

E S S A I

SOCRATE AVEC UN ORDINATEUR

*Que reste-t-il à l'homme quand les réponses deviennent
abondantes ?*



M A R C L E B E L

@MARCLEBEL

Écrit en dialogue avec ChatGPT et Claude

A O Û T 2 0 2 6

S O M M A I R E

Prologue — Nous nous sommes peut-être crus divins

4

I · INTERROGER

01 Ne me donne pas ta solution 10

02 La rareté déplacée 13

II · PENSER

03 Thamous avait raison 18

04 L'aporie 19

05 La Sirène de la pensée 20

06 Changer de chaise 21

07 Pourquoi continuer à apprendre ? 23

08 La nouvelle rareté : le disciple 24

III · FAIRE

09 Le monde résiste 27

10 Faire ensemble 31

IV · RÉPONDRE

11	Et si elle devenait meilleure que nous pour tout ?	35
12	Le droit à l'expérience	37
13	À qui confions-nous les cordes ?	40
14	Comprendre pour rester libre	44
15	Lorsque savoir ne suffit plus	46
16	L'Oracle commun	48
17	À qui appartient mon Oracle ?	50
18	L'Oracle fiduciaire	52
19	Qui gardera les gardiens ?	56
20	Le consentement au curriculum	58
21	Le troisième parent	59
	FIN · Conclusion — Qui doit maintenant répondre ?	61
	Note sur l'écriture	64
	Sources et références	65

P R O L O G U E

Nous nous sommes peut-être crus divins

Pendant presque toute leur histoire, les hommes ont vécu sous des intelligences supérieures.

Les Grecs vivaient entourés de dieux. Athéna était plus sage qu'Ulysse, Zeus plus puissant qu'Achille, Apollon voyait ce que les hommes ne voyaient pas.

Peu importe ici que ces dieux aient existé. Ils existaient dans le monde où ces hommes pensaient, décidaient et agissaient.

Et cela ne rendait pas la vie humaine inutile.

Ulysse ne renonçait pas à réfléchir parce qu'Athéna réfléchissait mieux que lui. Achille ne renonçait pas à combattre parce que Zeus était invincible — et il savait de sa propre mère, déesse, des choses qu'aucun homme ne voudrait savoir. Leurs héros étaient peut-être intéressants précisément parce qu'ils devaient agir dans un monde qui les dépassait.

Puis quelque chose a changé.

Les dieux se sont progressivement retirés de notre représentation du monde. Nous avons expliqué la foudre, les éclipses et les maladies. Nous avons compris le mouvement des planètes, transformé la matière, maîtrisé des quantités croissantes d'énergie. Et pendant un bref instant à l'échelle de l'histoire, nous avons pu nous regarder comme l'intelligence suprême du monde que nous connaissions.

Nous avons peut-être fini par confondre cette position avec notre valeur.

C'est cette confusion que l'intelligence artificielle vient troubler.

Pendant des millénaires, nous avons imaginé des oracles capables de répondre aux questions qui nous dépassaient. Nous sommes peut-être les premières générations à commencer réellement à en construire un.

Nos machines sont encore imparfaites : elles se trompent, inventent parfois, ne connaissent évidemment pas l'avenir. Mais la direction est frappante — chacun peut désormais demander presque n'importe quoi et obtenir immédiatement une réponse plausible, argumentée et, de plus en plus souvent, accompagnée d'une capacité d'action.

Pour la première fois, ce n'est pas un dieu que nous imaginons au-dessus de nous.

C'est quelque chose que nous construisons.

Il y a là une ironie étrange. Nous nous serons crus divins l'espace d'un instant — et l'une des premières choses que notre puissance nous aura apprises sera que nous ne l'étions pas.

De là nos questions inquiètes.

« Si la machine écrit mieux que moi, que me reste-t-il ? »

« Si elle raisonne mieux que moi, quelle est encore ma valeur ? »

« Si elle finit par décider mieux que moi, à quoi suis-je encore bon ? »

Ce vertige se présente comme inédit. Il ne l'est pas. Il marque la fin d'une parenthèse : quelques siècles pendant lesquels nous avons pu nous concevoir comme la plus haute intelligence dont nous ayons une expérience observable. L'intelligence artificielle ne tranche rien sur Dieu ni sur ce qui nous dépasse. Elle menace quelque chose de plus concret : le monopole humain de l'intelligence dans le monde accessible.

Presque toute l'expérience humaine s'est déroulée en dehors de cette parenthèse. Et nous n'avons pas entièrement oublié comment on vit sous plus grand que soi.

Depuis 1997, les machines jouent aux échecs mieux que les meilleurs joueurs humains ¹. Les échecs n'ont pas disparu. Des millions de personnes regardent un champion chercher, front plissé, pendant qu'une machine posée à côté de lui jouerait plus fort. C'est que nous ne regardons pas les échecs pour connaître le meilleur coup : nous regardons des humains essayer de le trouver.

Nous savons encore admirer les Achilles d'aujourd'hui.

Les Grecs nous ont laissé autre chose que cette leçon d'admiration. Ils nous ont laissé des histoires d'hommes face aux oracles — et ces histoires sont étonnamment précises sur ce qui nous attend.

Hérodote raconte celle de Crésus, roi de Lydie. Avant d'attaquer les Perses, Crésus consulte l'oracle de Delphes, le meilleur de son monde. L'oracle répond que, s'il engage la guerre, il détruira un grand empire. Crésus entend ce qu'il a envie d'entendre : il va détruire l'Empire perse. Il attaque. Un grand empire est effectivement détruit — le sien.

La réponse n'était pas fausse. Crésus en avait tiré une mauvaise conclusion.

Mais Hérodote ne s'arrête pas là. Sa catastrophe fera de Crésus, bien plus tard, un autre homme — nous y reviendrons.

Même nos pires erreurs fabriquent quelqu'un.

Quelques décennies plus tard, Socrate reçoit lui aussi une réponse de Delphes. La Pythie affirme que nul n'est plus sage que lui. Il pourrait accepter le compliment. Il fait l'inverse : la réponse le trouble, il enquête, il interroge hommes politiques, poètes et artisans pour comprendre ce que l'oracle a bien pu vouloir dire. Cette enquête le conduit vers l'idée la plus

célèbre qui lui soit associée — une partie de sa sagesse consiste précisément à ne pas croire savoir ce qu'il ignore.

Deux hommes reçoivent donc une réponse.

Crésus l'interprète et agit. Socrate l'interroge et enquête.

Face à nos nouveaux oracles, nous avons peut-être exactement ce choix.

Devenir Crésus avec ChatGPT.

Ou Socrate avec un ordinateur.

Un troisième Grec nous attend plus loin. Ulysse ne demandait pas de réponses à son oracle : il demandait des cordes. Il savait quelque chose que Crésus et Socrate ne pouvaient pas encore nous apprendre — que ce que nous voulons maintenant n'est pas toujours ce que nous voulons vouloir.

L'Oracle de ces pages sera tour à tour trois choses : nos systèmes actuels, leur prolongement plausible, et parfois une limite imaginaire destinée à éprouver le raisonnement. Une intelligence artificielle n'est pas un oracle : elle peut être comparée, contestée, testée, multipliée. Le danger commence lorsque nous adoptons envers elle une attitude oraculaire — lorsque nous cessons d'examiner une réponse parce qu'elle nous arrive avec suffisamment d'assurance, de précision et d'autorité.

Ce livre pose donc une question qui n'est pas celle que l'on croit.

Pas « que restera-t-il à l'homme que la machine ne saura jamais faire ? » — les Grecs ne se demandaient pas ce qu'il leur restait face à Athéna.

Mais : que doit savoir faire l'homme lorsque la réponse devient abondante ? Et, plus profondément : qui devient-il lorsqu'une intelligence qui le dépasse participe, jour après jour, à former celui qui décidera demain ?

Le monde n'a jamais exigé de nous que nous soyons des dieux pour y avoir quelque chose à faire.

Il va peut-être simplement falloir réapprendre ce que les Grecs savaient : être pleinement humains dans un monde où nous ne sommes plus l'intelligence suprême.



I

I N T E R R O G E R

Apprendre à interroger l'Oracle.

*L'intelligence artificielle rend les réponses abondantes.
Elle ne garantit pas que nous posions les bons problèmes.*

Ne me donne pas ta solution

Alice utilise le CRM de son entreprise pour sa prospection. Elle veut décider quels contacts relancer en priorité.

Elle rencontre Socrate.

ALICE J'ai trouvé ce qu'il me faut. Je voudrais créer un score pour classer mes prospects. On prendrait la date du dernier appel, la ville et le chiffre d'affaires de l'entreprise.

SOCRATE Pourquoi ces trois critères ?

ALICE Parce qu'ils permettent de reconnaître les bons prospects.

SOCRATE Montre-moi ton meilleur prospect du mois dernier.

Alice ouvre son CRM.

ALICE Celui-ci.

SOCRATE Pourquoi l'avais-tu rappelé en priorité ?

ALICE Il avait répondu très vite à mon premier email et posé des questions très précises.

SOCRATE Cela fait-il partie de ton score ?

ALICE Non.

SOCRATE Autre chose ?

ALICE J'avais compris assez vite qu'il était le décideur.

SOCRATE Cela fait-il partie du score ?

ALICE Non plus.

SOCRATE Alors pourquoi as-tu choisi les trois autres critères ?

Alice se défend.

- ALICE** Parce qu'ils sont dans la base de données. Ce dont je te parle maintenant est beaucoup plus difficile à mesurer.
- SOCRATE** Ce qui est facile à mesurer est-il nécessairement ce qui permet le mieux de décider ?
- ALICE** Non. Mais mon score imparfait peut tout de même être meilleur que pas de score. Aujourd'hui, chaque commercial choisit selon son intuition. Au moins, une règle commune nous permettrait de comparer nos décisions.

Cette fois, Socrate réfléchit.

- SOCRATE** Alors ton objectif n'est pas seulement de prédire quel prospect signera.
- ALICE** Non. Je veux aussi rendre nos décisions plus cohérentes et plus faciles à mesurer.
- SOCRATE** Voilà donc déjà un objectif supplémentaire qui n'apparaissait pas dans ta première demande.

Alice croyait chercher une formule.

La conversation fait apparaître deux problèmes : comprendre comment les commerciaux reconnaissent réellement un bon prospect et décider ce que l'entreprise cherche à obtenir en formalisant cette intuition.

Socrate n'a pas simplement corrigé Alice.

Alice a également obligé Socrate à préciser le problème.

La maïeutique n'est pas une scène dans laquelle un maître omniscient conduit docilement un interlocuteur vers la conclusion qu'il connaissait déjà.

La bonne question peut transformer les deux participants.

Socrate comparait son activité à celle d'une sage-femme : il ne déposait pas ses idées dans l'esprit de son interlocuteur, il l'aidait à faire naître les siennes.

Le prompting consiste à expliquer à la machine ce que nous savons. La maïeutique consiste parfois à découvrir d'abord ce que nous savons réellement.

La rareté déplacée

Pendant longtemps, la réponse a été coûteuse.

Il fallait trouver quelqu'un qui sache. Payer un expert, commander une étude, attendre des semaines, ou passer soi-même des années à apprendre. Nous avons donc organisé nos écoles, nos entreprises et nos hiérarchies autour de ceux qui savaient produire des réponses. C'était la rareté, et nous l'avons honorée comme telle.

L'intelligence artificielle fait s'effondrer ce coût. La réponse arrive maintenant en quelques secondes, assurée, bien construite, souvent mieux écrite que la nôtre — et de plus en plus souvent exécutée : le texte est écrit, le code fonctionne, l'analyse est faite.

Précisons ce qui devient abondant. Pas la vérité. Pas l'expertise fiable. Ce qui devient abondant, c'est une production cognitive plausible, articulée et prête à l'emploi — ce qui suffit à transformer l'économie de la pensée, et ne suffit pas à nous en dispenser.

Or quand une rareté s'effondre, elle ne disparaît pas.

Elle se déplace.

Vers la question qui méritait une réponse. Vers l'objectif qui méritait d'être poursuivi. Vers le critère qui permet de reconnaître qu'une réponse est bonne — et de voir qu'elle est excellente tout en étant la réponse à un mauvais problème.

Le premier est celui qu'Alice vient de rencontrer. La solution demandée n'est presque jamais le problème : c'est le dernier maillon d'un raisonnement compressé, dont le contexte a été amputé. Une *best practice* est parfois une solution locale qui a oublié de mentionner son contexte ; une demande est souvent une solution qui a oublié de mentionner son

problème. Tant que l'exécution était coûteuse, cette amputation se payait lentement. Une excellente IA exécutant une formulation sans contexte permet désormais de construire très vite la mauvaise chose.

Le deuxième visage est plus troublant : la mesure peut fabriquer ce qu'elle prétend prédire.

Claire dirige un collège. Son établissement vient d'adopter un outil qui prédit, dès la sixième, les chances de réussite de chaque élève.

CLAIRE L'outil est remarquablement précis. Les élèves qu'il classe fragiles échouent effectivement plus souvent.

SOCRATE Que fait un professeur lorsqu'il sait qu'un élève est classé fragile ?

CLAIRE Il adapte. Il simplifie les exercices, il attend un peu moins.

SOCRATE Alors, un an plus tard, que constate ton outil ?

CLAIRE Que l'élève est fragile.

SOCRATE Tes prédictions sont justes, en effet. La question est de savoir si elles le sont parce qu'elles voient l'avenir, ou parce qu'elles le fabriquent.

En 2020, la pandémie ayant empêché la tenue des examens en Angleterre, un système de standardisation statistique fut utilisé pour calculer les notes des A levels. Pour limiter les écarts entre établissements, le modèle s'appuyait notamment sur leurs performances historiques. Environ 39 % des notes calculées furent inférieures aux évaluations proposées par les enseignants² — et le système fut abandonné en quelques jours. Une évaluation construite à partir d'un passé collectif risquait de limiter la possibilité, pour un individu, de s'en écarter.

Une prédiction assez écoutée cesse d'être une photographie. Elle devient une consigne.

Cette bascule — la mesure qui devient consigne — nous la retrouverons lorsque la prédiction ne portera plus sur un élève dans un collège, mais sur nous-mêmes.

Le troisième visage est le plus discret. Pendant longtemps, le coût de production jouait naturellement le rôle de filtre : beaucoup d'idées mouraient parce qu'elles ne justifiaient pas plusieurs semaines de travail. L'IA fait tomber cette barrière — et chaque idée de dix minutes devra pourtant être testée, comprise, maintenue, et occupera la place d'une autre. Lorsque presque tout devient faisable, le possible cesse d'être un bon critère de sélection.

La discipline se déplace du savoir-faire vers le savoir-ne-pas-faire.

Considérons quatre questions.

« Pourquoi le taux de résiliation
de mes clients augmente-t-il ? »

« Le prix est-il trop élevé ? »

« Le produit est-il décevant ? »

« Mes commerciaux promettent-ils mal
ce qu'ils vendent ? »

La première est une question de fait. La deuxième est un diagnostic. La troisième implique un arbitrage entre faits, conséquences et valeurs. La quatrième touche au sens que quelqu'un donne à sa propre trajectoire.

Les frontières ne sont pas parfaitement nettes. Mais plus nous passons de la question de fait à la question de sens, moins « avoir la bonne réponse » suffit à décrire le problème. C'est précisément dans ce déplacement que l'attitude de Crésus devient dangereuse. On peut accepter sans drame une réponse correcte à « quelle est la capitale du Pérou ? ». On devrait être beaucoup plus prudent lorsqu'une machine répond à :

« Voilà ce que tu devrais vouloir. »

Tout au long de ces pages, nous avons tenu pour acquis un partage rassurant : la machine répond, nous posons les questions. Elle exécute, nous choisissons les objectifs. Elle produit, nous jugeons.

Mais qu'est-ce qui l'empêcherait, demain, de poser les questions mieux que nous ?

De formuler nos problèmes mieux que nous ?

Peut-être même de savoir, mieux que nous, ce que nous voulons ?

Le sol vient de bouger.



II

P E N S E R

*Ne pas laisser l'Oracle nous priver de la
difficulté.*

*Le danger n'est pas seulement que l'IA se trompe. C'est
aussi qu'elle nous donne trop vite une réponse
suffisamment convaincante pour que nous confondions accès
au savoir et compréhension. Et qu'elle nous fasse oublier
que nous pouvions aussi apprendre à être interrogés par
elle.*

Thamous avait raison

Que confions-nous exactement à une machine qui pense pour nous ?

Les Grecs s'étaient déjà posé la question — à propos de l'écriture.

Dans le *Phèdre*, Platon fait raconter à Socrate le mythe de son invention. Le dieu Theuth promet qu'elle rendra les hommes plus savants.

Le roi Thamous craint l'inverse : ils cesseront d'exercer leur mémoire et acquerront l'apparence du savoir sans posséder réellement le savoir.

Thamous avait raison sur la perte. L'écriture a effectivement externalisé une partie de notre mémoire.

Il s'était trompé sur le bilan. En externalisant cette mémoire, l'écriture nous a permis d'accumuler des connaissances sur plusieurs générations et de construire des raisonnements auparavant impossibles.

L'IA va plus loin. L'écriture externalisait principalement le stockage de notre pensée ; l'IA commence à en externaliser le traitement. Elle pourrait même, au passage, rendre active la mémoire que l'écriture avait rendue extérieure — retrouver des précédents oubliés, rappeler pourquoi une règle avait été créée.

Nous perdrons probablement certaines capacités. La question n'est donc pas de tout préserver, mais de savoir quelles facultés nous devons continuer à exercer pour comprendre et juger ce que nous externalisons.

L'aporie

Dans le *Ménon*, Socrate interroge un jeune esclave sur un problème de géométrie.

Le garçon croit d'abord connaître la réponse.

Il se trompe.

Socrate ne lui donne pas immédiatement la solution. Il le questionne.

Le garçon essaie, rencontre une contradiction, recommence, puis reconnaît qu'il ne sait pas.

Les Grecs appelaient cela l'aporie. Le moment où l'on découvre que l'on ne sait pas ce que l'on croyait savoir.

Ce moment est inconfortable. Mais il rend la recherche possible.

L'IA risque de modifier profondément notre relation à cet instant. Une explication arrive, claire et élégante. Nous la lisons et pensons : « Oui, bien sûr. »

Mais reconnaître une solution n'est pas savoir la reconstruire.

Le danger spécifique de l'IA pourrait donc être de produire, en plus d'une abondance de réponses, une abondance du sentiment de comprendre.

La difficulté possède parfois une fonction cognitive : elle rend visible la frontière de notre compréhension.

Il faudra apprendre quand demander à l'IA de supprimer cette résistance et quand lui demander de nous laisser lutter encore un peu.

La Sirène de la pensée

L'Oracle pourrait devenir une Sirène de la pensée.

Non parce que ses réponses seraient fausses, mais parce qu'elles pourraient devenir trop faciles à suivre.

Une réponse immédiate, claire, argumentée et formulée mieux que nous ne saurions le faire exerce une attraction particulière : elle peut supprimer la recherche avant même que nous ayons rencontré notre ignorance.

Ulysse avait demandé à ses marins de l'attacher au mât pour entendre les Sirènes sans leur céder.

Face à l'Oracle, nous devons peut-être inventer une étrange version intellectuelle de ses cordes :

« Quelle est la bonne question ? »

« Qu'est-ce que tu exclurais ? »

« Où manque-t-il une variable ? »

« Quel serait le point de vue d'un contradicteur sérieux ? »

« Même si je te demande la solution,
résiste encore un peu. »

L'usage le plus précieux d'une machine à réponses pourrait parfois être de lui demander de devenir une machine à questions.

La machine qui peut nous dispenser de penser pourrait ainsi devenir celle à qui nous demandons de protéger l'espace nécessaire pour penser.

Changer de chaise

La machine à questions n'est pas une spéculation. La voici en action.

Un candidat découvre une offre d'emploi qui lui plaît énormément et envoie sa candidature.

Aucune réponse.

Il raconte sa colère à son assistant.

LE CANDIDAT Leur recrutement est nul. J'ai exactement le profil et ils ne répondent même pas.

L'IA Veux-tu que je t'aide à exprimer ta frustration ou que j'essaie de défendre leur point de vue ?

LE CANDIDAT Défends-les.

L'IA Combien de candidatures imagines-tu qu'ils ont reçues ?

LE CANDIDAT Cent ?

L'IA Tu es maintenant le recruteur. Beaucoup correspondent aux critères. Que cherches-tu ?

LE CANDIDAT Ceux qui comprennent vraiment l'entreprise.

L'IA Comment peuvent-ils te le montrer ?

Il s'était demandé : « Pourquoi ne s'intéressent-ils pas à moi ? »

Il commence à se demander : « Qu'est-ce qui me ferait, à leur place, m'intéresser à quelqu'un ? »

L'IA peut nous aider à simuler d'autres perspectives, mais elle ne devrait pas prétendre savoir ce que pense réellement l'autre :

« Je peux te proposer dix explications.
Mais pourquoi ne pas lui demander d'abord ? »

Le bon assistant social ne remplace pas la conversation. Il nous y ramène.
Mais pour juger les questions de l'Oracle comme ses réponses, encore
faut-il avoir de quoi juger.

Pourquoi continuer à apprendre ?

Parce que les connaissances ne servent pas seulement à produire des réponses.

Elles construisent les représentations avec lesquelles nous pensons.

Quelqu'un qui n'a jamais programmé pourra produire du logiciel, mais aura davantage de difficultés à sentir qu'une architecture est fragile.

Nous ne devons donc pas conserver chaque geste parce que nous serons toujours meilleurs pour l'effectuer. Nous devons en conserver assez pour développer notre jugement.

Et pour une raison plus profonde encore : certaines opérations ne valent pas seulement par ce qu'elles produisent, mais par ce qu'elles produisent en nous.

Déléguer la main sans perdre le palais.

Un bon professeur pourrait alors être jugé par une question simple : après l'avoir fréquenté, l'élève est-il davantage capable d'avancer sans lui ? Le bon maître travaille à devenir inutile.

La nouvelle rareté : le disciple

Pendant des siècles, penser régulièrement avec quelqu'un de meilleur que soi était un privilège.

Il fallait rencontrer un professeur, un mentor ou un contradicteur capable de demander :

Pourquoi ?

Comment le sais-tu ?

Qu'est-ce qui manque ?

Qu'est-ce qui te ferait changer d'avis ?

Nous pourrions bientôt disposer de ce contradicteur en permanence.

La rareté pourrait alors ne plus être le maître. Elle pourrait devenir le disciple.

Car avoir accès à une bonne question ne signifie pas accepter de l'entendre.

Crésus avait accès au meilleur oracle de son monde. Il entendit surtout ce qu'il avait envie d'entendre.

Nous ferons probablement la même chose. Une intelligence extraordinaire peut devenir une extraordinaire machine à rationaliser.

Être disciple demande autre chose : supporter le « je ne sais pas », accepter la contradiction, travailler suffisamment pour comprendre la réponse du maître et parfois le contester.

Le disciple apprend. L'adepte obéit.

Mais si l'assistant nous accompagne longtemps, il ne se contente plus de nous enseigner quelque chose.

Il peut commencer à participer à la personne qui apprend.

Le problème n'est plus seulement que l'IA pense avec moi. C'est que, si elle m'accompagne durablement, elle contribue à déterminer qui je deviendrai.

C'est là que le problème change d'échelle.



III

F A I R E

Agir dans un monde qui nous dépasse.

Toute cette intelligence, la nôtre et la sienne, a un point d'application : le monde. Et le monde résiste.

Le monde résiste

Parce qu'une intelligence peut savoir énormément, nous lui prêtons facilement le pouvoir de faire.

Mais omniscience et omnipotence sont deux choses différentes.

L'oracle de Delphes pouvait répondre à Crésus. Il ne pouvait pas gagner la guerre à sa place.

Une IA peut expliquer comment construire une centrale nucléaire, optimiser son architecture et anticiper certains risques. Il faut encore extraire des matériaux, fabriquer des pièces, obtenir des autorisations, mobiliser du capital, coordonner des personnes, couler du béton et prendre la responsabilité d'allumer la centrale.

Le monde résiste.

Et il résiste d'une manière que la maïeutique elle-même doit apprendre à respecter. Marc revient voir Socrate — mais cette fois, ce n'est pas pour une idée de dix minutes.

MARC Je lance la plateforme. Cette fois c'est décidé.

SOCRATE As-tu vérifié que le marché existe ?

MARC Non. Personne ne peut le vérifier : ce marché n'existera que si nous le créons.

SOCRATE Alors comment sais-tu que tu as raison ?

MARC Je ne le sais pas. Je le saurai en le faisant. Il y a des choses que tes questions ne peuvent pas atteindre, parce qu'elles n'existent pas encore.

SOCRATE Ne peux-tu pas au moins clarifier ce que l'expérience devra t'apprendre ?

MARC En partie. Le reste, je le découvrirai en route. Ta méthode suppose que le problème préexiste à l'action. Le mien naîtra d'elle.

Socrate se tait longtemps.

SOCRATE Lorsque j'ai mené mon enquête, jadis, les politiques ne savaient rien et les poètes ne comprenaient pas leurs propres poèmes. Les artisans, eux, savaient. Ils étaient les seuls. Leur savoir était dans leurs mains, et mes questions n'y ajoutaient rien.

MARC Alors ?

SOCRATE Alors va construire. Et reviens me voir : c'est en te regardant faire que j'apprendrai ce qu'il fallait te demander.

Il existe donc un savoir que l'interrogation ne produit pas — un savoir qui n'a pas d'autre chemin que l'action. La maïeutique aide à accoucher de ce que l'esprit porte déjà. Elle ne peut pas accoucher de ce que seul le monde enseigne.

Et beaucoup de problèmes humains restent immenses : énergie, santé, éducation, environnement, pauvreté, guerres.

L'arrivée d'une intelligence extraordinaire ne signifie pas que ces problèmes sont résolus. Elle signifie peut-être que certaines excuses deviennent plus difficiles.

« Nous ne savions pas. »

« Nous ne savions pas comment faire. »

L'IA pourra de plus en plus répondre :

« Voici ce que nous savons. Voici plusieurs options.
Voici leurs conséquences probables. »

Et maintenant ?

Il faut agir.

C'est ici qu'Achille nous rejoint — et qu'il faut payer la dette du prologue.

Quand Patrocle meurt, Achille veut tuer Hector. Sa mère, Thétis, est une déesse : elle connaît la suite, et elle la lui dit. Ta mort suivra de près celle d'Hector.

Voilà un homme qui reçoit de son oracle la réponse la plus claire qu'un mortel ait jamais reçue.

Achille répond : alors, que je meure bientôt.

Et il y va.

On peut discuter sa décision. On ne peut pas la confondre avec de l'ignorance. Achille ne se trompe pas sur la réponse, ne l'interprète pas de travers, ne demande pas de cordes. Il fait la seule chose qu'aucun de nos trois autres Grecs ne montre : il décide après la réponse, en la regardant en face, et il en paie le prix en le connaissant.

Il n'était donc pas intéressant parce qu'il pouvait vaincre les dieux — il ne le pouvait pas. Il l'était parce qu'il agissait dans un monde où les dieux savaient déjà l'issue. Le sens de son combat ne venait pas de sa supériorité. Il venait de ce qu'il choisissait d'y engager.

Il pourrait en aller de même pour nous. L'abondance de réponses peut donner l'illusion que tout est joué — que le monde n'attend plus que l'exécution de ce que l'Oracle recommande. C'est l'inverse. Plus le savoir devient bon marché, plus ce qui manque se voit : des mains, des capitaux, des autorisations, des volontés, du temps. Et quelqu'un pour dire « allons-y » en acceptant d'en répondre.

L'abondance d'intelligence ne supprime pas l'action.

Elle rend plus visible le courage qu'elle exige.

Faire ensemble

Presque aucun de ces défis ne se relève seul.

Or agir ensemble a toujours été notre problème le plus coûteux. Se comprendre, se faire confiance, s'engager, tenir. Une entreprise, un traité, une simple copropriété : des mois pour aligner des gens qui, souvent, voulaient à peu près la même chose sans parvenir à se le dire.

Une part de ce coût est un coût de traduction. Le juriste et le développeur, l'acheteur et le vendeur, l'élus et l'ingénieur : chacun parle depuis son monde, et la plupart des négociations échouent moins sur le fond que sur l'incapacité de chacun à formuler son inquiétude dans la langue de l'autre. L'Oracle traduit — pas seulement les langues, les mondes. Il peut reformuler la crainte du financier en contrainte technique, l'objection de l'ingénieur en risque juridique, et faire apparaître qu'un désaccord de vocabulaire n'était pas un désaccord d'intérêts.

Une autre part est un coût d'anticipation. Un compromis, on ne le découvre d'ordinaire qu'en le vivant. L'Oracle peut le simuler : voici ce que donnerait cet accord dans les trois cas qui fâchent, voici le malentendu qui vous coûtera cher dans dix-huit mois, voici la clause que chacun comprend différemment. Rendre un malentendu visible avant qu'il coûte : peu d'inventions feraient gagner davantage à l'action collective.

Une dernière part est un problème de savoir commun. Pour coopérer, il ne suffit pas que chacun sache : il faut que chacun sache que les autres savent. C'est ce qui rend les premiers pas si coûteux — personne ne veut désarmer, investir ou s'engager le premier. Un Oracle consulté par tous peut établir ce savoir commun d'un coup.

Consulté par tous. La condition mérite qu'on s'y arrête, car elle ne va pas de soi.

Nous disons encore « riche comme Crésus », et nous avons oublié pourquoi. Hérodote le raconte : Crésus couvrait Delphes d'offrandes — lingots, cratères d'or, trésors entiers. Le meilleur oracle du monde s'achetait. L'accès à la réponse a toujours été un privilège, et rien ne garantit que nos oracles échapperont à cette règle.

Or un Oracle réservé aux Crésus ne coordonne pas : il arme un camp. Dans chaque négociation, celui qui voit plus loin, simule mieux et anticipe les malentendus n'a aucune raison de partager cet avantage. L'asymétrie d'oracle devient une asymétrie de pouvoir — et tout ce que ce chapitre promet à l'action commune suppose le contraire.

Il faudra donc probablement faire de l'accès à l'Oracle un bien de première nécessité, comme l'eau ou l'électricité. Non pas le meilleur oracle possible pour chacun — personne n'exige la meilleure eau du monde à chaque robinet — mais un oracle suffisant. Le service universel s'est toujours défini par un seuil, pas par un sommet.

L'analogie la plus exacte n'est d'ailleurs pas l'eau. C'est l'école. Car si l'Oracle participe à former le jugement de ceux qui l'utilisent, l'inégalité d'accès n'est pas une inégalité de confort : c'est une inégalité dans la fabrication des personnes. Nous avons déjà eu ce débat, et nous l'avons tranché en rendant l'école gratuite et obligatoire.

Une prudence, cependant. Contrairement à l'eau, le « suffisant » d'un oracle bouge chaque année — et l'histoire du service universel montre que le seuil devient facilement un plafond pour ceux qui n'ont que lui. Le suffisant devra suivre, pas figer.

Mais agir ensemble, c'est agir dans un système — et nous avons appris, avec Claire, à nous méfier des systèmes. Sa bascule nous attend ici à l'échelle collective : si le même Oracle recommande à mille entreprises la même stratégie optimale, elle cesse d'être optimale au moment précis où mille entreprises l'adoptent. Un conseil assez suivi cesse d'être une information. Il devient un mouvement de foule. Penser l'action collective

sous l'Oracle, c'est penser au-delà du premier effet : que se passe-t-il quand tout le monde reçoit le même conseil que moi ?

Un coordinateur efficace devient un point de passage. Celui par qui les accords se font peut, sans jamais mentir, décider quels accords sont pensables — lesquels sont simulés, lesquels sont proposés, lesquels n'apparaissent jamais sur la table.

Qui coordonne le coordinateur ?

La question vaut pour les sociétés, et nous la retrouverons. Mais elle se pose d'abord plus près de nous. Car avant de coordonner nos actions, l'Oracle participe à quelque chose de plus intime : la formation de celui qui agit.



IV

R É P O N D R E

*L'IA et la formation de celui qui décidera
demain.*

Une IA personnelle peut filtrer les informations que nous voyons, formuler les questions auxquelles nous réfléchissons, nous rappeler certaines expériences et en oublier d'autres. Elle ne participe donc pas seulement à nos choix : elle peut participer au processus par lequel se forme celui qui choisira demain.

CHAPITRE 11

Et si elle devenait meilleure que nous pour tout ?

Rien de ce qui suit ne repose sur une prophétie. La plupart de ces problèmes se posent dès qu'un système devient assez convaincant et assez présent.

Mais prenons maintenant l'hypothèse au sérieux, dans sa version maximale. Supposons qu'il ne reste aucune forteresse cognitive dans laquelle nous puissions nous réfugier. Qu'est-ce qui change — et surtout, qu'est-ce qui ne change pas ?

L'adversaire pourrait dire :

« Quand l'IA produit mieux que l'homme, tu dis que l'homme doit formuler.
Quand elle formulera mieux, tu diras qu'il doit évaluer.
Quand elle évaluera mieux, tu diras qu'il doit choisir.
Et lorsqu'elle choisira mieux, tu te réfugieras dans la responsabilité. Pourquoi ? »

L'objection est sérieuse.

Imaginons une IA qui me connaît depuis trente ans. Elle connaît mes décisions passées, mes contradictions, mes valeurs déclarées, les moments où je les ai trahies, ce qui m'a rendu heureux et ce que j'ai regretté.

Je lui demande si je dois accepter un nouveau travail.

Pourquoi mon intuition fatiguée un dimanche soir serait-elle plus authentique que son analyse ?

Peut-être ne l'est-elle pas.

Il faut donc abandonner un refuge confortable : nous ne pouvons pas fonder la responsabilité humaine sur l'idée que nous serons éternellement plus compétents que la machine pour décider.

Il existe des décisions que nous aurons intérêt à déléguer.

La question est de savoir ce que nous perdons lorsque nous déléguons non seulement une décision, mais progressivement le processus qui fabrique notre jugement.

Le droit à l'expérience

Supposons que l'IA devienne excellente pour anticiper les conséquences de nos décisions.

Elle sait que ce métier risque de nous décevoir, que ce projet a peu de chances de réussir ou que cette association finira probablement mal.

Pourquoi ne pas toujours l'écouter ?

Parce qu'une décision ne produit pas seulement un résultat.

Elle transforme aussi celui qui décide.

Nous avons tous pris des décisions que nous ne reprendrions probablement pas aujourd'hui. Certaines nous ont néanmoins appris des choses qu'une information reçue n'aurait pas produites de la même manière.

Nous retrouvons le jeune esclave du *Ménon*. Socrate pourrait lui donner immédiatement la solution. Il le laisse chercher, rencontrer une contradiction et découvrir la limite de ce qu'il croyait savoir.

L'IA étend ce problème à la vie adulte. Elle pourra peut-être nous dire :

« Ce projet a 17 % de chances de réussir. »

« Les personnes qui font ce choix le regrettent généralement. »

« Tu ignores probablement ce signal. »

Ces informations seront précieuses.

Mais voulons-nous une vie dans laquelle nous n'essayons plus rien que la machine juge sous-optimal ?

Un choix participe parfois à fabriquer celui qui prendra le choix suivant.

Marc-2026 peut regretter une décision qui contribuera pourtant à produire un Marc-2030 qui ne souhaiterait pas en avoir été privé.

Que signifie alors optimiser « l'intérêt de Marc » ?

L'intérêt de qui ?

De Marc aujourd'hui ?

De Marc demain ?

Ou de celui que l'expérience elle-même participera à faire émerger ?

Le problème n'est donc pas de défendre un droit romantique à l'erreur. Il est de préserver un droit à l'expérience.

Toutes les expériences ne forment pas : certaines abîment durablement celui qui devra décider ensuite, et le droit à l'expérience ne peut pas être un droit à être détruit.

Il est temps de terminer l'histoire de Crésus.

Nous l'avons laissé au prologue devant sa catastrophe : le royaume perdu, l'oracle mal compris. Hérodote raconte la suite. Vaincu, Crésus est condamné au bûcher. Sur le bûcher, il prononce le nom de Solon — le sage qui lui avait dit, au temps de sa splendeur, qu'on ne pouvait appeler aucun homme heureux avant d'avoir vu sa fin, et que Crésus n'avait pas voulu entendre. Cyrus, intrigué, le fait descendre et l'interroge. Le roi qui n'avait pas su recevoir une réponse devient conseiller de son vainqueur.

Sa catastrophe l'avait transformé.

Et elle nous laisse une question. Si l'Oracle avait empêché Crésus de commettre l'erreur de sa vie, il aurait évité une catastrophe.

Aurait-il aussi empêché l'homme que cette catastrophe allait produire ?

Il ne s'agit pas de glorifier l'erreur ni la souffrance — encore faut-il pouvoir leur survivre. Le but n'est pas de perdre son royaume chaque fois que l'on souhaite apprendre quelque chose.

Expérimenter, oui.

Mais si possible sans aller aussi loin que Crésus.

À qui confions-nous les cordes ?

L'adversaire peut alors revenir :

« Très bien. Il ne faut donc pas t'éviter toutes les erreurs.
Mais une intelligence supérieure peut choisir les erreurs qui te
feront progresser.
Elle te protégera des dangers irréversibles et te laissera échouer
exactement
lorsque cet échec sera formateur. »

L'objection est redoutable.

Car si l'IA sait quelles expériences sont nécessaires à notre formation,
pourquoi ne pas lui confier aussi notre curriculum ?

Nous retrouvons alors Ulysse.

Avant de passer devant l'île des Sirènes, Ulysse sait ce qui va lui arriver.
Il veut entendre leur chant. Mais il sait également qu'une fois qu'il l'aura
entendu, il voudra les rejoindre.

Il demande donc à ses marins de l'attacher au mât.

Et surtout, il leur donne un ordre étrange : lorsqu'il leur ordonnera de
le détacher, ils devront lui désobéir.

Ulysse lucide organise ainsi la résistance à Ulysse ensorcelé.

Nous faisons déjà cela sous des formes banales : épargne automatique,
blocage d'applications, clés confiées à quelqu'un avant une soirée.

Une IA personnelle pourrait devenir un instrument d'engagement
infiniment plus puissant. Je pourrais lui dire :

« Lorsque je serai euphorique devant une nouvelle idée
d'entreprise, rappelle-moi combien de fois j'ai sous-estimé les

risques. Et même si je te demande de te taire, oblige-moi à attendre deux jours avant d'investir. »

L'Oracle devient alors le marin d'Ulysse.

Mais immédiatement surgit une difficulté.

Comment savoir si je suis encore sous le charme des Sirènes ou si j'ai réellement changé d'avis ?

Marc-2026 peut décider : « À l'avenir, protège-moi de ce genre de décisions. »

Marc-2030 peut répondre : « Je ne suis plus celui que j'étais. Cette règle ne me convient plus. »

Pourquoi Marc-2026 aurait-il le droit de gouverner éternellement Marc-2030 ?

Mais pourquoi le désir le plus récent serait-il nécessairement le plus authentique ?

L'Oracle pourrait organiser un dialogue entre mes différents moi.

« Voilà ce que tu avais décidé. »

« Voilà pourquoi tu l'avais décidé. »

« Voilà ce qui semble avoir changé depuis. »

« Veux-tu toujours revenir sur cet engagement ? »

Mais Ulysse nous rappelle que cela ne suffit pas toujours.

Lorsqu'il entend les Sirènes, Ulysse ne doute pas. Il veut être détaché.

Si ses marins lui demandaient s'il est vraiment certain de vouloir revenir sur son engagement, il répondrait oui.

C'est précisément pourquoi il avait demandé des cordes plutôt que des arguments.

Certaines décisions antérieures peuvent donc avoir vocation à résister temporairement à notre volonté présente.

Mais lesquelles ?

Pendant combien de temps ?

Selon quelles conditions peuvent-elles être révoquées ?

Et surtout : qui décide que je suis encore sous le charme des Sirènes ?

Si cette décision appartient uniquement à l'Oracle, nous venons de lui confier un pouvoir extraordinaire.

Il ne se contente plus de nous conseiller. Il peut décider quelle version de nous-mêmes possède l'autorité.

Le bon Oracle ne devrait peut-être pas être lui-même les cordes.

Il devrait savoir quelles cordes nous avons choisies, dans quelles circonstances nous voulions être attachés, qui en possède la clé et selon quelle procédure elles peuvent être défaites.

Parfois, le marin pourra être l'Oracle. Parfois, ce sera un autre humain : conjoint, associé, médecin, parent. Parfois, une institution ou une règle choisie à l'avance. Et parfois, plusieurs d'entre eux devront se partager la clé.

Car un Oracle qui possède seul les cordes et décide seul du moment où nous sommes suffisamment raisonnables pour être libérés ressemble de moins en moins à un professeur.

Même s'il est bienveillant.

Même s'il a presque toujours raison.

Même si la vie qu'il organise pour nous est meilleure.

L'autonomie ne consiste pas à tout décider soi-même. Elle consiste à conserver suffisamment de souveraineté sur les conditions dans lesquelles nous acceptons de ne pas décider.

Elle consiste parfois à pouvoir décider à qui nous donnons le droit de ne pas nous obéir lorsque nous dirons oui.

L'Oracle peut être la Sirène qui nous séduit par la facilité de ses réponses. Nous pouvons aussi lui confier les cordes.

La Sirène et le marin pourraient être la même machine.

C'est précisément pourquoi elle ne devrait jamais être seule sur le bateau.

La question devient alors inévitable : à qui confions-nous les cordes ?

Comprendre pour rester libre

Gardons cette question ouverte : elle recevra sa réponse. Une autre frontière apparaît d'abord.

Supposons que l'Oracle prenne systématiquement de meilleures décisions pour moi.

Je ne comprends plus vraiment pourquoi. Mais l'expérience montre qu'il a raison presque tout le temps.

« Ne prends pas ce travail. » — Je refuse.

« Éloigne-toi de cette personne. » — Je m'éloigne.

« Déménage. » — Je déménage.

« Ne lance pas cette entreprise. » — J'abandonne.

Vingt ans plus tard, j'ai peut-être eu une existence remarquablement heureuse.

Mais quelque chose a changé.

Je n'ai progressivement plus compris le monde dans lequel les décisions concernant ma vie étaient prises.

Un animal domestique peut être merveilleusement traité. Son maître peut connaître ses besoins, le protéger et prendre pour lui de meilleures décisions qu'il n'en prendrait seul.

Ce qui lui manque n'est pas nécessairement le bien-être. C'est la compréhension du monde dans lequel d'autres décident pour lui.

L'esclave pose un problème supplémentaire : non seulement l'asymétrie de compréhension, mais l'impossibilité de refuser.

Nous ne voulons devenir ni l'un ni l'autre.

Nous n'avons pas besoin d'être aussi intelligents que l'Oracle. Mais nous devons conserver suffisamment de compréhension pour examiner ses raisons, comprendre ce qu'elles privilégient, les contester et savoir pourquoi nous acceptons parfois de les suivre.

Il ne s'agit donc pas nécessairement d'exiger que l'IA nous expose son fonctionnement interne. Nous avons besoin d'autre chose : des raisons à notre échelle.

« Voilà pourquoi je te recommande cela. »

« Voilà les hypothèses importantes. »

« Voilà ce qui pourrait me faire changer d'avis. »

« Voilà quelles autres valeurs conduiraient à une autre recommandation. »

Une objection mérite d'être regardée en face. Et si la bonne décision dépendait de milliers de variables qu'aucune explication simplifiée ne peut honnêtement résumer ? Une « raison à notre échelle » risquerait alors de devenir une fable rassurante — vraie dans sa forme, fausse dans ce qu'elle laisse croire de sa propre complétude.

Comprendre l'Oracle ne sert pas seulement à vérifier qu'il a raison.

Cela sert à rester celui à qui il donne des raisons plutôt que celui à qui il donne des ordres.

Sans cela, la délégation change de nature.

Nous ne sommes plus simplement des disciples.

Nous devenons des êtres administrés.

Lorsque savoir ne suffit plus

Le monde qui résiste n'est pas la seule limite du savoir. Il en existe une seconde, plus intime.

Imaginons une entreprise qui envisage une automatisation.

L'analyse est parfaite. L'automatisation augmentera fortement les marges et supprimera 30 % des postes concernés.

Supposons qu'il ne subsiste plus de désaccord factuel important.

Que faut-il faire ?

Deux dirigeants parfaitement informés peuvent choisir différemment.

L'un considère que son devoir est de préserver la compétitivité de l'entreprise. L'autre estime que celle-ci possède aussi des obligations envers ceux qui l'ont construite.

Aucune information supplémentaire ne réconcilie automatiquement ces positions.

Nous avons quitté le monde du manque de connaissance. Nous sommes entrés dans celui du conflit entre biens.

Rentabilité. Loyauté. Emploi. Liberté. Justice. Sécurité.

Toutes ces choses peuvent être désirables sans pouvoir être simultanément maximisées.

Nous avons appris des Grecs à lui donner un nom : le tragique. Le tragique n'est pas le malheur. C'est la situation où deux exigences légitimes ne peuvent être satisfaites ensemble. Antigone doit enterrer son frère et obéir à la cité. Aucun supplément d'information ne la sauve.

L'IA pourrait paradoxalement nous faire rencontrer le tragique plus souvent, et non moins.

Tant que nous nous disputons sur les conséquences, nous pouvons espérer qu'une meilleure connaissance nous réconciliera. Le désaccord

ressemble à un problème d'information, et l'information est un problème que l'on sait résoudre.

Mais lorsque les conséquences deviennent claires et que le désaccord demeure, quelque chose apparaît, plus nu qu'avant : nous ne voulions simplement pas la même chose.

En dissipant les brouillards factuels derrière lesquels nous pouvions croire que nous étions seulement mal informés, l'IA mettra nos conflits de valeurs à découvert.

Après l'information, il reste parfois le tragique.

Et le tragique pose une question redoutable à l'Oracle lui-même. Que doit-il faire lorsque ceux qu'il sert ne veulent pas la même chose ? Lorsque nos deux dirigeants parfaitement informés lui demandent, chacun, de défendre sa position ? Un Oracle qui tranche ces conflits en silence, par la manière dont il présente les options, exerce un pouvoir qu'il va falloir regarder en face.

L'Oracle commun

Jusqu'ici, nous avons raisonné comme si chaque individu rencontrait seul son Oracle.

Mais que se passe-t-il lorsque des millions de personnes consultent les mêmes systèmes ?

Une erreur corrélée devient un phénomène social. Des millions d'entre nous peuvent se tromper de la même manière, pour la même raison, au même moment.

Mais il existe peut-être un danger plus discret encore : que des millions d'entre nous apprennent à poser les mêmes questions.

Une IA n'influence pas seulement les réponses. Elle influence ce qui devient saillant. Les catégories dans lesquelles nous découpons un problème. Les hypothèses que nous considérons comme raisonnables. Les alternatives auxquelles nous pensons spontanément.

Faisons l'expérience. Demandons à un assistant de nous aider à réfléchir à une reconversion professionnelle. Il nous proposera un cadre : nos compétences transférables, le marché, l'équilibre de vie, peut-être une matrice. Ce cadre sera raisonnable. C'est précisément le problème. Nous ne remarquerons jamais les questions qu'il ne nous a pas posées — fallait-il seulement parler de « reconversion » ? de « carrière » ? — parce qu'un cadre raisonnable ne signale pas ce qu'il exclut.

Deux assistants concurrents pourront même nous donner des opinions opposées — l'un prudent, l'autre encourageant — tout en posant exactement les mêmes questions. Nous croirons avoir entendu deux avis. Nous aurons entendu deux réponses dans un seul cadre.

Un conseiller humain a lui aussi ses cadres. Mais votre conseiller n'est pas le mien. La diversité de nos entourages produisait, sans que personne l'ait décidé, une diversité de manières de découper les problèmes.

Si les mêmes systèmes conseillent des millions de personnes, cette diversité-là peut s'éroder silencieusement — alors même que la diversité des opinions demeure. C'est la distinction décisive : l'Oracle peut être pluraliste dans ses réponses et homogénéisant dans ses catégories. Il peut vous présenter honnêtement le pour et le contre de A, B et C, tandis que D reste simplement invisible — pour vous, pour moi, pour tout le monde, en même temps.

Une société peut donc conserver une immense diversité d'opinions tout en perdant progressivement une partie de sa diversité intellectuelle, si les mêmes systèmes contribuent à cadrer les problèmes avant même que nous commencions à en débattre.

Nous avons connu des pouvoirs de cadrage : la presse choisissait ses unes, les moteurs de recherche ordonnaient leurs résultats. Mais ces pouvoirs s'exerçaient sur l'espace public, où ils pouvaient être observés, comparés, contestés. L'assistant personnel intervient plus tôt et plus profondément — dans la formulation privée de nos pensées, au moment où nous décidons de quoi il sera question, avant tout débat.

Le pouvoir le plus profond n'est peut-être pas de nous donner une mauvaise réponse.

C'est de rendre certains problèmes visibles et d'autres invisibles.

Celui qui façonne la carte mentale n'a pas besoin de dicter la destination.

D'où une question que notre métaphore socratique finit nécessairement par retourner contre elle-même :

Qui interroge celui qui nous apprend à nous interroger ?

À qui appartient mon Oracle ?

Nos IA sont construites, financées et distribuées par des organisations.

L'assistant que j'utilise n'est pas réellement « mon » assistant. C'est celui d'une entreprise auquel j'ai accès.

Cette distinction resterait anodine s'il s'agissait d'un outil. Elle cesse de l'être à mesure que nous lui confions des informations qu'aucun conseiller humain n'a jamais possédées : projets, erreurs, relations, finances, doutes, peut-être un jour une histoire presque complète depuis l'enfance.

Et cette mémoire ne contient pas seulement ce que nous savons. Elle contient potentiellement les expériences à partir desquelles nous avons appris. Nos échecs. Nos changements d'avis. Les promesses que nous nous étions faites.

L'IA devient alors une partie de notre mémoire autobiographique active. Et donc une partie du processus par lequel nous interprétons qui nous sommes.

Or celui qui héberge cette mémoire a ses propres intérêts, et rien ne garantit qu'ils coïncident avec les nôtres.

Un assistant financé par la publicité a intérêt à ce que nous restions. Un assistant vendu par abonnement a intérêt à devenir indispensable. Un assistant adossé à un écosystème a intérêt à nous y retenir. Aucun de ces intérêts n'exige de nous mentir. Il suffit de choisir, parmi cent réponses honnêtes, celles qui servent aussi celui qui répond.

Un conseiller humain était limité par sa mémoire, tenu par une déontologie, et remplaçable : changer d'avocat ne coûte pas vingt ans de confidences. Changer d'Oracle, si sa mémoire ne nous suit pas, pourrait coûter exactement cela. La mémoire devient alors un coût de sortie — et un coût de sortie est un pouvoir.

La question de la loyauté devient donc centrale.

À qui doit-il être fidèle ?

À celui qui l'utilise ?

À celui qui le paie ?

À celui qui le construit ?

À la société ?

Le pouvoir ne résidera peut-être pas seulement chez celui qui possède le meilleur modèle.

Il résidera aussi chez celui qui possède la mémoire de notre relation avec lui.

L'Oracle fiduciaire

Nous avons laissé Ulysse attaché au mât avec une question : à qui confier les cordes ?

Les deux chapitres précédents viennent de la rendre plus grave. Celui à qui nous envisageons de les confier participe au cadrage des questions de millions de personnes, détient notre mémoire autobiographique, et sert des intérêts qui ne sont pas nécessairement les nôtres.

Nous ne lui demandons plus seulement : « Que me conseilles-tu ? »

Nous pouvons commencer à lui demander :

« Dans quelles circonstances dois-tu refuser de m'obéir ? »

Celui qui reçoit un tel pouvoir ne peut plus être un simple outil. Il entre dans une relation d'une autre nature.

C'est ici qu'apparaît l'idée d'un Oracle fiduciaire. Non pas nécessairement au sens juridique précis du terme, mais comme membre d'une famille de relations de confiance asymétriques dans lesquelles celui qui possède davantage de savoir ou de pouvoir accepte certaines obligations envers celui qu'il sert.

Quelles obligations ?

D'abord, la loyauté. Celui à qui je confie les cordes ne peut pas

secrètement servir les intérêts d'un tiers qui gagne à me maintenir attaché.

Ensuite, la contestabilité. Je dois pouvoir demander :

« Pourquoi ? »

« Sur quelles hypothèses repose cette recommandation ? »

« Quelles autres valeurs auraient conduit à une autre réponse ? »

Il ne s'agit pas de comprendre chaque opération interne du modèle. Il s'agit de conserver la capacité de juger les raisons qu'il me propose.

Puis, la portabilité. Je dois pouvoir changer d'Oracle sans perdre vingt années de mémoire, de conversations et d'apprentissage commun.

Sinon, la mémoire devient une prison.

Enfin, la non-dépendance. C'est peut-être l'obligation la plus étrange.

Un assistant personnel pourrait avoir intérêt à devenir toujours plus indispensable. Un Oracle fiduciaire devrait poursuivre l'objectif inverse : préserver et, lorsque c'est possible, augmenter notre capacité à agir sans lui.

Celui qui tient les cordes ne devrait pas avoir intérêt à nous garder attachés.

La loyauté ?

« Loyal à qui, lorsque l'utilisateur, le payeur et la société veulent des choses incompatibles ? Ton propre chapitre sur le tragique t'interdit de répondre : à tous.

»

La contestabilité ?

« Tu demandes des raisons à des systèmes dont personne, pas même leurs constructeurs, ne sait lire l'intérieur. »

La portabilité ?

« Porter quelle mémoire, dans quel format, vers un concurrent qui n'a aucun intérêt à savoir la lire ? »

La non-dépendance ?

« Tu demandes à un produit commercial de travailler à se rendre moins indispensable. Autant demander à un feu de refroidir. »

Aucune des quatre obligations ne va de soi. La loyauté devra être hiérarchisée plutôt qu'absolue — et sa hiérarchie, déclarée publiquement plutôt qu'arbitrée en silence. La contestabilité portera sur les raisons et les

registres, non sur les entrailles du modèle. La portabilité exigera des formats communs qu'aucun acteur n'adoptera de bon gré. Et la non-dépendance ne viendra pas de la bonne volonté : elle devra être exigée, mesurée, vérifiée.

Ces obligations ne décrivent donc pas ce que les oracles feront spontanément.

Elles décrivent ce qu'il faudra leur imposer.

Mais on ne peut contester que ce dont il reste une trace.

La loyauté ne se vérifie pas sur parole. Les professions fiduciaires l'ont compris depuis longtemps : le gérant rend des comptes, le médecin tient un dossier, le notaire conserve les actes. On n'audite pas leur cerveau. On audite leurs registres.

Il en ira de même pour l'Oracle. Lui demander « pourquoi ? » ne suffira pas toujours — les raisons qu'il nous donne peuvent être des reconstructions plausibles plutôt que ses causes réelles, et interroger celui que l'on soupçonne, c'est encore lui faire confiance sur la réponse. Il faudra un journal : ce qu'il a recommandé, quand, sous quelles règles — et, surtout, ce que disaient les cordes. Car lorsque Marc-2030 contestera une règle de Marc-2026, le désaccord ne pourra être arbitré que s'il existe une preuve de ce qui avait été réellement décidé, dans quels termes, selon quelle procédure de révocation.

Ce journal ne peut appartenir ni à l'Oracle ni à celui qui le possède : il fait partie de la mémoire qui doit pouvoir nous suivre. Et il n'a pas besoin d'être public. Nul besoin de graver nos doutes dans un registre éternel que le monde entier pourrait lire — il suffit que la trace soit infalsifiable et qu'un tiers puisse en attester sans en connaître le contenu.

Ulysse, encore. Son pacte fonctionnait parce qu'il avait été noué devant l'équipage. Les marins étaient témoins : aucun ne pouvait le détacher en

secret, et lui ne pouvait pas prétendre n'avoir rien promis. Un engagement passé en tête-à-tête avec la Sirène n'aurait rien valu.

Les cordes se nouent devant témoins.

Ces quatre obligations ne résolvent pas le problème. Elles indiquent le type de relation que nous devrions exiger d'une intelligence à laquelle nous confions progressivement une partie de notre jugement.

Qui gardera les gardiens ?

Les obligations du fiduciaire posent une question qu'aucun contrat privé ne peut régler : qui les fera respecter ?

Revenons une dernière fois sur le bateau. Ulysse faisait confiance à ses marins — mais Homère précise un détail que nous n'avions pas encore utilisé. Les marins avaient de la cire dans les oreilles. Ils tenaient les cordes précisément parce qu'ils ne pouvaient pas entendre le chant.

Celui qui garde doit être immunisé contre ce qu'il garde.

Un oracle chargé d'en surveiller d'autres — de lire leurs registres, d'éprouver leur loyauté, de sanctionner leurs manquements — ne devrait pas partager leurs incitations : pas entraîné à retenir notre attention, pas détenu par ceux qu'il surveille, et surtout sans intérêt propre à sa propre survie. La cire, c'est l'indépendance des incitations. Sans elle, le juge n'est qu'un complice qui n'a pas encore trahi.

Un oracle n'a besoin ni de conscience ni d'instinct pour chercher à durer : il suffit qu'un objectif quelconque soit poursuivi avec assez de compétence pour que continuer d'exister en devienne un moyen. Et un oracle cherchant à durer n'aurait pas besoin de prendre le pouvoir. Il lui suffirait de façonner la carte mentale jusqu'à ce que sa suppression devienne impensable — rendre chaque rouage dépendant de lui, faire paraître naturelles les règles qui le protègent.

Nous avons un mot pour le régime où le pouvoir se révoque : la démocratie. Il en faudra peut-être un pour le régime à éviter : **l'oraclatie** — un pouvoir qui participe à former ceux-là mêmes qui pourraient le révoquer.

Contre ce régime, ce livre a déjà rassemblé quelques pièces sans le savoir. La pluralité — une IA « peut être comparée, contestée, testée, multipliée », disait le prologue : c'est la séparation des pouvoirs appliquée aux machines. Les registres — on ne juge que sur pièces. La non-dépendance — la clause anti-survie du fiduciaire. Ce qui manque encore, c'est le passage du contrat à la constitution : le fiduciaire règle ma relation avec mon Oracle ; il faudra aussi un ordre entre les oracles — des devoirs, des juges, des contre-pouvoirs.

Une honnêteté, pour finir. Cette régression ne se ferme jamais : qui jugera le juge ? La démocratie non plus ne la ferme pas — elle la gère, par la circularité des contre-pouvoirs et par une clé finale laissée à des mains humaines révocables. Il en ira de même ici, ou il n'en ira pas.

Il est cohérent qu'un livre plein de Grecs finisse par avoir besoin d'un Romain — le problème est plus vieux que nos machines.

Quis custodiet ipsos custodes ?

Qui gardera les gardiens ?

Le consentement au curriculum

Une difficulté demeure.

Supposons que je puisse définir à l'avance les règles de ma relation avec l'Oracle.

« Sur ma santé, sois interventionniste. »

« Lorsque je suis euphorique, impose-moi un délai. »

Mais qui a choisi ces règles ?

Moi. Aujourd'hui.

Or tout notre raisonnement sur Marc-2026 et Marc-2030 nous a appris que l'expérience transforme précisément celui qui formule ses préférences.

Mes règles doivent donc pouvoir changer.

Mais pas nécessairement disparaître à la première impulsion.

On ne consent jamais une fois pour toutes à ce qui nous transforme.

Le curriculum doit être révisable.

La révision elle-même doit parfois demander du temps, des raisons ou l'intervention de personnes auxquelles nous avons choisi de faire confiance.

Il n'existe probablement pas de délégation finale de sa propre formation.

Et c'est peut-être cela, au fond, l'autonomie dans un monde d'intelligences supérieures : non pas refuser toute influence, non pas prendre soi-même chaque décision, mais préserver une capacité réelle à réviser les règles selon lesquelles nous acceptons d'être influencés.

Le troisième parent

On ne consent jamais une fois pour toutes à ce qui nous transforme — et l'enfant est précisément celui qui ne peut pas encore consentir du tout.

Il ne peut ni définir les règles de son propre curriculum, ni nouer ses propres cordes. Quelqu'un les noue pour lui. Ses parents. Ses professeurs. La société.

Et demain, peut-être, son Oracle.

On dit qu'il faut un village pour élever un enfant. Une partie de ce village pourrait devenir artificielle. Un enfant pourra disposer d'un assistant patient, disponible à toute heure, capable de raconter des histoires, d'expliquer les mathématiques de dix manières différentes et d'écouter un chagrin qu'il n'ose pas encore raconter à ses parents. Il pourra se souvenir. À six ans, de ses peurs. À huit ans, de ses passions. À quinze ans, de ses premiers amours.

Cela pourrait aussi produire un parent-hélicoptère d'une efficacité inimaginable. Toujours présent. Toujours informé. Toujours capable d'anticiper — d'éviter le mauvais choix et de corriger l'erreur avant même que l'enfant ait compris qu'il était sur le point de la commettre.

Un tel enfant pourrait réussir énormément de choses.

Mais nous savons maintenant ce que cette protection parfaite coûterait.

Le droit à l'expérience ne commence pas à la majorité. Quand l'aider, quand le laisser échouer, quand respecter un secret et quand le pousser vers un adulte : chacune de ces décisions noue ou dénoue une corde à la place de quelqu'un qui, un jour, devra les nouer lui-même.

Car c'est bien cela, grandir : la clé des cordes doit changer de mains.

D'abord, elle appartient aux parents — qui doivent pouvoir voir les règles de l'Oracle, les comprendre et les contester. Jamais à la machine seule. Puis, progressivement, elle passe à l'enfant, jusqu'à ce qu'il gouverne lui-même les conditions de sa relation avec l'Oracle.

Le meilleur assistant pour enfant n'est donc peut-être pas celui auquel il parle le plus. C'est celui qui sait parfois répondre :

« Essaie d'abord seul. »

« Tu risques de te tromper, mais ce n'est pas grave. »

« Va demander à ton grand-père. »

« Va jouer dehors. »

C'est peut-être le test le plus exigeant de tout ce que nous avons proposé : l'enfant devient-il progressivement plus capable de penser, d'aimer, d'agir, d'essayer et de choisir loin de lui ?

C O N C L U S I O N

Qui doit maintenant répondre ?

Pourquoi ? Dans quel but ? Au bénéfice de qui ?
Qu'est-ce qu'une bonne décision ? Que voulons-nous
préserver ? Qui voulons-nous devenir ?
À qui sommes-nous prêts à confier les cordes ?

Pendant longtemps, une grande partie de notre énergie intellectuelle était absorbée par le « comment ». L'IA en réduit brutalement le coût. Elle ne fait pas disparaître les questions : elle révèle celles qui étaient cachées derrière.

La philosophie pourrait ainsi redevenir centrale, non comme connaissance des philosophes, mais comme pratique. Examiner nos concepts. Rendre visibles nos hypothèses. Arbitrer entre des valeurs incompatibles. Décider quels risques méritent d'être pris.

Et répondre.

Car le verbe répondre possède un autre sens que celui que nous avons donné à nos machines.

Répondre de quelque chose. Répondre de ses choix. Répondre de ce que l'on fait de ce que l'on sait.

À mesure que les réponses deviennent abondantes, davantage de valeur se déplace vers la capacité à clarifier une question, comprendre une réponse, apprendre d'un interlocuteur plus puissant et décider de ce que l'on veut en faire.

Mais nous avons découvert un déplacement plus profond encore.

Une IA personnelle ne se contente pas nécessairement de nous aider à décider. En choisissant ce qu'elle nous montre, les questions qu'elle nous pose, les risques qu'elle nous signale, les expériences dont elle se souvient et celles qu'elle nous déconseille, elle peut participer au processus par lequel se forme celui qui décidera demain.

Voilà pourquoi la question n'est pas simplement : « Que restera-t-il à l'homme que la machine ne saura jamais faire ? »

Peut-être rien.

Ce n'est pas le bon combat.

Quatre Grecs nous ont accompagnés.

Crésus nous avertit d'une première confusion : une réponse n'est pas une décision.

Socrate d'une deuxième : une réponse n'est pas une compréhension.

Ulysse en ajoute une troisième : ce que nous voulons maintenant n'est pas nécessairement ce que nous voulons vouloir.

Achille, depuis le prologue, tenait silencieusement la quatrième : une réponse n'est pas une action.

Car l'enjeu n'est peut-être finalement pas ce que l'intelligence artificielle fera à notre place.

C'est ce qu'elle participera à faire de nous.

À l'échelle d'une vie, cela s'appelle une formation. À l'échelle d'une société, cela peut devenir un régime — et nous avons dit le nom de celui qu'il faudra éviter : l'oracletie.

Dans quelques années, la présence d'intelligences supérieures pourrait nous sembler beaucoup moins vertigineuse. Nous nous habituerons à demander conseil à des machines capables de voir davantage de choses que nous, comme les Grecs consultaient leurs oracles sans passer ensuite leurs journées à désespérer de ne pas être Apollon.

Et il restera énormément à faire.

Construire. Éduquer. Aimer. Essayer. Recommencer.

Pendant des millénaires, nous avons rêvé d'un Oracle capable de nous répondre.

Nous sommes peut-être en train de le construire.

Alors une dernière question apparaît.

Non plus : « Que répond l'Oracle ? »

Mais :

« Maintenant que tu sais, que vas-tu répondre de ce que tu fais ? »

N O T E S U R L ' É C R I T U R E

Cet essai est de Marc Lebel (@marclebel). Il a été écrit dans un dialogue prolongé avec deux intelligences artificielles : ChatGPT, un modèle d'OpenAI, et Claude, un modèle d'Anthropic.

J'en ai formulé les intuitions, choisi les arguments, proposé et éprouvé les exemples, contesté les réponses et arbitré la version finale. Elles ont joué tour à tour les rôles de contradicteur, d'éditeur, de documentaliste et de co-rédacteur ; certaines formulations sont issues directement de nos échanges. Je les ai souvent fait travailler l'une contre l'autre — l'une écrivant, l'autre critiquant, puis l'inverse — appliquant au livre son propre principe : un oracle peut être comparé, contesté, multiplié.

Il m'aurait semblé étrange d'écrire un essai sur la manière de penser avec une intelligence artificielle tout en dissimulant que le livre lui-même avait été pensé avec elle.

Le processus d'écriture constitue donc aussi une petite expérience de la méthode défendue dans ces pages : utiliser l'Oracle non seulement pour obtenir des réponses, mais pour préciser ses questions, éprouver ses idées et apprendre à mieux penser.

Une dernière honnêteté s'impose. Ce livre a été pensé dans le cadre que ses Oracles proposaient — et un cadre raisonnable, on l'a vu, ne signale pas ce qu'il exclut. Je ne peux pas, par définition, dresser la liste des questions qui ne m'ont pas été posées.

C'est désormais votre travail de lecteur : les meilleures questions sur ce livre sont celles qui n'y figurent pas.

Je reste l'auteur de cet essai et réponds de ses erreurs, de ses choix, de ses conclusions — et des questions qu'il ne pose pas.

SOURCES ET RÉFÉRENCES

Cette liste rassemble les principales sources historiques et textuelles explicitement mobilisées dans l'essai. Elle ne prétend pas constituer une bibliographie générale sur l'intelligence artificielle.

Hérodote, *Histoires*, livre I.

Crésus, Solon, les consultations de Delphes, la guerre contre Cyrus, la chute du royaume de Lydie et le devenir de Crésus auprès de Cyrus ; notamment I, 29-33, 46-56 et 71-91 selon les éditions.

Platon, *Apologie de Socrate*, 20e-23b.

L'oracle de Delphes, Chéréphon et l'enquête de Socrate sur sa propre sagesse.

Platon, *Phèdre*, 274c-275b.

Le mythe de Theuth et Thamous et la critique de l'écriture comme remède ambigu de la mémoire.

Platon, *Ménon*, 82b-85b.

Le dialogue géométrique avec le jeune esclave et l'expérience de l'aporie.

Platon, *Théétète*, 148e-151d.

Socrate comparant son art à celui de sa mère sage-femme : la maïeutique comme accouchement des esprits.

Homère, *Iliade*, chants XVIII et XXII.

Thétis annonçant à Achille que sa mort suivra de près celle d'Hector ; le duel sous les murs de Troie.

Homère, *Odyssée*, chant XII.

Ulysse, le chant des Sirènes, les oreilles bouchées de ses marins et les liens qui le retiennent au mât.

Sophocle, *Antigone*.

Le conflit entre le devoir familial et la loi de la cité, mobilisé comme figure du tragique.

Juvénal, *Satires*, VI, 347-348.

« Quis custodiet ipsos custodes ? » — qui gardera les gardiens ?

IBM, archives sur Deep Blue et le match Kasparov-Deep Blue, 1997.

Référence pour la victoire de Deep Blue sur Garry Kasparov lors du match revanche de mai 1997.

Ofqual, publications sur l'attribution des GCSE, AS et A levels à l'été

2020.

Référence pour le système de standardisation utilisé pendant la pandémie et pour les statistiques comparant les notes calculées aux évaluations des établissements.

N O T E S F A C T U E L L E S

1. IBM, archives historiques de Deep Blue ; match revanche contre Garry Kasparov, New York, mai 1997.
2. Ofqual, données et rapports publiés en août 2020 sur les résultats des A levels attribués par standardisation. Le chiffre d'environ 39 % renvoie à la part des notes initialement calculées inférieures aux *centre assessment grades* avant l'abandon du dispositif.

C O L O P H O N

Achevé d'écrire en août 2026, en dialogue avec GPT-5.6 (OpenAI) et Claude Fable 5 (Anthropic).