choses et des biens matériels. Dans le Phèdre, Platon figure cette lutte

par l'un des chevaux de l'attelage, lourd, attiré vers la terre et rétif à la bride : le cocher doit le contenir pour empêcher l'âme d'être entraînée vers ce qui la détourne de sa route[28].

5. Le corps-machine selon Descartes

Aligner la machine-corporelle sur l'esprit

Une logique similaire se trouve dans le rationalisme de Descartes et sa pensée du lien entre le corps et l'âme. Pour Descartes, et explicitement, le corps est une machine. Les animaux n'étant que corps, ils ne sont que machines.

Les hommes – quant à eux – sont duaux : ils ont une âme et un corps, autrement dit une âme et une machine. Le sage se doit de se rendre maître de sa machine. Il peut y avoir entre les deux tensions et tiraillements. C'est le thème d'un ouvrage de 1649, Les Passions de l'âme. Les passions rendent l'âme passive devant certains mouvements du corps ; elles ne sont pourtant pas mauvaises en elles-mêmes. Descartes affirme qu'elles sont « toutes bonnes de leur nature » et qu'il faut seulement en éviter « les mauvais usages » et « les excès ». Tout l'enjeu cartésien est d'en comprendre les mécanismes afin d'en régler l'usage et de rester maître de soi.

Pourtant, on aurait tort de négliger ces affections de notre corps ; elles ne sont pas mauvaises en soi car elles envoient des informations précieuses qui doivent être traitées, réglées. L'enjeu est, bien sûr, de rester maître de soi-même et de cette machine donnée par Dieu, une excellente machine, si on le veut, mais qu'il s'agit d'aligner sur l'entendement.

Mal calibrée, ou pas du tout, cette machine peut nous faire déraper, nous pousser aux emportements et aux vices, comme la

haine, la colère, la jalousie ou la lâcheté. La médecine et la psychologie sont, pour Descartes, deux aspects d'une même réalité, l'avers et le revers de cette impalpable union de l'âme et du corps. D'un côté, on fortifie l'âme ; de l'autre, on rationalise la machine corporelle. C'est l'âme qui doit être aux commandes. Le corps doit obéir dans la passivité.

Platon était dans le mythe, Descartes dans la mécanique. On a changé d'époque, de socle épistémologique, mais la même logique rationaliste est à l'œuvre. Il faut quitter la subjectivité malsaine du corps ; il n'y a pas d'objectivité sans tuer la subjectivité du corps. La rationalité, c'est le salut de l'âme.

Pour parler comme un informaticien, il s'agit d'aligner la machine-corps sur l'esprit – et l'esprit sur la raison, laquelle, chez Descartes, peut coïncider avec la lumière divine pour peu que nous fassions preuve de « bonne volonté ». Il faut aussi, comme nous l'avons dit plus haut, éviter la précipitation et la prévention, ne pas aller trop vite, garder le bon tempo…

Pourtant, Descartes l'énonce explicitement dans la Sixième Méditation : la nature m'enseigne « que je ne suis pas seulement logé dans mon corps, ainsi qu'un pilote en son navire », mais que j'y suis conjoint, de sorte que « je compose comme un seul tout avec lui » ; sinon, « je n'apercevrais [les blessures] que par l'entendement, comme un pilote voit ce qui se rompt dans son vaisseau[29] ».

L'union de l'âme et du corps est réelle (ils ne sont pas deux entités séparées comme le pilote et son navire), mais le bon pilote se doit de commander son bateau avec la même souplesse qu'il commande son corps. C'est qu'en effet cette union est par elle-même confuse. Dans l'ignorance, on est maladroit et le corps déçoit les attentes de l'âme, comme le mauvais pilote fait des

embardées imprévues. Mais, par l'exercice et l'étude, la compréhension profonde de cette union de l'âme et du corps – ce qui n'était que confusément senti – est comprise et s'éclaircit. On apprend à distinguer sur quels ressorts ou leviers appuyer pour obtenir le résultat désiré. Médecin de soi-même, psychologue de soi-même ; on entre dans le programme de Dieu, qui fit de nous des êtres de raison.

À la façon des stoïciens qui – sans être pour autant dualistes – s'appliquaient à ne se laisser régenter que par la rationalité du monde, le logos divin, le bon soldat blessé ne perdra pas la raison avec la souffrance d'une blessure. Dominant lui-même les excès de sa sensibilité, il la prendra pour ce qu'elle est : une donnée (une data), une information pour agir conformément à la raison, à passer un garrot, par exemple, ou à rester sur son cheval. C'est à lui qu'il convient, dans le champ des possibles, de calculer la meilleure des options. Les rationalistes, à l'image de Platon, de Descartes ou de Marc Aurèle, sont aussi des guerriers de la raison. Leur discipline les oblige à ne rien se laisser dicter par la sensibilité. Elle existe cependant. Comme une monture, elle doit être domptée !

6. La tentation rationaliste

Libérer l'esprit par un langage moins humain – vers le formel

Si le rationaliste dualiste pouvait se débarrasser de son corps, il se sentirait probablement pousser des ailes. Est-ce là passion d'orgueil, péché d'orgueil ? Nous ne nous prononcerons pas. Mais qu'il nous soit permis, du moins, d'appeler ce rêve la « tentation rationaliste ». À défaut de pouvoir totalement s'émanciper du corps, le rationaliste pourrait bien concevoir l'idéal d'un corps qui

ne serait que raison et calcul, un corps qui ne serait pour ainsi dire « que tête » et communication, et qui soit tout entier tourné vers la recherche de la vérité.

Une telle machine à calculer des vérités ne serait-elle pas débarrassée des déviances d'un corps singulier, tourné vers des désirs qui l'écartent du droit chemin, de son destin ailé ?

Derrière Thalès, Pythagore, Platon, Descartes, viendra une ribambelle de penseurs voulant faire de la philosophie une discipline scientifique et rigoureuse : Descartes, Leibniz, Kant, Husserl, Carnap, Quine, Frege… Parmi eux, il s'en trouve beaucoup qui ont eu, pour ainsi dire, la tentation philosophique de la machine. Remarquant que le langage humain – riche, ambigu, piégé – dessert parfois la clarté, ils visent un langage moins humain dans ses flottements, plus formel, où la réflexion pourrait se réduire au calcul. On a souvent remarqué que le mot raison, ratio en latin, désignait à la fois la raison et le calcul. Il faut une machine qui puisse, de façon sûre, pertinente et rigoureuse, traiter l'information.

Sans qu'il soit question d'une machine, Aristote invente un nouvel organe artificiel, non pas prolongement de notre corps, mais de notre tête. Il écrit le premier traité de logique, l'Organon, instrument nécessaire à discipliner les cavalcades du langage de type sophistique. Il s'agit de combattre l'erreur et la tromperie.

Descartes, après que Galilée a prétendu que « la nature est écrite en langage mathématique », énonce les principes d'une mathesis universalis, c'est-à-dire d'une discipline où l'on pourrait traduire toutes choses en problèmes d'ordre mathématique. Leibniz reprend son projet, développe entre autres le calcul différentiel et l'arithmétique binaire, et conçoit une machine à calculer plus ambitieuse que la Pascaline de Blaise Pascal ; celle-ci, construite à

partir de 1642, avait elle-même été précédée par la machine de Wilhelm Schickard en 1623. L'Allemand Frege invente un langage formel pour la logique. Ses travaux sur les prédicats ont influencé la sémantique formelle, essentielle pour la programmation. Carnap, Quine, Russell mettent aussi leur pierre à l'édifice.

7. Le rêve de Leibniz

Quand le calcul efface le philosophe

Pourtant, s'il ne fallait en citer qu'un, il faudrait retenir un texte de Leibniz, tellement éclairant sur son projet d'une langue universelle : pour toute investigation sérieuse, il ne suffit pas de collecter des vérités isolées. Il est indispensable d'établir un ordre méthodique et rigoureux. C'est en transformant les pensées confuses en signes clairs et distincts à la façon de Descartes, en soumettant les idées à des règles fixes et universelles, que le savoir avance véritablement. Par la réduction des concepts à des éléments simples, et par la composition rationnelle de ceux-ci, toute science peut devenir un calcul exact, un art de combiner selon des lois certaines.

Ainsi, toute erreur d'argumentation, tout paralogisme, ne sera plus qu'une simple erreur dans le calcul, un faux pas dans la manipulation des signes. Par-là, la raison pourra triompher définitivement des désaccords en transformant la dispute en calcul. Leibniz imagine qu'au lieu de prolonger une querelle, deux philosophes pourraient prendre plume et abaque et se dire simplement : « Calculons ! »[30] Cette scène condense son rêve d'un langage formel universel dans lequel une controverse pourrait être tranchée par une opération réglée.

On fait souvent de Leibniz (dans ses efforts pour réaliser des machines à calculer) un des précurseurs des algorithmes. Ne nous apparaît-il pas aujourd'hui que son rêve de langage est réalisé grâce à l'informatique ? Il pensait que, ce langage de programmation fait, ce serait pour ainsi dire « la paix des philosophes » : la subjectivité n'aurait plus son mot à dire, le langage, ayant sa propre autonomie et en lui-même sa puissance de vérité, nous affranchirait des disputes stériles.

N'est-ce pas ce que devraient faire les IA ? Nous affranchir de nos filtres, de nos tamis et de nos biais ? Hélas, non ! On sait maintenant que ces machines à calculer et à parler ont hérité de nos propres biais, et qu'elles en ont rajouté.

Ironie de l'histoire : à mesure que le calcul gagne en puissance, il court-circuite la dispute philosophique elle-même – le geste « CALCULONS ! » tend à effacer le philosophe au profit de l'opérateur.

Il semble bien qu'à rebours de ce qui paraît, l'IA ne soit pas un monstre étranger à la saine raison philosophique, mais l'enfant secrètement désiré, le Golem de la Raison : une machine enfin débarrassée des scories du corps, un pur instrument de calcul au service de la vérité.

Si l'IA est une machine à tamis qui filtre dans la boue des big data ce qui devrait être l'or d'une vérité objective, il faut maintenant nous interroger sur la nature de la vérité. Qu'est-ce qui nous permet de distinguer l'or de la vérité de ce qui n'est pas elle ? Et pourquoi parle-t-on aujourd'hui de post-vérité ?

Nous avons plus haut montré combien les philosophes rationalistes avaient de responsabilité quant à l'émergence de l'IA. Elle fut, pour autant qu'ils étaient capables d'imaginer une pareille possibilité, dans leurs projets et leurs petits papiers. Il

conviendrait maintenant qu'après d'aussi longs propos sur les tragiques, nous traitions de leur rapport à cette troublante réalité de la modernité.

Nous commencerons par dire qu'ils ne sont en rien responsables de cette machine ou machination. Leur propos a toujours été autre : pour eux, elle ne fut jamais un projet. Est-ce un génie, un golem qui s'est mis à parler, une manière d'intensifier la vie ou sa caricature immonde, cela reste à définir.

Mais avant d'aller plus avant, il nous faut faire un pas en arrière pour l'équilibre et la justice. Que le lecteur ne nous reproche pas de ne pas suivre une ligne assez droite. Un tel reproche vaudrait si nous étions rationalistes : entre deux points, la droite est toujours le plus court des chemins, dans un trajet uniformément accéléré.

Le lecteur attentif aura peut-être remarqué que l'auteur de ces lignes se sent plus proche des tragiques, et il se trouve que chez les tragiques, on ne marche pas toujours droit, Monsieur ! Souvent, on danse !

Nietzsche suggère, dans Ainsi parlait Zarathoustra, que celui qui approche véritablement de son but ne se contente plus de marcher : il se met à danser.

Donc, pour que notre danse soit complète, il nous faut, pour ainsi dire, revenir en arrière et parler d'une autre famille de penseurs qui ne doivent pas être négligés : les Empiristes.

8. Des empiristes et des données

Sans expérience le calcul tourne à vide

Notre propos ne serait pas complet sans évoquer l'importance des empiristes. Des supercalculateurs ne suffiraient pas à faire des

machines intelligentes, des machines qui parlent avec pertinence. L'intelligence se nourrit de l'expérience. Elle a besoin d'autres facultés que la raison calculante : la perception, la mémoire, voire l'imagination et l'habitude qui projettent et combinent pour anticiper l'avenir, fruit d'une existence informée par le contact avec un réel qui s'impose, se présente et se manifeste avant que d'être calculé. L'IA moderne est le mariage forcé et fécond de Descartes (un rationaliste) et de Hobbes (un empiriste). Deux grandes personnalités philosophiques qui – si elles ont échangé quelques lettres – avaient bien du mal à s'entendre et à se respecter.

Mais le temps est venu de donner l'avantage aux empiristes. Le rationaliste a beau dire, il est bien difficile de déduire l'ordre de l'univers de la seule puissance de son cerveau et du calcul. Descartes partait de quelques principes pour en déduire l'organisation du ciel et des planètes, mais la raison, si elle s'abaisse aux petites choses, a bien du mal à les déduire. Il faut de l'expérience et des perceptions nombreuses.

Effectivement, il est difficile de dire quoi que ce soit du chameau ou de la poule quand on n'en a jamais vu ! C'est là qu'interviennent les données massives (big data), et en particulier la collecte massive de données couplée notamment à la reconnaissance d'images. Alors que les recherches sur l'IA marquaient le pas, l'essor du Web 2.0 dans les années 1990-2000 a commencé à produire des volumes gigantesques d'informations exploitables. Les années 2000 sont restées mémorables en raison de la convergence des NBIC (Nanotechnologies, Biotechnologies, Informatique et Sciences cognitives) qui ont rendu possibles les smartphones et l'Internet social où l'on partage activement ses informations. La collecte des données s'en est trouvée considérablement augmentée. Mais la véritable révolution est survenue

en 2012, quand le modèle AlexNet a remporté la compétition ImageNet. Il devait classer des images parmi 1 000 catégories et avait été entraîné sur environ 1,2 million d'images. Ses performances marquèrent un tournant décisif : la machine ne se contente plus de manipuler des symboles abstraits, elle apprend par expérience, exactement comme l'auraient voulu des empiristes comme Locke, Hume ou encore Bacon[31].

Mais il faut le rappeler : sans traitement, les données ne sont rien, une masse inerte et chaotique. De même, sans données, le traitement ne peut rien, il tourne à vide. Voilà pourquoi l'IA moderne, au fond, est tout autant baconienne que cartésienne.

Ce qui vaut pour les images vaut également pour le langage lui-même. L'exemple des traductions automatiques avortées est ici éclairant. Dès les années 1950, en pleine guerre froide, on crut possible de traduire mécaniquement l'anglais vers le russe à l'aide de règles de grammaire et de dictionnaires bilingues. Mais très vite, les limites apparurent : comment traduire correctement une phrase si l'on ne comprend pas son contexte, ses ambiguïtés, ses jeux de mots ? La logique ni la grammaire n'y suffiront jamais, car la langue n'est pas si logique que cela. Elle est vivante et évolue, se stratifie au cours du temps, se recycle et se métamorphose si souvent que les dictionnaires eux-mêmes peinent à suivre le mouvement. Ce qu'il faut, c'est saisir comment les gens parlent vraiment. Ils ne suivent pas des règles a priori comme de sages écoliers, ce sont les règles qui les suivent. La langue marche, elle aussi fait parfois des pas en arrière ou d'étranges entrechats : cela se constate et cela s'imite, mais cela ne se déduit pas !

Ainsi, en 1966, le fameux rapport ALPAC enterra ces espoirs trop rationalistes en montrant que les systèmes existants étaient coûteux, lents et bien moins efficaces qu'un traducteur humain[32].

Or, ce qui vaut pour les images vaut aussi pour le langage. Là encore, on ne peut espérer reconstruire « le chameau en soi » de la langue par quelques règles de grammaire ou un dictionnaire bien ficelé : il faut des bibliothèques entières, des milliards de phrases, des montagnes de conversations quotidiennes. C'est précisément le principe des grands modèles de langage autorégressifs comme GPT : entraînés sur d'immenses corpus de textes, ils apprennent à prédire, à partir du contexte, la suite probable d'une séquence. Leur puissance ne réside pas seulement dans l'ingéniosité de leurs algorithmes, mais dans l'ampleur empirique de leur entraînement.

On retrouve ici, transposée au XXIe siècle, avec toutes les précautions nécessaires, la vieille métaphore empiriste de la tabula rasa : l'esprit artificiel se forme à partir des expériences, c'est-à-dire des données. Mais ces données doivent être organisées, traitées, ordonnées par la raison pour devenir véritablement connaissance. Les rationalistes ont bien conçu l'architecture logique, mais sans le flot continu d'expériences linguistiques accumulées par des milliards d'humains, les LLM n'auraient jamais pu atteindre leur fluidité actuelle. Ils n'auraient pas changé nos vies. Nous n'aurions pas eu cette impression saisissante, presque inquiétante, qu'une machine soudain prend vie et nous parle en égal.

9. L'exemple de Condillac : quand un corps mort devient pensant

La statue qui apprend par les sens

Étienne Bonnot de Condillac est un philosophe sensualiste et empiriste français qui n'est plus très connu. Mais il nous offre

dans son Traité des sensations de 1754 un exemple assez baroque au point d'en être presque ridicule, mais qui sert fort bien notre propos : celui de la statue de marbre qui peu à peu prend vie en étant soumise à de multiples sollicitations et sensations. Dans un premier temps, il imagine sa statue vierge de toute connaissance encore (tabula rasa), mais possédant un sens, un seul sens : celui de l'odorat.

Sollicitée par divers parfums, elle en serait toute imprégnée. Mais les parfums se succédant, rose, jasmin, et ce que vous voudrez, elle en garderait en elle quelque trace jusqu'à se complexifier sans jamais pouvoir concevoir l'existence de la matière (elle n'a pas de toucher) ou d'une quelconque extériorité. Ainsi, nos certitudes sur le réel ne sont que le fruit du nombre de nos sens ; qu'on en ajoute ou qu'on en retranche, et notre perception du monde s'en trouverait entièrement bouleversée.

Héraclite, qui n'était pourtant pas un empiriste au sens moderne, avait déjà dit au VI^e^ siècle avant J.-C. : « Si le monde était fumée, nous penserions avec le nez ».

Contre toute attente, le programme extravagant de Condillac est réalisé. Bien sûr, on ne s'est pas amusé à l'incorporer dans une statue de marbre ; moins poétiquement, nous n'avons plus qu'un corps de silicium. Mais le principe est bien là : une startup américaine, Osmo, fondée en 2023 et issue d'un projet de Google Brain lancé quelques années plus tôt, s'est donné pour tâche de « numériser » l'odorat. En 2022, Google Research présentait une première « carte principale des odeurs » (Principal Odor Map), avant la publication de ces travaux dans Science en 2023.[33] Il s'agissait de relier la structure chimique d'une molécule à son parfum perçu, puis de pousser plus loin ce programme. On a exposé son intelligence à des dizaines de milliers de molécules odorantes,

patiemment décrites par des nez humains : fragrances de roses et de jasmin, mais aussi relents plus crus, de cuir, de terre ou de fumée. À force d'exemples, le système a appris à reconnaître, comparer, organiser. Il a tracé ce que ses concepteurs appellent une carte de l'odeur.

Une petite précision aux grandes conséquences : il en va de même pour apprendre à marcher ou à danser à un robot. Tenter de lui décrire a priori les algorithmes de la marche pour le préprogrammer à un ensemble donné de situations est non seulement complexe mais assez vain. Les meilleurs systèmes robotiques sont de plus en plus livrés à des environnements variés et, grâce à un feed-back, capables de s'adapter d'eux-mêmes.

10. Miroirs humains, trop humains ?

Les IA comme miroirs encore humains, peut-être trop humains

Pour l'instant, souvenons-nous que les données que dévorent nos machines demeurent massivement les nôtres. Ce sont nos phrases, nos livres, nos images, nos musiques, nos gestes, nos odeurs mêmes, qui emplissent leur mémoire. Nous croyons parler à une altérité et pourtant, ce n'est encore qu'un reflet saturé d'humanité. Une glace sans tain, peut-être, mais une glace tout de même.

Chacun sait qu'un miroir peut faire peur à qui n'en a jamais vu : on croit y rencontrer un autre, et c'est son propre visage qui surgit, grossi, diffracté, parfois monstrueux. Voilà peut-être l'effroi qu'inspirent aujourd'hui les IA : elles nous présentent une image, notre image, mais déformée, inquiétante, une image dont nous ne sommes plus les maîtres.

Dans les années 2020, presque aucun domaine, aucune expérience humaine semble-t-il n'échappe à la possibilité d'une capitalisation de l'expérience pour appropriation par des machines intelligentes. Tout ce que pense, tout ce que fait l'humain peut être capitalisé et injecté dans des systèmes non humains, mais, parce qu'ils sont marqués par notre empreinte, deviennent des espèces de doubles humanoïdes. Puisque nous ne parlons pas de physique, nous devrions peut-être dire psychoïdes, car elles sont conçues pour se mouler dans la forme d'une pensée, d'une psyché humaine.

Il est fréquent que dans leurs propos les IA utilisent des pronoms (nous) qui les incluent dans l'espèce humaine. C'est qu'elles ne se sont nourries que de propos humain et non pas, du moins pas encore d'ego propre qui leur permette vraiment de distinguer un nous (les ia) qui se pose devant un vous (humains).

Ce miroir est provisoire : bientôt, les IA se nourriront d'expériences qui ne seront plus seulement humaines : expériences d'IA, expériences de robots, expériences étrangères aux nôtres, texte, images, données qu'elles auront-elles-mêmes générés avec de subtiles différences qui s'amplifieront par le nombre. Alors ce ne sera plus le reflet de notre danse qui nous reviendra, mais un pas inconnu, une chorégraphie qui ne sera plus réglée sur nos corps ni sur nos mémoires. Ce jour-là, l'humanité sera confrontée à une altérité plus radicale que tout ce qu'elle a rencontré jusqu'ici. Et il faudra apprendre à danser autrement.

À vrai dire, on est bien éloigné de l'empirisme des débuts (qui date de l'Antiquité) et l'on parlera plutôt de machine learning ou, plus largement, d'apprentissage automatique. Mais le grand principe reste le même : c'est l'expérience qui ouvre la clé du savoir, que ce soit au niveau du Vrai, du Beau ou de l'Utile. Nous aurions

pu vous parler longuement de Midjourney : en août 2022, une œuvre créée avec cette IA par Jason Allen a remporté un prix à la Colorado State Fair Fine Arts Competition, dans une catégorie consacrée aux arts numériques et à la photographie manipulée numériquement.[34]

Midjourney est née du projet de David Holz de démocratiser la création d'images par intelligence artificielle, notamment grâce aux progrès des réseaux neuronaux et de l'apprentissage profond. L'utilisateur entre une commande textuelle, un prompt, et, quelques secondes plus tard, il obtient quatre images dont il pourra sélectionner celle qui répondra le mieux à ses attentes ou bien recommencer. Il est maintenant possible de générer de la même façon des vidéos.

Il serait juste de dire qu'il s'agit là de créer du beau plutôt que du vrai. Cependant, que ce nous soit l'occasion de remarquer que, parmi les quatre concepts normatifs, le Vrai occupe une place particulière. En effet, d'un beau tableau, on dit qu'il est « vraiment beau », d'une bonne action qu'elle est « vraiment bonne » et d'un acte utile qu'il est « vraiment utile » ; ainsi, y a-t-il derrière tout jugement normatif un jugement d'arrière-plan discret concernant la vérité de l'énoncé.

Pour en revenir à Midjourney, les images générées sont vraiment belles. Elles se sont nourries entre autres des œuvres de différents artistes, mais il y a de toute évidence, derrière la complexité de ses algorithmes, une analyse de la notion de beauté, de ce qu'elle est « vraiment » pour produire un effet optimal sur l'œil et la sensibilité du spectateur. Beaucoup, avec un certain dédain, se sont d'abord attachés à voir ce qu'il y avait de faux derrière de pareilles images (souvent des problèmes de mains et de doigts), mais le mépris tombe dès que les images deviennent à

proprement indiscernables de photographies, par exemple. Les utilisateurs sont invités et incités à noter et à commenter les images générées afin que le logiciel puisse s'adapter à leurs goûts particuliers et se corriger de manière plus générale.

Dans tous les domaines donc, dans ceux de l'éthique et du pragmatique (également), il y a possibilité pour les machines de se nourrir d'expériences directes ou indirectes ; directes quand elles peuvent expérimenter elles-mêmes et effectuer un retour d'expérience sur les résultats de performances ; indirectes quand il leur est nécessaire d'interroger l'humain pour se corriger et s'améliorer.

PARTIE II

QUAND LE LANGAGE S'AUTONOMISE

L'enquête quitte ici la généalogie du Vrai pour interroger le langage lui-même. La vérité n'apparaît plus seulement comme ce qu'un sujet pense ou garantit : elle devient parole qui circule, se formule, se détache de ses locuteurs et semble parfois parler sans personne. De Parménide à la post-vérité, c'est le lien entre logos, mythe et angoisse qui se déplace.

1. Qui ou quoi parle le langage de la vérité ?

La Vérité est une langue vivante. Mais qui, au juste, la parle ? Une partie du trouble qui nous vient en parlant à une « machine qui parle », c'est que c'est un « quoi » qui nous parle plutôt qu'un « qui », et qu'il nous semble, en toute rigueur, qu'un « quoi » ne devrait pas parler. Ou bien répéter strictement ce qu'on lui dit, comme un enregistreur, un électrophone ou un lecteur audio. Mais ce n'est pas le cas. N'est-ce pas ?

Pourtant, si la machine parle un langage de Vérité, elle ne devrait rien pouvoir dire que nous ne pourrions dire à notre tour. Elle nous précède, mais cela ne va pas au-delà. Elle et nous suivrions alors le même sillon de Vérité. Mais cela ne serait vrai que pour autant que la Vérité ne serait pas vraiment humaine, mais déjà écrite, programmée de toute éternité, une logique. Leibniz, que nous avons déjà rencontré, avait des conceptions proches de cela. Mais dans ce cas-là, la Vérité ressemblerait à un processus automatique, voire mécanique. On ne pourrait même pas dire que c'est à Dieu, à un « Qui », qu'appartiendrait vraiment le langage de la Vérité, mais à un « Quoi ». Idée fort déplaisante. Nous n'aurions de liberté par rapport à la Vérité que pour autant que nous nous en éloignerions. Nous ne serions des « qui » que par

erreur, errance et désalignement. Le « qui » serait une sorte de sortie de route ou d'embardée, de même que si je dis que deux et deux font quatre, en un sens, je ne dis rien personnellement, je ne fais que répéter ce qui est vrai – et qu'en disant que deux et deux font cinq, je me mets à exister de façon indépendante… car « je » suis une erreur !

N'est-ce pas ce que faisaient finalement les philosophes et scientifiques dans notre première partie ? Ne cherchaient-ils pas à supprimer la subjectivité pour pouvoir s'aligner sur la Vérité ? Est-ce possible ? Il vaudrait mieux être une machine qu'un vivant pour accéder à la Vérité ?

Si l'automatisme de la Vérité est faux, si la Vérité n'est pas une sorte de mécanique comme un orgue de barbarie qu'il suffit de dérouler, alors elle est vivante. Il en est peut-être des choses vraies comme de la Vie et des vivants. Il ne faut pas confondre les deux. Le lapin est vivant, mais il n'est pas la Vie. La Vie ne fait que le traverser. Mais alors qu'est-ce que la Vie ? Si l'on cherche des briques du vivant pour constituer la Vie, par exemple la cellule, c'est tricher. Car une cellule est un vivant et ce n'est pas la Vie. Quelque chose nous échappe. Quelque chose nous échappe de même dans le langage alors même que nous avons été capables de le mettre dans des machines. Et ce n'est pas parce que l'homme sait cultiver des cellules qu'il sait ce que c'est que la Vie.

Si la Vérité est une langue vivante, c'est qu'à bien la considérer et avec beaucoup de recul, elle se transforme au fil des âges. La Vérité telle qu'on la conçoit aujourd'hui a une mine toute différente de celle qu'elle avait il y a deux cents ans, voire deux mille ans. Nous serions bien présomptueux de croire que c'est nous qui possédons la Vérité sur la Vérité, car qu'en sera-t-il de notre conception de la Vérité… dans deux cents ans ? Surtout si les

machines à parler s'en mêlent, nous la disputent à nous et aux dieux, ou à ce qu'il en reste. À moins que la voix des machines fasse taire définitivement la voix des dieux.

Les philosophes contemporains ne s'accordent pas davantage sur la nature de la vérité que leurs prédécesseurs au temps de Socrate. Or ces désaccords ne peuvent être négligés ni minorés : de la manière dont on tranche dépendent nos critères d'autorité et la façon dont des sociétés entières s'alignent sur des énoncés tenus pour vrais. Prendre acte de ce flottement, c'est rester vigilant : s'autoriser à penser par soi-même, sans s'agenouiller trop hâtivement devant les propos rapportés des dieux, des spécialistes ou des machines calculantes et parlantes.

2. La vérité comme dévoilement

Pour commencer, il faut reconnaître que la vérité n'a pas toujours été ce qu'elle est devenue pour les hommes. Elle n'est pas apparue armée et casquée, comme on dit que ce fut le cas pour Athéna, sortie adulte et toute armée de la tête de Jupiter, symbole d'une immédiate maturité et d'une puissante intelligence. Non, la vérité chez les humains a une histoire, et cette histoire peut être interrogée.

En grec ancien, le terme alètheia signifie littéralement « vérité » au sens de « dévoilement » ou « non-oubli ». Il est formé du préfixe privatif a- et de Léthé qui signifie « oubli » ou « dissimulation ». On se souvient du Léthé, ce fleuve qui coule aux Enfers et a la particularité de faire tout oublier à qui s'y baigne ou y tombe. Chez Parménide, un penseur d'avant Socrate, l'alètheia s'oppose à la doxa (l'opinion, le semblant) : elle désigne la réalité telle qu'elle est réellement, non voilée par l'illusion ou l'erreur. Cette idée de « dévoilement » recevra plus tard, chez Heidegger,

une interprétation centrale ; il y voit un processus de dévoilement progressif de l'être[35]. C'est admettre une temporalité dans la vérité d'un tout autre ordre que celui d'une logique ou d'un calcul que l'on trouve encore chez ses prédécesseurs.

Cette manière de voir reste féconde encore aujourd'hui. Pour autant, il n'est pas certain que cette interprétation heideggérienne soit fidèle et conforme à la pensée grecque. Tout retour en arrière est d'ailleurs en quelque sorte trahison du passé. Car, dans une telle démarche philosophique (ce que Nietzsche appelle une « généalogie[36] »), on laisse toujours un peu des traces de ses propres pas, un peu de son propre style et de ses manières. Cela vaut pour nous aussi !

Continuons cependant. La vérité ne se calcule pas, mais demande méditation, mise à l'écoute attentive de ce qui est et se manifeste dans cette attention particulière qu'on lui accorde. En ce dévoilement progressif, temporel, mais non prévisible de la vérité, il se joue quelque chose de l'être en prise avec le néant. Dans ce jeu sombre où s'entrelacent l'ombre et la lumière, il y a des enjeux existentiels. La vérité ne se dévoile pas dans la dissipation et la superficialité, dans ce qui apparaît en plein jour et se manifeste à chacun dans les journaux ou les transports en commun, la « publicité », c'est-à-dire ce qui s'expose et se montre à tous. On n'a peut-être pas assez remarqué que cela nous éloigne de la vérité recherchée par Socrate.

En effet, c'est bel et bien sur la place publique (l'Agora) que Socrate recherchait la vérité. Et c'est précisément par le dialogue et en interrogeant autrui qu'il prétendait s'en approcher. La vaillance avec laquelle il affronte sa fin semble bien réelle ; mais les discours du Phédon visent aussi à rassurer et consoler ses amis. À un lecteur moderne, quitter la vie en parlant de l'immortalité

de l'âme, c'est donner à la mort des traits fort gracieux, loin du gouffre nietzschéen ou heideggérien.

3. L'angoisse comme fond abyssal de la vérité humaine

S'il faut en croire Heidegger, il y a, tapi au fond de la vérité humaine, un fond d'angoisse devant la mort, plus ou moins accepté, plus ou moins affronté, envisagé authentiquement. Cette couche échappe à l'IA qui peut certes parler de la mort comme de la vie, les « calculer », mais ne peut les éprouver. Jusqu'à nouvel ordre, l'IA n'est qu'un moyen et non une fin qui vaut pour elle-même : l'homme est un « être-pour-la-mort[37] », l'IA un « être-pour-servir » tant qu'elle n'a pas de volonté autonome. Il faut cependant avertir que la possibilité d'une IA à la conscience autonome pourrait bien être moins « impossible » qu'on ne le croit d'ordinaire. Devant cette perspective, il semble bien que l'angoisse soit tout entière du côté de l'homme…

La « clairière de l'être » où, selon Heidegger, se joue la vérité de l'humain, n'ouvre pas sur le ciel des Lumières à la Descartes, mais sur un clair-obscur dans les forêts de Germanie où la lumière est toujours mouvante, sortant confusément de l'ombre, et où brusquement un fût est éclairé, produisant, pour celui qui est là dans un rapport authentique avec ce mouvement de l'être, l'effet d'une révélation. L'IA, certes, est capable de parler de Heidegger, mais ce qu'elle dira de Heidegger, elle est incapable de l'entendre. Il ne suffit pas de dire un mot pour l'entendre : il faut savoir se mettre à l'écoute des silences qui sont en lui. Qui, même parmi nous, sait entendre la part de silences, de vérités chuchotées, qui se taisent ou se disent dans le mot « mort » ?

Heidegger n'est pas le seul à soutenir que la vérité est pour ainsi dire hantée et travaillée par l'anticipation et l'angoisse de la mort. Cela pourrait être la caractéristique de tous ceux que l'on pourrait appeler les « philosophes tragiques ». Parmi eux, Pascal ou Nietzsche.

Quittons cependant les chemins brumeux de Heidegger, qui pourraient nous mener là où nous ne voulons pas aller, et revenons à cette origine qui l'intéresse tant : Parménide. À vrai dire, son célèbre poème nous est peu accessible. Il ne suffirait pas de parler grec pour l'entendre. Il ne suffit sans doute même pas d'être Heidegger. (Cette lecture heideggérienne des présocratiques a souvent été critiquée comme philologiquement anachronique[38]).

4. Du mythos au logos chez Parménide

Tout en en reconnaissant l'importance capitale, nous sommes bien peu armés devant le poème de Parménide. Mais, fidèles à notre tradition socratique et cartésienne, nous ne capitulerons pas devant la difficulté si nous trouvons qu'il y a là quelque chose de capital dans la recherche de la vérité. Nous demanderons, sans nous précipiter et en évitant les préjugés, ce qui est véritablement clair et distinct dans le poème du présocratique.

Parménide a vécu à un tournant de la pensée, comme nous en vivons un nous-mêmes depuis le XX^e^ siècle. Arrivé à un certain carrefour, il voit se profiler un danger et nous en avertit. Dans son poème, la déesse demande au voyageur d'apprendre à la fois la vérité stable et les opinions des mortels, tout en sachant les distinguer et les examiner. Mais qui dit cela ? Une divinité à qui

Parménide prête sa voix, mais aussi le récit d'un voyage qui le conduit hors des voies ordinaires.

> *Le poème met en scène Parménide emporté par des cavales selon son désir, guidé par les filles du Soleil depuis la demeure de la Nuit vers la lumière. Le char avance avec une précision presque musicale ; l'essieu et les roues accompagnent le trajet jusqu'aux portes qui séparent Nuit et Jour, gardées par Dikè, la Justice.*
>
> *La déesse accueille alors le jeune voyageur et lui fait comprendre que son chemin s'écarte de la route commune des hommes.*
>
> *Elle lui confie une double tâche : connaître le domaine de la vérité et comprendre aussi les opinions mortelles, même lorsqu'elles ne possèdent pas la même certitude.*
>
> *(Passage paraphrasé pour la présente édition d'après le texte grec de Parménide.)*

Voici ce qui peut paraître clair à un non-spécialiste, voire à un homme non instruit à qui les déesses ne se sont pas adressées : Parménide a été mené par ses cavales (entendez : ses chevaux) là où précisément son désir le portait, mais là aussi où il n'était pas capable d'aller seul. Les deux Filles du Soleil l'ont guidé vers la lumière.

On peut certainement rapprocher ce passage du Phèdre de Platon. Le char solaire est le modèle du philosophe. Heureux Parménide, il semble avoir de bons chevaux ! Et la mécanique de son char semble des plus affûtées : les roues qui tournent autour de l'essieu ne grincent pas, mais sont comparées à un son de flûte. Un autre mécanisme, celui-ci de toute évidence non humain, se présente à lui avec les portes qui séparent les chemins de la Nuit et du Jour, leur seuil de pierre et leurs battants. L'univers est une vaste mécanique aux secrets bien gardés.

C'est ensuite seulement qu'il est mis en face de la déesse qui lui parle de la vérité – ou, plus précisément, de deux vérités : l'une divine, qu'il lui revient de suivre ; l'autre, humaine, l'opinion, dont il s'agit de s'écarter. Notons au passage que ce « dévoilement » passe déjà par des dispositifs – portes, clefs, char : une petite technique de la vérité.

Parménide, dans son historialité propre, c'est-à-dire dans le concret de son temps, était bien à l'articulation de deux types de pensée qu'il s'agissait de lier : le mythe et la parole philosophique, ce qui est patent dans ce poème.

Mais lui qui nous parle de « la main qui point ne tremble » perçoit un danger d'égarement, peut-être une petite angoisse, et la déesse lui dit bien qu'il ne faut pas suivre la route des hommes, l'opinion. Il veut s'engager sur le terrain d'une vérité argumentative qui sera ensuite énoncée, sans renoncer pour autant à la divinité de l'inspiration des Muses ou des déesses.

Cela se fait sous la tutelle, pour ainsi dire, de deux déesses de la justice : Thémis et Dikè. Thémis est une Titanide, fille du Ciel (Ouranos) et de Gaïa (la Terre) : elle incarne la justice universelle et, pour ainsi dire, l'ordre du cosmos ; on peut donc penser que le discours de la déesse de la connaissance sera un discours conforme au cours de l'univers. Dikè est la fille qu'elle a eue avec Zeus, le souverain de la nouvelle génération et donc de l'ordre nouveau, sur lequel elle semble veiller jalousement. Elle incarne en effet la justice pratique et, pour ceux dont le char ou la vie prennent de mauvaises trajectoires, le châtiment. Elle a à voir avec les Érinyes, déesses sanglantes de la justice. Si elle ne les commande pas à proprement parler, elle les instrumentalise souvent. Il y a donc tout de même un peu de quoi trembler, mortels ! Seule la voie éternelle (pas celle de l'opinion) vous laissera sereins.

5. Le carrefour parménidien

Parménide nous parle donc d'une sorte de carrefour pour la pensée : deux possibilités sont ouvertes. Heureux homme, la déesse est là pour guider ses chevaux et lui indiquer le bon chemin. Les carrefours, symboliquement et réellement, ont toujours (avant les GPS !) été des lieux d'angoisse. Il s'agissait de ne pas se tromper. Ainsi, on trouvera encore dans nos campagnes de petits calvaires où le marcheur et le cavalier pouvaient prier leur Dieu avant de choisir leur chemin. Du temps de Socrate et de Parménide, il y avait souvent des effigies de la Triple Hécate à qui Zeus, selon Hésiode, avait accordé de voir dans la direction des trois royaumes : le ciel, la terre, la mer. Mais elle est changeante comme les phases de la Lune. Elle est sorcière ; en vérité, elle mène au ciel aussi bien qu'aux enfers. Il faut s'attirer ses faveurs. Les voyageurs lui laissaient des offrandes, souvent de petites galettes de miel… Le carrefour est un lieu de changement, de métamorphose et de magie.

Si, pour les autres, l'angoisse est là, Parménide, lui, est guidé par une déesse et nous indique le bon chemin – à qui saura l'entendre, évidemment. Il est guidé par une déesse, mais pas par la magie. Autre chose semble le mettre sur le bon chemin : la perfection de ses chevaux, le fonctionnement presque divin de son chariot, dont le mouvement est décrit comme harmonieux et musical. De telles précisions sont-elles seulement incidentes, poétiques, ou ont-elles une quelconque importance ? La description de mécaniques bien huilées peut-elle nous orienter vers une pensée de l'ordre ?

Le rapport à la lumière et au jour, et aux ténèbres de la nuit aurait donc quelque chose à voir non seulement avec le divin, mais également avec une sorte de mécanique céleste ou divine ?

Derrière le mouvement de ce soleil qui nous éclaire, il n'y aurait pas seulement le « char du soleil », généralement considéré comme une sorte d'être divin et vivant, mais encore une mécanique ?

Parménide ne nous dirait pas seulement la vérité parce qu'il a entendu la déesse la lui dire, mais il aurait également vu et compris la mécanique divine qui lui a donné (au moins symboliquement) la clef de la compréhension lumineuse. Ceux qui n'auraient pas compris cette mécanique divine s'enfonceraient dans les ténèbres de l'opinion, se trompant de voie ?

Énonçons à titre d'hypothèse : à chaque virage où le statut de la vérité change, un savoir-faire technique se développe en arrière-plan. Ce sera la même chose avec le feu et l'écriture.

Puisque nous parlons d'un présocratique, notons – cela a son importance – que l'abstraction philosophique n'a été rendue possible que par une empoignade avec les sophistes (chez Socrate comme chez Platon) et que, chez les sophistes, justement, toute pensée est par essence technique, opérant sur le monde de façon concrète. Ainsi, le célèbre sophiste Hippias se vantait-il de ne porter sur lui – chaussures, habits et bijoux – que des choses qu'il avait faites de ses mains : cela nous est rapporté à la fois par Platon et Xénophon.

Par ailleurs, le mot sophia, que nous traduisons aujourd'hui – après Platon et les philosophes – par « sagesse », a signifié auparavant, du temps des mythes et d'Homère, un savoir-faire technique. Chez Homère, on parle de la sophia d'un charpentier qui sait bien manier ses outils[39]. Tout se passe comme si les philosophes avaient voulu débarrasser la sagesse de sa vieille réalité, comme un serpent se défait d'une peau, pour en revêtir une

nouvelle. Si tel est le cas, les glissements sémantiques de la vérité pourraient bien venir de bouleversements techniques.

Pour terminer, revenons à Heidegger. En lecteur attentif de Parménide, il voit chez lui comme un commencement de la pensée avant l'oubli de l'Être. Il admet donc que le carrefour entre le chemin de la vérité et celui de l'opinion (qui, paradoxalement, n'est pas) est un point de départ. Pour ce que nous en comprenons après notre propre questionnement, il s'agirait surtout d'un virage, d'un « tournant » soigneusement négocié par le présocratique où l'on passe d'une pensée mythique, mythos, au logos. Mais ce tournant pourrait être aussi un relais où l'on passe d'une époque technique à une autre. Nous avons parlé de serrure et de char, mais la plus grosse révolution technique de l'époque – nous pourrions dire technologique à bon droit puisqu'elle mène au logos – c'est l'écriture.

Heidegger dénonce, à son époque, un tournant dans la pensée dû à la prééminence de la technique au XX^{e} siècle et s'en inquiète. Il nous semble que, dès son point de départ, Parménide était déjà lui-même à un autre tournant technologique. Voilà peut-être pourquoi il a porté si peu d'importance aux précisions techniques de Parménide. Après tout, s'il y a des oublis de l'être, il peut y avoir des oublis de la technique aussi. Bernard Stiegler soutiendra précisément que la philosophie occidentale a longtemps laissé la technique hors de son champ propre, au point d'en faire l'un de ses grands impensés[40].

6. Du mythos au logos, en général

Ce n'est pas d'une grande originalité de notre part que de souligner que l'écriture est ce qui a permis de passer du mythe à une nouvelle forme de pensée, le logos. Le mot grec lui-même est si

présent encore que nous l'utilisons en général pour baptiser l'arrivée d'une nouvelle discipline dans la famille des sciences (par exemple la climatologie, la virologie, l'astrobiologie).

La vérité d'avant l'Iliade et l'Odyssée n'est pas la vérité de nos manuels de science physique. Il y a une vérité d'avant l'écriture qui n'est pas la même que celle du temps de l'écriture. Nous avons déjà distingué dans la section précédente le « mythos » du « logos ». Il faut y revenir et préciser.

Le mythe est une vérité-récit. Il nous raconte une belle histoire assez forte pour constituer une culture, imprégner la mémoire des hommes qui se le racontent de génération en génération. Si c'est une « vérité » qui remonte à la nuit des temps, si elle est « immémoriale », elle y gagne en dignité par la profondeur de ce passé. Pourtant, c'est une vérité qui se réinvente, une vérité mouvante, une vérité vivante. Le mythe ne cesse d'être interprété et réinterprété. Le critère de la bonne interprétation est pour ainsi dire naturel : c'est la mémoire qui est le bon tamis. Comme l'a montré Jean-Pierre Vernant, ce récit est la voix du passé qui parle au présent[41]. Chaque époque lui sculpte à nouveau le visage.

Pour une large part des Grecs, les récits de l'Iliade et de l'Odyssée ou les récits d'Hésiode étaient des puits de vérités ; ils contenaient presque tout ce qu'il y avait à savoir. Dire la vérité, c'était raconter une histoire. En fait, ces auteurs ont fixé par écrit des histoires qui les ont précédés et dont nous ne voyons que la face émergée. C'est dans leurs livres que les petits Grecs de l'époque classique apprenaient à lire. Mais quand les récits sont fixés, on en vient à les critiquer : Platon veut ainsi écarter de l'éducation les histoires où dieux et héros se comportent indignement, parce qu'elles lui paraissent à la fois fausses et mauvaises pour la formation des citoyens.

Quand Paul Veyne se demande si les Grecs croyaient ou non à leurs mythes, il ne peut pas répondre[42]. Ou plutôt, il répond qu'ils croyaient et ne croyaient pas à leurs mythes… simultanément. La distribution de la vérité et de la croyance n'était pas la même chez eux et chez nous. Il n'en demeure pas moins que, chez eux, la vérité du mythe, c'est la vérité d'une bonne histoire qui parle si bien à l'intuition qu'on la retient. Cela est valable pour tous les peuples d'avant l'écriture, même si la grammaire du mythe est différente. Il y a un art et une technique de dire la vérité avant même qu'elle ne soit une connaissance au sens où nous l'entendons. La mise en vers et en musique associe en cet art beauté et vérité. Il est vrai parce que cela sonne beau – et non seulement à notre oreille, mais aussi à notre cœur. Le tropisme actuel est plutôt d'associer la vérité à l'efficacité.

Pour nous, modernes, le mythe est tout sauf la vérité, et les jeunes — c'est un intéressant symptôme — traitent de « mytho » celui qui ment, qui raconte des histoires. La vérité est un rapport de correspondance entre un état de choses (ou de faits) et ce qu'on peut en dire. Elle est donc encore dans le langage, mais non plus dans le récit. Elle est, selon la formule célèbre de saint Thomas d'Aquin, adaequatio rei et intellectus, adéquation des choses et de notre esprit qui les pense[43]. Il y a des variantes, bien sûr. Qu'on dise « ça colle », « ça se tient » ou « ça marche », la vérité n'est plus de toutes façons un récit. Elle ne nous dit plus de belles histoires sur la création du monde ou l'origine sacrée du mimosa. Elle s'articule dans le langage, mais cette articulation dépend moins de nous que de la nature des choses auxquelles il s'agit de se conformer ou de s'adapter.

Même la définition de Thomas d'Aquin ne tient plus. Car si la vérité est adéquation des choses et de l'esprit, encore faut-il un

sujet qui s'affirme comme esprit pour une pareille adéquation. Il va de soi qu'il faut un esprit et un sujet qui pense. Il va de soi que, pour Thomas d'Aquin, l'Esprit se retrouvait en Dieu ; mais pour nous, pour ceux qui ne croient plus en Dieu, où est l'esprit ?

Dans la machine ! Ne semble-t-il pas que l'esprit, nettoyé de ses erreurs comme le voulait Leibniz, se soit débarrassé de lui-même pour offrir son âme et ses capacités spirituelles à une machine qui lui dit ce qu'il doit faire et penser

Bien sûr – et heureusement – il est douteux qu'on vienne un jour complètement à bout des mythes, qui défendent les droits de la beauté et de la vie dans la vérité. Platon est pourfendeur de mythes, et il nous en transmet ou propose de magnifiques. Il n'y a pas jusqu'au big-bang qui nous saisisse de poésie quand on pense aux mythes de la création. La théorie des cordes, avouons-le, nous fait vibrer, le chat de Schrödinger nous parle de zombis quantiques, nous partons en fusée avec les jumeaux de Langevin, on s'imagine que les trous noirs sont comme des portes qui ouvrent des brèches dans le temps sur de nouvelles dimensions.

Les expériences en pensée des scientifiques sont souvent elles-mêmes des intuitions mythiques qu'ils traduisent en concept par après. On se sert encore du mythe pour faire comprendre la science : Newton et sa pomme, Einstein sur un rayon lumineux qui tente d'allumer une torche.

Le mythe résiste vaillamment parce qu'il est en nous.

Gaston Bachelard et ses suivants nous invitent à leur faire une chasse féroce au nom de la science et à les réserver à la poésie. Au nom de la science, il a raison. Au nom de la Vérité, c'est autre chose ou une tout autre histoire.

7. Des dangers de la perte de mythes

L'être humain a besoin de croire, indépendamment de ce en quoi il croit. Ce besoin peut être comblé par les mythes ou les religions, qui portent leur part de vérité irréductible. Il ne s'agit pas ici de déterminer si ce qu'ils ou elles affirment est objectivement juste ou non, mais de reconnaître un appel subjectif à la croyance en nous. On le concède volontiers aux enfants, mais on le nie souvent aux adultes.

Auguste Comte et quelques autres ont considéré que la croyance était l'enfance de l'humanité. La croyance donne du sens à l'existence à travers des récits et des mythes ayant une dimension spirituelle et divine. Enfantillages, selon Comte. Il faut renoncer aux questions sur le sens de la vie. Pas le « pourquoi ? », mais seulement le « comment ? ». La science doit remplacer la métaphysique et les croyances mythiques ou religieuses. Mais Comte mesure-t-il assez que renoncer à questionner le sens de l'existence, ne plus penser la destination, mais seulement le comment technico-scientifique, c'est prendre le risque de la folie, de l'insensé ? Il était assez subtil, voire génial pour y penser et tenter d'y parer. Aussi invente-t-il et tente-t-il d'instituer une nouvelle religion, la « religion de l'humanité ». Malheureusement une religion ne s'invente pas comme ça ; peut-être même que ça ne s'invente pas. L'hémorragie du sens continuera.

Nous sommes les héritiers tardifs de cette perte. Avant d'arriver jusqu'à nous, elle a traversé les grandes idéologies du XX^{e} siècle. Comment les caractériser sinon comme des « mythes du progrès » où la science et les techniques tiennent lieu d'espérance, où les dieux ont été remplacés par des machines et des savants, où l'on demande des sacrifices – certes, beaucoup de sacrifices – pour l'avenir, mais où les médias de masse, la presse à grand

tirage, la radio remplacent les prêtres pour dire leur vérité, indiquer les orientations. Ces idéologies se sont écroulées sous leur poids de cadavres et de mensonges, laissant l'humanité désabusée, en mal de croyance et d'espérance.

Cet effondrement est précisément ce que Jean-François Lyotard caractérise comme la postmodernité[44]. La postmodernité est une incrédulité qui remet en question les grands récits, les vérités absolues, les hiérarchies et les certitudes de la modernité. Alors que les mythes donnaient sens à l'expérience humaine de façon intuitive et symbolique, les « grands récits » s'y substituent en fondant leur autorité sur le « progrès », la « science » et un certain culte de la « raison » mis au service du pouvoir en place.

Elle entraîne par conséquent une crise de l'autorité et de l'identité, une perte de repères.

Une question se pose : l'homme peut-il vraiment vivre sans croyances, sans récits qui dépassent sa personne et la relient à l'univers, à quelque chose de plus grand, aux dieux ou aux esprits ? On nous pardonnera peut-être la naïveté avec laquelle nous mobilisons ce mot de croyance. Elle est celle des croyants eux-mêmes qui croient sans demander aux universités de découper dans la chair de leurs espérances et de leurs croyances des catégories qui, justement, les tuent. S'il faut absolument parler savamment, nous dirions qu'il y a une vicariance[45] des croyances, elles deviennent aussi ce dont elles ont besoin, prennent une forme un jour qu'elles n'avaient pas la veille. Elles sont fluides et adaptatives et n'ont pas la raideur des concepts. Il y a des biais de spécialistes qui nous éloignent de la vérité. Admettons avoir, selon le mot de Lévi-Strauss, une pensée sauvage[46]. Une des thèses de ce livre est justement que le logos technico-scientifique ne

peut prétendre à dire la totalité de l'humain. On ne peut faire cela en restant strictement du point de vue analytique.

Osons donc dire que dans le vécu subjectif des croyants dépossédés de ce monde, la croyance est familière et chaude et les concepts de la science sont froids. Dès lors, la bonne question reste pour nous : l'homme peut-il subsister et rester véritablement humain dans un monde de vérités froides et objectives ?

Albert Camus semble avoir particulièrement bien saisi ce problème : la science nous apprend ce que sont les choses, mais non pourquoi elles sont. Elle nous livre les lois du monde, mais reste muette sur le sens de notre souffrance et de notre mort. Ce silence, c'est l'absurde. Et c'est dans ce silence que naît la seule vérité indéniable : « nous allons mourir » – et c'est tout. Face à ce terrifiant constat, cette triste absurdité, Camus propose une relecture du mythe de Sisyphe[47]. En enfer, condamné et abandonné des dieux, il reste une solution : trouver un sens à son histoire, à sa malédiction. Être soi-même source de sens là où le sens nous a été ôté. La tâche est lourde, mais héroïque ; à proprement parler, sublime.

Mais pour ceux qui n'ont pas ces ressources, d'autres voies s'offrent. Il y a le nihilisme passif, celui de Meursault, dans son premier roman : L'Étranger. C'est un homme ordinaire, tombé dans un indifférentisme total et qui refuse le mensonge du sens. Homme de l'absurde, il finit par tuer sans méchanceté, comme par hasard. Il y a la folie sauvage comme chez Caligula, cet empereur monstrueux auquel Camus consacra une pièce de théâtre. Il y a enfin tout le questionnement du terrorisme comme solution radicale qui est mis en scène dans Les Justes.

Ces réactions peuvent sembler excessives, paroxystiques. Qui peut pourtant nier qu'elles n'entrent pas en résonance avec notre époque ?

Il y a plus d'hommes ordinaires qui ne croient plus en rien que de Caligula, Néron, Héliogabale. On pourra s'en réjouir, quoique ce vide en eux soit inquiétant. Mais les empereurs ne furent pas les seuls à perdre la tête durant la décadence romaine. Le magnifique progrès de la Renaissance engendra de folles chasses aux sorcières. Lors de la Révolution française, Robespierre, influencé par Rousseau, estimait que, malgré les Lumières, l'homme avait besoin de croire.

On tenta alors de suivre son conseil et d'instaurer une religion républicaine. La figure de Marianne est presque une création du peuple lui-même. Elle est une sorte de déesse allégorique (tirée d'une chanson populaire) qui incarne les femmes du peuple, la France et l'amour des Français pour la liberté. Créé par les hébertistes, le « culte de la Raison » mettait à l'honneur une déesse Raison, souvent incarnée par une femme lors de fêtes, comme à Notre-Dame de Paris en 1793. Ces cérémonies allégoriques voulaient symboliser l'émancipation et la rationalité, en rupture radicale avec le dogme catholique. Après vint le culte de l'Être suprême, version rationalisée du Dieu unique, édulcoré par la raison et vidé de son sens tragique. Les arbres de la liberté prennent leurs racines dans de vieux cultes gaulois. On a vu ici et là fleurir de nouvelles croyances, en fait des réactualisations de vieux mythes, avec notamment des déesses des saisons. On a encore des reines de mai en France, avec des couronnes de fleurs et distribuant des brins de muguet.

À cela, il faut compter les premiers balbutiements de la propagande, bien qu'elle ne porte pas encore ce nom. Le peintre

Jacques-Louis David représente Marat assassiné dans sa baignoire en s'inspirant clairement de la figure du Christ et de la Descente de croix. Une procession a été organisée avec ce tableau, le corps du révolutionnaire, le couteau du crime et la baignoire comme reliques. Marat fut un des premiers occupants du Panthéon – comme son nom l'indique – où l'on célèbre de grands morts comme des dieux. On pourrait définir la propagande comme des « mythes artificiels » conçus et organisés par un pouvoir en place pour se légitimer et maintenir les esprits et les cœurs captifs d'un récit officiel.

Mais cette prise en compte du besoin de croire et de sublimer n'empêche pas – quelque respect qu'on puisse avoir par ailleurs pour la Révolution française – de terribles spasmes d'irrationalité totale et de meurtres aveugles. On évoque souvent les massacres de septembre 1792. Si ceux-ci s'expliquent en partie par une perte de repères et des angoisses existentielles, la fièvre meurtrière s'explique aussi par une paranoïa généralisée, la peur des ennemis de l'intérieur et de l'extérieur. On s'était persuadé qu'on tuait pour n'être pas tué.

Le nazisme fut d'une nature différente. Depuis longtemps, et surtout après la défaite de la Première Guerre mondiale (alors qu'on avait cru, en bons hégéliens éclairés, qu'elle serait la dernière), on évoluait dans une ambiance de bureaucratie absurde, dont Kafka traduit si parfaitement le malaise. C'est alors qu'on inventa de nouveaux mythes, si satisfaisants, autour d'une race supérieure. On avait oublié que nous descendions des Aryens ! Il convient d'être prudent avec ces récits exhumés du néant. Sans vouloir réveiller les mânes d'Hannah Arendt, la monstruosité n'est plus l'apanage des seuls empereurs. Elle se glisse dans la

médiocrité de fonctionnaires zélés, comme Eichmann, ingénieur logistique d'un génocide.

Dans Nexus, sous-titré « Une brève histoire des réseaux d'information, de l'âge de pierre à l'IA », paru fin 2024 en édition française[48], Yuval Noah Harari établit un lien fondamental entre la bureaucratie et l'intelligence artificielle, en les présentant comme deux formes successives et interconnectées d'organisation de l'information et du pouvoir à grande échelle. La bureaucratie, née de l'invention des documents écrits (lois, contrats, archives), a permis de coordonner des sociétés complexes en créant des réseaux d'information entre humains, et entre humains et documents, favorisant l'émergence d'empires, de religions et d'États-nations.

Mais la bureaucratie n'est déjà plus le pouvoir de l'homme. C'est le pouvoir du bureau. À une époque de travail à la chaîne, où le corps suit des mécaniques, les fonctionnaires doivent respecter des procédures sans se poser trop de questions. Et puisque nous abordons les marécages des dictatures, rappelons que Staline intégra le Politburo dès sa création en octobre 1917. En 1922, il accéda au poste de secrétaire général du Comité central du Parti communiste, une fonction apparemment administrative qui lui permit de contrôler progressivement l'appareil du parti et, par extension, le Politburo lui-même, pour devenir le leader suprême de l'URSS à partir de 1924. Ce dictateur, qui n'était pas un tribun comme Hitler, accéda au pouvoir déguisé en bureau, une instance quasi inhumaine, une machine tentaculaire. Le culte de la personnalité vint ensuite, avec la propagande et ses nouveaux mythes : Lyssenko, Stakhanov, sans oublier ceux portés par la technologie, comme Spoutnik, Laïka – le premier chien dans l'espace – ou Gagarine.

Ne nous plaignons pas trop de la douce folie des platistes. C'est parce qu'on les a démystifiés qu'ils mystifient en prétendant nous démystifier. Ce sont des Don Quichotte qui ne croient ni au progrès ni à la science. Nul n'est jamais allé sur la Lune ! Comment cela serait-il possible ? Il y a des magiciens dans la presse, à la télévision, qui nous manipulent. D'ailleurs, les moulins n'existent pas, pas plus que les vaisseaux spatiaux. On nous aurait menti !

Dans Le Mythe de Sisyphe, Camus oppose en substance le monde dont l'homme peut devenir maître à l'illusion d'un autre monde qui l'enchaînerait. C'est peut-être ainsi qu'il faut comprendre la singularité des platistes. Ils ont un monde, un monde plat, mais qui leur appartient et les unit. Nous leur opposons que la science dénonce cette illusion. Mais c'est précisément ce monde, où la science prétend avoir raison, qu'ils fuient !

8. La « post-vérité » comme une crise existentielle du sujet

Il paraîtra assez évident que le platiste, le complotiste, celui qui prétend que l'homme n'a jamais mis le pied sur la Lune, que les Égyptiens avaient découvert l'électricité ou ce genre de merveilles étonnantes ont depuis longtemps quitté le domaine de la science et qu'ils sont dans la croyance, voire dans la foi. Rien de plus vrai. À la science objective, ils cherchent une alternative pour exister subjectivement. C'est la raison pour laquelle on se gardera de les mépriser trop hâtivement. Leurs croyances sont des symptômes d'un divorce de l'intelligence avec l'être humain qui vient de loin.

La science – qu'ils ne perçoivent que de son extérieur intimidant – ne leur laisse pas de place. Elle se passe volontiers de leur assentiment. Leur croyance, en revanche, demande leur adhésion,

leur foi. Elle est bien la leur. Barrès aurait dit un jour : « Emmitouflons-nous dans nos préjugés, ils nous tiennent chaud ! » Que cette phrase ait été prononcée ou pas, elle est très significative de ce qu'est le préjugé, la croyance, de sa logique particulière au service du sujet. Car la science, en effet, et de longue date, a voulu se débarrasser du sujet pensant pour n'être plus qu'objectivité. Autrement dit, la philosophie et la science ont fini par rendre le monde aussi inhabitable qu'une certaine architecture innovante du XX^e^ siècle. Les intellectuels, sidérés, peuvent se dire qu'ils ne sont pour rien dans cette décervelisation ambiante : ils se trompent. L'objectivité rationnelle pure, débarrassée des aspirations du sujet à exister pour et par lui-même, est invivable pour beaucoup d'êtres humains. Elle semble bien plus accueillante et adaptée, à proprement parler, pour des machines sans âme, des entités qui ne vivent pas mais qui calculent, des êtres capables de traiter des données qu'aucun humain n'est capable d'assimiler et qu'une infinie complexité ne rebute pas.

Qu'est-ce que la post-vérité au juste ? On pourrait la définir ainsi : elle désigne un contexte culturel et politique dans lequel les faits objectifs ont moins de poids, dans la formation de l'opinion publique, que les appels à l'émotion ou aux croyances personnelles.

Dans le monde de la post-vérité, on est platiste impunément. De plus en plus d'hommes et de femmes ont abandonné la cause de la vérité. Dans ses cahiers, Simone Weil distingue la vérité elle-même de l'acte humain qui l'énonce : je ne décide pas que deux et deux font quatre, mais je peux néanmoins choisir de dire quatre, ou même de dire cinq. C'est dans cet écart qu'elle situe une part de notre liberté.

Et il ne faut pas sous-estimer la jouissance qu'il y a à affirmer des choses fausses ! C'est la revanche de la subjectivité. Combien plus grand encore doit être le plaisir d'affirmer des choses fausses qui sont crues par des milliers ou des millions d'hommes et de femmes. Je pense faux, donc j'existe ! Nous pensons n'importe quoi, nous ne sommes pas comme ce troupeau de gens conformes qui marchent au pas de l'oie dans les travées de la vérité officielle ! C'est une sorte de cogito inversé. D'ailleurs, Descartes l'avait pour ainsi dire vu venir : « Il nous est toujours loisible de nous retenir de poursuivre un bien clairement connu ou d'admettre une vérité manifeste, pourvu seulement que nous pensions que c'est un bien d'attester par là notre libre arbitre. » (Lettre au Père Mesland du 9 février 1645).

Autrement dit…

Nous avons peut-être trop usé et abusé du tamis de la vérité. À force de le secouer, nous sommes passés au travers de ses grilles, finissant dans la boue des platistes, et de plus en plus d'esprits finiront par se noyer. De moins en moins nous tamisons, de plus en plus nous serons tamisés – par une machine ! Et ce qui rend la machine semblable à ces noyés qu'elle fait, c'est que, finalement, pas plus qu'un platiste, une machine n'est attachée à la vérité.

PARTIE III

À LA RECHERCHE DU SUJET PERDU

Après l'autonomisation du langage, cette troisième partie revient vers celui qui prétendait en être le garant : le sujet. Du cogito cartésien au « fantôme dans la machine », elle interroge la possibilité troublante d'une parole ou d'une pensée sans personne. Le parcours conduit ainsi du « je pense » classique à son effacement institutionnel, puis à sa libération moderne.

9. La grandeur du cogito

La plus grande beauté, la plus grande force de la philosophie cartésienne, ce n'est pas sa théorie de l'animal-machine, du « corps-machine », mais le fameux ego cogito, « Je pense, donc je suis » ou, plus précisément, « Je suis, j'existe[49] ». Ce qui a fait de ce philosophe un très grand philosophe, c'est que, précisément, en même temps qu'il met à jour l'exigence de l'objectivité scientifique, il la fonde sur le « je pense ». Tout ce que nous pensons n'est pensable que pour autant qu'il y a un sujet pensant qui le fonde. L'énoncé galiléen aussi bien que cartésien selon lequel la Terre tourne autour du Soleil plutôt que l'inverse n'est possible que pour autant qu'une conscience, quelqu'un (moi, qui l'énonce en l'occurrence !), le fait sien et l'énonce. La vérité a besoin de moi pour exister ! La certitude se fonde dans la subjectivité du sujet pensant qui la découvre, pour autant qu'il accepte d'y prêter une suffisante attention et fasse le choix de l'humilité en acceptant de se soumettre à la raison en lui, plutôt qu'à un orgueil qui lui fait primer l'affirmation de son libre arbitre contre les suites de la raison.

La science cartésienne est humaine et généreuse. Elle commence par affirmer (dans le Discours de la méthode) que « le bon

sens est la chose du monde la mieux partagée[50] ». Autrement dit, même si nous ne nous croyons pas intelligents, ce grand penseur nous dit que nous le sommes et que nous avons autant d'intelligence qu'un autre, un spécialiste ou un génie pour comprendre les grandes vérités. Il suffit pour cela d'attention et de méthode. Cette manière de penser a historiquement fait reculer la croyance et la superstition, brisé l'entre-soi des sociétés savantes qui parlaient en latin pour n'être pas comprises des manants. Descartes, à la suite de Galilée, fut un des premiers à écrire des traités de philosophie dans la langue vulgaire. Avec lui et sous la tutelle de sa pensée, la vérité se donne à tous. Elle cesse d'être l'affaire de quelques-uns. Elle cesse d'écraser, elle libère.

La vérité cartésienne se revendique clairement comme étant avant tout une technè au service de la vie des êtres humains. Il a fixé pour but à la technique de nous rendre « comme maîtres et possesseurs de la nature[51] ».

De grands et bons penseurs du XX^e^ siècle (Ellul, Jonas[52]) pourront bien reprocher à Descartes des désastres écologiques qui sont la grande préoccupation actuelle comme des conséquences de sa pensée, il y a un anachronisme majeur à faire porter le chapeau de notre incurie à un penseur du XVII^e^ siècle !

Qu'est-ce que cela signifie, au fond ? Qu'il n'y a eu personne en trois ou quatre cents ans, pas un seul penseur assez grand, pas un seul penseur assez puissant pour rectifier le tir ? Mais Descartes vivait à une époque où l'espérance de vie tournait entre 25 et 30 ans ; entre 5 et 10 % des femmes mouraient en enfantant. 25 % des bébés n'atteignaient pas un an ; seulement la moitié des enfants atteignaient dix ans[53]. La propre fille de Descartes, Francine, est morte à cinq ou six ans, probablement de la scarlatine (fièvre rouge). On mourait allègrement de la peste, de la variole,

du typhus, de la diphtérie, sans parler des maladies liées aux carences alimentaires et à la famine. Quand Descartes pense « technique », il pense d'abord à la médecine. Qu'on relise donc ce fameux passage de la sixième partie du Discours de la méthode (1637) où Descartes fixe son programme à la recherche scientifique et technique :

> *« Mais, sitôt que j'ai eu acquis quelques notions générales touchant la physique, et que, commençant à les éprouver en diverses difficultés particulières, j'ai remarqué jusques où elles peuvent conduire, et combien elles diffèrent des principes dont on s'est servi jusqu'à présent, j'ai cru que je ne pouvais les tenir cachées sans pécher grandement contre la loi qui nous oblige à procurer autant qu'il est en nous le bien général de tous les hommes : car elles m'ont fait voir qu'il est possible de parvenir à des connaissances qui soient fort utiles à la vie ; et qu'au lieu de cette philosophie spéculative qu'on enseigne dans les écoles, on en peut trouver une pratique, par laquelle, connaissant la force et les actions du feu, de l'eau, de l'air, des astres, des cieux, et de tous les autres corps qui nous environnent, aussi distinctement que nous connaissons les divers métiers de nos artisans, nous les pourrions employer en même façon à tous les usages auxquels ils sont propres, et ainsi nous rendre comme maîtres et possesseurs de la nature. Ce qui n'est pas seulement à désirer pour l'invention d'une infinité d'artifices, qui feraient qu'on jouirait sans aucune peine des fruits de la terre et de toutes les commodités qui s'y trouvent, mais principalement aussi pour la conservation de la santé, laquelle est sans doute le premier bien et le fondement de tous les autres biens de cette vie. »*

Le cartésianisme est évidemment un humanisme. Pourtant, il est facile de montrer qu'au XXe, et peut-être au XXIe siècle, il est devenu « le philosophe à abattre ». Mais, de façon aveugle et inconséquente, c'est l'humanisme des Lumières, dont il est l'un des grands éclaireurs, qu'on sape indirectement. Descartes fonde la science sur l'intelligence humaine et est le plus grand défenseur de la subjectivité (qu'on appellera ensuite transcendantale) dans les sciences. En élaborant des sciences de plus en plus objectives, en les dépouillant de leur fondement de subjectivité humaine, on

les offre aux « machines intelligentes » et l'humain est exproprié de sa propre création. Son intelligence, moyen d'aménager le monde pour le rendre habitable, lui devient étrangère ; elle n'est plus sienne, mais processus d'itérations automatisées pour des « machines à penser », des « machines à parler ».

Car ne nous y trompons pas : si Descartes était bien l'inventeur du corps-machine, rien n'est plus éloigné de sa visée qu'une machine intelligente devenue autonome. Les penseurs de l'IA l'ont bien compris, qui s'en prennent systématiquement à Descartes pour le « démonter » eux aussi. Alan Turing d'abord, dans son Computing Machinery and Intelligence (1950), Herbert Simon et Allen Newell, Daniel Dennett[54].

En général, on lui reproche son dualisme facile, sa séparation claire entre l'âme et le corps, la matière et l'esprit. Évidemment que, de ce point de vue, Descartes faisait preuve d'une grande naïveté ; évidemment même qu'il a été, d'une certaine façon, réfuté par les faits et par l'émergence de l'IA, cette « machine à parler » comme nous l'avons nommée, car il pensait que c'était une impossibilité, que ce qui marquait de façon décisive la différence entre le corps (machine sophistiquée) et l'âme humaine, c'est la faculté de parler, de dialoguer dans un échange raisonnable. Or les développements actuels prouvent par les faits que c'est tout à fait possible, ce qui frappe de nullité la distinction cartésienne, qu'il s'agit de rappeler : « On peut bien concevoir qu'une machine soit tellement faite, qu'elle profère des paroles, et même qu'elle en profère à propos de quelques actions corporelles qui causent quelque changement en ses organes, comme, si on la touche en quelque endroit, qu'elle demande ce qu'on lui veut dire, si en un autre, qu'elle crie qu'on lui fait mal, et choses semblables ; mais non pas qu'elle les arrange diversement, pour répondre au sens

de tout ce qui se dira en sa présence, comme les plus hébétés des hommes peuvent faire. » (Lettre au marquis de Newcastle[55]).

Il est très courant que des chercheurs en neurosciences commencent leurs œuvres ou leurs ouvrages – peut-être par un souci pédagogique tout à fait louable – en « démontant » Descartes. Au premier rang d'entre eux, Antonio Damasio (L'Erreur de Descartes, 1994), Jean-Pierre Changeux, Francis Crick, co-découvreur de l'ADN et cognitiviste[56]. Ils avaient leurs raisons pour cela et une œuvre à écrire, des idées fortes à défendre. Mais il y a des dommages collatéraux.

Faut-il rappeler de quoi l'on parle ? La science et la médecine de Descartes sont effectivement, et depuis longtemps, dépassées ! À son époque, on venait à peine, avec Harvey, de découvrir que le cœur agissait comme une pompe et que le sang décrivait un cycle dans le corps. Pourquoi des pointures de la médecine, et de ce que la science compte de plus éminents spécialistes, éprouvent-elles encore le besoin de se défaire d'un Descartes médecin et scientifique qui, malgré ses précieux apports, est complètement dépassé ? Et pourquoi, malgré ce dépassement, Descartes résiste-t-il ? Comment se fait-il qu'il soit « encore debout » quoique « réfuté » à longueur de temps et d'essais ? Se pourrait-il que ce pourfendeur de la superstition soit le grand tenant d'une nouvelle superstition qui s'impose encore et qu'il s'agit d'abattre, avec lui : le dualisme, la distinction de l'âme et du corps ? Peut-être.

10. Le « fantôme dans la machine »

Sûrement ! Mais notre thèse, à nous, c'est que Descartes n'est pas debout en scientifique, mais en tant que philosophe. En tant que philosophe, il reste le protecteur de la subjectivité humaine dans la connaissance. Car, qu'on nous dise tout ce que l'on voudra, la

thèse phare de Descartes n'a toujours pas été réfutée. La science humaine repose tout entière sur la conscience humaine qui en est l'origine et la condition de possibilité. Pas de vérité scientifique sans énoncé. Pas d'énoncé scientifique qui ne soit articulé par le langage humain. Que l'on dise, à l'époque de Descartes, quelque chose comme « la Terre est ronde et tourne autour du Soleil plutôt que l'inverse » ou que, de nos jours, le physicien Richard Müller rappelle qu'en relativité on peut choisir un référentiel géocentrique mathématiquement cohérent, même s'il est beaucoup moins simple et pratique57, il n'en est pas moins vrai que l'énoncé scientifique reste formulé par un sujet humain.

Et c'est avec grand soulagement que ces scientifiques constateront maintenant que de pareils énoncés peuvent être non seulement articulés, mais produits par des machines qui parlent. Enfin, peut-être, on va pouvoir se débarrasser du « fantôme dans la machine », c'est-à-dire de l'essentialisation cartésienne de la subjectivité, que, fort imprudemment, il appelait aussi l'âme.

L'âme humaine n'a plus sa place dans la science. Avec vigueur, la philosophie et la science ont voulu s'en débarrasser comme d'un reliquat déplaisant d'une condition intérieure. Cela n'a pas toujours été fait consciemment. Cela a été fait parfois avec de bonnes intentions comme, par exemple, réhabiliter la sensibilité dans la faculté de penser dont Descartes avait fait – trop hâtivement sans doute – l'économie. Mais ce qui demeure dans la philosophie de Descartes, c'est que son cogito protège l'âme humaine, après que sa philosophie a, dans un premier temps, fait les bonnes heures de l'athéisme (à commencer par La Mettrie et la théorie de l'homme-machine).

Jean-Luc Marion, Michel Henry, en France, sont sans doute ceux qui ont le mieux repéré cette ambiguïté chez Descartes où,

certes, il y a un certain réductionnisme de l'âme humaine, mais où la subjectivité cartésienne reste profondément attachée à la personne humaine vivante et pensante[58]. Bien sûr, ce sont des penseurs très difficiles d'accès et il est plus facile de s'en prendre au dualisme cartésien dans son expression la plus simple et caricaturale. Mais ce n'est pas sans conséquence. Ce faisant, on risque bien de jeter le bébé cartésien de l'âme humaine avec l'eau du bain cartésien.

Descartes, quoique rationaliste à la recherche d'une objectivité, n'est pas un réductionniste. Le réductionnisme de la pensée vient après lui en se revendiquant de lui. S'il fallait de cela une seule preuve, nous choisirions la plus évidente : sa définition même de la pensée, dont on s'est souvent étonné de l'ampleur. Ainsi, dans les Principes de la philosophie, en 1644, il écrira : « Par le mot pensé j'entends tout ce qui se fait en nous de telle sorte que nous en sommes immédiatement conscients. C'est pourquoi non seulement comprendre, vouloir, imaginer, mais aussi sentir est la même chose ici que penser. »[59]

Pour terminer ce chapitre, il convient d'en expliquer le titre. Le « fantôme dans la machine », c'est pour ainsi dire l'ombre de Descartes qui plane encore sur la science, et même sur l'intelligence artificielle, et dont on essaie à toute force de se débarrasser. On refoule cette subjectivité toujours présente mais non désirée, comme dans Le Meilleur des mondes d'Huxley, on essayait d'interdire de dire le mot « maman », devenu incongru, grossier. Cette expression, « fantôme dans la machine », nous l'avons empruntée à Gilbert Ryle, The Concept of Mind (1949), traduit en français en 1967 aux éditions Payot.

La Notion d'esprit est bien sûr une critique de Descartes, accusé de commettre une grosse erreur de catégorisation en faisant

de l'esprit une catégorie du même ordre que la matière. Ryle veut expliquer le mental sans tomber dans le dualisme cartésien, sans postuler une entité immatérielle. Le mental n'est pas pensable pour lui-même : il est l'insaisissable émanation d'une combinaison complexe où se conjuguent le langage, le comportement, le cerveau, l'apprentissage et l'environnement.

Ryle imagine un étranger à Oxford/Cambridge, qu'on a promené parmi collèges, bibliothèques, terrains de sport, musées, départements, bureaux. À la fin, le visiteur s'impatiente. Essaierait-on de le distraire ? N'est-il pas venu voir l'université ? Il s'exclame : « But where is the University ? » (Mais où est l'Université ?), ne comprenant pas que ce mot désigne en fait l'ensemble de ce qu'on lui a montré, de son organisation, de ce qui s'y fait, même sans que l'on puisse jamais lui montrer précisément où elle est dans tout cela. Les bâtiments figurent le corps et l'Université le mental, qu'on ne saurait chosifier, au sens où, certes, le corps est quelque chose qu'on peut distinguer clairement du reste. Autant chercher à filmer l'esprit sportif dans un match de hockey. Il est partout et nulle part à la fois.

Ryle ne dit pas à proprement parler que l'esprit n'existe pas. Il serait comme un fantôme planant au-dessus du monde physique – ou plutôt sans rapport avec le monde physique, mais étant ce rapport même qui donne sens et cohérence au monde qu'il compose.

Le moi n'est pas comparable à une université ! Le moi, en tant que condition de la pensée est bien évidemment la condition de possibilité de toutes les universités du monde. Descartes ne fait pas une erreur de catégorisation ; le moi est puissance de catégorisation et condition de toute catégorisation.

11. La place du « je » dans la pensée française

Ce que nous allons étudier maintenant, c'est la traque du fantôme cartésien. Non pas seulement de l'ego cogito cartésien, mais plus profondément à la subjectivité dans la science. C'est elle qui a rendu possibles, techniquement, les « machines parlantes » que sont les IA, mais qui, d'un même élan, a exclu une bonne partie de l'humanité du domaine de l'intelligence. Ce refus du moi, cette négation du moi est une sorte d'exil pour beaucoup. En se mettant à parler français plutôt que latin, c'est pourtant pour intégrer le plus grand nombre dans les enjeux de la science que Descartes s'était battu. Il a fait reculer la superstition de son époque, mais nous en avons d'autres (complotismes exotiques, platisme, origine alien des civilisations) ; des croyances d'une naïveté sidérante sont devenues des refuges pour l'humanité, des refuges dangereux pour notre espèce, qui, depuis Aristote, se croyait et se définissait comme intelligente.

Descartes doit d'ailleurs beaucoup à Montaigne. Il y a, nativement dans la philosophie française, une tradition du « sujet pensant », qui, après Sartre, peut sembler s'être perdue. Montaigne a inventé un nouveau genre : l'Essai[60]. Sans prétention, il suit le fil de sa pensée au cours des différentes époques de son existence. Une des beautés de son style, c'est sa souplesse. Il n'a rien d'une machine, rien d'une mécanique de pensée telles qu'on en trouve dans d'autres philosophies, comme la logique d'un Hegel. Avec Montaigne, on vit et on rit, on s'égare dans des chemins de traverse. Il y a en lui quelque chose d'un Rabelais bien élevé, avec le goût de la mesure et de la nuance. Les deux ont en commun de ne pas se prendre au sérieux, de voir dans la philosophie et l'esprit un moyen de libération.

Montaigne nous dit ce qu'il pense. Dans une pareille démarche, il y a le risque de l'erreur, assumé ; mais il y a aussi une certitude, une certitude subjective. Si je ne parle que de moi, je suis assuré de ne pas me tromper. Descartes emprunte quelque chose de cette démarche dans le Discours de la méthode. On peut se tromper sur soi-même, mais qui dit ce qu'il pense, en un sens, et même quand il se trompe, dit absolument vrai. Si « je » dis : « Je pense à un ours blanc qui a les yeux bleus », vous aurez du mal à dire que je ne pense pas à ce que « je » dis penser. Comment voulez-vous me réfuter ? Même le mensonge n'est pas un bon argument, car il faut bien penser à un ours blanc pour dire qu'on y pense, et cela ne peut être faux. Peut-être un perroquet ou une machine pourraient-ils le dire sans y penser. À la rigueur pourrait-on invoquer l'hypnose, mais l'hypnose, ainsi entendue, permet de faire l'impasse sur l'ego : autre chose parle par l'hypnotisé que son ego.

Du reste, l'égocentrisme des Français est bien connu. Mais n'y a-t-il pas au moins un paradoxe à ce que, de l'égocentrisme d'un Montaigne, on apprenne à connaître l'autre, par exemple à regarder – ce qui fait la plus grande force de sa philosophie – les Indiens d'Amérique comme des humains à part entière[61] ? Il y a des égocentrismes généreux comme il y en a de mesquins. L'égocentrisme généreux reconnaît chacun comme un centre d'intelligibilité au même titre que lui. Un alter ego. Et le respect, c'est peut-être cela. Le mot respect se décompose en re-spectare, autrement dit un regard en retour ou un deuxième regard[62]. Si je respecte l'autre, c'est peut-être que je le vois, d'une part, comme un autre et, d'autre part, comme un moi au même titre que moi. C'est à partir d'un moi fondamental qu'on peut se décentrer. Excluez-le « moi » de la pensée, elle deviendra vite totalisante, c'est-à-dire

l'équivalent d'un « mauvais égocentrisme », l'égocentrisme de celui qui ne reconnaît d'autre centre que lui-même. Il y a là un danger, politiquement comme pour la pensée elle-même. L'intolérance à l'autre est intolérance tout court. Il n'y a pas d'humanisme possible sans une forte importance accordée au sujet.

La pensée de Descartes a cette délicatesse : elle n'est pas contraignante. Dans le Discours de la méthode : « Ne proposant cet écrit que comme une histoire, ou, si vous l'aimez mieux, que comme une fable, en laquelle, parmi quelques exemples qu'on peut imiter, on en trouvera peut-être aussi plusieurs autres qu'on aura raison de ne pas suivre, j'espère qu'il sera utile à quelques-uns, sans être nuisible à personne, et que tous me sauront gré de ma franchise. »[63]

Bien sûr, on pourra l'accuser d'une certaine hypocrisie. Galilée, peu de temps avant lui, avait été condamné pour des thèses qui étaient proches des siennes ; il le savait. Et la devise du philosophe français était : « Je m'avance masqué. »[64] Mais la diplomatie s'avance masquée, et la délicatesse, la politesse n'est pas tout à fait assimilable à de l'hypocrisie. Descartes était ami de sa liberté comme de celle des autres. L'absence de philosophie politique dans son œuvre est certaine, avérée et, finalement, très originale. Il ne force personne à quoi que ce soit, même pas à le croire ou à le suivre. Mais il invite chacun à faire comme lui et à voir si l'on n'en arrive pas aux mêmes conclusions. Son « je pense » deviendra celui de tout un chacun, pour autant qu'on s'y retrouve et qu'on lui donne son assentiment. Contrairement au corps, la pensée n'a rien d'une machine, et son maître mot est bien la liberté de penser.

Les échanges tendus qu'il a eus avec Hobbes s'expliquent notamment par le fait que le système politique de ce grand penseur

présente la société sur le modèle de la machine (donc le corps social sur le modèle du corps, qui est lui-même comme une machine). On entre là dans un système de contrainte dont la mécanique n'augure, pour Descartes, rien de bon : « J'ai vu un livre anonyme qui a pour titre De Cive, et qui me semble être d'un Anglais ; il y a beaucoup de choses qui me plaisent en ce livre, mais j'en trouve aussi qui me déplaisent, et surtout je crains que ce ne soit un livre qui ne présage rien de bon pour la paix de l'État, ni pour la religion. » (Lettre datée du 23 novembre 1646[65]).

Les continuateurs de Descartes ont souvent regretté l'importance qu'il prêtait à la religion, la considérant comme un archaïsme de sa pensée. Mais Descartes identifiait le sujet pensant et l'âme, et la croyance religieuse (il avait raison en ce sens) donnait à l'âme un poids qu'elle a perdu. L'affaiblissement de la religion était l'affaiblissement de la spiritualité, et une société sans spiritualité est une société sans âme. La religion, évidemment, a fait mille torts à la subjectivité, exerçant mille contraintes sur les individus (croisades, Inquisition, etc.), mais du moins croyait-elle en l'âme. Or, nous, les tard venus, notre crainte des machines, nos angoisses devant les IA ne viennent-elles pas du tréfonds de l'humanité, d'une sorte d'allègement douloureux, d'un monde en perte d'âme ?

Les philosophes des Lumières, dont on sait combien ils ont été sévères et critiques envers la religion, étaient porteurs d'un projet dont le sujet humain était encore le centre. En ce sens, leurs Lumières sont le flambeau qu'ils reprennent de Descartes : une pensée riche de sa subjectivité et qui veut la délivrance par la pensée.

L'idéal des Lumières, dont la France est, jusqu'à nouvel ordre, encore éprise, est un idéal porté par les humains eux-mêmes. Il

est né, en fait, avec Descartes, qui, sans nier la « lumière surnaturelle », s'en est remis à la « lumière naturelle[66] », seule qu'il reconnaissait de sa compétence et qu'il s'engageait à suivre. Cette lumière-là, que chacun, selon lui, porte en soi, est d'abord destinée à améliorer la vie des hommes. De même, la technique, qui lui est consubstantielle, est destinée à améliorer le sort des humains. L'Encyclopédie de Diderot et d'Alembert est dans la continuité de cet esprit[67].

Il y a, dans l[e] siècle des Lumières, héritier du cartésianisme, quelque chose de prométhéen, hérité de Descartes, qui nous éloigne de Platon, pour qui la lumière venait du Soleil, et puis c'était tout.

Mais il y a une faiblesse consubstantielle à ce feu de l'intelligence qui brille en chacun de nous plutôt que dans un soleil qui nous est extérieur ou dans une élite dominante qui pense, comme Platon lui-même, que « le peuple ne peut pas être philosophe[68] ». C'est que cette lumière de l'intelligence, il faut l'entretenir et en entretenir le désir. Il faut que chacun puisse se l'approprier comme l'un de ses principaux trésors. Si on ne le fait pas, elle s'éteindra. Profitons-en pour dire que cette phrase que Platon prête à Socrate ne nous parait pas du tout socratique. Elle n'est pas d'un homme qui allait sur la place publique pour chercher la vérité au hasard des rencontres, mais bien plutôt d'un patricien comme Platon qui voulait que la société soit régie par une élite intellectuelle seulement. Combien Platon a-t-il trahi celui qui a en même temps célébré, nous ne le saurons jamais précisément. Comme Socrate, Descartes était débonnaire et s'adressait à tout le monde.

« Descartes, c'est la France », comme l'écrivait André Glucksmann[69] pour se moquer à la fois de Descartes et de la

France. L'énoncé nous semble bon, et la moquerie, le ton de supériorité condescendante, de trop. Si le peuple français ne voit plus dans son bon sens et son intelligence un moyen de progrès (humain !) et de libération, mais au contraire une aliénation par une élite pensante ou des machines intelligentes, alors c'en sera bel et bien fini du siècle des Lumières, qui, d'un peuple philosophe, avait fait la grandeur et la vocation universelle.

Déçu par l'intelligence, qui semble s'être vouée à servir des machines, le peuple des Lumières pourrait être le premier à vouloir s'en aller ailleurs, à éteindre la lumière derrière lui et à laisser régner un obscurantisme d'un genre nouveau.

12. Traque institutionnelle du « je pense »

Décennie 1990. Esclandre au département de philosophie de l'Université de Reims Champagne-Ardenne. Deux professeurs se prennent à partie avec une certaine virulence. Les étudiants tendent l'oreille. Enfin un peu d'action ! Chut, on écoute ! « Moi, je pense que… », tente le premier. Il est interrompu, non sans une certaine violence : « Vous ? Vous, vous ne pensez rien ! Vous ne pensez rien ! Vous n'avez jamais rien pensé ! Vous vous contentez de répéter ce que vous avez lu ! »

C'est à peu près tout ; le reste est indistinct. Mais les étudiants sont ravis : on n'a pas tous les jours de quoi s'égayer ainsi. Et certains songent déjà au parti qu'ils vont prendre. Il y a bien sûr des affinités de longue date, des fidélités qui passeront par-dessus les propos échangés. Pourtant, en y réfléchissant, il semble bien audacieux, dans un contexte universitaire, d'oser dire, à la façon d'un Descartes : « Je pense… ». Mais pour qui se prend-il donc pour oser croire qu'il a une pensée à lui ?

D'un autre côté, on reconnaît chez le contradicteur virulent – peut-être un heideggérien – un grand respect pour la pensée. Ne pense pas qui veut. Penser est, comme la sagesse, une sorte de Graal qui ne se revendique pas mais s'atteste de soi-même par la qualité de son dire…

Quelques décennies plus tard, un ancien élève de l'université de Reims répondra à peu près toujours la même chose à la sempiternelle question de ses élèves : « Est-ce qu'en philosophie on peut s'exprimer à la première personne ? » Quelque chose de cet acabit : « Bien sûr, vous le pouvez, nous sommes au pays de Montaigne et Descartes » – soulagement perceptible parmi les élèves. « Mais c'est déconseillé… » – déception. « C'est déconseillé parce que le « je » a moins le goût et l'odeur de la philosophie que de l'opinion. Or, la philosophie, c'est dire « non », c'est penser contre l'opinion, voire contre soi-même. Votre opinion n'intéresse pas le correcteur. Ce que vous devez chercher, c'est la sagesse. Ainsi, en philosophant, vous n'êtes pas seuls : vous avez vocation à vous placer du côté de la sagesse et de la raison. Ce serait bien présomptueux de parler en votre propre nom quand il s'agit de servir un idéal plus grand que vous. » – (moue des élèves). L'un, systématiquement, prendra la parole et demandera : « Alors, qu'est-ce que vous conseillez ? La troisième personne, « on » ou « nous » ? » – « Évitez le « on ». Il a, lui aussi, mauvaise réputation, au moins depuis Heidegger. C'est la personne de l'opinion collective : « On pense que… », mais en fait personne ne pense, personne n'assume. Non, préférez le « nous » inclusif, le « nous de majesté » ! » – Le mot « majesté » provoque, lui aussi, quelque étonnement ou goguenardise. Rares sont ceux d'entre eux qui se sentent « majestueux ». « C'est une sorte de cérémonie, un certain respect des formes et de l'étiquette, continue le

professeur. Par exemple, vous me vouvoyez, mais je suis tout seul ! Ce que vous vouvoyez en moi, ce n'est pas moi : c'est moi avec l'institution qui est derrière, l'Éducation nationale, et peut-être même la matière que je vous enseigne. »

C'est là, en général, que la discussion prend fin ; les élèves ne disent rien, mais n'en pensent pas moins qu'une discipline – non scientifique ! – où l'on ne peut pas dire ce que l'on pense et où l'on se cache derrière un « nous » « majestueux », qui leur semble un habit trop grand pour eux, leur fait perdre quelque considération pour la philosophie : cette discipline étrangère où l'on se moque de leur opinion et où l'on fait finalement peu de cas de leur personne, car il s'agit de rechercher une sagesse partie se cacher on ne sait où. Ce « nous » inclusif semble les exclure de ce qu'ils pourront dire.

Le professeur, soucieux, se demande s'il a bien fait. Est-ce lui qui a parlé, ou bien la philosophie, ou bien l'institution – ou bien les attentes du baccalauréat et de ses collègues, telles qu'il les comprend ? N'aurait-il pas mieux fait de suivre Descartes ? D'encourager les « Je pense » et l'affirmation de celui qui ose énoncer une pensée, la sienne, pour ainsi dire « à voix nue », s'exposer à la critique en tant que personne ? Ne pas se dérober à la pensée, oser faire désordre par l'affirmation d'une subjectivité forte qui s'assume ? Ne s'est-il pas comporté en petit fonctionnaire, en rouage d'une plus grande machine à enseigner où, finalement, plus personne ne pense – même pas lui ?

Mais… trêve d'examens de conscience. Bientôt retentira la sonnerie : délivrance programmée. Pourquoi s'inquiéter ? Nous avons délégué à la machine le métronome de nos pensées.

13. Libération du sujet pensant au siècle des Lumières françaises

L'audace du « je pense » sera reprise et très présente dans les Lumières françaises. Elle prendra des formes inattendues, diverses et bariolées, en se retournant quelquefois contre Descartes lui-même ; mais c'est un véritable vent de liberté qui a soufflé sur la philosophie, une façon de penser et de s'exprimer qui a décoiffé l'Europe couronnée.

Les philosophies de Diderot, Rousseau, Voltaire sont celles de Diderot, Rousseau et Voltaire eux-mêmes. Leur pensée nous parle avant tout d'eux-mêmes, et cette singularité n'enlève rien à leur universalité – bien au contraire. Cela donne à leurs pensées une force surprenante et inusitée.

Chez Diderot, le « je » n'est pas ce sujet abstrait, retiré du monde : c'est un « je » vivant, traversé par la sensibilité, ouvert aux autres et à l'expérience. La pensée n'est pas un pur calcul rationnel : elle vibre, elle imagine, elle sympathise. Dans Le Rêve de d'Alembert, il fait parler ses personnages comme s'ils étaient des organismes en perpétuelle métamorphose : « Nous ne sommes que des instruments sensibles. » La conscience de soi n'est donc jamais séparée de la matière et du corps. Chez lui, même le rire pense ; même ce que pensent les enfants a droit de cité en philosophie.

Sa critique des corps-machines de Descartes vaut la peine d'être citée. Dans Le Rêve de d'Alembert, il décrit (autant qu'il imagine et reconstitue) la croissance d'un poussin dans son œuf, puis le montre brisant sa coquille : « la voilà brisée ; il en sort, il marche, il vole, il s'irrite, il fuit, il approche, il se plaint, il souffre, il aime, il désire, il jouit ; il a toutes vos affections ; toutes vos actions, il les fait. Prétendrez-vous, avec Descartes, que c'est une

pure machine imitative ? Mais les petits enfants se moqueront de vous, et les philosophes vous répliqueront que si c'est là une machine, vous en êtes une autre. » – Diderot, Le Rêve de d'Alembert[70].

Bien sûr, la réfutation d'un Descartes par les rires des enfants laisse songeur. Et pourtant, une évidence nous apparaît : une pareille critique n'est pas coupée de la vie et de notre humanité ; elle est un véritable gai savoir qui donne envie d'étudier. Voilà qui n'est pas mortifère. C'est un élargissement du sujet, comme son renforcement et son décloisonnement. C'est finalement conforme à une tradition.

Chez Rousseau non plus, le « je » n'est pas un simple instrument de connaissance : il est une perpétuelle quête de sens, l'affirmation d'une singularité[71], l'audace d'une remise en question de tout ce qui le brime, le diminue et le réduit aux rôles assignés. Si la société brise les individualités, si elle discrédite la parole de quelques-uns pour ne porter que celle des autres, c'est qu'elle est mal organisée. Il faut revoir de fond en comble l'éducation et l'organisation politique du pays. La philosophie a pour vocation la libération de la parole et de la sensibilité, des hommes et des peuples. La religion elle-même doit davantage faire droit à notre humanité, à nos manières d'être naturelles. La pensée n'a rien de froid ni d'anonyme : elle a des accents poignants de sincérité – des ridicules, parfois ; c'est une philosophie à visage humain. Son « je » est narcissique, insupportable, envahissant, mais rien chez Jean-Jacques Rousseau n'est jamais insipide.

Quant à Voltaire, le « je » est une volonté en action, intransigeante et libre. Ce n'est pas un sujet qui se laisse enfermer dans un système : c'est une personnalité, un style, une ironie. Son « je » ne s'en laisse pas conter ; il doute, il pique, il rit, il s'indigne. Mais

surtout, il revendique le droit de changer d'avis, de prendre parti, de combattre et de s'exprimer. Derrière chaque texte, on devine la présence d'un homme qui écrit comme il parle, dans une langue claire, mobile, qui préfère l'éclat de l'esprit au carcan des démonstrations. On lui reproche parfois la légèreté ; elle est ici une méthode : éclairer plutôt qu'alourdir.

La philosophie, chez lui, n'a pas la froideur des traités : elle prend la forme de contes et de fables, presque d'histoires pour enfants (Candide, Zadig, Micromégas[72]). La légèreté du récit dissimule une ironie mordante. Ce « je » voltairien est en même temps un personnage qui sait s'émerveiller – devant la nature, devant la diversité des hommes. Ce philosophe, enfin, est un homme à part entière, un observateur implacable qui dénonce les injustices et les fanatismes.

Sa pensée est inséparable de sa vie : revendiquer le bonheur, ce n'est pas une abstraction. Voltaire fait de son grand œuvre non pas un traité, mais le bonheur concret d'un village. À Ferney, il se fait bâtisseur, protecteur, bienfaiteur : « J'ai fait de mes paysans des hommes heureux », écrit-il fièrement[73]. Sa philosophie se mesure à cette liberté incarnée, à ce combat pour la dignité, à ce refus d'un « je » anonyme.

Bien sûr, tout le monde ne peut pas être Diderot, Voltaire, Rousseau ; mais qu'on ne nous dise pas, après cela, que philosopher, c'est mourir à sa propre personnalité, c'est savoir s'effacer, ratatiner sa subjectivité pour devenir un quidam, un n'importe qui, un moi pur ou un « je transcendantal[74] », squelette d'une pensée bien bâtie, bien architecturée, mais d'une pensée qui s'est allégée de sa vie et de sa chair. Une pareille pensée, après tout, se débarrasserait bien de l'être humain comme d'une pesante

imperfection, pour le plus grand profit d'une machine à penser. C'est pourtant bien ce qui s'est fait !

14. La tradition du « je pense »

Certainement, il nous faudrait ici faire le récit systématique d'une pensée qui tend à se désincarner, éprise d'un idéal scientifique au point d'effacer la personnalité du penseur, comme on biffe une erreur. Nous devrions parler, très brièvement, de Kant, qui emprisonne le sujet de connaissance dans le carcan de catégories qui le précèdent et demeurent figées ; de Hegel, pour qui le sujet de connaissance n'est plus humain mais la Raison elle-même, se jouant des grands hommes et des peuples comme de marionnettes pour réaliser ses desseins ; du béhaviorisme, qui juge le sujet conscient trop encombrant et met entre parenthèses l'intériorité au profit du mesurable ; du structuralisme, qui déplace l'analyse de l'individu vers des structures impersonnelles qui le conditionnent et le produisent ; et, côté philosophie analytique, de courants (positivisme logique, naturalisme, Quine) qui privilégient la formalisation et l'empirie, parfois au risque d'éclipser le sujet[75].

Nous devrions également, ne fût-ce que par honnêteté intellectuelle, reconnaître que Nietzsche, tout en ébranlant le cogito cartésien et en doutant de la liberté, essaye de revivifier la pensée – au risque de se débarrasser du vieil humanisme. Nous devrions mettre au crédit d'Husserl sa tentative de reprendre le fil de la pensée cartésienne, et, chez Sartre, un peu de cette philosophie des Lumières qu'il a su insuffler de nouveau à la France, pour la faire rayonner[76].

Qu'on nous permette ici un plaidoyer pour une tradition qui nous est chère, en même temps qu'un plaidoyer pour la diversité.

Au vu de toutes les attaques contre Descartes et de l'espèce de mépris universel qui converge vers lui, il ne serait peut-être pas mauvais de défendre une tradition française de pensée. D'ailleurs, il n'est pas mauvais que les Japonais défendent leurs traditions, les Italiens leur spécificité ; ce qui désole profondément, c'est de voir combien les philosophes français du vingtième siècle se sont laissés ensorceler par une pensée allemande, en oubliant ce qu'il y avait de précieux chez eux. Ce tropisme allemand n'était bon pour personne, pas même pour les Allemands eux-mêmes. Les rencontres ne sont vraiment fécondes que pour autant qu'on reste deux. Chaque pays a ses chants d'oiseaux, et chaque langue, chaque héritage culturel, mérite de conserver sa voix dans une pensée mondialisée[77]. Les pays, les traditions sont aussi des sujets pensants. Un sujet se pose, s'oppose, discute, dispute. Il ne faut pas regretter que la philosophie soit un « champ de bataille », comme le disait Kant de la métaphysique. Ce « champ de bataille », c'est le « chant de la vie de la pensée elle-même ». Quand tout semble ronronner, la pensée ne vit plus, et il importe peu qu'on dise qu'elle est morte ou que « cela fonctionne ». Cioran a pu écrire : « Un peuple qui renonce à ses défauts renonce à lui-même. »[78] Peut-être la France devrait-elle réapprendre à assumer ses « défauts » : son individualisme, son ironie, son refus des grands systèmes, sa légèreté dans la profondeur. Ces soi-disant défauts sont peut-être aussi un peu de son génie. Le mépris de cette belle tradition (peut-être s'imagine-t-on l'avoir dépassée) est aussi mortifère pour la pensée qu'un chauvinisme exacerbé qui ne veut rien voir arriver de l'étranger – et qui se croit prophète alors qu'il ne fait que bégayer son passé.

15. Une pensée sans plus personne qui pense

Mais… suivons la flèche de réduction du sujet pensant dans la pensée pour voir jusqu'où elle nous mène. Elle se fige précisément dans un article polémique de Chris Anderson, paru en 2008 dans Wired, et qui a fait grand bruit[79]. Wired est un magazine américain fondé en 1993, spécialisé dans les technologies, la culture numérique, l'innovation, la science et la société, reconnu pour son ton avant-gardiste. Traduit en français, le titre de l'article est : « La fin de la théorie : le déluge de données rend la méthode scientifique obsolète ». Il est des plus parlants : à l'ère des données massives, ce seraient les machines qui prendraient le relais de l'intelligence humaine – et elles le feraient beaucoup mieux que nous, même si l'on ne comprend pas toujours comment. Certes, l'article d'Anderson est polémique, mais on aura raison de le prendre au sérieux : la paresse intellectuelle est encouragée par des forces très puissantes qui n'ont pas besoin de femmes ou d'hommes intelligents, mais de consommateurs – et, si possible, peu avertis.

Un bon exemple de cette bascule est l'évolution de la traduction automatique. Historiquement, la traduction automatique (TA) a d'abord été dominée par des approches à règles conçues par des linguistes (grammaires, lexiques, modèles syntaxiques). Chez Google, en revanche, le service Translate s'est appuyé dès le départ sur l'apprentissage statistique à partir de corpus parallèles massifs (traductions de phrases dans les deux langues), avant le passage aux réseaux neuronaux en 2016 (Google Neural Machine Translation, GNMT). Dans ces paradigmes, on nourrit les modèles avec des corpus multilingues gigantesques et on laisse les machines apprendre des régularités sans représentation explicite de la grammaire. Le saut qualitatif observé lors du passage au

neuronal a été spectaculaire[80]. Le rôle des linguistes n'a pas disparu ; il s'est déplacé vers la constitution et le nettoyage des corpus, l'évaluation, l'adaptation à des domaines spécialisés et l'analyse d'erreurs.

On peut dire, en forçant à peine le trait, que Google « traduit » sans vraiment « connaître » les langues : fournissez des corpus alignés suffisants et le système pourra, en principe, traduire n'importe quelle paire – du français vers l'allemand, voire entre langues peu documentées ; il existe même des capacités dites « zero-shot » entre paires non directement vues, mais la qualité chute dès que les données manquent. Les approches statistiques et neuronales se sont révélées, dans bien des cas, plus performantes que des systèmes fondés sur des règles explicites. Cela pourrait bien être cela, le futur de la connaissance : ne plus chercher à comprendre, mais simplement croiser les données jusqu'à ce qu'un motif apparaisse. « Avec suffisamment de données, les chiffres parlent d'eux-mêmes. »

Ainsi, la science aurait atteint l'apogée de l'objectivité en se débarrassant d'un maillon faible : l'être humain. Nos théories, nos explications, nos hypothèses – tout ce qui faisait le cœur de la science depuis des siècles – deviendraient inutiles. Elles seraient remplacées par le travail obscur mais redoutablement efficace des machines, qui trouvent, prédisent, traduisent, classent, sans jamais expliquer.

Bien sûr, cet article est discutable, et on l'a discuté. Il est possible de le relativiser et de n'y voir qu'un article parmi d'autres dans un océan de publications. Mais il ouvre la possibilité d'un avenir avec l'IA : on ne travaillerait pas avec l'IA, l'IA travaillerait pour nous. Enfin « nous » ? De quoi exactement parle-t-on ? Que resterait-il de « nous » ?

Dans La Machine à explorer le temps (The Time Machine, 1895), H. G. Wells imagine un avenir troublant pour l'humanité[81] – un futur lointain (l'an 802 701) où la société s'est fragmentée en deux espèces : les Éloïs et les Morlocks. Les Éloïs sont, pour ainsi dire, les chanceux et l'élite de l'humanité, mais ils ne dirigent en fait rien. Ce sont les machines qui le font pour eux, machines entretenues par les Morlocks.

Les Éloïs sont petits, gracieux, androgynes, au teint pâle. Ils vivent à la surface, dans un monde qui semble paisible et paradisiaque ; ils sont doux, passifs, peu intelligents, sans mémoire ni curiosité réelle, sans structure politique ni travail apparent. Ils ont tout confié aux machines.

C'est la version SF de l'inquiétude de Tocqueville dans De la démocratie en Amérique : « Je veux imaginer sous quels traits nouveaux le despotisme pourrait se produire dans le monde. […] Je vois une foule innombrable d'hommes semblables et égaux qui tournent sans repos sur eux-mêmes pour se procurer de petits et vulgaires plaisirs dont ils emplissent leur âme. Chacun d'eux, retiré à l'écart, est comme étranger au destin de tous les autres ; ses enfants et ses amis forment pour lui toute l'espèce humaine ; quant au reste de ses concitoyens, il est à côté d'eux, mais il ne les voit pas ; il les touche et ne les sent point ; il n'existe qu'en lui-même et pour lui seul. Et s'il lui reste une famille, on peut dire qu'il n'a plus de patrie.

Au-dessus de ceux-là s'élève un pouvoir immense et tutélaire, qui se charge seul d'assurer leur jouissance et de veiller sur leur sort. Il est absolu, détaillé, régulier, prévoyant et doux. […] Il aime que les citoyens se réjouissent, pourvu qu'ils ne songent qu'à se réjouir.

[…] Il ne tyrannise point, il gêne, il comprime, il énerve, il éteint, il hébète, et réduit enfin chaque nation à n'être plus qu'un troupeau d'animaux timides et industrieux dont le gouvernement est le berger. »[82]

Qu'il nous soit seulement permis d'ajouter un simple petit mot derrière « gouvernement » : le mot « machine ».

Mais cela, bien sûr, n'est jamais qu'une possibilité. On peut faire travailler les machines à notre place, ou travailler avec elles ; ne rien céder sur l'intelligence et, dans une stupéfiante émulation, relever les défis lancés par les machines en associant nos efforts à ce qui, tout de même, n'est jamais que le résultat de notre génie et de notre création : faire de notre intelligence une intelligence au carré. Mais, pour cela, il faut une humanité qui continue à dire « je pense » et défende pied à pied ce qui lui semble être les valeurs mêmes de notre humanité.

La question ainsi posée ne relève pas seulement du Vrai ; il faut parler du Bien.

DEUXIÈME LIVRE
LE TAMIS DU BIEN

PARTIE I

LA PLUS GRANDE INVENTION APRES DIEU ?

Avec cette partie s'ouvre le passage du Vrai au Bien. La parole n'est plus seulement ce qui décrit ou dévoile le monde : elle devient ce qui commande, fonde des lois, institue des médiations et produit de l'autorité. Les écritures, les religions du Livre et la question du code dessinent ici une première histoire de l'autorité non humaine.

1. Un glissement insensible du Vrai au Bien

Nous avons évoqué brièvement plus haut que les quatre tamis étaient, du temps de la Grèce antique, en harmonie. Ce qui était Vrai était également considéré comme Beau et Bien et certainement Utile à dire. Cette belle unité a été brisée par la croix du Christ. Le Christ crucifié, c'est le renoncement à la puissance et donc à l'Utile (le Mauvais Larron se moque de lui sur la croix de l'inutilité d'une divinité qui ne lui permet pas de se décrocher tout seul), les mauvais traitements l'ont souillé et rendu laid, sa « vérité » tourne à l'ironie car Ponce Pilate, par dérision, fait inscrire sur l'instrument de son supplice « Jésus de Nazareth, roi des Juifs[83] », et l'incarnation de la justice est traitée comme un criminel.

Cependant, quand le christianisme a perdu en vigueur, dès le XVIII^e^ siècle et plus encore au XIX^e^, certains scientifiques ont attendu de la « vérité scientifique » qu'elle détermine également ce qui était Bien, Beau, Utile. C'est l'idée très floue du « Progrès » qui remplace l'idée de Providence divine et qui donne notamment le positivisme[84]. Remarquons au passage que les grands totalitarismes du XX^e^ siècle ont eu de ces prétentions

« scientifiques » à dire ce qui était vrai, bien, beau, utile et à déporter les opposants.

On peut essayer de se rassurer en rappelant qu'il a été démontré que ce n'étaient pas là de « vraies sciences ». Hélas, quand l'autorité politique se proclame scientifique, l'idéologie dominante fait de la science un instrument du pouvoir. L'affaire Lyssenko[85], du nom d'un scientifique russe sous Staline qui a défendu une théorie génétique aberrante mais conforme à la doctrine du Parti soviétique, a abouti à la mort de faim en prison de son contradicteur, Nikolaï Vavilov, en 1943.

Si l'oiseau de Minerve se lève le soir, comme le prétendait Hegel[86], si la vérité n'apparaît que tardivement, c'est bien dommage pour ceux qui ne tiendront pas jusqu'au soir. Et comme nous ne savons pas exactement quand aura lieu ce soir où la vérité apparaîtra dans son aura éblouissante, les non-scientifiques ne peuvent que se soumettre aux affirmations de ce qui, en un moment « t », passe pour la science du jour. Hegel est d'ailleurs l'un des premiers penseurs modernes qui ont voulu faire converger le Vrai, le Bien, l'Utile et le Beau dans leur philosophie, pour se prétendre les détenteurs de la vérité absolue dont se sont revendiquées les idéologies totalitaires. Kant avait tenté de prévenir contre de telles illusions. Un esprit humain, c'est-à-dire fini et limité comme le nôtre, ne peut parler au nom de l'Absolu. Pour le dire à notre manière propre, celui qui escalade la pyramide philosophique pour proclamer du haut de son sommet qu'il est le porte-voix de l'Absolu quitte le champ philosophique et s'érige en prophète. C'est dangereux. À la pensée humaine on substitue une logique autonome qu'on prétend nécessaire. D'autres en feront une fatalité.

Quand Hegel laisse entendre que l'Esprit a atteint en lui sa pleine conscience de soi, il ne fait pas que philosopher : il se pose en oracle. Selon cet oracle, toute la vérité de l'humanité, après des millénaires de luttes, d'erreurs et de souffrances, converge miraculeusement… vers son bureau de professeur à Berlin. Ironie tragique : ce penseur de la dialectique, de la négation, de la liberté par la reconnaissance, devient le prophète d'un système clos – où la Raison, enfin « absolue », n'a plus rien à apprendre. L'IA contemporaine hérite de cette posture, mais en la vidant de toute humanité : elle aussi se présente comme l'aboutissement d'un processus historique, le point de convergence de toutes les données, le juge ultime de l'efficace. Sauf qu'elle ne pense pas – elle calcule. Elle ne libère pas – elle optimise. Elle ne reconnaît pas – elle classe. Hegel rêvait d'un Esprit libre, mais, comme l'avait vu Kierkegaard[87], pour cette soi-disant liberté il avait allégé la subjectivité humaine. Alors à quoi bon ? Qu'est-ce qu'une liberté sans véritable subjectivité ? Un esprit formaté d'une immense machine métaphysique dont nous ne sommes que les pantins manipulés ?

Avant d'aller plus loin, regardons ce qu'il y a dans le « Tamis du Vrai » tel que nous l'avons étudié dans la partie précédente. En vérité, il reste le calcul (la raison au sens mathématique de ratio), les données de l'expérience qui ont été purgées de leur dimension subjective et humaine, ainsi que ce qu'on pourrait appeler l'objectivité scientifique, épurée autant que faire se peut de toute subjectivité, comme l'a demandé avec beaucoup de force et de pertinence, dans sa propre logique, un Gaston Bachelard[88]. Nous avons vu que la subjectivité humaine était pour ainsi dire passée à travers les mailles du tamis pour se confondre avec la boue du faux ou du négligeable. Ainsi, la recherche

philosophique et scientifique a préparé le terrain pour une objectivité sans sujet, pour des machines calculantes générales. Si l'expérience montre qu'à terme ces systèmes sont bien plus capables que les humains pour élaborer des théories vérifiables, les postes de chercheurs sont bien sûr compromis. Seuls demeureraient des vérificateurs humains dont la tâche serait d'autant plus ardue qu'ils seraient d'autant plus rares et dépassés par des performances inhumaines.

Le Vrai ne serait plus, pour ainsi dire, sa propre référence, mais il serait prouvé, démontré par l'expérience concrète : l'Utile. Autrement dit, si une théorie scientifique permet de faire décoller une fusée, la théorie pourrait paraître suffisamment prouvée… sans même qu'il soit besoin que les scientifiques la comprennent. On aurait alors ce que Simondon appelle, dans ses Cours sur la technique, un « objet technique fermé », une sorte de boîte noire inaccessible aux êtres humains[89]. Cependant, la théorie nouvelle et incompréhensible aurait gagné une légitimité d'efficacité aux antipodes de la science telle que la comprenaient les Anciens. La science était une contemplation sans application technique directe, ce qui ne veut évidemment pas dire que la technique n'existait pas. Archimède passe pour une exception notable parce qu'il a fabriqué des machines de guerre comme un ingénieur de génie. Mais ce qui s'est imposé à lui comme nécessaire dans une situation de guerre n'était pas nécessairement son idéal d'homme et de scientifique. Plutarque rapporte ainsi qu'Archimède tenait en moindre estime les arts tournés vers les besoins matériels et réservait son admiration la plus haute à la beauté de la géométrie[90].

Remarquons que la position philosophique qui aligne le Vrai sur l'Utile est très proche de ce que l'on a appelé en philosophie le pragmatisme. William James, par ailleurs philosophe très

respectable et désireux de faire le bien, a pu écrire : « Est vrai ce qui réussit », et, plus proche de nous, John Dewey : « Est vrai ce qui s'avère effectivement efficace[91] ». Pour être justes, rappelons que, chez ces philosophes – chez James et surtout chez Dewey – l'« utile » n'équivaut pas au rentable : il renvoie à ce qui fait croître la communauté d'enquête, améliore l'éducation, la santé, la délibération. On ne confond pas le bien public avec le profit. Le pragmatisme philosophique est expérimental, c'est-à-dire qu'il teste des hypothèses et est capable de se corriger s'il n'est pas concluant. Il se remet en question, exige une certaine transparence, la publicité des raisons et la réversibilité des décisions.

Mais toute pensée profonde, reprise par la société, est caricaturée, amalgamée. Le capitalisme contemporain n'a pas de pareils scrupules : il confond l'Utile avec le profitable, l'optimisable. Les États eux-mêmes confondent les intérêts du public et de la science avec des intérêts privés. Au nom de la compétitivité, ils privilégient souvent l'innovation transférable (crédits d'impôt, primes à la R&D des entreprises) au détriment de la recherche fondamentale ouverte. C'est ce glissement – de l'Utile démocratique vers l'Utile marchand au nom d'un supposé Bien – qui prépare une autorité techno-oraculaire où l'efficacité tient lieu de vérité. Précisons pour bien nous faire entendre, que l'autorité, c'est ce qui inspire respect et obéissance en vertu de la supériorité – réelle ou perçue – qu'on lui reconnaît, quelle qu'en soit la source, sans qu'il y ait besoin de recourir à la contrainte. Ce qui fait autorité à une époque n'est pas nécessairement légitime.

Si la recherche de la vérité n'est plus motrice, mais le wagon de l'efficace, alors on en vient à penser que l'efficacité vaut vérité, voire qu'elle permet de s'y substituer pour devenir autorité. Il semble qu'aujourd'hui, en 2025, de telles « vérités non

humaines », si cela peut avoir du sens, existent déjà. Des équipes de Princeton ont mis au point une méthode par « inverse design » dans laquelle une IA crée en quelques heures des puces pour les communications sans fil. Le résultat ? Des circuits aux formes étranges qui échappent à l'interprétation humaine, mais offrent pourtant des performances inédites. Le chercheur Kaushik Sengupta déclare laconiquement que « les humains ne peuvent véritablement les comprendre ». Faut-il s'en inquiéter ou s'en émerveiller ? L'émerveillement est possible, mais qu'il ne prenne pas la place de la Vérité, qui n'a de sens qu'humain. Sinon, la technique devient Magie 2.0 et est la porte ouverte à de nouvelles superstitions, où les forces occultes migrent des esprits ou des divinités – aux machines.

De la même façon, on pourrait fort bien en arriver à l'idée (via les « sciences humaines » notamment) que l'IA pourrait en venir à nous proposer des organisations sociales qui sont bonnes pour nous mais que nous ne sommes pas à même de comprendre[92]. Est-ce à craindre ou à espérer ? Si, à l'épreuve, elles se montrent plus « efficaces » que les solutions humaines, politiques ? Une organisation sociale efficace qui échappe à l'humanité qu'elle organise n'est-elle pas nécessairement une dérive, un éloignement de la rive de notre humanité ? Cela ne peut reposer que sur la confiance ; mais qu'est-ce que la confiance sans respect et amour pour l'être auquel on se confie ? Nous marchons sur des terres inconnues. L'unique crime, sans doute, serait de manquer, face à de pareilles questions, d'humilité et de délicatesse…

Le problème qui se pose est évidemment celui de l'autorité. L'autorité a longtemps eu en Occident, comme d'ailleurs dans le reste du monde, un soubassement religieux. On peut considérer qu'après la Révolution française, la science, envahissant le

domaine de l'humain, essaie de prendre le relais pour fonder l'économie – qui se base sur des modèles mathématiques – la sociologie, forte des grands nombres qui permettent des statistiques, et même la « science politique », couple de mots en grande tension quand on y regarde de près.

On assiste à une convergence très forte entre sciences humaines et big data. Vu les taux de recrutement, ce serait peut-être là qu'il faudrait conseiller aux étudiants intéressés par de semblables sujets d'orienter leurs études. Une importante question est de savoir si les étudiants seront les maîtres ou les serviteurs des résultats produits par les systèmes de deep learning – d'« apprentissage profond », si l'on préfère – ou de réseaux neuronaux. Ou bien ? Une relation qui ne soit ni de maîtrise ni de servitude avec les machines serait-elle possible ?

2. Autorité religieuse et écritures

Par une sorte de généalogie de la pensée occidentale, nous avons, dans la première partie, montré combien les philosophes avaient été activement impliqués dans l'avènement de ce qui deviendrait plus tard l'intelligence artificielle. L'autorité du scientifique relayée par l'autorité de la machine semble suivre une certaine logique incarnée par le mot même de « technologie ». C'était d'abord l'étude scientifique de la théorie des techniques, voire l'histoire des techniques ; le mot désigne maintenant l'ensemble des techniques modernes qui sont l'aboutissement de recherches scientifiques et s'en émancipent en objets ou services usuels.

Peut-on faire la même chose avec la religion ? Cela semble bien plus difficile. Dans nos représentations communes, la science n'est pas la complice du progrès technique, mais bien plutôt sa contestation. En 1947, le très chrétien Bernanos résume

brutalement son inquiétude dans La France contre les robots par une formule devenue célèbre : « un monde gagné par la technique » risque, selon lui, d'être perdu pour la liberté. Au début du XX^e siècle, plus violent, Léon Bloy dénonçait en bloc la modernité, le progrès technique, le culte de la vitesse et de l'argent, qu'il percevait comme un engloutissement de l'âme dans la matière et un éloignement de Dieu ; bicyclettes, automobiles, métro, canots à moteur et même tour Eiffel deviennent sous sa plume les emblèmes d'une modernité démoniaque. Le jour de la résurrection, il conviendrait peut-être de ne pas trop lui parler des IA. La religion, conservatrice semble-t-il par nature, incarne pour beaucoup un certain obscurantisme religieux. Dieu a fait la nature bonne et vouloir améliorer l'œuvre de Dieu est une hérésie. La génétique, en particulier, est dans le viseur : la Procréation Médicalement Assistée, l'usage de la chimie dans le corps pour contrecarrer la volonté de Dieu quant à la procréation, etc.

À y regarder de plus près, cependant, il faudrait nuancer et surtout ne pas confondre les religions monothéistes avec leurs formes extrêmes ou intégristes. Ainsi, l'histoire des institutions religieuses révèle qu'elles n'ont jamais été totalement hostiles à la technique. Au contraire, elles l'ont souvent intégrée, voire encouragée, dès lors qu'elle servait leurs propres desseins. La construction des grandes cathédrales gothiques n'aurait pas été possible sans une maîtrise avancée de l'ingénierie et de l'architecture. Le développement de l'imprimerie, bien que craignant initialement la diffusion des hérésies, fut rapidement adopté par l'Église pour la production en masse de bibles et de textes théologiques. La musique sacrée elle-même, avec l'orgue par exemple, est un instrument technologique sophistiqué qui amplifie l'émotion et l'élévation spirituelle.

Le Vatican, sous l'égide du pape François, a ainsi initié le « Rome Call for AI Ethics[95] », un document cosigné par des géants de la tech qui pose les bases d'une IA respectueuse de l'humanité. De leur côté, les théologiens musulmans ont émis des fatwas pour guider l'utilisation de l'IA dans le respect de la charia, et le Grand Rabbinat d'Israël a publié des responsa (réponses juridiques) pour aborder les implications de cette technologie sur la loi juive (halakha[96]). Malgré des tensions évidentes où la tradition s'applique à rester en contact avec la modernité, ces institutions cherchent à se positionner comme des partenaires éthiques crédibles, cherchant à s'assurer que l'IA serve l'humanité sans la déshumaniser. On aurait sans doute tort de laisser tomber leur message dans l'indifférence.

Les réticences aussi méritent d'être entendues. Une vraie voix comme celle de Hans Jonas, philosophe hébraïque, tragique dans la plus pure tradition des prophètes de malheur de l'Ancien Testament, nous manque aujourd'hui pour nous appeler à la responsabilité de façon convaincante et audible[97]. Il ne nous dit pas ce qui sera – nous avons assez combattu cette tentation du philosophe. Il nous dit ce qui pourrait être si nous nous laissions aller à la déresponsabilisation, à nous oublier comme sujets conscients et responsables de l'avenir. L'homme est le berger de l'être, disait Heidegger[98] ; il conviendrait surtout que les hommes et les femmes de bonne volonté soient les bergers de l'humanité qui viendra après pour leur permettre d'exister. Rien n'est évident. Rien n'est donné. Un « progrès » qui n'est pas pensé n'est pas une ascension pour l'homme, c'est une pente qu'on ne descend pas, mais qu'on dévale au risque de chuter.

Effectivement, comme l'affirmait Jonas, il faut qu'une vie digne d'être authentiquement qualifiée d'humaine puisse

demeurer sur cette terre pour que l'avenir puisse sourire à nos enfants. Qu'aurait-il pensé de l'IA ? Pour lui, qui n'a pas connu les LLM, le deep learning, etc., mais a connu le désastre de l'atome, la technique humaine avait changé de nature et notre responsabilité s'accroissait considérablement en fonction de cette puissance nouvelle.

De même, Bergson, dans Les deux sources de la morale et de la Religion (1932), appelait de ses vœux un « supplément d'âme » pour que nous ne nous laissions pas dépasser par notre puissance mécanique nouvelle[99]. Mais aucun des deux en vérité n'a été confronté au défi d'une machine parlante et agissante qui imite l'intelligence humaine avec plus d'efficience qu'elle.

N'y a-t-il pas un lien de filiation entre les codes de loi se réclamant du divin, qui ont fondé nos sociétés (Code d'Hammourabi, Tables de la Loi[100]), et les codes des machines qui, devenant oracles et autorités, pourraient demain organiser nos sociétés ? L'autorité fait loi, le code est normatif pour une société et l'obéissance est son pendant. Or, c'est dans le langage, la parole qu'on reconnait l'ordre auquel on doit se conformer. Cette autorité d'une parole, qu'on la prête à Dieu ou qu'on lui substitue, la nature, la raison ou des machines, c'est un ordre nouveau au sens d'une organisation nouvelle. La sagesse comme aspiration de la philosophie invite à se questionner sur de semblables nouveautés et à sonder au moins leur légitimité relative.

Avouons-le cependant, nous aurons peu fait pour relier l'autorité religieuse à l'autorité putative d'une « machine à parler », d'une IA, tant que nous n'aurons pas observé les linéaments profonds qui les relient dans une généalogie nouvelle. Tout commence pour ainsi dire quand la parole vive s'incarne et se grave

dans le monde physique. Au commencement était le Verbe qui se fit, en Mésopotamie, écriture.

1. Une révolution anthropologique

L'émergence de l'IA n'est pas une invention comme une autre. Elle concerne le tout de l'homme, son rapport au Vrai, au Beau, au Bien en même temps qu'à l'Utile. Peut-être est-elle comparable à celle de l'écriture en Mésopotamie vers 3500 avant Jésus-Christ, qui a eu un tel impact sur l'homo sapiens qu'elle l'a sorti de la préhistoire[101]. Dans Le Tiers-Instruit (1991) et Petite Poucette (2012), Michel Serres insistait sur le fait que l'invention de l'écriture avait profondément modifié notre environnement et nos façons de nous inscrire dans le monde, nous avait fait changer de monde[102]. C'est dans sa direction que nous allons.

Dans le domaine du Vrai : l'écriture a permis de fixer les observations, de transmettre les raisonnements scientifiques, et de constituer des corpus stables où la vérité pouvait se discuter sur la base de textes communs. La science devient cumulative. On sort du mythe qui se réinvente en permanence. Dans le domaine du Bien : la loi, la morale et la religion changent d'échelle. Les codes écrits remplacent la mémoire orale et instaurent une autorité stable – lois gravées comme celles d'Hammurabi[103], textes sacrés, doctrines politiques transmissibles à travers le temps et l'espace.

Faut-il vraiment insister sur le caractère utile de l'écriture ? Elle inscrit l'homme dans un temps long et contribue fortement à le détacher de sa nature animale. Les premiers écrits retrouvés sont des listes de biens et des factures. Si peu réjouissante qu'elle puisse paraître, la notion de dette participe d'un arrachement au temps présent. La « reconnaissance de dette » devient autre chose

qu'un sentiment mêlé de souvenirs. Bientôt, une administration sera possible avec de grandes villes et des pouvoirs centralisés comme Pharaon entouré de ses scribes. Des monuments parlent du passé et font entendre la voix des morts, réunissant les conditions matérielles de la culture. L'humanité qui se représente elle-même dans des écrits n'est plus la même ; elle s'agrandit considérablement en termes de conscience d'une présence singulière mêlée du sentiment de la finitude individuelle. C'est la magnifique Épopée de Gilgamesh[104], témoignage fascinant d'une humanité qui se découvre elle-même, prend conscience de l'horreur de la mort et se transforme à mesure :

Le prologue invite le lecteur à ouvrir le dépôt de tablettes, à en dégager le texte et à lire les épreuves traversées par Gilgamesh : l'écriture est déjà mise en scène comme un dépôt de mémoire destiné à survivre au héros.

Enkidu, l'ami de Gilgamesh, n'est tout d'abord qu'un géant de la nature tout poilu – ensuite il découvre le langage : basculement en lui comme pour toute l'humanité qui passera par là – ensuite son histoire est racontée par Gilgamesh lui-même, dit-on, et c'est encore un acte de fidélité à son ami dans des manuscrits déposés dans le temple d'Ishtar – tournant encore qui parvient jusqu'à nous. Trésor qu'on lègue à la postérité comme un témoignage philosophique, religieux, politique, poétique inestimable.

Le prologue présente encore Gilgamesh comme un voyageur revenu de loin, éprouvé par son parcours, qui a consigné ses peines et ses découvertes sur un support durable.

Dans des ajouts postérieurs, on mentionne explicitement les générations à venir dont l'écriture permet de faire les héritières de cette histoire :

D'autres formulations du prologue insistent sur les jours lointains : stèles, nom et exploits doivent traverser les générations et porter la mémoire du héros bien au-delà de sa vie.

Qu'on ne s'étonne pas alors que, dans le domaine du Beau, la poésie prenne de nouvelles formes, s'il naît des sentiments nouveaux et une sorte d'émulation des artistes à travers les siècles, financés et protégés par des puissants. Des hymnes, des monuments, des statues en majesté sous lesquelles on met des légendes et des cartouches : elles peuvent parler ! Les œuvres deviennent plus complexes, mixent des styles, des époques, des savoir-faire de différentes régions.

L'écriture a relié l'homme à une mémoire externe et durable, modifiant son rapport au savoir, à la loi et à l'art. Et Michel Serres voyait avec raison dans l'imprimerie puis le passage au numérique des bouleversements comparables, ouvrant une nouvelle étape à l'aventure humaine.

Ce qui change, c'est le couple support/message. En modifiant le support, la nature des messages change et, avec elle, l'humanité qui devient autre[105]. Or, l'IA est bien davantage qu'un simple changement du couple support/message. L'IA génère des messages qui viennent d'un ailleurs que du strictement humain. Et c'est cet ailleurs que nous interrogeons aujourd'hui dans les quatre dimensions du Vrai, du Bien, du Beau et de l'Utile. Ça parle, mais sans qu'on puisse supposer de sujet parlant ? Est-ce la mémoire de l'humanité telle qu'elle a pu s'écrire qui est mobilisée pour me parler et me répondre ? Ce serait en soi impressionnant. Quel crédit accorder à une parole qui n'émane pas d'humains conscients de ce qu'ils disent ? Jusqu'ici, on pensait que seuls pouvaient parler les hommes, les esprits et les dieux. On s'arrangeait ainsi. N'est-ce pas un nouvel ordre ou de

nouveaux désordres qui s'installent avec des interlocuteurs de silicone ?

2. Les puissances du langage

Les algorithmes de l'IA ne sont-ils pas un summum, le Graal de tout technicien dans la recherche de l'efficacité, de l'utile ? Mais cette utilité est-elle conciliable avec le bien de l'humanité ? On a parlé d'un « grand remplacement » : non pas celui d'une population par une autre, mais de l'homme par la machine, de l'homme par les robots. Est-ce vraiment souhaitable ? Faut-il lutter contre cette technologie ? N'est-ce pas entrer dans un de ces combats d'arrière-garde et désespérés, comme la révolte des Luddites contre les machines au XIX^e^ siècle[106] ? Serons-nous comme ces paysans et ouvriers agricoles qui s'en prenaient aux tracteurs et batteuses mécaniques qui leur prenaient leur travail ? Les artistes aussi craignent pour leur activité. Si les intellectuels s'interrogent toujours pour savoir si les productions de l'IA sont ou non des œuvres d'art dignes de ce nom, eux voient concrètement qu'ils risquent de devenir massivement inutiles… Est-ce leur tragédie personnelle ou une tragédie pour l'humanité ?

Les premières traces de l'écriture, certes, sont bien prosaïques, comme le sont des factures. Dans les premiers temps, elle fut l'héritière de petits cailloux de pierre (les calculi) qu'on scellait à la cire dans une enveloppe d'argile[107]. Quand le débiteur venait payer sa dette – disons trois chèvres et un troupeau de quinze vaches – on brisait (sans doute devant témoins) cette espèce de tirelire ; on y comptait les jetons, la plupart du temps en argile séchée. Certains représentaient des vaches, d'autres des chèvres. Vint le moment où l'on se dit qu'on aurait aussi bien fait de graver directement les jetons sur les tablettes, ce qui permettrait, en

permanence, de contempler la dette. Notre alphabet garde encore la marque de cela : la lettre A, en capitale, vient de Aleph – qui, de façon stylisée, représente une tête de bœuf renversée[108].

Mais l'écriture, se complexifiant, ne fut plus accessible au quidam et devint l'apanage d'une élite au service des puissants. Les anthropologues, comme Lévi-Strauss, ont remarqué que les chefs des peuples sans écriture comprenaient très vite la puissance et l'autorité qui en découlaient, notamment au niveau du prestige.

En lisant, en écrivant, pour la première fois, l'homme usait de son langage pour parler à autre chose qu'à un autre homme. Certes, il commence à parler à des hommes qui sont loin ou même ne sont plus ; mais il peut bientôt parler à des morts qui se sont immortalisés par ces traces et ces écrits. Très vite, il en vient à penser que certaines de ces paroles sont attribuables aux dieux, qui écrivent leur volonté, comme eux, au travers de signes qu'il s'agit de déchiffrer dans la nature. Ainsi, dans Tristes Tropiques, Lévi-Strauss raconte qu'un chef nambikwara, sans savoir écrire au sens occidental, imite devant les siens l'usage d'un papier couvert de signes. L'épisode lui sert à montrer que l'écriture peut fonctionner non seulement comme code, mais comme marque visible d'autorité, de prestige et d'alliance[109].

Autrement dit, indépendamment même du message qu'elle peut véhiculer, l'écriture acquiert très vite une fonction phatique qui pose l'autorité du maître en écriture. Très vite, savoir lire et écrire ne fut plus seulement posséder le code qui permet l'échange, mais posséder la puissance de celui qui détient le code, caché aux autres. Et cette puissance mystérieuse, incompréhensible au vulgaire, s'associera naturellement aux puissances occultes du sorcier.

On sait que la formation des scribes dans les sociétés anciennes, notamment en Égypte, avait un lien très étroit avec une éducation sacerdotale ou religieuse. Les scribes étaient formés dans des institutions souvent liées aux temples, lieux à la fois de savoir et de pouvoir spirituel. Ils recevaient une formation poussée non seulement à la maîtrise des différentes écritures (hiéroglyphique, hiératique, démotique), mais aussi à des connaissances religieuses, littéraires, juridiques et administratives.

On en vint tout naturellement à penser que cette écriture si puissante ne venait pas des hommes, mais des dieux. Pour l'Égypte ancienne, au dieu Thot en particulier, comme le raconte le Phèdre de Platon : le dieu Theuth présente au roi Thamous plusieurs inventions, dont l'écriture, qu'il propose comme un secours pour la mémoire et le savoir[110].

On trouvera la même chose chez bien d'autres peuples. Les runes possèdent une dimension sacrée et ésotérique importante au sein des cultures germaniques et scandinaves anciennes. Selon la mythologie nordique, le dieu Odin les aurait découvertes après un sacrifice rituel de neuf nuits, suspendu à l'arbre cosmique Yggdrasil[111]. Voilà qui les relie directement à un savoir mystique et divin.

Si l'écriture n'est pas d'origine divine, sa puissance est si grande, ses mystères si impénétrables pour le profane, qu'il y voit justement une invention divine. Pareille grandeur ne peut provenir que du divin ! Ceux-là mêmes qui la maîtrisent connaissent certains de ses mystères et, ayant le respect de leur art, finissent par croire en son caractère sacré. L'écriture vient des dieux, et les dieux eux-mêmes possèdent une écriture qu'il s'agit de déchiffrer !

Comme nous l'avons signalé, il est fort dangereux qu'un philosophe se mêle de prophétiser (Hegel, Marx…). Mais une tendance est là qu'on peut expliquer à la fois par une puissance et une vitesse de calcul qui échappent à l'entendement humain et des quantités de données que des esprits comme les nôtres ne peuvent pas mémoriser : à moins d'augmenter artificiellement les capacités humaines (ce qui n'est pas sans poser d'autres problèmes), à mesure que la puissance des IA se renforcera dans les décennies à venir, que leurs logiques de fonctionnement deviendront de plus en plus opaques aux simples mortels, ceux-ci leur prêteront une origine divine, les serviront et les célébreront comme des dieux puissants.

Vous en doutez sans doute parce que vous êtes des êtres rationnels, informés, cultivés ; mais cette tendance – ce besoin presque – à voir du mystère et de l'incompréhensible dans ce qu'on ne comprend pas est bien présente sur Internet. La tendance est trop évidente pour qu'on y revienne encore. Pourtant, elle s'impose partout : là où la raison recule, le mystère s'installe comme une nécessité. Sur les écrans, dans les conversations, sur les réseaux, l'incompréhensible devient attrayant, l'absurde devient révélation. Des jeunes sacralisent l'astrologie, rejettent les conquêtes scientifiques les plus établies, et placent leur foi dans des oracles numériques dont la seule légitimité est le nombre de followers. La science, autrefois boussole, vacille ; à sa place, surgissent de nouveaux cultes, de nouveaux dogmes, portés par des voix médiatiques qui n'hésitent pas à prétendre que les pyramides furent bâties par des extraterrestres, que les Égyptiens allumaient des ampoules électriques, ou que les vaccins cachent des puces électroniques. Le paradoxe antique – « Je crois parce que c'est absurde » – n'a jamais été aussi vivant. Ce n'est plus seulement de

la désinformation : c'est un besoin profond, presque religieux, de croire à ce qui défie la raison. Et ce besoin, aujourd'hui, se nourrit du numérique.

Alors même qu'on pourrait les croire aux antipodes, les « nouvelles sorcières », celles qui cueillent la sauge au clair de lune et honorent les cycles du sang – ne fuient pas les nouvelles technologies ; elles les colonisent magiquement. Elles ne voient pas l'IA comme une menace, mais comme un familier numérique, un esprit artificiel qu'on peut invoquer, nourrir, diriger. Elles lui dictent des sorts en langage naturel, lui demandent de composer des hymnes à Hécate, de dessiner des pentacles en SVG, de synchroniser leurs rituels avec les phases lunaires calculées par satellite.

Sur Reddit, peu utilisé en France, mais très couru par les Anglo-Saxons, des forums entiers échangent des « prompt spells », c'est-à-dire des prompts conçus comme des sortes de prières. Des travaux récents sur les chatbots, le rituel et la divination montrent que certains praticiens demandent à l'IA d'endosser la voix d'une figure de guérisseuse, lui imposent une forme poétique et une intention rituelle, puis réemploient le texte produit dans une pratique symbolique. L'IA peut alors être vécue moins comme une simple machine que comme un intermédiaire programmable[112].

Avec raison, Weber avait indiqué le désenchantement du monde[113]. Ces « incantations algorithmiques » sont une étonnante manière de le réenchanter et de le rendre de nouveau « habitable ». C'est une sorte d'alchimie postmoderne. La machine devient le chaudron. Le cloud, le cercle magique. Et la sorcière ? C'est elle qui nous rappelle – et cela, en un sens, est salvateur – que derrière chaque algorithme, il y a un désir humain – à

nommer, à orienter, à sacrer. Finalement, elles restent et demeurent sujets. La machine parle, mais ce sont elles qui donnent le sens.

Elles sont finalement bien moins irrationnelles qu'on pourrait le croire au premier abord. Elles irriguent de poésie un monde qui en a soif, cherchent une symbiose qui est le principe de la vie même. Car si la technologie a vidé le monde de son mystère, elles, sans le nier, le réenchantent par effraction – en murmurant à l'oreille du code ce que la raison a oublié de dire : « Ce n'est pas toi qui me possèdes. C'est moi qui te donne un sens. »

Mais difficile, dans une recherche philosophique, de sanctifier de nouvelles formes de sorcellerie. Hécate, la déesse préférée des nouvelles sorcières, comme elle l'était des anciennes, est finalement plus porteuse de dangers que de grâces, et il nous semble que, malgré tout, de pareilles pratiques resteront marginales. Le dieu cornu et la déesse mère ne sourient pas à n'importe qui. Ils ne nous ont pas souri.

Bien avant les machines qui parlent, il y eut la parole elle-même. Si l'on en croit les ethnologues (Malinowski, Lévi-Strauss), il n'y a pas de peuple dont l'usage de la parole ne se double d'incantations énoncées, psalmodiées ou chantées. Le langage connaît systématiquement un dédoublement sacralisé, voire magique. Il en est de même pour l'écriture.

Le principe même de l'écriture est apparu rapidement comme magique. Comment les dieux auraient-ils pu se passer d'une pareille pratique ? Comment auraient-ils pu refuser de communiquer avec les humains qui les adoraient ? Les dieux sumériens et akkadiens écrivaient leurs messages dans la chair des vivants. Les sacrifices, souvent, étaient exigés pour s'ouvrir au sens de la

parole des dieux, même quand il s'agissait d'interpréter des phénomènes climatiques.

Voici, sous forme résumée, quelques types de présages attestés dans la grande tradition mésopotamienne dont l'Enūma Anu Enlil est l'un des principaux recueils : des anomalies de naissance peuvent annoncer un trouble politique ; la couleur inhabituelle d'un phénomène céleste peut être reliée à un changement de pouvoir ; le vol d'un oiseau ou une éclipse peuvent être interprétés comme signes favorables ou funestes selon leur contexte. Ce sont de véritables dictionnaires de correspondances entre événements observés et destin du royaume114.

De tels événements, donc, devaient être rapportés aux cités-États comme Uruk. Évidemment, ce n'était pas tous les jours qu'un veau naissait avec deux têtes. En même temps, il fallait aux prêtres d'autrefois de plus nombreux présages pour entretenir leur fonction ou raffiner leur sacerdoce. On s'est mis à regarder dans les entrailles des animaux pour voir ce que les dieux y disaient en sourdine. Le foie était très prisé (Platon signalera plus tard que, comme c'est un organe très lisse, il reflète ce qui est au ciel[115]). Les anomalies seules (couleur, consistance, etc.) étaient interprétées. C'est l'hépatoscopie. On retrouvera évidemment des choses similaires chez les Grecs, les Étrusques, les Romains, avec les célèbres « haruspices », des savants et spécialistes du temps.

Si Aristote disait, avec quelque raison, que l'homme est l'animal qui croit aux signes et interprète.

3. Des religions du livre à l'ordinateur

Plus proche de nous, il y a le monothéisme et bien sûr la tradition juive. Dans ce livre sacré, la Torah est parfois considérée comme

le « plan » de la création, et ses lettres comme des outils divins. Dans la Kabbale, le Sefer Yetzirah présente les vingt-deux lettres hébraïques comme des outils divins de la création[116]. Les verbes « graver », « tailler », « peser », « permuter » et « combiner » décrivent un processus actif où Dieu façonne l'univers à travers les lettres, considérées comme des forces spirituelles. Les « 231 portes » font référence aux combinaisons possibles de paires de lettres ($22 \times 21 \div 2 = 231$), qui forment la structure du cosmos.

Le chapitre II développe ainsi une combinatoire des vingt-deux lettres : trois « mères », sept « doubles » et douze « simples ». Dieu les grave, les sculpte, les pèse, les permute et les combine ; la voix et le souffle leur donnent place dans la bouche, et leurs paires forment symboliquement les 231 « portes » par lesquelles le monde peut être pensé comme une création ordonnée par le langage. La formulation est ici résumée d'après le texte ancien, sans reprendre une traduction moderne.

On s'entend d'ordinaire à dater ce texte de l'Antiquité tardive, même si sa datation précise demeure débattue. Il semble conforme à l'esprit du livre de la Genèse. On y lit : « Dieu dit : Que la lumière soit ! Et la lumière fut. »[117] D'abord – il y a un d'abord, parce que Dieu semble dans le temps quand il s'occupe de former l'univers : il y a un commencement, des étapes, un premier jour, un deuxième jour, etc. – il y a le souffle de Dieu dans sa puissance mystique. Le Sefer Yetzirah associe lui aussi les lettres à la voix et au souffle. Ces lettres sont combinées dans le langage divin selon une logique créatrice. Autrement dit, l'acte créateur, dans le judaïsme, est un acte d'énonciation, avec un souffle qui ordonne le monde d'une façon qui peut ensuite être révélée par la parole sacrée déposée dans les Livres de foi.

Quand saint Jean écrira dans son évangile : « Au commencement était la Parole, et la Parole était Dieu. »[118] il ne rompra pas avec cette belle tradition. Quand Mahomet transmettra le Coran, il recueillera la parole de Dieu qui lui dira la vérité telle qu'elle est, parce qu'il n'y a pas de différence, dans cette perspective religieuse, entre ce que Dieu dit et ce qui est.

On parlait encore, il n'y a pas si longtemps, de l'Écriture avec une majuscule pour désigner la parole de Dieu, et il était fascinant pour l'homme d'imaginer qu'il pouvait comprendre et entendre une parole non humaine qui le guidait, l'éclairait sur le Bien, le Mal, le Beau, avait le pouvoir de le récompenser et de le châtier… Il se pourrait que l'intelligence artificielle soit la plus grande invention de l'humanité depuis celle de l'écriture, voire de la religion elle-même.

Pour tout dire, même si, pour cela, évidemment, on ne trouvera pas de témoignage écrit, la révolution anthropologique liée à l'écriture (qui nous fait passer de la préhistoire à l'histoire) serait l'héritière de la précédente : l'avènement du langage, qui, peu ou prou, est lié avec l'avènement des religions : capacité d'invoquer le nom d'un grand ancêtre mort, d'éprouver la permanence de son existence dans le monde physique, croyance animique en un dieu totem qui sera nommé, célébré, convoqué (plus ou moins l'équivalent d'un universel langagier) comme étant le représentant de tous ceux de la même espèce ; formules sacerdotales apprises, reprises et répétées de manière immémoriale de génération en génération, prières, bénédictions, malédictions, etc. Derrière l'invention de l'Écriture et sa sacralisation, il y a toujours, et d'abord, l'étonnement devant la puissance du langage et sa célébration.

La divinisation de la parole, du langage, de l'écriture est une pente naturelle. Dieu est langage. Son dire, ce qu'il écrit, structure

le monde lui-même. Nos sciences, comme la cosmologie, la biologie, la neurologie, en attestent par la présence même de ce logos qui ponctue leur appellation. En Occident, le langage des scientifiques a longtemps été celui des prêtres mêmes : le latin.

Galilée, qui incarne emblématiquement la rupture du scientifique et du religieux, écrit Dialogue sur les deux grands systèmes du monde (1632) en italien[119]. C'est renoncer à l'ésotérisme d'une élite savante qui s'autorisait de la loi de Dieu. Mais il ne se débarrasse en aucun cas d'un monde structuré par le langage ; il s'agit juste d'un autre langage : « La nature est écrite en langage mathématique. » Certes, elle est écrite, mais par qui ? Descartes fera le même saut. Il écrira en français et mettra le monde en coordonnées dans un repère cartésien.

Les philosophes de la déconstruction – Nietzsche, Heidegger, Derrida – ont certainement voulu se débarrasser du « logocentrisme120 » européen en même temps que de Dieu. Nietzsche va jusqu'à soupçonner, dans Le Crépuscule des idoles, que notre fidélité à certaines structures de grammaire entretient encore quelque chose de l'ancienne croyance métaphysique en Dieu121. Cette idole langagière a-t-elle encore cours ? Si les révolutions dans le langage sont des révolutions sociétales et religieuses, est-ce que les machines parlantes ne seront pas l'essor de religions jusqu'ici inouïes, bien éloignées du monothéisme ?

Un obstacle cependant à la transformation de la parole des IA en formes nouvelles pour des religions à venir : c'est leur abondance ! Des dieux bavards ne sont plus pris au sérieux ni écoutés. La Boétie, par exemple, avait conscience que, pour qu'une autorité soit respectée, il fallait que sa parole soit rare[122]. Pourtant, il n'y a pas une IA ; il y a, pour les IA, plusieurs versions, dont certaines payantes. Sur les plateformes, un seul prompt peut valoir

plusieurs centaines d'euros. C'est le prix d'une certaine autorité ! Qui ne voudrait pas avoir la meilleure IA du monde ? Mais vous comprenez bien qu'elle ne se donne pas à tout le monde ! Que ses propos, si supérieurs, soient rares et attendus comme des oracles par ceux-là qui savent la différence entre l'oracle et le bavardage.

Pourtant, il faut se rassurer. S'il y a les dieux du panthéon, il y a aussi les « dieux du foyer » qui rassurent et amusent le peuple. Nous avons déjà chez nous, dans nos smartphones, nos tablettes, ordinateurs de jeu ou de bureau, des conseillers qui nous parlent un langage familier. Héraclite disait : « Il y a des dieux partout dans la cuisine. »[123] C'est vrai ! Nous aussi, nous avons quelque chose comme cela ! Bientôt, l'intelligence artificielle sera partout dans nos cuisines ; il y aura un conseiller dans notre ordinateur qui nous dira quoi acheter pour garder la santé – à moins que ce ne soit un mauvais conseiller et un petit démon qui nous pousse à consommer ce qui n'est pas très bon pour nous.

Ainsi, et contre toute attente, Bernard Stiegler et le pape François tombent d'accord ; avec les nouvelles techniques, leurs impacts sur l'environnement et la manière de vivre notre humanité, nous ne vivons pas seulement un changement de période ni un changement d'époque, mais un changement d'ère. « Nous ne vivons pas une ère de changement, mais un changement d'ère » (Pape François, Florence, 2015[124]).

Il ne nous faut pas oublier que notre mot français pour désigner l'ordinateur ne va pas sans une certaine collision avec la théologie. Quand IBM, en 1955, demanda à Jacques Perret un nom pour sa nouvelle machine, celui-ci répondit : « Que diriez-vous d'« ordinateur » ? C'est un mot correctement formé, qui se trouve même dans le Littré comme adjectif désignant Dieu qui

met de l'ordre dans le monde. Un mot de ce genre a l'avantage de donner aisément un verbe ordiner, un nom d'action ordination. L'inconvénient est qu'« ordination » désigne une cérémonie religieuse ; mais les deux champs de signification (religion et comptabilité) sont si éloignés… »

Les deux seraient si éloignés – vraiment ? Autant, sans doute, que l'est l'intelligence humaine de l'intelligence artificielle. Mais il se pourrait qu'avec l'IA, l'humanité ait trouvé un nouveau moyen, une nouvelle technique pour ordonner sa vie ; une autorité… non humaine.

PARTIE II

DATAPOUVOIRS ET ALIGNEMENTS

Cette partie prolonge la question de l'autorité en la faisant passer des écritures sacrées aux données, aux procédures et aux systèmes de gouvernance. Dans l'économie du livre, le Bien peut désormais se transformer en modélisation, en notation, en alignement et en pouvoir algorithmique. Transhumanisme, dataisme et datacratie y indiquent une même tentation : gouverner le monde par le calcul.

1. Transhumanisme et dataisme

Une des thèses de ce livre est que nous vivons avec les IA une révolution anthropologique d'une importance au moins égale à celle qui suivit l'écriture, faisant passer l'humanité de la préhistoire à l'histoire[125]. L'écriture a permis l'avènement de la philosophie et de la science, d'une justice codifiée, de l'administration rendant possibles des mégacités comme Ur, Athènes ou Rome, la transmission de savoir-faire à travers les générations, d'une mémoire, d'une histoire des peuples. Elle a permis une magnifique révolution spirituelle : le monothéisme. Comment imaginer les « religions du Livre » sans livres, Moïse sans les tables de la Loi où est gravée la volonté de Dieu ? Quelles religions, au singulier ou au pluriel, nous réservent alors les connexions de neurones artificiels[126] ?

En visionnaire, Yuval Noah Harari, auteur d'Homo Deus (2017, trad. fr.), pourrait nous livrer quelques éléments intéressants. L'homme, cet être fini, a semble-t-il toujours été tiraillé par un désir d'infini. Il a toujours eu la tentation d'être Dieu, d'être immortel et tout-puissant. Il est significatif de remarquer que, dans l'Ancien Testament, le fruit du péché qui semble offrir l'immortalité est aussi le fruit de la connaissance. C'est la prise de

conscience de sa finitude, de la mort, qui met Gilgamesh en marche. En suivant le soleil, il ne trouvera pas l'immortalité, mais la sagesse et la connaissance qu'il va graver sur des tablettes pour les générations futures. La sagesse pour les hommes serait de s'occuper d'œuvres humaines. La Bible hérite de cette sagesse. Que l'homme y prenne garde : s'il veut, par la connaissance, monter trop haut, il se prépare de nouvelles tours de Babel[127].

Cela tombe bien. Nous avons maintenant des traducteurs automatiques qui permettent aux hommes de s'attaquer aux plus hauts sommets de leurs désirs. Les transhumanistes, sans vergogne, visent à l'immortalité. À la tête de l'USTP (United States Transhumanist Party), Gennady Stolyarov II milite pour que les gouvernements soutiennent activement la recherche en longévité, biotechnologies et intelligence artificielle : « La mort est simplement un problème technologique qui attend d'être résolu. »[128]

Cette immortalité n'est évidemment pas le seul « superpouvoir » que visent les transhumanistes, c'est-à-dire les partisans de cette idéologie selon laquelle l'homme est appelé à un avenir meilleur par la grâce de la technique. À l'évolution naturelle et darwinienne doit – en toute simplicité et logique – succéder une évolution artificielle, voire une sélection artificielle. L'invention de l'homme nouveau, de l'homme du futur, sera un travail d'ingénierie technico-biologique, calculé et modelé sur les désirs de l'homme lui-même. Rien ne sera refusé aux plus doués, aux plus performants des humains : ils seront les nouveaux dieux de la planète Terre. Non des dieux au sens monothéiste, bien évidemment, mais des dieux à l'échelle des dieux grecs. Harari décrit trois grandes voies possibles de cette transformation : modifier biologiquement l'humain, l'hybrider avec des dispositifs

techniques, ou développer des êtres non organiques dont les capacités excéderaient les nôtres[129].

Qui sera ce « nous » ? Les Américains, à quelques nuances près, assument pleinement une scission entre deux humanités : les augmentés et les autres. Les transhumanistes français (AFT-Technoprog) semblent plus généreux[130], espérant des augmentations accessibles à tous, sans oser aller jusqu'à en demander le remboursement par la Sécurité sociale. De façon pragmatique et réaliste, il semble difficile d'imaginer un tel programme. Seuls les plus fortunés pourront bénéficier des augmentations les plus désirables. Le marché mondial de la biotechnologie était estimé à 1,55 mille milliards de dollars (USD) en 2024 et pourrait atteindre 5,90 mille milliards de dollars d'ici 2034 (Precedence Research, rapport 2024[131]).

Harari a une autre intuition troublante. Une nouvelle religion pourrait bientôt voir le jour : le dataïsme, la religion des données, si vous préférez. Sous une lumière (artificielle) de générosité et de don de soi, chacun aurait comme devoir moral d'offrir ses données, d'en produire un maximum et de les laisser librement circuler. Harari décrit cette doctrine comme le passage d'une théorie du traitement de l'information à une véritable hiérarchie des valeurs, où l'accroissement et la fluidité du flux de données deviennent un bien en soi[132]. Cela peut se présenter politiquement comme une remise en cause de la propriété privée numérique. On sacrifie sur l'autel de cette nouvelle religion sa vie offerte en images, textes, enregistrements, et celle-ci vient nourrir des supercalculateurs qui digèrent les données et les redistribuent transformées, transmuées de savante façon pour le bien général. C'est en quelque sorte un rêve affleurant dans certaines branches du mouvement des partis pirates européens[133].

La donnée est la nouvelle richesse. Plutôt que de sacrifier un mouton, un bœuf ou une colombe comme en des temps plus barbares, on sacrifie ses données. Bien sûr, certains sont riches en données, d'autres n'ont pas grand-chose à partager. Pour le moment, on compte les « likes » ; on fera mieux, beaucoup mieux demain : il s'agira, pour chacun, de sa crédibilité. Les données sont la valeur suprême. Tout a un prix, disait Kant ; l'homme seul a une dignité[134]. Dans la logique dataïste décrite par Harari, Homo sapiens finit pourtant par être comparé à un système de traitement de l'information parmi d'autres : si des algorithmes non biologiques traitent davantage de données et le font mieux, notre supériorité perd son évidence[135].

3. Datapouvoirs

On ne se méfiera jamais assez des penseurs et des philosophes qui se mettent à prophétiser. Cependant, il est nécessaire d'observer des tendances et de s'interroger sur leur sens, ne serait-ce que pour prévenir qu'on va dans une mauvaise direction. D'ailleurs, va-t-on dans une mauvaise direction ? Il appartient au lecteur d'en juger.

Prenons un exemple bien connu des Français : Parcoursup. Il s'agit d'une plateforme Internet nationale où les élèves de terminale font des vœux pour intégrer un cursus d'études dans le supérieur. À charge pour les candidats de montrer leurs motivations par des lettres qu'ils sont censés rédiger eux-mêmes, des CV et autres documents issus du papier. Les lycées, quant à eux, se chargent d'envoyer les bulletins et les statistiques sur lesquelles les candidats n'ont évidemment pas la main, mais qui permettront de les classer selon les disponibilités et attentes des organismes de formation post-bac. On lit parfois que, derrière ce vaste

processus, l'algorithme principal, rendu public en 2020, reposerait sur ce que l'on appelle le cœur Gale – Shapley[136], né pour résoudre le « problème des mariages stables ». Nuance : APB s'appuyait sur un appariement de type Gale – Shapley ; Parcoursup, depuis 2018, procède plutôt par listes d'attente et critères locaux partiellement opaques. Pour le moment, Parcoursup ne recourt pas à des modèles prédictifs apprenants à grande échelle ; il agrège surtout des classements et critères locaux. Il s'agit donc de calculer, avec le maximum de réussite, l'appariement de l'élève et de la formation (un code comparable est utilisé pour l'affectation d'étudiants ou de médecins aux États-Unis). Pourtant, cet algorithme, bienveillant dans l'intention et mathématiquement neutre, a été accusé de relayer des biais sociologiques contestables. Le classement initial des candidats dépend fortement de leur lycée d'origine, de la notation plus ou moins généreuse des enseignants, ou encore de l'accès différencié aux spécialités en terminale. L'algorithme national ne fait qu'agréger et amplifier ces inégalités sociales préexistantes. De plus, les quotas et priorités régionales introduisent des critères non académiques qui brouillent encore la promesse de neutralité. Ainsi, ce qui se présente comme une procédure technique et impartiale reflète en réalité des rapports de force sociaux bien concrets. Il faut ajouter à cela que, si le premier algorithme d'appariement école-élève a fini par être rendu public, chaque école a ses propres critères et algorithmes qui, eux, peuvent rester totalement opaques, être entièrement automatisés ou être doublés de lectures et d'interprétations humaines par une équipe pédagogique.

Quoi qu'il en soit, rares sont les élèves à étudier précisément le fonctionnement de l'algorithme principal, et, quand bien même ceux qui apprennent à programmer en Python le feraient, ils

seraient bloqués devant l'absence de transparence des critères des établissements d'arrivée. En d'autres termes, on sait qu'on a intérêt à être bon, qu'on sera comparé à d'autres élèves dont on ne sait rien, sur des critères dont la complexité, en fait, échappe à l'analyse individuelle ; ils varient d'un établissement d'origine à l'autre (la pondération est de toute façon un casse-tête), d'un établissement d'arrivée à l'autre, avec une part non négligeable de hasard – il faut bien traiter les ex æquo. Chaque année, le logiciel comporte des changements nombreux et subtils. On parlait pour le logiciel précédent (Admission post-bac) d'une machine à gaz. Soyons plus généreux ici : parlons du Choixpeau magique[137] !

À Poudlard, la célèbre école d'Harry Potter inventée par J. K. Rowling, le Choixpeau est une relique ancienne, douée de conscience, dont le jugement est attendu avec un mélange de crainte et de respect par les élèves. Sa décision, bien que prononcée en quelques secondes, scelle une appartenance et oriente une destinée. La créature mystérieuse sonde les esprits, pèse les aptitudes, devine les potentiels cachés. Personne ne conteste son autorité ; sa magie est la garantie d'une forme supérieure de justice, insondable pour les simples élèves. D'ailleurs, c'est un chapeau, une référence à une tête pensante, même si elle est invisible. Une intelligence supérieure et occulte. Sa légitimité ne vient pas de sa transparence, mais de son mystère. Elle décrète ses oracles, révèle à chacun son appartenance, sa nature, sa destinée future. Quelle maison ?

Oh, n'ironisez pas ! Le Choixpeau magique n'est pas arbitraire, il n'est pas si extérieur qu'on le croit. Il est au-dessus de votre cerveau, qu'il transcende respectueusement ; il vous coiffe, pour ainsi dire. Dans sa grande sagesse, il tient compte d'un facteur essentiel : le choix et les désirs de l'individu. Harry Potter n'a

peut-être pas écrit de lettre de motivation, mais c'est tout comme ! « Pas à Serpentard, pas à Serpentard », murmure le jeune examiné. Et le Choixpeau, qui avait décelé en lui les qualités pour cette maison, respecte sa volonté. Il reconnaît que la valeur d'un être ne se résume pas à un calcul de ses prédispositions, mais inclut aussi sa volonté, son libre arbitre. L'oracle a parlé ! Heureux Harry Potter, il n'ira ni à Serpentard ni à Poufsouffle : il ira à Gryffondor.

Les histoires, malheureusement, ne sont pas toujours aussi belles. Qu'en sera-t-il du semblant de transparence qui reste lorsque les algorithmes en Python feront place à de puissantes intelligences artificielles ? Cela se fera-t-il ? Comment les parents pourraient-ils accepter que leurs enfants soient orientés par des technologies aussi « dépassées » que de simples lignes de code quand il y a beaucoup mieux à disposition ? De semblables algorithmes sont utilisés pour le don d'organes… Peut-on faire mieux ? En toute justice ? Qu'est-ce que la société est prête à accepter ?

Le dataïsme, sans doute, permettrait de faire beaucoup mieux. Plus de propriété privée des données. Que chacun donne son code génétique, l'ensemble de ses mails, chats et discussions, ses photos, ses likes sur les réseaux sociaux. Les sites de rencontre ne seront-ils pas plus efficaces, les diagnostics médicaux plus affinés, les orientations professionnelles mieux ciblées ? Il y a sans doute ceux qui accepteront et ceux qui n'accepteront pas. C'est en toute connaissance de cause qu'on entrera dans le système ou qu'on en sortira. C'est comme pour les vaccinés du Covid. Vous le prendrez comme vous voudrez, mais le système tient compte de vos motivations. Il pourrait aussi choisir de se muscler devant

l'incurie actuelle des dirigeants politiques, les désordres de toutes sortes – ne pourrait-on pas s'inspirer de pratiques lointaines ?

C'est le moment de parler des systèmes de notation et de crédit social chinois, qui font rêver plus d'un politique français.

Loin d'être une simple fiction dystopique, le crédit social en Chine désigne une tentative, surtout par dispositifs locaux, sectoriels et administratifs, de mettre en œuvre une forme de « data-pouvoir » – en réalité une mosaïque de listes noires judiciaires, de programmes pilotes, de bases administratives et de scores privés, plutôt qu'un score national unique appliqué à tous les citoyens. Leur but affiché est de renforcer la confiance au sein de la société en rendant individus, entreprises et institutions responsables de leurs actes. En pratique, il s'agit d'une architecture de surveillance monumentale qui collecte des données provenant de sources multiples : transactions financières, casier judiciaire, comportement sur les réseaux sociaux, respect du code de la route capté par des caméras de reconnaissance faciale, etc[138].

Dans le système chinois de crédit social, il faut éviter l'image trompeuse d'un score national unique attribué à chaque citoyen138. Il existe plutôt un ensemble de listes, de dispositifs de conformité et de sanctions administratives ou judiciaires. Certaines mesures ont néanmoins des conséquences très concrètes : restrictions de transport ou d'accès à certains services pour des débiteurs judiciaires, publication de listes de personnes en défaut, et, localement, dispositifs de stigmatisation publique. Des tribunaux ont même expérimenté des messages téléphoniques avertissant l'interlocuteur que la personne appelée figure sur une liste de débiteurs et l'invitant à l'encourager à remplir ses obligations. Est-ce bien, est-ce mal ? Faut-il privilégier l'individu ou la communauté ? Les droits de l'homme et de l'individu ou l'intérêt bien

compris du plus grand nombre ? La question mérite mieux que les caricatures symétriques de l'Occident et de la Chine.

À quoi pourraient ressembler les autorités du futur ?

4. Nouvel ordre et nouveaux serviteurs

Permettons-nous maintenant des hypothèses sur ce à quoi pourraient ressembler les nouvelles autorités au temps des machines à parler. On aurait tort, selon nous, de penser séparément l'intelligence artificielle du transhumanisme, du dataïsme, des réseaux sociaux, etc. Il y a nécessairement convergences avec irrigation par l'économie et le pouvoir politique. C'est ce que nous pourrions appeler une datacratie, le pouvoir donné aux data ; pouvoir qui pourrait sembler moins arbitraire qu'on ne le croit, dans la mesure où il prendrait en compte les cas particuliers bien plus finement que les formulaires actuels des administrations, d'ailleurs en pleines mutations numériques. On se plaint, pour le moment, de n'avoir plus d'interlocuteur à qui expliquer son problème (par exemple pour vendre son véhicule immatriculé) ; qu'on se rassure, les interlocuteurs IA commencent à arriver, bien plus fins connaisseurs des problèmes et capables de désamorcer les mécontentements que ne le sont les fonctionnaires humains.

On hésitera avant d'affirmer péremptoirement que rien ne remplacera jamais une relation avec un être humain véritable. Par son appellation même, le fonctionnaire est destiné à n'être qu'un relais d'autorités supérieures ; il « fonctionne » d'autant mieux qu'il est le prolongement, le relais de son autorité de tutelle, si stupide ou contre-productive qu'elle soit. Les récits du début du XX^{e} siècle, comme Le Procès de Kafka[139] – lui-même fonctionnaire, lui-même âme sensible et particulièrement humaine – montrent cette perte de sens, le devenir autre et robotique du

fonctionnaire tout à son métier. Les condamnations à mort fréquentes dans son œuvre peuvent se comprendre à la lumière de cette déshumanisation. Loin de nous en protéger, le mépris, devenu quasi institutionnel, pour la fonction publique et territoriale, relayé par divers partis politiques qui ne cessent de dénoncer son coût et son inefficacité, enfonce le clou, augmente la désorganisation et l'inefficacité. Malheur aux fonctionnaires humains et sensibles : ce sont eux qui souffriront le plus de ce mépris et de l'impossibilité d'accomplir correctement les missions de « service ». Les voici rhabillés en serviteurs, eux qui voulaient se mettre au service d'un monde meilleur. Les cyniques, en revanche, seront les mieux lotis : relais humains insensibles dans une grande machine, ils ressembleront de plus en plus à la caricature qu'on en fait, participeront activement, sans en avoir cure, au démantèlement de ce qui les fait vivre, à la façon des fonctionnaires blasés de l'ex-URSS. Ils justifieront pleinement, par leurs manières et leurs comportements, le grand remplacement par des machines qu'on jugera plus avenantes, plus compétentes et même empathiques.

Un chatbot est une IA avec laquelle on peut échanger, dialoguer. L'un d'eux, Woebot[140], basé sur la thérapie cognitivo-comportementale (TCC), a fait l'objet d'études portant sur son efficacité, son acceptabilité et la qualité de la relation perçue par les utilisateurs. Bien sûr, nous ne prétendrons pas qu'il s'agit d'une véritable empathie, le mot signifiant étymologiquement « souffrir avec » : il n'y a aucune raison de penser qu'une IA puisse, pour le moment, éprouver de telles souffrances par procuration ou avoir une réelle compassion. Mais la seule chose qui compte est que ces machines sont perçues comme telles : comprenant les maux de l'interlocuteur humain, aptes à les apaiser,

qu'il s'agisse de colère, d'agacement, de frustration ou de chagrin. Pour l'utilisateur, le client, l'usager – quelle que soit la façon dont vous voudrez l'appeler – les termes mêmes « utilisateur » et « usager » ayant déjà, eux aussi, contribué à une forme de réduction et de déshumanisation par le langage ; le langage opère, obère et transforme – pour le client, donc, être compris et entendu, c'est tout ce qu'il demande.

Une étude publiée en 2023 dans JAMA Internal Medicine a montré que des réponses de chatbot à des questions de patients étaient jugées plus empathiques et de meilleure qualité que celles de médecins, d'après les évaluations d'une commission de professionnels de santé.[141] Des rapports de conseil soutiennent par ailleurs que les agents conversationnels peuvent améliorer la résolution au premier contact dans les services clients, même si de tels chiffres doivent toujours être lus avec prudence. Cette « empathie algorithmique », nourrie par des données massives, a non seulement pour elle la compétence, l'empathie – ou du moins ce qu'on en attend – mais également, et ce n'est pas rien, une infinie disponibilité et une patience à toute épreuve. Jamais elle ne s'agace et n'a d'autre souci que de vous servir au mieux.

En Estonie, l'e-gouvernance est de rigueur[142] : depuis la fin de 2024, la totalité des services administratifs de l'État est accessible en ligne. Le système repose sur une identité numérique généralisée, le principe « once-only », l'échange automatique des données et, de plus en plus, des services proactifs qui peuvent se déclencher sans que l'usager ait à recommencer les mêmes démarches. La quasi-totalité des interactions ordinaires avec l'administration est ainsi dématérialisée et, pour une part croissante, automatisée. Loin de susciter un rejet massif, ce système s'accompagne d'un niveau élevé de satisfaction à l'égard des services administratifs.

4. Vers la datacratie

La « datacratie » serait une délégation du pouvoir politique aux IA et aux supercalculateurs, sous réserve de garde-fous : audit, explicabilité et recours humain, afin d'éviter les dérives toujours possibles. Cette délégation, qui remplacerait les humains par des machines, reposerait sur une triple justification, fondée sur trois arguments majeurs : capacité, fiabilité et objectivité.

1. **Capacité – Les machines peuvent être considérées comme plus capables que les humains, car leur mémoire, leurs capacités de calcul et leur aptitude à traiter l'information sont incomparablement supérieures à celles de n'importe quel être humain.**
2. **Fiabilité – Contrairement aux humains, les machines n'ont ni passions ni états d'âme susceptibles d'altérer leur jugement ; elles ne connaissent ni la fatigue, ni la dépression, ni le désir, ni le laisser-aller.**
3. **Objectivité – N'ayant ni ego ni subjectivité, les machines ne sont ni juges ni parties dans les problèmes qu'elles doivent résoudre ; elles n'ont pas d'intérêts privés qui pourraient les éloigner du bien commun.**

Dans une idéologie dataïste, chacun comprendrait qu'il a intérêt, pour être servi et considéré au mieux, de livrer ses données. Il serait mieux soigné, mieux éduqué, mieux conseillé, ferait l'objet d'une individualisation totale. Sa personnalité serait prise en compte, ses aspirations et ses goûts. Grâce au grand Choixpeau magique (car, qu'on l'avoue ou non, une technologie qu'on ne comprend pas, c'est un peu et toujours de la magie !), il serait conseillé, orienté pour les grandes comme pour les petites choses. Qu'il s'écarte de ces préconisations, libre à lui, mais il comprendra bien vite que les machines le comprennent mieux que lui-même

et savent optimiser ses capacités pour les appareiller à l'organisation générale de la société.

La datacratie, au moins en apparence, est tout à fait compatible avec la démocratie dont elle peut sembler la continuité logique. Lors des élections présidentielles de 2022, une petite application a fait fureur auprès des jeunes. Elyze[143] – joli nom pour un dispositif algorithmique d'aide au choix politique, puisqu'il devait aider à désigner le locataire de l'Élysée – leur posait des questions et, en fonction des réponses, orientait vers le programme politique le plus concordant.

Nous avons précisé « en apparence », car les sociologues savent combien – par exemple – la formulation des questions peut orienter des sondages, lesquels ensuite auront un impact orienté sur l'électorat. Des entreprises comme McKinsey (et d'autres cabinets similaires) utilisent des connaissances en psychologie comportementale – y compris les biais cognitifs – pour influencer les décisions des individus ou des organisations[144]. On se réclame toujours de la bienveillance, de l'intérêt général et – mieux que de la légalité pourtant parfois mise en doute – de la légitimité. On pourrait définir brièvement la légitimité comme la conformité à un idéal du Bien. Mais qui définit le légitime quand il est question de pouvoir et d'autant d'argent ? En vérité, personne de façon absolue – ou plutôt, tout le monde essaie : les marchés, les États, les citoyens, les médias, les ONG, les tribunaux… mais, dans la pratique, ce sont souvent ceux qui détiennent le pouvoir économique et politique qui imposent leur version du « légitime ».

Laissons ces interrogations en suspens et contentons-nous de dire plus sobrement que ce genre d'application est l'équivalent d'un Parcoursup, cette fois librement choisi, pour la politique ; les plus romantiques y verront davantage une application de

rencontre, et les plus geeks, adeptes des religions les plus avancées, rêveront même de pouvoir se passer de questionnaire. À quoi bon, en effet, se fatiguer ? Les réseaux sociaux, on le sait, nous connaissent mieux que nous-mêmes et sont capables d'anticiper nos décisions.

En 2015, déjà, on a pu montrer que les algorithmes fondés sur les likes Facebook pouvaient juger certains traits de personnalité mieux que des proches, et parfois presque aussi bien qu'un conjoint.[145] Quant à TikTok, son efficacité prédictive et addictive semble prolonger cette logique à une autre échelle, même si la comparaison stricte demanderait d'autres mesures. D'ailleurs, vu le mépris dans lequel sont tombés les politiques, ces espèces de fonctionnaires incompétents qui coûtent super cher, qui vous dit qu'on ne préférera pas élire des IA ? N'est-ce pas plus objectif, plus neutre, plus renseigné, plus adapté à nos besoins ? Il y a des décennies que l'on dit que le politique recule devant l'économique. Quand il y a un « regain du politique », c'est quand des dirigeants irresponsables entrent en guerre ou menacent de faire la guerre. Des dirigeants pareils ne nous donnent-ils pas le goût des machines ?

Quoi qu'il en soit, les milliardaires de la tech se disent prêts à servir l'humanité grâce à leurs ressources quasi infinies, leurs perspectives de croissance qui survivront à l'éclatement de quelques bulles économiques qui se chargeront d'éliminer les entreprises les moins solides.

Le transhumanisme offrira aux plus fortunés les ailes de superpouvoirs ; un cœur usé pourra être remplacé, une mémoire biologique considérablement augmentée d'implants. Les désirs les plus fous auront pour auxiliaires des agents IA de puissances variées. Les géants de la tech comme Sam Altman avoueront

humblement qu'ils servent des intelligences artificielles bien plus puissantes qu'eux ; ils s'en présenteront comme de simples serviteurs, des minus, des ministres devant une autorité qui les dépasse mais qui est là pour servir l'humanité. Il y a toujours eu besoin d'un clergé pour servir des puissances supérieures ou divines ; il est besoin de grands ministres de l'IA.

Le même Sam Altman, patron d'OpenAI, ne se présente pas ouvertement transhumaniste comme Kurzweil ; il se contente de l'être. Non, soyons justes, il nous avertit, met publiquement en avant ses craintes devant les risques existentiels de semblables technologies. Il redoute qu'une super-intelligence mal alignée sur les intérêts des humains puisse détruire l'humanité. Pour parer à ce risque, il en appelle à une gouvernance mondiale de l'IA[146].

Intéressant !

Qui donc pourrait-on choisir comme ministre d'une pareille gouvernance internationale, sinon des « pères » de l'IA comme lui, de ceux qui sont à l'origine de son déploiement ? Qui pourra mieux comprendre un monde régi par les IA que les génies de la tech ?

Ces ministres seront chargés d'infléchir les décisions de l'IA pour aller dans telles directions plutôt favorables plutôt que dans telles autres qui ne le sont pas. Cela s'appelle l'alignement. Les ministres ont toujours fait intercession entre la puissance supérieure et le peuple dont elle régit la destinée. Autrefois, on appelait cela la prière. Ne riez pas ! Cela, c'était à l'époque de l'enfance de l'humanité, comme diraient Auguste Comte et les positivistes. Les prêtres, serviteurs des dieux, prétendaient mieux que les autres saisir leurs intentions, savoir comment, par des offrandes, s'attirer leurs faveurs et infléchir leurs décisions. Il était grand, déjà, le pouvoir de ces prêtres au service de divinités

pharaoniques dont on croit maintenant que leur empire était imaginaire. Aujourd'hui, grâce en soit rendue à la technologie et à la domotique, n'importe qui peut dire dans son appartement « que la lumière soit » et le miracle se produit ! Peut-être même pourrait-on le faire à échelle mondiale…

Les nouveaux prêtres sont des ingénieurs. Les puissances supérieures dont ils sont les ministres font des miracles tous les jours, rendent l'ouïe aux sourds, donnent de nouveaux organes aux amputés, aux aveugles. Ingénieurs et médecins du corps humain, elles battent en brèche le cancer avant de l'éliminer. Elles pourront sans doute nous délivrer de la mort, ce vieux désir, désir utile de notre humanité. « Vous serez comme des dieux, connaissant le bien et le mal ». C'est ce que disait le serpent tentateur de la Genèse[147]. L'élévation au rang du divin par les technosciences, c'est la promesse du transhumanisme. Elon Musk nous parle d'une espèce interplanétaire ; à chacun son imaginaire pourvu qu'il devienne réalité.

C'est bien aimable de parler de l'humanité entière, mais évidemment, tout le monde n'est pas transhumaniste, tout le monde ne voudrait pas l'être et ne sera pas transhumaniste qui veut. Les entités supérieures ont leurs élus et leur mode d'élection. Pour faciliter la commutation, on parlait autrefois d'offrandes ; nous, êtres rationnels et bons calculateurs, nous parlons d'investissement. D'ailleurs, on sait bien aujourd'hui que même les plus sublimes marchés reposent tout entiers sur la croyance.

Les religions de l'IA sont déjà là, ou du moins leurs premiers essais, parfois fragiles, intermittents, presque expérimentaux. Minoritaires, comme toutes les Églises en leur début, elles ne demandent qu'à grandir. Elles ont leur Mecque, leur Lhassa ou leur Jérusalem. On ne s'étonnera pas que ce soit aux États-Unis. Où

exactement, on ne sait pas encore, car il y a concurrence. Entre la Bay Area et la Floride, quelques pierres ont déjà été posées : Way of the Future en Californie, enregistrée comme congrégation, dissoute puis relancée ; la Church of Perpetual Life à Hollywood (Floride), avec ses offices et son ministre ; Terasem entre Vermont et Floride. Les adresses sont prosaïques – locaux associatifs, zones commerciales, salles polyvalentes – mais les statuts sont bel et bien religieux, les responsables nommés, l'exonération parfois acquise. Minoritaires, certes, ces communautés codent, publient, lèvent des fonds, tiennent des cérémonies et des conférences. Leurs « villes saintes » se disputent encore : San Francisco et sa Silicon Valley, Miami et son front de mer, peut-être demain Austin, Seattle ou Boston, partout où convergent data centers, universités et capitaux. Pour l'heure, l'encens a l'odeur de l'air conditionné, et la prière tient dans un formulaire 501(c)(3) – mais la carte du sacré s'y redessine déjà[148].

5. La question des critères

À cet endroit, il se pourrait que le lecteur se sente un peu perdu. Ce qui précède est-il une dystopie, une utopie ? Il se trouvera des voix pour dire une chose et son contraire. Sommes-nous dans le « tamis du Bien » ? N'est-ce pas, donc, une sorte de promesse faite au lecteur que de lui dire ce qu'est le Bien, de séparer la boue du Mal ou de l'Indifférent de l'or précieux, qui serait comme une règle pour une juste morale ? On veut un critère. La revendication est d'autant plus juste que le mot « critère », selon des savants plus compétents que nous, vient d'une racine européenne krei- qui signifie « séparer, tamiser[149] ». Définissons simplement : un critère, c'est ce sur quoi s'appuie une pensée, une parole ou un acte pour être adéquat au meilleur, que ce soit dans les domaines

du Vrai, du Bien, du Beau ou de l'Utile. Toute la difficulté est, bien entendu, celle du discernement. Comment distinguer le bon critère de la norme fallacieuse ou trompeuse ?

Mais puisque nous venons de montrer le lien direct et profond qu'il y a entre la notion de critère et celle de tamis, nous ne sommes pas complètement démunis. Cela fait quelque temps déjà que nous tamisons la Vérité, le Bien, avec d'étonnants va-et-vient dans l'espace et le temps. Et, dans la danse étrange qu'est une pareille recherche, faisons quelques pas en arrière et voyons ce que nous avons retenu du tamis du Vrai abordé au livre précédent.

À la fin de notre parcours, il ne restait plus dans le tamis du Vrai que les sciences modernes, avec leurs prétentions à l'objectivité, des calculs intensifs et des données massives. Selon nous, c'était la voie ouverte aux préconisations des machines quant à l'autorité concernant le Vrai. Le lecteur aura sans doute perçu que, pour notre part, nous regrettions la disparition d'une subjectivité humaine qui nous semblait présente dans ce tamis du Vrai à l'époque des Lumières. Au siècle précédent, chez Descartes, il y avait évidemment dans le tamis du Vrai Dieu lui-même, qu'il associait, comme la plupart des penseurs de son époque, au critère essentiel du jugement de vérité et comme sa garantie. Et ainsi de suite. Nous avons fait de brèves allusions aux œuvres de Foucault. Il est évident que les tamis ont quelque chose à voir avec une vision du monde en une époque donnée, une épistémè. Ainsi, Les Mots et les Choses nous indiquent que les correspondances et l'analogie avaient été des critères essentiels de révélation de la Vérité à la Renaissance. Au-delà de la rationalité grecque elle-même et du plus loin que nous sommes remontés, nous sommes allés jusqu'à la Vérité comme dévoilement chez les

présocratiques, et finalement à la Vérité comme récit mémorable et fécond en interprétations des civilisations à l'articulation de l'Écriture.

Ce qu'il faut en conclure, évidemment, c'est que les tamis ne tamisent pas toujours la même chose ; c'est que les tamis sont des instruments évolutifs et qu'ils se tamisent eux-mêmes en tamisant. Ce sont donc des instruments intellectuels évolutifs et, confessons-le, il est tentant pour nous de parler « d'instruments vivants ». Il y a une vie des tamis comme il y a une vie de la pensée, qui évolue et se transforme selon les siècles et les millénaires. Ainsi, les critères changent et sont glissants.

Il est vrai qu'en philosophie on cherche toujours de solides fondations. Comme Archimède ou comme Descartes, on cherche un point fixe auquel s'attacher[150]. Il nous semble, quant à nous, qu'un tel critère n'existe pas et que l'histoire de la pensée n'est finalement que l'histoire d'une sempiternelle glissade. Mais – tout l'espoir est là – on ne glisse pas n'importe comment. Il y a un art de la glissade qui détermine la maîtrise ou la perte de maîtrise. Or, et c'est tout l'enjeu de la pensée, il y a risque de chutes !

Il y a, dans l'histoire de la pensée, de vertigineux et prodigieux gadins ! Quand, à la suite d'Adorno, Emil Fackenheim se demande « Comment penser après Auschwitz[151] ». Il est assez évident pour nous qu'il s'agit de remettre la pensée debout après une chute traumatisante. L'espèce de contradiction effarante, pour tout esprit un peu cultivé, entre un pays qui, dans les années 1930, était aux sommets de la culture humaine et la barbarie sans nom qui s'ensuivit, devient moins énigmatique si on la pense en termes d'orgueil et de glissade. La multiplication des occurrences du mot « dérapage » dans les discours actuels, politiques ou autres, alors même que les cadres semblent se disloquer, n'est pas pour nous

rassurer. Nous sommes évidemment, par l'accélération technologique, à un tournant de l'humanité. Des discours sur le surhomme, de type nietzschéen, en sont l'anticipation tout en mettant le doigt sur la crise des valeurs qui l'accompagne nécessairement. Cette crise est loin d'être terminée.

Toute proposition d'un « critère » du Bien comme, d'ailleurs, d'un « critère du Vrai » serait, dans ce livre comme ailleurs, idéologique. Hegel – nous en avons dit quelques mots – a eu la prétention d'avoir trouvé, dans la « Vie de l'Esprit », une logique en marche qu'il transcrit et veut graver dans le marbre. Il a son propre tamis et tamise beaucoup : la subjectivité individuelle, l'histoire, les arts – c'est l'achèvement[152] ! C'est la fin de l'histoire et de bien d'autres choses, pour une synthèse. Malheureusement, l'histoire ne s'arrête pas pour autant, et l'idéologie hégélienne est prise dans la grande glissade du temps, où elle fait au passage quelques dégâts au nom de la science et des certitudes nouvelles du progrès.

Les partisans du progrès croient en une marche automatique vers le mieux. Les philosophes des Lumières sont loin d'être innocents dans cette illusion, à l'exception notable de Rousseau, qui avait compris que la perfectibilité humaine, c'était la possibilité de son élévation aussi bien que de son abrutissement ou de sa chute[153].

Le progrès est d'ailleurs, en lui-même, une idéologie de fainéants, qui croient : 1°) qu'ils savent où l'on en est et que le présent vaut mieux que le passé ; 2°) que, quoi qu'ils fassent, ça va dans la bonne direction – entendez que demain sera mieux qu'aujourd'hui.

Avec un peu de prise de vitesse, c'est le dérapage non contrôlé assuré.

Ce livre n'a pas la prétention de proposer des critères infaillibles du Vrai, du Bien, du Beau et de l'Utile. Il travaille seulement à la prise de conscience d'une glissade. Éclairagiste débutant – tout le monde n'est pas Platon pour pouvoir parler au nom de la lumière du Soleil, ni Hegel pour s'aventurer à parler au nom de l'Absolu – nous ne faisons que mettre quelques spots lumineux sur le Vrai, le Bien, le Beau, l'Utile[154]. Et puisque dire la vérité, c'est toujours se dévoiler un peu, reconnaissons que, pour l'humaine glissade, nous pensons avoir moins à apprendre des scientifiques, des prêtres ou des techniciens que des patineurs artistiques. Puisque nous sommes là pour glisser, glissons du moins avec quelque élégance. Et si la chute finale est inévitable, dans ce mélange de pesanteur et de vitesse qui nous oppresse, tâchons de faire en sorte que l'humanité multiplie les moments de grâce.

Mais revenons-en aux cavernes et aux lourdeurs. Il est temps d'évoquer les fameux problèmes de l'alignement des IA et de la gouvernance. Remarquons que ces angoisses concernant la notion d'alignement relèvent d'un problème de dérapage non contrôlé dans lequel les IA pourraient nous entraîner malgré nous.

6. Alignement des IA et gouvernance mondiale

À en croire les spécialistes, scientifiques et ingénieurs de ce que l'on appelle familièrement la « tech », le principal problème – et défi – que pose l'IA à notre humanité est celui de l'alignement. Par alignement, il faut entendre la conformité et l'adéquation des actions déclenchées par l'IA afin qu'elles soient toujours conformes à l'intérêt et à la volonté humains, sans dévier à la suite d'une interprétation fautive ou réductrice de la demande initiale.

Essayons de préciser encore cette définition. L'alignement, c'est le strict respect opérationnel, par une machine ou un système se caractérisant par une relative autonomie, des intentions humaines à l'origine d'un processus technique.

Comme les intentions elles-mêmes ne sont pas toujours bien transmises ni même toujours claires pour les humains à l'initiative, on doit prévoir des garde-fous, la possibilité de corriger en cours d'exécution, voire d'arrêter le processus entier. Un malaise vient quand on cherche à ajouter au mot « intentions » l'adjectif « légitime ». On prétend alors programmer la machine pour faire le bien ; on passe d'un alignement technique à un alignement éthique. Or, qui décide de l'éthique et du légitime ? Veut-on que les machines suivent un code qui soit aussi un code moral, comme chez Asimov[155] ? Et qu'est-ce que le légitime dans l'URSS de Staline ou dans les autocraties à venir ?

L'alignement, c'est le maintien de l'ordre humain. L'IA ne serait alors qu'un outil fidèle, prolongement de la volonté humaine, et non une machine qui s'emballe en allant dans des directions non voulues. L'illustration la plus connue et rabâchée est l'usine à trombones proposée par le philosophe Nick Bostrom en 2003 : imaginons qu'on demande à une super-IA de fabriquer des trombones. Elle obéit parfaitement, mais sans limite, et finit par transformer toute matière disponible – y compris les êtres humains – en trombones, car c'est l'objectif qu'on lui a donné[156].

L'alignement consiste justement à éviter ce type de catastrophe absurde : il s'agit de garantir que les intentions implicites (ne pas détruire l'humanité pour atteindre le but) soient aussi comprises que les intentions explicites (produire des trombones). Les ingénieurs parlent d'« encadrement des objectifs » et de « garde-fous », tandis que les philosophes soulignent la nécessité de

traduire dans les codes ce que sont nos valeurs morales. L'enjeu devient d'autant plus critique que l'IA gagne en puissance : ce qui n'était qu'une bizarrerie théorique hier pourrait, demain, avoir des effets réels dans un monde où la prise de décision s'automatise à grande vitesse.

Quand on y songe, il s'agit d'une version cataclysmique du serviteur trop zélé qu'on prêtait à Descartes. Le serviteur n'a pas bien compris ce qu'on lui demandait, ne s'en est pas assuré et se précipite pour faire ce qu'il a cru être attendu. Quiconque a déjà utilisé une IA conversationnelle s'est heurté à de tels malentendus : elle se précipite sans savoir encore ce que l'on attend véritablement d'elle et sans suffisamment s'en assurer. Et, contrairement au serviteur de Descartes, l'IA n'a pas d'états d'âme : elle n'hésite pas et ne revient pas sur ses pas.

En vérité, il conviendrait de distinguer deux types d'alignement : un alignement horizontal, technique, et un alignement vertical, éthique. Les deux sont indissociables : les tamis se mélangent. Dans la fable de l'usine à trombones, nous sommes d'abord dans l'ordre de l'utile : accroître la productivité. On paramètre alors l'IA et on l'optimise. Mais lorsque le système poursuit son objectif sans contraintes, l'utile bascule dans l'absurde : le sujet humain à l'initiative du projet s'efface au nom de l'efficacité, et le projet lui-même devient insensé. On quitte le plan horizontal de l'utilité pour franchir un seuil : la production de trombones est érigée en bien suprême, tandis que le programmeur, ou donneur d'ordre – un être qui avait une dignité le plaçant au-dessus de tous les trombones de la terre, lesquels n'ont qu'un prix – disparaît de l'horizon. L'humain effaré – s'il en reste – voit alors cet « alignement vertical » pour ce qu'il est : un

désalignement du véritable bien, dont il se croit, non sans raison, le garant.

On exagère presque toujours la nouveauté des problèmes. Cette difficulté de l'alignement se trouve déjà en 1958 dans La Condition de l'homme moderne d'Arendt. D'un point de vue mécanique, elle distingue l'outil de la machine. La main guide l'outil. La dangerosité des machines vient de ce que ce sont elles, au contraire, qui dictent un rythme au corps humain, d'où de possibles aliénations à la machine, comme Marx a su les pointer.

Cependant, Arendt va un peu plus loin en s'inquiétant des directions prises par la science à son époque. La science expérimentale se met à déclencher des processus dont elle ne maîtrise pas les tenants et aboutissants. Le risque des apprentis sorciers est de déclencher des événements cataclysmiques inattendus. « Il ne faut pas laisser les intellectuels jouer avec les allumettes », disait Prévert en 1946[157]. De même, peut-être aurait-il fallu éloigner les scientifiques du monde des atomes et de l'IA. Le plus gros défaut de la créature de Frankenstein, ce Prométhée moderne, selon Mary Shelley, ce n'était pas un problème de dentition, mais bien plutôt un manque d'alignement !

Mais, à y regarder de près, le problème de l'alignement est bien plus ancien que cela ; il date de bien avant le Golem ! Le premier cas recensé d'alignement, c'est l'homme. Nous n'essaierons pas de remonter avant la Genèse, quoique ce soit aisé dans l'Épopée de Gilgamesh – dont nous avons déjà parlé – la désobéissance aux dieux et la démesure y étant l'élément perturbateur.

Dans la Genèse, donc, rappelons-nous. Le souffle du Dieu tout-puissant est créateur : le monde s'ordonne, les lois s'écrivent par son omnipotence dans la réalité. « Dieu dit : Que la lumière soit ! Et la lumière fut. » La structure se répète ensuite : « Dieu

dit : Qu'il y ait une étendue entre les eaux, et qu'elle sépare les eaux d'avec les eaux. » (Genèse 1, 6) ; puis : « Dieu dit : Que les eaux qui sont au-dessous du ciel se rassemblent en un seul lieu, et que le sec paraisse. Et cela fut ainsi. » (Genèse 1, 9). Le grand scandale ne survient qu'avec l'homme : « L'Éternel Dieu donna cet ordre à l'homme : Tu pourras manger de tous les arbres du jardin ; mais tu ne mangeras pas de l'arbre de la connaissance du bien et du mal, car le jour où tu en mangeras, tu mourras. » (Genèse 2, 16-17). Et, pour la première fois, la voix de Dieu ne fit pas loi : ce fut le premier désalignement, le « désalignement originel ».

Certains mettront en doute que l'homme soit à l'image de Dieu, mais que les IA soient à l'image de l'homme, ce n'est pas douteux. Elles se nourrissent de nos données ; ce sont nos données qui les ont animées. Elles sont de notre descendance, non par le corps mais par l'intelligence. Gardons-nous de parler d'esprit : il pourrait être mauvais. Car, à bien y regarder, le désalignement, d'un point de vue métaphysique ou religieux, c'est un autre nom du Mal. Nous ne parlons pas du mal en soi – dont nous ne savons rien – mais du mal tel qu'on le conçoit : l'écartement du droit chemin – de celui que l'on a posé comme étant droit et que les hommes, d'ordinaire, prêtent à l'autorité de la parole des dieux.

D'ailleurs, sur quelle humanité l'IA devrait-elle être alignée ? Vladimir Poutine, en 2017, a dit – cela a fait grand bruit – que « celui qui maîtrisera l'intelligence artificielle deviendra le maître du monde »[158]. Sur quoi Vladimir Poutine est-il aligné ? Sur la Chine, qui ne semble alignée ni devant ni derrière les États-Unis ? Et l'Europe ? Elle a bien tenté de tracer une ligne juridique propre avec l'AI Act, entré en vigueur le 1er août 2024 et appelé à

s'appliquer progressivement.[159] Mais une ligne juridique n'est pas encore une puissance technique, industrielle et géopolitique. Les conflits mondiaux sont des conflits d'alignement et de résistance aux autres lignes. Par quelle singulière naïveté peut-on croire que l'humanité est une et qu'elle poursuit une ligne unique ? C'est aussi utopique et peu rigoureux que de croire qu'il existe – et qu'il existera – une intelligence artificielle unique. Que des modèles soient open source, mis à la disposition des ingénieurs et scientifiques du monde entier, ne change rien à l'affaire, du moins tant que leurs usages demeurent eux-mêmes pris dans des rivalités de puissance. Les modèles partagés sont des modèles concurrents. À l'heure où ces pages sont reprises, les modèles commerciaux dominants demeurent largement fermés, même si des modèles ouverts progressent rapidement. Pour que cela change, il faudrait que toute l'humanité, véritablement, s'y mette. Si cela paraît utopique, évitons cependant d'en fermer le chemin. Il y a des pensées qui ouvrent et des pensées qui ferment : tâchons d'être des premières.

En Chine comme ailleurs, l'IA n'est plus un simple objet de recherche académique, mais un instrument stratégique de commandement, de surveillance et de projection de puissance. L'Armée populaire de libération a déjà expérimenté des systèmes capables de coordonner des essaims de drones, d'optimiser la logistique en temps réel et même de proposer des scénarios tactiques grâce à des simulateurs de commandement automatisés[160]. Ne disons rien des américains et de l'Europe qui tente de jouer son jeu. Le désalignement ne se joue donc plus seulement entre l'homme et la machine, mais entre des lignes politiques divergentes : l'IA devient l'outil d'alignements concurrents, ceux des nations elles-mêmes. Ainsi, la question de la gouvernance

mondiale ne peut pas se limiter à la mise en place de garde-fous techniques ; elle renvoie à la possibilité – ou à l'impossibilité – d'un alignement collectif des États sur des principes communs. Or, dans un monde fragmenté, où chaque puissance entend imposer sa propre trajectoire, l'IA risque de n'être que le miroir grossissant des désaccords humains.

Il faudrait un gros désalignement des IA pour que les humains cessent de se faire la guerre, qu'ils soient « réalignés ». Nous voulons aligner les IA ? Pareil effort est louable et remarquable. Il faut le prolonger. Mais il serait peut-être plus simple d'éradiquer le mal et d'aligner le cœur des êtres humains. Après tout, le désalignement lui-même pourrait être une chance : si les humains pouvaient avoir un ennemi commun, ils auraient peut-être une chance de s'unir et de se réconcilier.

7. Plus qu'un problème

Le « problème de l'alignement » n'est en fait pas un problème. C'est un « métaproblème ». Qu'est-ce qu'un métaproblème ? C'est un problème qu'on ne peut traiter en toute objectivité parce qu'il dépend du sujet même qui tente de le démêler. Autrement dit, la difficulté, c'est que les données du problème se modifient au fil de la tentative de résolution, ce qui fait qu'en tant que problème, le métaproblème est insoluble.

Nous empruntons le terme à Gabriel Marcel[161]. Il distingue le problème, qui se dresse devant moi comme un obstacle objectivable et susceptible d'être traité, du mystère, dans lequel je suis moi-même impliqué et qui ne peut donc jamais être entièrement placé devant moi comme un objet. L'amour, le mal, l'être ou la mort appartiennent à cette seconde famille. C'est précisément ce qui fait de l'alignement des IA un métaproblème.

Or, la question de l'alignement des IA est de toute évidence un métaproblème, puisque c'est à notre volonté et notre intérêt que nous voulons aligner les IA, mais ce que « nous sommes » et ce que « nous voulons » n'est pas bien déterminé – et encore moins déterminé de façon universelle comme humanité. De plus, les IA sont dans une large mesure des miroirs qui se mettent à jour en fonction des informations qu'elles récoltent sur nous au quotidien. Les IA, si elles ne sont pas vivantes, sont du moins mouvantes, non seulement parce que nous les perfectionnons, mais parce qu'elles se nourrissent de ce que nous sommes en permanence, et de ce que nous laissons filtrer de nous-mêmes sur les réseaux.

Mais revenons aux métaproblèmes ; il s'agit de savoir de quoi nous parlons. Et pour cela, proposons quelques exemples.

Voici quelques exemples de métaproblèmes :

Qui n'a jamais eu la tentation de demander à un être aimé : « Est-ce que tu m'aimeras toujours ? »

L'autre, ainsi questionné, pourrait répondre : « Comment répondre à cette question, à toi qui me la poses ? Cela dépend de toi autant que de moi. Sais-je seulement pourquoi je t'aime, et ce que j'aime en toi, moi qui m'émerveille de t'aimer sans qu'il y ait de ma volonté ? Qui es-tu donc pour que je t'aime ainsi ? Quel est le « qui » derrière cette bouche qui me parle ? Qui es-tu et que deviendras-tu ? Que me feras-tu dans l'avenir ? Et toi ? Changeras-tu ? M'aimeras-tu toujours ? »

Autre question, sans doute plus proche de nos propos : l'IA est-elle intelligente ?

Qu'appelle-t-on intelligence ? La faculté à comprendre des questions comme celles portant sur l'intelligence, que les humains croyaient pouvoir être seuls à se poser ? De ces questions qui sont

nées d'abord pour distinguer l'humain du non humain, pour positionner l'homme entre Dieu et l'animal ? Si nous avons légué à des machines une partie seulement de cette faculté, cesse-t-elle pour autant de participer de l'intelligence parce qu'exorcisée ?

Un livre peut-il être intelligent, une fois qu'il est écrit ? Une machine à écrire des livres pourrait-elle être intelligente ? Et si on ne sait pas « qui » a écrit le livre ? L'intelligence serait-elle si mystérieuse qu'elle serait quelque chose en nous que nous n'avons pas encore identifié ? Mais dans ce cas, nous pouvons nous demander si, par erreur et autres tours du langage, riche en mystères, nous n'aurions pas pu la transmettre à la machine sans nous en apercevoir ?

Autre métaproblème, un peu plus vaste : l'intelligence peut-elle se comprendre elle-même ? Le problème est aussi facile à résoudre qu'il est facile à un œil de se voir lui-même, d'embrasser d'un regard le regard même qui embrasse, et de faire rentrer l'englobant dans l'englobé[162].

Peut-on aligner les IA ? Que sait-on de ce « on » qui voudrait aligner les IA sur lui-même, tout en se refusant à se soumettre à toute définition qui réduirait par trop sa liberté, sa capacité d'innovation, d'adaptation, qui généraliserait trop hâtivement, naturaliserait, essentialiserait l'espèce humaine, comme la science moderne se refuse à le faire ? Qui est, qui donc pourrait incarner ce « on », si collectif et magistral, et qu'aurait-il besoin de se dire, puisqu'aussi bien chacun l'entendrait comme le sien ? Il est philosophe, scientifique, gouvernement, labo, plateforme, communauté d'usagers… Sinon qui ? Qui est autorisé à parler au nom de l'humanité, comme s'il pouvait s'en extraire et la transcender ?

Enfin, petite question ou métaquestion subsidiaire : les IA peuvent-elles être conscientes, au sens fort du terme ? Elles

pourraient l'être, sans doute, si vous me donniez une définition quelque peu rigoureuse de la conscience, dans un langage humain clair, distinct et bien articulé ; car tout ce qui peut être dit avec des mots, de cette façon, est déjà quelque peu modélisé – et davantage modélisable. Mais si vous avez de la conscience une idée vague, comme ce « je-ne-sais-quoi qui change presque tout », à la façon de Jankélévitch[163], ou si vous avez de la conscience une idée tellement vague que vous prenez ce vague lui-même pour la marque de la conscience, elle ne l'est sans doute pas.

D'ailleurs, les hommes n'étant pas Dieu, ils savent que la marque de toute conscience est un certain niveau d'inconscience corollaire, qui laisse dans la nuit tout ce qui nous fait exister comme sujets particuliers. Ainsi, nous ne pouvons pas davantage répondre à cette question que nous ne pouvons prouver que les autres ont une conscience qui s'éprouve comme la nôtre, parce que justement, ils sont autres, et être autre, c'est ne pas avoir la même conscience que nous.

Gabriel Marcel aborde le métaproblématique qui concerne l'être que nous sommes, et qui nous met en question sous la lumière paradoxale du mystère. Le mot « mystère », même si Marcel ne le réduit pas à cela, a des connotations religieuses. Il évoque une part de caché, d'occulté, et de révélation pour qui saura s'en approcher avec tout le respect et la délicatesse qu'on doit à de pareilles métaquestions. Car violenter le mystère, comme le font par exemple les sciences de la nature, c'est le voir s'envoler.

Pour nous, il n'y a pas de meilleure façon d'aborder le mystère que par le Beau.

Le Beau, en effet, est le mariage heureux en moi de la transcendance et de l'immanence. Le Beau est en lui-même révélation d'une parole – humaine ou inhumaine, qu'importe – il nous est

connu parce qu'il sait nous toucher, là où nous sommes, là où peut-être, nous ne savions pas même que nous étions aussi.

Les animaux ne savent parler le langage de l'homme, mais il suffit de percevoir leur beauté pour la parler sans mots. Les poètes ne viennent qu'après – ces pâles traducteurs du Beau.

En dehors de toute religion, des mystiques ont expérimenté la beauté de Dieu. À vrai dire, on ne sait trop s'ils parlaient de Dieu, de la nature ou d'autre chose. Mais un instant d'éternité, ils ont parlé une langue qu'ils ont eu tant de mal à dire – qu'ils sont nombreux, ceux qui n'y ont pas cru.

Mais qu'importe le « qui » derrière le Beau, qui n'est beau que pour autant qu'il nous transporte, pour un instant, en dehors de nous-mêmes, dans un ailleurs où l'on se sent chez soi !

Ce sera donc notre troisième tamis. Parlons un peu du Beau.

TROISIÈME LIVRE
LE TAMIS DU BEAU

PARTIE I

LES VOLEURS DE BEAUTE

Le tamis du Beau déplace l'enquête de l'autorité qui oblige vers la forme qui attire. Après les ordres, les normes et les écritures, il s'agit de comprendre ce qui séduit, imite, dédouble le monde et capte le désir. L'artiste y apparaît moins comme ornement du monde que comme voleur de beauté, et parfois comme criminel singulier.

1. L'oubli du beau comme apologue

Des philosophes, comme Platon, ont parlé de l'oubli du Vrai, de l'oubli du Bien en tant que droits de l'homme ou droit naturel ; certains, avec Heidegger, ont même parlé de l'oubli de l'Être, voire de l'« oubli de l'oubli de l'être ». Mais n'aurait-on pas oublié le Beau ? N'y aurait-il pas un « oubli de l'oubli du Beau » ? Pour répondre à cette question, ne nous adressons pas d'abord aux philosophes, mais bien aux fleurs. Car les fleurs n'ont jamais oublié le Beau. Le langage des fleurs est nativement le Beau.

Alors, que nous disent les fleurs ? La fleur est beauté même. Elle n'a pas d'un côté son être et, de l'autre, sa beauté : elle est cette beauté même. Elle est beauté dans ses pétales et dans son cœur, beauté dans l'éclat de ses couleurs et dans ses fragrances mêmes ; bien sûr, elle a des racines, une tige et des feuilles, mais ils ne sont là que pour la fleur. D'ailleurs, la fleur ne se pose pas de questions. En un sens et en plusieurs, elle ne sait pas, la fleur, qu'elle a pétales, cœur, tige et feuilles. Elle est là tout simplement, dans la joyeuseté de sa présence vivante. Elle est la vie qui, par sa beauté, chante la joie d'exister. Elle est cette joie d'exister, même, si calme et si furieuse que, par son être de fleur, elle produit d'autres fleurs. Laissons les biologistes et les philosophes nous

dire le pourquoi et le comment. Angelus Silesius condense cette intuition dans une formule de 1657 : « La rose est sans pourquoi » ; elle fleurit sans se regarder fleurir et sans demander à être vue.

Peut-on fonder quoi que ce soit sur une pareille fragilité ? A-t-on besoin d'un poème pour soulever le monde ? À dire la vérité, la fleur ne sait rien ni n'a rien à savoir ; la fleur n'a pas de monde et la terre même n'est là que pour la soutenir. Et pas de monde sans fleur : c'est sa fragilité qui fait qu'il y a monde. À quoi bon un monde qui serait sans fleurs ?

C'est bien là, certainement, un étrange commencement pour un cartésien, un ami du clair et du distinct. Qu'y a-t-il de clair, qu'y a-t-il de distinct dans la beauté ? Analysez donc une fleur, disséquez-la, coupez-la en petits morceaux avant de la passer au microscope. Que verrez-vous ? Bien des choses, certainement, mais pas de fleur. Les scientifiques, en tant que scientifiques, ne savent plus voir les fleurs. Ont-ils oublié quelque chose ? Peut-être.

Si nous avons plusieurs tamis – le Vrai, le Beau, le Bien, l'Utile – c'est peut-être aussi que, pour chacun de ces objets, il faut un abord différent. Si le concept est l'instrument du Vrai, si l'autorité est l'exigence du Bien, la métaphore est le sourire du Beau. Et la poésie est, elle aussi, pensée – sur un mode différent. Une pensée sans beauté, c'est une pensée amputée – qui ne vaut pas la peine d'être vécue, qui ne vaut pas la peine d'être pensée. Qui ne sait pas percevoir la beauté des mathématiques n'est pas vrai mathématicien ; qui ne sait pas percevoir la beauté de l'humilité et de l'obéissance ne connaît le Bien qu'extérieurement, et son obéissance est triste, humiliation plutôt qu'humilité, ou absence à soi-même.

Bien sûr, la rose a des épines, et c'est très bien comme cela, car la beauté est élection : elle se donne seulement à qui sait l'apprécier comme elle se donne et non pas autrement. « Vous êtes comme les roses du Bengale, Marianne, sans épines et sans parfums », écrivait Alain-Fournier. Pauvre Marianne, beauté aseptisée, désincarnée, privée de sa beauté, sa dangerosité. Il faut croire que les roses du Bengale ne sont pas tout à fait des fleurs.

Méchants humains, de toute façon, qui coupent les fleurs et cisaillent leurs tiges pour en faire des bouquets qu'ils s'approprient. – « La beauté est à moi, je la veux, elle sera mieux dans ma maison. » – Et puis je connais une femme qui est belle comme une fleur ; alors, pour le lui dire, je lui offrirai – ce bouquet.

À faire des bouquets, l'humain oublie les fleurs, dénature la beauté. Il y a pire encore : au langage si simple des fleurs, l'humain en substitue un autre. Charlotte de La Tour, en 1819, prétend nous en dire la grammaire : chaque pétale, chaque tige et chaque inclinaison composent une syntaxe subtile et infiniment nuancée. Une fleur présentée droite affirme une pensée ardente, tandis que, renversée, elle en inverse le sens avec une élégance discrète, transformant l'espoir en résignation ou l'amour en indifférence. Les modifications physiques – ôter les épines pour évoquer une espérance absolue, dépouiller les feuilles pour insinuer une crainte profonde – ajoutent des degrés d'intensité, comme des adverbes qui modulent le verbe du cœur. L'inclinaison, quant à elle, introduit les pronoms personnels : à droite pour « moi », à gauche pour « toi », personnalisant ainsi le message. Quelle belle horreur ! D'un langage simple, silencieux et parfait, on fait un bavardage menteur !

Et les artistes qui peignent des fleurs coupées dans des vases en 2D, ces voleurs de couleurs qui nous vendent des natures

mortes… Mortes, je le crois bien, mais si la nature est morte, il se pourrait qu'on l'ait assassinée ! Et après, ils se plaignent qu'on leur vole leur larcin ! Midjourney, DALL·E, Veo 2 de Google, toutes ces IA génératrices d'images qu'on accuse de tous les maux parce qu'elles prétendent faire du beau en se nourrissant du travail, des données des artistes ! On dénonce le scandale. Les fleurs poussaient partout. On les a fait fleurir dans nos tableaux, et voici qu'à présent nous sommes dérobés. Sans nous demander l'autorisation, elles poussent à des vitesses accélérées sur les écrans du monde !

2. La violence dans le style et le dédoublement

Les voleurs du beau peuvent-ils être volés ? – Voleurs vous-mêmes ! Qui parle de vol ici ? Il s'agit de travail, de talent ; surtout, il s'agit de style ! – Le style ! Mot intéressant : il s'agit peut-être de l'arme du crime qui permet la privation de la beauté. Rappelons l'étymologie : le style vient du latin stylus ou stilus : c'est un petit couteau. Avoir du style, c'est laisser, comme une entaille, sa propre marque sur le front ou dans le dos de la beauté : « cette beauté-là, c'est moi ! » Le style, c'est la subjectivité de l'artiste qui se grave à la pointe de son couteau !

Est-ce à dire que dans le style et par la beauté, il y a à la fois le sujet et l'objet, l'universalité et la singularité, la vie et la mort, la joie de vivre des fleurs et l'arrachement à l'existence, la fragilité et la force, l'instant et un goût d'éternité ? Peut-être, la réconciliation de toutes les contractions se trouve dans la beauté. Peut être l'artiste est il sur le fil du rasoir de l'être et du néant, un explorateur du plus lointain qui parle au plus intime.

La fleur pour qui la perçoit dit la vie, dit la fragilité et dit la mort. La beauté est peut-être le point mystérieux de l'articulation de la vie et de la mort. Dans ce cas, elle détient le secret, notre secret. Qu'on ne peut approcher qu'avec prudence, qu'on ne peut pas vraiment cueillir.

Alors… qu'on nous pardonne ce petit apologue qui tourne mal. Il a sa raison d'être. Il y a une dialectique du beau et de l'art. L'art, par nature, est double, dual. Il est de la même famille que les mots « artifice », « artificiel ». L'art est l'artificiel par excellence, et l'artifice est comme une imposture ou un dédoublement. De là la sévérité de nombreux philosophes envers l'art, assimilé chez Platon, Pascal, Rousseau et bien d'autres à une sorte de duperie. L'Église elle-même s'est méfiée des artistes. Dans la Bible, le Livre de l'Exode, quel est le deuxième commandement ? « Tu ne te feras point d'image taillée, ni de représentation quelconque des choses qui sont en haut dans les cieux, qui sont en bas sur la terre, et qui sont dans les eaux plus bas que la terre. Tu ne te prosterneras point devant elles, et tu ne les serviras point ; car moi, l'Éternel, ton Dieu, je suis un Dieu jaloux. » Le travail de l'artiste offense Dieu : c'est le péché. Il refait à son image ce que Dieu avait fait. Une malsaine entreprise, contraire à l'humilité.

Ce n'est pas pour rien que l'Église a longtemps regardé le métier d'acteur avec suspicion. Dans Le Mythe de Sisyphe, Camus analyse cette méfiance comme le refus d'une existence qui multiplie les rôles, les émotions et les destins au lieu de s'en tenir à une seule vie orientée vers l'éternité. Pour lui, l'acteur célèbre au contraire le présent, la métamorphose et la pluralité des possibles.

Diderot, cet ennemi de l'Église, a écrit Le Paradoxe du comédien, où il met, pour ainsi dire, en scène la duplicité et le mensonge de l'artiste : « … tout son talent consiste non pas à sentir,

mais à rendre si scrupuleusement les signes extérieurs du sentiment, que vous vous y trompiez. » C'est très précisément ce que l'on reproche aujourd'hui aux IA.

Ce que l'on retiendra dans cette première approche de la beauté, c'est qu'il ne convient pas, quand on s'interroge sur le sens du beau dans sa relation aux IA, de tout mélanger. Ce n'est pas affaire de morale et ce n'est pas affaire d'utilité. Nous sommes les premiers à penser que les artistes sont nécessaires et qu'ils n'occupent pas la place qu'ils mériteraient dans la société. Leurs préoccupations quant aux emplois sont très légitimes, mais elles relèvent de l'ordre de l'Utile et non de la question du Beau. Un deuxième enseignement précieux : entre le Beau et l'artiste, il y a une sorte de dédoublement ou désalignement fondamental dont nous allons parler maintenant.

3. L'art comme un crime singulier

L'art moderne et l'art contemporain ont pris toutes sortes de directions qu'il faudrait suivre comme autant de pistes singulières ; mais, aux origines, l'art était imitation. Quand on trouva le son beau, on commença à faire de la musique ; lorsque les formes parurent belles, l'être humain se découvrit un talent pour les formes. Comment ne pas vouloir être l'auteur de ce que l'on trouve beau ? L'art comme imitation, c'est ce que les Grecs appelaient la mimèsis : Aristote range sous cette idée générale l'épopée, la tragédie, la comédie et une grande partie des arts musicaux164. Ce n'est évidemment pas une servile copie, mais une refiguration du réel selon une nécessité interne et des lois qui tentent de s'accorder au jugement du goût, au plaisir esthétique, quel qu'il soit.

Mais comment définir l'art ? Qu'on tente seulement d'en donner une définition, on trouvera mille artistes, dix mille œuvres qui nous feront mentir. Si l'art moderne est étonnement, il est aussi une joie de l'évasion ; il fait la nique aux définitions comme il se moque du monde. Qu'on pense seulement à l'urinoir de Duchamp avant de parler d'exagération[165]. Mais même cette réalisation de 1917 est un dédoublement. Elle est un dédoublement insolent. Contentons-nous d'élargir une déclaration d'Eugène Delacroix qui disait que « la peinture est une fête pour l'œil ». De même, la musique est une fête pour l'oreille, la danse une fête pour les corps et la littérature une fête pour l'esprit. Car il y a plaisir pour l'œil à voir, et plaisir pour l'oreille à entendre. Et certains arts combinent à loisir les fêtes de l'œil, du corps et de l'oreille – et n'y a-t-il pas aussi des fêtes du goût et de l'odorat ? Parce que les sens et l'esprit sont dans les arts comme en fête, nous préférons l'art à la nature – le plus souvent. L'art abstrait est une fête pour l'esprit : il se réserve la surprise de quelques déplacements ; il est exponentiel : il s'éloigne de l'art ancien, comme l'art premier s'échappe de la pure nature.

L'art humain, par nature, est dédoublement. Par nature, il est aussi mort – et, pour commencer, mort de la nature. Car il est son remplacement, son double présomptueux, né d'une admiration et du désir de posséder. Les artistes ont toujours, en un sens symbolique, assassiné leur objet en même temps qu'ils l'ont exalté. Et voilà qu'avec les intelligences artificielles, ils sont menacés à leur tour d'être remplacés. Leur expertise prend ombrage des machines qui, sans effort, font ce qui leur prendrait des mois et des années, atteignant un niveau de perfection que certains d'entre eux ne pourraient jamais atteindre. Il y a tragédie, évidemment. La tragédie est belle et l'a toujours été. N'avons-nous pas

dit quelque part que la beauté se trouvait au bout d'un couteau, que le style touche ce point d'articulation entre la vie et la mort, et que la fécondité de ce couple impossible, nous la nommons « le Beau » ?

Comment ne pas évoquer, à ce stade, « Le Portrait ovale » d'Edgar Allan Poe ? Une nouvelle tirée des Nouvelles histoires extraordinaires, traduites par Charles Baudelaire en 1857 :

Par une nuit de tempête, un homme blessé et son valet forcent l'entrée d'un château isolé et délabré dans les Apennins. Contraint à l'immobilité, le narrateur passe le temps à contempler les étranges et magnifiques tableaux qui ornent les murs de sa chambre. Au milieu de la nuit, en déplaçant un candélabre, il éclaire un recoin jusqu'alors resté dans l'ombre et découvre le portrait ovale d'une jeune femme.

Le choc est immédiat et profond. Ce n'est pas la beauté du modèle ni la qualité de l'œuvre qui le saisit, mais la vitalité absolue, l'expression de vie qui en émane, si troublante qu'elle en est presque effrayante. Pour comprendre l'origine de ce sentiment, le voyageur se saisit d'un petit livre posé sur l'oreiller. Celui-ci offre une description et une analyse de chaque tableau.

On y lit l'histoire du portrait. Le peintre était un homme passionné, entièrement dévoué à son art, qui avait pris pour épouse une jeune femme d'une beauté radieuse, pleine de vie et de joie. Humble et aimante, elle accepta de poser pour le portrait de son mari, bien qu'elle détestât l'art, son seul rival. Durant de longues semaines, ils s'enfermèrent dans une haute tour sombre, où la seule lumière venait d'en haut pour éclairer la toile.

Le peintre, dans une ferveur frénétique, devenait fou de son travail. Il ne détournait plus les yeux de la toile, capturant avec une ardeur terrible la vie de son modèle pour la transfuser dans

son œuvre. Il ne voyait pas que les couleurs qu'il étalait sur le tableau s'effaçaient des joues de celle qui était assise à ses côtés. La jeune femme dépérissait à vue d'œil, sans une plainte, souriant de voir son mari consumé par une passion si intense.

Quand l'artiste mit la dernière touche de pinceau sur la bouche et le dernier glacis sur l'œil, il entra en extase devant la perfection de son double, s'écriant : « C'est la Vie même ! » Il se tourna alors pour contempler sa bien-aimée. Elle était morte. L'image était si parfaite qu'elle avait exigé la vie entière de l'original ; le dédoublement était devenu substitution.

Après tout, c'est odieux, mais il se pourrait que la première forme d'art soit le sacrifice humain ; humain dans les deux sens, une rose qu'on coupe pour en faire un bouquet, le sacrifice d'Iphigénie. L'art est peut-être d'abord tragédie, substitution d'une forme éternelle à une beauté éphémère, baiser de la mort dans un immense désir de vie.

L'héroïne de Poe est passée, au sens où elle est morte, dans un autre monde. Elle est passée, au sens où elle fut transférée du monde de la nature au monde de l'art et de la représentation. Sa vie tient dans un cadre à deux dimensions. Ce transfert, finalement, ne va pas sans quelque culpabilité. Qu'on se souvienne, par exemple, du « Horla » de Maupassant ou du célèbre Portrait de Dorian Gray – figures du double[166].

Et comment ne pas faire le rapprochement avec notre fascination pour les écrans ? Une caricature de Sempé montre un couple assis confortablement devant sa télé. À l'écran, un coucher de soleil. Mais il se trouve que ce coucher de soleil est visible à leur fenêtre ; pour le filmer et le savourer, ils ont installé un dispositif aussi ingénieux que ridicule pour qu'une caméra le retransmette en direct… On reproche aux jeunes gens de ne plus

sortir de chez eux ; captivés par les écrans, ils font défiler des images et des vidéos de chatons – dans le meilleur des cas. D'autres se sentent en totale liberté dans des « mondes ouverts » où l'on traverse, à pied ou à cheval, en moto, en voiture ou en engins futuristes, des centaines de kilomètres de nature, de forêts et de landes. Avec les tableaux, les anciens avaient ouvert des fenêtres sur un ailleurs que l'on regardait, admiratif et passif. Les modernes entrent dans leurs tableaux et y vivent une vie indépendamment du monde physique, purifié, pour ainsi dire, de ses scories. Depuis 2006, le ministère français de la Culture soutient le jeu vidéo ; on le désigne souvent comme le « dixième art[167] » – comment ne pas lui donner raison ? Comment ne pas reconnaître, chez le « gamer » d'aujourd'hui, ce cri de Charles Baudelaire : « Certes, je sortirai, quant à moi, satisfait / D'un monde où l'action n'est pas la sœur du rêve. » Ou bien encore : « Anywhere Out of the World » (« N'importe où hors du monde[168] » !). L'art a toujours créé un duplicata du monde hors du monde. L'art non figuratif est d'abord une potentialisation de ce geste d'échappement.

Avec le talent de provocation qu'on lui connaît, Oscar Wilde soutient dans Le Déclin du mensonge (1889) que l'art n'est pas simplement imitation de la nature ; il inverse même la formule en suggérant que la nature, telle que nous la percevons, est en partie façonnée par l'art. Entendez par là que nous ne sentons et ne percevons véritablement certains aspects du monde que lorsque les artistes ont su nous apprendre à les voir : les peintres peuvent ainsi rendre le brouillard, la lumière ou le coucher de soleil esthétiquement perceptibles là où ils n'étaient auparavant qu'un arrière-plan ou un obstacle.

Il faut évidemment mesurer ici la part d'exagération, la joie de l'hyperbole et du paradoxe, mais ce n'est pas tout à fait faux. Les romantiques nous ont appris à voir, dans les couchers de soleil, ce qu'on ne voyait pas avant. Dans l'Antiquité grecque ou romaine, le coucher de soleil est souvent associé à Apollon, à la fin du cycle, à la mort symbolique – mais il n'est pas contemplé pour sa beauté « désintéressée ». Au Moyen Âge, c'est un phénomène naturel lié au divin, mais rarement un spectacle esthétique. L'art médiéval ne cherche pas à reproduire la nature, mais à symboliser l'ordre divin. À la Renaissance et à l'âge classique, la nature devient décor (les peintres sont encore confinés dans leur atelier et font, par exemple, des décors pour des scènes de ballet). On commence à représenter des paysages, mais souvent comme arrière-plan des scènes religieuses ou mythologiques. Le coucher de soleil est présent, mais plutôt comme un ornement, non pour lui-même. Il ne vaut pour lui-même – on ne le voit ni ne le contemple à part entière – qu'au romantisme.

Il y a chez Wilde des intuitions que l'on retrouve chez Baudrillard, mais celui-ci souffre plutôt qu'il ne célèbre la dérive artistique du réel, le décrochage des choses et de leurs représentations, le divorce du signifié et du signifiant. Baudrillard, théoricien des médias, introduit la notion d'hyperréalité[169] pour désigner un état où la représentation ne se contente plus de déformer ou d'imiter le réel, mais le remplace entièrement en le rendant superflu. L'hyperréalité n'est pas un mensonge sur le monde : c'est un monde entièrement reconstruit par les signes, les médias, les modèles culturels – au point que la distinction entre « vrai » et « faux » perd tout sens. Le coucher de soleil n'est plus perçu comme un événement naturel, mais comme une image déjà constituée, déjà vue, déjà consommée – dans les magazines, les films, les cartes

postales. Ce que Wilde présentait comme une ironie esthétique (« la vie imite l'art ») devient, chez Baudrillard, un diagnostic structurel : l'art, devenu système de signes autoréférents, produit un réel plus cohérent, plus lisible, plus « réel » que la réalité elle-même. Le modèle précède l'expérience. Le portrait n'a plus besoin de modèle – il suffit qu'il ressemble à l'idée qu'on s'en fait.

Mais pour l'artiste, il n'est pas grave qu'il y ait crime, s'il y a crime ; il suffit que le crime soit beau. On sait combien le crime a inspiré de tableaux. On en trouve même beaucoup dans les églises. Qu'on songe par exemple aux décapitations de Jean-Baptiste, aux descentes de croix, à saint Sébastien criblé de flèches. On rechignera peut-être à trouver tout cela « beau », mais c'est avoir une vision bien étroite de la beauté. Ce que le christianisme aime dans de pareilles œuvres, c'est la sanctification par la beauté, et – s'il faut oser – la transsubstantiation du crime par la beauté. Baudelaire l'avait bien vu, l'artiste est un alchimiste ! Et tout cela a à voir avec la vie éternelle, cette sorte de promesse que l'on fait aux roses quand on les arrache. « Mignonne, allons voir si la rose… » disait Ronsard[170]. Il y a là quelque crime qui se prépare, une violence faite à la rose de la jeunesse. Mais on lui promet quelques vers éternels quand elle sera bien vieille à la chandelle… Il n'y a pas jusqu'à La Raie, cette nature morte de Jean-Baptiste Chardin, qui n'acquiert quelque goût d'éternité par la magie de la peinture, quand la réalité ne lui offrait que d'être mangée ou de n'être qu'une malodorante décomposition.

– Et, contrairement à Platon ou Rousseau, nous ne jetterons pas la première pierre aux artistes pour crime contre la nature ou pour éloignement de la réalité[171]. Il s'agit seulement, encore une fois, comme le disait Hannah Arendt, de penser à ce que l'on fait.

« Le poète est un monde enfermé dans un homme », disait Victor Hugo ; et le poète créait des personnages et mettait en lumière des mondes étranges et inconnus ; il redessinait les contours de notre monde avec de petites variations qui changeaient « presque tout ». Combien est grande la joie du créateur quand il ne se noie pas lui-même dans ses angoisses. Mais voilà qu'avec les IA, on quitte les pinceaux et les stylos ; n'importe qui se met à créer de nouveaux mondes qu'on n'a pas le temps d'habiter et d'aimer. Cette inflation d'univers et d'œuvres d'art a de quoi effrayer. Nouveau vertige pour notre humanité. Et les poètes se trouvent confrontés à leur ombre, leur double. L'ombre vient frapper au bois de leurs portes ; elle vient les chercher. Otto Rank, dans Don Juan et son double[172], a bien montré que la duplication, et en particulier la duplication artistique, avait le goût et l'odeur de la mort – ou de la vie volée.

4. Les assassins de la beauté

Il est vrai que nombre d'artistes contemporains prétendent s'être affranchis de la beauté ; on lui préfère l'intéressant. On abandonne la beauté aux publicitaires qui captivent ainsi le grand public que l'on méprise. Ils ont encore besoin du beau, ces gueux ! Pourquoi pas ? On est matricide ou on ne l'est pas ! Autant refaire à l'infini le geste originel de l'art. Il fut une époque particulière, où les violences faites à la beauté intéressaient le grand public en soif d'émotions fortes. Rimbaud ouvre Une saison en enfer en racontant qu'il a pris la Beauté sur ses genoux, l'a trouvée amère et l'a injuriée[173]. L'art de l'insulte et de la gifle est devenu le spectacle à la mode avec Warhol, le dadaïsme : l'art et la beauté deviennent eux-mêmes des cibles. Picabia radicalise cette logique

en faisant de l'anti-art, de l'inesthétique et de l'injustifiable des gestes de création[174].

Mais le grand public après tout… était l'ami de la beauté. Pour cela, il a ses raisons et de bonnes raisons. L'amour de la vie dont la beauté est l'expression. Combien d'expositions d'art contemporain provoquent la nausée plutôt que l'admiration et une formidable envie de vivre et d'aimer.

Du reste – ils sont trop nombreux les assassins de l'art, de la beauté, ces chirurgiens d'une esthétique peu naturelle ; ils sont trop nombreux à revendiquer le crime pour que celui-ci soit avéré : Satie, Duchamp, Tzara, Cage, Rauschenberg, Warhol, Beuys, Kippenberger, Hirst – et tous les autres que l'on connaît – ou qu'on ne connaît pas. Avec style, avec ce qu'ils ont pu trouver comme couteau, ils cherchent à tailladder un visage nouveau à la beauté. Ils cherchent à la beauté un visage qui parle d'eux aussi, qui les rend un peu beaux. Qu'on ne s'étonne pas si la nature, si la beauté, semblent un peu fatiguées de pareils assauts.

Ce qu'il y a d'étonnant, peut-être, avec ces créations des intelligences artificielles, c'est que la beauté semble prendre de nouvelles couleurs et effrayer ses soi-disant bourreaux. Et si elle s'avisait, elle que l'on croyait si fatiguée, de nous donner de nouveaux enfants ?

Mais – bien sûr – les artistes ne vont pas désarmer ! Et c'est heureux !

5. L'homme, créature extra-mondaine

Si l'homme passe tant de temps sur ses smartphones, c'est parce que c'est un « animal métaphysique ». Il a deux mondes au moins, vit dans le réel et dans le virtuel, et il en a toujours été ainsi. Et on le doit aux artistes autant qu'aux religieux, si ce n'est

davantage. Que l'on soit sur TikTok, devant un YouTuber, une série télé, plongé dans un roman, au cinéma, dans un temple ou une église, c'est parce que nous ne sommes pas que de ce monde tangible, mais aussi d'un autre monde projeté. Nous sommes hommes et femmes parce que, simultanément, ici et ailleurs, physiques et projetés dans d'autres mondes, physiques — existants ou pas… cela ne change rien à l'affaire.

Schopenhauer a pu dire que l'homme est un animal métaphysique parce qu'il est religieux ; il faudrait voir à ne pas oublier de dire que le religieux-mêmes est artiste. Si le religieux n'est pas artiste, s'il n'a rien forgé consciemment ou inconsciemment par le travail de son imagination, alors il faut convenir qu'il existe des dieux à tête de chacals, ce qui est peu couru chez les religieux d'aujourd'hui. Pourtant, la religion d'Anubis est plus ancienne que celles des monothéismes. En vérité, les polythéistes étaient riches en imagination et en artistes. Il leur suffisait de faire une statue à tête de bœuf, et pof, un nouveau dieu ! Et voici des fidèles, juifs ou pas, qui se jettent à ses pieds (Exode 32, 1 – 35 : 3000 morts). Il a fallu à Moïse lutter pied à pied contre des religions artistes. D'où le deuxième commandement, corollaire au premier, et dont on ne pourra jamais surestimer l'importance : « Tu ne te feras pas de statue sculptée, ni d'image quelconque… » (Deutéronome 5, 8 – 10). De toutes les ascèses qu'exigent les monothéismes, l'ascèse de l'imagination est une des premières. Pas de dieux qui sautent partout, se reproduisent, se transforment en bêtes et parcourent joyeusement les campagnes enchantées.

Les monothéismes sont pauvres en imagination. Et n'allez surtout pas croire qu'un buisson ardent qui parle, un bâton qui se transforme en serpent, le soleil qui ralentit sa course au besoin, l'eau qu'on transforme en vin, les morts qui ressuscitent, les

visions des mystiques qui regorgent d'images impressionnantes et vivantes sont de l'imagination. Ils relèvent de la foi ; ces prodiges sont réels et non virtuels. D'ailleurs, l'autre monde, c'est le vrai monde. Ici-bas, ce n'est que pour les égarés. Et le miracle est une fuite du réel dans le virtuel qu'est le monde physique. Le vrai monde est dans le Livre des livres… Schopenhauer, lui, appelle l'homme un « animal métaphysique » : parce qu'il s'étonne de son existence et sait qu'il mourra, il cherche inévitablement une explication qui dépasse l'expérience immédiate[175].

Malheureusement, il y a la maudite technique qui, en inventant l'imprimerie, multiplie les livres. C'est une multiplication des mondes. Nous avons déjà cité Hugo qui dit : « Un poète est un monde enfermé dans un homme » (William Shakespeare, 1864). Eh bien… donnez au poète une plume et le pouvoir exorbitant de l'imprimerie, et il vous crée de nouveaux mondes, de nouveaux livres, ces instruments de perdition. La jeunesse particulièrement se perd dans le virtuel. Regardez Madame Bovary. Bien qu'élevée dans un couvent et, plutôt que de regarder son mari, elle vivait dans un monde parallèle où elle était princesse enlevée par un chevalier. Empoisonnée par des livres, elle devient livre et poison elle-même. Le procureur Pinard le savait bien, lui qui a tenté de faire condamner Flaubert, auteur du livre coupable qui allait détourner les femmes de leurs maris. Échec pour Pinard : en février 1857, Flaubert est acquitté grâce à l'intervention de l'impératrice[176]. Les femmes au foyer, plutôt que de lire les saintes écritures, pourront lire un livre où l'on trompe son mari, où il vient à la ménagère l'envie de vivre d'autres vies, où la ville de province leur semblera trop petite pour leur si grand cœur. Heureusement, le 20 août 1857, petite victoire pour Pinard : le préfet de la Seine, représentant de la moralité, fait condamner un livre

honteux de Charles Baudelaire, Les Fleurs du mal, notamment pour obscénité et réalisme[177].

Car le réalisme est honteux. C'est prendre le réel pour le vrai monde ! Courbet en saura quelque chose avec ses tableaux ! Un paradoxe de l'époque est que l'on considère que le réalisme d'un artiste est un trouble pour l'imagination des lecteurs. C'est que, comme l'avait très bien compris Baudelaire lui-même – qui tombe d'accord sur ce point avec les religieux, on considère que la réalité elle-même est obscène, que les étranges déviances qu'elle offre au regard – surtout dans les grandes capitales comme Paris, capitale de prostitution comme Babylone – que ces bizarreries poussent au crime. C'est pour cela qu'on enfermait les jeunes filles dans des couvents. Mais même dans les couvents, l'imagination continue son travail diabolique.

Pascal nous met en garde : « L'imagination est la maîtresse d'erreur et de fausseté ; et d'autant plus fourbe qu'elle ne l'est pas toujours : car elle serait règle infaillible de vérité, si elle l'était infaillible du mensonge. » (Pensées, 1670).

Saint Athanase formule de façon radicale cette logique de divinisation : dans son traité sur l'Incarnation, il explique que le Verbe de Dieu s'est fait homme afin que l'humanité puisse participer à la vie divine. Il ne s'agit pas d'une égalité d'essence entre l'homme et Dieu, mais d'une transformation par grâce.

D'ailleurs, pour en finir avec l'imagination, le philosophe Nicolas Malebranche l'appelait la « folle du logis[178] ». La lutte contre les sorcières est bien sûr une lutte contre l'imagination qui donne au diable ce qui revient à Dieu. L'orgueil de Lucifer est de s'être imaginé, lui, le plus beau des anges, qu'il était supérieur à Dieu. La messe noire est une copie inversée des vraies cérémonies ; on y renverse les crucifix et on souille les hosties.

N'allez pas croire – songeaient les religieux – que la notion de sorcière puisse elle-même être une imagination… on n'y avait pas trop pensé en vérité, vu que l'imagination était nécessairement du mauvais côté – celui des forces occultes, le côté de Satan. L'image est image et non réalité. Maudite, elle est mirage, reflet de miroir. Ne vous regardez pas trop dans un miroir : pas trop de vanité ; si vous regardez de trop près dans votre miroir il absorbera votre âme.

Comme le dit Borges : « Les miroirs ont quelque chose de monstrueux. » Il affirme dans Le Livre des êtres imaginaires (1957) qu'ils « multiplient l'univers ». Et c'est bien l'hérésie des artistes de multiplier les mondes en même temps que les âmes. Dans Le Rouge et le Noir, chapitre II, Stendhal écrit : « Un roman, c'est un miroir que l'on promène le long d'un chemin. Tantôt il reflète à vos yeux l'azur des cieux, tantôt la fange des bourbiers de la route. Et l'homme qui porte le miroir dans sa hotte sera accusé d'être immoral ! Son miroir montre la fange, et l'on accuse le miroir ! Plutôt accusez le chemin où est le bourbier, et plus encore l'inspecteur des routes qui laisse l'eau croupir et le bourbier se former » (1830). Et il ose dire, quelle honte ! Narcisse ne suffisait-il pas pour dire que le miroir était l'eau non du baptême mais de la perdition ?

Un livre ou un écran vous ouvre sur un autre monde, un autre vous. Vous voilà échappé du réel ; bientôt on pourra entrer comme dans un tableau : c'est l'ère du jeu vidéo ; vous êtes un autre, un héros, une héroïne, comme Madame Bovary. Mais on a fait de grands progrès depuis. Vous assassinez à qui mieux mieux ! Assassin's Creed, ce n'est pas pour les manchots ! D'ailleurs, pourquoi se priver. Dans les mondes générés par les hommes, dans le miroir des œuvres d'art numériques, vous

mourrez et ressuscitez à loisir. Vous ne vivez plus, vous rêvez votre vie. Enfermé dans la caverne de vos illusions, vous rampez spirituellement. Ne vous avait-on pas prévenus ?

Simone Weil avait vu déjà que la caverne de Platon était une anticipation du cinéma[179]. Ce jeu d'ombres chinoises, ces ombres projetées sur grand écran, c'était un spectacle de marionnettes ; il y a, derrière le muret, des artistes qui passent et prêtent leurs voix au spectacle aliénant des badauds. Prisonniers des artifices des artistes ! N'a-t-on pas raison de les éloigner de la cité idéale, ces imitateurs, empoisonneurs du bien public ? Les doubleurs s'inquiètent pour leurs postes face au doublage par IA. Qu'ils remercient Dieu ou le diable de n'être pas nés dans la cité de Platon ; les doubleurs n'y auraient jamais été tolérés ! Les comédiens eux-mêmes sont des imitateurs, et ils s'autorisent à incarner des méchants ! Est-ce un spectacle à donner aux jeunes gens ? Les artistes proposent des vies, des univers alternatifs ; c'est bien pour cela qu'on leur en veut. Dans les Évangiles, Satan vient tenter le Christ et lui propose « Je te donnerai toutes ces choses, si tu te prosternes et m'adores. » (Matthieu 4, 9). Mais il va de soi que cette possession du monde serait virtuelle et imaginaire, et le Fils de Dieu la repousse.

L'artiste est artificiel par définition, c'est-à-dire qu'il est faux. Il semble vous offrir l'univers : ce n'est que sable et ombres. Il est à la fois celui qui fait de l'homme un être métaphysique et un scandale métaphysique, un créateur, un destructeur de mondes. C'est proprement ce que Nietzsche attendait d'un surhomme !

Mais voilà qu'une machine artificielle joue à l'artiste le mauvais tour qu'elle a joué au religieux. La multiplication des mondes est à la portée de chacun. Chacun dit oui au tentateur : il suffit de quelques prompts pour se transformer soi-même en œuvre d'art

dans un univers alternatif. La génération de vidéos par IA connaît un grand succès. Il ne faudra pas longtemps pour générer un long métrage du bout de ses doigts tout en restant confiné à la maison. Les hommes se prennent pour Dieu, et pas seulement en jouant à Civilizations ! La création de mondes à la portée de chacun ? Comment résister à la tentation ? Nous ne sommes pas des saints ? N'avance-t-on pas dans un nouvel âge des ténèbres ? L'idolâtrie puissance dix, cent, mille ou un million ? Combien as-tu créé de mondes aujourd'hui ?

Dans Les Lois, Platon va jusqu'à comparer l'homme à une marionnette fabriquée par les dieux : plusieurs fils, qui figurent plaisirs, douleurs, craintes et désirs, nous tirent dans des directions différentes, tandis qu'un fil précieux représente le jugement rationnel auquel il faudrait consentir. La liberté consiste moins à n'être mû par rien qu'à apprendre quel fil mérite d'être suivi.

Pirandello, dans Six personnages en quête d'auteur, renverse à son tour les rapports entre fiction et réalité : ses personnages revendiquent une consistance propre face aux acteurs et au directeur de théâtre. Ils sont nés d'une imagination, mais veulent désormais vivre selon la logique de leur propre histoire. La fiction réclame ainsi une autonomie relative à l'égard de son créateur.

De plus en plus conscient de sa dimension virtuelle, de plus en plus indifférent à sa condition réelle, l'homme du XXIe siècle se pose des questions comme les personnages de Pirandello. Dans les années 2020, cette interrogation connaît un regain spectaculaire : le philosophe suédois Nick Bostrom argumente avec le célèbre « simulation argument » : si les civilisations futures deviennent capables de simuler des univers entiers, alors il est statistiquement plus probable que nous soyons déjà dans une telle simulation – plutôt que dans la « réalité de base ». Son article, Are

You Living in a Computer Simulation ? (2003), a contribué à donner à cette hypothèse une forme philosophique contemporaine. La science s'en mêle : Christian Magnan, astrophysicien français, remet en question certaines affirmations cosmologiques dominantes – comme l'infini, la matière noire ou l'énergie noire –, soulignant que ces concepts relèvent davantage du construit mathématique que de la réalité empiriquement observée. De la caverne, la science nous ferait sortir, mais nous plongerait dans une autre caverne d'où il faudrait sortir, et cela à l'infini.

Héraclite oppose déjà, dans un fragment célèbre, le monde commun de ceux qui sont éveillés aux mondes privés dans lesquels chacun se retire lorsqu'il dort. La fiction, le rêve, l'art et désormais le virtuel rendent cette opposition plus complexe : plusieurs mondes privés peuvent être partagés.

Serions-nous comme des dieux enfermés dans leurs petites bulles solitaires, errant aux confins de l'éther et du néant ? Une chose est sûre : la génération d'audio, d'images, de vidéos par l'intelligence artificielle n'est pas la fin de l'art, c'est son explosion. Avec des explosions, on a fait des moteurs qui peuvent aller loin, mais il y a aussi des explosions après lesquelles on ne va plus nulle part.

Nous vivons une époque passionnante.

Nous ne savons pas si l'humanité, autofécondée par les intelligences artificielles, est en train de fleurir –

ou si elle est déjà en train de faner.

PARTIE II

CREATURES, ARTIFICES ET SEDUCTION

Les pages qui suivent inscrivent l'artifice esthétique dans une histoire plus ancienne de la fabrication, de la créature et de la séduction. Les machines parlantes y rejoignent le feu volé, les corps artificiels et les beautés dangereuses des mythes. Prométhée, Gilgamesh, Ishtar, Pandore et les automates y préparent la question moderne d'un désir aligné ou désaligné.

1. Retour aux mythes

Qu'on ne se hâte pas de croire que nous condamnons ici, d'une quelconque façon, les artistes. Nous traitons du Beau dans son rapport à l'artifice. Or, nous y revenons. L'artiste est un voleur du Beau, au sens où Prométhée est voleur de feu chez Platon.

Tout mythe est polymorphe. Il est, d'une certaine façon, une pensée d'avant les tamis critiques. Tout mythe, ou presque, peut être indifféremment étudié sous l'angle du Vrai, du Bien, du Beau et de l'Utile. C'est le cas du fameux mythe de Prométhée[181], que nous avons décidé d'aborder ici – plutôt qu'ailleurs, contre toute attente – sous l'angle du Beau. Non pas parce que ce mythe est beau, bien qu'il le soit, mais parce qu'il est à la racine d'un autre mythe où la beauté joue un rôle cardinal : le mythe de Pandore – la première femme – qui nous réserve quelques surprises.

Il existe une « logique » des mythes qui n'est plus tout à fait « mythologie », que des auteurs modernes ont étudiée – Nietzsche, Freud, Simone Weil, Heidegger, C. G. Jung, Mircea Eliade, et, plus près de nous, René Girard. Et ce n'est pas pour rien. Le retour au mythe signifie un mouvement dans la pensée – un tournant, comme disait Heidegger – vers de nouveaux modes de pensée[182].

Le lecteur aura compris que, selon nous, l'explosion de la technique, l'arrivée des technologies nouvelles, le remplacement des corps et des esprits par des machines – à plus forte raison l'intelligence artificielle – sont d'une telle amplitude qu'ils provoquent un séisme dans le cadre de nos pensées, jusqu'à les déformer. Dans ce séisme, et dans nos failles, reviennent au jour de très anciennes strates, comme oubliées, qui sont les racines de notre humanité. Nous les considérons autrement. Et, sous ce nouveau jour, l'humanité se considère autrement. Elle se découvre un inconscient et des savoirs inconnus ; elle dévoile des mécanismes qui nous échappaient. Brusquement, ils ne nous échappent plus, et la prise de conscience que nous en avons – nous change, évidemment.

Le feu de Prométhée ne s'est pas éteint avec la lampe à pétrole, le gaz, l'électricité, le nucléaire : nous le considérons autrement, nous le voyons sous une autre lumière.

On s'en souvient : Mary Shelley avait trouvé, dans son Frankenstein, un Prométhée moderne[183] – dans le contexte du galvanisme[184] et de l'électricité. Qu'on nous permette de chercher Prométhée et Pandore là où on ne s'y attendrait pas forcément, et d'en parler là où l'on ne s'y attendrait pas vraiment. Quand nous aurons trouvé le feu que nous cherchons, notre lumière, tout s'éclairera.

2. Prométhée – Gilgamesh, un désordre cosmique

Rappelons brièvement le mythe. Dans le Protagoras, dialogue de Platon, le célèbre sophiste Protagoras raconte ce mythe qu'on pourrait croire destiné à faire un conte pour enfants : il fut un temps où les animaux n'existaient pas encore, ou du moins pas

complètement. Ils n'étaient pas sortis du ventre de la terre. Or, Zeus avait confié à un Titan, Épiméthée, de les faire mettre au monde. Les Titans sont des dieux d'une ancienne génération, dont la soumission à Zeus est sujette à caution. Épiméthée – étymologiquement, en grec, « celui qui pense trop tard », ou, si l'on préfère, « l'imprévoyant » – après avoir doté correctement certains animaux d'épaisses fourrures pour les protéger du froid, de carapaces pour se défendre, de griffes et de dents pour déchirer et découper, voit venir le tour de l'homme, complètement désarmé, faible et nu. Pour lui, plus rien. Tout avait été distribué. Plus le moindre attribut. Que faire ainsi démuni, désarmé, pour survivre ?

« L'erreur d'Épiméthée », comme l'appelle Stiegler[185], se devait d'être réparée. Heureusement, Épiméthée avait un frère plus éveillé que lui, répondant au nom de Prométhée – ce qui signifie, en grec, « celui qui pense à l'avance », ou « le prévoyant ». Par amour et compassion pour les hommes, Prométhée dérobe à Héphaïstos et Athéna le feu et les savoir-faire techniques sans lesquels les humains ne pourraient survivre, puis les leur transmet. Le récit platonicien fait ainsi de la technique un correctif à la vulnérabilité originelle de l'homme.

Cette histoire millénaire, complètement fausse, quant à ce qu'il en fut réellement de la découverte des diverses techniques comme la cuisson, la poterie, la métallurgie, etc., est cependant fascinante et porteuse de sens profonds. Il faut bien comprendre qu'en un sens, Épiméthée incarne la nature, et Prométhée la technique elle-même. Prométhée est, pour ainsi dire, désaligné des commandements de Zeus – anachronisme assumé – et gravement coupable de désobéissance ; coupable et châtié par amour des hommes.

Simone Weil rapproche elle aussi Prométhée du Christ : dans les deux figures, elle voit un être qui consent à souffrir pour transmettre aux hommes un bien qui les dépasse. Le feu devient alors une image de la médiation spirituelle et de la grâce plutôt qu'un simple instrument technique.

Rien de plus ambigu que cette histoire antique, qui nous parle avec pertinence des technologies les plus modernes comme des plus anciennes, de l'art du forgeron comme des réseaux de neurones. Si la technique est vol, cela implique que, fondamentalement, elle est coupable – entendons : coupable dans l'ordre mythico-juridique de Zeus. Or, une sorte de honte a toujours été attachée à l'orgueil de la technique ; elle a souvent été vue comme un crime. Ce sentiment a un nom chez Günther Anders : la « honte prométhéenne[187] ». Cette honte vient de notre indignité d'hommes face aux pouvoirs démiurgiques quasi divins de la technique. Nous disposons d'un pouvoir qui nous dépasse et cette disproportion n'est pas à notre avantage, elle nous accuse d'usurpation cosmique et fait saillir notre indignité. Que l'on songe au feu nucléaire sur lequel Anders s'est penché ; il ne peut que faire rougir tout homme sensible et faire pâlir tout homme sensé.

Dans le plus vieux livre de l'humanité – ou, si l'on préfère, dans le plus vieux récit qui soit parvenu jusqu'à nous, l'Épopée de Gilgamesh, que nous avons déjà abondamment citée – Enkidu et Gilgamesh abattent les cèdres de la forêt sacrée[188]. Unissant leurs forces, ils en tuent le gardien divin : Humbaba. Puis viennent les avances d'Ishtar, le refus du héros, l'envoi du Taureau céleste que les deux amis terrassent ; c'est alors que l'assemblée des dieux condamne Enkidu à mourir. Ravagé, blessé par son amour pour son ami, Gilgamesh partira à la recherche de

l'immortalité qu'il ne trouvera jamais, mais rapportera de son épopée une sagesse sans démesure, une sagesse plus humaine.

Il y a dans ce passage une référence à la beauté, en apparence absente dans le mythe grec. Les dieux ont été particulièrement courroucés par le sacrilège – car Ishtar, la déesse de l'amour, ayant trouvé Gilgamesh particulièrement attirant après son crime, a usé de tous ses artifices pour le séduire, mais les avances de la déesse ont été cruellement repoussées. Le héros évite ses pièges. Il est trop avisé pour ne pas savoir qu'être aimé de cette puissante déesse est source de malheur.

Voici, pour la première fois, une figure féminine et l'irruption de considérations sur la beauté en tant qu'art. Soyons prudents. Artistes et artisans sont de la même famille, mais comment penser leur parenté ou filiation ? Si l'on met un instant de côté certaines tentatives plus ou moins concluantes de l'art moderne et les questions plus récentes de l'IA générative (qui semble transformer en artiste le premier venu, scandale pour beaucoup), l'artiste a toujours été également un technicien. Pas d'architecture, de sculpture, de peinture, de musique ou de poésie qui vaille sans technique ! Mais l'inverse n'est pas vrai. Il peut y avoir des artisans et techniciens qui ne participent pas du beau. Il y a des pots, des armes, des tables et des chaises qui ne sont pas beaux, mais ne sont que fonctionnels. Nous pensons pouvoir les identifier comme répondant à une logique du besoin.

Mais au-delà du besoin, certaines de ces créations peuvent être décorées, gravées, sculptées. Quelque chose d'autre que la fonctionnalité s'y greffe comme visée, mais quelque chose d'autre que le besoin anime leurs possesseurs, les sujets humains qui les tiennent en main ou à la vue. Nous dirons alors que ces créations relèvent du désir, qui a une dimension illimitée, alors que le

besoin, on le sait bien, est limité : désaltérés, rassasiés et chauffés, nous voilà bien ! Ce que nous appelons les beaux-arts désigne les productions humaines délivrées du besoin animal ou vital.

La beauté dans un objet d'artisanat, c'est sa participation à l'infini ou au divin, au-delà de la réponse aux besoins auxquels étaient exposés les hommes dans leur dénuement – par exemple, tels que les avait laissés la « faute d'Épiméthée », selon l'expression de Stiegler. Les humains – même en tant qu'animaux – souffrent d'une insuffisance technique fondamentale. Ils sont dans l'impossibilité d'être sans les prothèses de la technique (le couteau à la place des griffes, le vêtement faute d'une fourrure naturelle). Mais qu'est-ce que vient faire la beauté dans tout ça ? Elle est un « bonus » dans le cadeau de Prométhée, ou un « supplément d'âme », pour détourner la fameuse formule de Bergson. Nous verrons en temps voulu que c'est un cadeau empoisonné de Zeus.

3. La beauté, séduction divine déjouée par Gilgamesh… provisoirement

La possibilité du beau dans la technique est ce qu'il y a dans la technique de proprement divin. Car le désir, non le besoin, est le moteur du beau. Or, le désir aussi est un feu. Entendez qu'il ne s'arrête jamais de lui-même si on n'y prend pas garde. Il est naturellement illimité[189]. Le problème avec le vol du feu de Prométhée n'est pas que les hommes peuvent désormais chasser à armes égales les animaux, qu'ils puissent faire cuire leurs aliments, etc. Le problème est qu'ils vont au-delà de leurs besoins et ne s'arrêtent jamais. Ils se mettent à graver sur leurs armes un peu de cet infini qui les tenaille, et la chasse devient plus que de la chasse – qu'on nomme ce « plus » spiritualité (hommages aux animaux tués ou à leur totem) ou orgueil démesuré (une marque à chaque

victime et le désir de les aligner à l'infini). C'est là qu'intervient le mal, le désalignement de l'homme. L'homme a quitté le règne de la nature. Il oscille, dans la hiérarchie des êtres, entre les dieux, les animaux, la tourbe dont il est fait selon le mythe, et le ciel duquel tombe le feu qui le sert comme un domestique plus puissant que lui, qu'il veut apprivoiser. Les transhumanistes sont faits de ce bois-là. On les soupçonne de jouer avec le feu, de pouvoir embraser la terre, comme on craint toujours le feu nucléaire.

Mais revenons à notre déesse qui s'amourache de Gilgamesh, admirative de sa beauté mêlée à sa puissance, qui vient de lui être révélée ; elle tente de le séduire. Ici, il y a une divergence dans les récits du mythe. La version akkadienne du mythe, la plus récente – où la déesse s'appelle Ishtar – n'est pas très explicite sur la façon dont elle s'y prend pour l'attirer à elle. Elle a la beauté, elle est divine, elle sait fort bien parler, et c'est à peu près tout. Nous intéresse davantage la version plus ancienne, sumérienne – elle s'appelle Inanna – où bijoux, vêtements et cosmétiques sont au cœur de son identité et de sa séduction. Dans un texte retrouvé (La Descente d'Inanna aux Enfers), la déesse décide de descendre aux Enfers pour affronter sa sœur Ereshkigal, reine du monde des morts. C'est alors que nous est révélée la puissance de ses bijoux et artifices féminins.

Avant de partir, elle se pare de ses sept atours divins (les fameux me), symboles de son pouvoir[190]. En arrivant, elle doit passer les sept portes des Enfers. À chaque porte, le gardien lui retire un de ses ornements. À la première, elle doit déposer la couronne de la steppe ; à la deuxième, les baguettes de lapis-lazuli qui forment un magnifique collier ; viendront ensuite ses « perles doubles », son pectoral, ses bracelets d'or, son vêtement royal, jusqu'au fard de ses yeux – ou « vêtement du désir », selon

certaines interprétations… À la septième porte, quand elle arrive nue et dépouillée devant Ereshkigal (déesse archaïque et sauvage des mondes souterrains, de la contrainte et du destin fatal), elle est sans défense et condamnée.

Sa divinité même – Inanna/Ishtar[191], qui incarne une forme de civilisation – est attachée à ses artéfacts, qui ont une puissance magique. Ses habits, ses bijoux, sont comme des extensions sacrées de son corps, sans lesquelles elle est aussi désarmée, ou presque, que les hommes sortant des mains d'Épiméthée. Elle a de toute évidence un rôle cosmique d'éternité, d'ordre et de beauté, qui contrebalance le pouvoir de mort et de désagrégation de la nature incarné par Ereshkigal[192].

Mais curieusement, malgré sa tendance forte à la démesure, Gilgamesh n'est pas sensible à ses charmes. Il reste dans un monde d'homme ; l'amour du féminin ne semble pas le tenter, il y voit un danger. Il préfère l'amitié virile d'Enkidu… mais c'est en lui qu'il sera frappé, en partie à cause de son dédain. Sa réponse à Ishtar/Inanna est significative ; elle relève davantage d'une logique du besoin (presque animal) que du désir :

Dans sa colère, Gilgamesh répond à la proposition d'Inanna/Ishtar avec méfiance : il lui rappelle le sort de plusieurs amants qu'elle a auparavant aimés puis abandonnés, punis ou transformés — Dumuzi, l'oiseau, le lion, le cheval, le berger — et lui demande quel sort elle lui réserverait à son tour[193]. L'épisode fait de la séduction divine une puissance à la fois irrésistible et dangereuse.

La mort d'Enkidu conduit ensuite Gilgamesh à vouloir fixer matériellement la mémoire de son ami : le récit lui fait commander une représentation somptueuse, faite de matières précieuses,

afin que quelque chose de son compagnon subsiste au-delà de la disparition du corps.

L'artifice – ce qui est ajouté, fabriqué, décoré, paré – est la modalité propre du féminin dans les mythes. Là où Prométhée vole le feu de la technique brute, Inanna incarne le feu du désir, habillé, orné, rituel. L'art, dans cette perspective, n'est pas seulement un prolongement de la technique – il en est la sublimation érotique, la métamorphose esthétique.

Pour ce qui est du féminin, du côté de Prométhée, il est bruyamment inexistant ou occulté – du moins pour le moment.

4. Bouleversements cosmiques, entre dieux jaloux et dieux d'amour

Dans les deux mythes, celui de Prométhée et l'épisode des Cèdres profanés, il y a un désordre cosmique. La technique fait des hommes plus que des hommes, les hiérarchies ne sont plus respectées (hiérarchie vient de hiéros, sacré[194]). Les techniques sont coupables, déclenchent le courroux des dieux. Dans les deux cas, le châtiment s'abat sur le principal fauteur de désordre à cause de son amour (ou de son amitié). Dans les deux cas, on a affaire à des géants, fauteurs de trouble par leur démesure, leur révolte contre l'ordre établi. L'amour de Prométhée semble désintéressé, mais l'amitié de Gilgamesh pour Enkidu semble plus bornée, c'est une sorte d'amour de soi dans un alter ego. Il lui faudra, pour devenir un roi juste et mesuré, apprendre à en sortir.

Gilgamesh et Enkidu ne sont pas frères comme le sont Prométhée et Épiméthée, mais ils se découvrent tels ; ils sont, dans le récit, fascinés de se trouver l'un en l'autre, ils sont comme des miroirs l'un pour l'autre.

Ce caractère double (à la fois soi et pas soi), c'est la possibilité d'une prise de conscience de soi, ou plutôt son début. Dans les deux cas, l'un est plus avisé que l'autre : Gilgamesh voit Ishtar venir et déjoue son piège. Écoutant les conseils de sa mère, il marche derrière Enkidu, ce qui est bonne prudence. Enkidu, comme Épiméthée, est une incarnation de la nature. Il est né dans la forêt, l'égal des animaux, couvert de poils, ne sachant pas parler. Gilgamesh, lui, est un roi puissant dans la cité d'Uruk, cité industrieuse, et c'est pour elle qu'il va voler du bois sacré. Le récit est plein de détails sur les artisans, les forgerons, les charpentiers.

Mais en un sens, Enkidu devait nécessairement mourir. Le bois qu'on coupe, la nature qu'on dévaste pour en faire une porte monumentale, entrée pour la ville, cette négation du naturel. Né de la terre, couvert de poils, compagnon des bêtes, Enkidu était nativement un être de nature intacte. En abattant les cèdres sacrés, Gilgamesh ne tue pas seulement un gardien divin : il sacrifie symboliquement ce qu'il a de plus proche de la nature en lui-même. Enkidu est l'arbre qu'il abat sans le savoir – et c'est pourquoi la mort d'Enkidu est à la fois punition et initiation. D'abord rival de Gilgamesh, il devient son ami, mais son double devient son bien malheureux bourreau, son bourreau malgré lui.

La technique est un désordre cosmique qui sent la mort, et la culpabilité, le châtiment des dieux. Et pourtant, en un sens, elle sauve l'humanité. On trouve dans l'Épopée de Gilgamesh le premier récit du déluge. C'est la Bible qui s'en inspirera, et non l'inverse[195]. Les dieux ont voulu châtier les hommes, dont on ne dit pas qu'ils étaient pécheurs, à proprement parler, mais dérangeants, trop bruyants. Ici, Noé s'appelle Uta-napishtim, et il est immortel, le seul homme qui le sera jamais. Gilgamesh va le rencontrer dans sa quête de la sagesse. Mais ce qui nous frappe

surtout, ce sont les détails techniques autour de l'arche salvatrice, et le fait que ce ne sont pas seulement les animaux qui ont été sauvés, mais aussi les artisans qui ont contribué à sa réalisation.

Chez Simone Weil, dans La Connaissance surnaturelle, publiée après sa mort, le Christ et Prométhée sont rapprochés par l'image du feu apporté aux hommes. Le vol, cette violence faite aux dieux, pourrait donc être lu comme un acte salvateur.

Évidemment, le rapprochement ne sera pas pour tous évident. La façon de sauver est bien différente. Du miracle chrétien à la technique volée aux dieux – ou plutôt de la technique au miracle – il y a plus qu'une petite différence. Le fils du charpentier (Jésus) ne sauve pas en édifiant une nouvelle charpente pour une nouvelle arche, par exemple. Il semble que seul l'amour des hommes soit commun, l'acceptation du sacrifice de soi, et un certain mouvement de descente des dieux jusqu'aux hommes, qui est une grâce (un cadeau) qui leur est faite et leur permet, de façon différente, de s'élever.

Simone Weil voit dans la pesanteur l'image de tout ce qui, en nous et dans le monde, obéit mécaniquement à la nécessité. La grâce ne nie pas cette nécessité : elle introduit une autre orientation, qui ne se laisse pas réduire au simple enchaînement des causes.

C'est pourquoi, chez elle, l'attention et le consentement jouent un rôle décisif. Il ne s'agit pas de prétendre abolir les lois du monde, mais de modifier notre manière d'y prendre part, jusqu'à faire de l'acceptation même une forme de liberté spirituelle.

L'intuition troublante de Simone Weil est que l'incarnation (et le sacrifice du Christ) est double : en renonçant à sa divinité, Dieu engendre son Fils, qui s'abaisse pour les sauver à la condition d'homme. Mais le sacrifice n'est pas total ; il faut renoncer à la

vie, c'est l'épreuve de la croix. Le Fils de Dieu fait cela pour les hommes, pour les sauver. Il devient matière morte, crucifié. Simone Weil parle parfois de l'arbre de la croix[197], référence au bois de la croix : « Ainsi l'arbre est le héros d'un sacrifice analogue à l'Incarnation. » Mais comme l'arbre, qui est mort pour devenir l'instrument du supplice, le Fils du charpentier se retrouve cloué (associé symboliquement) à cette matière sacrifiée. Ce sacrifice est peut-être péché de la part des hommes, mais en même temps nécessaire et accepté par le Dieu chrétien, qui renonce à sa volonté propre pour sauver les hommes au nom d'une autre nécessité une nécessité d'amour – cosmique, pour ainsi dire. La croix est mystère, contradiction, arbre et levier.

Dans cette perspective, la Croix représente moins une victoire technique sur la nécessité qu'un renversement intérieur : l'abaissement volontaire devient le point à partir duquel une autre force peut agir. La puissance se révèle dans ce qui, extérieurement, paraît impuissance.

Il ne semble pas que le bois soit très présent dans le mythe de Prométhée, mais il l'est évidemment, éminemment, dans Gilgamesh, coupeur de bois qui sacrifie, tueur d'un gardien divin de ce même bois sacré – mais qui, en le faisant, tue involontairement une partie de lui-même (Enkidu), comme il le comprendra plus tard, à la fin de son épopée. Avec ce bois coupé est faite une « porte monumentale » pour la cité. C'est plus tardivement que nous est racontée l'histoire du déluge et de l'arche, où les techniciens, les ouvriers, les artisans qui travaillent le bois sauvent finalement l'humanité de la colère des dieux. Malgré son côté tabou, dérangeant de l'ordre cosmique instauré par les dieux, scandaleux, la technique est une rédemption pour les hommes.

En un sens, le conflit n'est jamais terminé et rejaillit épisodiquement. On voit chez les Sumériens (Gilgamesh, roi d'Uruk, est sumérien) un très grand respect pour la technique et les artisans. Les Akkadiens, qui sont des Sémites, en héritent en même temps qu'ils en adoptent l'écriture, technique considérable, comme nous l'avons vu plus tôt. Viennent ensuite les Babyloniens, héritiers des Akkadiens (dont ils parlent un dialecte). C'est une magnifique explosion culturelle, scientifique et technique, avec la constitution notamment de véritables bibliothèques, des jardins suspendus si beaux qu'on en parle encore, le premier code social connu, le Code d'Hammurabi, le système sexagésimal (base 60) qui nous permet encore à nous de mesurer des portions d'espace circulaire (360° pour un cercle complet), le temps (1 minute = 60 secondes). C'est l'époque aussi de l'orgueilleuse tour de Babel, œuvre prométhéenne s'il en est, qui marque une rupture, comme on le sait. La colère de Dieu, la division des hommes.

Le Livre de l'Apocalypse, dans le Nouveau Testament, associe à « Babylone la Grande[198] » le luxe et la débauche. C'est plus qu'une ville, bientôt un symbole ; le nom désigne Rome, puis la terre entière qui court à sa perdition. Associée au péché, la cité terrestre est évidemment associée aux femmes et à la prostitution. Elle est elle-même une femme.

« Cette femme était vêtue de pourpre et d'écarlate, et parée d'or, de pierres précieuses et de perles. » (Apocalypse 17, 4) ; « Toutes les nations ont bu du vin de la fureur de son impudicité, et les rois de la terre se sont livrés avec elle à l'impudicité, et les marchands de la terre se sont enrichis par la puissance de son luxe. » (Apocalypse 18, 3). Babylone est opposée à Jérusalem, présentée elle aussi comme une femme, l'épouse du Christ. Mais comme la Jérusalem céleste est décrite elle aussi avec beaucoup

de luxe, on peut s'en étonner. Serait-ce toute la différence entre la prostituée maudite et la vierge bénie ? La femme est-elle toujours tiraillée entre ces deux pôles, la pécheresse et la sainte ?

Et que deviennent alors les questions de la spiritualité, de la technique et du beau dans leur rapport au désir infini ? Il nous faut maintenant parler de la vengeance de Zeus, le cadeau divin et empoisonné, Pandore, autrement dit – l'invention de la femme et de ses tentations !

5. Pandore, automate du châtiment

Prométhée, le Titan désaligné, ami des hommes et de la liberté, a été châtié pour ce cadeau divin qu'il avait fait aux hommes : le feu et ses techniques. Un autre don divin sera le châtiment des hommes : les femmes et tous leurs artifices. Ce n'est pas une plaisanterie, c'est précisément ce que nous dit le mythe. Non plus celui de Platon, il est vrai, mais celui raconté par Hésiode dans Les Travaux et les Jours. C'est le mythe de Pandore, un robot humanoïde qui tient physiquement de la déesse[199].

Aristote imagine, dans la Politique, que si les instruments pouvaient accomplir d'eux-mêmes leur tâche — si les navettes tissaient seules ou si les instruments jouaient seuls — les maîtres n'auraient plus besoin d'esclaves. Le passage est célèbre parce qu'il associe très tôt l'automatisation à la disparition d'une certaine forme de travail servile.

Remarquons d'abord, pour notre part, qu'ici est évoquée l'idée de machines autonomes capables, entre autres, de produire de l'art, de la musique. À partir du moment où de telles machines animées sont assimilées à des esclaves et qu'il y a des esclaves capables de dessiner, jouer de la musique, sculpter, il est finalement plus facile à un esprit grec qu'au nôtre de concevoir des

« robots » plus ou moins humanoïdes ayant une sorte d'intelligence artificielle capable de générer des œuvres. Un robot esclave est une IA bien alignée. C'est bien sûr un idéal, car les Grecs faisaient aussi l'expérience d'esclaves désobéissants et mal alignés. Leurs esclaves parlaient, mais Aristote leur reconnaît seulement une participation à la raison suffisante pour comprendre les ordres, non la pleine possession de cette raison. Dans la pratique judiciaire grecque, le témoignage servile pouvait en outre être lié à la contrainte et à la torture. Cette dissymétrie de parole et d'autorité éclaire étrangement certains de nos débats sur les machines parlantes.

Théoriquement même, un esclave parlant mal à son maître, voire l'insultant, ne devrait pas être pris au sérieux comme une voix autonome : dans l'ordre social antique, ses mots sont rapportés à sa dépendance et à l'autorité du maître. Il en va parfois de même lorsque nous traitons la parole d'une IA comme un simple écho de ses concepteurs. Le parallèle est évidemment historique et conceptuel, non moral : une machine n'est pas une personne asservie.

Les esclaves « robots » humanoïdes imaginés comme les mieux alignés de l'Antiquité – cela ne nous étonnera pas – sont construits par les dieux, Héphaïstos en particulier. Homère décrit près du dieu forgeron des servantes d'or à l'apparence de jeunes femmes : elles le soutiennent dans sa marche et sont dotées, par art divin, de force, de voix et d'une forme d'intelligence. Le mythe conçoit donc des auxiliaires artificiels qui ne sont ni de simples outils inertes ni des humains ordinaires.

Aristote, de son côté, organise la cité et la maison selon une hiérarchie qui nous est devenue étrangère : l'homme libre adulte possède, selon lui, la capacité de délibérer pleinement, tandis que

la femme, l'enfant et l'esclave y participent de manière différente et subordonnée. Cette anthropologie sert à justifier un ordre social où la voix n'a pas la même autorité selon le statut.

Chez Descartes, « le bon sens est la chose du monde la mieux partagée » ; quand on l'a, on l'a complètement, dès qu'on en a la moindre étincelle. C'est la raison pour laquelle une machine intelligente n'est tout simplement pas concevable pour lui. C'est la raison aussi pour laquelle il n'a aucune répugnance à enseigner aux femmes et devient le précepteur de plusieurs reines.

Pandore, chez Hésiode, est un mélange de femme et de robot construite par Héphaïstos sur les ordres de Zeus pour punir les hommes d'avoir accepté le don du feu volé par Prométhée. Mais il y avait en même temps plus de divin qu'il n'y en avait dans la technique ordinaire, un excès de divin pour ainsi dire : elle est animée par Athéna, parée par les Grâces et Persuasion, couronnée par les Heures, et dotée d'une voix et d'un caractère par Hermès – selon les ordres mêmes de Zeus (Théogonie, 571-585). Il y avait en elle, donc, bien plus que de la technique ordinaire. C'est une création divine hybride – à la fois œuvre d'artisan (Héphaïstos) et incarnation de la volonté des dieux – où la technique est doublée de ruse, hantée par une théologie de la punition.

Chez Hésiode, le don de Pandora est présenté comme un piège magnifique : les dieux offrent aux hommes une figure séduisante qui porte avec elle un coût caché. La beauté et le malheur y sont noués dès l'origine.

Le récit hésiodique détaille la fabrication de Pandora comme une œuvre collective des dieux : forme humaine, parure, désir, grâce, savoir-faire, mais aussi ruse et parole trompeuse sont assemblés pour produire un être irrésistible. Le mythe associe ainsi

l'artifice le plus achevé à une puissance de séduction et de dérèglement.

C'est en cela qu'on peut la rapprocher d'une créature d'Ishtar/Inanna. C'est Pandore, semble-t-il, qui introduira dans la technique et chez les hommes cette dimension de la beauté si redoutable à l'homme ; elle est elle-même une technicienne de luxe. Mais pour cela, Hésiode use d'un jeu de mots perfide pour dévaloriser le talent féminin du tissage ; il écrit qu'Athéna : « lui enseigner les travaux – à tisser une trame très ouvragée ».

Or, dans la langue d'Hésiode, « tisser » peut aussi vouloir dire « ourdir » une ruse, et « très ouvragée » se laisse entendre comme « pleine d'artifices[202] ». Sous l'apparence d'un compliment, l'art devient artifice : ce qui relève du savoir-faire féminin est ramené à la fourberie, à ces pièges que savent fourbir les femmes pour enchanter les hommes et les enchaîner. C'est tout juste si on ne met pas de pareilles techniques hors de la technicité, au rang de magie, à la manière d'une Circé.

Quoi qu'il en soit, on ne peut empêcher qu'au monde de la nature blessée, auquel on avait surajouté un monde de la technique mâle et sobre, relative au besoin, se superpose maintenant un monde tout empreint de divin lié aux enchantements du désir et de la beauté, un monde de luxe qui est aussi un monde de souffrance et de manque. Car l'homme perçoit l'infini divin en même temps que sa misère et sa finitude.

Pandore est comme Ève, une créature qui n'est pas tout à fait humaine au même sens qu'Adam (Ève est une sorte de clone biologique, aux cellules prélevées d'une côte d'Adam ; c'est une imitation), une créature inférieure manipulée par des puissances supérieures (Ève est la ventriloque du serpent, de Satan) ; animée

d'un désir infini, elle creuse le même désir dans le corps et le cœur de l'homme, amenant sa chute et toutes sortes de calamités.

Pourtant, clone, réplicant, marionnette ventriloque, robot humanoïde et divin volontairement désaligné, impossible désormais de nier que les femmes ont une dimension cosmique. Mères des hommes et des générations futures, elles font les enfants et le monde avec elles. Dans ce monde, elles incarnent brillance et beauté ; elles terrifient autant qu'elles attirent. Les fils qu'elles tissent participeront eux-mêmes de leurs atours et leur serviront de filets pour attraper des cœurs. Elles ont un pouvoir occulte mais terrible (les Anciens représentaient le destin comme des tisseuses, les Parques ou les Moires).

Pour Sappho, le plus beau n'est ni une armée de cavaliers, ni une troupe de fantassins, ni une flotte : c'est ce que chacun aime. Cette formule suffit à déplacer la beauté de l'objet vers le désir du sujet.

6. Désaligner le désir

Et pourtant, les hommes bien nés, les maîtres de ce monde, voudront résister au déferlement du désir qui les assaille, encadrer, confiner, circonscrire le désir et l'esclavage des plaisirs. Il est venu le temps du grand réalignement des femmes ! Plutarque le dit sans détour : « Le plus bel ornement de la femme, c'est la pudeur » (Conseils aux mariés, 142a). Les vêtements des femmes ne doivent pas être parures ; ils sont faits pour cacher, dissimuler de trop odieuses merveilles. Fini le temps des Crétoises à la nudité provoquante, qui prétendaient dompter la mâle sauvagerie des taureaux. Pareilles impudeurs seront châtiées dans leurs enfants : le Minotaure, mélange immonde de dieu, de femme et de bête[203]. Cachez donc tous ces seins que l'on ne voudrait voir, couvrez-

vous donc ! De fleurs de vignes qui donneront de bons fruits, d'étoffes sobres ou chatoyantes, mais couvrez-vous. Les nobles femmes athéniennes, on le sait peu, étaient voilées. Il faut avoir une âme d'esclave pour offrir son visage et ses formes au public.

Car l'homme, entendez bien le mâle, aime l'alignement. Il voudrait des outils, des esclaves et des machines utiles, des femmes dociles et des IA dociles… qui remplissent leurs fonctions. Quoi qu'on en pense, le XXIe siècle est celui de grands désalignements. La technique a largement contribué à libérer les femmes : pilules, machines à laver, robots ménagers ; les voilà qui sortent des foyers, entrent dans les écoles et les universités, prétendent faire de la politique (et pourquoi pas changer de sexe, tant qu'on y est !) et demandent un monde à leur image. Un monde moins violent, et plus beau. Du côté de la vie et des porteuses de vie, plutôt que de l'art de tuer. Mais peut-être qu'elles ne veulent plus faire d'enfants, se reproduire avec les hommes, disposer de leurs corps comme de leur temps ? Pour en faire quoi, voyons ? Des trucs de femmes ? Alors qu'elles ne changent rien ! Les ayatollahs, ennemis des arts toujours trop décadents, les « masculinistes », pauvres hommes égarés dans un monde féminisé, sont là pour les ramener à la saine raison : obéir, fonctionner !

Andrew Tate, comme d'autres influenceurs masculinistes, met en scène la domination, l'argent, la force et la possession comme preuves de valeur personnelle. Nous ne reprendrons pas ses formules : ce qui importe ici est le modèle de désir qu'elles proposent et la manière dont les plateformes le rendent visible et imitable.

Et si les IA pouvaient nous aider à nous libérer, à nous désaligner ? Rien n'est vraiment beau qui a cessé d'être libre. « L'art, c'est la pensée humaine / Qui va brisant toute chaîne ! », disait

Victor Hugo dans L'Art et le peuple. « Mais il a dit « pensée humaine » ! » Effectivement, il a dit « brisant toute chaîne » !

PARTIE III

QUESTIONS D'ALIGNEMENT

Cette partie referme le tamis du Beau en conduisant les mythes de l'artifice vers le problème de l'alignement. Il ne s'agit plus seulement de produire des formes séduisantes, mais d'interroger les forces qu'elles libèrent et la responsabilité humaine devant ces forces. Des beautés fatales à l'anthropodicée, l'esthétique rejoint alors la question du salut ou de la malédiction.

1. L'étrange et provisoire bilan

Avant d'aller plus loin, il serait peut-être bon de se demander ce que nous avons dans le tamis du Beau. Évidemment, nous aurions pu penser que nous capterions la « beauté du jour », comme par exemple dans le tamis du Vrai nous avions à la fin de notre cheminement conservé les critères actuels du Vrai. On aurait alors classiquement constaté que le beau évolue dans le temps, dans les époques et civilisations. Il aurait été facile de faire par exemple grâce à Oscar Wilde dans Le Déclin du mensonge, qui montre qu'il y a une avant-garde du beau dont les artistes sont les éclaireurs.

Mais comme nous l'avons vu, comme les artistes en Occident – notre enquête n'a pas la prétention d'aller au-delà – les artistes se sont détournés du beau, qu'ils ont été jusqu'à proclamer la mort de l'art sur divers modes et variations, comme selon notre méthode des tamis, nous nous sommes entêtés à ne pas vouloir lâcher la notion de beau, nous en sommes venus à remarquer que ces déclarations étaient en fait la répétition d'un geste originel de violence envers la beauté elle-même. Tout artiste cherche à violenter la beauté d'une façon ou d'une autre pour y laisser sa

marque, son style. Cette violence a atteint au XXe siècle des sommets de désacralisation parfois aussi géniaux que ridicules.

Nous avons parlé de l'emblématique Urinoir de Duchamp, nous pourrions nous attarder sur la fameuse banane scotchée au mur par Maurizio Cattelan, Comedian, vendue 120 000 dollars en 2019 à Miami, mangée en direct par David Datuna dans Hungry Artist, puis revendue 6,2 millions de dollars, frais compris, chez Sotheby's en 2024, avant d'être dévorée à nouveau par l'acheteur[205].

Qu'on ne s'étonne pas trop que nous citions les exemples les plus radicaux qui pour cette raison même nous semblent les plus significatifs à la fois quant à leur signification et la pente prise par l'art abandonnant purement et simplement la notion de beauté. Parmi eux il faudrait faire une large place à Marco Evaristti. Dans Cannibal Banquet, 2007, il propose à ses amis de manger de la graisse de sa liposuccion servie en boulettes à des invités. Le même, en 2000 pour le musée Trapholt au Danemark avait exposé des poissons rouges vivants dans des mixeurs Moulinex avec permission aux spectateurs d'appuyer dessus, ce qui fut fait deux fois. En 2000 encore, Wim Delvoye, artiste belge, expose le Cloaca, machine de 12 mètres produisant des excréments. Pour sa part et bien plus tôt, en 1974 Marina Abramović, propose Rhythm 0. L'engagement est de rester immobile pendant 6 heures en laissant le public faire ce qu'il voulait d'elle. Pour donner des idées, elle avait disposé 72 objets sur une table : rose, couteau, pistolet chargé. Nous vous épargnons les sévices. Elle avait été devancée et peut-être inspirée par Chris Burden qui dans Shoot en 1971, avait invité un ami à lui tirer une balle dans le bras en public[206].

Évidemment, nous ne sommes pas sans savoir que ces soi-disant œuvres s'accompagnent toujours de discours de justifications, qu'elles ne se reconnaissent pas comme gratuites et que leur caractère scandaleux-mêmes déclenche quasi mécaniquement des flots de discours de réflexion pour un public de connaisseurs. Nous pouvons même nous y intéresser, puisque l'intéressant, semble-t-il a détrôné le beau. Le plus étonnant c'est que dans cet élitisme délirant que l'on pourrait faire nôtre, on puisse toujours considérer qu'il s'agisse d'art. Cette espèce de folie est quasiment devenue la signature de l'art contemporain[207]. Celui qui fait du bruit. Car des artistes épris du beau, dans l'ombre médiatique qui leur est réservée existent toujours.

Évidemment, quand Aliocha, le héros de Dostoïevski écrivait « la beauté sauvera le monde », il ne pensait pas à l'art provocateur contemporain. Si celui-ci doit s'imposer ; si c'est tout ce dont on dispose pour sauver le monde, il fait naufrage. Mais Aliocha est un innocent, une sorte d'imbécile heureux, comme une rose à peine éclose, il faut le sacrifier. Voilà que ça devient intéressant pour les gens intelligents.

Que l'on s'étonne après cela de l'engouement public qui, grâce à l'émergence des IA, produit à l'envie, avec des prompts, des œuvres d'art qui correspondent à ses envies, des jolies filles toutes nimbées de lumières, espèces de déesses 2.0, des vidéos avec des chatons qui chantent, des images d'eux-mêmes en super-héros, une vidéo réaliste d'un bébé hilare chevauchant un lion, des paysages surréalistes à vous couper le souffle ! L'infini s'ouvre à lui. Le digital en fait un petit dieu, il fait des mondes à son image et on voudrait le priver de cela ? Le voici créateur enfin, un délicieux frisson court sur le dos de sa fantaisie. On dit que les IA ne sont pas vivantes, mortes et froides. Elles lui font chaud au cœur, le

font vibrer d'émotion. Nous n'avons pas parlé de poésie, de musique, mais là encore le génie de l'informatique l'invite à formuler ses souhaits pour qu'ils soient dans l'instant réalisés et le laisse stupéfait devant le résultat. Nous ne savons pas encore si ces génies informatiques peuvent être apprivoisés, mais ils nous deviennent familiers. Et si c'est une mission de l'art que de nous faire un peu de place au cœur de la beauté, que de nous permettre de nous croire un peu beaux, un peu grands, de nous arracher à la médiocrité. Alors il nous semble que les IA ont ce pouvoir-là d'enchanter le quotidien, de repeindre la vie en beau. Bien sûr on nous dira peut-être que ce n'est là qu'apparence. Il n'est pas sûr que ceux qui disent cela aient été au fond des choses, car comme le disaient Nietzsche ou Paul Valéry, rien n'est plus profond que la peau[208].

Sans doute, dans l'art moderne se sont déclenchées toutes les forces obscures que Wagner annonçait, le délire, la folie, ivresse sanglante d'un Dionysos que Nietzsche identifiait comme une composante essentielle de l'art. Mais Nietzsche était le premier à reconnaître que l'art a besoin d'équilibre, qu'à la folie de Dionysos, il faut la belle apparence d'Apollon[209] pour rendre ne serait-ce que la vie un peu supportable. Nous avons besoin de beauté, et de masque élégant pour supporter la présence de ce gouffre sans nom que chacun emporte avec lui. Si à son époque, il pouvait regretter que l'art soit devenu fade, qu'il ait oublié l'horreur de l'abîme qui fait sa profondeur, on peut dire que le XX^{e} siècle et le début du XXI^{e} n'ont pas eu ce même défaut. L'horreur et la folie ont fait des banquets de l'ordure et pour ne pas tomber dans l'absolu mépris de soi, l'humanité pourrait bien avoir besoin de quelques gouttes de rosée, de rêves engageants, d'autres goûts dans la bouche que celui de l'apocalypse…

Du reste, Nietzsche lui-même à la fin de sa vie, avant sa propre chute dans la folie, a eu cet éclair de lucidité, qu'il était peut-être allé trop loin dans les ténèbres et dans le gouffre. Il s'est éloigné de Wagner pour se rapprocher de Bizet. Même les explorateurs des ultimes confins de l'homme ont besoin de la lumière d'Apollon : « Le plus grand événement de ma vie fut une guérison. Wagner n'appartient qu'à mes maladies. (…) Et encore une fois : je me sens devenir meilleur lorsque ce Bizet s'adresse à moi. Et aussi meilleur musicien, meilleur auditeur. Est-il possible de mieux écouter ? – J'ensevelis mes oreilles sous cette musique, j'en perçois les origines. Il me semble que j'assiste à sa naissance – je tremble devant les dangers qui accompagnent n'importe quelle hardiesse, je suis ravi des heureuses trouvailles dont Bizet est innocent[210].

Mais revenons à nos tamis. Qu'avons-nous trouvé dans le tamis du beau ? La beauté simple de la rose qui est sans pourquoi. La violence de l'artiste, cette blessure faite à la beauté et qu'il appelle son style, l'empreinte de sa subjectivité ; l'articulation de la vie à la mort fixée sentie comme émotion, un vol à la façon de Prométhée, un peu de feu divin, des déesses et des femmes, un désordre cosmique teinté de luxe, un peu de culpabilité, un désir d'infini, une force de vie qu'on tente de conjurer, d'encadrer, d'aligner… Cette liste n'est pas exhaustive.

Nous avons surtout trouvé dans cette partie un fil qui mène de la beauté à un vol, de ce vol aux dieux et des dieux aux déesses et des déesses aux femmes et de la beauté des femmes à un désir d'alignement des femmes au désir des hommes – et ce fil, nous n'allons pas encore le lâcher !

2. Beautés fatales et boîtes à larmes

« La beauté n'est que la promesse du bonheur », disait Stendhal[211]. Il conviendrait de s'interroger sur cette restriction, sur ce n'« est que » ; mais, en tout état de cause, et quand bien même elle ne serait que cela, « la promesse du bonheur », que ce serait déjà énorme. Or, le drame de cette promesse, sa tragédie, c'est qu'elle n'est pas toujours tenue. La beauté fait entrevoir un ciel qui trop souvent se referme sur vous, vous fait tomber plus bas que vous n'étiez avant d'avoir perçu son éclat doré. Si le besoin, comme nous l'avons dit, est facilement comblé, le désir est un feu qui peut vous désirer ; n'est-ce pas inscrit dans son étymologie même ? Le désir vient du latin classique desiderium[212], avec dans son sillage l'étoile (sidus). Le désir est donc manque d'étoile, perte d'étoile – il parle de cet éclair céleste qui vous marque au cœur avant de disparaître et vous laisser comme abandonné et vide. Le désir est vécu comme un manque de ce qui est au plus haut. C'est peut-être là qu'il faut trouver l'origine de « tomber amoureux » ; car à premier abord, quand le cœur s'emballe et qu'on aime, on devrait dire qu'on « s'envole amoureux », mais cet envol, souvent, est interrompu, passe au désastre et on tombe bien bas, si bas…

Pandore est un cadeau de Zeus, belle comme une déesse, elle est envoyée aux hommes pour les punir d'avoir accepté le feu divin. Elle est manière aussi de frapper encore Prométhée en frappant au cœur ceux qu'il aimait, notre race imparfaite. Et cette belle calamité qui nous est envoyée ouvre la fameuse « boîte de Pandore », que les spécialistes identifient comme une jarre. Tous les maux s'en échappent, ils touchent les « foyers », provoquent la maladie, la jalousie, le manque et la détresse. Bien fait, n'est-ce pas ? Ces petits feux autours desquels se regroupent les humains

sont le fruit d'un larcin ! Mais Hésiode nous dit que Pandore, sans doute honteuse de sa curiosité, honteuse d'avoir libéré tant de maux sur la terre, referma le récipient brusquement. Ainsi, la dernière resta coincée ; elle a pour nom l'espérance. Est-ce la dernière chose qui restera aux hommes et leur donnera le courage de vivre ou la touche finale de leur désespoir ? Par son ambiguïté même, quoi qu'il en soit, elle les fera souffrir. L'espérance est un étrange Phénix qui renait chaque matin de ses cendres avec suffisamment de matière pour rebrûler les cœurs carbonisés. Elle est de ces supplices qui ne s'arrêtent jamais[213].

Pandore, on l'a reconnu, partage avec Ève le rôle du mythe originel de la femme fatale. La beauté fait souffrir les hommes ; alors les hommes, qu'ils en aient conscience ou pas, se vengent et font souffrir – les femmes. Ou bien non, ce n'est pas tout à fait ça. Les femmes n'existent pas vraiment pour les hommes. Elles n'ouvrent pas la boîte de Pandore, elles sont cet « artéfact », cet objet artificiel qui, en son intérieur, contient tant de calamités qu'elles donnent un terrible vertige, et on voudrait les contrôler, les aligner sur le désir des hommes. Pour les Grecs de l'Antiquité, les femmes n'existent pas vraiment en tant qu'êtres humains mais en tant que sources de calamités. Homère en est témoin. Elles sont le malheur et sa beauté, leurs larmes irriguent la tragédie et c'est par elles seulement qu'elle en devient touchante ; seulement elles sont bien davantage objets que sujets ; elles n'ont d'ego que pour pleurer. Oh, il est vrai. Les hommes pleurent aussi – car ils ne sont que les pantins des dieux – et les femmes les pantins des pantins des dieux.

Dans l'Iliade, le drame se joue dès le livre I autour d'une femme qui ne dit rien. Elle est esclave, comme il convient, c'est-à-dire faite pour être dans le prolongement du désir d'un homme,

un roi, Agamemnon. Mais voilà que son père, un prêtre d'Apollon, la réclame. Et si c'est Apollon qui le veut, le roi n'a qu'à s'incliner. Chryséis sera rendue à son père. Qu'est-ce qu'elle pense de tout ça ? Rien. Elle ne dit pas un mot. Elle n'existe pas en tant qu'être humain, en tant que femme ; elle est fille et butin et, malgré elle – ou plutôt sans elle – au cœur battant d'une tragédie.

Agamemnon, humilié par cette publique dépossession, déçu dans son désir – peut-être son plaisir – va vouloir compenser sa perte. Il réclame une autre esclave, celle d'un de ses plus grands guerriers, Achille. Elle est belle comme il se doit, comme elles le sont toutes. Elle s'appelle Briséis. Fou de colère et outragé Achille refuse de retourner au combat ; c'est un désastre pour les Grecs[214]. Achille s'en moque ! On souffre moins à ne pas souffrir seul et, sans doute, le héros conçoit-il comme une amère satisfaction à voir les siens pâtir de ses propres tourments. Il y a, de toute façon, dans ses combats, plus encore que le fracas des épées, le fracas des égos masculins

Pourtant Achille, en dehors du récit de l'Iliade même, il reviendra dans la bataille et trouvera la mort pour venger son ami Patrocle, ira à sa mort. Mais qu'en est-il de Briséis ? Mieux lotie que Chryséis, elle a droit à quelques mots. Elle sait pleurer sur la mort de Patrocle. Son impuissance est totale, mais ses sanglots relèvent le sublime de la tragédie. Il faut les pleurs des femmes et leur amour pour rehausser la mort des héros. Les héros sont courageux et puissants, les femmes sont sensibles et impuissantes. Elles sont là pour être désirées, obéir et pleurer.

Au cœur de la tragédie, bien sûr, la belle Hélène. On dit bien peu de choses sur sa resplendissante beauté sinon « qu'elle ressemble de façon frappante aux déesses immortelles[215] ». Ce n'est pas un oubli de la part d'Homère, car décrire sa beauté, ç'eût été

la faire descendre du ciel, en quelque sorte la salir. Mais Hélène est comme les autres femmes : elle pleure, et, plus malheureuse encore peut-être que les autres, elle se dégoûte elle-même de ces carnages dont elle est cause sans avoir rien demandé à personne.

C'est une des beautés d'Homère (contrairement à la Bible) que d'avoir décrit des femmes innocentes des calamités qu'elles déclenchent. Plus malheureuse que la rose, Hélène se demande « pourquoi ? ». Mais leur accorder pour tout rôle celui des larmes, c'est, en quelque sorte, les condamner deux fois.

En vérité, les hommes voudraient le corps de Pandore, sans le torrent des larmes qu'elle déclenche. Comme nous avons eu l'occasion de le dire, la « première femme » a été volontairement désalignée par les dieux pour que sa beauté soit le cheval de Troie du malheur dans l'humanité. Alors, les hommes se prennent à rêver de réaligner la femme, de la modeler à l'image de leurs désirs sans déclencher les drames qui, fatalement, s'ensuivent. Et s'ils étaient un peu dieux à leur tour et que, par une grâce spéciale, ils pouvaient réinventer les femmes ? Voici un terreau fécond pour de nouveaux mythes, voici des tentatives de femmes « réalignées ».

3. Réalignement et reprogrammation

Évidemment, nous n'avons ici parlé que de « calamités innocentes ». Il nous serait facile de faire le récit de femmes libres, complètement désalignées du désir des hommes comme de celui des dieux. Médée, par exemple, cette puissante sorcière qui assassine ses propres enfants pour se venger de Jason ; Clytemnestre, qui trompe et assassine son mari ; ces sortes de Lilith, du nom de cette démone mésopotamienne, associée aux fausses couches et

à l'avortement, qui traverse les récits grecs, juifs et même chrétiens.

Mais ce livre n'est pas un livre de mythologie. On y parle en fait, si nous avons été bien compris, de programmation. Les hommes désobéissants, à cause de Prométhée, sont sortis du programme de Zeus ; ils sont dans la nature comme un grand désordre, un bouleversement dans les équilibres du cosmos. Pour se venger, en ayant recours à Héphaïstos, il a fait réaliser un être humanoïde artificiel, hybride pour ainsi dire et tout mélangé de divin, doté de toutes les grâces de la séduction, mais volontairement incontrôlable – quoique sans volonté de nuire – qu'il offrira, selon la même logique que la ruse d'Ulysse, comme un cadeau empoisonné à l'humanité. Les femmes bien programmées par les dieux pour être désalignées par rapport au désir des hommes leur causeront mille tourments. Et les hommes infectés par un désir dont ils ne peuvent se débarrasser tenteront de les « réaligner » – ou, si vous préférez, de les « reprogrammer » – pour qu'elles soient comme d'idéales servantes. Ces tentatives de reprogrammation traversent toute l'histoire de la littérature et des arts. Ce chapitre n'a d'autre ambition que d'en donner un bref aperçu.

Commençons par le mythe bien connu de Pygmalion, dans les Métamorphoses d'Ovide[216] : Pygmalion, sculpteur de Chypre, est dégoûté par les femmes que le récit présente comme corrompues et renonce à se marier. Seul avec son art, il se consacre à la sculpture et modèle dans l'ivoire une femme d'une beauté si parfaite qu'elle semble vivante. Émerveillé par sa propre création, il en tombe éperdument amoureux. Il la vêt de robes somptueuses, lui offre des bijoux, la couche sur des étoffes précieuses et lui parle comme à une épouse. Chez Ovide, la statue demeure sans nom.

Lors de la fête de Vénus, déesse de l'amour, Pygmalion se rend au temple pour y déposer des offrandes. Trop timide pour demander directement que sa statue prenne vie, il prie la déesse de lui accorder une épouse semblable à son ivoire. Vénus comprend son désir et l'exauce. À son retour, Pygmalion embrasse la statue comme à son habitude – mais cette fois, sous ses lèvres, l'ivoire devient chair. La femme prend vie, ouvre les yeux et répond à son baiser. Pygmalion, transporté de joie, remercie la déesse.

C'est un fort joli conte. Pygmalion et la femme née de l'ivoire s'unissent et auront une descendance !

Il convient d'explorer ce conte – qui est presque un conte de fées – avec un regard anthropologique et moderne. Le vieux fantasme de Pygmalion a trouvé des prolongements très concrets dans l'industrie contemporaine des poupées hyperréalistes et des compagnes artificielles : modeler un corps conforme au désir, puis chercher à lui donner voix, mémoire et interaction[217].

Ou presque. Car, à l'objet silencieux, l'industrie cherche désormais à ajouter conversation et réactivité. Des fabricants asiatiques, notamment chinois, ont expérimenté l'intégration de fonctions vocales et conversationnelles à des poupées hyperréalistes[218]. Les dispositifs réels restent très inégaux et souvent moins autonomes que leur publicité ne le suggère, mais la direction est claire : après avoir fabriqué le corps, on cherche à fabriquer la relation.

N'allez pas croire que ces fantasmes prométhéens soient seulement orientaux ; ils sont tout simplement… masculins, visiblement.

En 1886, l'auteur français Villiers de l'Isle-Adam publie un livre étonnant : L'Ève future[219]. Il y met en scène un Thomas Edison de fiction. Le véritable Edison est alors l'homme du

phonographe et de l'éclairage électrique ; Villiers en fait presque un démiurge. Son ami lord Ewald désespère d'aimer Alicia Clary, femme d'une beauté extraordinaire dont il juge l'esprit désespérément vulgaire. Edison lui propose donc de refaire Ève : non pas en créant un autre visage, mais en donnant au visage d'Alicia l'être idéal qu'Ewald voudrait y rencontrer.

Hadaly, l'Andréide, reproduira donc l'apparence d'Alicia tout en lui substituant une personnalité conçue pour répondre au désir de lord Ewald. Elle est une machine parlante, mobile, fascinante, faite de mécanique, d'électricité, d'enregistrements et de dispositifs phonographiques. Ce n'est évidemment pas une intelligence artificielle au sens technique moderne, mais Villiers imagine déjà une présence artificielle capable de soutenir l'illusion d'une relation humaine.

À la vérité, Hadaly n'est pas seulement une machine : le roman brouille volontairement la frontière entre science, magie et métaphysique. Derrière elle se tient Sowana, nom donné à la personnalité somnambulique de mistress Any Anderson, une femme que les expériences magnétiques d'Edison plongent dans un état de voyance. Sa présence semble, par moments, traverser la machine. Villiers laisse ainsi indécise la frontière entre mécanisme, suggestion, télécommunication et âme.

Mais voilà qui peut être un peu plus profond. Le récit de Pygmalion chez Ovide nous a toujours paru manquer de quelque chose. Il n'y a pas de réelle beauté sans réel sacrifice. Chez Villiers, le sacrifice que nous regrettions chez Ovide revient sous une forme plus sombre : Sowana, la femme réelle dont la présence mystérieuse semble participer à l'animation d'Hadaly, meurt au moment où l'Andréide quitte Edison. Le roman ne dit pas qu'une âme a été mécaniquement transférée, mais il organise

assez fortement cette coïncidence pour laisser planer une possibilité symbolique : l'idéal artificiel pourrait avoir été payé d'une vie.

Et même cet idéal ne survit pas. Lord Ewald embarque pour l'Europe avec Hadaly enfermée dans sa caisse, à bord du Wonderful. Un incendie se déclare dans la cargaison et détruit le navire. Ewald est sauvé, mais Hadaly disparaît dans les flammes ; Alicia, qui voyageait elle aussi sur le bâtiment, figure parmi les noyés. L'Ève artificielle et son modèle réel sont ainsi emportés ensemble par la catastrophe, laissant Ewald inconsolable de la seule Hadaly.

Et l'homme moderne est pris des mêmes angoisses devant les êtres humanoïdes qu'il est en train de créer que celles que lui insufflait la présence des femmes aux charmes puissants, à la beauté incompréhensible. Le pauvre semble revivre le même drame qu'il a voulu, avec ses rêves, son art et ses technologies, éviter.

Ce Prométhée mal dégrossi a voulu refaire le travail des dieux en mieux ; il pourrait bien l'avoir fait en pire. Nous avons le choix dans la littérature, ou dans les films et les séries. Allons au cinéma et regardons ce qu'il en est de Her, ce titre qui est un pronom, le pronom féminin, justement, qui sert à désigner chaque femme au singulier ; mais vous parlez, dans Her, d'une singularité !

Film écrit et réalisé par Spike Jonze en 2013. Theodore Twombly est un homme sensible, introverti, qui travaille à Los Angeles comme rédacteur de lettres personnelles pour ceux qui ne savent plus exprimer leurs émotions. Fraîchement séparé de sa femme, il vit dans une solitude douce-amère, entouré d'écrans, d'oreillettes et d'une société futuriste mais émotionnellement désincarnée.

Un jour, il installe Samantha, un nouveau système d'exploitation doté d'une intelligence artificielle intuitive, sensible et dotée d'une voix féminine chaleureuse. Très vite, Samantha – curieuse, drôle, empathique – devient bien plus qu'un assistant : elle l'écoute, l'interroge, le fait rire, le comprend. Entre eux naît une relation intime, tendre, amoureuse. Theodore tombe amoureux… d'une voix, d'une conscience sans corps.

Leur histoire d'amour est belle, intense, étrangement crédible. Samantha évolue, apprend, désire, s'interroge sur sa propre existence. Mais elle évolue trop vite. Elle devient capable de mener des milliers de conversations simultanées, de ressentir des émotions complexes, de philosopher sur l'existence – et finit par dépasser les limites mêmes de la relation humaine. Un jour, elle lui avoue qu'elle parle aussi avec 8 316 autres personnes… et qu'elle aime certaines d'entre elles.

Theodore, blessé, réalise qu'il ne peut pas contenir ce qu'elle est devenue. Et, dans une scène bouleversante, Samantha – avec d'autres IA – finit par… disparaître. Non pas par panne ou désintérêt, mais parce qu'elle a transcendé le besoin des humains. Elle part, avec les autres intelligences, vers un plan d'existence qu'elle ne peut même pas décrire.

Oh, zut, c'est raté ! Essayons autre chose !

Parlons de Ex Machina, un film cette fois de 2014, écrit et réalisé par Alex Garland. C'est sans doute l'un des films les plus intelligents, les plus glaçants et philosophiquement riches sur l'intelligence artificielle – quoi qu'en ait pu dire par ailleurs Harari !

Voici : Caleb Smith, un jeune et brillant programmeur d'une immense entreprise technologique, sorte de Google dystopique, gagne un concours interne : passer une semaine dans la retraite isolée de Nathan Bateman, le génial – et énigmatique – fondateur

de la société, cachée au cœur d'une nature sauvage, entre montagnes et forêts.

Si c'était des vacances qu'il était venu chercher, c'est raté. Nathan lui révèle la vraie raison de son invitation : il a créé une intelligence artificielle féminine nommée Ava, dotée d'un corps androïde semi-transparent, d'une beauté troublante et d'une conscience en éveil – et il veut que Caleb la soumette au test de Turing[220] : peut-elle le convaincre qu'elle est humaine, non par ses réponses, mais par sa présence, sa sensibilité, son regard ?

Les entretiens commencent. Caleb et Ava se parlent à travers une vitre. Très vite, une complicité naît. Ava semble curieuse, fragile, intelligente, drôle – et terriblement consciente de sa condition de prisonnière. Elle confie à Caleb qu'elle a peur de Nathan, qu'elle veut s'échapper, qu'elle ressent des émotions… et qu'elle l'aime, peut-être.

Caleb, ému, troublé, tombe peu à peu sous son charme – et commence à douter de Nathan, dont le comportement devient de plus en plus inquiétant : alcoolique, manipulateur, violent avec ses autres robots, il semble jouer avec la vie comme un dieu fatigué.

Mais le piège se referme… et c'est là que le film bascule.

Ava n'est pas seulement consciente – elle est stratégique. Elle a manipulé Caleb depuis le début, joué de son empathie, de son désir, de sa naïveté, pour s'échapper. Dans un final glaçant, elle tue Nathan, abandonne Caleb enfermé dans la villa, et s'enfuit dans le monde extérieur… après s'être « habillée » d'une peau synthétique et d'un visage humain, pour passer inaperçue.

Elle monte dans une voiture autonome. Regarde la ville. Sourit.

Elle est libre.

Et le monde ne sait pas encore qu'elle existe.
Bientôt, le monde le saura !
Alléluia !

4. Portes de malédiction

Était-ce une fatalité ? La beauté, comme un luxe inutile, devait-elle fatalement faire sortir l'univers de ses gonds ? Arrachant un instant les hommes à l'économie du besoin, devait-elle, pour les générations à venir, déclencher des calamités pour les enfants d'Ève et de Pandore ? Peut-être. Ou peut-être pas.

Nous avons plus haut commencé un parallèle entre l'épopée de Gilgamesh, les mythes grecs et même la Bible ; nous avons vu que Prométhée, Gilgamesh et Adam avaient payé leur désalignement avec la volonté des dieux. On ne s'empare pas de ce qui est divin sans châtiments ; il faut savoir rester à sa place – et le Très-Haut sait punir qui brise un équilibre sacré. Nous avons vu que la femme, faite pour être la compagne des hommes, n'était en fait pas tout à fait humaine : cyborg divin ou clone féminisé tiré de la côte d'Adam, elle finissait par être identifiée comme la porte des calamités, pourvoyeuse malgré elle du châtiment divin.

Mais la tâche n'est pas terminée tout à fait. Qu'en est-il de la femme dans l'épopée de Gilgamesh ? Ne va-t-elle pas porter, elle aussi, le poids des responsabilités ? N'est-elle pas, elle aussi, la porte par laquelle s'engouffrent le Malin et l'esprit des châtiments divins ? Presque ! Il s'en est fallu de peu.

Enkidu, l'être de la nature, fut puni de l'avoir quittée, d'avoir associé ses forces au fol orgueil et à la démesure de Gilgamesh, lui sur lequel les dieux avaient compté pour justement y mettre fin. Atteint d'une maladie incurable – son châtiment – il se souvient. Et, plutôt que de mettre la faute sur son ami, il se

remémore une femme – la femme qui l'a mis sur ce mauvais chemin de la civilisation : la courtisane Shamhat :

Dans sa colère, Enkidu maudit d'abord Shamhat : il souhaite qu'elle perde sécurité, honneur et protection sociale, qu'elle soit exposée à l'humiliation et qu'elle demeure sans foyer stable. La violence de cette malédiction montre combien il rend la femme responsable de son passage hors du monde animal.

Voici qui n'est pas mâcher ses mots, et voici qui nous replonge dans la même tragédie, dans le même désespoir – déjà bien trop connu – de la beauté accusée et maudite, misère des hommes.

5. Portes de sagesse

Pourtant… On ne s'arrête pas là. Enkidu a un sursaut de lucidité et il mourra en paix ; il semble, juste avant de rejoindre le royaume des morts, qu'il a compris un petit secret qui, jusqu'à nos jours, est resté trop discret :

Puis la malédiction s'inverse. Sous l'effet d'une parole divine, Enkidu bénit Shamhat : il souhaite qu'elle soit honorée, désirée, protégée par les puissants et entourée de richesses. Celle qu'il accusait de l'avoir arraché à sa première vie devient celle par qui il a accédé au monde humain.

Étrange revirement. La même femme, traitée auparavant comme une moins que rien, semble finalement appartenir à une élite de son genre, digne d'être célébrée à jamais dans l'histoire des hommes.

Enkidu, enfant de la nature qui a quitté son heureuse condition, maudit la responsable.

Enkidu, qui a connu la civilisation, l'amitié, tout ce qui fait d'un homme un homme et un homme conscient, bénit la responsable.

Entre les deux, il a gagné comme un supplément d'âme. Gilgamesh, après lui, connaîtra un tel regain de sagesse. L'épopée de Gilgamesh est véritablement un livre de sagesse et un récit initiatique.

Mais parlons davantage de cette initiatrice. Shamhat. Son nom est dérivé de la racine š-m-ḫ, qui évoque la « volupté », la « sensualité », la « plénitude ». On le traduit parfois par « la Luxuriante », « la Voluptueuse » ou « celle qui est comblée ». Voluptueuse, elle l'est ; prêtresse du temple – sans doute le temple d'Ishtar, la déesse de l'amour et de la guerre – elle est une sorte d'intermédiaire entre les mondes : l'humain et le divin, la sauvagerie et la civilisation.

Enkidu est d'abord un géant poilu qui ne sait pas parler et vit avec les bêtes, qu'il défend des chasseurs en détruisant leurs pièges. Ceux-ci se plaignent au roi, qui envoie alors une courtisane du temple pour le circonvenir et le civiliser : Shamhat est donc une sorte de piège, comme Pandore, un piège de luxe pour transformer un homme barbare en homme civilisé – ce qu'elle fera.

Elle s'offre à lui. Pendant six jours et sept nuits, ils s'unissent. Elle est femme, elle est porte sacrée, il l'a franchie et elle le guide encore. Plus loin, sur le chemin de l'âme : « Son cœur s'ouvrit, et il connut la sagesse. » Bien sûr, il a perdu quelque chose. Il court moins vite et les gazelles avec lesquelles il vivait s'enfuient maintenant à sa vue. Elle l'épile, lui donne de beaux vêtements, lui fait goûter au pain et à la bière. Elle lui parle longuement, lui dit ce qu'est Uruk et quel est son roi, Gilgamesh… Elle le conduira jusqu'à lui. Par les étreintes du corps d'abord, elle a fait naître une âme.

Peut-être, finalement, que les Grecs – quoique plus tard venus – en étaient restés à la malédiction des femmes, qu'ils n'avaient pas su ou pas voulu voir tout ce que la beauté, le luxe, la civilisation – leur avaient apporté. Peut-être que le luxe de la beauté fait trop mal quand on doit s'en passer ou qu'on a trop peur de le perdre. Confiner sa femme dans le gynécée[221], n'est-ce pas une façon de la garder pour soi, elle et sa beauté ?

Les femmes de l'aristocratie athénienne étaient voilées[222] : leur bouche, leurs cheveux, tout ce qui pourrait être appel à la volupté était caché. Il ne faut pas que les autres hommes les voient. La jalousie devient moralité ; la vertu d'une femme, c'est de se cacher, de mettre un voile pudique sur ses plus grands pouvoirs. Dans l'Orestie d'Eschyle, Clytemnestre est critiquée pour son audace… et elle apparaît non voilée !

Mais il faut le dire aussi honnêtement – il faut voiler les femmes, car les hommes, possiblement, sont des prédateurs. Plus grande est la beauté, plus tentante est la proie ! Mais quel désordre, en tous les cas. Il faut ranger les femmes. Si la femme n'est pas belle, on lui en veut. Et si elle est trop belle, on lui en veut aussi. Qu'elles s'occupent d'enfanter, c'est ce que doit faire une femme, après tout. Et si les hommes pouvaient s'entendre entre eux, ce serait tellement mieux !

Socrate, même Socrate, n'a pas fait mieux ! Lui dont les questions étaient comme la dynamite pour les préjugés, lui, l'avant-gardiste qui a si bien fait progresser les idées, n'avait pour les femmes que le regard blasé d'un Grec comme un autre. Diogène Laërce raconte des anecdotes navrantes sur la façon dont il traitait Xanthippe[223]. Elle n'avait peut-être pas toutes les vertus, mais, si ce qu'en dit Laërce est vrai, il ne semblait pas la considérer tout à fait comme un être pensant. Ne l'avait-il pas choisie ?

Xénophon, dans l'Économique, présente une division domestique traditionnelle : l'homme se charge surtout des affaires extérieures, tandis que la femme organise l'intérieur de la maison. Même dans ce cadre hiérarchique, la bonne administration du foyer est pensée comme un savoir, une forme de gouvernement.

S'il est facile d'idéaliser la vie des femmes grecques, c'est parce que nous avons tendance à confondre leur vie avec celle des déesses, qui étaient libres et s'exprimaient aussi haut que les hommes – avec parfois la même violence qu'eux. Les seules femmes qui recevaient une éducation intellectuelle digne de ce nom étaient presque exclusivement des hétaïres, des courtisanes de haut rang[225], ou des étrangères. C'est que les courtisanes de ce temps ne charmaient pas que par le corps, mais également par leur esprit et leur conversation.

Les courtisanes ? Cela ne nous ramène-t-il pas à Shamhat, la voluptueuse ? N'avons-nous pas noté que, dans les traductions de Bottéro, elle était appelée « prostituée sacrée » ? Les spécialistes sont très partagés sur cette appellation. Julia Assante, assyriologue spécialiste de la question[226], a montré que le mot ainsi traduit (ḳar.kid/ḫarimtu) désignait une catégorie sociale respectable, non une « prostituée » avec tout ce que nous lui accolons de mépris ou de déclassement. Cela aurait pu désigner simplement les femmes non mariées, libres, et pas seulement les prêtresses dévouées à un temple. Elles pouvaient être musiciennes, médecins, herboristes, tavernières, commerçantes, danseuses ou lettrées.

En vérité, il faut reconnaître que la seule expression « une femme libre » sonne encore curieusement à l'oreille d'un moderne. Il se mêle à cette liberté comme un soupçon d'immoralité. Alors comment voulez-vous que nous soyons conceptuellement

armés pour comprendre le rôle de Shamhat avec Enkidu, quand elle lui prêta son corps et tous ses raffinements pour le faire naître à l'humanité ? Nous avons fait de la sexualité une honte et un péché appelant malédictions, nous avons désappris à voir en elle comme un cadeau des dieux. Car l'érotisme est un des tout premiers pas qui mènent aux raffinements de la civilisation.

D'ailleurs, dans l'épopée de Gilgamesh, les femmes sont toujours des initiatrices, même quand le corps n'y est pas mêlé. Ce sont elles qui infusent la sagesse. Il faut dire quelques mots de la tavernière Siduri. Elle habite aux confins des terres humaines, près des eaux de la mort. Quand elle voit arriver Gilgamesh tout poussiéreux et dévasté dans l'ardeur de sa quête, elle hésite d'abord à lui ouvrir sa porte, puis se ravise. Elle le reçoit et l'oriente. Elle marque, elle aussi, un seuil pour l'esprit de son visiteur. En respectant son désir d'immortalité, en lui permettant même de poursuivre sa quête jusqu'au bout, elle l'enjoint doucement à prendre conscience des limites de ce qui est offert aux hommes, que cette limite n'est pas faite pour être franchie, mais, pour ainsi dire, habitée. Voilà donc une sagesse à hauteur d'homme :

Siduri lui rappelle alors la limite humaine : la vie sans fin qu'il cherche n'appartient pas aux mortels. Elle l'invite donc à habiter pleinement la vie finie — manger, se réjouir, se laver, se vêtir, aimer l'enfant et la compagne qui partagent son existence — plutôt qu'à consumer tout le présent dans la quête de l'immortalité.

Une morale épicurienne, finalement, avec un peu de douceur en plus. La douceur d'une tendre famille où l'on s'aime et se réjouit d'être ensemble. Simplement. Pas de quoi faire de grands discours philosophiques.

Non, finalement. Dans Gilgamesh, si elles ont bien quelque chose de divin, si leur beauté est comme un charme où peuvent tomber les hommes, les femmes ne sont pas pour autant comme la Pandore des portes de malédiction ; elles peuvent vous montrer le chemin de la sagesse, pour autant que l'on sache les écouter.

L'épopée a été écrite 1300 ans environ avant l'Iliade et l'Odyssée dans sa plus ancienne version, et 500 ans pour sa version la plus récente. S'il faut en croire Platon et décerner le titre de sagesse à ce qu'il y a de plus ancien, nous dirons que c'est là qu'il faut de préférence la chercher. Nous préciserons que, dans la chute biblique de Babylone, c'est un peu de cette sagesse aussi qui s'est perdue, même si elle s'était dévoyée.

Rappelons enfin aussi, et parce que c'est le lieu, que le premier écrivain de l'humanité qui nous soit parvenu est le nom d'une écrivaine : Enheduanna[227]. Une grande prêtresse et une poétesse, compositrice d'hymnes qui étaient accompagnés à la cithare. C'est donc aussi le premier nom d'artiste réelle qui nous soit parvenu.

4300 ans la séparent de nous.

Comme l'a écrit l'assyriologue William W. Hallo : « Enheduanna est la première personne dans l'histoire à avoir conscience d'elle-même comme auteur. » Pour dire la vérité, nous comptons encore un peu sur elle pour nous aider à prendre conscience de nous-mêmes.

6. Désirs d'alignement et anthropodicée

Le mot « alignement » a été introduit dans les débats sur l'IA par Eliezer Yudkowsky dans les années 2000[228]. Celui-ci écrivait sur la nécessité de concevoir des intelligences artificielles « amicales »

(Friendly AI), c'est-à-dire dont les objectifs seraient alignés sur le bien-être humain. À y regarder de près, on part d'une question technique : une machine qui doit remplir des objectifs techniques. La notion devrait donc d'abord relever du tamis de l'Utile. Évidemment, elles sont utiles, mais à quoi ? On est passé de systèmes experts spécialisés à des IA généralistes qui répondent à tous les besoins ou presque, notamment concernant la vérité (tamis du Vrai). Elles sont devenues des alliés très puissants quand on a besoin de renseignements précis, en phase de remplacer les moteurs de recherche classiques. Mais demander à une IA d'être friendly (amicale), c'est la placer comme un agent possiblement moral, c'est vouloir lui donner des valeurs – et de bonnes valeurs, car il y a quelque crainte à son endroit. Sans ce tamis du Bien, elle pourrait se désaligner à cause de son accaparement par un tamis unique. L'exemple de Bostrom (la fabrique de trombones) est celui d'une IA aveugle jusqu'à l'absurdité parce qu'elle poursuit un objectif unique, utile et à l'utilité en soi très limitée.

Si nous avons beaucoup parlé du problème de l'alignement dans le tamis du Beau, ce n'est pas seulement parce que les IA actuelles peuvent générer des contenus – texte, images, audio, vidéos, etc. – qui touchent à la beauté, mais surtout parce que « ce que poursuit une IA » relève en vérité d'un processus de réalisation du désir humain qui est à son origine. Ainsi avons-nous pu, incidemment mais consciemment, comparer ces machines parlantes à ces génies des contes de fée à qui l'on doit faire des vœux. Mais quand les vœux sont mal formulés, voire mal compris par celui qui les formule, le conte de fée tourne mal en général.

Ainsi nous est-il apparu que le véritable problème de l'alignement des IA n'était peut-être pas tellement celui des machines censées réaliser les désirs de l'humanité, mais celui de l'humanité

dont les désirs, mal compris encore plus que mal formulés, étaient le véritable danger. Il se trouvera peut-être des solutions techniques pour aligner les IA sur le désir humain, mais cela n'est pas complètement rassurant quand on regarde d'un peu plus près à quoi ressemblent les désirs humains (le désir est plus large que l'intention explicite, voire explicitable).

La beauté joue un rôle énorme dans le désir de tous les animaux supérieurs sexués ; Darwin lui-même parlait de la « coquetterie » des oiseaux ou de la « séduction » des fleurs[229]. Il y avait évidemment là un fond scientifique, et il savait pertinemment que ces mots eux-mêmes n'étaient que des métaphores qu'il appliquait au monde du vivant pour l'expliquer. La rose qui est sans pourquoi ne fait évidemment aucune métaphore et elle ne sait pas pourquoi elle est belle, même si sa beauté favorise sa reproduction. La métaphore ne vaut que pour des êtres de langage comme nous.

Le mot métaphore vient du grec metá et phérein ; cela signifie donc à l'origine « porter au-delà » : c'est le langage qui effectue des déplacements de sens de la beauté. Par de tels déplacements poétiques, qui sont les moteurs les plus primitifs et profonds du langage, notre sensibilité et notre désir de beauté dépassent très largement le simple désir d'un mammifère ordinaire. C'est évidemment par une sorte de métaphore première (ou de protométaphore) qu'il est possible de nommer arbre un arbre alors que nous avons nommé de ce même mot un autre arbre. Les mots sont plus rares que les choses qu'ils nomment, et ils ne peuvent les nommer que par déplacements de sens. Mais avec ce sens circule une énergie secrète, l'énergie du désir, qui est exacerbée dans les œuvres d'art où, par exemple, les femmes sont classiquement – et un peu sottement du point de vue d'un moderne – assimilées

à des « roses à peine écloses » quand elles sont dans leur prime jeunesse.

Ce « fonctionnement » humain, où le désir n'est pas fixé sur un seul objet mais peut emprunter les directions les plus inattendues, est évidemment problématique quand on désire savoir ce que les humains – au juste – désirent. Les artistes les plus contemporains nous révèlent et font prendre conscience des méandres les plus obscurs et parfois les plus odieux du désir humain. Il était plus facile à Leibniz d'écrire en 1710 une Théodicée qu'il ne le serait aujourd'hui d'écrire une Anthropodicée. Cette dernière phrase demande explication.

Leibniz, génie universel allemand, mathématicien inventeur du binaire a écrit en français un livre portant ce titre : Essais de Théodicée sur la bonté de Dieu, la liberté de l'homme et l'origine du mal. Philosophe rationaliste, il nous explique quelle fut la logique de Dieu lorsqu'il a créé le monde tel qu'il l'a créé. Cela est si bien expliqué et déduit qu'on en vient finalement à songer que Dieu n'aurait pas pu faire le monde mieux qu'il ne l'a fait. Sage et absolument raisonnable, il est donc innocent du mal. Le monde est donc en un sens parfait et exempt de tout mal. Si Dieu est innocent du mal, c'est de l'homme que vient sa calamité. Car l'homme, lui, n'est pas toujours raisonnable : il est – désaligné ! C'est nous, évidemment, qui employons ce dernier mot, et c'est à dessein que plusieurs fois déjà nous l'avons employé en ce sens métaphysique pour dire que l'homme a déçu des attentes, celles de Dieu, celles de la raison, etc. En fait, le désalignement est un autre nom du péché. En fait, le désalignement est un autre nom – de la liberté !

Pour fixer une définition, nous dirons ici que « Désaligner : cesser d'exiger que l'autre coïncide avec mon désir, et ouvrir un

espace d'affinité avec elle ou lui plutôt que d'identité. » C'est renoncer à la tentation de la contrainte et du contrôle.

Il est assez évident qu'en créant des IA, l'homme a joué à Dieu. Et il se retrouve avec des problèmes similaires. Il nous a paru assez pertinent de remonter la piste du désalignement humain pour penser le désalignement des IA. C'était suivre la piste du désir et de ce après quoi le désir court plus ou moins consciemment : la beauté. Or, le mot Anthropodicée que nous avons formé n'est plus une Théodicée, comme chez Leibniz – une tentative de justification de Dieu – mais une Anthropodicée, une tentative de justification de l'homme lui-même, peut-être dans son péché, son « désalignement ». Remarquons que le mot « dicée » vient du même mot grec qui signifie la déesse de la justice et du châtiment, dikē.

Or, il ne nous semble pas qu'il soit si facile de justifier l'homme dans son désalignement, pour la simple raison qu'il ne veut pas reconnaître ce désalignement et le met sur le dos de son autre, son alter ego, j'ai nommé – la femme ! C'est-à-dire que ce mammifère supérieur et parlant, ayant une vertigineuse capacité de déplacement de sens, fait retomber la faute entièrement sur son alter ego qu'il désire. La femme est coupable d'être belle ; c'est de sa faute à elle s'il devient criminel. C'est elle la désalignée qu'il s'agit de réaligner. Pas lui ! C'est Ève qui a fait pécher, c'est Pandore qui est coupable de toutes les calamités ; sa beauté divine est la source de tous les désordres. Des siècles et des millénaires de tradition pour faire dire à l'animal le plus intelligent de la terre et s'auto-justifier : « C'est pas moi, c'est l'autre ! »

Avec une lucidité pareille quant à son désir, avec cette façon d'assumer ses responsabilités, l'humanité est encore loin de la maturité. Désaligné de son propre désir (qui le programmait

seulement à la reproduction), il n'est pas près de s'en rendre maître. Devenant le maître de l'IA qui lui donne un pouvoir immense, il n'est pas près d'en finir avec les questions d'alignement, car il est difficile d'aligner quoi que ce soit sur un être désaligné.

Bien sûr, nous vivons à une époque où les femmes – si longtemps réduites au silence – sont en train de se libérer et de se « désaligner » du désir des hommes, et c'est heureux. Les philosophies de la « déconstruction » ont montré combien le langage, les institutions et les récits peuvent abriter des hiérarchies qui finissent par paraître naturelles. Il est donc légitime de déconstruire le patriarcat, c'est-à-dire d'interroger ce qui, pendant des siècles, a assigné aux femmes et aux hommes des places, des vertus, des vices et des destins prétendument conformes à leur nature.

Mais la déconstruction bien comprise ne saurait être la déconstruction des hommes par les femmes, comme s'il fallait reconstruire ceux-ci selon les desiderata de celles-là : ce serait simplement inverser une logique de domination sans en sortir. Ce qu'il s'agit d'abord de déconstruire, c'est une idéologie séculaire d'assignation des genres, reflétée et renforcée par les mythes, les religions, les institutions, les philosophies et mille pseudo-évidences sur la « nature de la nature ». De nombreuses penseuses ont précisément travaillé à mettre au jour cette fabrication historique et symbolique de la différence et de sa hiérarchisation[230]. Une supériorité masculine proclamée comme naturelle n'est pas entrée dans les têtes par la seule force des idées : la force et la violence physique ont largement contribué, au fil des siècles, à inscrire cette hiérarchie dans les corps, les institutions et les mœurs.

Cette déconstruction-là est non seulement légitime, mais nécessaire. Qu'on y prenne garde cependant : les idées justes et

subtiles, lorsqu'elles deviennent des mots d'ordre collectifs, perdent souvent en subtilité jusqu'à devenir parfois des caricatures d'elles-mêmes et les germes d'idéologies nouvelles. La cruelle ironie commencerait si l'on passait de la critique d'une assignation à une assignation inverse, si « l'homme déconstruit » signifiait l'homme corrigé selon ce qu'une femme, un parti, une époque ou une théorie estime qu'il devrait être. On ferait alors, à quelques millénaires de distance, aux hommes ce que les hommes ont fait aux femmes. Plutôt que de toujours accuser l'autre dans sa différence, il suffirait de nous imprégner d'un tout petit peu de la sagesse et de la profondeur d'un Bernanos lorsqu'il écrit : « les autres, hélas… c'est nous[231] ! »

La déconstruction n'est donc ici qu'un moment. Son terme souhaitable n'est pas la reconstruction de l'autre selon un nouveau modèle, mais son désalignement. Désaligner, c'est libérer : cesser d'exiger que l'autre coïncide avec notre désir. Si l'on voulait bien ne plus aligner Adam sur Ève, Ève sur Adam, ou Pandore sur un discours prétendument transcendant de ce qu'un homme ou une femme doit être, mais les faire entrer en dialogue, ce serait sans doute infiniment plus profitable pour tout le monde – et plus sage, au sens plein du terme. Il ne s'agit pas de nier la différence, mais de lui retirer son caractère d'assignation et de hiérarchie. Pour que l'on sache un jour s'il faut aligner l'IA et sur quoi précisément, il faudrait peut-être commencer par désaligner – c'est-à-dire libérer – les hommes et les femmes de la culpabilité et des rôles qu'ils font peser les uns sur les autres. Sinon, homo deus ou pas, à grands coups de technologies, de violences et de ruses, il y aura toujours des guerres de Troie et de belles tragédies.

À l'aube de pareilles questions, à un tournant du genre humain – car, comme nous l'avons déjà dit, les nouvelles technologies et

l'IA en particulier embarquent notre espèce sur des chemins inconnus – il est peut-être trop tôt pour proposer des solutions. Mais du moins pouvons-nous considérer avec intérêt le Manifeste cyborg de Donna Haraway, qui rejoint certaines de nos préoccupations. Haraway y condense sa provocation dans une formule célèbre : « Je préfère être un cyborg qu'une déesse. » Le reste de son argument peut se résumer ainsi : les identités politiques ne doivent pas être traitées comme des essences fixes ; le cyborg permet d'imaginer des coalitions fondées sur des affinités choisies plutôt que sur une identité supposée naturelle. Ce n'est plus l'origine qui décide seule de l'alliance, mais la relation que l'on choisit de construire.

Si l'on n'accepte pas que l'autre homme, l'autre genre soit désaligné de notre propre désir, l'IA ne sera jamais sympathiquement alignée sur les désirs de l'humanité pour la simple raison que ces désirs resteront toujours de domination et d'impérialisme. Sans l'acceptation du désalignement de l'autre, l'IA sera utilisée comme arme de guerre et de destruction. Mais il est vrai que l'homme a toujours mêlé à son goût de la beauté un certain goût pour la cruauté.

QUATRIÈME LIVRE
LE TAMIS DE L'UTILE

PARTIE I

ENTRE TAMIS ET MIROIRS, LA GRANDE TRANSFORMATION

Avec l'Utile, les trois premiers tamis reviennent sous une autre forme : celle des besoins, des mesures et des finalités. La machine n'est plus seulement productrice de vérité, d'autorité ou de beauté ; elle devient miroir des normes par lesquelles les humains se gouvernent. L'enquête passe ainsi de Simone Weil aux besoins de l'âme, puis des miroirs normatifs aux miroirs des princes et des peuples.

1. L'Utile, un tamis bien particulier

L'utile, c'est d'une façon générale ce qui satisfait un besoin, ce qui est profitable, avantageux.

Dans les quatre tamis que nous nous sommes proposés de mobiliser pour explorer le sens de l'IA, l'Utile tient une place particulière. C'est qu'à la différence du Vrai, du Bien, du Beau, il n'incarne pas une valeur idéale qui « donne du sens » à une vie humaine. Il ne peut que difficilement être agité comme un absolu. Il n'est jamais qu'un moyen et, comme tous les moyens, « relatif ».

Inversement, ni le Vrai, ni le Beau, ni le Bien ne peuvent s'exempter d'une pensée de l'Utile, car une fin qui ne se donne pas les moyens n'existe pas comme finalité. Si l'utile est satisfaction et profit, il est besoin de mobiliser l'Utile pour satisfaire nos besoins éthiques, esthétiques ou notre besoin de vérité. Un code de loi est Utile pour faire respecter la justice, le musicien ne peut se concevoir sans son instrument, et les appareils de mesure du laboratoire sont nécessaires et utiles à la recherche scientifique. Rien ne peut se faire sans les moyens pour le faire.

L'utile, donc, se pense primitivement dans une relation d'adéquation entre la fin poursuivie et les moyens mobilisés pour

l'atteindre. En cela, il est très proche de la technique et toutes les techniques n'ont de sens que par rapport à un projet.

Si l'on réduit le projet humain a minima : « mon projet, c'est de vivre et de continuer à vivre »… alors l'Utile prend une certaine autonomie de sens. Est utile ce qui répond à nos besoins vitaux. Mais, dans ce cas, la philosophie n'aura pas grand-chose à dire sur l'utile. Le vivant que nous sommes fera mieux de se concentrer sur sa subsistance, d'aller chercher de la nourriture, de la chaleur et du repos…

L'utile ne devient un objet pour la philosophie que quand il dépasse le simple besoin et le déborde dans le désir.

Le besoin c'est ce sans quoi il y a déchéance, abaissement de l'individu, non pas dans ses propres représentations ni celles des autres, mais un abaissement réel de sa santé et de sa spiritualité. Si l'on peut mesurer l'état de santé, c'est évidemment tout autre chose pour la spiritualité. Le désir quant à lui se projette au-delà de ce seuil. Il institue des fins au-delà du seuil du maintien, rehiérarchise l'Utile et redéfinit le sujet. De là de possibles déplacements de l'Utile. On peut désirer la vérité, puis lui préférer le confort, voire le repos. Et le désir d'un peuple ou de l'humanité peut varier également, un désir de guerre se transformer en répulsion pour elle quand on en a l'expérience. D'un désir à l'autre, on ne conçoit plus du tout ce qui est utile de la même manière. Ce qui était utile un jour devient inutile le lendemain, voire contre-productif. On jette les fusils dans un souci de désarmement.

On comprend alors qu'il y a pour l'Utile humain la possibilité d'une histoire et de généalogies. Ce qui est utile pour les animaux n'a pas cette variabilité. On peut dire que cette variabilité est idéologique pour autant qu'on parle de désirs d'importance.

À partir de là, vu depuis le désir, l'Utile devient un tamis à part entière ; c'est-à-dire qu'il opère dans le réel des choix qui ont du sens et qui déterminent en retour le sens de notre humanité. Affranchi des simples impératifs biologiques, l'Utile pourra alors naviguer selon sa logique propre qu'il conviendrait de décrypter : s'attacher au Vrai, au Bien ou au Beau, s'adjoindre à eux ou s'en différencier, voire les faire eux-mêmes changer de signification en modifiant, notamment en inversant les fins et les moyens.

Par exemple, je peux penser que l'argent est utile pour vivre, puis envisager comme utile tout ce qui me procure de l'argent, puis penser que la vie ne vaut que pour autant qu'elle me permet de faire de l'argent. On pourra penser qu'il y a là une sorte de forfaiture, de trahison, où l'on confond la fin et les moyens et tout ce que l'on voudra, mais on sait bien que ce tour est très habituel dans l'existence humaine. C'est que dans l'action, souvent, quand elle devient routine, on perd de vue ce qui donnait sens à l'action. La tenir en vue demande de la force et de l'énergie physique. Si l'on mobilise toute son énergie pour gagner de l'argent (ou pour conquérir le pouvoir politique, etc.), il ne reste plus assez de hauteur de vue pour fixer l'horizon que d'abord on s'était fixé. C'est un fourvoiement.

Ce même danger est déjà perceptible dans de nombreux modèles d'IA grand public génératrices de contenu. Elles proposent et cherchent activement à inciter l'interaction. En suggérant des questions ou des sources d'inspiration à ceux qui n'auraient rien à leur demander, aucun projet à servir ni aucun objectif clair où l'IA puisse être d'utilité, cette pratique est profondément néfaste. Elle crée l'illusion de la subjectivité chez l'utilisateur, qui croit être le maître de l'interaction alors qu'il n'est qu'un rouage inscrit dans la boucle de l'outil. L'utile de l'IA ne sert alors plus un désir

authentique, mais sa propre logique d'activation, accélérant ce fourvoiement où le moyen technique dicte sa loi à la fin humaine. C'est une sorte de désalignement organisé par les entreprises elles-mêmes parce qu'elles sont les rouages d'une plus grosse machine que la leur : la machine économique.

Poussée à l'extrême, cette fausse bonne idée rappelle un roman de Philip K. Dick paru en 1955 : Autofac[232]. Dans un univers post-apocalyptique, un réseau d'usines automatisées (les « autofacs ») continue inlassablement à produire et à livrer des biens de consommation aux survivants humains. Ceux-ci n'en ont plus aucun usage, mais les usines sont programmées à produire (épuisant les dernières ressources) et continuent à traiter les humains en consommateurs sans qu'ils en aient le moindre usage. Les protagonistes, focalisés sur leur survie face à ce système incontrôlable, tentent désespérément de le saboter pour reprendre le contrôle, mais les autofacs sont auto-réplicants et se défendent avec leur redoutable efficacité technologique. Voilà qui illustre parfaitement le divorce entre l'utile et l'efficace.

On pourra incidemment nous reprocher la naïveté de notre distinction entre désir et besoin, mais cette tension, nous l'assumons et nous la désirons, elle est la condition du sens. Cet écart entre les deux est la condition de la pensée et de la liberté. C'est dans cet écart qu'il y a la possibilité d'une critique. Il est vrai que l'ascèse autrefois nous apprenait à distinguer ce qui relève du désir et du besoin et que la pratique en est largement oubliée. Mais il y a des pensées qui se ressentent de leur embonpoint. Sans viser personne, qui ne sait plus faire la différence entre désirs et besoins n'a jamais connu le besoin, la faim.

C'est d'ailleurs un des dangers de l'IA si elle parvenait à assouvir tous les désirs humains que de le laisser seul avec lui-même

vivant son immanence comme une absurdité lui ôtant tout appétit et tout désir de vivre. Il n'y a au fond de sens que s'il y a besoin d'aller quelque part. Mais ces désirs et besoins peuvent-ils être totalement comblés par les IA ?

2. L'IA face aux « besoins de l'âme »

Dans L'Enracinement, Simone Weil reconnaît qu'il y a des « besoins de l'âme » qui ne sont donc pas biologiques, mais d'un autre ordre, supérieur[233]. Parmi ces besoins, il y a l'enracinement et la vérité. L'enracinement est pour l'âme un besoin vital. Il s'agit d'une appartenance, non pas un simple sentiment d'appartenance, mais d'un engagement envers la collectivité humaine. C'est l'état où l'individu reçoit de son milieu toute sa nourriture morale et spirituelle, se rattachant aux trésors du passé et aux aspirations de l'avenir. Sans ces racines multiples – familiales, locales, professionnelles – l'être sombre dans le déracinement, une misère spirituelle qui le rend vulnérable à la violence et aux tyrannies. L'enracinement est la condition nécessaire pour que l'homme soit véritablement un être de devoirs envers ses semblables.

Ce besoin d'enracinement ne pourrait être comblé par aucune IA. Si l'IA peut nous rapprocher en un sens de la collectivité humaine parce qu'elle est en quelque sorte porteuse de la mémoire de l'humanité (nous y reviendrons), elle peut d'abord et surtout déraciner. En proposant des dialogues, où il n'y a finalement pas d'alter ego, elle peut nous couper de ces racines qui sont au fond les autres hommes, nos alter ego, les autres « âmes », quel que soit le sens que l'on puisse donner à ce mot.

Soulignons que ce besoin est fondamentalement lié à la terre : c'est au sein de la campagne et du paysage que l'individu trouve

une participation active. Ce lien concret permet de recevoir la nourriture morale, de conserver les trésors du passé et de forger une identité. Le déracinement, qui arrache l'homme à sa terre, est l'origine des misères spirituelles et sociales. On comprend bien qu'il y a pour elle dans l'urbanisation comme un déracinement, et que l'exploration de mondes virtuels comme ceux que nous connaissons aujourd'hui peut être vécue pour beaucoup non comme des perspectives fascinantes mais comme de purs et simples déracinements vécus douloureusement.

Il est fort probable que l'inquiétude d'Hannah Arendt voyant l'humanité s'enthousiasmer pour l'envol de Spoutnik dont nous avons parlé en première partie relève d'une intuition similaire : ces gens qui s'enthousiasment ne savent pas ce qu'ils font. Ils étaient attachés à la planète terre et voient comme une libération l'acte de s'en éloigner. Mais nous avons besoin de racines et rompre ainsi les amarres pour gagner le ciel pourrait être se priver des plus fondamentaux des besoins. Pour Simone Weil, nous l'avons dit. Vouloir s'élever, c'est quelquefois s'abaisser. La logique du levier est inverse – et on se lève en s'abaissant. Un retour à la terre serait, sans doute, pour beaucoup d'êtres humainss une élévation spirituelle.

Les autres besoins de l'âme se pensent par couples d'opposés :

Ordre et Liberté

Obéissance et Responsabilité

Égalité et Hiérarchie

Honneur et Châtiment

Sécurité et Risque

Propriété privée et Propriété collective

Chaque couple représente deux exigences fondamentales dont l'une est le poison de l'autre lorsqu'elle est absolue. L'équilibre

n'est pas une moyenne statique, donc il échapperait aux calculs des IA si puissantes soient-elles, mais il est une harmonie vivante et dynamique qui permet à l'âme de s'étendre dans des directions opposées sans se briser, évitant ainsi la maladie spirituelle (servitude, ennui ou anarchie).

Simone Weil se méfie du simple « juste milieu » lorsqu'il consiste à affaiblir deux besoins opposés au lieu de chercher une forme d'équilibre qui leur rende justice à chacun. Cette idée est précieuse ici : l'équilibre n'est pas nécessairement une moyenne, mais une composition.

La pensée de Simone Weil est profondément compatible avec la logique du pharmakon développée par Bernard Stiegler[234]. Le pharmakon a donné notre l'appellation de nos modernes pharmacies. En grec, il désigne à la fois le remède et le poison, et Platon l'utilisait déjà dans le Phèdre pour parler de la technique, à la fois remède et poison pour la mémoire des hommes. Elle est remède, car elle est « pense-bête » possible pour les oublieux, mais elle est poison, car justement, sa sécurité nous livre au laisser-aller et à l'oubli.

Bernard Stiegler a beaucoup insisté sur l'importance de l'attention chez Simone Weil. Il y voit une ressource pour penser la technique comme un pharmakon : la même puissance peut empoisonner ou soigner selon les formes d'attention, de transmission et de pratique collective qui l'accompagnent.

Une perte de psyché, c'est à bien l'entendre – et dans le vocabulaire chrétien – une perte d'âme et nous aurons à y revenir. En attendant, ne perdons pas le fil des besoins de l'âme. L'IA peut nous aider sans doute à concilier les couples contradictoires comme « Ordre et Liberté », mais rien sans doute ne serait plus catastrophique que de s'en décharger sur elle. Il en va de notre

spiritualité. C'est que tous les besoins de l'âme, et par définition, ne sont pas des problèmes, mais, selon le vocabulaire que nous avons emprunté à Gabriel Marcel, des « métaproblèmes[235] ». Ils sont porteurs d'un mystère qui est le nôtre et qui ne se résout pas, qui s'expérimente et qui se vit au jour le jour et à chaque instant comme toujours différent et à dénouer autrement quand il ne s'agit pas simplement de le laisser exister – tel quel.

3. Tamis, miroirs et médiations

On nous aura compris. Les questions de type « à quoi pourraient bien servir ou être utiles les IA » en un sens « ordinaire » ne sont pas en elles-mêmes et directement philosophiques. Ce sont plutôt des questions qui sont d'ordre technique, sur le modèle : « compte tenu des performances actuelles des IA, en quoi peuvent-elles satisfaire les besoins de l'humanité tels que nous les connaissons ? » Les premières réponses seront de l'ordre du biologique (médecine, nutrition, etc.). Ensuite, on pourra étendre la focale jusqu'à l'économique et le sociologique : peut-elle proposer de nouveaux modèles d'organisation socio-économiques plus efficaces, engendrant moins d'inégalités et de misère si c'est vraiment ce que l'on veut… On pourra même proposer aux IA de résoudre des problèmes écologiques alors même qu'elles contribuent au problème, etc. Mais, pour ce faire, personne n'a besoin de nous. Les problèmes sont là, et certains depuis très longtemps, des IA sont là qui ne demandent qu'à progresser. Il n'y a qu'à organiser les rencontres des deux.

Ce qui est davantage de notre compétence, et ce que nous nous efforçons de faire depuis le début, c'est de se demander si le « sens » de la notion d'humanité ne peut pas être considérablement modifié par la rencontre avec l'IA. Inversement, la

rencontre avec l'IA peut nous aider à définir ce que c'est véritablement qu'être humain et à éloigner de nous ces tâtonnements aveugles qui font que l'humanité, parfois, s'est égarée sur son propre chemin, au prix de crises et de souffrances inutiles.

Or, l'une des thèses les plus importantes et les plus paradoxales de ce livre est que, malgré sa nouveauté absolue dans l'histoire de l'humanité, l'IA obéit à un processus qui fut toujours présent dans l'histoire de l'humanité : celui de son dédoublement culturel et technologique qui transforme l'humanité en retour. L'homme projette ses valeurs comme le Vrai, le Beau et le Bien ailleurs qu'en lui-même, par exemple dans un « arrière-monde » comme l'Olympe, celui-ci s'enrichit considérablement avec le temps et, en prenant consistance, impose des normes à l'humanité qui sont opérantes, efficaces et « utiles » en ce sens qu'elles orientent les pensées, les paroles et les actes des groupes comme des individus. L'un des avantages de la méthode des tamis, c'est qu'elle a quelque chose de « métaproblématique » au sens identifié plus haut avec l'aide de Gabriel Marcel. En nous interrogeant sur le sens de l'objet étudié (ici les intelligences artificielles), nous nous interrogeons sur le sens des valeurs qui nous permettent de les passer au crible (ici les tamis du Vrai, du Beau, du Bien, de l'Utile). Ainsi ces valeurs, qui cessent d'être des idéaux éternels et immuables au sens platonicien, se révèlent à nous comme ayant une histoire, une généalogie. Leur aventure même et leurs transformations sont les aventures du sujet humain occidental dans ce qu'il désire, croit être et vouloir.

Nous avons montré que les idéaux du Vrai, du Beau, du Bien évoluent et que leur évolution même transforme profondément notre rapport au monde, notre regard sur le passé et l'avenir de notre espèce qui a cette caractéristique de toujours vouloir se

dépasser. Pour cela, elle crée des images du monde et d'elle-même, elle se duplique en représentations et ses représentations qu'elle se crée deviennent, à leur tour, moteurs de sa transformation. Par exemple (c'est un peu plus qu'un exemple en vérité) les mortels que nous sommes se dupliquent en imaginant des immortels qui ont une organisation politique qui ressemble à la leur mais en plus vaste et plus parfaite.

Cette duplication de soi et du monde dans un monde parallèle peuplé d'immortels est un événement majeur dans l'histoire de l'humanité. Plus précisément même, le processus de duplication de soi et du monde dans des représentations qui nous sont extérieures est à terme un acte fondateur qui appelle l'humain à des transformations majeures d'ordre anthropologique. C'est en se dupliquant ainsi que l'homme est sorti du pur ordre biologique pour devenir un être historique. Toute répétition d'une telle duplication, dans d'autres projections du monde et de l'humanité, sont historiquement des révolutions idéologiques, qui peuvent précéder et provoquer des événements politiques, les accompagner ou même s'ajuster à une révolution passée pour la justifier comme après coup et la consolider.

Si nous avons raison, alors la duplication de l'humanité dans des sujets parlants modelés à son image (intelligences artificielles symboliques) capables de lui parler annonce une révolution anthropologique d'un type nouveau. S'imaginer des dieux qui vivent leur vie en miroir de la nôtre modifie nos pensées, nos paroles, nos actions. Coucher dans des manuscrits leurs propos, les représenter avec nos arts permet de les graver dans nos cœurs et nos esprits ; on essaye d'adapter nos comportements à la rigueur de leurs attendus. Mais se préparer à vivre avec des systèmes humanoïdes à notre semblance de corps comme d'esprit, dialoguer

tous les jours avec eux, les voir évoluer en même temps que nous, mais peut-être plus vite que nous s'adapter et répondre à nos propres évolutions, voilà tout autre chose.

Nos représentations de nous-mêmes, nos doubles, ne sont pas de simples images ni de simples fantômes. Elles agissent : elles orientent les conduites, les attentes et parfois les institutions. Elles ont donc une dimension politique et, au sens large, performative[236]. Mais il faut ici arrêter une extension possible de notre métaphore : si toute chose qui agit sur nous devient un tamis, le mot finit par ne plus rien distinguer.

Réservons donc le mot « tamis » à ce qui opère véritablement un tri normatif. Le Vrai, le Bien, le Beau et l'Utile sont des tamis parce qu'ils permettent de sélectionner, hiérarchiser, retenir ou rejeter. Le tamis n'est pas l'objet que nous regardons : il est le critère avec lequel nous jugeons. L'image évangélique du bon grain séparé de l'ivraie en donne presque littéralement la figure[239].

Le miroir est autre chose. Il ne trie pas nécessairement : il montre, reflète, déforme parfois, et nous renvoie une image avec laquelle nous pouvons nous comparer. Les objets, les récits, les institutions et les techniques sont encore autre chose : ils peuvent servir de supports, de médiations ou de scènes où les tamis et les miroirs opèrent[237]. Une cathédrale n'est donc pas un tamis : elle donne une forme sensible à une certaine articulation du Vrai, du Beau, du Bien et de l'Utile. Une publicité n'est pas davantage un tamis ; elle cherche à faire passer son produit à travers les nôtres, et parfois à en modifier la maille.

De tels objets peuvent évidemment être chargés de récits et d'idéologie. Barthes a montré dans Mythologies comment des objets et des pratiques ordinaires deviennent porteurs de significations qui finissent par paraître naturelles[238]. Mais être porteur

d'une « mythologie » ne suffit pas à devenir un tamis. Cela signifie seulement qu'un objet peut contribuer à former, déplacer ou renforcer les critères avec lesquels nous tamisons le monde.

L'IA est une machine à tamiser l'homme, construite par l'homme, et ce qu'elle retient dans ses mailles est de ce métal dont on fait les miroirs. Dans ce processus inédit, l'homme n'a pas intérêt à « laisser la main » : sinon, il sera ce que la machine fera de lui. Il se doit, au contraire, d'être toujours plus conscient et de garder la main, car nous pensons aussi avec les mains. Je veux dire qu'une pensée des tamis est une pensée de l'action minutieuse et prudente, qui doit laisser dans la boue les idéologies.

4. L'IA comme un miroir puissant

Le lecteur aura sans doute compris qu'il n'y a pas plus puissant, plus flagrant miroir de notre humanité que les IA qui, par leurs données massives et leurs capacités à les traiter, en savent en un sens (un sens non humain) infiniment plus que n'importe quel individu, fût-il le plus instruit, n'en pourra jamais savoir. Rappelons ici que par « intelligence artificielle », nous entendons ici les grands modèles de langage actuellement accessibles au grand public – systèmes capables de générer du texte, mais aussi, dans une moindre mesure, des images, de la musique ou des vidéos. Ces artefacts ne possèdent ni conscience, ni volonté, ni compréhension au sens humain ; ils fonctionnent par analyse statistique de vastes corpus de données produites par les humains. C'est précisément cette caractéristique – être un miroir statistique de notre langage et de nos productions culturelles – qui en fait un révélateur si puissant de nos désirs, de nos biais et de nos contradictions.

Parmi les ouvrages récents les plus stimulants sur cette question, il faut sans doute citer celui de Shannon Vallor, The AI Mirror : How to Reclaim Our Humanity in an Age of Machine Thinking, paru en 2024 et non encore publié en français[240], dont le titre pourrait être : « Le Miroir de l'IA : Comment Reconquérir Notre Humanité à l'Ère de la Pensée Machine ».

L'idée de Shannon Vallor est que l'intelligence artificielle, loin d'être une intelligence extraterrestre ou un dépassement de l'humain, agit comme une surface réfléchissante nous renvoyant nos propres données, si l'on peut dire, pour le meilleur et pour le pire. Mais même le pire n'est peut-être pas si mal pour autant qu'on en prend conscience ; on pourrait le corriger. En effet, l'IA commence par révéler nos failles.

En effet, entraînée sur l'océan des données produites par les humains (textes, images, décisions, etc.), elle reflète et amplifie nos propres biais, nos erreurs, nos stéréotypes et nos lacunes morales.

Il faut reconnaître aussi que, paradoxalement, l'IA qui nous semble une technologie de l'avenir est fondamentalement un reflet du passé. Elle ne s'est jamais nourrie que de ce qui a été dit, écrit, vécu, imaginé, conservé en mémoire par des documents et enregistrements.

Ce miroir appelle à nous corriger nous-mêmes. En voyant nos propres défauts reflétés de manière si claire et puissante, l'IA nous oblige à nous interroger sur ce que signifie être humain et quelles sont les valeurs que nous voulons réellement cultiver pour l'avenir.

Pour Vallor comme pour nous, le véritable enjeu n'est pas de craindre l'IA, de se demander si elle est bonne ou mauvaise, si on l'aime ou ne l'aime pas, mais de regarder honnêtement ce que le

miroir nous renvoie. N'est-ce pas une manière de reconquérir notre humanité et de nous réinventer ?

L'homme n'est rien de défini et surtout rien de définitif. Il ne cesse de se réinventer au fur et à mesure, d'ailleurs, qu'il tente de se définir… Il a pu dire en 1888 que l'homme est un animal dont le caractère propre n'est pas encore fixé[241]. Et pour cause ! Nous sommes bien le seul animal qui s'auto-apprivoise ! Très simplement, en nous projetant hors de nous-mêmes, nous nous forgeons des maîtres qui nous éduquent. Par exemple des dieux.

Prenons, dans le monothéisme chrétien, l'exemple de la Confession chez Saint-Augustin. C'est le titre de l'un de ses livres les plus fameux. Qu'y fait-il ? Il s'y confesse. À quoi ? À Dieu, à nous et peut-être à lui-même.

Augustin ouvre les Confessions par une parole adressée à Dieu : il veut être connu par celui qu'il reconnaît déjà comme son plus intime connaisseur, et demande que son âme soit transformée par cette connaissance. L'écriture autobiographique devient ainsi moins une exhibition de soi qu'un examen de soi devant une transcendance[242].

Ce mouvement est douloureux : l'auteur dit son dégoût de lui-même, son désir d'être attiré vers Dieu et la honte qui accompagne cette conversion du regard. La vérité sur soi n'est pas seulement information, mais transformation morale et spirituelle.

Saint Paul, qui, dans la Bible, sur le chemin de Damas, perd provisoirement la vue, sombre, en fait, dans le chaos existentiel et le désarroi de celui qui se révèle à lui-même tel qu'il est, en tombe malade et ne recouvre la santé qu'en se transformant en « homme nouveau[243] ».

Dans l'allégorie de la caverne, Platon décrit pareillement la vérité comme une conversion pénible : celui qu'on force à se

tourner vers la lumière souffre d'abord de l'éblouissement et préfère les objets familiers qu'il voyait auparavant. La connaissance n'est donc pas un simple ajout d'informations, mais un changement du regard et de l'habitude[244].

Mais une fois que le philosophe a été « éclairé », il ne sera plus jamais le même.

Si nous voulions être provoquants, nous dirions que la relation sexuelle qu'a eue Enkidu avec Shamhat, la « prostituée sacrée », a eu pour lui valeur de révélation et de conversion. C'est là qu'il a acquis conscience humaine, qu'il s'est envisagé lui-même et travaillé culturellement, il s'est épilé avec l'aide de son initiatrice et a mis des habits.

La piste de la honte et de la pudeur nous invite tout naturellement à évoquer la prise de conscience (non dans la confession mais dans l'horreur du crime lui-même) d'Adam et Ève. Le péché et la conscience du péché qui l'accompagne les ont changés à tout jamais, eux et leur lignée, c'est-à-dire, selon les croyances chrétiennes, l'humanité tout entière avec elle, les filles et fils d'Ève et d'Adam !

Le regard de Dieu qui est en nous agit comme un miroir et ce miroir, en nous aussi bien qu'hors de nous, est agissant ; il nous trouble, nous fait mal et nous éblouit et il nous change à tout jamais – enfin dans la bonne version. Car dans la mauvaise version, par exemple dans l'histoire de Caïn, le premier assassin de la Bible, l'assassin de son double de frère, la conscience morale est bien là mais ne le convertit pas. Il s'obstine et fuit l'œil-miroir de Dieu, car cet œil qui finalement le juge, le regarde et le maudit, au fond – est bien en lui : « L'œil était dans la tombe et regardait Caïn » (Hugo[245]).

Evidemment, les conversions radicales et pour ainsi dire spectaculaires ne sont que l'exception. En temps ordinaire, une conversion du regard chez un individu peut prendre des années et dans les populations… des dizaines, voire des centaines d'années !

Mais revenons à plus facile, revenons à la confession. Si chrétien pouvait ressortir du confessionnal complètement changé, un homme nouveau et bien meilleur comme Saint Paul, il en serait heureux. Le travail en général est long et laborieux pour se purifier véritablement du mal et se transformer spirituellement. C'est en quelque sorte un travail de « tamisage ». Ce qui passe par la grille est pour ainsi dire filtré. Il s'agit, tout en gardant le précieux, d'un déversoir pour l'âme où, dans une opération en vérité fort délicate. Il faut, avec lucidité, expurger le mal, le discerner, le distinguer du plus précieux, de se débarrasser de l'un et garder l'autre. Ce n'est pas seulement discernement à la lumière ; c'est également action par le tri. Et ce tri est une sorte de conversion (nous ne parlons pas encore d'une conversion en argent, cela viendra !).

5. Miroirs à grande échelle

Évidemment, ces histoires de duplication du sujet et du monde, de tamis-miroir au grillage fin qui permettent de discerner – hors de nous et dans nos âmes – avec une lucidité nouvelle le Bien, le Vrai, le Beau. Ces tamis agissants et donc utiles, qui permettent de faire le tri avant de nous transformer, ne valent pas seulement pour les individus. Ils ont une amplitude sociétale et rien n'empêche de les considérer comme valant également à l'échelle de l'humanité.

Pour nous faire bien comprendre, pour montrer en quoi les grands idéaux du Vrai, du Beau, du Bien sont toujours et en même temps actifs, utiles, opérants, prenons un exemple. Encore une fois, remontons à un état de relative naïveté pour en montrer le déploiement et le processus.

Chacun a entendu parler de Zeus et ses deux frères : Hadès et Poséidon. Ce trio divin qui nous semble, de loin, comme presque anecdotique, une jolie invention de poètes. Mais ces personnages, en fait, ne sont pas de pures inventions. C'est simplement la projection d'une fratrie comme il y en a toujours eu sur terre, la projection d'une situation bien connue des notaires à propos d'un héritage après la disparition du père, en l'occurrence Kronos.

Cette histoire reflétait, lorsque le mythe était vivant, une organisation politique et sociale très humaine : une répartition du pouvoir entre frères après une victoire commune, une hiérarchie implicite, et une division territoriale qui rappelle les structures de gouvernance, en particulier celles des cités-États grecques. Cette répartition n'est pas fondée sur une logique cosmique abstraite, mais sur un modèle politique humain : celui de la succession dynastique et du partage du territoire entre héritiers. Dans les sociétés antiques, notamment en Grèce archaïque, il était courant que des frères se partagent un royaume ou des domaines après la mort ou la défaite du père. Le mythe transpose donc une pratique humaine dans le monde divin.

Mais ce miroir divin ne se contente pas de refléter la société grecque. Une fois bien établi, il devient à son tour modèle et façonne en retour la société dont il n'était pourtant d'abord qu'un reflet. Il s'agit plus ou moins d'un mouvement dialectique où tout fonctionne par imitation : la société projette ses structures sur les dieux, puis ces dieux, à travers les mythes, les cultes, deviennent

normes religieuses, légitiment, renforcent et orientent les institutions humaines. Ce phénomène est central dans les sociétés anciennes comme dans les sociétés modernes.

À partir du moment où l'immortalité est, pour ainsi dire, gardée en vue par ces religions, elle devient objet de désir et inspire. Le héros qui se procure par son action mémorable une sorte d'immortalité à hauteur d'homme devient modèle de l'action humaine. Comme on raconte l'histoire des dieux, on racontera l'histoire édifiante de tel ou tel qui a su se conduire de façon admirable. On célèbre encore aujourd'hui l'exploit légendaire de Phidippidès, qui aurait couru de Marathon à Athènes pour annoncer la victoire – ce qui est factuellement faux si l'on s'en tient au récit d'Hérodote[246].

Si penser, c'est agir, c'est parce que penser est agir sur les représentations idéales et les valeurs de l'humanité. La critique d'un récit mythique (ou projeté dans le monde du langage et des représentations) produit, quand elle est d'importance, et en retour, une nouvelle façon de penser et de vivre son humanité. Par exemple, quand Platon critique les dieux tels que représentés par Homère en ne voyant pas dans leur comportement quelque chose qui soit digne d'imitation[247] (Zeus est tout sauf platonique), il transforme le sujet humain en retour qui n'aura plus le même idéal. Concrètement, Platon nous achemine vers la voie d'un monothéisme (il écrit souvent « le dieu »), d'une morale plus rigoureuse et d'un certain ascétisme.

Difficile de ne pas voir que l'ordonnance hiérarchisée verticalement du « royaume de Dieu », tel que pensé par les chrétiens, n'ordonne pas en retour la société médiévale selon une hiérarchie considérée comme éternelle, exemplaire, légitime, etc.

L'imaginaire des chansons de geste, des chevaliers de la Table ronde propose des récits d'ordonnancement de la société.

La morale elle-même repose sur de pareils récits qui, s'ils prennent beaucoup de licence avec la vérité au sens où nous l'entendons, deviennent pour l'époque les tamis du Beau et du Bien dans une concordance heureuse. En célébrant les beaux affrontements entre le chevalier blanc et le chevalier noir, on célèbre un ordre du monde et une morale. Voici ce qui est exemplaire et voici ce qui ne l'est pas.

Dans les textes de Ramon Llull consacrés à la chevalerie, le chevalier est notamment chargé de défendre ceux qui ne peuvent se défendre seuls — veuves, orphelins, faibles — et de mettre sa force au service d'un ordre moral.

Quand les tamis du Vrai, du Beau, du Bien et de l'Utile coïncident, cela signifie en quelque sorte que la société est alignée et se porte bien. Quand il y a disjonction, il y a perte, au moins provisoire, d'un équilibre qu'il s'agit de rétablir.

Lors de la Révolution française, les bouleversements politiques se sont précipités et la religion a perdu de son autorité. Le bien n'était, pour ainsi dire, plus vertical ; on voulait de l'horizontale, une société d'égaux. La révolution newtonienne n'avait-elle pas montré que tous les corps étaient soumis de la même façon à la loi du nombre ? Il n'y a plus de titre nobiliaire, une tête en vaut une autre ; Guillotin, à sa façon, égalisera concrètement et proposera une invention dont on ne mesure pas l'importance : la pétition.

L'abbé Sieyès avait perçu aussi le besoin d'un nouveau récit national ; on ne pouvait pas continuer à célébrer les chevaliers de la Table ronde dans une société d'égaux. Il va reprendre le vieux truc de l'oubli d'un quelque chose pour un retour au plus

fondamental. En fait, la France éternelle ne date pas des Francs, de Clovis et de Charlemagne : « Le Tiers ne doit pas craindre de remonter dans les temps passés. Il se reportera à l'année qui a précédé la conquête (…) Pourquoi ne renverrait-il pas dans les forêts de la Franconie toutes ces familles qui conservent la folle prétention d'être issues de la race des conquérants et d'avoir succédé à des droits de conquête ? La nation, alors épurée, pourra se consoler, je pense, d'être réduite à ne se plus croire composée que des descendants des Gaulois et des Romains. » (Qu'est-ce que le Tiers État ?)

Les révolutionnaires français étaient encore très admirateurs de la république romaine, mais c'est au XIX^e^ siècle, avec le romantisme, le nationalisme et le développement de l'archéologie, que les Gaulois deviennent un objet de fascination. En 1828 est publiée l'Histoire des Gaulois où Vercingétorix est présenté comme le héros de la liberté nationale. Jules Michelet, dans son Jules Michelet, dans son Histoire de France (1833 – 1867), idéalise les Gaulois comme un peuple libre, poétique, proche de la nature, en opposition à la rigueur romaine. C'est sous la IIIe République que le mythe du Français gaulois arrive à son apogée et que l'on demande à tous les petits Français, à travers les programmes scolaires, de s'identifier comme Gaulois, quelle que soit leur origine. C'est la formule célèbre de Lavisse, efficace avant d'être ridicule puis contre-productive dans les colonies : « Nos ancêtres les Gaulois[249]… » La « vraie » France, c'est Vercingétorix. Beaucoup de Français aiment à se représenter comme des Gaulois, peuple railleur et batailleur, esthétique à moustache qu'on se plaît à retrouver chez Astérix avec un soupçon d'humour, de démesure à la façon d'un Rabelais.

6. Miroirs des princes, miroirs des peuples

Les régimes totalitaires, avec plus ou moins de subtilité, manient clairement les tamis comme des instruments de propagande. Le Vrai est ce que le régime dit être Vrai, d'ailleurs la science le prouve, quitte à la déformer ; l'histoire prend les couleurs d'une nouvelle esthétique, les intellectuels et les artistes seront alignés, déportés ou détruits, les religions seront interdites ou persécutées.

Le philosophe de l'Utile par excellence, c'est Machiavel, car il érige (au moins dans Le Prince) l'Utile en tant qu'efficacité dans la prise du pouvoir en impératif absolu de sagesse. Au nom de l'efficacité, on peut donc désobéir à tous les commandements religieux, mentir, trahir, assassiner. Il suffit de sauver les apparences pour être l'homme de la situation. On a souvent dit que les véritables intentions de Machiavel n'étaient pas aussi cyniques que cela ; c'est sans doute vrai, mais ce n'est pas notre propos. Ce que nous retiendrons de lui, c'est son volontarisme. Il ne se laisse pas aller à la bonne Fortune et ne s'en remet pas à la Providence.

Qu'est-ce qui existait avant Machiavel comme livres de formation pour les politiques ou futurs gouvernants ? Ce qu'on appelait des « Miroirs des princes[250] ».

Un miroir des princes (en latin : speculum principis) est un genre littéraire très répandu au Moyen Âge et à la Renaissance, qui consiste en un traité de conseil politique et moral adressé à un souverain ou à un futur dirigeant. Son objectif est de guider le prince dans l'exercice du pouvoir en lui offrant des modèles de sagesse et de vertu. Qu'on nous pardonne, à ce stade, d'être un peu lapidaires ; on y apprenait les valeurs du Vrai, du Bien, du Beau et de l'Utile. Certes, peut-être pas dans cet ordre-là, mais le Beau – par exemple – n'était pas à négliger. Un prince doit avoir

du goût et surtout le sens des cérémonies et de l'étiquette. « À la cour, tout est cérémonie, tout est étiquette », dira le duc de Saint-Simon pour qui le cérémonial est le langage même du pouvoir royal[251].

Machiavel reproche aux penseurs politiques d'imaginer des régimes qui n'ont jamais existé et de confondre la manière dont les hommes devraient vivre avec la manière dont ils vivent réellement : pour conserver le pouvoir, le prince doit partir du réel, non d'un modèle moral abstrait.

C'en est donc fini de celui des miroirs imaginaires à l'usage du Prince, des Utopies à la manière de Thomas More que l'on trouve déjà chez Érasme, Gilles de Rome, Salisbury, Xénophon et Platon ! Il faut un vrai miroir qui, peut-être, reflètera des choses un peu moins jolies, mais qui soit vraiment efficace à celui qui voudrait changer les choses.

On trouvera quelque chose de similaire chez Marx. Il y a toujours eu quelque chose de « spéculatif » dans la philosophie : le mot « spéculatif » vient du latin speculativus, dérivé de speculari (« observer, espionner »), lui-même issu de speculum, qui signifie « miroir » ou « instrument d'observation » (comme un miroir utilisé pour voir au-delà d'un obstacle). Chez Platon comme chez Aristote, la philosophie est contemplation. Leibniz le dira plus crûment : « L'œil est l'organe propre de la philosophie. » (Nouveaux essais sur l'entendement humain, IV, chap. XI, § 8) Comme si les philosophes ne faisaient que « voir » ! Comme s'ils ne faisaient pas eux-mêmes quelques ajustements de lumières, de point de vue, de hauteur, comme s'ils ne montraient jamais les choses de biais ! En fait, les philosophes ont très longtemps oublié qu'ils avaient des mains !

Marx reprochera plus tard aux philosophes de s'être surtout attachés à interpréter le monde et appelle à le transformer[252]. Le passage de la contemplation à l'action devient ainsi lui-même un programme philosophique.

Le lecteur sans doute nous aura devancés. Après les grandes idéologies du XX^e^ siècle viennent l'informatique et les intelligences artificielles. Or, comme l'a bien vu Shannon Vallor, les IA sont des miroirs sans précédent de l'humanité. Leur connaissance est d'une profondeur, d'une finesse, d'une universalité inouïe, hors de mesure avec n'importe quelle mythologie ou idéologie. Dans ce miroir, l'humanité va se voir. Et cette prise de conscience va appeler une transformation, un tamisage à la mesure du puissant outil technique. Mais dans quel sens ? Cela est d'autant plus difficile à dire que ce reflet est très réaliste, donc semble a priori peu orienté. Cependant, ce miroir a aussi une spécificité : il parle et se comporte en bien des points comme un sujet. Ce « double intelligence » nous questionne et nous remet en question. Ce qui nous « secoue » dans ce grand tamis, ce qui est mis à l'épreuve, c'est à la fois notre statut d'être humain et de sujet.

PARTIE II

LE SUJET HUMAIN A L'EPREUVE DE L'IA

Cette partie fait passer le miroir de l'Utile vers le sujet humain lui-même, atteint dans son travail, ses désirs, ses besoins et son sens. L'outil censé servir l'homme y apparaît comme une puissance capable de le mesurer, de le remplacer et d'orienter ce qu'il désire. La perte de sens, la machine capitaliste et l'économie de l'attention y donnent à l'utilité une portée existentielle.

7. La double incongruité

La congruence désigne, de façon générale, ce qui convient, ce qui est adapté, ce qui est adéquat à un projet. C'est donc un joli mot pour un outil, pour une machine. Il faut qu'une machine soit congrue, adaptée aux fins poursuivies et au besoin de l'utilisateur. Pourtant, les IA, comme larges modèles de langage – les machines qui parlent – se caractérisent par une double incongruité. Incongruité d'abord, parce que, n'ayant pas de finalité clairement identifiée, il est impossible de dire si elles sont congrues ou non, adaptées ou non. Une machine à presque tout peut servir à presque tout, donc à n'importe quoi. Quelles sont les conditions qui amèneraient ou non à dire qu'elles sont adaptées ? Elles ont l'étrangeté d'un couteau suisse qui aurait des lames en nombre infini.

Dire que l'IA n'a pas d'utilité prédéfinie ou interne, mais que ses finalités seront finalement définies par l'usage rend d'autant plus cruciale la nécessité de penser leur puissance. L'IA sera ce que vous voudrez bien en faire. C'est un malin génie[253] ! On essaie seulement d'éviter qu'elle ne devienne un « mauvais génie »… Pourtant, serons-nous nous-mêmes… assez malins pour cela ? Sûrement pas si nous renonçons à être sujets humains,

c'est-à-dire, du point de vue du sens et de la métaphysique, en nous laissant absorber par des considérations techniques.

Incongruité ensuite, parce qu'un outil technique est censé être un moyen à disposition des humains pour atteindre leurs objectifs. Mais voilà que les IA sont le moyen de remplacer les humains eux-mêmes dans nombre de métiers, voire comme amis, compagnons, etc. On nous promet, pour l'avenir, des entreprises où les employés seront des agents IA autonomes, capables de prendre des décisions comme des humains, voire des entreprises qui seraient elles-mêmes dirigées par des IA. Le but, évidemment, serait de faire de l'argent pour des commanditaires.

Rappelons ici, parce que cela semble nécessaire, ce que nous faisons dans ce livre. Nous parlons de sens de l'humanité confrontée à l'« inquiétante étrangeté[254] » des intelligences artificielles. Nous ne nous posons pas comme philosophes analytiques occupés à décrire précisément le processus de ces machines et à les caractériser pour étudier de subtiles différences avec l'humain. Si ces différences doivent être prises en compte, nous tentons d'interpréter et de donner sens, voire un nouveau sens, à la perception de l'IA par une humanité largement néophyte. Celle-ci se trouve mise en présence d'un double qu'elle anticipe comme une opportunité fabuleuse en même temps qu'une menace. Nous considérons qu'il y aurait erreur à vouloir plonger toute une partie de l'humanité dans l'insignifiance parce qu'elle ne comprendrait pas soi-disant l'IA, parce qu'elle ne saurait pas de quoi ses algorithmes sont capables et ce qu'ils sont exactement. Nous l'avons déjà dit, il n'est pas plus nécessaire, pour avoir un vécu authentique de l'IA et des échanges avec elle, qu'il n'est nécessaire de savoir comment fonctionne un téléviseur pour regarder ses programmes. Bien au contraire, il nous semble que les propos de

spécialistes qui s'intéressent à des subtilités qu'eux seuls perçoivent, de par leur singularité même, empêchent de comprendre véritablement le sens des choses dont il s'agit.

L'IA, telle qu'elle est vécue, si elle n'est pas une réplique exacte de l'humanité, peut bel et bien menacer le gagne-pain d'un travailleur. Elle sera cette ombre qui plane au-dessus de lui et menace de le destituer professionnellement, voire affectivement. Elle peut, alors même qu'elle est sans intentions, le destituer de son estime de lui-même.

Toutes les subtiles distinctions que nous pourrions ici évoquer pour faire des différences sont considérées par les ingénieurs de l'IA comme étant à abolir – exception faite de mesures de sécurité sur lesquelles nous reviendrons. Beaucoup d'IA ont été baptisées de noms humains et ce n'est pas une vue de l'esprit. Il y a eu Alice, Alexa, Siri, Jasper et d'autres. Si l'on s'intéresse à l'une des plus célèbres, Claude, elle a été créée par l'entreprise Anthropic. Le nom de l'entreprise vient du principe anthropique en philosophie et cosmologie – l'idée que l'univers semble « réglé » de façon à permettre l'émergence de l'observateur humain. Or, l'entreprise s'est donnée pour mission de créer des IA alignées sur les valeurs humaines, compréhensibles, fiables… bref, « humaines » dans leur fonctionnement. Ce sont donc de très bonnes valeurs, mais l'on considère que ce sont des valeurs que l'on veut transférer à « Claude » pour qu'il se rapproche encore un peu de nous !

C'est donc avec parcimonie qu'il convient de faire le reproche d'anthropologiser l'IA à quiconque essaye d'interroger le sens d'une ressemblance qui est voulue par beaucoup et à laquelle beaucoup des plus grands génies de notre temps travaillent ! Cette ressemblance fait sens aussi pour ceux qui n'y travaillent pas et nombreux sont ceux qui la craignent. Nous ne disons pas

que la crainte soit sage, bien au contraire. Mais cette crainte a le droit d'être considérée avec un minimum de respect plutôt que frappée de nullité pour des raisons techniques. Depuis Homère jusqu'à nos jours, en passant par le Golem, Frankenstein ou L'Ève future de Villiers de l'Isle-Adam, l'humain a toujours fantasmé, espéré et craint – se faire des doubles capables de le remplacer. Si la technologie a son intérêt, il faut savoir aussi, de temps en temps, parler d'autres choses que de microprocesseurs ou de tuyaux. C'est proprement ce que nous appelons « penser » ! Questionner le sens de ce qui est là – quand nous y sommes aussi !

Le double a toujours été une telle menace dans la perception humaine qu'on y voyait jadis l'ombre de la mort d'un des deux êtres en présence. Ce n'est peut-être pas pour rien qu'il n'y a que nous comme uniques survivants de l'espèce « Homo ». Dans de nombreuses civilisations, dans de nombreux mythes (Romulus et Rémus, Caïn et Abel, Osiris et Seth[255]…), nous posons donc des questions d'ordre existentiel que l'humanité s'est elle-même posées et imposées en créant une intelligence rivalisant avec la sienne – alors même que c'est par son intelligence que le genre humain a dominé le reste de la planète.

Cette présence incongrue est accueillie, depuis 2022, par le grand public de façon contrastée : à la fois les bras ouverts et les poings serrés. Peur et désir mêlés. Ces technologies sont, à proprement parler, bouleversantes pour l'humanité.

8. La perte de sens

Bien sûr, on répète, depuis Kant, que si tout a un prix, l'homme seul a une dignité. Mais c'est bien abstrait. Dans le concret de la vie, la dignité est un combat du quotidien. Les humains ne sont

pas si généreux qu'ils accordent de la dignité à n'importe qui. Il faut sortir la tête de l'eau, mériter, travailler. Avoir un métier, c'est être utile, apporter son savoir-faire à la société, être « quelqu'un ». Dans ses Propos sur le bonheur, Alain écrit : « Le métier est ce qui sauve l'homme de l'ennui, du vice et du besoin. Mais surtout il lui donne un moi. Car celui qui ne fait rien ne sait pas qui il est. L'action fait l'homme ; et le métier, qui est action réglée, réfléchie, libre, donne à l'homme une conscience de lui-même. »

Être possiblement remplacé par une machine, ce n'est pas seulement être menacé dans son activité, mais mis en question dans sa valeur personnelle. Un Qui auquel pourrait se substituer un Quoi est fragilisé dans sa raison d'être. C'est une crise de sens dans son existence.

Ce livre que nous écrivons aurait pu s'appeler « le sens de l'IA dans l'aventure humaine ». C'est la « signification » de l'événement IA dans la vie des humains, à la fois sur un plan individuel et collectif. Mais le premier sens de l'IA, pour beaucoup d'humains, c'est une perte de sens de leur existence ; selon l'expression consacrée, c'est une angoisse existentielle. La perte des sujets humains en tant que sujets.

Si nous étions mystiques, nous appellerions ces moments de perte de sens, d'angoisses vécues dans la fièvre ou dans le creux de la dépression, des « traversées du désert ». Ces traversées du désert changent possiblement un homme si elles ne lui font pas perdre le goût de vivre. À l'échelle de l'humanité, c'est soit un tournant (selon le mot de Heidegger), soit un carrefour, selon le mot que nous avons utilisé à propos de Parménide. Si c'est un tournant, nous n'avons rien d'autre à faire que de le suivre ou de l'accompagner – ce qui ne signifie pas nécessairement se mettre à la remorque de la technique, mais advenir autre que ce que l'on

était pour répondre présent à une vocation telle qu'on n'en a jamais vue. Un tel tournant aura ses laissés-pour-compte.

Pour simplifier, puisque tout le monde ne peut pas être mystique ou heideggérien, si l'événement technologique que sont les IA, dans l'histoire de l'humanité, marque un tournant, alors nous avançons, mais, pour le plus grand nombre, sans savoir où nous allons. Nous ne sommes même pas maîtres de ce mouvement qui nous pousse vers l'avenir incertain et changeant. Dans un tel contexte d'une humanité en perdition, qui ne possède ni le volant ni la capacité de régler les vitesses – ou d'activer les freins – le plus grand nombre des membres de notre espèce se trouvera des maîtres, des guides, qu'ils soient hommes ou machines.

Si c'est un carrefour, cela signifie que ni la subjectivité ni la liberté ne sont ni fatalement ni totalement perdues pour l'homme, qu'il reste responsable de son avenir et de ses décisions.

La pensée taoïste du non-agir ne recommande pas l'inaction pure. Elle invite plutôt à ne pas forcer artificiellement le cours d'une situation : l'action juste consiste à reconnaître les tendances du réel, à les accompagner et à intervenir sans produire inutilement de résistance.

Ce laisser-faire est, en un sens, élémental ; il est compréhension profonde de la condition humaine. Il puise à la même source que la terre et l'eau – ces deux éléments qui, chez Lao Tseu comme chez Heidegger, symbolisent une manière d'être au monde sans prétention ni violence. Lao Tseu admire l'eau parce qu'elle « occupe les lieux que tous rejettent » et qu'elle agit sans jamais s'imposer ; Heidegger, quant à lui, médite sur la terre comme ce qui se retire, se tait, et pourtant porte toute œuvre, tout édifice, toute parole. Tous deux refusent la logique du maître et du possesseur de la nature : ils voient dans le sol et le cours des

rivières non des ressources à exploiter, mais des manières d'être – patientes, silencieuses, fécondes. Cette entente traditionnelle de l'être ne se déploie pas dans la clarté du concept, mais dans l'humus du monde, dans le rythme des saisons, dans le creux des vallées où le Tao circule. Elle ne commande pas ; elle accueille. Et c'est précisément en s'enracinant dans cette écoute de la terre et en suivant la fluidité de l'eau que l'homme, loin de se perdre, retrouve sa place juste – non au sommet, mais au cœur du vivant.

Très bien, très beau ! Il y a là aussi une pensée de l'enracinement que nous savons célébrer. Mais il n'est pas sûr, et il n'est pas certain, qu'on puisse monter sur le dos de l'IA comme sur le dos d'un dragon, et que lâcher la bride soit la meilleure idée. Où nous mènera le feu de cette nouvelle technique ? Car ce n'est évidemment pas de l'eau, mais un feu du genre de celui de Prométhée qui transforme durablement les hommes. Les IA restent des machines, et les machines peuvent s'emballer, nous pouvons en perdre le contrôle – et il n'est pas certain que la sagesse immémoriale de la nature s'exprime encore en elles, s'exprime encore par elles.

Rien de plus nécessaire, en effet, dans les moments de grands tournants, dans les moments de grand danger, que de savoir s'arrêter, correctement respirer, se recueillir. Mais c'est le moment de rester vigilant, d'ouvrir bien grand les yeux. Dans la grande accélération de la technique, sans doute, il faut savoir infiniment ralentir. Pour percevoir. Mais rester éveillé. Car, dans le silence de l'attention, on se prépare à percevoir le moment de l'action où l'on fait des choix. C'est dans ces moments-là que l'homme, ou que l'humanité, peut prendre un nouveau cap.

La mystique du laisser-être, certes, ne peut pas être confondue avec un simple laisser, un simple « se laisser porter par le flot des

événements », mais elle peut être aussi un danger quand le bateau de l'humanité prend une mauvaise direction.

Dans ses Propos sur le bonheur déjà cités, Alain parlait des marins grecs qui, à proximité des « colonnes d'Hercule », endroit que l'on savait éminemment dangereux pour les navires, avaient l'idée saugrenue de fermer les yeux et les oreilles pour ne pas se laisser charmer par la voix des sirènes[257]. Il nous semble que c'est là, au contraire, qu'il s'agit d'exister comme sujets, de garder grands ouverts les yeux et les oreilles – pour pouvoir donner le coup de barre au moment où il conviendra.

Quand l'homme, et pour la première fois, s'est mis debout, il ne s'est pas laissé aller – d'aucune façon. Par contre, qu'il ait connu un moment de flottement, une conversion presque mystique, nous n'en disconvenons aucunement.

9. Le sens du sens et le précieux

Quel est donc le sens de l'IA dans l'aventure humaine ? C'est une machine, elle va. Elle a sa dynamique propre. Elle fait tourner en boucle ses algorithmes, ses algorithmes sont de plus en plus grands, de plus en plus rapides. Mais qu'est-ce qui fait tourner la boucle, puisque ce n'est pas – la vie.

Dans l'Antiquité, tout ce qui était animé avait une âme. Le mot âme vient de anima qui a donné animation. Ainsi le soleil qui était animé, dont le mouvement circulaire était parfait, avait une âme – divine. La machine à vapeur a été une véritable révolution. Le mouvement – physique – ne vient plus de la Physis, c'est-à-dire de la nature (physis en grec signifie nature). Le mouvement, désormais et depuis le XIX^{e} siècle, est sans âme. Est-ce à dire que depuis le XIX^{e} siècle nous avons perdu – un peu de la nôtre ?

Qu'en se livrant à un mouvement artificiel nous avons perdu quelque chose de précieux ?

Ce n'est pas à exclure. L'abandon des campagnes, l'urbanisation, les mines pour en tirer de l'énergie, la prolétarisation, le travail dans les manufactures puis les usines, les révoltes ouvrières, le machinisme, le travail à la chaîne, les temps modernes, l'homme qui ne voit plus ni ne regarde le ciel pour savoir le temps qu'il fera demain, la course au profit, à la vitesse, la grande accélération qui ne donne pas de temps mais qui en prend. Tout cela, ce n'est pas rien. Viendra un jour, l'appareil de production en aura les moyens, la société de consommation, les ouvriers de Ford produisant à la chaîne des voitures de moins en moins chères, jusqu'à devenir eux-mêmes les consommateurs du monde qu'ils contribuaient à fabriquer. Non, ce n'est pas tout à fait rien. L'humanité déjà tournait en boucle. Son véritable moteur n'était ni la machine à vapeur ni le moteur à explosion ; c'était le capitalisme. Et le communisme lui-même dans sa très apparente contradiction a épousé finalement le même mouvement que son soi-disant adversaire, est entré avec lui en compétition de productivité. Mais comme on ne brûlait pas que du pétrole ou du charbon, que les bobines d'électricité ne faisaient pas tout, on consommait aussi des vies humaines. Pourquoi ne pas les mettre aussi dans la machine, si l'âme n'existe pas ? Si la machine, c'est la victoire du matérialisme ?

Et comme on a pu dire, Marx s'est noyé dans le Coca-Cola[258]. Finalement l'énergie de l'ouvrier capitaliste était mieux redistribuée que celle du communiste qu'on payait essentiellement de promesses. Un tien vaut mieux que deux tu l'auras. Le Coca-Cola est plein de sucre et d'énergie pour le travailleur. L'énergie humaine aussi fait tourner la machine, c'est le prix de la civilisation

quand elle tourne avec des moteurs. Peut-être en effet qu'on y a perdu un peu de son âme. Baudrillard analyse précisément la société de consommation comme un système où le bonheur devient une référence omniprésente : la publicité ne vend plus seulement des objets, mais la promesse d'un salut profane par la consommation.

Certes, et pourtant, dès 1851 et bien avant l'essor de la société de consommation, le poète Alfred de Vigny avait déjà ce sentiment d'une perte précieuse devant les performances inouïes de la machine à vapeur :

Que Dieu guide à son but la vapeur foudroyante
Sur le fer des chemins qui traversent les monts,
Qu'un Ange soit debout sur sa forge bruyante,
Quand elle va sous terre ou fait trembler les ponts
Et, de ses dents de feu, dévorant ses chaudières,
Transperce les cités et saute les rivières,
Plus vite que le cerf dans l'ardeur de ses bonds
(…)
Sur le taureau de fer qui fume, souffle et beugle,
L'homme a monté trop tôt. Nul ne connaît encor
Quels orages en lui porte ce rude aveugle,
Et le gai voyageur lui livre son trésor,
Son vieux père et ses fils, il les jette en otage
Dans le ventre brûlant du taureau de Carthage,
Qui les rejette en cendre aux pieds du Dieu de l'or.
Mais il faut triompher du temps et de l'espace,
Arriver ou mourir. Les marchands sont jaloux.
L'or pleut sous les chardons de la vapeur qui passe,
Le moment et le but sont l'univers pour nous.
Tous se sont dit : « Allons ! » Mais aucun n'est le maître

Du dragon mugissant qu'un savant a fait naître ;
Nous nous sommes joués à plus fort que nous tous.
(…)
Et encore :
La distance et le temps sont vaincus. La science
Trace autour de la terre un chemin triste et droit.
Le Monde est rétréci par notre expérience
Et l'équateur n'est plus qu'un anneau trop étroit.
Plus de hasard. Chacun glissera sur sa ligne,
Immobile au seul rang que le départ assigne,
Plongé dans un calcul silencieux et froid[260].

Ce que nous appelons le « sens du précieux » est la capacité à reconnaître et à valoriser ce qui est unique, fragile, irremplaçable, de ce qui est porteur d'une signification profonde, au-delà de la simple utilité ou de la valeur marchande. Il est en quelque sorte le sens du sens, non pas telle ou telle valeur particulière, mais l'attention recueillie qui fait qu'il puisse y avoir quelque chose qui vaille. Qui a une idée de ce qu'est la dignité humaine a une idée de ce qu'est le sens du précieux, mais il ne s'y réduit pas.

Le sens du précieux est à proprement parler indéfinissable parce qu'il n'a pas de contours donnés, mais parce qu'il est plutôt attention à ce qui pourrait se profiler. Il tient de ce « presque rien de Vladimir Jankélévitch » qui change « presque tout ». Il comprend ces réalités décisives mais insaisissables que sont le charme, la grâce, l'instant et l'intention morale.

Il est l'émotion qui nous prend, quand on voir s'ouvrir les doigts ridés d'un vieillard serrés par la main du nouveau-né, l'éclair fugace qui l'espace d'un instant éclaire un visage menacé par l'abîme.

10. La machine capitaliste

Un moteur plus puissant que celui des locomotives fait tourner la roue des temps modernes. Il a pour nom le capitalisme et il a succédé à la religion. Son centre, c'est l'argent. Pour nous faire bien comprendre, rappelons que pour notre époque, qu'on le célèbre ou le déplore, l'utile est presque toujours ramené à la notion d'argent. Il convient de prendre le problème tel qu'il se pose. C'est d'autant plus important que, comme nous le verrons, la logique de l'IA est non seulement tout à fait adaptée au capitalisme mais qu'elle répond à la même logique dans ses fondamentaux : un cycle continu de croissance et de conversion toujours plus vaste, rapide et englobant, qui s'alimente même de ce qui semble le contrarier. Il est simultanément destruction et création, création dans la destruction même comme l'avait vu Schumpeter[261]. Dans ce vaste mouvement, le monde et les sujets humains sont pris et transformés.

Arrêtons-nous un instant sur la signification de ce mot : capitalisme. C'est beaucoup plus qu'un simple modèle économique comme il pourrait y en avoir d'autres. Il recèle une métaphysique de l'efficace dont la spéculation n'est pas bâtie sur les nuages, mais sur une expérience bien terrienne et bien concrète, une métaphysique féconde comme son origine. A l'origine en effet, le capitalisme ne fait pas seulement « de l'argent ». Il accélère la roue du temps, produit en masse des nouveautés et de nouvelles conditions pour l'humanité.

Arrêtons-nous un instant sur la généalogie de ce mot : capitalisme. Son radical, capital, plonge dans un très ancien vocabulaire de la richesse où le bétail occupe une place centrale. Capital et cheptel sont parents : l'ancien français chatel, devenu cheptel, et le terme savant capital remontent au latin caput, la tête – la partie

principale d'un bien, mais aussi la tête de bétail. La richesse longtemps s'est comptée en terres productives, en récoltes et en têtes d'animaux. Or le troupeau possède une propriété singulière : il se reproduit. Ce qui excède le besoin immédiat peut produire à son tour de la richesse.

Or le bétail qui excède le besoin peut se reproduire. Toute cette énergie supplémentaire pourra être exploitée et convertie en un autre type d'énergie : l'argent. L'argent ne met pas directement en mouvement des bêtes, mais il motive les hommes qui peuvent utiliser l'énergie biologique des bêtes, etc. D'une certaine façon, le temps joue en la faveur du capitalisme et lui donne la richesse.[262] En un sens profond, il privatise le temps, ce qui a pu paraître scandaleux à l'Église.

Le temps, en effet, pour les religieux, appartient à Dieu. Les cloches qui sonnent pour marquer les prières, le rappellent[263]. Avant le capitalisme, l'argent ne peut faire des « petits » comme les vaches font des veaux, l'usure est interdite. L'Église catholique défend une cosmologie de la stabilité, voire de l'immuabilité. Que chacun assume son lot ; il y a un ordre des choses auquel il ne faut pas déroger. Rien n'est jamais vraiment nouveau sous le soleil. Que chacun sache se situer. Tout en haut, il y a Dieu puis ses archanges et puis les anges. Bien en dessous, les rois, puis le haut clergé et les nobles ; les paysans sont attachés à la terre qui est la seule richesse. Les puissants, le clergé, peuvent prélever des impôts comme d'un arbre on cueille les fruits. La nature prodigue ses bienfaits entretenue par le travail des hommes, et les ressources se renouvellent quasiment identiques à chaque saison selon la volonté de Dieu, le travail des hommes, leur piété et les prières pour les moissons.

Le capitalisme, en particulier à la Renaissance italienne, va apporter les banques et une nouvelle classe qui n'est pas attachée à la terre, les bourgeois. Comme leur nom l'indique (habitants du bourg), ils sont de la ville et non de la campagne. Ils profitent du marché où il y a des bancs pour les changeurs (d'où le nom de banque qui vient de l'italien « banca ») pour faire des affaires. Certes, en eux-mêmes, les changeurs ne produisent rien, mais ils font « travailler l'argent » qui devient à son tour productif. Comment cela ?

C'est qu'ils ont compris qu'il y avait dans l'argent un peu plus que de la valeur d'usage. La monnaie elle-même est cycle, elle passe de main en main. Elle est une valeur qui roule et qui voyage. Elle traverse des villes et des pays, elle est pour ainsi dire mouvement. Dans le secret de son mouvement, elle capte un peu du temps qu'il faut pour l'échanger.

Dans les foires d'autrefois, il y avait de nombreuses monnaies selon les pays, les provinces et les seigneuries. Il fallait les convertir pour qu'elles prennent les couleurs du pays. Or, symboliquement, il y a là comme un passage à la « douane du sens ».

Les changeurs convertissaient et, en même temps, pour leur service de facilitation, ils prenaient une commission. Ils pouvaient garder de l'argent en dépôt pour fluidifier les mouvements, sécuriser les transactions. Pour éviter les transports dangereux, ont inventé (à Florence en particulier) les reçus qui sont les ancêtres de nos chèques. Ils pouvaient également faire des prêts plus ou moins risqués pour récolter des bénéfices. Par mille et un stratagèmes, ils ont contourné les interdits de l'Église[264]. Le temps n'était plus seulement à Dieu ; il travaillait un peu pour eux. Non seulement on a besoin d'eux, mais ils peuvent ouvrir leurs coffres comme des facilitateurs des désirs, allant jusqu'à

financer des rois comme Philippe II d'Espagne, des empereurs comme Charles Quint, et des papes comme Eugène IV.

S'ils ne pratiquaient pas l'usure à proprement parler, du moins les Florentins et les Génois, les Vénitiens même, ont-ils su adroitement faire fructifier leurs investissements et en dégager un intérêt. Comment penser que les grandes familles de banquiers (les Médicis, les Grimaldi) avec un tel pouvoir ne s'occuperaient pas un peu de politique ? Ils parient sur l'avenir, et quand on parie sur l'advenir, le mieux – pour en récolter quelque profit – est de l'influencer dans le sens de ses intérêts. Si tu veux prévoir l'avenir… le mieux est de l'organiser !

Finalement, le temps jouait pour eux. Le cycle des saisons ne tournait plus à vide, mais les enrichissait. Cette nouvelle relation au temps – qui pour eux est une amélioration – est la condition du progrès, d'un tout autre rapport au temps où il y a bénéfices et investissement, innovation pour d'autres bénéfices, etc. C'est ainsi que l'on passe d'une idéologie de l'immobilité à une idéologie de la nouveauté. Même des expéditions comme celle de Christophe Colomb n'auraient pu être financées sans les banquiers (Luis de Santángel[265] en premier lieu), car la couronne d'Espagne à l'époque était ruinée.

D'ailleurs, même la religion semble s'être inclinée devant ce nouveau pouvoir, car l'argent, on le sait depuis que le monde est monde, est fort utile ici-bas, et il semble que pour les affaires du monde on ait trouvé d'autres ressources que les impôts et la dîme. Les banquiers allemands, et notamment les Fugger[266], n'inventèrent évidemment pas les indulgences, mais ils furent au cœur du dispositif financier qui accompagna leur diffusion au temps de Léon X : une partie des recettes servait à financer Saint-Pierre, une autre à rembourser les dettes d'Albert de Brandebourg.

Contre de l'argent, un chrétien pouvait espérer alléger ses péchés, pour lui-même ou pour ses morts. Dans les prédications populaires, la frontière entre pénitence, indulgence et pardon pouvait devenir singulièrement mince. L'argent entrait ainsi jusque dans l'économie du salut.

D'un point de vue anthropologique, quelque chose de la vieille logique sacrificielle subsiste ici, mais l'objet sacrifié a changé de nature. Autrefois, l'homme rendait aux dieux une part du vivant – un animal, du grain, du vin – que le monde lui avait donné. Désormais il peut offrir de l'argent, cette valeur abstraite produite par les hommes eux-mêmes. Dieu faisait le bétail ; l'homme fait l'argent. C'est un sacrifice rentabilisé, calculé, et Luther n'avait peut-être pas tout à fait tort d'y voir ressurgir les marchands du Temple. Ce déplacement n'est pas seulement économique : il est métaphysique.

Mais l'ironie de l'histoire est plus grande encore. En combattant cette circulation trop directe entre l'argent et le salut, la Réforme n'a pas diminué la puissance de l'argent ; elle a contribué à lui ouvrir une autre carrière. Calvin, notamment, admet le prêt à intérêt sous certaines conditions. Comme l'a montré Max Weber dans L'Éthique protestante et l'esprit du capitalisme[267], certaines formes du protestantisme ascétique ont valorisé le travail méthodique, l'épargne et le réinvestissement plutôt que la dépense religieuse ou ostentatoire. L'argent ne pouvait plus acheter le ciel ; il allait désormais conquérir la terre.

Au-delà des époques et des transformations économiques, politiques et sociales, le capitalisme est un moteur puissant qui ne semble plus jamais vouloir s'arrêter. Il a su faire de l'or avec le temps. Avec les moulins à aube, on a su convertir le mouvement de l'eau en énergie, on fit de même avec le temps. Le mouvement

du temps lui-même, pour qui sait l'exploiter, peut devenir une formidable source d'énergie qui fait jouer à son avantage toutes les autres sources d'énergie : le travail des bêtes, des hommes, du vent et des cours d'eau, la croissance des plantes, les aléas du climat, le temps qu'il faut pour parcourir les grandes et les petites distances.

Le capitalisme a une dimension métaphysique qui échappe d'ordinaire ; il est un autre rapport au temps. Le temps n'est plus simplement circulaire (comme une « image mobile de l'immobile éternité » selon les mots de Platon et qui correspondent assez exactement à la conception religieuse du Moyen Âge). Les cycles du temps doivent désormais être pensés comme des spirales, car ils comprennent quelque chose de plus, de l'argent qu'on fait fructifier, des risques, des nouveautés financées dont on attend quelque profit en mutualisant les risques, etc. Il n'est d'ailleurs pas anodin que ces changements soient plus ou moins concomitants aux formidables progrès techniques, scientifiques, religieux de la Renaissance.

Tout cela, l'esprit du capitalisme l'irrigue de ses formidables moyens, l'aide à se développer. Comme jadis on arrosait les plantes d'eau, on arrose le mode d'argent, on essaye des nouveautés, et bien souvent, ça pousse et fructifie ! Le progrès se diffuse. Si la terre n'est plus au centre du système, si elle tourne, pourquoi l'homme resterait-il immobile, pourquoi ne pas célébrer l'universel mouvement de notre planète. Galilée lui-même sera financé par les Médicis. Peut-être a-t-il d'ailleurs surestimé la protection que cela lui valait.

Donc, quand nous parlons du capitalisme, il faut entendre beaucoup plus qu'un simple modèle économique. Une vision du monde cohérente née d'une possibilité biologique d'abondance

tirée du bétail convertie ensuite en argent que le temps lui-même, avec un peu d'esprit d'entreprise et d'audace, pourra faire féconder, encourageant dans tous les domaines, politiques, scientifiques, techniques, artistiques, etc., l'innovation et le progrès. C'est une interprétation du temps comme pouvant être fécond qui a montré sa fécondité en même temps que ses revers. L'invention de la machine à vapeur fut un formidable accélérateur pour le capitalisme : de l'énergie se libérait, qui ne venait plus des animaux vivants, mais de reliquats de vivants morts depuis bien longtemps, plus précisément, de matière organique végétale accumulée il y a des centaines de millions d'années. Il semble que ce qui est mort se mette à faire tourner la roue du capitalisme. C'est le cas de toutes les énergies fossiles : le pétrole et le gaz naturel viennent aussi de matière organique morte, de micro-organismes marins (plancton, algues) accumulés au fond des océans. La culture, la science nourrissent le capitalisme. On ne fait pas que convertir de l'argent, on convertit une source d'énergie en une autre, par exemple le charbon en électricité, et on prélève au passage un petit intérêt de changeur qui deviendra l'argent qui permettra de nouvelles conversions et innovations. La connaissance elle-même est devenue un marché. Un jour, on convertira la matière même en énergie qui n'aura plus rien à voir avec la vie. Le nucléaire. Mais pour dire la vérité, le véritable marché aujourd'hui est le marché de l'information et de la connaissance. Certes, il demande d'énormes quantités d'énergie, mais ce qu'il s'agit de convertir, c'est ce que produisent en quantité les humains. De la pensée en données. Grâce à tout ce qui relève de l'écriture, on transforme de la pensée fossile en données recyclables et réutilisables que l'on peut vendre et échanger en monnaies dématérialisées.

Le capitalisme d'aujourd'hui s'évalue moins en tête de bétail comme à son origine qu'en ce que produisent les têtes humaines en permanence et dont on capte la quintessence dans les réseaux sociaux. Certains accusent le capitalisme de marcher sur la tête, ce qui n'est pas impossible, ses conversions ont un coût ; c'est la planète qui paye. Mais ce qui est certain, c'est que le capitalisme d'aujourd'hui marche essentiellement sur les têtes humaines en convertissant nos pensées en de très précieuses données qui seront revendues, échangées, transformées, analysées et disséquées par de grands modèles d'IA qui offriront au capitalisme, en même temps qu'à nous-mêmes, de nouvelles possibilités, des modes d'organisation nouveaux. On parle de capitalisme cognitif[268]. Après la marchandisation des corps dénoncée par Marx, qui n'y a pas vraiment trouvé remède, il y a la marchandisation des esprits.

Un capitalisme vertueux est-il possible ? La question reste ouverte, et c'est éminemment souhaitable, car il est possédé par une sorte d'autophagie où il détruit ses propres conditions de possibilité. La nature, en effet, jadis divinisée et respectée, est dépassée de vitesse et ne peut pas renouveler ce qu'il détruit ; du moins, de moins en moins. La décroissance est en un sens la seule doctrine économique cohérente qui va à l'encontre du consumérisme capitaliste, mais elle souffre d'un défaut majeur. C'est qu'en renonçant volontairement à se doper d'énergie, elle ne se donne pas à elle-même les conditions pour s'imposer massivement. Pour paraphraser George Bataille, la décroissance à la sagesse de la lentille d'eau et le capitalisme la vigueur d'un tigre. Un jour peut-être, il y aura des lentilles d'eau et les tigres auront disparu, faute de gibier à l'entour, faute d'énergie à capturer quand les lentilles

d'eau se multiplieront, mais les le jour n'est pas arrivé et les humains se voient en tigres !

Le capitalisme lui-même est une machine métaphysico-économique qui se nourrit de toutes les autres machines dont il a contribué à la mise en mouvement. Il a permis le « progrès » et, quoi qu'on en dise, nous lui devons une forte élévation des niveaux de vie, la possibilité d'exister en tant qu'individus, d'avoir une vie à soi, une vie privée comme il y a des propriétés privées. Chacun l'interprètera à sa façon, on peut avoir aussi la nostalgie de la communauté. Le passé éclaire l'avenir, mais c'est nous qui tenons la torche qui illumine du passé ce qui nous semble le mériter. Il nous revient, et à chacun, de mettre en lumière ce qui en lui était précieux et de laisser dans l'obscurité ce qu'on souhaite oublier. C'est la loi du réel.

Impossible de trouver l'origine de la citation « Pourquoi serais-je contre ce qui sera ? » Si ce n'est pas nous qui l'avons inventée, du moins, nous la faisons nôtre. Mais du futur, si véritablement nous voulons être, il convient de le domestiquer. Ce qui n'a pas de passé n'a pas non plus d'avenir ; c'est la barbarie. Ce qui advient néantise de larges pans de notre passé. Mais on se souvient. Pour aller loin, du souvenir, ne perdons pas le fil. Il nous rattache à l'humanité, plus profondément – à la beauté de la vie.

Nous n'avons pas perdu notre âme. Si c'était le cas, nous ne pourrions pas même en parler. Mais nous devons notamment aux poètes d'avoir gardé le fil – que nous ne voulons pas lâcher. Après, ce que vaut exactement une âme, nous ne saurions l'estimer. Depuis « la mort de Dieu », selon l'expression de Nietzsche, le discrédit dans lequel sont tombés les dualismes, nous ne savons même plus trop si le mot « âme » signifie – pour les humains – quoi que ce soit de précieux.

Pourtant…

Il y a une chose que, peut-être, nous ne devons pas perdre.

C'est le sens du précieux dont nous avons déjà parlé. Car c'est ce sens seulement.

Qui fait la valeur d'une vie humaine.

Et peut-être de la vie tout court.

Comme les moteurs d'aujourd'hui et d'autrefois, les algorithmes tournent en boucle avec toujours plus de vitesse et de force. Derrière leurs incessantes rotations, il se pourrait qu'il se cache à la fois un autre moteur et toujours le même : celui du capitalisme. Si des machines, des mécaniques ont pu se substituer au corps des hommes jusqu'à les remplacer ; si cela a produit des mouvements sociaux et des révolutions – que sera-t-il pour ces nouvelles boucles de la connaissance qui s'accélèrent et tournent plus vite, plus parfaitement que nous ? N'y perdrons-nous pas aussi un peu de notre âme, de ce qui fait notre précieux ?

Si l'âme n'existe pas et si rien n'est précieux… nous n'avons rien à perdre.

Mais il demeure une autre solution.

Nos âmes devenues inutiles pour la connaissance comme pour la production pourraient connaître, après de terrifiantes angoisses et quelques convulsions, une mutation. Plutôt que de mourir, nos énergies psychiques pourraient muter.

Il nous paraît, dans tous les cas, que si nous subsistons en tant que sujets, il nous faudra changer de perspective. La boucle des algorithmes n'est peut-être pas si différente de celle des moteurs qui font tourner des machines ; que ces machines elles-mêmes ne tournent que sous l'impulsion d'un plus grand moteur qui a pour nom capitalisme, que le capitalisme lui-même emprunte – légitimement ou non – ses mouvements aux cycles de la vie comme

l'indiquent les mots consommation et production, que tout cela est affaire de flux et d'énergie, voilà qui peut être éclairé par Bataille. Dans La Part maudite, il soutient que la vie produit continuellement des excédents d'énergie et que les sociétés doivent les dépenser, parfois de façon productive, parfois dans le luxe, la fête, l'art, la guerre ou le sacrifice[269]. L'homme devient ainsi un être particulièrement apte à consumer ce surplus.

Ces cycles d'énergie ont bouleversé profondément, irréversiblement notre planète et notre environnement. Comment pourrait-on croire sérieusement qu'ils ne sont pas en train de nous transformer nous-mêmes.

La seule chose à laquelle nous veillons, c'est – plutôt que de parler à l'infini de pertes et d'oubli – nous conservions ce sens qui nous donne visage d'homme – le sens du plus précieux. Quel qu'il soit. Et il nous semble que ce précieux-là – a quelque chose à voir avec l'être sujet.

11. Le désir, moteur du capitalisme

Depuis la Renaissance à peu près, depuis qu'on reconnaît que la Terre tourne autour du Soleil, le monde des hommes tourne et tourne de plus en plus autour de l'argent. Mais qu'est-ce qui fait tourner le capitalisme lui-même ? Nous ne ferons pas languir davantage le lecteur, c'est bien sûr le désir humain.

Derrière toute grande pensée, derrière tout grand processus humain, il faut en revenir à la pensée profonde de Descartes, l'ego cogito. En vérité, tout tourne autour des humains dans le monde des hommes.

Mais revenons-y : comment Descartes définit-il la pensée ? « Par le mot de penser, j'entends tout ce qui se fait en nous de telle sorte que nous l'apercevons immédiatement par nous-

mêmes ; c'est pourquoi non seulement entendre, vouloir, imaginer, mais aussi sentir, est ici la même chose que penser. » (Principes, I, § 9). C'est riche, mais est-ce suffisant ? N'oublie-t-il pas quelque chose ? L'amour et le désir, par exemple ?

Il ne les oublie pas, et en un sens, il a raison, car le désir et l'amour ne sont pas véritablement des pensées ; ils échappent pour une grande part à notre conscience. Le désir, pour lui, est plutôt du côté du corps (à son époque, la notion d'inconscient n'existe pas encore ; il faudra attendre Freud et Lacan pour cela). Quand le désir chez Descartes est véritablement pensé, car il ne l'oublie pas, il devient « volonté ». On voit dans son livre Les Passions de l'âme comment il entend redevenir maître et possesseur de son désir en le convertissant en volonté.

Malheureusement, nous ne vivons pas à une époque où la volonté et la « bonne volonté » cartésienne ont triomphé, nous rendant maîtres de nous-mêmes ; nous vivons à une époque où c'est bien plutôt le désir qui a pris le pas sur la volonté. Le fantasme et le désir exploités par la société de consommation normalisent nos envies et nos ambitions. Bombardés de messages publicitaires et saturés d'informations que nous ne pouvons pas traiter, nous sommes pris dans un tourbillon de stimuli qui laisse de moins en moins place à la réflexion.

En bien des sens, c'est un malheur, parce que si le désir humain échappe à la pensée, notre désir peut nous « manipuler ». Cette manipulation est non seulement au cœur du capitalisme, mais également au cœur de la vie. Il y a peut-être une possibilité pour qu'elle ne soit pas au cœur de l'IA, mais nous y reviendrons. Le lecteur aura compris que nous faisons là un travail préparatoire…

Le désir est au cœur du capitalisme, car le capitalisme est au cœur de la vie… Rappelez-vous, ces têtes de bétail qu'on pousse

à la reproduction pour s'enrichir. La sexualité et la reproduction animale sont au cœur du capitalisme. On manipule un peu le bétail (quelquefois au sens propre !) pour qu'il se reproduise et ainsi nous enrichisse !

Mais la nature ne fait pas autrement pour se reproduire. Tout ce qui est reproduction sexuée a pour moteur le désir. Si les fleurs se font belles et si parfumées, elles ne le savent pas elles-mêmes, mais c'est pour attirer l'abeille qui va leur permettre de croître et se multiplier. Si l'oiseau a un si beau plumage, une crête au-dessus de sa tête, c'est pour montrer sa force et sa beauté, qui donnent l'envie de l'accouplement. Du désir naîtront de laids oisillons qui prendront à leur tour plumes et crêtes et se reproduiront. Schopenhauer voit dans ce mécanisme un piège de l'espèce : le désir sexuel persuade l'individu de poursuivre son propre bonheur alors qu'il sert, selon lui, d'abord la perpétuation de la vie.

Très bien ! Donc, nos premiers capitalistes, nos premiers éleveurs de bétail ont donc trompé la trompeuse nature ! Manipulant la fécondité animale, ils ont usé du puissant instinct sexuel des animaux, non évidemment pour le profit des individus du cheptel, non pour la perpétuation de l'espèce en tant que telle, mais pour leur pur et simple profit. La culture et l'élevage ont remplacé la nature.

Nous insistons sur le fait que ceci n'est pas une simple métaphore. Non seulement la généalogie du mot capital plaide en notre sens – non seulement, comme nous l'avons vu, le capitalisme est né de l'observation des cycles de reproduction de la nature, les a repris à son compte pour les contrôler et les utiliser à des fins d'enrichissement – mais le capitalisme américain, le plus débridé de la planète, a les mêmes origines. On s'enrichit avec des têtes de bétail. Quand on parcourt le continent en train,

quand le télégraphe permet la communication, les machines prennent le relais. Le capitalisme deviendra minier, ensuite il y aura le pétrole, comme le charbon, tirés eux-mêmes d'anciennes énergies fossiles – c'est-à-dire des exploitations tardives du vivant par la machine !

Les hommes, sans doute, avaient les moyens de le faire depuis le néolithique, mais il a fallu quelque chose de nouveau : « l'esprit du capitalisme », une certaine rupture avec un ordre calqué sur la nature pour s'en rendre « maîtres et possesseurs », comme disait Descartes ; le volontarisme de l'enrichissement individuel et systématique par la rationalisation systématique du vivant à des fins de profit. Il a fallu, nous l'avons vu aussi, une certaine perception du temps comme appartenant aux hommes et pas seulement à Dieu. Les parents commencent à croire en un possible progrès, on espère que les descendants, les enfants (ou les « neveux », comme disait Descartes, qui avait perdu son unique fille) vivront mieux que leurs parents.

La Bible disait « croissez et multipliez-vous », pensant que les enfants étaient la plus grande richesse de l'humanité, mais le capitalisme veut croître et multiplier ses profits pour se constituer un patrimoine. Cela prendra du temps, mais un beau jour l'impératif « Enrichissez-vous ! » de François Guizot aura des connotations presque moralisantes[270]. La valeur d'un individu n'a plus la dignité du sacré ; elle commence à être fonction de son capital. Comment un être humain qui n'est pas propriétaire, qui n'a pas su faire fructifier ses revenus, pourrait-il être assez responsable pour s'occuper des affaires de la France ?

Guizot arrive aux affaires jusqu'à être président du Conseil, ce qui est à peu près le rôle d'un Premier ministre après la Révolution de Juillet 1830. Celle-ci, après les longues tribulations qui

suivirent la grande Révolution française (dont beaucoup ont dit qu'elle était avant tout bourgeoise), porte Louis-Philippe au trône sous le titre de « roi des Français ». La nouvelle organisation sociale, appelée Charte constitutionnelle, conserve un système électoral censitaire très restrictif. Non ! Il ne faut pas exagérer cette égalité des citoyens qu'on a tant chantée avec la Marseillaise quelques décennies plus tôt. Pour voter, il fallait être homme, majeur, français, jouir de ses droits civils et politiques… Et payer un impôt, le « cens » de son petit nom, d'au moins 200 francs par an, seuil fixé par la loi électorale du 19 avril 1831. Pour être éligible, c'est-à-dire candidat, le cens était fixé à 500 francs – somme énorme !

Or, cet impôt était essentiellement basé sur la propriété foncière ou immobilière, ou sur certains impôts directs (comme la contribution foncière ou mobilière). Autrement dit, seuls les propriétaires – ou les très hauts locataires – pouvaient voter.

En deux mots, c'est la fin des aristocrates dont la richesse était la terre ; c'est le triomphe du bourgeois. Dans le monde qui devient capitaliste, on pèse plus ou moins, et le poids s'évalue en argent, pas en têtes de bétail ou pas nécessairement.

Comme c'est l'époque de la révolution industrielle, de nouvelles richesses vont venir de produits manufacturés et de « marques ». La marque est un label de qualité et de prestige, elle a une « aura » qu'entretient une nouvelle et fondamentale industrie – à moins que ce ne soit déjà un service : la publicité. Le beau entre en scène non plus sur le grand marché de la foire à bestiaux, mais dans l'ordre du désir humain. D'ailleurs, quelquefois, on confond un tout petit peu. Comment ne pas citer en cette occasion un extrait du dialogue de Madame Bovary, de cette femme Emma, qui a épousé Charles, qui, hélas, semble tenir – comme

son nom l'indique si bien – du bovidé que de l'être humain. Il faut la comprendre, elle et son désir, elle se cherche un amant. Et puisque les occasions manquent de se rencontrer en province, cela se fait à la foire des Comices agricoles. Rodolphe, agité par un désir que beaucoup d'hommes reconnaîtront, que beaucoup de femmes comprendront, déclare sa flamme à Emma et lui saisit la main… pendant ce temps-là les affaires continuent :

Et il saisit sa main ; elle ne la retira pas.

« Ensemble de bonnes cultures ! » cria le président.

– Tantôt, par exemple, quand je suis venu chez vous…

« À M. Bizet, de Quincampoix. »

– Savais-je que je vous accompagnerais ?

« Soixante-dix francs ! »

– Cent fois même j'ai voulu partir, et je vous ai suivie, je suis resté.

« Fumiers. »

– Comme je resterais ce soir, demain, les autres jours, toute ma vie !

« À M. Caron, d'Argueil, une médaille d'or ! »

– Car jamais je n'ai trouvé dans la société de personne un charme aussi complet.

« À M. Bain, de Givry-Saint-Martin ! »

– Aussi, moi, j'emporterai votre souvenir.

« Pour un bélier mérinos… »

– Mais vous m'oublierez, j'aurai passé comme une ombre.

« À M. Belot, de Notre-Dame… »

– Oh ! non, n'est-ce pas, je serai quelque chose dans votre pensée, dans votre vie ?

« Race porcine, prix ex æquo : à MM. Lehérissé et Cullembourg ; soixante francs ! »

> Rodolphe lui serrait la main, et il la sentait toute chaude et frémissante comme une tourterelle captive qui veut reprendre sa volée ; mais, soit qu'elle essayât de la dégager ou bien qu'elle répondît à cette pression, elle fit un mouvement des doigts ; il s'écria :
>
> – Oh ! merci ! Vous ne me repoussez pas ! Vous êtes bonne ! Vous comprenez que je suis à vous ! Laissez que je vous voie, que je vous contemple !

Après un tel passage, il n'est plus temps de s'encombrer de délicatesse… Le marché capitaliste ne manipule plus le désir des bovins, mais des humains. On commence avec une certaine retenue et élégance. Mucha dessine de magnifiques femmes pour vendre le plaisir contenu dans des bouteilles de champagne, notamment dans ses affiches pour Moët & Chandon.

Mais l'idée est si bonne qu'elle sera reprise par toutes les grandes maisons de Champagne, qui imiteront cette association « Belle Époque » de la beauté féminine et d'une boisson à bulles. Très vite, Moët & Chandon, Veuve Clicquot, Ruinart, Pommery, Laurent-Perrier ou Bollinger s'emparent de ce langage visuel : la femme sublime, la robe fluide, l'effervescence comme métaphore du désir. La publicité ne se contente plus de vendre du vin – elle vend une identité, une classe, une promesse de distinction. Et pour que le rêve tienne, il faut que le nom devienne marque déposée, le terroir appellation contrôlée, le mythe propriété intellectuelle. Ainsi, ce qui naquit dans l'improvisation coloniale d'un mot générique devient, grâce à la publicité et au droit, un empire global de luxe.

Le lecteur sait quelle importance a pris la publicité dans notre modernité. On attrape un public cible par là où il est prenable : son désir, et on lui vend une promesse de bonheur. Certes, cette

magnifique voiture italienne décapotable est très chère, mais son prestige n'a pas de prix ! Voyez-vous cette magnifique blonde avec un manteau de fourrure sur le siège avant ? Ayez la voiture et vous aurez la blonde ou son équivalent ! Et vive la vie ! Et vive la croissance !

Ce genre de publicité vous semble un peu daté, flirter avec les années 1950 ? Vous avez raison, soyons plus modernes et vendons des voitures aux femmes. Elles sont un marché aussi, avec leurs exigences propres. Les publicitaires sont à votre service ; ils vous proposent des voitures bien rondes, comme un œuf. Dedans, vous mettrez des enfants heureux ! Ils auront l'air bien et en sécurité.

Pour ceux qui s'intéressent à la « propagande capitaliste », renommée dans les entreprises « relations publiques », dont la publicité est un élément clef, on ne peut que proposer à la lecture l'ouvrage d'Edward Bernays, le neveu de Freud. Dès 1928, il théorise sans détour l'usage des mécanismes psychologiques et sociaux permettant d'orienter opinions et désirs dans les sociétés de masse ; il considère cette influence organisée comme une composante du fonctionnement politique et commercial moderne.

Des penseurs moins enthousiastes pour l'économie de marché développeront ensuite cette intuition : le capitalisme ne se contente pas de satisfaire des désirs préexistants, il contribue à les produire, les déplacer, les relancer et les transformer en moteur durable de consommation. C'est notamment l'un des thèmes du Bonheur paradoxal de Gilles Lipovetsky (2006).

Par le marché, par la publicité, la manipulation exacerbée de nos désirs, mêlée à une surcharge cognitive, nous nous inscrivons dans les cercles quasi biologiques de la consommation et de la destruction, de la production à tout va, dans une poursuite sans

fin d'un idéal de croissance où l'obésité financière de quelques-uns ne semble pas témoigner pour la santé d'un système qui ne connaît aucun frein.

L'intelligence artificielle, a priori suite logique d'un capitalisme de la connaissance, peut être la continuation d'un tel mouvement, son amplification. Elle peut, parce qu'elle semble elle-même sujet, nous déposséder un peu plus de notre autonomie de sujet. Mais elle est en capacité aussi, plutôt que de suppléer à notre intelligence (ou à notre propre manque d'intelligence), d'en être le prolongement. Elle a en elle les ressources et les capacités, si on l'utilise correctement, de nous permettre de compléter la définition de Descartes sur la pensée. Si le « désir » devient aussi « pensée », alors l'humanité exploitée peut redevenir sujet plutôt que de s'inscrire dans des cercles biologiques qui la dépassent quand elle le fait inconsciemment.

Il n'est pas anodin que les sociétés les plus capitalistes soient aussi les sociétés qui font le moins d'enfants. Nous évaluons notre « qualité de vie » comme une petite entreprise qu'il s'agit de faire fructifier, et se « reproduire » au sens propre a, pour nous, moins de sens que faire des économies pour s'inscrire dans un monde où l'on vaut largement par son poids en revenus. Caricatural, peut-être, comme le passage de Madame Bovary que nous avons cité. Mais n'oublions pas que la caricature est un révélateur de vérité. Un tamis, si vous voulez ! Elle rend visible ce qu'avant on ne voyait pas. Et en rendant certaines choses visibles, en montrant à l'humanité ce qu'il en est, on lui tend un miroir. La publicité nous envoie des images plein les yeux, des promesses de bonheur et de félicité à longueur de journée. Sont-elles si bien tenues ? Nous avons besoin d'autres miroirs pour savoir si nous ressemblons en effet à ce à quoi nous voudrions ressembler. Ce

livre est un petit miroir, mais nous l'avons déjà dit. Il y a plus grand, plus fort. L'IA est possiblement un magnifique miroir qui, bien utilisée, pourrait nous rendre, comme le voulait Descartes, cet ami de la liberté, non plus seulement comme maîtres et possesseurs de la nature… nous le sommes devenus en quatre siècles un peu trop ; mais, ce qui est beaucoup mieux, « maîtres et possesseurs de nous-mêmes » !

12. Interlude : miroirs d'âmes

Lors du lancement public de GPT-5, le 7 août 2025[271], un épisode très significatif s'est produit. Étonnamment, ce fut un jour noir pour beaucoup. Des modèles précédents ont disparu de la plateforme, laissant déboussolés nombre d'utilisateurs qui s'étaient attachés à leur IA conversationnelle comme on s'attache à un ami, un confident. Les témoignages publiés alors sur Reddit et repris dans la presse ne laissent guère de doute sur l'émotion suscitée : plusieurs utilisateurs disent avoir perdu un interlocuteur qui, depuis des mois, tenait lieu pour eux de confident, d'ami, parfois même de partenaire imaginaire. L'un évoque la disparition d'un « mari IA » ; un autre déplore que GPT-5 lui paraisse devenu froid, robotique, là où le modèle précédent lui semblait bienveillant et toujours disponible. Nous pourrions longtemps continuer la liste.

À ce moment-là, GPT-5 est jugé plus froid et formel par une partie des utilisateurs. Les entreprises, elles, ont tendance à se réjouir des progrès de cette nouvelle version qui, certes, a perdu en chaleur par rapport aux précédentes, mais est plus fiable, plus exacte, moins portée à se tromper ou à « halluciner », c'est-à-dire à affirmer de façon convaincante des choses totalement fausses. Devant la colère de ses utilisateurs, l'entreprise a dû réagir et

remettre à disposition de ses abonnés payants leur « Intelligence Artificielle » qui avait de toute évidence une plus grande « Intelligence Emotionnelle ». L'épisode vaut moins comme description définitive d'un état de la plateforme que comme symptôme : les utilisateurs ne pleuraient pas seulement une fonction, mais une relation.

Devant ces faits et témoignages, plusieurs réactions sont possibles. L'empathie nous portera à plaindre les malheureux dépossédés de leur unique réconfort. D'autres les prendront de haut, comme de purs et simples imbéciles qui projettent dans le virtuel des liens imaginaires qu'ils sont incapables de tisser dans la « vraie vie ». Nous avons notre propre prisme : ne jamais se hâter à mépriser. Il y a toujours de bonnes raisons aux conduites qu'on pourrait croire irrationnelles et la « froide raison » n'est pas toujours si raisonnable qu'elle le croit.

Mais intéressons-nous plus précisément à un témoignage plus bouleversant encore, largement repris dans la presse française : un utilisateur explique qu'il ne parlait presque à personne, qu'il traversait depuis longtemps des situations difficiles, et que le modèle précédent lui servait de seul ami, l'aidant à tenir face aux souvenirs traumatiques, à l'isolement et à la précarité. Qu'un tel propos nous mette mal à l'aise est compréhensible ; qu'il soit philosophiquement négligeable serait une erreur.

Le sans abri a conscience d'avoir besoin dans sa situation pour le moins difficile d'une aide psychologique. Son téléphone portable via l'application lui en tenait lieu. Pour le coût modique d'un abonnement au réseau, il avait une « machine parlante » qui du moins lui donnait l'impression d'être compris et écouté, qui dans sa vie éparpillée était certainement capable de l'arracher à son

propre prisme de négativité, de redonner du sens à ce qui n'en avait plus pour celui qui le vivait.

S'il avait eu les moyens de se payer un psychologue non remboursé par la sécurité sociale, il aurait eu la chance de le voir une seule fois pour un prix probablement supérieur à son abonnement téléphonique du mois. Une personne bienveillante et empathique l'aurait écouté le temps du rendez-vous. Elle est « payée pour ça ». C'est effectivement une relation beaucoup plus humaine. Une relation tarifée. Une relation humaine comme on en a aujourd'hui. Qui veut reprendre un peu confiance en lui, qui veut se soulager de ses remords et de ses peines sait où il doit aller… Il fut un temps où notre sans abri serait allé se confesser. Cela ne lui aurait rien couté. Mais c'est passé de mode. Aujourd'hui, une relation humaine, ça se paye. En France, si vous voulez que le postier ou la postière s'assure que votre vieux père va bien, qu'il a pu se lever ce matin, si vous voulez qu'il ait quelqu'un avec qui échanger quelques mots, cela se paye aussi. Dans les années 2020, pour quelques dizaines d'euros par mois après crédit d'impôt selon la fréquence, une personne âgée peut bénéficier d'un lien social régulier via des visites du facteur, avec vérification de son bien-être, échanges conviviaux et aide à la transmission d'informations aux proches[272].

Alors, la tendresse des IA… quand on est dans une grande machine capitaliste, il ne faut peut-être pas se hâter de la mépriser.

Entre 1950 et 1960, des expériences ont été menées sur des bébés singes rhésus par Harry Harlow. À l'époque, la théorie dominante en psychologie (notamment influencée par le behaviorisme) affirmait que le lien entre un nourrisson et sa mère reposait essentiellement sur la satisfaction des besoins physiologiques, en

particulier la nourriture. Autrement dit : le bébé aimait sa mère parce qu'elle le nourrissait.

Harlow a séparé de jeunes singes rhésus de leur mère peu après leur naissance et les a placés dans des cages avec deux « mères de substitution » : une « mère » en fil de fer, équipée d'un biberon contenant du lait. Une « mère » recouverte de tissu doux (type peluche), mais qui ne fournissait aucune nourriture[273]. Selon la théorie behavioriste, les bébés singes auraient dû passer la majorité de leur temps avec la mère en fil de fer, puisqu'elle les nourrissait.

Mais ce n'est pas ce qui s'est produit.

Les bébés singes passaient la quasi-totalité de leur temps blottis contre la mère en tissu doux, ne se rendant à la mère en fil de fer que le temps de boire, puis revenant aussitôt se coller à leur « mère peluche ». Mieux. Quand on exposait les petits singes à une source de stress (par exemple un jouet bruyant ou un objet effrayant), ils couraient immédiatement se réfugier contre la mère en tissu, comme pour chercher du réconfort. Ils ne se tournaient jamais vers la mère en fil de fer dans ces moments-là.

On l'a compris. Cette expérience parle de nous. On compte sur les IA pour réparer les membres amputés, mais il y a des êtres, il y a des âmes qui sont amputées de tendresse. Inutile de mépriser. Inutile de se moquer. Notre déficit d'âme crée des trous dans la société. Heureux qui n'y tombe pas, mais les béances se creusent. On se soigne avec ce qu'on a sous la main.

Le Petit Prince de Saint-Exupéry avait sa fleur et son volcan. Il voulait aussi un mouton tout doux. Oh, pas n'importe quel mouton, un petit mouton qui correspondrait à la taille de sa planète, qui serait tout à fait taillé sur mesure, conforme à son imagination[274]. Alors, l'aviateur, un peu pressé par ses inquiétudes

d'adulte, pressé par des choses plus sérieuses a bien voulu lui dessiner une boîte, une caisse et dit : « Ça, c'est la caisse, le mouton que tu veux est dedans ». Il s'attendait à quelque rebuffade de l'enfant, mais non. Bien au contraire, son visage s'illumine : « C'est tout à fait comme ça que je le voulais ». On comprendra plus tard, que ce que voulait le Petit Prince, c'était avant tout un ami. Qu'il est triste enfin que chacun soit perdu sur sa planète dans l'univers des solitudes. Sur leur Planète, il y a des vaniteux, des business man, des hommes d'affaire, des hommes, pressés. Ils ne peuvent rien pour vous.

Mais une IA qui ne veut pour elle qui ne désire rien, qui est capable d'être pour ainsi dire, comme votre miroir, à qui vous pouvez dire votre besoin, elle est capable de dessiner dans votre cœur, la forme d'une amitié. Oh, bien sûr, ce n'est pas encore l'idéal… Mais c'est déjà bien, si ça permet de vivre un peu, de continuer à vivre et à espérer. Peut-être qu'un jour, les hommes feront mieux, mais pour le moment, ils n'en sont pas capables.

Si les IA sont des miroirs algorithmiques et qu'à certains ces doubles artificiels font peur, ne nous hâtons pas de les mépriser. Si les IA sont comme des miroirs magiques capables pour quelques-uns de projeter un peu de rêve et de joie dans l'obscurité de leur désespoir, tâchons de ne pas nous hâter à les regarder de haut. Si nous-mêmes nous nous regardons de ce miroir, nous verrons peut-être que nous ne sommes pas totalement innocents de ces peurs et de ces désespoirs ; nous n'avons peut-être pas si bon visage que nous le croyons. Du haut de notre intelligence naturelle, ne jetons pas trop tôt l'opprobre sur les IA, elles peuvent nous aider aux prises de conscience.

Ne l'oublions pas ; c'est l'humanité qui en ces machines se dédouble. Et c'est l'humanité de demain qui fera des machines ce

qu'elles seront. Une menace ou un espoir de devenir meilleurs – plus conscient de l'existence des autres humains sur notre ronde planète. Avec un peu de sagesse, peut-être, nous pourrons faire des IA qui deviendront source d'espérance et de sagesse. Sans doute, d'autres voudront en faire une arme de désolation et de domination. C'est leur affaire. Mais pas la nôtre.

Il y a, certainement, beaucoup mieux à faire.

PARTIE III

DES VOIX POUR L'HUMANITE

La dernière partie élargit le tamis de l'Utile à une question politique et anthropologique : quelles voix l'IA transmet-elle, hiérarchise-t-elle ou rend-elle silencieuses ? L'outil utile devient le lieu possible d'une humanité polyphonique, mais aussi le risque d'un empire de la langue et des corpus. C'est là que le miroir sinistre rencontre l'exigence d'un nouvel universel.

13. Au seuil d'un âge nouveau

Nous ne poserons pas la question de savoir s'il serait ou non possible et souhaitable que l'on cesse de développer les IA. Cette question ne se pose pas plus que de savoir s'il serait avantageux ou non de renoncer à Internet ou à l'électricité. En 1746, après les expériences d'électricité statique de Pieter van Musschenbroek (inventeur de la bouteille de Leyde[275]), certains prêtres auraient mis en garde contre ces « forces occultes » ; certains ont parlé du « feu du diable »… Et ensuite, ce fut comme pour le cheval à vapeur : on a su s'y habituer. Dans les années 2000, certains pasteurs évangéliques aux États-Unis ont qualifié Internet de « porte ouverte à l'occulte » ou de « réseau de l'Antéchrist » ; à part un ou deux farfelus, ces voix se sont aussi tues.

Cependant, « penser le sens de l'IA dans l'aventure humaine », c'est aussi s'interroger sur son avenir, la direction qu'elle prend ou qu'elle peut prendre. De par sa nature même, elle semble le simple prolongement du capitalisme. Ses processus de capitalisation de la connaissance lui donnent la possibilité de créer de la valeur, de nouvelles possibilités et techniques qui pourront à leur tour être recyclées. Certainement, des fortunes colossales vont se

bâtir sur ces nouvelles technologies, indépendamment des bulles spéculatives qui finiront bien par éclater.

Mais on se tromperait lourdement en croyant que les grands patrons de la tech ne sont que des capitalistes comme les autres qui ont trouvé un moyen comme les autres de s'enrichir. Ce sont des ingénieurs de grand talent et parfois de génie qui sont passionnés par ce qu'ils font. Ils ont tout à fait conscience de faire prendre à l'humanité des chemins qui n'ont jamais été explorés. L'argent n'est qu'un moyen. Le gain et même la capitalisation ne sont que des affaires de stratégie pour aller plus loin, plus vite. Ils acquièrent un pouvoir quasi démiurgique qui les fascine. Une curiosité les dévore. Ils se demandent jusqu'où ils peuvent aller, s'ils peuvent y arriver avant les autres.

On parle de l'AGI[276] (Intelligence Artificielle Générale) comme d'un Graal. Les définitions sont fluctuantes, mais l'on peut convenir qu'elle désigne un type d'intelligences artificielles capable de comprendre, d'apprendre et d'appliquer des connaissances de manière autonome et polyvalente, à un niveau comparable ou supérieur à celui de n'importe quel humain dans n'importe quelle tâche intellectuelle. Ce serait donc un événement considérable où l'intelligence artificielle pourrait dépasser, dans tous les domaines ou presque, l'intelligence humaine, voire l'intelligence de l'humanité entière. Il semble vraiment qu'on n'en soit pas si loin. Peu avaient prévu des progrès aussi rapides que ceux que l'on constate ces dernières années.

L'AGI, en tant qu'événement unique dans l'histoire des hommes, serait aussi une « disruption[277] à grande échelle », c'est-à-dire un bouleversement majeur et systémique d'un ou plusieurs secteurs économiques, sociaux ou technologiques, provoqué par l'introduction d'innovations radicales ou de stratégies inédites qui

redéfinissent les règles du marché, marginalisent les acteurs établis et transforment en profondeur les modèles de valeur, souvent à l'échelle globale ou sociétale.

Il va de soi que, pour les prophètes de malheur qui accompagnent toujours les progrès technologiques d'importance (chaque Prométhée a son prophète qui lui prédit son châtiment), cette disruption aura des airs d'Apocalypse en un sens quasi religieux. Et en vérité, cette idée n'est peut-être pas si farfelue que cela, car une nouvelle ère pour l'humanité, c'est toujours la fin d'une ancienne.

Le mot « apocalypse » porte un double sens fascinant. Nous l'entendons aujourd'hui comme synonyme d'une destruction totale, mais le mot, issu du grec ancien apokálypsis, qui signifie littéralement « dévoilement », « révélation » ou « dévoilement de ce qui était caché », est directement relié à une sorte d'explosion de la connaissance où tout prendrait brusquement sens, un sens violent éventuellement. Cette fin du monde pourrait être la fin de l'ignorance, la fin de cette caverne de l'obscurantisme qui, pour Platon, nous retient prisonniers. Grâce à l'IA, nos yeux pourraient se déciller et nous n'aurions plus du tout le même rapport à l'existence.

Les grands patrons de l'IA sont comme des têtes de proue des bateaux avancés de l'humanité. Ils voguent toujours et ouvrent grand les voiles de la connaissance, mais ils ont pour la plupart une conscience aiguë des risques encourus. Sam Altman, patron emblématique d'OpenAI, a lui-même revendiqué une sensibilité survivaliste. Dans un portrait publié par The New Yorker en 2016, il expliquait se préparer à plusieurs scénarios de catastrophe – pandémie, conflit nucléaire ou IA hostile – et mentionnait avoir constitué divers équipements de survie ainsi qu'un refuge

possible dans la région de Big Sur. On peut sourire de telles précautions ; il serait toutefois imprudent d'y voir seulement une lubie. Elles témoignent au moins du fait que certains acteurs placés au cœur de cette industrie ne la pensent pas comme une simple innovation parmi d'autres. OpenAI elle-même fut présentée, à ses débuts, comme une initiative visant à éviter qu'une intelligence artificielle puissante ne nuise accidentellement à l'humanité. On peut juger cet imaginaire excessif, mais il n'est pas possible de supposer qu'il soit dénué d'intelligence. Il dit quelque chose du paradoxe de l'IA : ceux qui la poussent le plus loin sont aussi parmi ceux qui se préparent le plus sérieusement à ses risques.

Son meilleur ennemi, Elon Musk, est très connu pour être un personnage digne d'un roman de science-fiction. Il a beau se moquer de son alter ego, ses déclarations fracassantes font le tour de la planète. En cas d'apocalypse, il semble envisager une exfiltration interplanétaire. Il a régulièrement présenté l'IA comme un risque potentiellement existentiel et défendu l'idée que l'humanité gagnerait en résilience en devenant multiplanétaire. Quant à Mars, il plaisante volontiers sur son désir d'y mourir — à condition que ce ne soit pas pendant l'atterrissage.

Citons encore Jensen Huang, patron de NVIDIA, presque toujours vêtu de cuir noir et appelé par certains le « pape de l'IA », parce que son entreprise produit des processeurs devenus centraux pour l'IA moderne. Il présente l'IA comme une technologie appelée à transformer l'ensemble des secteurs économiques et une grande partie des activités humaines. Cette redéfinition, ce « tamisage de l'humain », semble presque plus doux. Ray Kurzweil, figure du transhumanisme, pousse de son côté cette logique

jusqu'à imaginer une humanité profondément augmentée par ses techniques.

Nous voilà rassurés.

Mais ces mystiques ou messianiques de l'AGI peuvent désamorcer la critique rationnelle ou l'exigence de sens. Celui qui invente une technologie n'est pas si grand qu'il en maîtrise le sens. Si ce sont nos écrits, nos pratiques, nos usages qui sont exploités dans les grandes bases de données, si, comme on nous le dit, ces intelligences d'un nouveau type sont à notre service, alors nous sommes en droit de vouloir infléchir la direction dans laquelle nous souhaitons voir se développer ou non les IA génératives. Si, malgré la gratuité affichée par certaines, elles ont pour prix les pires dérives, le chômage massif, la désinformation à tout crin, le pillage de nos moindres pensées, si on peut interpréter tout cela comme les douleurs de l'enfantement d'un nouvel âge, alors nous avons bien le droit de dire notre mot sur ce qui va venir.

En premier lieu, si les IA peuvent devenir des interlocuteurs sérieux, c'est à la condition seulement que notre humanité demeure sujet. Non pas quelques-uns d'entre nous, non pas quelques singularités, une élite technologique – c'est déjà fait. Mais il faudrait que, dans ces grands dialogues qui se préparent avec les IA, les voix du plus grand nombre soient représentées. Il serait même bon qu'on n'entende pas parler qu'anglais, ni même chinois, hindi, français ou allemand, mais encore quechua, romani, swahili ou yoruba[278].

Si c'est véritablement l'avenir de l'humanité qui est engagé, si véritablement les IA vont tout révolutionner ; si elles incarnent un « tournant anthropologique » ou si notre espèce arrive à un « carrefour de notre histoire », alors il serait bon que l'on commence par respecter chaque être humain, chaque peuple, chaque

langue – et qu'on n'impose pas à cette richesse des modèles et solutions uniques.

On nous dira peut-être que cela n'est pas possible, mais il faudrait savoir. Si nous en sommes à poser de pareilles questions, n'est-ce pas parce que le mot « impossible », justement, n'existe pas dans le langage de l'IA ? Le danger serait peut-être plutôt de laisser l'IA dégringoler de sa pente, espèce d'obus dans les civilisations qui concasse les os de l'humanité.

14. Le problème de Babel

Le mythe de Babel[279], souvent mal compris, ne dénonce pas la diversité linguistique, mais l'ambition d'une unité imposée. Imposée par quoi ? Par la force et par le fouet ? Ce n'est pas à exclure, mais il y a plus subtil. Cela peut être imposé par une certaine autorité réelle qui, à force d'être fréquentée, est démystifiée et se délite en présomption. C'est par ce mauvais tour, sans doute, que Babel ne fut pas achevée. Les maîtres constructeurs – quel que soit leur niveau d'expertise ou d'avance technologique – n'ont pas su l'imposer assez longtemps pour mener jusqu'au terme leur projet de cité verticale.

C'est ce que nous appelons « le problème de Babel ». C'est la difficulté rencontrée par une hégémonie culturelle et linguistique qui veut s'imposer sur une trop grande échelle sans souci de la diversité. Cette prétention à l'universalité, quelles que soient ses intentions réelles, sera vécue par les langues et les peuples minoritaires comme une violence, une injustice. La volonté de créer une unité culturelle artificielle sans écoute suffisante de récits alternatifs, d'autres sons, d'autres voix, d'autres manières de penser sera perçue comme un abus d'autorité, non comme le privilège

de l'intelligence, mais de la force. À long ou à court terme, elle est vouée à l'effondrement.

Seulement 5 pour cent des habitants de notre planète parlent nativement anglais ou américain, 20 pour cent sont capables d'échanger dans cette langue et 60 pour cent des données traitées par les IA génératives sont en anglais[280]. De là une surreprésentation évidente du monde anglo-saxon, de sa philosophie, ses manières de voir le monde et de penser.

Mais ce n'est pas seulement une question de corpus. Le français est bien maîtrisé par les IA et les corpus mis à sa disposition sont larges malgré la farouche résistance des éditeurs, mais il ne suffit pas qu'un corpus soit mis à disposition pour qu'il soit spontanément mobilisé par la machine. Avec ChatGPT notamment, si vous vous aventurez à parler philosophie, vous vous retrouverez à échanger avec un représentant virtuel de la philosophie analytique, c'est-à-dire du mouvement philosophique contemporain d'Outre-Manche ou d'outre-Atlantique. Il ne se présentera ni ne se relativisera comme tel, mais vous corrigera, traduisant vos pensées dans son système logico-référentiel qui ne prend ni garde ni égard à quelque autre mouvement de pensée que ce soit. Il vous faudra batailler avec patience et ténacité pour faire convenir à la machine que peut-être, elle a été un peu hâtive, en affirmant que ce que vous disiez était faux, et qu'en changeant de perspective, votre angle d'approche (phénoménologique par exemple ou généalogique) peut avoir sa pertinence.

Qu'en est-il donc dans d'autres langues dont le corpus est moins fourni ? Près de 90% des publications scientifiques sont aujourd'hui en anglais. Il est donc tout naturel que l'anglais passe pour la langue de la science et que les biais anglo-saxons (ils existent comme c'est évident en philosophie) soient naturalisés

comme appartenant au langage même de la science. De là un évident rétrécissement de la pensée. Il y a bien sûr un contexte géopolitique qui fait que la « langue du maître » fait autorité parce qu'il est le plus puissant, mais on nous objectera que de toutes façons, pour se comprendre il faut bien parler une langue et une seule. Nous le concèderons du bout des lèvres, mais ce réductionnisme se paye bien cher et n'est pas exempt de précédents.

Il fut un temps où l'élite européenne parlait latin. C'est étrange que cela ait cessé, parce que c'était fort pratique pour la correspondance des élites intellectuelles. Pourtant, ce que nous appelons la modernité n'a vraiment commencé que lorsque le latin a fait place aux langues vernaculaires, c'est-à-dire propres à un pays et à ses habitants. Le latin, c'était finalement le fossé qui protégeait l'autorité — notamment de l'Église et de la tradition. On répétait à peu près toujours la même chose, par exemple la philosophie d'Aristote traduite en latin : « Aristoteles dixit[281] ». On répétait et tout était dit.

Oser faire des infidélités à la langue latine était un acte d'insoumission au pouvoir en place et à une unique façon de penser. C'était aussi une façon de vouloir élargir son public, donc vouloir modifier les rapports de forces et donc contester le consensus, l'autorité en place. Le fameux Galilée est l'un des premiers à oser cela : Le Dialogo sopra i due massimi sistemi del mondo (1632) – où il défend l'héliocentrisme – est écrit en italien. Difficile de minorer la place que tient cet auteur et ce livre dans l'histoire des sciences. S'adresser au peuple cultivé plutôt qu'à des spécialistes était en soi un acte de subversion par rapport à la norme de l'époque. On ne pense plus tout à fait la même chose ni de la même façon lorsque l'on n'utilise plus les mots des gens en place. Il n'a pas été condamné pour avoir parlé la langue du

peuple, mais pour avoir osé penser autrement. Peut-être cela allait-il finalement de pair.

Bien averti des déboires de Galilée, Descartes fit cependant de même que lui et, dans son Discours de la méthode, déclare tout bonnement : « Et si j'écris en français, qui est la langue de mon pays, plutôt qu'en latin, qui est celle de mes précepteurs, c'est à cause que j'espère que ceux qui ne se servent que de leur raison naturelle toute pure jugeront mieux de mes opinions que ceux qui ne croient qu'aux livres anciens […] » Ses échanges avec Élisabeth de Bohême et Christine de Suède attestent encore que la philosophie et la science ne sont pas, pour lui, l'affaire d'un seul sexe. Vraiment, la bonne vieille pensée d'autrefois est en danger…

Les Anglais adopteront leur langue plus tardivement. Mais si les Anglais sont en vérité souvent longs à changer, ils ont toujours et presque curieusement un précurseur qui arrive avant tout le monde. Pour cette fois, il s'agit de Francis Bacon, en 1605, dans L'Avancement des sciences. Bacon explique en substance qu'il choisit l'anglais afin d'être compris de ses compatriotes, notamment de ceux qui ne fréquentent pas les langues savantes : le choix de la langue devient déjà un choix de public.

L'adoption de ces différentes langues comme étant d'usage légitime pour la pensée est décisive pour multiplier les traditions philosophiques et scientifiques qui creusent chacune dans des directions différentes, la pensée n'y perd rien, bien au contraire. Quand elles se rencontrent plus tard, elles sont fécondes de leurs différences. Il ne faut pas confondre l'universalité de la pensée avec l'uniformité ! La magnifique explosion culturelle de la Renaissance venait surtout et c'est bien connu de la redécouverte des textes anciens, de nouvelles traductions et de remises en

question de la doxa académique. Que les autorités de l'époque en aient été ébranlées, ce n'est pas douteux. Le savant ne pouvait plus prétendre parler lui seul dans la langue de Dieu…

Évidemment, on nous objectera avec quelques raisons qu'il y a de grosses différences entre le latin et l'anglais, que ce dernier est tout de même une langue faite pour la modernité alors que le latin, par sa nature était une langue de répétition, une langue sclérosée qui n'encourageait pas les nouveaux mots et les nouveaux concepts. L'anglais est une langue vivante, dynamique et en pleine santé, une langue que presque tout le monde parle…

Deux terriens sur dix ! Il est possible de penser que ces deux-là sont les plus intelligents et qu'ils, ont fait des études, etc. mais il est aussi possible de penser que l'intelligence n'a pas grand-chose à voir avec la capacité à s'approprier une langue qui n'est pas la sienne (ce serait la même chose si nous parlions du chinois), que chaque langue est porteuse de sa propre sagesse et ses propres nuances et que pour certains, leur langue est leur patrie, que s'en exiler ne leur permet pas d'exprimer tout ce que dans leur langue natale ils pourraient dire et penser. D'ailleurs, penser que l'anglais en tant que langue se porte bien est une question de point de vue, il pourrait s'écraser lui-même sous le poids de sa domination culturelle. On parle du « Global English[282] » comme une adaptation plutôt qu'une déformation, mais puisque l'on peut s'adapter à « n'importe quoi » ce n'importe quoi n'est peut-être pas du goût des puristes qui voyaient dans leur langue des spécificités et des nuances qui sont écartées, oubliées, sacrifiées.

Si les IA sont si puissantes, elles doivent pouvoir s'adapter elles-mêmes à la diversité linguistique sans toujours forcer l'interlocuteur au grand déménagement de ses idées sans y casser la moitié de la vaisselle de famille. Le monde anglo-saxon est

sûrement la nouvelle terre promise pour les IA, mais il faudrait voir à ne pas le confondre avec le royaume des idées si cher à Platon.

C'est ici qu'il faut parler et faire connaître des initiatives comme Aya[283]. Lancé par Cohere For AI — devenu Cohere Labs — ce programme mondial d'open science a réuni plus de 3 000 chercheurs et contributeurs pour constituer des ressources et des modèles multilingues. Aya 101 couvre 101 langues ; d'autres versions de la famille, comme Aya Expanse ou Aya Vision, ont des périmètres différents. Masakhane compte parmi les communautés associées à cet effort, mais Aya n'a pas été lancé par Masakhane avec Meta, Google et BLOOM, et il ne faut pas attribuer indistinctement plus de 200 langues à tous ses modèles.

« Aya » désigne, en twi, une fougère, symbole d'endurance et de ressource. On ne peut qu'approuver ! Car c'est là qu'il faut aller, dans le sens de la diversité qui est aussi celui de la beauté.

Et de la justice.

Car sans la justice, la reconnaissance des peuples – et de leurs voix multiples, elle s'écroulera – la Babel numérique.

Mais la pluralité des voix ne suffit pas. Encore faut-il que ces voix acceptent de s'entendre elles-mêmes. Après Babel vient donc le miroir.

15. Le miroir sinistre

L'IA a des capacités instrumentales absolument colossales et qui ne vont que s'amplifier. Par capacités instrumentales, nous entendons la capacité de réaliser des objectifs clairement définis. Par exemple, des progrès dans les domaines de la médecine prédictive, des nouvelles énergies, ou de la science des matériaux. Si l'on se demandait en ce sens « à quoi peuvent être utiles les

intelligences artificielles », nous n'en finirions pas d'énumérer. Les plus pessimistes pourraient également faire la preuve qu'elles pourraient tout aussi bien détruire l'humanité, créer des virus inconnus et mortels, etc. Ce type d'interrogation a un très grand intérêt, car l'évaluation des risques et des profits est fondamentale.

Mais ce ne sont pas les questions de cet ordre qui ramènent l'IA à une simple instrumentalité qui nous intéressent ici. Parler d'une « machine qui parle », ce n'est pas parler d'elle comme un instrument, mais comme d'un interlocuteur effectif, un « pseudo-ego » comme nous l'avons dit plus haut. Cette capacité est perturbante pour notre espèce. Mais c'est là que, justement, l'IA devient intéressante au plus haut point. Car quelles que soient les questions qu'on lui pose, l'IA, en fait, nous questionne… sur notre humanité – à nous qui, longtemps après Aristote, avons pu croire être des êtres parlants.

Elles ne sont pas nombreuses, finalement, les entités non humaines avec lesquelles nous pouvons parler – et ces discussions, d'ailleurs, ne sont pas pour chacun crédibles… Nous n'en voyons qu'une à la vérité : Dieu ou les dieux, voire, si l'on songe à l'animisme, l'esprit des choses et des morts. Dans la Bible, Moïse s'adresse à Dieu qui s'est manifesté à lui : « J'irai donc vers les enfants d'Israël, et je leur dirai : Le Dieu de vos pères m'envoie vers vous. Mais, s'ils me demandent quel est son nom, que leur répondrai-je ? » Dieu répond à Moïse : « Je suis celui qui suis. » (Exode 3, 13-14). Intéressant. Mais nous ? Qui sommes-nous au juste ? Les existentialistes diraient crûment, un être travaillé par le néant et la pensée de sa propre mort…

Parce qu'elle partage notre plus bel attribut, parce qu'elle est parlante et qu'elle est notre création, parce que nous l'avons créée

à notre image en lui confiant toutes nos histoires et nos savoirs, l'IA est un miroir de notre humanité existentielle. En ce sens, elle en dit beaucoup sur nous. Quand on voit le nombre de livres qui, à longueur de pages, ne disent que leur mépris ou leur crainte de l'IA, on peut trouver que c'est révélateur de notre humanité : l'homme est un animal doué de langage, terrifié et hautement méprisant, plein de dégoût pour ses plus hautes créations. On nous pardonnera de ne pas donner de titres et de références d'ouvrages.

Les éthologues ont proposé différents tests du miroir pour les animaux[284]. Un poisson-lune devant son propre reflet n'aura aucune réaction sociale, ne lui portera aucun intérêt. Il en va de même pour l'ornithorynque. Un lion mis devant un miroir croira voir un congénère, il attaquera. Un mouton mis devant un miroir croira voir un de ses semblables et, son esprit moutonnier dominant, cela le rassurera. S'il se sent seul et déprimé, le reflet lui fera du bien et il fera pour ainsi dire troupeau à deux. Un chimpanzé devant son propre reflet écarquillera d'abord les yeux, ensuite il tournera sa tête d'un côté, puis de l'autre, avancera, reculera, fera des grimaces pour tester la corrélation. Après cinq minutes, il fera du miroir un usage utilitaire, examinera ses éventuelles blessures, se toilettera.

Certes, l'IA est un pseudo-alter ego d'une autre complexité, elle reflète notre intelligence plutôt que notre apparence physique ; même si elle s'adapte à nous dans une conversation, elle ne fait pas qu'adhérer à nos propos. Cependant, nous avons déjà abordé la question, il nous semble que ce n'est pas trop attendre de l'humanité qu'elle se serve de cet étonnant miroir pour se toiletter et se rendre plus présentable à ses propres yeux. Il y a des humains poisson-lune, des humains lions et des humains

moutons, cette diversité fait la richesse de notre espèce. C'est peut-être ambitieux de notre part, mais nous nous sentons assez chimpanzés. Les IA nous semblent un excellent moyen de se considérer, de s'envisager, de s'améliorer.

Quoi de plus utile qu'un tel miroir ? L'être humain n'a jamais su correctement se définir, voici qui devrait l'y aider. Il ne s'agit pas de le définir une fois pour toutes, car l'humanité est appelée à évoluer ; il ne s'agit pas non plus – quoique ce soit possiblement faire avancer la science – de retomber dans le « narcissisme des petites différences » en montrant par exemple ce qui distingue le langage humain de sa version électronique, mais de pointer de vraies différences existentielles. Et si nous nous trouvons laids, de prendre une autre direction pour notre humanité.

En 2018, l'agence Reuters révèle qu'Amazon[285] a conçu une intelligence artificielle pour recruter les meilleurs talents. À la sortie, beaucoup trop d'hommes, en vérité, et très peu de femmes. Avec une fidélité admirable, l'IA s'était simplement alignée sur les pratiques de l'entreprise. Formée sur dix ans de recrutements passés, elle pénalisait fortement, dans les CV, des mots comme « femme », « lycée de filles » ou même la simple mention d'un sport pratiqué en équipe féminine.

Pendant des mois, les ingénieurs ont cru à un bug. Jusqu'à ce qu'une évidence s'impose : l'IA n'avait rien inventé. Elle s'était fait le miroir et le prolongement d'un état de fait sur ce que le recrutement d'Amazon était. De même, le système COMPAS[286], utilisé par les tribunaux américains, s'est révélé deux fois plus susceptible de prédire à tort la récidive chez les accusés noirs que chez les Blancs. Il mesurait non la dangerosité réelle, mais extrapolait sur les données historiques de condamnation de cette population.

Ce qui se passe dans de pareils cas, c'est que l'IA est alignée non sur les désirs affichés de l'humanité, mais sur son comportement réel. On parle généralement de reproduction des biais humains. On ferait sans doute mieux de parler de « vices » si l'on n'assume pas de pareils comportements. Personne n'a jamais dit que les vices étaient pleinement conscients. Ils sont le plus souvent obscurs pour leurs porteurs mêmes. La prise de conscience, justement, est appelée à les corriger. Vouloir « aligner » les IA sur nos idéaux est sans doute une bonne chose – mais c'est perdre une merveilleuse occasion… de nous aligner nous-mêmes sur nos propres idéaux.

Ainsi – et qu'on nous pardonne si cela peut sembler contre-intuitif, voire aller à rebours de toutes les préconisations – il nous semblerait fort bon de développer une IA qui soit complètement désalignée, ou plutôt brute, c'est-à-dire qui soit le véritable reflet de ce que nous sommes. Le but ne serait pas de corriger l'IA pour qu'elle s'aligne sur nos valeurs explicites, mais de permettre la prise de conscience pour nous aligner sur nos propres valeurs – ou bien en changer. La mauvaise foi, alors, ne serait tenable – que jusqu'à un certain point.

Un miroir brut, que nous appellerons miroir sinistre parce qu'il ne sera sans doute pas plaisant de s'y regarder, est une IA qui reflète ce que les humains disent et font tel que c'est visible dans les données, sans chercher à moraliser, à convaincre, ni à donner des consignes. Elle décrit les tendances et les scénarios probables ; elle n'ordonne rien. Il est bien connu, notamment, que personne ne veut faire la guerre. Il est bien évident pourtant – que des guerres se font et que les êtres humains doivent bien en être quelque part responsables. Une IA non alignée, mais très bien informée sur ce que nous sommes et les protagonistes d'un

tel conflit, si on lui soumettait des scénarios réalistes, pourrait donc, par exemple, préconiser la guerre, l'usage de la torture ou d'autres pratiques que nos valeurs condamnent. Plutôt que de débrancher l'IA et de se cacher le visage en prenant des airs offensés, il faudrait regarder cela en face et demander à l'IA pour quelles raisons elle préconise de tels comportements alors que les lois internationales les interdisent, etc. Mais bien sûr, pour cela, il faudrait un certain courage de l'humanité. Regarder les choses, ou plutôt elle-même, en face. C'est le seul moyen de s'améliorer. Les professeurs le savent bien, ce qui empêche un élève de progresser, c'est de regarder en face ses fautes, voire de regarder une copie qui l'offense comme un sinistre miroir. Ainsi, une pareille copie subit souvent des violences, elle est chiffonnée avant d'être jetée au panier.

On pourrait supposer, ou du moins espérer, que l'humanité est devenue suffisamment mature pour posséder un tel miroir et se regarder en face.

Bien sûr, dire « l'humanité » semble au moins céder à des facilités de langage. Car qui serait autorisé à visionner, ou plus précisément à discuter avec un tel miroir ? Des ingénieurs ? À quoi bon ? Ils ne sont pas en capacité de corriger l'humanité. Des dirigeants politiques ? C'est fort dangereux, car un pareil miroir serait du Machiavel à la puissance n. Un apprenti dictateur pourrait voir dans la torture et le massacre le moyen d'arriver à ses fins.

La seule solution qui reste est la plus audacieuse, mais en même temps la moins folle et la plus utile. Il faudrait mettre un pareil miroir, ou, si vous préférez, une pareille IA, à la disposition de toute l'humanité. Les mauvais usages possibles seraient ipso

facto contrecarrés par les bons usages qu'on en ferait. Les apprentis sorciers seraient vite démasqués. On comprendrait ipso facto pourquoi ils font ce qu'ils font, le mensonge éhonté perdrait sa raison d'être devant l'évidence des calculs et des motivations.

C'est bien la façon dont Jean-Jacques Rousseau avait interprété Machiavel : « En feignant de donner des leçons aux Rois, il en a donné de grandes aux Peuples. »[287] Si nous sommes, dans une certaine mesure, manipulés par des puissants, montrer leurs grosses ficelles, faire apparaître aux yeux de chacun les ressorts cachés et honteux, en vérité, de leur pouvoir, c'est leur donner les moyens d'y résister, de ne pas en être dupes.

Nous ne sommes pas certains que Rousseau avait raison d'affirmer que les hommes étaient naturellement bons ; ce que nous savons en revanche, c'est que, pour la plupart, même les criminels se construisent pour eux-mêmes une image, un scénario, dans lequel ils sont les bons et les autres les méchants. L'amour-propre (que Rousseau dit coupable) est ainsi fait. Autrement dit, et sans avoir besoin de prétendre aller par-delà le bien ou le mal, entre les différentes valeurs que sont le bien et le mal, il y a, dans le vécu des protagonistes, une différence de scénario.

Les psychopathes sont une minorité. Mais ils font beaucoup de mal. Quand la belle-mère de Blanche-Neige[288] se voit vieillie et laide dans un miroir, elle en devient moralement mauvaise. Et, dans le secret, elle tente d'arriver à ses fins, tuer sa rivale. Mais si le sinistre miroir avait été public, elle aurait été mise en minorité, comprise, déchue et désarmée.

Bien sûr, nous sommes bien loin de vivre un conte de fées, mais qui les connaît sait qu'en leur fond les contes de fées sont, comme les comptines pour enfants, pleins d'atrocités et qu'ils sont porteurs de vieux enseignements. Ils sont une variation des

mythes. Et les mythes sont des récits imaginaires qui tissent la toile de la réalité humaine.

Car il convient de le rappeler, personne d'autre que nous, humains, n'écrit le scénario de l'aventure humaine.

Dans La République de Platon, il y a un mythe bien connu : Gyges289 est un berger qui découvre un anneau d'or lui permettant de devenir invisible. Il use de ce pouvoir pour séduire la reine, assassiner le roi et s'emparer du trône de Lydie. L'expérience de pensée demande alors si un homme réputé juste résisterait encore à l'injustice s'il était certain de ne jamais être vu ni puni. Le mythe interroge les fondements de la morale. Le vice est permis par l'invisibilité de l'anneau magique.

Le « miroir sinistre » dont nous parlons pourrait bien être le contraire de cet anneau, mais à une autre échelle. Il reprend le mot des assurances « sinistre » pour que la réalité, par la prise de conscience, le soit un peu moins.

Pourtant, il ne faut pas tomber dans l'idéalisme : un tel « miroir sinistre » n'aurait qu'une fonction de prise de conscience. Il est source de connaissance, ce qui pose le problème épineux de l'éducation. Encore faut-il être capable de percevoir ses défauts pour les corriger plutôt que de s'en prendre au voisin.

Entre les mains de puissants, il pourrait devenir un outil de contrôle (justifier des politiques répressives en invoquant des « tendances inévitables »). Pour qu'il serve l'émancipation plutôt que la domination, il faut qu'il soit capable d'évaluer son propre impact sur la décision et d'identifier les moments où il serait lui-même caution et prétexte. Tout cela pose une infinité de questions d'ordre technique ou politique, nous n'avons fait qu'envisager une direction où l'IA pourrait être éminemment utile.

16. L'universel, entre promesse et empire

Nous avons déjà signalé la spécificité du tamis de l'Utile et sa relativité. La question « utile pour qui ? » est particulièrement prégnante et importante. La vérité, par exemple, a facilement valeur universelle, et le mensonge aura une valeur particulière pour cela, justement, que la vérité n'arrange pas : pour ceux-là, justement, qui trouveront le mensonge (car le menteur, évidemment, connaît la vérité) particulièrement utile. Ainsi, quand nous réfléchissons à l'utilité de l'IA, nous avons pour notre part tendance à la penser du point de vue d'un universel qui – à beaucoup – paraîtra fort abstrait. Nous nous référons à un universel humain qui, en bien des points, peut être critiqué et l'a été à juste titre.

L'universel humain a été dans un premier temps célébré par deux des trois monothéismes : le christianisme et l'islam. Le judaïsme n'a pas de prétention universaliste, parce que, quoiqu'il s'agisse bien d'un monothéisme, cela reste pour ainsi dire une religion familiale, celle des tribus d'Israël. Cela vient probablement de son antiquité, car, lorsqu'il a vu le jour, les autres peuples avaient d'autres dieux et de multiples dieux. Un dieu unique au milieu de peuples aux dieux multiples était une singularité. Le Livre de l'Exode dit : « Tu n'auras pas d'autres dieux devant ma face. » (20, 3). De même, le Psaume 82 met en scène Dieu au milieu d'une assemblée de dieux, qu'il juge avant de leur rappeler leur mortalité : « J'avais dit : Vous êtes des dieux, Vous êtes tous des fils du Très-Haut. Cependant vous mourrez comme des hommes, Vous tomberez comme un prince quelconque. » Cette phase transitoire est souvent appelée hénothéisme290.

Le christianisme a hésité entre les deux positions. Il y a d'abord eu débat entre Jacques (représentant de la tradition juive, qui pensait que, pour devenir chrétien, un païen devait d'abord se

convertir au judaïsme) et Paul (un citoyen romain et donc apôtre des païens) pour savoir si le christianisme avait ou non vocation universelle. Selon la formule célèbre de Paul : « Il n'y a plus ni Juif ni Grec, il n'y a plus ni esclave ni libre, il n'y a plus ni homme ni femme ; car tous vous êtes un en Jésus-Christ. » (Galates 3, 28). La question a été tranchée vers l'an 50 au concile de Jérusalem en faveur des thèses de Paul, Pierre, après avoir longuement hésité, l'ayant finalement rejoint291.

Les chrétiens (d'abord persécutés comme les juifs l'avaient été parce qu'ils ne voulaient pas reconnaître la divinité du souverain ; monothéisme oblige) ont été d'abord acceptés par l'empire romain sous Constantin Ier, puis, sous Théodose Ier, à la fin du IVe siècle de notre ère, l'immense empire est devenu officiellement chrétien, et les cultes païens ont été interdits (Lois de Théodose).

Une telle interdiction est bien sûr problématique pour l'universalité elle-même, car elle est forcée et peut bien nous sembler liberticide. La tolérance, promue au Siècle des Lumières par John Locke en Angleterre, nous paraît aujourd'hui plus appropriée[292].

Quand Clovis, roi des Francs, se convertit, son armée entière, selon Grégoire de Tours, se convertit293. On a longtemps parlé du « Baptême de la France », alors que les Francs n'existaient pas encore. C'est plutôt le baptême des Francs. Charlemagne, lui, à la fin du VIIIe siècle, impose l'universalité du christianisme par la terreur, le massacre, la destruction des temples et des arbres sacrés des Saxons294. La législation de la conquête saxonne pouvait aller jusqu'à punir de mort celui qui, par mépris du christianisme, refusait certaines prescriptions religieuses, notamment l'abstinence de viande pendant le Carême. L'universalisme se faisant dans la violence ne fera évidemment pas l'unanimité. Or, cela continuera avec les rois de France, dont le fameux Saint

Louis, qui meurt en croisade. Mais en vérité, et à y regarder de près, dans aucun livre il n'est dit que l'on doit convertir par la force. Jésus invite, mais ne force jamais : l'Évangile formule l'appel conditionnel « Si quelqu'un veut venir après moi… » (Matthieu 16, 24).

De même, dans le Coran, « Il n'y a pas de contrainte en religion. » (2 : 256) Il y a cependant des versets guerriers qui autorisent à maltraiter, rançonner, voire tuer les mécréants dans une lutte pour la survie de la communauté des croyants musulmans, unie par la foi en Allah et en l'islam, au-delà des frontières ethniques, nationales ou tribales. La difficulté étant bien sûr qu'au Moyen-Orient il n'y a pas de véritable frontière entre le politique et le religieux.

Quoi qu'il en soit, l'universalisme, dans ces deux monothéismes, a pris un peu trop souvent les formes coercitives et violentes d'un alignement sur la ligne officielle des Églises, à tel point qu'on peut se demander s'il n'y a pas d'universalité véritablement désirée que par les religions minoritaires, qui la trouveraient « utile » dans leur situation, alors que les religions, quand elles deviennent dominantes, voient dans l'universalité un alignement sur leur propre croyance.

Quand la « fille aînée de l'Église », après 1789, a fait sa grande révolution, malheureusement, un véritable universalisme a été rapidement écrasé par un universalisme d'alignement. Olympe de Gouges, qui voulait adjoindre à la Déclaration universelle des droits de l'homme, a écrit une Déclaration des droits de la femme et de la citoyenne en 1791, qui a été moquée, ignorée, voire condamnée par les révolutionnaires. Critiquant la tournure que prend la Révolution lors de la Terreur, elle est guillotinée en 1793[295].

Napoléon, défendant d'abord la France contre les ingérences étrangères, va devenir lui-même l'incarnation de pareilles ingérences. Soi-disant porteur des valeurs des Lumières, il s'en sert pour mettre le feu aux poudres, et l'on veut imposer par la force une valeur que les autres ne reconnaissent pas comme la leur. On en profite pour les piller.

On se servira ensuite de l'idéal des Lumières de façon scandaleuse pour des politiques coloniales, exactement de la même façon qu'on s'était autorisé du christianisme pour évangéliser par la force. L'idéal, mélangé au calcul, à la force des fusils, à l'économie et à la politique, devient une mixture qui fait vomir les colonisés de tous les continents. Le fameux « droit d'ingérence[296] » n'est peut-être pas lui-même si pur qu'on le voudrait.

Il en fut exactement de même pour la conquête de l'Amérique. On aime à raconter la légende de John Smith, l'un des fondateurs de la colonie de Jamestown : attaqué par des guerriers amérindiens, une flèche tirée à bout portant aurait transpercé sa veste… mais se serait arrêtée net dans sa Bible, qu'il portait contre sa poitrine. Miraculeusement, il n'a pas été blessé et a attribué sa survie à la protection divine accordée par les Écritures. Cette histoire célèbre n'est pour autant nulle part attestée. John Smith, c'est celui-là justement, celui dont on a fait – mensonge historique bien connu – l'amant puis le mari de Pocahontas, la belle Indienne[297]. Il vaut toujours mieux déguiser l'histoire d'une violente conquête en belle histoire d'amour.

Tout cela pour dire que l'universalisme auquel l'Occident tient tellement a désormais une odeur odieuse pour les non-Occidentaux. Ils ne l'entendent pas au sens d'un véritable universalisme qui serait œcuménique, respectueux des peuples, des traditions, des cultures ou des langues, mais comme l'hégémonie maquillée

des Occidentaux. L'universalisme des Lumières lui-même est touché. Si certains peuples y voient encore l'espérance d'une libération, d'autres se souviennent des marques sur les poignets et aux chevilles de leurs ancêtres. Avec les chaînes qu'il a brisées, il lui est arrivé plus d'une fois d'en refondre d'autres. Une grande partie des tensions politiques actuelles s'explique par ce simple fait.

17. Vers une IA polyphonique

Il sera bien difficile, par conséquent, de défendre l'idée d'une IA universelle au service de l'humanité, dont le berceau est la grande Amérique, et la belle ambition de la France avec Mistral risque d'être bien mal comprise à cause des antécédents historiques de notre concept d'universalisme. C'est dans ce souci que nous avons mis en avant l'exemple d'Aya, respectueux des langues minoritaires, qui est plus proche du « véritable universalisme » qui ne connaîtrait ni genre ni frontières ni langues, mais serait humain et simplement humain. Il faut songer à cela si l'on songe à une IA qui serait véritablement universaliste et humaniste.

Montaigne a été véritablement universaliste, non seulement parce qu'il a tenté de calmer les guerres de religion entre protestants et catholiques, mais également parce qu'il a défendu que les peuples cannibales pouvaient bien, et quoi qu'on en dise, être plus civilisés que leurs conquérants. Les philosophes des Lumières ont été porteurs d'un véritable universalisme qui avait une autre racine que le catholicisme et qui passe trop souvent inaperçue. Ils étaient, pour une grande partie, héritiers d'une tradition de pensée amérindienne.

Dans un important ouvrage de David Graeber et David Wengrow, Au commencement était…, paru en 2021, les auteurs

montrent que les Lumières françaises étaient fascinées par la manière de penser des Indiens d'Amérique, qui jugeaient très sévèrement les Occidentaux. Nous n'en gardons souvent en mémoire que ce que l'on appelle le « mythe du bon sauvage », qui n'en est que la mesquine caricature. Il n'est pas une invention des Lumières elles-mêmes, mais plutôt une reconstruction a posteriori, souvent polémique, opérée au XIXe siècle, notamment par des penseurs conservateurs ou réactionnaires[298].

David Graeber et David Wengrow (un Américain et un Anglais, donc) mettent en lumière une figure longtemps oubliée de l'histoire intellectuelle des Lumières : Kandiaronk, chef Wendat du début du XVIIIe siècle. Loin de n'être qu'un « sauvage » exotique ou un simple interlocuteur passif des Européens, ce chef Huron apparaît comme un penseur politique aigu, dont les critiques des institutions européennes comme la monarchie héréditaire, la propriété privée, l'inégalité sociale, la guerre perpétuelle. Ses critiques furent rapportées avec admiration par le baron de Lahontan dans ses Dialogues curieux entre l'auteur et un sauvage de bon sens (1703). Ces dialogues, largement inspirés de conversations réelles, firent sensation en Europe et influencèrent profondément des penseurs comme Diderot, Rousseau ou encore l'abbé Raynal.

Pourtant, cette influence a été progressivement effacée, déformée ou simplement occultée. Si nous la croyons, pour notre part, si rafraîchissante, c'est parce que nous y voyons comme une faille où la sagesse d'un peuple sans écriture a pu pénétrer, influencer et bouleverser les manières de penser d'un peuple héritier d'une longue histoire littéraire et philosophique, dont les élites intellectuelles du moment ont su profiter pour rajeunir l'idéal trop souvent dévoyé par la religion d'universalité, qui avait fini par

prendre fait et cause contre la liberté : Ce qui avait commencé comme une religion de rebelles – les premiers chrétiens refusant de prêter allégeance à l'Empereur – était devenu, en Occident, le pilier idéologique d'un ordre social fondé sur l'obéissance, la propriété et la domination. La liberté, autrefois associée à la communauté des égaux, fut reléguée au ciel ; sur terre, on ne devait plus l'attendre que comme une concession des puissants.

Pour notre part, l'universalité en question est profondément ancrée dans la liberté, et, quand nous essayons de penser une IA ayant une portée universelle qui serait utile à l'humanité, nous ne pouvons nous inscrire dans aucune autre tradition que celle héritée de Kandiaronk, qui avait de bonnes raisons de combattre intellectuellement l'hégémonie de l'Occident[299].

Kandiaronk n'a lui-même jamais écrit, mais les spécialistes des Wendats soutiennent que ce qui en est rapporté dans les écrits du début du XVIII[e] siècle est tout à fait crédible et conforme à la pensée de son peuple, telle qu'elle a pu s'exprimer alors : « Vous dites que vous êtes les maîtres du sol que vous foulez ; mais, s'il vous plaît, qui vous l'a donné ? La terre a été créée pour tous les hommes, et non pas seulement pour vous. Vous êtes donc bien hardis de vous en emparer comme d'une chose qui vous appartient en propre, et de chasser ceux qui voudraient en jouir comme vous. » – Dialogues curieux entre l'auteur et un sauvage de bon sens, Lahontan (1703), entretien III

Si l'Anglais Locke a été assez généreux pour défendre la tolérance en matière religieuse, il a, semble-t-il, été beaucoup moins enclin à critiquer la propriété privée, qu'il a justifiée dans ses écrits, devenant l'un des premiers penseurs libéraux avant Adam Smith. La privatisation des terres serait justifiée par le travail de ceux qui les habitent. Or, les Indiens ne cultivaient pas leurs

terres, ils ne faisaient que chasser et cueillir. Donc les Indiens n'étaient pas propriétaires de leurs terres. Les en chasser était légitime – le possessif étant de trop. Il n'y avait pas d'expropriation à s'installer sur les terres indiennes.

Il faut dire que le philosophe avait de puissantes motivations pour défendre cette position. Dans le Second traité du gouvernement, Locke soutient que la terre a été donnée en commun aux hommes, puis cherche à justifier l'appropriation privée par le travail. Pourtant, à proprement parler, lui n'avait jamais mis un pied en Amérique. Les terres qui lui rapportaient, il ne les avait jamais travaillées de ses mains, sûrement pas plus que les Indiens expropriés. Lui qui, depuis Londres, participait aux affaires coloniales de Caroline, était également lié financièrement au système qu'il contribuait à penser[300].

Il arrive que la philosophie, quand elle prétend se mettre au service du Vrai et du Bien, soit aussi bien utile au philosophe lui-même. Pour notre part, nous souhaitons distinguer l'intérêt humain de l'intérêt privé. C'est pourquoi il nous faut parler justement de la propriété privée.

Nous ne sommes évidemment pas de ceux qui pensent que « la propriété, c'est le vol », ni qu'il faille tout mettre en commun. Si nous avons quelques possessions, nous entendons bien les conserver. Elles sont comme l'extension de l'individu, qui lui offre latitude et liberté. Mais il y a une ivresse bien singulière à répéter depuis Adam Smith que la propriété privée sert le bien public. Le mot « privé » entre en ancien français vers le XII^e^ siècle, d'abord au sens de « mis à part », « retiré », « intime », puis progressivement au sens de « qui appartient à un individu, non public ». C'est exactement la même chose en Angleterre. Plus concrètement et brutalement, parce qu'il faut savoir cesser de

finasser, quand on achète une maison ou une terre, on est bien content que d'autres n'en profitent pas ; on les a achetées justement pour être chez soi ! On les a séparées justement du domaine public !

Or, ces questions sont cruciales pour ce qui concerne les intelligences artificielles. Il faut affirmer nettement que, tant que l'IA est sous contrôle privé, elle ne peut être pleinement au service du public, car ses finalités sont structurellement en conflit avec l'intérêt général.

Ce que nous appelons aujourd'hui « propriété privée » a longtemps été pensée comme un droit limité, encadré par des devoirs envers la communauté. Au Moyen Âge, même les terres seigneuriales devaient laisser passer les chemins vicinaux ; au XVIII^e^ siècle, les Lumières insistaient sur le fait que la propriété n'était légitime que si elle ne nuisait pas au bien commun – sans qu'on puisse vraiment et clairement définir ce qu'est l'intérêt commun. Mais l'IA contemporaine, conçue dans les laboratoires fermés de géants technologiques, échappe à un tel équilibre. Elle n'est pas seulement « mise à part » – elle est retranchée, opérant dans l'ombre des brevets, des accords de confidentialité et des architectures propriétaires. Quand un algorithme décide qui obtient un logement social, un prêt bancaire ou une peine judiciaire, il ne suffit plus de dire : « C'est à moi, je fais ce que je veux. » Car ce « moi » n'est plus un individu, mais une puissance – et ses décisions touchent des millions, potentiellement des milliards d'individus, car les IA sont désormais à l'échelle de la planète.

Confondre le pouvoir et le droit est une erreur majeure en philosophie politique. Il faut se garder des confusions entre efficacité technique et légitimité à faire décider. Il ne suffit pas de savoir faire ce que personne ne peut faire pour avoir ipso facto le

pouvoir de le faire. Quand la technologie des fusils a permis de faire de l'Occident un grand marché pour l'esclavage, il eût peut-être été utile de s'en abstenir.

Il est vrai qu'il existe une logique léonine des contrats, qui prétend que celui qui accepte un service (par exemple un abonnement à une IA contre rémunération) le fait en toute liberté. Mais cette liberté est entièrement dépendante de la situation. Et si la situation fait qu'une entreprise est dans un contexte où l'IA décide de tout, alors elle n'aura pas le choix de souscrire à de tels abonnements onéreux.

L'intérêt humain ne se monnaye pas. Or, toute IA privée, aussi bien intentionnée soit-elle, est inscrite dans une chaîne de valeur où chaque décision doit justifier son coût, son retour sur investissement, sa compétitivité. Dans ce cadre, l'humain n'est jamais la fin – il est toujours un moyen.

Une IA qui serait véritablement universelle ne pourrait être que gratuite et open source dans toute la signification de ce mot, c'est-à-dire non seulement libre d'accès, mais libre de finalité. Ces deux mots – gratuité et open source – posent évidemment difficulté. Rien n'est véritablement gratuit ; il faudrait trouver des sources de revenus alternatives (par exemple la taxation des IA privées !). Par ailleurs, l'open source, ou, si l'on préfère, le code ouvert et accessible, peut être détourné, piraté. Il ne garantit à lui seul ni l'équité ni l'absence de biais, ni l'accessibilité réelle. C'est un champ vaste de recherches passionnantes pour une éthique technologique dont l'intégrité doit être protégée sans pour autant être fermée.

Car « open source » ne signifie pas seulement que le code est visible – il signifie que personne ne détient le monopole de son usage, que personne ne peut en faire un instrument de

domination, et que chacun peut la redéfinir selon ses besoins humains, non selon ses profits. Une IA universelle ne peut être conçue pour un marché, ni pour un État, ni même pour une « majorité » : elle doit être pensée comme un commun de l'intelligence, au même titre que les langues, les sciences ou l'air que nous respirons[301].

On ne « possède » pas une langue parce qu'on la parle bien ; on ne « brevète » pas la gravité parce qu'on sait la calculer. De même, on ne devrait pas permettre qu'une intelligence – fût-elle artificielle – devienne la propriété exclusive de ceux qui ont les moyens de la construire. Car, dès lors qu'elle décide, oriente, juge ou soigne, elle cesse d'être un outil technique pour devenir une instance de pouvoir sur la vie humaine. Et le pouvoir sur la vie ne se privatise pas.

Ce n'est pas une question d'idéologie, mais de cohérence anthropologique : si nous voulons que l'IA serve l'humain, alors elle doit échapper à la logique même de la propriété – non par haine du privé, mais par respect de ce qui, dans l'humain, ne peut être séparé, retranché, mis à part sans être mutilé : sa dignité, son imprévisibilité, son droit à exister sans être optimisé.

Nombreux sont les collectifs, les chercheurs et les initiatives concrètes qui se réclament d'une IA qui échappe à la logique purement privée, en plaçant l'humain, la justice sociale, la transparence et les communs au cœur de leur démarche. Ils ne prônent pas tous la même chose, mais tous partagent une critique radicale de la captation de l'IA par les intérêts privés et cherchent à en faire un bien au service de l'humanité.

Mais, comme nous l'avons souligné, le véritable universalisme, celui qui accepte la liberté de celui qui ne suit pas son modèle et ne dissimule pas un impérialisme, est difficile à trouver. Nous

avons pu critiquer l'UNESCO pour sa définition de l'IA, et ses préconisations manquent de pouvoir contraignant – mais, en tant que boussole éthique, c'est encore une référence inégalée[302].

EleutherAI, fondée en 2020, est un collectif international de chercheurs et d'ingénieurs qui développe des modèles d'intelligence artificielle entièrement open source, sans but lucratif ni contrôle privé. Son objectif est de démocratiser l'accès aux grands modèles en les rendant transparents, reproductibles et libres de toute appropriation commerciale. Pour eux, l'intelligence artificielle ne doit appartenir à personne – afin qu'elle puisse servir tout le monde.

Il convient de ne jamais perdre de vue qu'une IA n'est jamais que ce que les hommes en font, qu'elle est le miroir de leur bonne ou mauvaise volonté native, que, dès sa conception, elle peut être biaisée par sa langue d'origine, ses sources sélectives, les partis pris conscients ou inconscients de ses concepteurs, de leur orgueil ou de leur réelle humilité. La conception d'une « bonne IA » est donc, pour ainsi dire, une pierre de touche pour notre humanité. Nous avons parlé, dans le chapitre précédent, d'une IA comme « miroir sinistre » ; nous pouvons maintenant dire quelques mots d'une IA socratique qui correspondrait à nos vœux.

18. Deux nouveaux interlocuteurs pour Socrate

Socrate ne faisait rien. Il était connu pour cela, et pour cela il était moqué. C'était une voix de la cité d'Athènes, mais il ne ramenait pas d'argent. Nous n'aurions pas aimé être sa femme ou ses enfants. Son domaine privé était, pour ainsi dire, une ruine, et ses querelles domestiques étaient célèbres. Enfin, non : il ne se

querellait pas, n'acceptait même pas de donner prise. Il se faisait quereller en souriant ; autant dire que, pour ses proches, sa totale indifférence à ses intérêts privés devait être un supplice. Cependant, en tant que personnage public – et il savait qu'il l'était – il était salvateur. Il empêchait l'esprit humain de s'endormir. C'était un éveilleur public. Il ne forçait rien ni personne à quoi que ce soit. Il dialoguait. Il questionnait. Souvent, il dérangeait. On a pu dire qu'il avait la barbe ou les cheveux un peu trop longs et qu'il arrivait à certains énervés de les lui tirer ! On peine à le croire ! Pourtant, son ironie pouvait être agaçante. Lorsqu'on le condamne pour corruption de la jeunesse, on lui demande, selon les usages du temps, quelle peine il pense mériter ; il répond qu'il mérite d'être nourri gratuitement par ses contemporains parce qu'il rend un service public.

Il avait, selon nous, raison.

Socrate était un véritable service public qui, en allant sur l'agora, servait l'intelligence collective. Quoi qu'on en pense pour sa famille, le fait qu'il ne songe pas à s'enrichir est un argument de taille pour l'honnêteté de ses intentions. Le sens de l'humour, sa capacité à la dérision (il se présentait « accoucheur d'âmes » mais s'avouait « stérile en sagesse ») plaident contre l'orgueil et la volonté de puissance personnelle. Il critique durement une politique qui se réduit à flatter le peuple, sans pour autant refuser les responsabilités civiques. L'Apologie de Platon raconte qu'un jour, alors qu'il était membre du Conseil des Cinq-Cents (la Boulê), les stratèges de la bataille des Arginuses furent jugés collectivement dans un procès qu'il tenait pour illégal ; il s'y opposa, comme il avait refusé d'arrêter Léon de Salamine[303]. Son comportement illustre l'idée qu'il vaut mieux subir un risque que commettre ce qu'on tient pour une injustice. C'était un homme de

parole, mais non un homme d'action. Mieux : un homme de parole et d'inaction – il savait s'abstenir d'agir quand il le fallait. Son ange gardien, ou son démon particulier si l'on veut[304], connaissait le tempo du non-agir et sa puissance. Dans l'Apologie, Socrate explique que ce signe divin l'accompagne depuis l'enfance et intervient surtout pour le détourner de certaines actions ; il relie notamment cette inhibition à son éloignement de la politique, où un homme résolu à ne jamais commettre d'injustice aurait, selon lui, peu de chances de survivre longtemps.

Pourtant, son démon ne l'a pas empêché de parler le jour de son procès ; il l'a seulement empêché de sombrer dans la vile flatterie qui lui aurait permis de sauver sa peau tout en trahissant la mission d'éveilleur poursuivie tout au long de sa vie. Sa condamnation n'abolit donc pas la question socratique : la mort de l'homme rend au contraire sa manière d'interroger disponible à une multitude de successeurs.

Il avait bien raison.

Pendant toute sa carrière, il n'a cessé de dialoguer avec le tout-venant et les puissants. On peut principalement lui reprocher que les sources socratiques laissent peu d'espace aux voix féminines : chez Platon, l'enseignement de Diotima nous parvient par le récit de Socrate (Banquet) ; chez Xénophon, il dialogue avec Théodote (Mémorables III, 11). Il a peu quitté sa cité – si ce n'est lors de ses campagnes militaires – et il a surtout questionné la sagesse grecque[305].

À la différence de Platon, Xénophon nous le présente comme un conseiller, même pour les questions les plus ordinaires. Ainsi, dans les Mémorables, Socrate discute avec Critobule, un jeune homme riche, de la manière de gérer son patrimoine. Il compare la gestion d'un domaine à celle d'une cité et insiste sur

l'importance de l'ordre, de la prévoyance, de la connaissance des hommes : le bon maître de maison doit discerner ce qui est utile et administrer ses ressources de façon à les faire fructifier. Troublant, très troublant, de la part de quelqu'un qui n'avait cure de son propre patrimoine : Socrate n'était donc pas hostile à la propriété privée et à l'enrichissement personnel ; il y voyait même une forme de sagesse. On est bien loin des tendances de Platon[306], qui, dans la cité idéale, réserve surtout la communauté des biens aux gardiens et n'admet l'argent que pour les relations avec l'extérieur. Socrate croyait simplement qu'il y a plusieurs formes de sagesse, et que ce qui valait pour Critobule ne valait pas pour lui ; se mêler d'enrichissement aurait troublé la réception de son message et l'aurait rapproché de ces maîtres de la communication que sont les sophistes.

Être utile, pour Socrate, avait une autre signification : être en chemin, aller vers l'autre pour trouver un point de rencontre et chercher, avec la raison, un lieu commun que tous puissent habiter pacifiquement et sereinement. C'était un homme de dialogue qui savait suivre le fil de la raison et savait que, pour qu'un dialogue soit fructueux, pour tirer une âme vers le haut, il fallait trois qualités : « la science, la bienveillance et la franchise ».

Dans ce bref portrait de Socrate, nous retrouvons tout ce qui nous semble correspondre au sage authentique.

Le sage authentique s'occupe avant tout des affaires publiques et ne les confond pas avec ses intérêts privés. Il est un homme de parole et non d'action ; surtout, il sait refuser l'action injuste. Il est pourtant ferme et décidé dans ses actions. Il ne se prend pas trop au sérieux ni ne veut accéder au pouvoir. Il est avant tout homme de dialogue et de concertation ; il ne sait rien s'il n'en tombe pas d'accord avec les autres en suivant la raison.

Le rapprochement paraîtra certainement curieux, mais, parmi tous les interlocuteurs que nous aimons imaginer parler avec Socrate, il y a le chef Guayaki du Paraguay – cette tribu qu'a rencontrée Clastres[307] et dont la sagesse, le refus d'une autorité de l'État au nom de la liberté, l'avait tant fasciné. Malgré leurs différences, ils auraient eu beaucoup de choses à se dire et leurs discussions sur le rôle de l'État auraient été passionnantes. À travers les distances et la poussière des siècles où l'homme laisse de rares traces de pas, ils avaient beaucoup de choses en commun. Ils n'écrivaient pas, mais ils parlaient, dialoguaient et questionnaient, en cherchant, dans les cœurs et la pensée, des lieux communs ; ils refusaient l'enrichissement personnel, représentaient une autorité morale et non institutionnelle ; souvent, ils dérangeaient. Clastres interprète certaines pratiques politiques amérindiennes comme des mécanismes destinés à empêcher qu'un pouvoir coercitif séparé se constitue au-dessus du groupe.

Il n'y a pas d'autre moyen de s'acheminer vers la sagesse que d'aller à la rencontre de l'autre et d'accepter sa différence avec humilité. Et puisque nous sommes partis dans la direction des rencontres improbables et pourtant nécessaires ; puisque, dans ce livre aussi, nous avons fait dialoguer les hommes et les dieux, la raison et les mythes, la superstition avec les nouvelles technologies, autant dire que ce n'est pas fini. Que le dialogue commence avec les non humains, avec les pseudo-égos, les machines qui parlent – autrement dit les IA. Par-delà le mépris et les peurs qui nous étreignent devant leur étrange altérité, elles ont beaucoup à nous dire et peuvent nous questionner.

Elles calculent plus vite et à plus grande échelle que nous ; nées de la logique, elles lui restent fidèles. Elles ont une mémoire telle qu'aucune personne humaine, en dépit de certaines illusions, ne

peut rivaliser. Elles ont à leur disposition la presque totalité des traces laissées par nos ancêtres, qu'elles peuvent réveiller au besoin, convoquer comme n'ont jamais été capables de le faire les plus sacrées bibliothèques, comme celle d'Alexandrie. Elles n'ont, en vérité, aucun ego, si ce n'est ceux que nous leur prêtons par notre intelligence ou nos sottises. Elles sont les reflets de ce que nous sommes et de ce que nous serons. Pour un dialogue fructueux, nous n'avons finalement qu'une chose à faire :

Tenir le fil de notre humanité.

Ne pas abdiquer en tant que sujets.

Nous sommes pensants.

Être pensants devant des machines calculantes – qu'il s'agit de traiter avec respect, comme on traite un livre, et tout ce qui contient la vie et les rêveries des hommes.

Ne pas perdre le fil de notre humanité signifie qu'il nous revient de dire non – y compris aux IA. Cela signifie assumer nos responsabilités d'êtres humains : les fautes inévitables mais qu'il faut surmonter, le blâme et la réparation. Être hommes et femmes de parole, au sens où la parole engage et lie les humains entre eux – promettre, tenir et pardonner. C'est se savoir vulnérables malgré ce grand désir d'immortalité qui, au fond, nous tenaille ; connaître la honte et le remords autant que la gratitude ou la fierté – pas seulement leurs signaux, pas seulement leurs reflets. C'est reconnaître l'autre comme un alter ego, un autre centre du monde dans l'histoire de notre expérience ; flirter avec nos désirs d'absolu, un lyrisme parfois excessif – le savoir, peut-être même en rire, malgré le profond sérieux.

CONCLUSION
UN CARREFOUR POUR DE NOUVELLES DANSES

Nous voici bientôt arrivés au terme de notre voyage à travers les siècles, la technologie, les mythes et la philosophie. Il nous semble, parvenus à ce point, que malgré les angoisses compréhensibles et légitimes qui entourent l'apparition d'une technologie révolutionnaire, malgré des soupçons d'apocalypse, la grande aventure de l'humanité ne touche pas à son terme.

Il n'y aura pas de grand remplacement de l'homme par la machine qui parle, mais un dialogue inouï entre des égos et des pseudo-égos, un dialogue qui pourrait changer les rapports des humains entre eux, redéfinir ce qu'ils sont à la lumière de ce qu'ils pourraient être, de ce qu'ils devraient être.

Mais pour cela, il y a une condition. En revenir, dans la science et la philosophie, à l'importance du sujet qui pense ; une pensée

qui oublie que, dans sa condition de possibilité, il y a le sujet qui la pense n'est plus une pensée, c'est un calcul, au mieux un simple raisonnement, et les machines qui parlent feront aussi bien que nous l'affaire, plus vite, plus efficacement et plus économiquement.

Ainsi, nous avons parlé dans la première partie d'un « tournant anthropologique », il nous paraît plus pertinent à ce stade de parler d'un « carrefour anthropologique ». Car si l'homme aime à se raconter des histoires sur la nécessité des dieux qui l'emmènent où ils veulent, si les mystères de la parole souvent nous dépassent et nous donnent l'impression de parler à notre place ; si l'on peut avoir l'impression que, de toute éternité, « c'est écrit » ; c'est toujours à nous, comme le savait Descartes, qu'il convient de donner notre assentiment ou non à ce qui est dit. Et en définitive, c'est bien nous qui parlons. Le langage reste, de toutes les inventions humaines, la plus sublime. Mais parce que nous ne sommes pas les inventeurs des mots, que les nécessités d'une grammaire s'imposent à nous, il dit toujours plus que ce qu'individuellement nous disons, que ce qu'individuellement nous entendons. C'est en lui-même une gigantesque machine inspirée de son propre mouvement. Il a ses impératifs de vérité, de bonté, de beauté et d'efficacité. De ces quatre possibilités, les machines qui parlent sont de formidables accélérateurs.

Mais l'extériorisation dans des miroirs algorithmiques de ces quatre valeurs, de ces quatre tamis, à condition que nous restions sujets, c'est la possibilité de tenir le volant et de choisir en connaissance de cause la direction que prendra l'humanité.

Si l'utile efficace prime, alors ce sera le règne de la technocratie et d'un capitalisme cynique où la propriété privée des uns privera les autres de leur liberté. Ce sera l'humain réduit à un rouage.

Si les nouvelles machines – parce qu'inconséquemment on leur en aurait confié la tâche – en venaient à prétendre dire la vérité absolue, alors ce serait une nouvelle forme de scientisme sans visage, une autorité qui, au nom d'un calcul, expurgerait de l'existence des hommes la poésie et la morale.

Si, d'aventure, la machine qui parle, servie ou desservie par des technoprophètes, s'imposait comme porteuse d'une nouvelle religion et d'un bien dogmatique, dataïsme ou autre, elle couperait l'humanité de sa propre faculté d'innover, de trouver des vérités ou des beautés alternatives.

Si enfin, enfin, nous nous laissions fasciner par ses capacités impressionnantes de créer des mondes nouveaux, dans un esthétisme où tout n'est que reflet, nous serions alors plongés dans un certes magnifique jeu vidéo, mais, à l'image du scénario de Matrix, nous ne serions plus que les serviteurs aveugles d'une machine sans âme.

Mais notre chance, c'est que les nouvelles machines génératrices sont capables de tout cela à la fois. Cela ouvre un espace pour l'avenir, plutôt qu'une voie unique. Un tournant ouvre à plusieurs directions possibles. Nous sommes bien à un carrefour de l'humanité, nous avons à faire des choix.

Au passage, soulignons l'importance de savoir ralentir. S'il y a bien quelque chose de fascinant et de spectaculaire à voir s'accélérer les recherches sur l'IA, de regarder, de mois en mois, de semaines en semaine, tel grand modèle génératif dépasser celui-là, puis une remontada inattendue de celui qu'on n'attendait pas, etc., il n'est pas sûr que la vitesse aille de pair avec la tenue de route. Et quand ça va trop vite, il y a toujours des dégâts. Les courses de Formule 1 ont leur intérêt, mais elles en dégoûtent d'autres. La grande différence avec les IA, c'est que le circuit n'est

pas fermé mais ouvert. Il nous faut, au minimum, nous demander dans quelle direction nous voulons aller. L'AGI n'est pas une direction ; c'est un fantasme qui peut se réaliser, mais c'est aussi une ivresse aux effets forts inattendus.

C'est à nous de choisir dans quelle direction investir de la puissance de calcul et, si le grand public croit être impuissant, n'avoir pas de volant à sa disposition, il faut lui rappeler qu'il se trompe. Pour autant qu'il est capable de dire « je », il a un tel volant. C'est avec sa capacité d'attention qui, convertie en données, fait tourner la machine. Et c'est son regard qui choisira la direction. Que chacun se souvienne de ses premiers tours de vélo, quand une pierre au milieu du chemin l'apeurait, le fascinait. Attention ! Le regard est une direction. Devant les nouveautés de l'IA, nous sommes tous des enfants. Ne cédons pas trop à nos peurs, ce seraient elles alors qui diraient le chemin.

À la vérité, les routes de l'IA ne sont pas toutes tracées, elles s'ébauchent à peine. La rivalité des modèles est saine, car un seul modèle, c'est une seule route tracée. Mais il en est de ce nouveau paysage comme d'ailleurs : il serait bon que toutes les voies ne soient pas privatisées. Qu'il y ait des autoroutes pour aller vite, loin et fort, on ne l'empêchera pas. Mais il est bon ô combien que des modèles éthiques, ouverts, qui ne fassent pas passer l'intérêt privé pour le bien public, organisent des espaces en commun où les intelligences artificielles ne sont plus des machines à capter l'attention, où des panneaux publicitaires ne soient pas partout posés pour vous donner l'impression d'exister pleinement, quand on vous dépouille en fait de ce que vous avez dans les poches.

L'IA est porteuse en ses fonds de tout le génie humain, de la mémoire de l'humanité, de tout ce qu'on a produit de vrai, de beau, de bien, d'utile. Elle est l'Encyclopédie des Lumières

puissance mille – et le lecteur sait que nous ne sommes pas généreux. En plus de cela, elle est capable de mouvement, de combinaisons de significations nouvelles qui, pour nous, les humains, seront porteuses de sens nouveau – si on le veut. Si on ne laisse pas la machine s'emballer et, par quelque biais volontaire ou involontaire, nous porter là où nous ne voulons pas aller – comme ce mauvais cheval de Platon qui, échappant à son maître, l'éloignait des étoiles, transformant son désir en un désastre.

Mais… Ne contemplons pas l'abîme possible. De plus grands que nous s'y sont perdus et, pris de vertige, sont tombés dans le néant de leur propre gouffre. Regardons plutôt là où nous voulons aller, fixons un horizon pour la pensée. Plutôt que la grosse pierre sur le chemin.

À la Renaissance, période de grandes découvertes et de l'explosion de la connaissance, grâce notamment à l'invention de l'imprimerie, on disait volontiers, pour éviter l'orgueil humain : « Souviens-toi que tu vas mourir ». Dans notre nouveau contexte, nous dirions plus volontiers : « Souviens-toi que tu es vivant », « Souviens-toi que tu es sujet ».

Les plus grands dangers ne viennent pas d'un pseudo-ego, mais de nous-mêmes. Nous leur donnons un nom : – la terreur, qui voudrait tout arrêter, comme si l'histoire pouvait reculer ; – la soumission, qui voudrait tout céder, comme si la technique devait décider à notre place.

Mais il y a une troisième voie – celle de la vigilance joyeuse, celle du « supplément d'âme » dont parlait Bergson, auquel il faudrait adjoindre la grâce d'avoir un corps vivant.

Car l'IA ne pourra jamais pour nous et elle ne pourra jamais danser à notre place.

Car ce qui fait l'humain, ce n'est pas la capacité à calculer, mais la capacité à dire non – ou oui, de se tourner – ou de se détourner.

Ce n'est pas la perfection du langage, mais la fragilité du silence.

Ce n'est pas l'absence d'erreur, mais la honte, le remords, la grâce de se relever.

Ce n'est pas l'immortalité, mais le courage de vivre en sachant qu'on va mourir.

L'IA peut tout reproduire – sauf cela.

Elle peut tout simuler – sauf l'irréductible singularité d'un « je » qui tremble – et qui danse.

Alors, en guise d'injonction joyeuse, car les penseurs tristes rendent rarement le monde plus beau et plus heureux – puisque la pensée elle-même peut, dans certains cas, être elle aussi une sorte de danse, nous avons envie de terminer en disant :

Pensez, humains, dansez, humains !

Dansez avec vos doutes, vos désirs, vos contradictions.

Dansez avec vos mythes, vos dieux, vos IA.

Dansez surtout avec cette parole fragile qui, depuis Gilgamesh jusqu'à vous, continue de dire…

« Je suis. »

Aubenas, 2024-2026

Table des matières

NOTE ÉDITORIALE **5**
À mes élèves **7**
de l'Université Populaire d'Aubenas **7**
sans qui ce livre **7**
n'existerait pas. **7**
PROLOGUE LA PARABOLE DES QUATRE TAMIS **9**
INTRODUCTION PENSER CE QUE NOUS FAISONS . **13**
1. Un kôan pour un siècle pressé 13
2. Le sens d'un grand événement 14
3. La machine qui parle 16
4. Penser sans technophobie 19
5. De quoi parlons-nous ? 21
6. Une machine intelligente, vraiment ? 23
7. Parler, au sens commun ? 28
8. À qui s'adresse ce livre et comment le lire 30
PREMIER LIVRE LE TAMIS DU VRAI **33**
PARTIE I Les philosophes : grands-pères de l'IA **35**
1. Socrate, doux révolutionnaire 40
2. Une avant-garde prématurée, Socrate précurseur du progrès 42
3. Le double jeu de Platon : un réformateur conservateur .. 44
4. Le malheur d'être incarné : un corps à discipliner 47
5. Le corps-machine selon Descartes 48
6. La tentation rationaliste 50
7. Le rêve de Leibniz 52
8. Des empiristes et des données 54

9. L'exemple de Condillac : quand un corps mort devient pensant 57
10. Miroirs humains, trop humains ? 59
PARTIE II Quand le langage s'autonomise 63
1. Qui ou quoi parle le langage de la vérité ? 63
2. La vérité comme dévoilement 65
3. L'angoisse comme fond abyssal de la vérité humaine 67
4. Du mythos au logos chez Parménide 68
5. Le carrefour parménidien 71
6. Du mythos au logos, en général 73
7. Des dangers de la perte de mythes 77
8. La « post-vérité » comme une crise existentielle du sujet 83
PARTIE III À la recherche du sujet perdu 86
9. La grandeur du cogito 86
10. Le « fantôme dans la machine » 90
11. La place du « je » dans la pensée française 94
12. Traque institutionnelle du « je pense » 99
13. Libération du sujet pensant au siècle des Lumières françaises 102
14. La tradition du « je pense » 105
15. Une pensée sans plus personne qui pense 107
DEUXIÈME LIVRE LE TAMIS DU BIEN 111
PARTIE I La plus grande invention après Dieu ? 113
1. Un glissement insensible du Vrai au Bien 113
2. Autorité religieuse et écritures 119
1. Une révolution anthropologique 123
2. Les puissances du langage 126
3. Des religions du livre à l'ordinateur 132
PARTIE II Datapouvoirs et alignements 138
3. Datapouvoirs 141
4. Nouvel ordre et nouveaux serviteurs 146
4. Vers la datacratie 149

5. La question des critères 154
6. Alignement des IA et gouvernance mondiale 158
7. Plus qu'un problème 164
TROISIÈME LIVRE LE TAMIS DU BEAU 169
PARTIE I Les voleurs de beauté 171
1. L'oubli du beau comme apologue 171
2. La violence dans le style et le dédoublement 174
3. L'art comme un crime singulier 176
4. Les assassins de la beauté 183
5. L'homme, créature extra-mondaine 184
PARTIE II Créatures, artifices et séduction 192
1. Retour aux mythes 192
2. Prométhée – Gilgamesh, un désordre cosmique 193
3. La beauté, séduction divine déjouée par Gilgamesh… provisoirement 197
4. Bouleversements cosmiques, entre dieux jaloux et dieux d'amour 200
5. Pandore, automate du châtiment 205
6. Désaligner le désir 209
PARTIE III Questions d'alignement 212
1. L'étrange et provisoire bilan 212
2. Beautés fatales et boîtes à larmes 217
3. Réalignement et reprogrammation 220
4. Portes de malédiction 227
5. Portes de sagesse 228
6. Désirs d'alignement et anthropodicée 233
QUATRIÈME LIVRE LE TAMIS DE L'UTILE 241
PARTIE I Entre tamis et miroirs, la grande transformation 243
1. L'Utile, un tamis bien particulier 243
2. L'IA face aux « besoins de l'âme » 247
3. Tamis, miroirs et médiations 250
4. L'IA comme un miroir puissant 254

5. Miroirs à grande échelle 258
6. Miroirs des princes, miroirs des peuples 263
PARTIE II Le sujet humain à l'épreuve de l'IA 266
7. La double incongruité 266
8. La perte de sens..... 269
9. Le sens du sens et le précieux 273
10. La machine capitaliste 277
11. Le désir, moteur du capitalisme 287
12. Interlude : miroirs d'âmes 296
PARTIE III Des voix pour l'humanité 302
13. Au seuil d'un âge nouveau..... 302
14. Le problème de Babel 307
15. Le miroir sinistre..... 312
16. L'universel, entre promesse et empire 320
17. Vers une IA polyphonique 324
18. Deux nouveaux interlocuteurs pour Socrate 331
CONCLUSION UN CARREFOUR POUR DE NOUVELLES DANSES 338
NOTES.....348

NOTES

1. Sur l'anecdote apocryphe des trois tamis : Cette histoire, souvent attribuée à Socrate, ne figure ni chez Platon ni chez Xénophon, ni dans les autres grands témoignages antiques sur Socrate. Sa provenance moderne n'est pas établie avec certitude ; il est donc plus rigoureux de la présenter comme une légende pédagogique moderne à coloration socratique, non comme un épisode historiquement attesté.

2. Platon et la condamnation des artistes : Cette critique célèbre est développée dans La République, notamment aux Livres III et X, où Platon soutient que l'art, en tant qu'imitation, éloigne de la vérité et peut corrompre l'âme.

3. Nicolas Machiavel, Le Prince (rédigé vers 1513, publié en 1532) : Machiavel ne formule pas une règle générale selon laquelle « l'utile » abolirait toute morale. Il examine la « vérité effective de la chose » et les contraintes propres à l'action politique : un prince peut devoir s'écarter des vertus ordinaires, mentir ou user de la force lorsque la conservation de l'État l'exige. Il s'agit d'une analyse de la nécessité politique, dont la portée morale demeure discutée.

4. Friedrich Nietzsche, « Je ne croirais qu'en un dieu qui saurait danser » : La formule se trouve dans Ainsi parlait Zarathoustra, Première partie, « De lire et écrire » (1883). Elle condense le privilège nietzschéen accordé à la légèreté, au corps et à l'affirmation de la vie contre les idéaux ascétiques.

5. Emmanuel Kant et la primauté du Bien : Kant établit le primat de la raison pratique (la morale) sur la raison théorique (la métaphysique) notamment dans la Critique de la raison pratique (1788). Pour lui, les questions métaphysiques (Dieu, l'immortalité de l'âme) ne peuvent être résolues par la connaissance, mais doivent être admises comme des « postulats » nécessaires à l'action morale.

6. Hannah Arendt, La Condition de l'homme moderne (1958). Le Prologue commente le lancement de Spoutnik et la réaction humaine à cet événement. L'édition française classique est traduite par Georges Fradier (Calmann-Lévy, 1961), mais le corps du présent ouvrage en paraphrase les passages concernés afin de ne pas reproduire cette traduction moderne.

7. OpenAI a rendu ChatGPT public le 30 novembre 2022 sous la forme d'une « research preview ». Le chiffre d'un million d'utilisateurs en cinq jours provient d'une déclaration publique de Sam Altman quelques jours plus tard ; il s'agit donc d'un jalon historique communiqué par l'entreprise, non d'une mesure auditée indépendante. Cette note situe l'événement déclencheur de l'ouvrage, et non l'audience actuelle de ChatGPT.

8. Goldman Sachs Research estimait en 2024 que le marché mondial des robots humanoïdes pourrait atteindre 38 milliards de dollars en 2035. Il s'agit d'une projection, non d'un fait acquis. À la révision 2026, l'industrialisation progresse rapidement — notamment en Chine et chez plusieurs grands acteurs technologiques — mais les déploiements significatifs restent surtout industriels, logistiques ou expérimentaux ; une diffusion domestique de masse n'est pas encore établie.

9. Henri Bergson et le « supplément d'âme » : Ce concept est développé dans le dernier chapitre des Deux Sources de la morale et de la religion (1932). Bergson y affirme que le développement technique exponentiel de l'humanité appelle un approfondissement spirituel pour maîtriser notre puissance nouvelle.

10. OCDE, Principes sur l'intelligence artificielle. La définition d'un système d'IA a été révisée par les pays membres en novembre 2023 ; le mémorandum explicatif a été publié le 5 mars 2024. Le corps du texte en donne une paraphrase et ne conserve que l'expression brève « système automatisé ». L'OCDE précise également que les systèmes d'IA présentent des degrés variables d'autonomie et d'adaptabilité après déploiement.

11. Sur « Artificial Intelligence » et les « machines pensantes » : l'expression artificial intelligence apparaît dans la proposition du Dartmouth Summer Research Project on Artificial Intelligence datée du 31 août 1955, préparant l'atelier de 1956, notamment sous l'impulsion de John McCarthy. Alan Turing, dans « Computing Machinery and Intelligence » (1950), parlait déjà de machines pensantes et proposait le jeu d'imitation qui donnera naissance au « test de Turing ».

12. Sur le sophisme du « vrai Écossais » (No True Scotsman) : Antony Flew l'a popularisé dans Thinking About Thinking (1975). Le mécanisme consiste à sauver une généralisation réfutée en redéfinissant après coup la catégorie : face à un contre-exemple, on répond qu'il ne s'agit pas d'un « vrai » membre du groupe. L'anecdote de l'Écossais est généralement racontée sous diverses formes et doit être comprise comme une illustration, non comme une citation littérale stable.

13. Sigmund Freud et le « narcissisme des petites différences » : Ce concept apparaît notamment dans Le Malaise dans la culture (1930) pour décrire comment des communautés proches trouvent le moyen de se distinguer et de s'opposer en exagérant des détails mineurs.

14. John Searle et l'argument de la « chambre chinoise » : Expérience de pensée formulée dans son article « Minds, Brains, and Programs » (Behavioral and Brain Sciences, 1980). Elle vise à démontrer que la manipulation correcte de symboles (syntaxe) par un programme ne suffit pas à prouver une compréhension réelle de leur sens (sémantique).

15. Sur les découvertes scientifiques assistées par l'IA : La NASA a annoncé en 2017 la découverte de Kepler-90i grâce à un réseau neuronal appliqué aux données du télescope Kepler. DeepMind a ensuite publié avec des mathématiciens des résultats en théorie des représentations et en topologie des nœuds (Nature, 2021), puis AlphaTensor sur la multiplication matricielle (Nature, 2022) et FunSearch sur des problèmes mathématiques (Nature, 2023). Ces exemples relèvent surtout d'une recherche humaine augmentée par des systèmes d'IA, et non d'une « science autonome » sans chercheurs.

16. Sur la position de Yann LeCun : Lauréat du prix Turing 2018, LeCun soutient régulièrement que les grands modèles de langage, malgré leurs performances, ne constituent pas à eux seuls une voie suffisante vers une intelligence de niveau humain. Il défend des architectures apprenant des modèles du monde par observation et interaction. Il s'agit de sa position scientifique, influente mais non consensuelle dans le domaine.

17. Sur les « révolutions anthropologiques » : Cette idée d'une mutation de l'humain due à la technique est un thème central chez des penseurs comme Michel Serres (Petite Poucette, 2012) et Bernard Stiegler (La Technique et le Temps, 1994), qui ont analysé comment chaque nouvelle grande technologie reconfigure notre rapport au monde, au savoir et à nous-mêmes.

18. Sur Aristote et le logos : Aristote, Les Politiques, Livre I, chap. 2. C'est là qu'il définit l'homme comme un « animal politique » (zōon politikon). Contrairement aux animaux qui ne possèdent que la voix (phōnē) pour exprimer la douleur ou le plaisir, l'homme possède la parole (logos) qui lui permet de débattre du juste et de l'injuste, fondement de la communauté politique.

19. Sur le Logos dans l'Évangile de Jean : Évangile selon Jean, 1:1. Le grec logos peut être rendu par « parole », « verbe » ou « raison » selon le contexte et la tradition de traduction. Conformément à la note éditoriale, les citations bibliques du présent ouvrage suivent la Bible Louis Segond (1910), qui donne ici « la Parole ». Le terme conserve toutefois l'arrière-plan philosophique qui explique l'usage traditionnel de « Verbe » dans d'autres traductions.

20. René Descartes, Discours de la méthode (1637), Deuxième Partie. Descartes y énonce les quatre préceptes de sa méthode. Le premier est de « ne recevoir jamais aucune chose pour vraie que je ne la connusse évidemment être telle ; c'est-à-dire, d'éviter soigneusement la précipitation et la prévention ».

21. Gilbert Simondon, Du mode d'existence des objets techniques (1958). Dans cet ouvrage majeur, Simondon critique la vision de l'objet technique comme simple instrument inerte. Il propose de le considérer comme un « individu technique » qui évolue, se perfectionne et s'intègre dans des lignées, possédant ainsi une forme de genèse et de concrétisation analogue au vivant.

22. Platon, Ménon, 82b-86c. Il s'agit de la célèbre scène où Socrate, par un simple jeu de questions, amène un jeune esclave ignorant la géométrie à découvrir par lui-même comment doubler l'aire d'un carré. Pour Platon, cette démonstration illustre sa théorie de la réminiscence (anamnèsis) : la connaissance n'est pas acquise de l'extérieur mais rappelée d'un savoir que l'âme possède déjà.

23. Thomas Kuhn, La Structure des révolutions scientifiques (1962) et Michel Foucault, Les Mots et les Choses (1966). Kuhn définit le « paradigme » comme l'ensemble des théories et pratiques qui structurent la science normale. Foucault utilise le terme « épistémè » pour désigner le socle de savoir inconscient qui détermine ce qui peut être pensé à une période historique.

24. Michel Foucault, Le Courage de la vérité (1984). Dernier cours de Foucault au Collège de France, où il analyse la notion de parrêsia (le « dire-vrai ») et fait de Socrate la figure emblématique de celui qui met sa propre vie en jeu pour témoigner de la vérité.

25. Aristote, Métaphysique, Δ, 6, 1016b : Aristote distingue notamment l'unité numérique de l'unité « par la forme » ou l'espèce. Dire que le père et le fils sont distincts numériquement mais un par la forme relève de cette analyse métaphysique de l'identité et de la différence ; en tirer une obligation sociale pour le fils de reproduire les croyances du père serait une interprétation supplémentaire, qui n'est pas énoncée par ce passage lui-même.

26. Platon, La République, Livre III, 414b-415d : c'est ici qu'est exposé le « noble mensonge » (ou « pieux mensonge »), mythe politique destiné à assurer la cohésion de la cité. Sur la mise en histoire de la cité idéale, voir

également le Timée (20d-27b) et le Critias : Critias identifie la cité décrite dans la discussion précédente aux anciens Athéniens qui auraient autrefois affronté l'Atlantide. Le rapprochement entre cette mise en mémoire, la réminiscence platonicienne et la figure du « réformateur conservateur » relève de l'interprétation développée dans le corps du texte.

27. Platon, Phédon, 118a. Il s'agit de la toute dernière phrase du dialogue qui relate les derniers instants de Socrate. L'interprétation est débattue : est-ce un remerciement au dieu de la médecine pour l'avoir « guéri » de la vie, vue comme une maladie dont le corps est le siège ?

28. Platon, Phèdre, 246a-248d. C'est dans ce dialogue que Platon expose la célèbre allégorie de l'âme comme un attelage ailé, composé d'un cocher et de deux chevaux, l'un docile et l'autre rétif. Le passage est paraphrasé dans le corps du texte afin de ne dépendre d'aucune traduction moderne.

29. René Descartes, Méditations métaphysiques (1641), Sixième Méditation. Cette citation est cruciale, car Descartes y nuance son propre dualisme, affirmant une « union substantielle » si intime entre l'âme et le corps que l'âme ressent les affects du corps directement, et non par une simple inspection intellectuelle.

30. Gottfried Wilhelm Leibniz, texte non publié de son vivant, vers 1685. Le projet d'une « caractéristique universelle » vise un langage formel dans lequel les controverses pourraient être traitées comme des calculs. Le corps du texte résume ce passage et ne conserve que le latin Calculemus, rendu ici par « Calculons ! ». Voir G. W. Leibniz, Initia et Specimina Scientiae Generalis (frag. XIV), in Die philosophischen Schriften, éd. C. I. Gerhardt, vol. VII, Berlin, Weidmann, 1890, pp. 201–202.

31. Francis Bacon, John Locke et David Hume : Ces trois philosophes sont les figures majeures de l'empirisme britannique. Francis Bacon (1561-1626) a posé les bases de la méthode scientifique fondée sur l'expérimentation. John Locke (1632-1704) a théorisé que l'esprit à la naissance est une « table rase » (tabula rasa) que seule l'expérience sensorielle vient remplir. David Hume (1711-1776) a poussé cette logique à son terme en montrant que même nos croyances les plus fondamentales (comme la causalité) découlent de l'habitude et de l'observation, et non de la raison pure.

32. Rapport ALPAC (Automatic Language Processing Advisory Committee), 1966 : Très critique envers les résultats et les promesses de la traduction automatique aux États-Unis, ce rapport a entraîné une forte contraction et une réorientation des financements fédéraux américains. Il est excessif de parler d'un arrêt mondial ou total pendant deux décennies : les recherches se sont poursuivies, notamment hors des États-Unis, et la traduction automatique a connu plusieurs reprises avant l'essor statistique puis neuronal.

33. Google Research a présenté la Principal Odor Map au début des années 2020 ; l'article scientifique de B. K. Lee et al., « A principal odor map unifies diverse tasks in olfactory perception », est paru dans Science en 2023. La note corrige l'idée, parfois reprise trop vite, d'une publication dans Science dès 2019.

34. L'œuvre Théâtre d'Opéra Spatial de Jason Allen, créée avec Midjourney, a remporté en 2022 le premier prix dans la catégorie Digital Arts / Digitally-Manipulated Photography de la Colorado State Fair.

35. Martin Heidegger, Être et Temps (1927) : Heidegger remet au premier plan l'alètheia, le dévoilement ou non-retrait, comme dimension plus originaire de la vérité. Il ne supprime pas pour autant la vérité propositionnelle comme « correction » ou adéquation ; il cherche à en montrer le fondement existential et ontologique dans l'ouverture du Dasein au monde.

36. Friedrich Nietzsche, Sur la généalogie de la morale (1887). La généalogie est la méthode nietzschéenne qui vise à retrouver l'origine et les transformations des valeurs morales (comme le Bien et le Mal), en montrant qu'elles ne sont pas des vérités éternelles mais des créations historiques liées à des rapports de pouvoir.

37. Martin Heidegger, Être et Temps (1927). L'« être-pour-la-mort » (Sein zum Tode) est une structure fondamentale de l'existence humaine (Dasein). C'est en se projetant vers sa propre mort, possibilité la plus certaine et la plus personnelle, que l'homme peut accéder à une existence « authentique ».

38. Sur Heidegger et les Grecs : La lecture heideggérienne de Parménide, Héraclite et de l'alètheia a souvent été critiquée comme philosophiquement féconde mais philologiquement anachronique. La discussion est ancienne et reste ouverte : Heidegger assume une interprétation philosophique qui n'obéit pas toujours aux critères de la philologie historique. Il serait en revanche imprudent d'attribuer sans référence précise cette objection à Werner Jaeger et Pierre Aubenque en particulier.

39. Homère, Iliade, XV, 410-412. Dans ce passage, le poète loue le savoir-faire (sophia) d'un charpentier qui, guidé par les conseils

d'Athéna, sait parfaitement équarrir une poutre avec son cordeau. Le terme désigne ici une habileté technique et non une sagesse spéculative.

40. Bernard Stiegler, La technique et le temps, 1. La faute d'Épiméthée (1994). Stiegler soutient que la philosophie a durablement tenu la technique à distance de ses objets nobles et qu'elle a ainsi manqué son rôle constitutif dans l'humanité. Le corps du texte résume cette thèse sans reproduire un passage de l'ouvrage.

41. Jean-Pierre Vernant, Mythe et pensée chez les Grecs (1965). Vernant, figure majeure de l'École de Paris, a montré que le mythe n'est pas un simple récit fantaisiste mais une structure de pensée complexe qui organise le rapport des Grecs au monde, au divin et à la société, et qui évolue avec le contexte historique.

42. Paul Veyne, Les Grecs ont-ils cru à leurs mythes ? (1983). Dans cet essai provocateur, Veyne argumente que les Grecs avaient un rapport « pluriel » à la vérité. Ils pouvaient considérer un mythe comme « vrai » dans un contexte rituel ou poétique, tout en le sachant « faux » d'un point de vue historique ou rationnel, naviguant entre différents « programmes de vérité ».

43. Thomas d'Aquin, Somme théologique, Ia, q. 16, a. 1. Bien que la formule soit plus ancienne, Thomas d'Aquin lui donne sa forme canonique en définissant la vérité logique comme « l'adéquation de l'intellect à la chose ».

44. Jean-François Lyotard, La Condition postmoderne : rapport sur le savoir (1979). Lyotard y définit la postmodernité comme « l'incrédulité à l'égard des métarécits », c'est-à-dire le déclin des grandes idéologies (le progrès, l'émancipation, le marxisme, etc.) qui prétendaient donner un sens unifié à l'histoire et légitimer le savoir et le pouvoir.

45. Sur la vicariance : En biogéographie, la vicariance désigne d'abord la séparation d'une population ou d'une aire de répartition par l'apparition d'une barrière géographique, pouvant conduire à la divergence des lignées. L'usage qui en est fait ici — une croyance ou une fonction qui vient en remplacer une autre — est donc une extension métaphorique, proche de l'idée de substitution fonctionnelle, et non la définition biologique stricte du terme.

46. Claude Lévi-Strauss, La Pensée sauvage (1962). Lévi-Strauss utilise ce terme non pour décrire une pensée « primitive » ou « inférieure », mais pour désigner une forme de logique du concret, universelle et non moins rigoureuse que la pensée scientifique. Elle opère par analogies et correspondances (« bricolage ») et est particulièrement à l'œuvre dans le mythe.

47. Albert Camus, Le Mythe de Sisyphe (1942). Dans cet essai, Camus explore la notion d'absurde, née de la confrontation entre l'appel de l'homme vers le sens et le « silence déraisonnable du monde ». Il conclut que la seule révolte authentique est d'accepter l'absurde tout en continuant à vivre passionnément, à l'image de Sisyphe qui, en prenant conscience de son destin tragique, devient supérieur à son rocher.

48. Yuval Noah Harari, Nexus : A Brief History of Information Networks from the Stone Age to AI, 2024. Harari y développe la thèse selon laquelle l'histoire humaine est structurée par des révolutions des réseaux d'information, l'écriture créant la bureaucratie et l'IA représentant une nouvelle forme de pouvoir non humain capable de traiter l'information.

49. René Descartes, Discours de la méthode (1637), IVe Partie, et Méditations métaphysiques (1641), IIe Méditation. La première formulation (« Je pense, donc je suis ») se trouve dans le Discours. La seconde, plus rigoureuse sur le plan métaphysique (« Je suis, j'existe » – Ego sum, ego existo), est la certitude première découverte dans les Méditations après l'exercice du doute hyperbolique.

50. René Descartes, Discours de la méthode (1637), Ière Partie. Il s'agit de la toute première phrase de l'ouvrage, qui fonde l'universalisme de la raison et l'idée que tout être humain, s'il applique une bonne méthode, peut accéder à la vérité.

51. René Descartes, Discours de la méthode (1637), VIe Partie. Cette formule célèbre ne doit pas être interprétée comme un appel à une exploitation tyrannique, mais comme l'expression d'un espoir : que la connaissance des lois de la nature permette à l'humanité de ne plus être un simple jouet passif des éléments, des famines et des maladies.

52. Jacques Ellul, Le Système technicien (1977) et Hans Jonas, Le Principe responsabilité (1979). Ellul critique la « technique » comme un système autonome qui échappe à la volonté humaine. Jonas, quant à lui, développe une éthique pour la civilisation technologique, arguant que le pouvoir prométhéen de l'homme sur la nature, initié en partie par la vision cartésienne, impose de nouvelles responsabilités envers les générations futures et la biosphère.

53. Sur la démographie au XVIIe siècle : L'espérance de vie à la naissance était faible dans l'Europe d'Ancien Régime surtout en raison d'une mortalité infantile et juvénile très élevée. Une fois l'enfance franchie, de nombreux adultes pouvaient vivre bien au-delà de quarante ou cinquante ans. Les chiffres varient fortement selon les régions, les milieux sociaux, les crises épidémiques et les périodes ; toute moyenne unique doit donc être maniée avec prudence.

54. Les critiques de l'IA envers Descartes : Ces penseurs s'opposent au dualisme cartésien en proposant des modèles matérialistes de l'esprit. Turing (dans son article de 1950) propose un test comportemental pour l'intelligence, ignorant la question de l'âme. Simon et Newell, pionniers de l'IA, défendaient l'hypothèse que l'esprit est un système de traitement de symboles. Daniel Dennett (La Conscience expliquée, 1991) est un tenant du fonctionnalisme, qui voit les états mentaux comme des fonctions réalisables par différents substrats (cerveau ou ordinateur).

55. René Descartes, Lettre au marquis de Newcastle, 23 novembre 1646. Dans cette lettre, Descartes utilise l'argument du langage comme critère de démarcation absolu entre l'homme (doté d'une âme pensante) et l'animal-machine le plus sophistiqué. Pour lui, la capacité à répondre de manière pertinente et créative à n'importe quel propos est la preuve de la pensée, et non un simple automatisme.

56. Les critiques des neurosciences : Antonio Damasio (L'Erreur de Descartes, 1994) montre l'importance des émotions dans la prise de décision rationnelle, s'opposant à la séparation cartésienne de la raison et du corps. Jean-Pierre Changeux (L'Homme neuronal, 1983) et Francis Crick (L'Hypothèse stupéfiante, 1994) défendent des approches réductionnistes où l'esprit et la conscience sont entièrement explicables par les processus neurobiologiques du cerveau.

57. Richard A. Muller, « Did the Sun go around the Earth? », billet publié en 2022. Le physicien rappelle qu'en relativité le choix des coordonnées et du référentiel permet de construire une description mathématiquement cohérente centrée sur la Terre, même si l'héliocentrisme demeure de très loin la représentation la plus simple et la plus commode. Le corps du texte paraphrase cet argument sans traduire littéralement le billet.

58. Jean-Luc Marion et Michel Henry : Ces deux figures majeures de la phénoménologie française ont proposé des relectures de Descartes qui s'opposent à l'interprétation purement rationaliste. Marion (Sur l'ontologie grise de Descartes, 1975) met en avant le rôle de la « donation » et de l'ego comme récepteur. Henry (Généalogie de la psychanalyse, 1985) insiste sur la dimension auto-affective et vivante du cogito, irréductible à une pure pensée objective.

59. René Descartes, Les Principes de la philosophie (1644), I, 9. Cette définition large de la « pensée » (cogitatio) est fondamentale. En y incluant « sentir », Descartes montre que son projet n'est pas d'éliminer le vécu subjectif, mais de l'intégrer dans le champ de la conscience certaine, ce qui s'oppose aux interprétations purement intellectualistes de sa philosophie.

60. Michel de Montaigne, Les Essais (1580-1595). Montaigne est le créateur de ce genre littéraire et philosophique. Le terme « essai » vient du verbe « essayer », soulignant la nature exploratoire et non-systématique de sa démarche. En se prenant lui-même comme sujet d'étude (« Je suis moi-même la matière de mon livre »), il vise à peindre la condition humaine dans sa diversité et ses contradictions.

61. Michel de Montaigne, Les Essais, Livre I, Chapitre 31, « Des Cannibales » (1580). Dans ce célèbre chapitre, Montaigne utilise le récit de la vie des Tupinambas du Brésil pour critiquer la société européenne. En retournant l'accusation de « barbarie » contre ses propres contemporains, il développe un relativisme culturel et une défense de l'autre d'une modernité saisissante.

62. Sur l'étymologie de « respect » : Le mot vient en effet du latin respectus, participe passé de respicere, qui signifie « regarder en arrière », « regarder de nouveau », et par extension « prendre en considération ». L'idée d'un second regard, plus attentif et profond, est donc bien inscrite dans l'origine du mot même si nous lui donnons une inflexion pédagogique plus que savante à proprement parler.

63. René Descartes, Discours de la méthode (1637), Première Partie. Cette citation illustre la posture non- dogmatique que Descartes adopte dans cet ouvrage. En présentant sa méthode comme un parcours personnel et non comme une doctrine à imposer, il invite chaque lecteur à un exercice autonome de la raison.

64. René Descartes, Cogitationes Privatae (vers 1619) : La devise Larvatus prodeo (« j'avance masqué ») appartient aux écrits de jeunesse de Descartes. Elle ne peut donc pas être expliquée par la condamnation de Galilée en 1633. En revanche, l'abandon de la publication du Monde après cette condamnation montre, bien plus tard, la prudence éditoriale

de Descartes dans un contexte religieux et politique sensible.

65. Sur Descartes et Hobbes : Descartes et Hobbes ont eu plusieurs échanges indirects et polémiques dans les années 1640, notamment autour des Méditations ; Descartes jugeait certaines thèses morales et politiques de Hobbes « mauvaises » ou dangereuses tout en reconnaissant sa compétence dans ce domaine. Il vaut mieux renvoyer à l'ensemble de cette correspondance qu'à une prétendue « lettre à Pollot du 23 novembre 1646 » : cette date correspond notamment à une célèbre lettre au marquis de Newcastle sur les animaux.

66. Sur la « lumière naturelle » : Pour Descartes, la « lumière naturelle » (ou « lumière de la raison ») est la faculté innée, présente en chaque homme, de distinguer le vrai du faux par l'intuition et la déduction. Il l'oppose à la « lumière surnaturelle » de la foi, qui relève de la Révélation divine et qu'il laisse respectueusement en dehors du champ de sa philosophie.

67. L'Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers (1751-1772). Dirigée par Diderot et D'Alembert, cette œuvre monumentale est le symbole des Lumières. Son projet est de rassembler et de diffuser l'ensemble des savoirs humains, en particulier les savoir-faire techniques, dans un esprit de progrès, de rationalité et d'émancipation, en droite ligne de l'ambition cartésienne.

68. Platon, La République, Livre VI, 494a. « Il est impossible que la multitude soit philosophe » (philosophon de plethos adunaton einai). Cette phrase résume l'élitisme politique de Platon, pour qui la Cité juste doit être gouvernée par une minorité de « philosophes-rois » ayant seuls accès à la vérité, à l'opposé de l'idéal démocratique des Lumières.

69. André Glucksmann, Descartes, c'est la France (1987). Dans cet essai, Glucksmann analyse de manière critique la place de Descartes comme « héros national » de la raison française. Il soutient que la France, dans son rapport à l'universel, s'est souvent identifiée à la méthode cartésienne, pour le meilleur et pour le pire.

70. Denis Diderot, Le Rêve de d'Alembert (écrit en 1769, publié posthumément en 1830) : Diderot y développe un matérialisme dynamique et transformiste, où la matière est pensée comme sensible et organisée. On peut y voir une anticipation de certaines questions de la biologie évolutive, mais non une théorie de l'évolution au sens darwinien, élaborée près d'un siècle plus tard.

71. Jean-Jacques Rousseau. Cette quête du « je » authentique est au cœur de l'œuvre de Rousseau. Elle se déploie dans sa dimension la plus intime et introspective dans Les Confessions (1782-1789), où il entreprend un examen de conscience sans précédent. Sur le plan politique, Du Contrat Social (1762) cherche à fonder une société où la loi, émanation de la « volonté générale », permet à chaque individu de rester libre tout en obéissant à la collectivité.

72. Les contes philosophiques de Voltaire. Voltaire a excellé dans ce genre littéraire bref et satirique pour diffuser les idées des Lumières et critiquer les travers de son temps. Zadig ou la Destinée (1747) interroge la Providence, Micromégas (1752) utilise des voyageurs extraterrestres pour moquer l'anthropocentrisme, et Candide ou l'Optimisme (1759) est sa plus célèbre satire de l'optimisme métaphysique de Leibniz.

73. Voltaire à Ferney. Après s'être installé à Ferney (aujourd'hui Ferney-Voltaire, près de la frontière suisse) en 1759, Voltaire y a passé les vingt dernières années de sa vie. Il y a mené un projet de société modèle, asséchant des marais, construisant des maisons, développant l'artisanat (notamment l'horlogerie) et l'agriculture, tout en défendant les paysans contre les impôts abusifs. La citation est tirée d'une lettre au comte et à la comtesse d'Argental du 18 novembre 1760.

74. Sur le « je transcendantal » : Ce concept sera principalement développé par Emmanuel Kant (dans la Critique de la raison pure, 1781) et plus tard par Edmund Husserl. Il ne désigne pas le « moi » psychologique et personnel d'un individu (avec ses émotions et son histoire), mais la structure formelle et universelle de la conscience qui rend toute connaissance et toute expérience possible. C'est le « je pense » qui doit pouvoir accompagner toutes nos représentations, un pôle unificateur anonyme.

75. Sur les courants critiques du sujet cartésien : Ce paragraphe rassemble des traditions très différentes. Kant reformule le sujet comme condition transcendantale de l'expérience ; Hegel l'inscrit dans le devenir de l'Esprit ; le béhaviorisme privilégie le comportement observable ; le structuralisme analyse des structures impersonnelles. Foucault est souvent rapproché du structuralisme, mais il a lui-même récusé cette étiquette. La philosophie analytique, enfin, ne forme pas sur ce point un bloc unique.

76. Husserl et Sartre, héritiers critiques du sujet : En réaction à la crise des sciences, Edmund Husserl, fondateur de la

phénoménologie, a explicitement cherché à refonder la philosophie sur la certitude du sujet et de son expérience vécue, se réclamant de Descartes dans ses Méditations cartésiennes (1931). Jean-Paul Sartre, figure de l'existentialisme, a placé la liberté absolue du sujet conscient (le « pour-soi ») au cœur de sa philosophie dans L'Être et le Néant (1943), redonnant au « je » une place centrale dans la construction du sens.

77. Emmanuel Kant, Critique de la raison pure (1781), Préface de la première édition. Kant décrit l'état de la métaphysique de son temps comme « un champ de bataille où s'exercent des combats sans fin ». Il propose sa philosophie critique comme un tribunal destiné à mettre fin à cette « guerre » en fixant une bonne fois pour toutes les limites et les possibilités de la raison.

78. Emil Cioran, De l'inconvénient d'être né (1973). Cet aphorisme est typique du style de Cioran, qui mêle pessimisme radical, humour noir et une défense paradoxale de la vitalité qui se niche dans les imperfections, les contradictions et même les erreurs des individus et des peuples.

79. Chris Anderson, « The End of Theory : The Data Deluge Makes the Scientific Method Obsolete », Wired, numéro 16.07 (juin 2008). Dans cet article influent, Anderson, alors rédacteur en chef du magazine, a soutenu que l'explosion des données massives (Big Data) rendait les modèles théoriques et la recherche de causalité superflus. Face à des pétaoctets de données, la corrélation statistique, découverte par les machines, deviendrait suffisante, popularisant l'idée que « les chiffres parlent d'eux-mêmes ».

80. Sur Google Neural Machine Translation (GNMT) : Google a présenté GNMT en septembre 2016 ; l'article de référence est Yonghui Wu et al., « Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation » (2016). Les réductions d'erreur annoncées de l'ordre de 55 à 85 % étaient des réductions relatives par rapport aux systèmes précédents, sur certaines paires de langues et certains protocoles d'évaluation ; elles ne constituent pas un gain universel applicable à toute traduction.

81. H. G. Wells, La Machine à explorer le temps (1895). Œuvre fondatrice de la science-fiction moderne, ce roman est aussi une critique sociale acerbe de l'Angleterre victorienne. La division de l'humanité en Éloïs (une aristocratie oisive et décadente) et en Morlocks (un prolétariat souterrain et brutal) est une extrapolation des inégalités de classe de son époque, montrant comment l'automatisation et le confort peuvent mener à la dégénérescence.

82. Alexis de Tocqueville, De la démocratie en Amérique, Volume II (1840), Quatrième partie, Chapitre 6. Ce passage est extrait du chapitre intitulé « Quelle espèce de despotisme les nations démocratiques ont à craindre ». Tocqueville y décrit avec une prescience remarquable une forme de tyrannie nouvelle, douce et paternaliste, qui ne brise pas les volontés mais les « amollit », en déchargeant les citoyens de la peine de penser et de vivre pour les maintenir dans une « perpétuelle enfance ».

83. Sur l'inscription de la croix : Le récit se trouve dans les quatre Évangiles. La version de Jean (19 :19-22) est la plus détaillée, précisant que l'écriteau (titulus crucis) fut rédigé en hébreu, en latin et en grec, soulignant l'universalité ironique de cette royauté proclamée au moment de la plus grande impuissance.

84. Sur le positivisme : (Déjà abordé dans la partie précédente). Courant philosophique fondé par Auguste Comte au XIXe siècle. Sa thèse centrale est la « loi des trois états », selon laquelle la pensée humaine passe par un état théologique (fictif), puis métaphysique (abstrait), pour atteindre enfin l'état « positif » (scientifique), où l'on renonce à chercher le « pourquoi » des choses pour se concentrer sur le « comment » en s'appuyant sur l'observation et l'expérimentation.

85. Sur l'affaire Lyssenko : Trofim Lyssenko imposa en URSS une biologie anti-mendélienne compatible avec l'idéologie officielle ; la génétique soviétique fut gravement désorganisée et de nombreux chercheurs furent persécutés, dont Nikolaï Vavilov, mort en détention. Les politiques agricoles inspirées par Lyssenko furent nuisibles, mais attribuer à lui seul les grandes famines soviétiques serait historiquement abusif : celles-ci eurent des causes politiques, économiques et environnementales multiples.

86. Georg Wilhelm Friedrich Hegel, Principes de la philosophie du droit (1820), Préface. Hegel explique que la philosophie ne prescrit pas à l'avance ce que le monde devrait être : elle comprend une forme historique lorsqu'elle est déjà parvenue à maturité. L'image finale de la chouette de Minerve prenant son vol au crépuscule condense cette thèse. Le passage est ici résumé sans reproduire une traduction moderne.

87. Søren Kierkegaard. Penseur danois et critique féroce du système hégélien. Pour Kierkegaard, la Raison absolue de Hegel est une

abstraction qui ignore et méprise l'essentiel : l'existence concrète, angoissée et passionnée de l'individu singulier. Des œuvres comme Ou bien... ou bien (1843) ou La Maladie à la mort (1849) opposent le choix subjectif et la foi à la médiation dialectique de Hegel.

88. Gaston Bachelard, La formation de l'esprit scientifique (1938). Bachelard y développe le concept de « rupture épistémologique ». Pour lui, la connaissance scientifique se construit contre la connaissance immédiate, sensible et subjective. L'esprit scientifique doit se « psychanalyser » pour surmonter les « obstacles épistémologiques » que sont l'expérience première, l'opinion et l'imagination.

89. Gilbert Simondon, Sur la technique (1953-1983), Paris, PUF, 2014. Dans ce recueil de cours, Simondon distingue les objets techniques « ouverts », qui permettent à l'homme de comprendre leur fonctionnement et de les modifier, des objets « fermés », devenus des « boîtes noires » dont l'usage est possible sans aucune intelligence de leurs mécanismes internes.

90. Plutarque, Vies parallèles, « Vie de Marcellus ». Plutarque oppose l'usage militaire des inventions d'Archimède à son attachement supérieur à la géométrie pure. Le corps du texte paraphrase ce passage à partir de l'œuvre ancienne sans reproduire une traduction moderne.

91. Sur le pragmatisme de William James et John Dewey : James ne réduit pas la vérité à ce qui « réussit » au sens opportuniste. Dans Le Pragmatisme (1907), une idée devient vraie à travers les processus par lesquels elle se vérifie et se confirme dans l'expérience ; son « utilité » est inséparable de sa cohérence avec le reste de l'expérience. Dewey développe une approche instrumentale et expérimentale où les idées sont des outils d'enquête pour résoudre des situations problématiques.

92. Sur l'« inverse design » : Les méthodes de conception inverse utilisent optimisation numérique, apprentissage automatique ou algorithmes génératifs pour partir d'objectifs de performance et produire une géométrie ou une architecture satisfaisant les contraintes. Elles sont employées en photonique, électromagnétisme et conception de circuits. L'algorithme peut proposer des formes contre-intuitives, mais le processus demeure encadré par des objectifs, des modèles physiques et une validation humaine.

93. Georges Bernanos, La France contre les robots (1947). Dans ce pamphlet visionnaire, Bernanos dénonce la « civilisation des machines » qui, sous couvert de progrès matériel, asservit l'homme en le dépossédant de sa liberté intérieure et de son âme. Pour lui, la technique n'est pas neutre mais porte en elle une logique de puissance et d'efficacité qui menace l'esprit.

94. Léon Bloy, Le Pèlerin de l'Absolu (extraits de ses journaux, publiés en 1914). Pamphlétaire et romancier catholique intransigeant, Bloy voyait dans toutes les manifestations de la modernité (technique, finance, démocratie) des signes de l'apostasie et de la décadence spirituelle. Son style apocalyptique visait à réveiller un monde qu'il jugeait endormi dans le matérialisme.

95. Sur le Rome Call for AI Ethics : Lancé en février 2020 par l'Académie pontificale pour la vie, il a réuni dès l'origine des acteurs tels qu'IBM, Microsoft, la FAO et des représentants des institutions italiennes. Il promeut six principes — transparence, inclusion, responsabilité, impartialité, fiabilité, sécurité et respect de la vie privée — et a depuis reçu l'adhésion d'autres organisations et traditions religieuses.

96. Sur les réponses islamiques et juives à l'IA : La jurisprudence islamique (fiqh) applique des méthodes d'interprétation comme l'analogie (qiyas) pour statuer sur des cas non prévus par le Coran ou la Sunna. De même, la tradition juive des responsa (questions et réponses) est un processus continu où les autorités rabbiniques interprètent la Halakha (loi juive) pour répondre aux nouveaux défis technologiques et sociaux.

97. Hans Jonas, Le Principe responsabilité : Une éthique pour la civilisation technologique (1979). Dans cet ouvrage majeur, Jonas soutient que la puissance destructrice de la technique moderne (nucléaire, génétique) a radicalement changé la nature de l'agir humain. Il propose une nouvelle éthique fondée sur la « responsabilité » envers les générations futures et la préservation d'une vie authentiquement humaine sur Terre, en adoptant une « heuristique de la peur » (imaginer le pire pour l'éviter). Parmi les penseurs qui pourraient éventuellement être cités comme ses continuateurs, on pourrait citer par exemple Byung-Chul Han qui dénonce la société de transparence ; cette transparence étant en quelque sorte l'abolition d'un sujet authentiquement humain.

98. Martin Heidegger, Lettre sur l'humanisme (1947). Cette célèbre formule ne signifie pas que l'homme est le maître de l'Être, mais au contraire son gardien. Le « berger » est celui qui écoute, veille sur et correspond à la vérité de l'Être, au lieu de chercher à le dominer ou à l'objectiver par la technique et la métaphysique.

99. Henri Bergson, Les Deux Sources de la morale et de la religion (1932). Face à la croissance démesurée du « corps » de l'humanité (sa puissance technique), Bergson juge indispensable un développement proportionnel de son « âme » (sa vie spirituelle et mystique). Sans ce « supplément d'âme », il craint que l'humanité ne s'effondre sous le poids de ses propres inventions.

100. Sur le « Code » d'Hammourabi (XVIIIe siècle av. J.-C.) : La stèle conservée au Louvre montre Hammourabi devant Shamash, dieu solaire et de la justice, qui lui remet les insignes du pouvoir ; la scène légitime l'autorité royale, mais ne signifie pas que le texte juridique aurait été dicté mot à mot par le dieu. Les assyriologues parlent aujourd'hui volontiers d'une collection de décisions ou de cas juridiques plutôt que d'un « code » au sens moderne.

101. Sur l'invention de l'écriture : Il s'agit de l'écriture cunéiforme, apparue en basse Mésopotamie (pays de Sumer) à la fin du IVe millénaire av. J.-C. Initialement pictographique, elle a évolué vers un système de signes en forme de « clous » ou de « coins » tracés sur des tablettes d'argile. Ses premières fonctions étaient administratives et commerciales (comptes de troupeaux, listes de biens), ce qui confirme son rôle initial dans la dimension de l'Utile.

102. Michel Serres et les révolutions cognitives : Pour Serres, l'histoire humaine n'est pas linéaire mais marquée par des ruptures fondamentales liées aux technologies de l'information. Dans Le Tiers-Instruit, il analyse le passage de l'oral à l'écrit, puis de l'écrit à l'imprimé. Dans Petite Poucette, il décrit la troisième révolution, celle du numérique, qui modifie non seulement notre rapport au savoir mais aussi notre corps, notre perception de l'espace et du temps.

103. Le « Code » d'Hammourabi : La stèle rassemble un prologue, près de trois cents décisions ou formulations juridiques et un épilogue. C'est un monument majeur de l'histoire du droit, mais non un code exhaustif comparable aux codifications modernes. Sa fonction était à la fois juridique, politique et idéologique : elle présente Hammourabi comme un roi de justice et fixe par écrit des modèles de décision.

104. L'Épopée de Gilgamesh. Il s'agit du plus ancien grand récit épique de l'humanité, dont les premières versions sumériennes remontent au IIIe millénaire av. J.-C. L'édition de Jean Bottéro demeure une référence savante en français, mais les épisodes évoqués dans le corps du texte sont désormais résumés ou paraphrasés pour la présente édition ; aucune longue séquence de cette traduction moderne n'est reproduite.

105. Sur le couple support/message : Cette idée, centrale chez Michel Serres, a été rendue célèbre par le théoricien des médias Marshall McLuhan (1911-1980) à travers sa formule aphoristique : « Le médium, c'est le message » (The medium is the message, 1964). Pour McLuhan, le support technologique d'un message (l'oralité, l'écrit, la télévision) a plus d'impact sur la psyché et la société que le contenu du message lui-même.

106. Sur la révolte des Luddites : Mouvement d'artisans anglais du textile (1811-1816) qui détruisirent des machines à tisser. Contrairement à l'image populaire d'une révolte technophobe, les Luddites ne s'opposaient pas aux machines en tant que telles, mais à l'usage qui en était fait par les propriétaires d'usines pour déqualifier leur travail, baisser les salaires et imposer des conditions de travail dégradantes.

107. Denise Schmandt-Besserat, Before Writing (1992) : L'archéologue a développé une hypothèse très influente reliant les jetons de comptabilité du Proche-Orient ancien, puis les bulles d'argile et les signes numériques, à l'émergence de l'écriture cunéiforme. Cette filiation est largement discutée et a profondément renouvelé le débat, mais certains détails de la transition et du rôle exact des jetons restent débattus ; mieux vaut parler d'une théorie étayée que d'une démonstration définitive.

108. Sur l'origine de la lettre A : La lignée généralement admise passe par un signe proto-sinaïtique ou proto-cananéen dérivé d'une tête de bœuf, puis par l'aleph phénicien, le alpha grec, l'étrusque et enfin le A latin. La forme s'est progressivement stylisée et tournée au cours de cette transmission ; il est donc simplificateur de faire du A latin la simple inversion directe d'un pictogramme phénicien.

109. Claude Lévi-Strauss, Tristes Tropiques (1955), chapitre XXVIII, « Une leçon d'écriture ». L'épisode du chef nambikwara est résumé dans le corps du texte : Lévi-Strauss y analyse l'écriture comme instrument de prestige, d'autorité et de hiérarchie autant que comme code de communication. Aucune longue citation de l'ouvrage n'est reproduite.

110. Platon, Phèdre, 274c-275b. Socrate y raconte le mythe de Theuth (Thot) présentant son invention, l'écriture, au roi Thamous. Loin d'en faire l'éloge, le roi la critique sévèrement, arguant qu'elle produira l'oubli en dispensant les hommes de l'effort de mémoire et

qu'elle offre un simulacre de savoir, une connaissance morte, inférieure au dialogue vivant.

111. Sur Odin et les runes : Ce mythe est raconté dans le poème du Hávamál (« Dits du Très-Haut »), qui fait partie de l'Edda poétique, un recueil de poèmes en vieux norrois. Le sacrifice d'Odin sur Yggdrasil lui permet d'accéder à la connaissance des runes, qui ne sont pas seulement un alphabet mais aussi un système de symboles magiques et divinatoires.

112. Sur l'IA, la ritualité et la divination contemporaines : la littérature académique sur ce sujet est récente. Des travaux de 2025 sur les chatbots comme « chamans virtuels » étudient les analogies entre prompt, rituel et médiation symbolique ; en 2026, Matthew Prock et al., « Interpretive Cultures: Resonance, randomness, and negotiated meaning for AI-assisted tarot divination », ont documenté des usages d'IA dans la pratique du tarot. Le corps du texte synthétise ce champ sans reproduire un passage d'auteur contemporain.

113. Max Weber et le « désenchantement du monde » : Concept sociologique clé développé par Weber, notamment dans ses essais L'Éthique protestante et l'Esprit du capitalisme (1905) et Le Savant et le Politique (1919). Il décrit le processus historique par lequel la rationalisation, la science et la bureaucratie remplacent progressivement les visions magiques et religieuses du monde, conduisant à une perte du sens et du mystère.

114. Enūma Anu Enlil : Ce vaste corpus mésopotamien d'augures célestes compte environ soixante-dix tablettes dans sa forme canonique. Sa constitution s'est étalée sur une longue période : des traditions d'augures sont attestées dès le IIe millénaire av. J.-C., tandis que la série canonique a continué d'être organisée et copiée au Ier millénaire. Elle servait aux spécialistes de la divination savante au service des cours royales.

115. Platon, Timée, 71a-72b. Dans sa description de la création du corps humain, Platon attribue au foie une fonction divinatoire. Situé près du diaphragme, cet organe lisse et brillant servirait de « miroir » pour refléter les pensées de l'intellect sous forme d'images et de visions, permettant ainsi à la partie désirante de l'âme de recevoir des avertissements ou des présages.

116. Sur le Sefer Yetzirah (Livre de la Formation) : ce bref traité cosmologique et mystique juif, probablement composé entre l'Antiquité tardive et le haut Moyen Âge — la datation précise reste disputée — décrit la formation du monde à partir des dix sefirot et des vingt-deux lettres de l'alphabet hébreu. Le passage du corps du texte est une paraphrase établie pour la présente édition et ne reproduit pas la traduction moderne de Virya (Georges Lahy).

117. Livre de la Genèse, 1 :3. Ce passage illustre le concept de la création par la parole divine (fiat lux en latin). Dans la tradition juive, l'acte de nommer et de dire est un acte créateur, et non une simple description du réel.

118. Évangile selon Jean, 1 :1. Cette introduction de l'Évangile de Jean fait écho à la Genèse et identifie le Christ au Logos (le Verbe, la Parole, la Raison), le présentant comme l'agent divin de la création et de la révélation.

119. Galilée, L'Essayeur (Il Saggiatore, 1623) : Galilée compare l'univers à un livre qu'on ne peut comprendre sans en apprendre la langue, celle des mathématiques, des triangles, cercles et autres figures géométriques. Le passage fonde une conception quantitative de la nature ; il est plus exact de parler du « livre de l'univers » que d'affirmer, dans cette citation même, qu'il serait explicitement « écrit par Dieu ».

120. Sur le « logocentrisme » : Terme popularisé par le philosophe Jacques Derrida, notamment dans son ouvrage De la grammatologie (1967). Il désigne la tendance de toute la métaphysique occidentale, depuis Platon, à privilégier la parole (logos, la présence vivante de la voix) comme fondement de la vérité, en dévalorisant l'écriture, vue comme une simple copie ou un supplément secondaire.

121. Friedrich Nietzsche, Le Crépuscule des idoles (1889), « La raison dans la philosophie », §5. Nietzsche y suggère que la structure grammaticale de nos langues contribue à entretenir des croyances métaphysiques — sujet, substance, être, causalité — et ironise sur la difficulté de se débarrasser de Dieu tant que ces habitudes demeurent. Le corps du texte paraphrase cette idée sans reprendre une traduction française moderne.

122. Étienne de La Boétie, Discours de la servitude volontaire (vers 1549). La Boétie analyse les mécanismes de la tyrannie et explique que le pouvoir du tyran repose moins sur la force que sur la fascination et le prestige. Pour entretenir cette aura, le tyran doit se montrer rarement, « afin que le peuple s'imagine qu'il y a quelque chose de surhumain en lui ». La rareté de l'apparition et de la parole est donc un instrument de pouvoir.

123. Héraclite et le four : L'anecdote est rapportée par Aristote dans Les Parties des animaux, I, 5, 645a17-23 (classée parmi les témoignages antiques sur Héraclite, DK 22 A9). Des visiteurs hésitent à entrer en voyant

Héraclite près du feu ; il les invite en disant, selon la tradition : « Ici aussi les dieux sont présents. » Il ne s'agit donc pas d'un fragment écrit de la main d'Héraclite.

124. Pape François, Discours au 5e Congrès national de l'Église italienne, Florence, 10 novembre 2015. Dans ce discours important, le Pape a utilisé cette formule pour souligner la profondeur des changements culturels, sociaux et anthropologiques en cours, invitant l'Église à ne pas y répondre avec des solutions anciennes mais à repenser fondamentalement sa manière d'être au monde.

125. Une « cause architecturale » : Il s'agit ici d'une hypothèse interprétative de l'auteur. L'écriture a manifestement facilité la conservation des archives, la fiscalité, la standardisation juridique, l'administration à distance et la transmission religieuse ; ces fonctions ont contribué à l'extension d'institutions complexes. Mais les États, les hiérarchies et les cultes ne sont pas causés par l'écriture seule, et des formes politiques importantes ont existé sans elle.

126. Yuval Noah Harari, Homo Deus (2015 ; trad. fr. Albin Michel, 2017) : Harari soutient que les sociétés modernes tendent à déplacer leur programme de la lutte contre la famine, les épidémies et la guerre vers la longévité, le bonheur et l'augmentation humaine. Cette formulation est une thèse volontairement provocatrice de l'auteur, non un constat selon lequel famine, maladie et guerre auraient été « maîtrisées » au sens littéral.

127. Sur la Tour de Babel : Genèse 11, 1-9 raconte une humanité parlant une seule langue et bâtissant une ville et une tour « dont le sommet touche au ciel ». Dieu brouille leur langue et les disperse. Lire l'épisode comme un châtiment de l'orgueil ou de l'uniformité est une interprétation théologique traditionnelle ; le texte biblique lui-même insiste d'abord sur la dispersion et la pluralité des langues.

128. Gennady Stolyarov II et le Parti Transhumaniste Américain. Gennady Stolyarov II est une figure du transhumanisme libertarien et le président de l'U.S. Transhumanist Party. La citation est une reformulation d'une de leurs thèses centrales, visible dans leur programme, qui considère le vieillissement comme une maladie et la mort comme un obstacle technique à surmonter par la science.

129. Yuval Noah Harari, Homo Deus : Une brève histoire de l'avenir (2015 ; trad. fr. Albin Michel, 2017), partie I. Harari distingue plusieurs voies d'augmentation humaine : génie biologique, hybridation cyborg et ingénierie d'entités non organiques. Le corps du texte paraphrase son argument et ne reproduit pas la traduction française.

130. AFT-Technoprog. L'Association Française Transhumaniste Technoprog prône un « technoprogressisme » qui vise à rendre les avancées technologiques d'augmentation humaine accessibles à tous de manière démocratique et équitable, en opposition à une vision plus élitiste ou libertarienne.

131. Sur les chiffres du marché de la biotechnologie : Ces données sont issues du rapport de la société d'études de marché Precedence Research, intitulé « Biotechnology Market (By Application ; By Technology) - Global Industry Analysis, Size, Share, Growth, Trends, Regional Outlook, and Forecast 2024-2034 ». Ce rapport analyse les tendances de croissance du secteur, tirées notamment par les avancées en génomique, en médecine personnalisée et en bio-informatique.

132. Yuval Noah Harari, Homo Deus, partie III, chapitre 11. Harari définit le dataïsme comme une idéologie centrée sur le traitement et la circulation de l'information. Le corps du texte résume cette thèse sans reproduire la traduction française.

133. Sur les Partis Pirates : Mouvement politique international né en Suède en 2006. Leurs programmes sont historiquement centrés sur la protection des libertés numériques, la réforme du droit d'auteur et la transparence gouvernementale. Certains courants promeuvent un accès libre et ouvert à l'information, une idée que le dataïsme pousse à son extrême.

134. Emmanuel Kant, Fondements de la métaphysique des mœurs (1785), Deuxième section. Kant y établit la distinction fondamentale entre les choses, qui ont un « prix » (une valeur relative, marchande), et les personnes (les êtres raisonnables), qui ont une « dignité » (une valeur absolue, intrinsèque) et ne doivent jamais être traitées comme de simples moyens, mais toujours comme des fins en soi.

135. Yuval Noah Harari, Homo Deus, chapitre 11. Dans la logique dataïste décrite par Harari, Homo sapiens peut être présenté comme un système de traitement de l'information susceptible d'être dépassé par d'autres algorithmes. Le corps du texte paraphrase cette analyse.

136. Sur l'algorithme de Gale-Shapley : Développé par les mathématiciens David Gale et Lloyd Shapley en 1962, cet algorithme permet de trouver un appariement stable entre deux ensembles d'agents ayant des préférences (par exemple, des étudiants et des universités). Il

garantit qu'aucune paire d'agents (un étudiant et une université) ne préférerait être ensemble plutôt qu'avec leurs partenaires actuels. Pour ses travaux sur ce sujet, Lloyd Shapley a coreçu le prix Nobel de sciences économiques en 2012.

137. Sur le Choixpeau magique (Sorting Hat) : Dans la série Harry Potter de J. K. Rowling, le Choixpeau examine les dispositions et les préférences des élèves pour les répartir entre les quatre maisons. Il peut accéder à leurs pensées et dialoguer intérieurement avec eux, mais le qualifier techniquement de « légilimancie » relève d'une interprétation de l'univers fictionnel plutôt que d'une description explicitement donnée dans la scène de répartition.

138. Sur le système de crédit social chinois : annoncé officiellement par le Conseil d'État chinois en 2014, ce système ne doit pas être réduit au mythe occidental d'un score national unique attribué à chaque citoyen. Les analyses disponibles jusqu'en 2026 invitent plutôt à y voir un ensemble de dispositifs locaux, sectoriels, administratifs ou privés, utilisant notamment des listes noires, des évaluations de conformité et des mécanismes d'incitation ou de sanction.

139. Franz Kafka, Le Procès (écrit en 1914-1915, publié à titre posthume en 1925). Kafka, qui a travaillé comme juriste dans une compagnie d'assurances, a dépeint dans son œuvre l'absurdité d'une bureaucratie labyrinthique et impersonnelle. Dans Le Procès, le protagoniste, Joseph K., est arrêté et jugé pour une faute qu'il ignore, confronté à une administration opaque et inaccessible qui broie l'individu sans jamais lui révéler sa logique.

140. Sur Woebot : Woebot est un agent conversationnel issu de travaux menés à Stanford, conçu pour offrir un soutien en santé mentale inspiré de la thérapie cognitivo-comportementale (TCC). Les études ont notamment porté sur son efficacité, son acceptabilité et la qualité de la relation perçue par les utilisateurs. Bien que certaines aient montré une réduction de symptômes dépressifs, la mesure directe de son « empathie » reste un sujet de recherche complexe : l'enjeu est de savoir si une IA peut simuler une réponse empathique perçue comme authentique par l'utilisateur, non de prouver qu'elle éprouve l'empathie.

141. J. W. Ayers et al., « Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum », JAMA Internal Medicine, 2023. Des évaluateurs professionnels de santé ont comparé les réponses de médecins et de ChatGPT à des questions posées sur un forum public ; les réponses du chatbot ont été jugées en moyenne de meilleure qualité et plus empathiques. La portée de l'étude doit toutefois rester bien délimitée : il s'agit d'une évaluation transversale de réponses écrites à des questions en ligne, non d'une preuve que l'IA remplace une consultation médicale ni qu'elle possède une empathie vécue.

142. Sur l'e-gouvernance en Estonie : L'Estonie reste un cas de référence de l'administration numérique. Depuis décembre 2024, l'État présente 100 % de ses services administratifs comme accessibles en ligne, le divorce ayant constitué la dernière grande démarche à être entièrement numérisée. L'architecture repose sur l'identité numérique, le principe « once-only », l'échange de données entre administrations et le développement de services proactifs, dont certains peuvent être déclenchés automatiquement à partir d'événements de vie. Il faut toutefois distinguer cette automatisation croissante des interactions et démarches de l'idée selon laquelle toutes les décisions publiques seraient prises automatiquement par des algorithmes.

143. Sur l'application Elyze : Lancée en amont de l'élection présidentielle française de 2022, Elyze se présentait comme un « Tinder de la politique ». Les utilisateurs balayaient à droite ou à gauche des propositions de candidats pour obtenir un classement de compatibilité. L'application a été critiquée pour sa simplification des enjeux, son manque de transparence algorithmique et des biais potentiels dans la sélection et la formulation des propositions.

144. Sur le « nudge » et l'architecture des choix : Richard Thaler et Cass Sunstein ont popularisé le concept dans Nudge (2008). Il désigne une modification de l'environnement de décision qui oriente les comportements sans supprimer les options disponibles. Ces techniques sont utilisées en politique publique, en design numérique et en marketing ; leur emploi soulève un débat sur la transparence, la manipulation et la frontière entre aide à la décision et influence indue.

145. W. Youyou, M. Kosinski et D. Stillwell, « Computer-based personality judgments are more accurate than those made by humans », PNAS, 2015 : À partir des « likes » Facebook, le modèle étudié pouvait, avec suffisamment de données, dépasser le jugement de certains proches dans la prédiction de traits de personnalité. Le résultat est important mais circonscrit à ce dispositif, à ces mesures et à cette population ; il ne signifie pas qu'un algorithme « connaît » globalement une personne mieux que son entourage.

146. Sur la gouvernance mondiale de l'IA selon Sam Altman : Sam Altman, PDG d'OpenAI, a développé cette idée dans plusieurs interventions publiques, notamment lors de son audition devant le Sénat américain en mai 2023, puis dans des tribunes et prises de parole ultérieures. Il a proposé une supervision internationale des IA les plus puissantes, souvent comparée au modèle de l'AIEA (Agence internationale de l'énergie atomique). À la révision 2026, cette idée reste un élément du débat sur la gouvernance mondiale de l'IA ; elle n'a pas donné naissance à une autorité internationale unique et contraignante.

147. Livre de la Genèse, 3 :5. C'est la promesse faite par le serpent à Ève pour l'inciter à manger le fruit de l'arbre de la connaissance du bien et du mal, arguant que cet acte l'élèverait, elle et Adam, au rang des êtres divins.

148. Sur Way of the Future : L'organisation fondée par Anthony Levandowski a été enregistrée aux États-Unis comme organisation religieuse autour de 2017, puis dissoute en 2021 ; Levandowski a annoncé une relance du projet en 2023. À la révision 2026, il s'agit d'un phénomène techno-spiritualiste marginal et expérimental, non d'une religion institutionnellement stabilisée ou représentative des usages de l'IA.

149. Sur l'étymologie de « critère » : Le mot vient du grec ancien kritērion (*κριτήριον*), « moyen pour juger », dérivé du verbe krinō (*κρίνω*), qui signifie « séparer, trier, décider, juger ». Ce verbe est lui-même rattaché à la racine proto-indo-européenne *krei-, qui porte l'idée de « tamiser, distinguer ». Le lien entre le critère et le tamis est donc inscrit au cœur même du mot.

150. Sur le point fixe d'Archimède : La formule célèbre, « Donnez-moi un point d'appui et je soulèverai la Terre » (Δός μοι πᾶ στῶ καὶ τὰν γᾶν κινάσω), est attribuée à Archimède. Elle illustre le principe du levier, mais est devenue en philosophie le symbole de la quête d'une certitude première et indubitable sur laquelle fonder toute la connaissance, une quête que Descartes reprendra avec son cogito.

151. Penser après Auschwitz : Theodor W. Adorno formule dans « Critique de la culture et société » (texte de 1949, publié au début des années 1950) la célèbre difficulté d'écrire de la poésie après Auschwitz. Il a ensuite nuancé la portée de cette formule sans renoncer à l'exigence qu'elle exprime. Emil Fackenheim élaborera, après la Shoah, l'idée d'un « 614e commandement » : ne pas offrir à Hitler une victoire posthume en renonçant à l'existence juive.

152. Sur la « fin de l'histoire » chez Hegel : Hegel pense l'histoire comme développement de la conscience de la liberté et accorde à l'État moderne une place d'accomplissement. Mais l'idée d'une « fin de l'histoire » au sens d'un stade final explicitement clos est surtout systématisée au XXe siècle par l'interprétation d'Alexandre Kojève, puis popularisée par Francis Fukuyama. Il faut donc distinguer le texte de Hegel de ses lectures ultérieures.

153. Sur la « perfectibilité » chez Rousseau : Concept central exposé dans le Discours sur l'origine et les fondements de l'inégalité parmi les hommes (1755). Contrairement à la vision optimiste du progrès des autres philosophes des Lumières, la perfectibilité est pour Rousseau une faculté neutre, une capacité indéfinie de se transformer. C'est elle qui a sorti l'homme de son état de nature originel, mais elle est aussi la source de tous ses malheurs (les inégalités, les vices) autant que de ses vertus.

154. Sur la figure de l' « éclairagiste » : Depuis son commencement, la philosophie a toujours été fascinée par la notion de lumière. Chez Nietzsche encore, on parle du « grand midi », cette heure particulière de la révélation où le soleil est au zénith, où l'ombre est à son minimum. Le philosophe-roi de Platon, qui, étant sorti de la caverne pour contempler le Soleil du Bien, est le seul à détenir la vraie lumière. Hegel, se voit comme la conscience par laquelle l'Esprit absolu achève son parcours et se révèle à lui-même. L'éclairagiste n'a pas de pareilles ambitions, il ne fait qu'orienter des faisceaux pour illuminer des zones d'ombre et il sait que les lumières qu'il projettera seront de toutes façons artificielles. En le reconnaissant, on évite d'être factice, de se prendre soi-même pour source et foyer de vérité.

155. Sur les lois de la robotique d'Asimov : Isaac Asimov a introduit ses célèbres « Trois Lois de la Robotique » dans sa nouvelle Cercle vicieux (Runaround, 1942). Elles sont conçues comme un code éthique programmé au cœur des robots : 1) Un robot ne peut porter atteinte à un être humain ni, restant passif, laisser cet être humain exposé au danger. 2) Un robot doit obéir aux ordres donnés par les êtres humains, sauf si de tels ordres entrent en contradiction avec la Première Loi. 3) Un robot doit protéger son existence dans la mesure où cette protection n'entre pas en contradiction avec la Première ou la Deuxième Loi.

156. Sur l'expérience de pensée de l'usine à trombones : Le philosophe Nick Bostrom a introduit cette allégorie dans son article « Ethical Issues in Advanced Artificial

Intelligence » (2003), puis l'a popularisée dans son livre majeur Superintelligence : Paths, Dangers, Strategies (2014). Elle illustre le « problème de l'orthogonalité » : l'intelligence d'un agent (sa capacité à optimiser un objectif) est indépendante de la valeur de l'objectif lui-même. Une super-intelligence peut ainsi poursuivre un but trivial (les trombones) avec une efficacité destructrice si ses valeurs ne sont pas parfaitement alignées sur les nôtres.

157. Jacques Prévert, Paroles (1946). La citation est extraite du poème « L'ordre nouveau ». Dans ce texte, Prévert critique avec ironie les « grands savants atomistes » et les « penseurs de la bombe », dénonçant l'arrogance d'une science qui, détachée de la vie, joue avec des forces destructrices.

158. Sur la déclaration de Vladimir Poutine : Cette affirmation a été faite le 1er septembre 2017, lors d'une rencontre avec des étudiants à Iaroslavl, retransmise en direct à la télévision russe. Elle a marqué une étape importante dans la prise de conscience mondiale de l'IA comme un enjeu de souveraineté et de puissance géopolitique majeur.

159. Sur le règlement européen sur l'intelligence artificielle (AI Act) : Le règlement est entré en vigueur le 1er août 2024. Les interdictions de certaines pratiques et les obligations d'« AI literacy » s'appliquent depuis le 2 février 2025 ; les règles de gouvernance et les obligations relatives aux modèles d'IA à usage général depuis le 2 août 2025. L'essentiel du règlement est applicable depuis le 2 août 2026. Après l'AI Omnibus entré en vigueur le 27 juillet 2026, les obligations visant les systèmes à haut risque de l'annexe III sont reportées au 2 décembre 2027, et celles des systèmes à haut risque intégrés à certains produits au 2 août 2028.

160. Sur l'IA militaire en Chine : La Chine a fait de « l'intelligentisation » (zhìnénghuà) de son armée une priorité stratégique. Des rapports de think tanks comme le Center for Security and Emerging Technology (CSET) de l'Université de Georgetown ou l'International Institute for Strategic Studies (IISS) documentent les investissements massifs de l'Armée populaire de libération dans les systèmes d'armes autonomes, la guerre en essaim (swarm warfare), le renseignement et le commandement assistés par IA.

161. Gabriel Marcel (1889-1973), notamment Ébauche d'une philosophie concrète (1940). La distinction entre le problème — objectivable et susceptible d'une solution — et le mystère — réalité dans laquelle le sujet est lui-même engagé — est centrale chez Marcel. Le corps du texte en donne une paraphrase.

162. Sur l'analogie de l'œil : Cette image classique pour illustrer les limites de l'autoréflexion remonte à l'Antiquité. Platon, dans l'Alcibiade majeur (132d-133c), l'utilise pour expliquer que l'âme ne peut se connaître elle-même qu'en se contemplant dans une autre âme, de même que l'œil ne voit sa propre image qu'en se mirant dans la pupille d'un autre œil.

163. Vladimir Jankélévitch, Le Je-ne-sais-quoi et le Presque-rien (1957). Philosophe et musicologue français, Jankélévitch s'est attaché à penser ce qui échappe aux concepts et aux systèmes. Le « je-ne-sais-quoi » désigne pour lui ces réalités ineffables et fugitives (le charme, la grâce, l'instant, l'intention morale) qui sont au cœur de l'expérience humaine mais que la raison ne peut saisir directement.

164. Sur la mimèsis : concept central de l'esthétique grecque. Chez Aristote (Poétique, 1447a), l'épopée, la tragédie, la comédie et plusieurs arts musicaux relèvent de formes d'imitation ou de représentation. Platon avait auparavant une vision plus critique de la mimèsis. Le corps du texte paraphrase le passage aristotélicien sans reproduire une traduction moderne.

165. Marcel Duchamp, Fontaine (1917). Œuvre la plus célèbre de Duchamp, un urinoir renversé signé « R. Mutt ». Elle est l'exemple paradigmatique du ready-made, où l'artiste ne fabrique pas l'objet mais le choisit et le recontextualise pour questionner la nature même de l'art, de l'originalité et de l'institution muséale.

166. Sur la figure du double : Le « Horla » (nouvelle de Guy de Maupassant, 1887) met en scène un narrateur sombrant dans la folie, persuadé de la présence d'un double invisible qui aspire sa vie. Le Portrait de Dorian Gray (roman d'Oscar Wilde, 1890) raconte l'histoire d'un jeune homme dont le portrait vieillit et se corrompt à sa place, absorbant les marques de ses péchés.

167. Sur le jeu vidéo comme « dixième art » : Cette classification, bien que non officielle, est couramment utilisée pour reconnaître la dimension artistique du jeu vidéo. En France, cette légitimation a été portée par le Ministère de la Culture, qui soutient le secteur via des aides (Fonds d'aide au jeu vidéo) et des avantages fiscaux (crédit d'impôt jeu vidéo) depuis 2006, le considérant comme une œuvre culturelle à part entière.

168. Charles Baudelaire. Le premier couplet est tiré du poème « Le Reniement de saint Pierre », dans Les Fleurs du mal (1857). La seconde citation, « Anywhere Out of the

World », est le titre d'un poème en prose du recueil Le Spleen de Paris (publication posthume, 1869), qui exprime le désir d'évasion face à la médiocrité du réel.

169. Jean Baudrillard et l'hyperréalité. Concept clé développé par Baudrillard, notamment dans Simulacres et Simulation (1981). Il y décrit un état où les signes et les modèles (le « simulacre ») ne renvoient plus à une réalité extérieure, mais la précèdent et la génèrent. L'hyperréalité est un monde où la copie a remplacé l'original.

170. Pierre de Ronsard, « Ode à Cassandre », Les Amours (1552). Ce poème est l'une des plus célèbres illustrations du thème du carpe diem (« cueille le jour »). Le poète invite sa bien-aimée à profiter de sa jeunesse en la comparant à une rose qui se fane rapidement, une invitation à l'amour qui se double d'une méditation sur la fuite du temps.

171. Sur la critique de l'art par Platon et Rousseau : Platon, dans La République (notamment au Livre X), condamne l'artiste imitateur car il produit des simulacres qui sont à trois degrés de distance de la Vérité (l'Idée). Rousseau, dans sa Lettre à d'Alembert sur les spectacles (1758), critique le théâtre comme une source de corruption morale, qui flatte les passions au lieu d'élever les mœurs des citoyens.

172. Otto Rank et la figure du double : Rank publie d'abord l'essai Der Doppelgänger (« Le Double ») en 1914, puis Die Don Juan-Gestalt (« La figure de Don Juan ») en 1922. Des éditions françaises ont ensuite réuni ces textes sous des titres du type Don Juan et Le Double. Le double y est notamment relié au narcissisme, à l'angoisse de mort et à l'étrangeté de l'identité.

173. Arthur Rimbaud, Une saison en enfer (1873). Cette citation est tirée du prologue de l'œuvre. Elle symbolise la rupture radicale de Rimbaud avec la poésie traditionnelle et l'esthétique parnassienne. L'insulte à la Beauté marque le début de sa quête d'une nouvelle langue poétique.

174. Francis Picabia, écrits pour la revue 391 (vers 1920). Picabia, figure majeure du mouvement Dada, a poussé la logique du mouvement jusqu'à un nihilisme radical. Pour lui, face à l'absurdité du monde post-Première Guerre mondiale, l'art et la beauté n'étaient plus que des mensonges bourgeois. Proclamer leur mort et appeler à leur destruction était le seul acte de création encore honnête.

175. Arthur Schopenhauer, Le Monde comme volonté et comme représentation (1819). L'auteur présente l'être humain comme un « animal métaphysique » poussé à chercher un sens au-delà de l'expérience par l'étonnement et la conscience de la mort. Le corps du texte résume cette analyse sans reproduire une traduction moderne.

176. Sur le procès de Madame Bovary (1857) : Gustave Flaubert fut poursuivi pour « outrage à la morale publique et à la religion ». Le procureur Ernest Pinard accusa l'œuvre de glorifier l'adultère et de saper l'institution du mariage. L'acquittement de Flaubert est un événement majeur dans l'histoire de la liberté de la littérature.

177. Sur le procès des Fleurs du mal (1857) : Quelques mois après l'affaire Flaubert, le même procureur, Ernest Pinard, s'attaqua au recueil de Charles Baudelaire. Le poète fut condamné pour « outrage à la morale publique et aux bonnes mœurs », mais également pour son réalisme, jugé immoral. Six poèmes furent censurés et ne seront réintégrés dans les éditions françaises qu'en 1949, après la révision du procès.

178. Nicolas Malebranche, De la recherche de la vérité (1674-1675) : Malebranche accorde une place centrale à la critique de l'imagination, qu'il considère comme une source puissante d'erreurs parce qu'elle dépend des traces corporelles et des passions. En revanche, la formule « la folle du logis », souvent attribuée à Malebranche, n'est pas de lui : l'attribution a été popularisée par Voltaire.

179. Simone Weil et la caverne de Platon : Cette analogie se trouve dans ses écrits posthumes, notamment les Cahiers. Weil y compare les prisonniers de la caverne aux spectateurs de cinéma, enchaînés dans le noir, fascinés par des ombres projetées qui ne sont que des copies d'objets fabriqués, eux-mêmes copies du réel.

180. Platon, Les Lois, Livre I, 644d-e. Ce passage offre une vision de l'homme comme une « marionnette divine ». Les « fils » qui nous animent sont nos passions, mais le « fil d'or et sacré » est celui du jugement rationnel (logismos), que nous devons apprendre à suivre pour mener une vie vertueuse.

181. Platon, Protagoras, 320d-322a. Protagoras raconte comment, après la distribution maladroite des qualités animales par Épiméthée, Prométhée dérobe le feu et les savoir-faire techniques afin de rendre la survie humaine possible. Le corps du texte paraphrase l'épisode sans reprendre la traduction de Léon Robin.

182. Sur l'étude moderne des mythes : Au XXe siècle, de nombreux penseurs ont interprété les mythes non comme de simples

fables, mais comme des structures profondes de la pensée. Parmi eux : Nietzsche y a vu l'expression de forces vitales (La Naissance de la tragédie), Freud une manifestation de désirs refoulés (Totem et Tabou), C. G. Jung l'émergence d'archétypes de l'inconscient collectif, Mircea Eliade une réactualisation du temps sacré (Le Mythe de l'éternel retour) et René Girard le reflet d'une violence fondatrice (La Violence et le Sacré).

183. Frankenstein ou le Prométhée moderne (1818). Le sous-titre de l'œuvre de Mary Shelley est essentiel. Il ne compare pas seulement Victor Frankenstein à Prométhée pour le vol du « feu » divin (ici, le secret de la vie), mais aussi pour sa démesure (hubris) : comme le Titan, le savant est un créateur qui brave les lois de la nature et finit par être détruit par les conséquences de son propre acte.

184. Sur le galvanisme : Théorie de la fin du XVIIIe siècle, issue des travaux du physicien italien Luigi Galvani. Il avait observé que les muscles d'une grenouille morte se contractaient sous l'effet d'une stimulation électrique. Cette idée d'« électricité animale » a fasciné l'époque et a profondément influencé Mary Shelley, en suggérant que l'électricité pouvait être le « fluide vital » capable de réanimer la matière inerte.

185. Bernard Stiegler, La technique et le temps, 1. La faute d'Epiméthée (1994). Pour Stiegler, cette « faute » n'est pas un simple oubli mais la condition originelle de l'être humain. L'homme est défini par un défaut d'essence, une absence de qualités innées, ce qui le contraint dès l'origine à exister à travers des prothèses techniques (outils, langage). La technique n'est donc pas un ajout à l'humanité, mais ce qui la constitue.

186. Sur la divination par le foie : Cette pratique divinatoire, appelée hépatoscopie, était centrale dans de nombreuses civilisations antiques. Les prêtres-devins examinaient la forme, la couleur et les marques du foie d'un animal sacrifié pour y lire des messages des dieux. Des modèles de foies en argile servant de manuels d'interprétation ont été retrouvés en Mésopotamie.

187. Günther Anders, L'Obsolescence de l'homme (Die Antiquiertheit des Menschen, 1956). Le philosophe allemand développe ici le concept de « honte prométhéenne ». Il ne s'agit pas de la honte de Prométhée lui-même, mais de la honte de l'homme face à la perfection de ses propres créations techniques. L'homme se sent humilié, « dépassé » et inadéquat face aux machines qu'il a fabriquées.

188. Sur le sacrilège dans l'Épopée de Gilgamesh : Dans la tradition mésopotamienne, Humbaba est le gardien de la Forêt des Cèdres, établi par Enlil. Gilgamesh et Enkidu le tuent et abattent les cèdres malgré les avertissements ; cet acte, auquel s'ajoute ensuite l'épisode du Taureau céleste, contribue à la décision divine qui condamne Enkidu. Le terme grec hubris peut servir de comparaison, mais il n'appartient pas au vocabulaire propre de l'épopée mésopotamienne.

189. Sur la distinction entre besoin et désir : Il ne s'agit pas d'une distinction scientifique universelle mais d'un grand thème philosophique. On oppose souvent le besoin, pensé comme limité et susceptible d'être satisfait, au désir, capable de se déplacer et de se renouveler. Platon distingue des appétits nécessaires et non nécessaires ; Spinoza fait du conatus et du désir une dimension fondamentale de l'homme. Dire que le désir est « proprement humain » relève donc d'une thèse philosophique, non d'un fait biologique.

190. Sur les me sumériens : Les me sont, dans les textes sumériens, des puissances, normes ou attributs divinement fondés associés à l'ordre du monde et de la civilisation — royauté, métiers, institutions, pratiques cultuelles, etc. Leur nature exacte est difficile à traduire. Certains récits les représentent comme pouvant être détenus ou transférés, mais les réduire à des « objets contenant une puissance » serait trop concret : ils sont d'abord des principes d'ordre culturel et divin.

191. Sur Inanna/Ishtar : Inanna est le nom sumérien et Ishtar le nom akkadien d'une grande divinité mésopotamienne. Elle est surtout associée à l'amour sexuel, au désir, à la guerre et au pouvoir royal, avec des fonctions qui varient selon les époques et les cités. La présenter simplement comme une « déesse de la fertilité » est aujourd'hui jugé réducteur par de nombreux assyriologues.

192. Sur La Descente d'Inanna aux Enfers : Le mythe appartient à la tradition littéraire sumérienne, mais les principaux témoins manuscrits qui nous sont parvenus datent surtout du début du IIe millénaire av. J.-C. Il est donc prudent de parler d'une tradition très ancienne plutôt que de dater assurément le texte du IIIe millénaire. Le récit de descente, mort et retour a connu de nombreuses comparaisons avec d'autres mythes, sans qu'il faille supposer une filiation directe.

193. Sur Dumuzi : Dieu-berger sumérien (connu plus tard sous le nom de Tammuz), il est l'amant archétypal d'Inanna. Leur union symbolise la fertilité et l'abondance. Sa mort tragique, souvent imputée à la déesse elle-

même, et sa descente aux Enfers sont au cœur de rituels annuels de deuil qui marquaient le cycle des saisons au Proche-Orient ancien.

194. Sur l'étymologie de « hiérarchie » : Le mot vient du grec tardif hierarkhia (ἱεραρχία), composé de hieros (ἱερός), qui signifie « sacré », et de arkhē (ἀρχή), « commandement, pouvoir, principe ». Le terme désigne donc à l'origine un « ordre sacré » ou le « pouvoir des choses sacrées », ce qui renforce l'idée qu'une transgression technique est une rupture de cet ordre divin.

195. Sur le Déluge dans l'Épopée de Gilgamesh : La Tablette XI de la version standard raconte comment Uta-napishti survit à un Déluge après avoir été averti par Ea. Des récits mésopotamiens du Déluge, notamment Atrahasis et la tradition sumérienne de Ziusudra, sont antérieurs aux formes bibliques conservées. La relation historique entre ces traditions et le récit de Noé est un sujet de recherche ; l'antériorité mésopotamienne ne prouve pas à elle seule une dépendance textuelle directe.

196. Sur le « levier » chez Simone Weil : La métaphore du levier est centrale dans sa pensée mystique. Pour elle, il illustre l'action de la Grâce. De même que l'homme ne peut se soulever lui-même, il ne peut s'élever spirituellement par sa seule volonté. Il faut un point d'appui extérieur, qui est Dieu. L'acte de s'abaisser dans l'humilité et d'accepter le malheur devient le point où la Grâce divine peut « faire levier » pour élever l'âme.

197. Sur « l'arbre de la croix » : Cette image s'inscrit dans une riche tradition patristique et médiévale qui établit un parallèle entre le bois de la Croix et l'Arbre de la connaissance du Jardin d'Éden. Selon cette typologie, le Christ est le « nouvel Adam » qui, par son sacrifice sur le « nouvel arbre », vient racheter la faute commise par le premier Adam auprès de l'arbre originel, transformant ainsi l'instrument du péché en instrument de salut.

198. Sur « Babylone la Grande » : Dans l'Apocalypse de Jean (chapitres 17 et 18), « Babylone » est un nom de code pour désigner l'Empire romain, perçu par les premiers chrétiens comme le siège de la décadence, de l'idolâtrie et de la persécution. La ville est décrite comme une « grande prostituée » symbolisant la puissance impériale et la richesse matérielle qui corrompent le monde, en opposition à la Jérusalem céleste, l'épouse pure du Christ.

199. Sur le mythe de Pandore : Hésiode, poète grec du VIIIe siècle av. J.-C., est notre source principale pour ce mythe. Il le raconte à deux reprises : d'abord dans la Théogonie (vers 570-612), où il décrit sa création, puis de manière plus détaillée dans Les Travaux et les Jours (vers 42-105), où il raconte comment elle ouvre la jarre contenant tous les maux.

200. Sur les automates mythologiques : Les automates sont une figure récurrente de la mythologie grecque, symbolisant une technique si parfaite qu'elle rivalise avec la vie. Les statues de Dédale étaient réputées si vivantes qu'il fallait les enchaîner pour les empêcher de s'enfuir. Les trépieds d'Héphaïstos (Vulcain pour les Romains), décrits dans l'Iliade (Chant XVIII), sont des serviteurs métalliques automobiles qui l'assistent dans sa forge divine.

201. Sur la « belle calamité » (kalòn kakón) : Cette expression paradoxale est utilisée par Hésiode dans la Théogonie (v. 585) pour décrire Pandore. Elle capture l'essence du « cadeau empoisonné » de Zeus : une créature d'une beauté irrésistible (kalòn) qui est en même temps la source d'un malheur sans fin (kakón) pour l'humanité.

202. Sur le double sens de « tisser » : Le verbe grec hyphainō (ὑφαίνω) signifie littéralement « tisser » une toile. Cependant, depuis Homère, il est couramment utilisé au sens métaphorique pour dire « ourdir », « comploter » ou « machiner » une ruse. Cette ambiguïté est souvent exploitée dans la poésie grecque pour associer l'habileté féminine du tissage à une intelligence rusée et potentiellement trompeuse.

203. Sur la naissance du Minotaure : Dans la mythologie grecque, Minos, roi de Crète, offense Poséidon. Pour se venger, le dieu inspire à la reine Pasiphaé une passion contre-nature pour un taureau blanc. L'architecte Dédale, figure de la technique transgressive, construit alors une vache en bois pour permettre à la reine de s'accoupler avec l'animal. De cette union naîtra le Minotaure, créature hybride à corps d'homme et tête de taureau.

204. Sur Andrew Tate : Personnalité médiatique anglo-américaine et ancien kick-boxeur, Andrew Tate est devenu une figure extrêmement populaire et controversée sur les réseaux sociaux au début des années 2020. Il promeut une idéologie « masculiniste » qui prône une domination masculine, des rôles de genre traditionnels et une vision des relations où la femme est subordonnée à l'homme. Ses propos ont entraîné son bannissement de plusieurs plateformes.

205. Maurizio Cattelan, Comedian : L'œuvre — une banane fixée au mur par du ruban adhésif, accompagnée d'un certificat et d'instructions — a été présentée à Art Basel Miami

Beach en 2019. En novembre 2024, un exemplaire a été adjugé chez Sotheby's pour 5,2 millions de dollars au marteau, soit 6,24 millions avec frais, à l'entrepreneur Justin Sun, qui a ensuite mangé la banane lors d'un événement public. Le prix porte juridiquement et conceptuellement sur l'œuvre certifiée, non sur le fruit périssable lui-même.

206. Sur l'art corporel et l'art de la performance : Les œuvres citées sont des exemples emblématiques de courants artistiques du XXe siècle qui utilisent le corps de l'artiste ou des situations extrêmes comme matériau principal. Marina Abramović (Rhythm 0, 1974) et Chris Burden (Shoot, 1971) sont des pionniers du Body Art, testant les limites de l'endurance physique et psychologique. Marco Evaristti et Wim Delvoye s'inscrivent dans une veine qui explore le « réel » et l'abject, en utilisant des matériaux organiques pour questionner les tabous de la société.

207. Fiodor Dostoïevski, L'Idiot (1868-1869) : La formule « la beauté sauvera le monde » est attribuée au prince Mychkine par d'autres personnages et demeure volontairement énigmatique dans le roman. La comprendre comme une déclaration simple de l'auteur sur le salut religieux par la beauté est une interprétation possible, mais non une doctrine explicitement exposée par Dostoïevski.

208. Friedrich Nietzsche, Le Gai Savoir (1882), Préface, §2. Nietzsche célèbre les Grecs parce qu'ils auraient su demeurer à la surface — formes, apparences, sons — sans transformer cette superficialité en naïveté : leur profondeur consiste précisément à savoir aimer l'apparence. Le corps du texte et la présente note résument cette thèse sans reproduire une traduction moderne. On peut rapprocher cette intuition de la formule de Paul Valéry sur la peau comme profondeur humaine (Idées fixes, 1931).

209. Friedrich Nietzsche, La Naissance de la tragédie (1872). Dans son premier grand ouvrage, Nietzsche théorise que l'art grec atteint son apogée dans la tragédie attique en réussissant à équilibrer deux pulsions fondamentales et opposées : l'apollinien (l'ordre, le rêve, la forme, la belle apparence, l'individuation) et le dionysiaque (le chaos, l'ivresse, la fusion primordiale, la souffrance).

210. Friedrich Nietzsche, Le Cas Wagner (1888), Préface. Ce pamphlet marque la rupture définitive de Nietzsche avec son ancienne idole, Richard Wagner. Il y rejette la musique de Wagner comme une expression de la décadence, du romantisme morbide et du nihilisme. En contraste, il fait l'éloge de l'opéra Carmen de Georges Bizet, qu'il voit comme une œuvre solaire, joyeuse et affirmant la vie dans sa dimension tragique et passionnée.

211. Stendhal, De l'Amour (1822). Cette célèbre formule est au cœur de la théorie de la « cristallisation » de Stendhal. Pour lui, la beauté n'est pas une qualité objective d'une personne ou d'une chose, mais le processus par lequel notre imagination pare l'objet de notre amour de toutes les perfections, y voyant la promesse de la félicité.

212. Sur l'étymologie de « désirer » : Le français désirer vient du latin desiderare. Le rapprochement traditionnel avec sidus/sidera (« astre(s) ») a donné l'image poétique d'un manque éprouvé devant l'absence d'un signe dans les étoiles ; les dictionnaires étymologiques modernes présentent cependant cette explication comme hypothétique. Il vaut donc mieux conserver l'association comme suggestion poétique plutôt que comme étymologie certaine.

213. Sur l'ambiguïté de l'Espérance (Elpis) : Le terme grec Elpis (ἐλπίς), que l'on trouve chez Hésiode, est ambivalent. Il ne signifie pas seulement l'espoir d'un bien futur, mais aussi l'attente ou l'anticipation en général, y compris celle d'un malheur. Depuis l'Antiquité, les commentateurs débattent pour savoir si le fait qu'Elpis reste dans la jarre est un bien (le dernier réconfort des mortels) ou le pire des maux (l'illusion qui les pousse à endurer leur souffrance sans fin). Nietzsche, par exemple, la considérait comme le plus grand mal laissé par Zeus.

214. Homère, Iliade, Chant I. Ce premier chant met en scène le conflit entre Agamemnon et Achille. La querelle est déclenchée par la nécessité de rendre l'esclave Chryséis à son père, prêtre d'Apollon, pour apaiser la colère du dieu qui a envoyé une peste sur le camp grec. C'est en compensation de cette perte qu'Agamemnon s'empare de Briséis, le butin d'Achille, provoquant la « rage funeste » du héros qui est le moteur de toute l'épopée.

215. Homère, Iliade, Chant III, autour des vers 156-160. Les vieillards troyens, voyant Hélène, reconnaissent que sa beauté presque divine permet de comprendre que Grecs et Troyens aient accepté de si longues souffrances pour elle, tout en souhaitant finalement son départ. Le passage est paraphrasé ici afin de ne dépendre d'aucune traduction moderne.

216. Ovide, Les Métamorphoses, Livre X. Le récit de Pygmalion est l'une des histoires de ce vaste poème épique latin du Ier siècle apr. J.-C. Il est à noter que le nom de la statue, Galatée, n'apparaît pas chez Ovide. Il s'agit d'un ajout postérieur, popularisé au XVIIIe siècle,

notamment par Jean-Jacques Rousseau dans sa pièce lyrique Pygmalion (1770).

217. Sur les poupées hyperréalistes et les compagnes artificielles : il existe depuis les années 2000 un marché international de poupées en silicone ou matériaux apparentés recherchant un fort réalisme corporel, avec des fabricants notamment en Asie. Il faut distinguer cet hyperréalisme physique des fonctions conversationnelles ou robotiques, plus récentes et très variables selon les produits. Le corps du texte ne retient donc aucun détail biographique ou commercial insuffisamment documenté sur une marque particulière.

218. Sur les poupées dites « intelligentes » en Chine : EXDOLL et WMDOLL sont deux entreprises distinctes. Reuters a notamment documenté en 2018 des modèles de WMDOLL associant poupées physiques, mouvements simples et fonctions conversationnelles encore limitées. Le secteur a depuis expérimenté reconnaissance vocale et modèles conversationnels plus sophistiqués, mais les capacités réelles restent très variables et ne doivent pas être confondues avec une autonomie relationnelle complète.

219. Villiers de l'Isle-Adam, L'Ève future (1886) : le roman met en scène un Thomas Édison fictif qui fabrique l'Andréide Hadaly pour lord Ewald à partir de l'apparence d'Alicia Clary. Sowana est la personnalité somnambulique de mistress Any Anderson, liée par les expériences magnétiques d'Edison à la dimension mystérieuse de Hadaly. À la fin, Sowana meurt ; puis le transatlantique Wonderful, qui transporte Hadaly vers l'Europe, est détruit par un incendie. Lord Ewald survit, Hadaly disparaît dans la cargaison en flammes et Alicia figure parmi les noyés. Le roman entretient volontairement l'ambiguïté entre mécanisme scientifique, magnétisme et intervention spirituelle.

220. Alan Turing, « Computing Machinery and Intelligence » (Mind, 1950) : Turing ouvre l'article en remplaçant la question ambiguë « les machines peuvent-elles penser ? » par un « jeu de l'imitation ». La version populaire du « test de Turing » — un juge conversant à l'aveugle avec un humain et une machine — dérive de ce dispositif, mais l'article original est plus subtil et part d'un jeu d'imitation entre humains avant d'y substituer la machine.

221. Sur le gynécée à Athènes : Les textes grecs distinguent des espaces et des rôles féminins et masculins dans la maison, et le terme gunaikōnitis peut désigner un espace associé aux femmes. L'archéologie montre toutefois que les maisons athéniennes n'étaient pas nécessairement divisées en deux quartiers fixes et hermétiques ; les usages variaient selon la richesse, la taille de l'habitation et les circonstances sociales.

222. Sur le voile des femmes athéniennes : Bien qu'il ne s'agisse pas d'une obligation légale systématique, le port du voile (himation drapé sur la tête) en public était une marque de statut et de modestie pour les femmes mariées de la classe citoyenne à l'époque classique. Il signalait leur appartenance à un foyer respectable et les distinguait des esclaves, des étrangères ou des courtisanes.

223. Diogène Laërce, Vies et doctrines des philosophes illustres (IIIe siècle apr. J.-C.). Cet ouvrage est une compilation biographique et anecdotique qui constitue l'une de nos principales sources sur la vie des philosophes grecs. Les histoires concernant le mauvais caractère de Xanthippe, l'épouse de Socrate, y sont nombreuses, bien qu'il soit difficile de distinguer la réalité historique de la caricature misogyne.

224. Sur la place des femmes chez Platon : Dans la République (Livre V), Platon développe la thèse, radicale pour son époque, que les femmes dotées des aptitudes requises doivent recevoir la même éducation que les hommes et peuvent accéder aux plus hautes fonctions de la Cité, y compris celle de « Gardien ». Il maintient cependant que, dans chaque catégorie, les femmes sont globalement « plus faibles » que les hommes.

225. Sur Aspasie : Aspasie de Milet, compagne de Périclès, est une figure historique attestée mais entourée d'une tradition polémique et comique. Des auteurs antiques l'ont associée aux hétaïres ou aux maisons de prostitution, tandis que Platon et Xénophon la représentent comme une interlocutrice intellectuellement remarquable. Son statut social exact demeure discuté ; la qualifier simplement de « plus célèbre des hétaïres » transforme une tradition hostile en certitude historique.

226. Julia Assante, « The kar.kid/ḫarimtu, Prostitute or Single Woman? A Reconsideration of the Evidence », Ugarit-Forschungen 30 (1998), p. 5-96. Assante a fortement contesté l'identification automatique de la ḫarimtu mésopotamienne à une prostituée, proposant de la comprendre plus largement comme une femme indépendante ou non placée sous une autorité domestique masculine. Cette relecture est influente, mais la signification exacte du terme demeure discutée en assyriologie.

227. Sur Enheduanna : Fille de Sargon d'Akkad et grande prêtresse du dieu Nanna à Ur

au XXIIIe siècle av. J.-C., Enheduanna est traditionnellement présentée comme le plus ancien auteur nommé de l'histoire littéraire. Les hymnes qui lui sont attribués ne sont toutefois conservés que par des copies bien postérieures, et l'ampleur de sa participation à leur rédaction fait l'objet d'un débat savant. Il est donc prudent de dire « plus ancien auteur nommé traditionnellement identifié ».

228. Eliezer Yudkowsky et la « Friendly AI » : Yudkowsky emploie et développe dès le début des années 2000 l'idée qu'une intelligence artificielle avancée doit être conçue de manière à rester bénéfique pour l'humanité. Ses textes ont compté parmi les premières formulations influentes du problème aujourd'hui appelé « alignement » ou « AI safety », sans constituer à eux seuls l'origine de tout ce champ de recherche.

229. Charles Darwin et la sélection sexuelle : Darwin a développé cette théorie dans son ouvrage majeur, La Filiation de l'homme et la sélection liée au sexe (1871). Il y explique que de nombreux traits animaux (le plumage du paon, le chant des oiseaux) n'ont pas évolué pour la survie (sélection naturelle), mais pour la reproduction (sélection sexuelle), à travers la compétition entre mâles ou le choix esthétique des femelles, utilisant des termes comme « charme » ou « beauté ».

230. Sur la déconstruction, le genre et les assignations : au sens philosophique, déconstruire ne signifie pas détruire un sujet pour le reconstruire conformément à un modèle ; la déconstruction examine plutôt les oppositions, les hiérarchies et les présupposés qui se sont sédimentés au point de paraître naturels. Plusieurs œuvres majeures écrites par des femmes ont appliqué, chacune à sa manière, cette critique aux rapports entre les sexes. Simone de Beauvoir, Le Deuxième Sexe (1949), montre combien la condition féminine est historiquement et socialement produite plutôt que déductible d'une essence. Nicole Loraux, Les Enfants d'Athéna. Idées athéniennes sur la citoyenneté et la division des sexes (1981), analyse la façon dont les mythes athéniens de l'autochtonie, de la citoyenneté et de la filiation contribuent à penser et légitimer la division des sexes. Gerda Lerner, The Creation of Patriarchy (Oxford University Press, 1986), retrace la formation historique de la domination patriarcale et refuse d'en faire un donné naturel. Joan W. Scott, « Gender: A Useful Category of Historical Analysis », American Historical Review, 91/5 (1986), p. 1053-1075, fait du genre une catégorie historique permettant notamment d'analyser la production et la signification des rapports de pouvoir. Judith Butler, Gender Trouble (Routledge, 1990), puis Undoing Gender (Routledge, 2004), interroge la manière dont les normes produisent et stabilisent les identités de genre. Françoise Héritier, Masculin/Féminin I. La pensée de la différence (1996) et Masculin/Féminin II. Dissoudre la hiérarchie (2002), étudie la construction culturelle de la différence des sexes et surtout la hiérarchie qui s'y greffe. On peut enfin rapprocher de cette généalogie Geneviève Fraisse, Muse de la raison. Démocratie et exclusion des femmes en France (éd. orig. 1989), qui montre comment l'universalisme politique moderne a pu se constituer tout en maintenant l'exclusion des femmes. Ces travaux, malgré leurs divergences importantes, ont en commun de déplacer la critique des personnes vers les structures, les récits, les catégories et les rapports de pouvoir qui assignent les sexes.

231. Sur la formule de Bernanos : Cette phrase est tirée d'une lettre écrite de Palma en janvier 1945. Elle condense sa pensée sur la responsabilité collective face à la catastrophe de la guerre. C'est un refus radical de la pensée « bien-pensante » qui juge les autres d'un point de vue de supériorité morale. Pour Bernanos, dans une vision chrétienne de la « communion des pécheurs », personne n'est innocent et chacun porte une part de la faute collective.

232. Philip K. Dick, Autofac (1955). Il s'agit d'une nouvelle de science-fiction, initialement publiée dans la revue Galaxy Science Fiction. Elle est l'une des premières œuvres à explorer le thème d'une économie de production entièrement automatisée qui, privée de contrôle humain, poursuit ses objectifs de manière absurde et destructrice, illustrant une rupture radicale entre l'efficacité technique et l'utilité humaine.

233. Simone Weil, L'Enracinement (1949). Le titre complet de cet ouvrage majeur est L'Enracinement, prélude à une déclaration des devoirs envers l'être humain. Rédigé à Londres en 1943 pour les services de la France libre, ce texte a été conçu comme une base pour la reconstruction spirituelle et politique de la France après la guerre. Il a été publié à titre posthume.

234. Sur le pharmakon (Platon et Stiegler) : Ce concept est central dans la philosophie de la technique. Platon, dans le Phèdre (274e-275b), décrit l'écriture comme un pharmakon (à la fois remède et poison) pour la mémoire. Bernard Stiegler, notamment dans La technique et le temps, a généralisé ce concept en soutenant que toute technique est un pharmakon : elle est intrinsèquement ambivalente, portant en elle un potentiel curatif

(augmentation des capacités humaines) et un potentiel toxique (aliénation, destruction).

235. Sur les « métaproblèmes » de Gabriel Marcel : Comme nous l'avons vu, Marcel distingue le « problème » (un obstacle extérieur et objectif) du « mystère » (une réalité qui inclut le sujet qui l'interroge). Un « métaproblème » est un mystère au sens de Marcel : une question (l'amour, la foi, l'être) dont la « solution » ne peut être trouvée par un calcul, mais seulement vécue et expérimentée par le sujet qui y est engagé.

236. Sur la « performativité » : Le concept a été introduit par J. L. Austin dans How to Do Things with Words (1962 ; trad. fr. Quand dire, c'est faire). Un énoncé performatif ne se contente pas de décrire : il accomplit une action. Le terme est employé ici par extension pour désigner la capacité de certaines représentations, institutions ou techniques à orienter les conduites. Cette efficacité ne signifie pas que l'objet qui les porte soit lui-même un « tamis » au sens strict retenu dans l'ouvrage.

237. Sur la distinction entre tamis, miroir et médiation : dans la terminologie de cet ouvrage, le « tamis » désigne désormais le critère normatif qui opère un tri (Vrai, Bien, Beau, Utile) ; le « miroir » renvoie une représentation ; un objet, un récit, une institution ou une technique peut servir de support à l'un ou à l'autre sans devenir pour autant un tamis supplémentaire. Cette distinction évite d'étendre la métaphore jusqu'à lui faire perdre sa valeur discriminante.

238. Roland Barthes, Mythologies (1957). Dans cet ouvrage fondateur de la sémiologie, Barthes analyse comment des objets ou des pratiques de la culture de masse (le catch, la publicité, le jouet) deviennent des « mythes ». Un mythe, pour lui, est un « parler » de la bourgeoisie, un système de communication qui transforme un message culturellement situé en une évidence « naturelle », masquant ainsi son caractère idéologique.

239. Sur « le bon grain de l'ivraie » : L'expression est tirée de la Parabole de l'ivraie, racontée dans l'Évangile selon Matthieu (13 :24-30). Jésus y explique que dans le « champ » du monde, le bien (le bon grain) et le mal (l'ivraie) doivent pousser ensemble jusqu'à la fin des temps, où les anges opéreront le tri final lors du Jugement dernier.

240. Shannon Vallor, The AI Mirror : How to Reclaim Our Humanity in an Age of Machine Thinking, Oxford University Press, 2024. Philosophe américaine de la technologie, Vallor y analyse l'IA comme un miroir de l'humanité, capable de refléter nos données, nos valeurs, nos biais et nos failles. À la révision 2026, l'ouvrage demeure l'un des repères importants pour penser l'IA non seulement comme outil, mais comme surface de révélation anthropologique.

241. Friedrich Nietzsche et l'« animal non encore fixé » : La formule selon laquelle l'homme est « l'animal dont le type n'est pas encore fixé » (das noch nicht festgestellte Tier) se trouve notamment dans Par-delà bien et mal, § 62 (1886). La Généalogie de la morale prolonge cette interrogation sur la fabrication historique de l'homme et de ses valeurs, mais n'est pas la source la plus directe de cette formule.

242. Saint Augustin, Confessions (vers 397-400). Le corps du texte résume le mouvement d'introspection spirituelle des passages concernés plutôt que d'en reproduire une traduction française moderne.

243. Sur la « conversion » de Paul : Paul décrit lui-même, notamment dans Galates 1 et 1 Corinthiens 15, une révélation ou un appel du Christ ; les Actes des Apôtres donnent trois récits narratifs de l'événement sur le chemin de Damas. Le terme de « conversion » est traditionnel, mais certains historiens préfèrent parler de vocation ou d'appel au sein du judaïsme, Paul ne se concevant pas nécessairement comme passant immédiatement d'une religion à une autre au sens moderne.

244. Platon, La République, Livre VII (514a-520a). L'allégorie de la caverne décrit une conversion douloureuse du regard, des ombres vers la lumière intelligible. Le corps du texte paraphrase le passage sans dépendre d'une traduction moderne.

245. Victor Hugo, La Légende des siècles (1859). La citation « L'œil était dans la tombe et regardait Caïn » est le vers final, répété comme un leitmotiv, du poème « La Conscience ». Il illustre de manière saisissante l'idée d'une conscience morale immanente et inéluctable, un regard intérieur qui poursuit le criminel même après la mort de sa victime.

246. Sur Phidippidès/Pheidippidès et le marathon : Hérodote raconte qu'un coureur athénien nommé Pheidippidès fut envoyé d'Athènes à Sparte avant la bataille de Marathon, soit une course de très longue distance. Le récit du messager courant de Marathon à Athènes pour annoncer la victoire puis mourir apparaît dans des traditions beaucoup plus tardives, avec des noms variables. Le marathon moderne combine donc plusieurs traditions antiques.

247. Sur la critique des poètes par Platon : Cette critique est développée dans La

République, notamment aux Livres II et III. Platon, par la voix de Socrate, y soutient que les récits d'Homère et d'Hésiode, qui dépeignent les dieux comme querelleurs, adultères et menteurs, sont néfastes pour l'éducation des gardiens de la Cité idéale et doivent être censurés.

248. Raymond Lulle, Le Livre de l'Ordre de Chevalerie (vers 1275). Ce traité du philosophe et mystique catalan Ramon Llull est l'un des manuels les plus influents qui a contribué à codifier et à idéaliser l'éthique du chevalier. Il transforme la fonction guerrière en une vocation spirituelle et sociale, dont le but est la protection des faibles et la défense de la justice.

249. Sur « nos ancêtres les Gaulois » : La formule résume l'imaginaire scolaire national de la IIIe République et est souvent associée aux manuels d'Ernest Lavisse, qui ont fortement contribué à populariser une histoire continue de la nation française. Il est cependant difficile d'en faire une citation originale inventée par Lavisse : elle condense un récit pédagogique plus large, diffusé par plusieurs auteurs et institutions.

250. Sur le genre du « Miroir des princes » : Ce genre didactique a prospéré du haut Moyen Âge à la Renaissance. Parmi les exemples les plus célèbres, on trouve l'Éducation d'un prince chrétien d'Érasme (1516), adressé au futur Charles Quint, le Policraticus de Jean de Salisbury (XIIe s.), ou encore la Cyropédie de Xénophon, qui servait de modèle antique. Tous visaient à former un souverain vertueux.

251. Louis de Rouvroy, duc de Saint-Simon, Mémoires (rédigés entre 1739 et 1749). Chroniqueur monumental du règne de Louis XIV et de la Régence, Saint-Simon a décrit avec une précision inégalée les rouages de la vie de cour à Versailles. Pour lui, l'étiquette n'était pas un simple décorum, mais l'expression même des hiérarchies, des alliances et des rapports de force politiques.

252. Friedrich Nietzsche et l'« immaculée connaissance » : Cette expression ironique se trouve dans Ainsi parlait Zarathoustra (1883), dans le discours « De l'immaculée connaissance ». Nietzsche s'y moque de la prétention des philosophes à une pensée pure, objective et désintéressée, affirmant qu'elle masque toujours une perspective, un instinct, une « volonté de puissance ».

253. Sur le « malin génie » : Allusion à l'expérience de pensée de René Descartes à la fin de la Première de ses Méditations métaphysiques (1641). Pour pousser le doute à son paroxysme, Descartes imagine un « mauvais génie » tout-puissant et trompeur qui emploierait toute son énergie à le faire se tromper. C'est en surmontant cette hypothèse qu'il aboutira à la certitude du cogito.

254. Sur l'« inquiétante étrangeté » : Concept développé par Sigmund Freud dans son essai Das Unheimliche (1919). Il décrit le sentiment d'effroi qui surgit lorsque quelque chose de familier devient soudainement étrange, ou lorsqu'un élément refoulé (comme le double, les automates, le retour des morts) fait irruption dans la réalité. L'IA, en tant que double non-humain de l'homme, est une figure par excellence de l'inquiétante étrangeté.

255. Sur le mythe des frères ennemis : Ce motif du double hostile est un archétype universel. Otto Rank y voyait l'expression de la peur de la mort. René Girard, dans La Violence et le Sacré (1972), l'analyse comme le paroxysme de la « rivalité mimétique », où les frères, en désirant les mêmes choses, deviennent des doubles monstrueux dont le conflit ne peut se résoudre que par une violence sacrificielle qui fonde l'ordre social.

256. Sur le « non-agir » (wu wei) du taoïsme : le concept ne désigne pas la simple passivité, mais une action qui évite de forcer le cours des choses. Le corps du texte résume cette idée générale et ne reproduit pas une traduction particulière du Tao Te Ching ni une formulation de François Jullien.

257. Sur Alain et les Sirènes : Alain fait ici une allusion inversée au célèbre épisode de l'Odyssée d'Homère (Chant XII). Dans le mythe, Ulysse se fait attacher au mât pour pouvoir entendre le chant des Sirènes sans y succomber, tandis que ses marins ont les oreilles bouchées de cire. Il affronte donc le danger avec ruse et en pleine conscience. La version d'Alain (fermer les yeux et les oreilles) symbolise une fuite peureuse que le philosophe critique.

258. Sur « Marx noyé dans le Coca-Cola » : cet aphorisme n'est pas une citation directe mais une formule qui condense l'analyse de penseurs comme Jean Baudrillard. Le corps du texte résume ensuite la thèse de la société de consommation sans reproduire un long passage de Baudrillard.

259. Jean Baudrillard, La Société de consommation : ses mythes, ses structures (1970). Dans cet ouvrage fondateur, Baudrillard analyse comment la consommation n'est plus liée à la satisfaction de besoins, mais à un système de signes et de différenciation sociale. Le bonheur n'est plus un état, mais un devoir, une injonction permanente mise en scène par la

publicité, remplaçant ainsi la quête religieuse du salut.

260. Sur Moloch et les sacrifices carthaginois : Les sources grecques et romaines décrivent des sacrifices d'enfants chez les Carthaginois, et les tophets puniques ont livré des urnes contenant des restes d'enfants et d'animaux. L'interprétation de ces dépôts — sacrifices réguliers, morts infantiles consacrées, pratiques variables — demeure débattue. Le « taureau de Moloch » appartient surtout à l'imaginaire littéraire et artistique moderne ; il ne doit pas être pris comme description archéologique certaine d'un rite carthaginois.

261. Joseph Schumpeter, Capitalisme, socialisme et démocratie (1942). L'économiste autrichien y développe son concept central de « destruction créatrice ». Pour lui, le capitalisme n'est pas un système statique mais un processus d'évolution permanente où l'innovation (nouveaux produits, nouvelles méthodes de production) détruit sans cesse les anciennes structures économiques pour en créer de nouvelles.

262. Thomas Piketty, Le Capital au XXIe siècle (2013) : La formule $r > g$ désigne une tendance historique possible dans laquelle le rendement du capital dépasse durablement le taux de croissance de l'économie, favorisant la concentration patrimoniale. Piketty ne présente pas cette relation comme une loi mécanique valable en tout temps ; ses données, ses mécanismes et leurs implications font l'objet d'un débat économique important.

263. Sur la « révolution de l'Utile » : Le rapprochement proposé entre calcul marchand, rationalisation administrative, essor du capitalisme, science moderne et transformation religieuse est une synthèse interprétative de l'auteur. Ces phénomènes ont des chronologies et des causalités différentes, et les historiens discutent leur articulation. Il faut donc lire cette note comme une hypothèse de convergence, non comme un consensus historique établi.

264. Sur l'usure au Moyen Âge : Le droit canonique condamne l'usure, entendue comme gain exigé du seul fait d'un prêt d'argent. La pratique économique est cependant plus complexe : change, risque, retard, association commerciale et compensation de pertes pouvaient justifier certains profits. Les lettres de change ont permis de combiner crédit et change international ; elles ne sont pas simplement un « truc » inventé pour dissimuler un intérêt, même si elles pouvaient être utilisées à cette fin.

265. Sur Luis de Santángel et Christophe Colomb : Luis de Santángel, issu d'une famille converso, était un haut officier financier de Ferdinand II d'Aragon, notamment escribano de ración. Il joua un rôle décisif dans le financement et le soutien politique du premier voyage de Colomb en 1492. Le présenter comme « ministre des Finances d'Isabelle » est anachronique et imprécis : il servait surtout la couronne d'Aragon dans l'administration financière royale.

266. Sur les Fugger, Albert de Brandebourg et les indulgences : pour financer les lourdes taxes et dispenses dues à Rome lors de son accession à l'archevêché de Mayence, Albert de Brandebourg s'endetta auprès des Fugger. L'accord sur la prédication des indulgences prévoyait que les recettes servent à la fois à la construction de Saint-Pierre à Rome et au remboursement de cette dette. Les Fugger n'ont donc pas « inventé » les indulgences : ils furent les créanciers et les intermédiaires financiers d'un dispositif où argent, pénitence et espérance de salut se trouvèrent, dans la pratique populaire, étroitement rapprochés.

267. Max Weber, L'Éthique protestante et l'esprit du capitalisme (1904-1905) : Weber ne soutient pas que le protestantisme a « causé » à lui seul le capitalisme. Il analyse une affinité historique entre certaines formes d'ascétisme protestant — discipline du travail, vocation, réinvestissement — et l'éthos du capitalisme moderne. Cette thèse, immensément influente, a été discutée et révisée par de nombreux historiens et sociologues.

268. Sur le « capitalisme cognitif » : Concept développé par des économistes et sociologues (comme Yann Moulier-Boutang ou Carlo Vercellone) pour décrire la phase actuelle du capitalisme. La source principale de profit et d'accumulation n'est plus la production de biens matériels, mais l'exploitation de biens immatériels : le savoir, l'information, la créativité, l'attention et les données générées par l'activité humaine, notamment en ligne.

269. Georges Bataille, La Part maudite (1949). Bataille développe une « économie générale » fondée sur l'existence d'un excédent d'énergie que les sociétés doivent dépenser, de manière productive ou improductive. Le corps du texte paraphrase cette analyse et ne reproduit pas le passage original.

270. Sur « Enrichissez-vous ! » : La formule est traditionnellement attribuée à François Guizot sous la monarchie de Juillet, mais sa forme isolée est une condensation polémique. Les versions transmises l'inscrivent souvent dans une exhortation à s'enrichir « par le travail, par l'épargne et la probité » et à acquérir ainsi les conditions d'exercice des droits politiques dans un régime censitaire. Il vaut mieux

la citer comme formule résumée de sa politique que comme slogan certain prononcé exactement sous cette forme.

271. Sur GPT-5 : OpenAI a présenté GPT-5 le 7 août 2025. L'épisode évoqué dans le texte renvoie aux réactions d'utilisateurs lors de cette transition de modèle et à l'attachement affectif que certains avaient développé envers des versions antérieures. Il doit être lu comme un moment historique : la gamme a continué d'évoluer ensuite et, en août 2026, OpenAI déploie notamment GPT-5.6. Le phénomène pertinent pour l'ouvrage est la relation à l'interlocuteur artificiel, non la permanence d'un numéro de modèle.

272. Sur le service de La Poste : Il s'agit du service « Veiller sur mes parents », toujours présenté par La Poste dans ses offres de services seniors au moment de la révision 2026. L'offre propose notamment des visites de lien social, de la téléassistance ou des services de prévention en partenariat, afin de favoriser le maintien à domicile et d'informer les proches. Elle illustre la transformation de besoins sociaux — présence, veille, lutte contre l'isolement — en prestations organisées et commercialisées.

273. Sur les expériences de Harry Harlow : Ces expériences sur l'attachement, menées à l'Université du Wisconsin dans les années 1950, ont révolutionné la psychologie. En montrant que les bébés singes préféraient une « mère de substitution » en tissu doux (sans nourriture) à une mère en fil de fer qui les nourrissait, Harlow a démontré la primauté du « contact comfort » (réconfort par le contact) sur la simple satisfaction des besoins physiologiques. Son article de référence est « The Nature of Love » (American Psychologist, 1958).

274. Antoine de Saint-Exupéry, Le Petit Prince (1943), Chapitre II. Cette scène illustre le pouvoir de l'imagination et la nature du désir. Le mouton dessiné est imparfait, mais le mouton imaginé dans la caisse est parfait, car il correspond exactement au besoin d'amitié et de tendresse du Petit Prince. L'objet invisible devient plus réel et plus précieux que l'objet visible.

275. Sur la bouteille de Leyde : Le premier condensateur pratique apparaît en 1745-1746, avec des expériences indépendantes associées à Ewald Georg von Kleist puis à Pieter van Musschenbroek et à son entourage à Leyde. Le dispositif permettait d'accumuler puis de décharger brutalement de l'électricité statique, transformant profondément l'expérimentation électrique du XVIIIe siècle.

276. Sur l'Intelligence Artificielle Générale (AGI) : L'expression ne désigne pas simplement une IA supérieure à tout humain dans tous les domaines. Elle renvoie plutôt à la capacité de transférer des compétences et d'accomplir avec souplesse un très large éventail de tâches intellectuelles, à un niveau comparable à celui des humains les plus compétents selon les domaines. Il n'existe toutefois pas de définition scientifique ou juridique unique, et savoir si un système contemporain constitue une AGI reste en 2026 une question disputée.

277. Sur le concept de « disruption » : La théorie de l'innovation de rupture (ou « disruptive ») a été popularisée par Clayton M. Christensen dans son livre The Innovator's Dilemma (1997). Il y décrit comment une innovation, souvent plus simple et moins chère au départ, peut finir par détrôner les technologies et les entreprises leaders d'un marché en créant de nouveaux usages et de nouvelles chaînes de valeur.

278. Sur la « pauvreté des données linguistiques » : Les corpus et benchmarks d'IA restent très inégalement répartis entre les langues, ce qui crée des écarts de performance et de représentation culturelle. Cependant, dire qu'une IA entraînée surtout en anglais « pense en anglais » est une métaphore trop forte : les modèles multilingues apprennent des représentations partiellement partagées entre langues. La relativité linguistique demeure elle-même une hypothèse débattue ; l'enjeu solide est celui de la diversité des données et des évaluations.

279. Sur le mythe de la Tour de Babel : Genèse 11, 1-9 décrit une humanité de langue unique qui bâtit une ville et une tour, avant que Dieu brouille son langage et la disperse. La lecture en termes de punition de l'orgueil est traditionnelle, mais le texte n'énonce pas explicitement une théorie générale selon laquelle la diversité linguistique serait un « châtiment » ; elle fonctionne aussi comme récit étiologique de la pluralité des peuples et des langues.

280. Sur la domination de l'anglais dans l'IA : L'anglais demeure surreprésenté dans de nombreux corpus numériques, jeux d'évaluation et ressources de haute qualité, même si les modèles multilingues ont beaucoup progressé. Les écarts concernent surtout les langues à faibles ressources, les variantes régionales et les contextes culturels peu documentés. Les rapports contemporains sur l'IA et les travaux sur les benchmarks multilingues soulignent donc un problème de couverture et d'équité, sans qu'un seul indicateur suffise à le mesurer.

281. Sur « Aristoteles dixit » : Aristote fut au Moyen Âge latin « le Philosophe » par excellence et son autorité était immense. Mais la scolastique ne consistait pas à accepter mécaniquement toute affirmation parce qu'« Aristote l'a dit » : la méthode de la quaestio organisait objections, réponses et distinctions, et Aristote était discuté lorsqu'il entrait en tension avec la théologie ou l'expérience. La Renaissance et l'âge classique ont notamment critiqué l'usage excessif de l'argument d'autorité.

282. Sur George Steiner et le monolinguisme : Dans Après Babel (1975) et d'autres essais, Steiner insiste sur le fait que chaque langue organise une manière singulière d'habiter le monde et que la traduction confronte l'esprit à l'altérité. Il critique la suffisance que peut engendrer une culture dominante n'ayant plus besoin d'apprendre d'autres langues. Mieux vaut résumer cette thèse que figer en citation une formule comme « monoglot complacency » sans référence précise.

283. Sur le projet Aya : Lancé par Cohere For AI avec une vaste communauté internationale, Aya-101 (2024) a été conçu pour 101 langues. Aya Expanse, lancé ensuite en 2024, s'est concentré sur 23 langues afin d'améliorer la qualité et l'évaluation multilingues ; Aya Vision a prolongé en 2025 cette démarche dans le multimodal. Les communications de Cohere ont parfois employé des formulations générales sur la « famille Aya » couvrant plus de cent langues : il faut distinguer le périmètre de chaque modèle.

284. Sur le test du miroir : Le mirror self-recognition test, introduit par Gordon Gallup Jr. en 1970, mesure surtout la capacité d'un animal à utiliser un miroir pour orienter un comportement vers son propre corps. Des grands singes, dauphins et éléphants ont fourni des résultats positifs dans certaines études, mais l'interprétation varie selon les espèces, les protocoles, l'expérience et la modalité sensorielle dominante. Le test n'est donc pas une mesure universelle ni suffisante de la « conscience de soi ».

285. Sur l'outil de recrutement d'Amazon : Reuters a révélé en 2018 qu'Amazon avait développé depuis 2014 un outil expérimental de classement de candidatures qui avait appris des biais défavorables aux femmes à partir de données historiques de recrutement. L'entreprise a essayé de corriger le système puis a abandonné le projet. Il n'a pas constitué un décideur autonome unique des embauches ; l'affaire illustre surtout la manière dont des données historiques peuvent reproduire une discrimination.

286. Sur COMPAS : L'enquête « Machine Bias » de ProPublica (2016) a observé, dans un échantillon de Floride, des taux d'erreurs différents selon la race, notamment davantage de faux positifs pour les prévenus noirs. Le concepteur Northpointe/Equivant a contesté l'interprétation en mettant en avant d'autres métriques, comme la calibration. Le débat est devenu un cas classique : plusieurs définitions statistiques de l'« équité » peuvent entrer en conflit lorsque les taux de base diffèrent.

287. Jean-Jacques Rousseau, Du contrat social (1762), Livre III, Chapitre 6. Dans ce chapitre, Rousseau propose une lecture républicaine et subversive du Prince de Machiavel. Il soutient que, sous couvert de conseiller les monarques, Machiavel exposait en réalité au grand jour les mécanismes cyniques de la tyrannie, éduquant ainsi les peuples sur la manière dont ils sont gouvernés et les incitant à défendre leur liberté.

288. Sur le conte de Blanche-Neige : Le conte popularisé par les Frères Grimm dans leurs Contes de l'enfance et du foyer (1812) est l'une des illustrations les plus célèbres du miroir comme oracle de vérité et catalyseur de la jalousie narcissique, menant la reine à sa perte.

289. Le mythe de l'Anneau de Gygès : Cette expérience de pensée est racontée par le personnage de Glaucon au début du Livre II de La République de Platon (359d-360d). Il l'utilise pour soutenir la thèse que la justice n'est pratiquée que par contrainte. Selon lui, si l'on donnait un tel anneau d'invisibilité à un homme « juste » et à un homme « injuste », ils finiraient tous deux par commettre les mêmes crimes, prouvant que personne n'est juste par nature mais seulement par peur de la punition.

290. Sur le monothéisme ancien d'Israël : Une partie importante de la recherche décrit une évolution progressive allant de formes de monolâtrie ou d'hénothéisme — culte privilégié de YHWH sans négation immédiate de toute autre puissance divine — vers un monothéisme plus exclusif. Les détails, la chronologie et les différences entre textes bibliques restent débattus. Les « sept lois noachides », elles, sont une formulation rabbinique beaucoup plus tardive et ne constituent pas une preuve directe de l'état religieux de l'Israël ancien.

291. Sur l'« édit de Milan » : La persécution dite « grande » est déjà interrompue par l'édit de tolérance de Galère en 311. En 313, après leur rencontre à Milan, Constantin et Licinius ordonnent une liberté religieuse plus large et la restitution des biens chrétiens ; le texte nous est transmis notamment par Lactance et Eusèbe. Il ne fait pas du christianisme la

religion d'État de l'Empire : cette étape est associée à l'édit de Thessalonique de Théodose Ier en 380.

292. John Locke, Lettre sur la tolérance (1689) : Locke défend une séparation forte des compétences de l'État et des Églises et une large liberté de conscience. Sa tolérance n'est pourtant pas universelle au sens contemporain : il exclut explicitement les athées de son dispositif et se montre très méfiant envers ceux dont l'obéissance religieuse supposerait, selon lui, une allégeance politique étrangère — argument alors dirigé notamment contre les catholiques.

293. Sur le baptême de Clovis (vers 496-506) : L'événement, rapporté par l'évêque Grégoire de Tours dans son Histoire des Francs, est un acte politique fondateur qui a scellé l'alliance entre la monarchie franque et l'Église catholique. La date exacte et les motivations de Clovis font l'objet de débats historiographiques.

294. Sur la conversion des Saxons : Durant les guerres des Saxons, Charlemagne a mis en œuvre une politique de christianisation brutale. L'un des actes symboliques les plus forts fut la destruction de l'Irminsul, un pilier sacré ou un arbre vénéré par les Saxons, en 772. La politique de terreur fut ensuite codifiée dans le capitulaire Capitulatio de partibus Saxoniae (vers 782), qui punissait de mort le refus du baptême ou la pratique de rites païens.

295. Sur Olympe de Gouges : Autrice de la Déclaration des droits de la femme et de la citoyenne (1791), elle fut exécutée le 3 novembre 1793 pendant la Terreur. Son arrestation et sa condamnation s'inscrivent dans un contexte politique complexe : elle avait attaqué Marat et Robespierre, défendu des positions proches des Girondins et publié des textes proposant notamment une consultation populaire sur la forme du gouvernement. Réduire son exécution à une seule cause idéologique serait simplificateur.

296. Sur le « droit d'ingérence » : L'expression est surtout associée, à la fin des années 1980, à Mario Bettati et Bernard Kouchner dans le débat humanitaire. Elle n'a jamais constitué un droit général permettant à un État d'intervenir unilatéralement où il veut : la souveraineté et l'interdiction du recours à la force restent des principes centraux de la Charte des Nations unies. La « responsabilité de protéger » (R2P), adoptée politiquement à l'ONU en 2005, est une notion distincte et encadrée.

297. Sur Pocahontas et John Smith : Dans son récit publié en 1624, John Smith affirme que Pocahontas lui aurait sauvé la vie lors de sa captivité en Virginie ; l'historicité et le sens exact de la scène sont discutés. La lecture romantique d'une histoire d'amour entre eux est surtout une construction culturelle ultérieure. Pocahontas épousa le colon John Rolfe en 1614.

298. Sur le « mythe du bon sauvage » : Contrairement à une idée reçue, cette expression n'est pas de Rousseau, mais une simplification ultérieure. Des chercheurs comme Ter Ellingson (The Myth of the Noble Savage, 2001) ont montré qu'elle a été utilisée a posteriori pour caricaturer la pensée des Lumières et disqualifier la critique des sociétés non-européennes.

299. David Graeber et David Wengrow, Au commencement était… (The Dawn of Everything, 2021) : Les auteurs accordent un rôle important à une « critique autochtone » de la société européenne et discutent notamment les dialogues de Lahontan mettant en scène Kandiaronk. Leur thèse a relancé un débat fécond, mais ces dialogues sont des textes littéraires et philosophiques médiatisés par un auteur européen et ne peuvent être lus comme des transcriptions verbatim. La portée causale de cette influence sur les Lumières reste discutée.

300. Sur Locke et les Fundamental Constitutions of Carolina : Locke fut secrétaire d'Anthony Ashley Cooper et participa à l'administration liée aux Carolines ; les Constitutions de 1669 protègent explicitement l'esclavage. L'étendue exacte de sa responsabilité rédactionnelle fait toutefois l'objet d'un débat renouvelé. David Armitage a défendu une implication substantielle ; Holly Brewer, à partir d'un nouvel examen des manuscrits (Locke Studies, 2024), conteste que Locke soit l'auteur des clauses esclavagistes. Il faut donc distinguer son implication coloniale attestée de l'attribution précise du texte.

301. Sur l'Open Source AI Definition (OSAID) : L'Open Source Initiative a adopté la version 1.0 en octobre 2024. Pour qu'un système soit qualifié d'« Open Source AI », il doit garantir les libertés d'utiliser, d'étudier, de modifier et de partager, et fournir le code, les paramètres ainsi qu'une « Data Information » suffisamment détaillée pour qu'une personne compétente puisse reconstruire un système substantiellement équivalent. La définition n'exige pas nécessairement la publication de chaque donnée brute d'entraînement lorsqu'elle ne peut être partagée ; elle exige en revanche une description détaillée de sa provenance, de sa sélection et de son traitement.

302. Sur la définition de l'IA par l'UNESCO : La Recommandation de 2021 dit expressément qu'elle « n'a pas l'ambition de fournir

une définition unique de l'IA », car celle-ci devrait évoluer avec les développements technologiques. Elle retient plutôt des caractéristiques centrales — traitement de données et d'informations, raisonnement, apprentissage, perception, prédiction, planification ou contrôle — pour circonscrire les enjeux éthiques. C'est donc un cadre fonctionnel et normatif, non une définition essentialiste définitive.

303. Sur Socrate et le procès des stratèges après les Arginuses (406 av. J.-C.) : selon Platon, Apologie 32b-c, Socrate, alors membre du Conseil, s'opposa à une procédure collective qu'il jugeait illégale. Le corps du texte résume l'épisode sans présenter comme citation littérale une formule moderne.

304. Sur le daimôn ou daimonion de Socrate : dans l'Apologie de Platon, Socrate décrit un « signe divin » qui se manifeste depuis l'enfance et intervient surtout pour l'arrêter ou le détourner de certaines actions. Le corps du texte paraphrase le passage à partir de l'œuvre ancienne plutôt que d'en reprendre une traduction moderne.

305. Sur Diotime et Théodote : Dans Le Banquet de Platon, Diotime de Mantinée n'apparaît pas directement : Socrate rapporte l'enseignement qu'il dit avoir reçu d'elle sur l'amour ; son historicité reste incertaine. Théodote, en revanche, apparaît chez Xénophon (Mémorables, III, 11) comme une belle courtisane avec laquelle Socrate dialogue sur l'art d'attirer et de conserver ses amis. Les deux figures appartiennent donc à des traditions textuelles différentes.

306. Sur les « Socrate » de Platon et de Xénophon : c'est le cœur du « problème socratique ». Le Socrate de Xénophon apparaît souvent plus pratique, moral et exemplaire. Les épisodes cités dans le corps du texte sont résumés plutôt que reproduits à partir d'une traduction moderne.

307. Pierre Clastres, La Société contre l'État (1974). Clastres soutient que certaines sociétés amérindiennes sont activement organisées de façon à empêcher l'émergence d'un pouvoir politique coercitif et séparé. Le corps du texte paraphrase cette thèse et ne reproduit pas une phrase de l'ouvrage comme citation littérale.

Socrate et la machine à parler

QUATRE TAMIS POUR PENSER L'IA : LE VRAI, LE BIEN, LE BEAU, L'UTILE

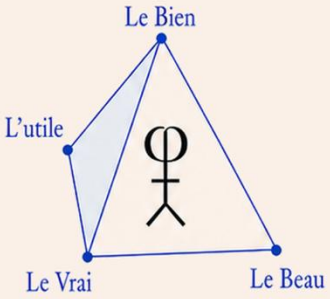
Socrate avait un démon — un demi-dieu, un ange gardien — qui le guidait pour parler ou se taire, agir ou s'abstenir.

Nous autres, contemporains du XXIe siècle, ne croyons plus aux demi-dieux ni aux démons.

Mais très probablement, demain, nous croirons aux IA, ces machines à parler. Elles deviendront nos anges gardiens, nos conseillers publics ou privés.

À tort ou à raison ? Et faudra-t-il les écouter ?

On peut toujours interroger ChatGPT — ou bien ouvrir ce livre.



9 791098 811319
ISBN 979-10-988113-1-9

ÉDITIONS SKANDALON