

IA : l'extinction qui voulait éclipser le séisme

Société

Par Philippe Lemoine

Publié le 16 octobre 2026

Entrepreneur et essayiste, président du Forum d'Action Modernités

De l'innovation à l'apocalypse, il n'a fallu que quelques mois pour que l'intelligence artificielle générative soit présentée comme une menace pour la survie de l'humanité dans son ensemble. Mais les prophéties les plus sombres ne servent-elles pas ici à cacher les véritables enjeux de la mutation en cours ?

Le 9 septembre 2026, un chercheur, Jacob Coxon, démissionnait d'Open AI en expliquant que l'Intelligence Artificielle, telle qu'elle était développée, « pourrait tous nous tuer d'ici une décennie ». Dans la foulée, un de ses collègues déclarait que la probabilité que l'IA conduise à une extinction de l'humanité était supérieure à 10%. Le 23 septembre, devant le Conseil de sécurité de l'ONU, Sam Altman, le CEO d'Open AI, déclarait « Peu importe qu'on estime le risque

de catastrophe à 10%, 1%, 12%, ou 0,1% : aucun de ces niveaux n'est le moins du monde acceptable ». Que faut-il penser de ces déclarations ? L'Intelligence Artificielle peut-elle conduire à l'extinction de l'humanité ?

À peine posée, c'est comme si la question était à moitié oubliée. Il faut dire que, depuis quelques mois, on n'arrête pas de parler de l'IA. Toutes les semaines, il y a une nouvelle actualité sur laquelle chacun a son opinion. Depuis quelques jours, en France, on s'écharpe sur le point de savoir si le best-seller *C'était ça ou mourir* de Thélyson Orélien a été écrit ou non par une IA et, dans ce cas, si c'est vraiment grave. Les tribunes et les déclarations prolifèrent. Demain, ce sera autre chose. On dirait que cette discipline a le chic pour savoir faire parler d'elle. Dès les années 1970, le grand mathématicien, fondateur de la théorie des automates, Marco Schützenberger, disait en privé que « l'IA, c'est l'arrivée de la société du spectacle dans la science ». Sans cesse, dans le tourbillon qui a été créé, un sujet chasse l'autre. Mais, tout de même ! On nous annonce un risque d'extinction de l'humanité ? Cela vaut la peine de s'arrêter un moment pour savoir quoi en penser...

En 1668, Jean de la Fontaine publiait le premier volume de ses *Fables*. La troisième fable s'intitulait « La Grenouille qui se veut faire aussi grosse que le bœuf ». Elle raconte l'histoire d'une grenouille qui, ayant vu un bœuf, est envieuse de sa stature. Elle « s'étend, et s'enfle, et se travaille pour égaler l'animal en grosseur ». Le résultat ? Elle « s'enfla si bien qu'elle creva ». La morale ? « Le monde est plein de gens qui ne sont plus sages. / Tout bourgeois veut bâtir comme les grands seigneurs. / Tout petit Prince a des Ambassadeurs. / Tout Marquis veut avoir des Pages ». À la manière de La Fontaine, on pourrait écrire une fable de l'IA. Son titre serait « L'Extinction qui voulait éclipser le Séisme ». À l'époque actuelle, celle qui consacre ce que Jean-Pierre Dupuy a nommé « le catastrophisme éclairé », le paradoxe est que l'affirmation de la vraie puissance semble ne commencer qu'en côtoyant l'Apocalypse^❶. On ne devient plus,

en effet, un enjeu majeur par la simple accumulation des capitaux et des honneurs. Il faut, en plus, rivaliser avec les plus grandes menaces et figurer tout en haut des priorités du monde. C'est ce qu'illustre notre fable sur l'IA, en commentant tour à tour les trois termes : extinction, éclipse, séisme.

Extinction : des risques mortels à conjurer

Extinction tout d'abord : existe-t-il vraiment un risque que l'IA en rajoute par rapport aux perspectives de la sixième extinction de masse, cette extinction massive des espèces qu'anticipe l'écologie, entièrement due à l'activité humaine, à la pollution, au changement climatique, à la déstabilisation de tous les écosystèmes naturels ? Par une étrange pulsion suicidaire, à peine consacré maître abusif de l'Anthropocène, l'Homme aurait inventé l'outil qui conduirait à sa propre extinction. Le pronostic est effrayant, plus déstabilisant encore que celui de la singularité technologique, cette émergence prochaine d'une super-intelligence qui surpasserait celle des hommes. Lorsque Ray Kurzweil avait avancé cette hypothèse, en 2005, il parlait du risque d'être dépassé mais aussi de l'opportunité trans-humaniste qu'auraient les hommes de s'hybrider et de devenir des hommes augmentés. Nous n'en sommes plus là ! Le risque aujourd'hui serait celui de l'anéantissement, de la disparition, de l'extinction.

Quelle forme pourrait avoir ce scénario catastrophe ? Curieusement, les prophètes ne sont guère diserts sur ce plan. On peut imaginer deux choses. La première serait qu'une IA très puissante parvienne à franchir tous les garde-fous qui entourent les armes de destruction massive, comme les bombes atomiques, ou, pire encore, qu'elle pénètre dans un laboratoire et parvienne à bricoler une arme biologique rapidement contagieuse et massivement létale, dont nul ne parviendrait à décoder les gènes. Perspective terrifiante, en effet ! L'autre

scénario d'épouvante n'est pas celui d'une mort physique de tous les êtres humains, mais celui d'une disparition insidieuse de tout ce qui leur confère une humanité. Bien au-delà de l'économie de l'attention qui avait caractérisée l'ère des réseaux sociaux et des moteurs de recherche, les robots conversationnels s'imposeraient comme compagnons, comme confidents et, en définitive, comme maîtres de notre esprit. Soumis aux algorithmes et privés d'une autonomie de jugement, les corps humanoïdes seraient-ils encore des hommes ?

Ces deux risques existent, en effet. Mais jusqu'à quel point constituent-ils une rupture avec ce que nous connaissons déjà ? Avec les explosions d'Hiroshima et de Nagasaki, le monde est entré dans une ère nouvelle, celle où les hommes ont acquis le pouvoir de détruire la planète. Nous nous savons désormais menacés par la folie des Docteurs Folamour et des kamikazes atomiques anonymes. Depuis le XXe siècle également, nous avons appris que la propagande et la terreur pouvaient fonder des régimes totalitaires où l'on extermine à grande échelle des millions d'hommes, de femmes et d'enfants, sans que la barbarie suscite le rejet et la désobéissance de la plupart des humains. Le nazisme puis le stalinisme ont certes été vaincus ; la tentation de Big Brother, elle, elle n'a pas disparu. Nous sommes au siècle suivant, mais 1984 n'est pas mort.

Les risques majeurs de l'IA sont donc ceux d'une amplification importante des dangers qui pèsent d'ores et déjà sur nous. La vraie nouveauté, c'est qu'on paraît délaissé les deux armes qui permettent de faire face à ce genre de périls : le Droit et la Science. Le Droit pour responsabiliser les hommes dont le rôle est établi dans tous les dérapages ; et la Science pour évaluer, par avance, la dangerosité des outils. Mais ceux qui parlent le plus fort ne veulent ni de l'un ni de l'autre. Ce sont en effet les patrons des plus grosses sociétés d'IA qui alertent sur la possibilité que leurs produits mènent à une destruction de l'humanité. Et on les laisse dire ! Ils déclarent vouloir freiner la course mais ils perfectionnent sans cesse leurs modèles, en

espérant que leur attelage puisse les mener rapidement vers une introduction en Bourse fracassante. À juste titre, un journaliste a écrit que le responsable d'un laboratoire pharmaceutique qui continuerait à promouvoir un de ses produits, après avoir reconnu qu'il peut décimer la population, serait arrêté et jeté en prison^②. Pourquoi en va-t-il autrement avec l'IA ? On accepte peu à peu qu'il n'y ait pas d'autre règle que celles de l'autorégulation et pas d'autre compétence que celle de la recherche privée.

L'urgence est de rétablir la force du Droit et l'autorité de la Science. L'Europe pourrait montrer la voie avec un grand texte sur l'IA, assis sur quelques principes clairs et sur des sanctions dissuasives, et avec la création d'un « CERN de l'IA », un grand centre public européen, chargé de la recherche fondamentale sur l'IA de frontière et de l'évaluation des nouveaux systèmes qui sont mis sur le marché. Avec de telles mesures, nous serions à même de ramener les menaces d'apocalypse au niveau des risques que l'humanité a appris à gérer. Vivrions-nous pour autant dans un monde apaisé ? Pas vraiment. Il faut en effet prendre la mesure de ce que tout ce brouhaha sur l'Apocalypse tente d'éclipser.

Éclipse : une catastrophe peut en cacher une autre

Pourquoi les grands patrons de la Tech américaine en viennent-ils à sonner le branle-bas de combat et à tirer sur la sonnette d'alarme ? Parce qu'ils ont peur. Ce qui les terrorise ce n'est pas de se retrouver dans la situation d'un Docteur Frankenstein qui serait dépassé par la puissance incontrôlable de sa créature. C'est d'être perché au sommet d'un édifice immense qui menace de s'écrouler.

Au début de l'année, j'avais publié un article sur « le séisme qui vient ». Il rappelait que le numérique était actuellement plongé dans un cycle d'hyper-concentration du capital,

d'hypercentralisation des données et d'hyper-exploitation des ressources. Si l'on considère les cinq premières capitalisations mondiales, elles pèsent 17000 milliards de dollars aujourd'hui (contre 1400, au lendemain de la crise de 2008), elles sont toutes dans le numérique (contre 2) et toutes américaines (contre 3, à l'époque). Les cinq entreprises en question (Nvidia, Meta, Apple, Microsoft, Amazon) ont su s'enrichir, en captant le puissant mouvement de centralisation des données qu'a déclenché l'ouverture du Web, ce standard qui a permis à 6 milliards de personnes de se connecter à Internet sans formation et avec des outils très simples (les smartphones), mais au prix d'une terrible régression : le placage de la vieille architecture client-serveur, verticale et hiérarchisée, sur un réseau révolutionnaire qui favorisait une communication horizontale et pair-à-pair. Et ce cycle du numérique a impliqué une hyper-exploitation des ressources, qu'il s'agisse des ressources naturelles (eau, énergie et terres rares), des ressources humaines (taylorisation tertiaire et travailleurs du clic) ou des ressources informationnelles (non seulement les fake news mais les datas de seconde main, produites par les IA et peu-à-peu, majoritaires sur le Web).

Ce système est fragilisé par ses excès mêmes. On parle d'une « bulle de l'IA » mais la crise boursière à venir va bien au-delà de l'IA et implique l'ensemble du numérique. Les valorisations n'ont plus rien de rationnel et sont de plus en plus soutenues par des interdépendances circulaires où la surcapitalisation des uns permet d'investir dans les autres, en échange de commandes que les seconds s'engagent à passer... Quand cela va-t-il péter ? Nul ne le sait. Dans un an, dans six mois, la semaine prochaine ? Ce qui est sûr, c'est que ce ne sera pas une simple crise. Si nous avons parlé de séisme, c'est que c'est toute une logique économique qui va exploser en vol. L'idée de fonder la valeur sur la centralisation des données va buter sur l'appauvrissement de l'humus informationnel du Web. Comme un sol trop labouré et trop gorgé de pesticides qui

devient infertile, le monde du numérique centralisé devient celui des datas infertiles.

Pour sauver leur empire, les Dieux du numérique ont alors misé sur une carte qui leur paraissait pouvoir conforter leur assise : l'Intelligence Artificielle. Une technologie puissante, fascinante, et de surcroît à même de légitimer les positions acquises. Son développement par apprentissage profond dans un océan de données ne justifiait-il pas l'hypercentralisation des datas ? Les coûts faramineux de mise au point des LLM, des grands modèles de langage, ne donnaient-ils pas un sens à l'hyper-concentration du capital ? La compétition dans l'IA générative ne donnait-elle pas un sens à la course à l'appropriation des ressources et à la construction de data-centers toujours plus gigantesques ? Google dégaina le premier et parvint, dès 2014, à modéliser l'univers du langage dans les équations du calcul vectoriel. Puis, en 2015, Sam Altman fédéra des investisseurs, dont Elon Musk, autour de l'idée de contrer Google et de créer une compagnie qui se voulait non-profit, OpenAI. Mais, dès 2019, l'association se transforme en entreprise à but lucratif, accueillant Microsoft à son capital et enclenchant la guerre avec Musk. Des équipes quittent alors Google pour fonder Anthropic et reprendre derrière Dario Amodei le combat d'une IA plus éthique tandis qu'Amazon, Microsoft puis l'ancienne maison-mère, Google, en deviennent actionnaires. Tout un petit théâtre de haines, de jalousies et de participations croisées se tisse ainsi, sur l'Olympe du numérique, au nom de l'IA et de la perpétuation de l'hyper-capitalisme numérique.

Mais, patatras ! Ce qui devait être le sauveur se révèle brusquement comme l'accélérateur de la chute. Dans l'encyclique *Magnifica Humanitas*, le pape Léon XIV a comparé une certaine IA, l'IA des puissants, à une tour de Babel, lui opposant la possibilité d'une IA édifiée par en bas, à l'image de la muraille de Jérusalem, reconstruite par ses habitants^③. La comparaison est très juste. De même que la tour de Babel n'a pas résisté à la folie des grandeurs et s'est écroulée, de même

l'IA de la Big Tech parachève de fragiliser la montagne de capitaux et de datas, accumulés par les géants. Trois phénomènes mènent inexorablement au séisme. L'incapacité à rentabiliser les capitaux investis, tout d'abord : des sommes folles ont été engagées pour développer les IA de frontière, sans modèle économique suffisant et en diffusant des versions gratuites, vite utilisées par des milliards de personnes, mais avec un marché mondial de la publicité bien trop petit pour équilibrer une technologie ayant un spectre d'usage si puissant. La nature de l'IA par ailleurs qui est en elle-même une technologie décentralisatrice puisqu'elle vise à rendre accessible de nombreux pouvoirs, à commencer par le pouvoir de programmer que vont acquérir les 6 milliards d'internautes. Cette bascule vers la décentralisation sera accélérée par l'intérêt même de certains acteurs comme le plus puissant d'entre eux, Nvidia, qui fournit les puces de ces systèmes et qui poursuit l'objectif de transformer tous les PC du monde en une porte d'entrée vers l'IA.

Le troisième phénomène, c'est la progression rapide de l'IA chinoise qui est précisément fondée sur des systèmes beaucoup plus légers, mobilisant moins de capitaux et consommant moins de ressources et présentant surtout l'avantage d'être *open-source* ou *open-weight*, c'est à dire d'être moins opaque qu'un logiciel propriétaire. Très différente de celle des robots conversationnels, cette vision de l'IA est adaptée à l'optimisation du *manufacturing*, à la sophistication des machines, à l'autonomisation des drones et des automobiles. Elle répond plus facilement aux attentes des entreprises de tous les secteurs d'activité, même aux États-Unis où de plus en plus d'entreprises s'équipent d'une IA chinoise. Là où l'IA des géants bute sur la difficulté de trouver un marché à sa taille, cette IA plus modeste s'achemine vers la rentabilité. C'est dans ce contexte que les Goliaths font monter le son d'un risque d'Apocalypse. Menacés par le séisme qui vient, ces dirigeants oublient l'ultra-libéralisme pour se regrouper autour du Président Trump et attendent de lui des mesures

protectionnistes. Protectionnistes de l'humanité sur le long terme, peut-être, mais surtout, au plus vite, protectionnistes de leur trône qui s'affaisse.

Séisme : l'opportunité d'un autre numérique

Sommes-nous condamnés à l'alternative entre deux catastrophes ? L'horizon est moins fermé que cela. On peut même espérer qu'il s'ouvre à nouveau. La déjà longue histoire du numérique a vu se succéder des cycles, des périodes d'un quart de siècle environ, où les enjeux se déplacent et où on assiste à une redistribution des cartes. C'est ainsi qu'au cycle de l'ordinateur (1936-1960) a succédé le cycle de l'informatique de gestion (1960-1984) puis celui d'Internet et de l'informatisation de la société (1984-2008). Depuis cette date, nous sommes dans un quatrième cycle, celui de la Data. Or, ce qu'annoncent ces perspectives de séisme, c'est l'ouverture d'un nouveau cycle, pouvant être fondé sur d'autres principes que ceux de l'hyper-capitalisation, de l'hypercentralisation et de l'hyper-exploitation des ressources.

Par un retour du balancier, sommes-nous pour autant assurés de voir triompher à court terme, en 2032 peut-être, une période où le numérique se conjuguerait à nouveau avec décentralisation, émancipation et *empowerment* de tous ? Non, évidemment. Ce que l'histoire montre, c'est qu'un changement de logique suppose l'action conjointe de trois forces : une crise de l'ordre ancien ; une bifurcation possible autour d'une nouvelle technologie ; une volonté politique. Le cycle actuel de la Data était né de la crise financière de 2008 ; le nouveau cycle naîtra de l'effondrement des Goliaths que j'appelle le Grand Séisme. Le cycle actuel avait été conçu dans la matrice du Web et du smartphone ; le cycle futur naîtra de l'intelligence artificielle, bientôt dopée par l'ordinateur quantique. Pour qu'un virage ait lieu, il faut que s'adjoigne à ces forces en mouvement,

celle d'une volonté politique. Existe-t-il une chance de la voir émerger ?

Cette dernière partie va se jouer sur deux scènes. La première est géopolitique : l'espoir serait bien mince si l'on en restait à une histoire du numérique seulement dominée par les États-Unis et la Chine. Comment espérer d'eux un changement de logique par rapport au capitalisme de surveillance ou au totalitarisme idéologique ? Il faut absolument que se dessine une troisième voie. Pour cela, l'Europe ne manque pas d'atouts et elle les renforcerait par un grand texte sur l'IA et par l'institution d'un « CERN » qui deviendrait leader de la recherche publique sur l'IA. Mais il faut aller au-delà, soutenir l'inventivité des communautés open source, définir des communs publics, réunir des capitaux, encourager des entreprises. Tout cela réussira d'autant plus sûrement que l'Europe saura nouer de vraies alliances stratégiques. Une alliance avec l'Inde serait à rechercher en priorité : son programme « Indian Stack » est un vrai succès et entraîne dans son sillage nombre de pays du Sud. Face aux États-Unis et à la Chine, l'alliance du plus grand marché du monde et du pays le plus peuplé de la planète serait loin d'être un poids négligeable.

L'autre scène sur laquelle il faut miser, c'est celle des opinions publiques et de la société civile. Mais est-ce réaliste d'attendre d'elles qu'elles pensent quelque chose de l'avenir du numérique ? Dans le monde entier, les personnes ont mis peu de temps pour se connecter à Internet et encore moins de temps pour adopter massivement les outils de l'IA. Croit-on vraiment que cette vitesse témoignerait qu'elles sont bêtement manipulées et aliénées ? On peut parier au contraire sur le fait qu'elles ont su reconnaître ce que la technologie pourrait leur apporter et qu'elles sont majoritairement à même de s'interroger et de s'exprimer sur les enjeux d'avenir. À l'orée d'un nouveau cycle, les thèmes de débat ne manquent pas, à commencer par les risques sur l'emploi dans certaines professions et les risques de déséquilibre de certaines régions

où se concentrent les data centers. À l'inverse, des thèmes majeurs d'espoir méritent d'être connus et débattus, qu'il s'agisse non seulement de progrès médicaux ou scientifiques mais aussi d'un passage à l'échelle d'enjeux plus sociétaux comme ceux de l'économie circulaire, du dialogue intergénérationnel, de la société inclusive ou de la démocratie augmentée.

Plus importante encore qu'une bifurcation technologique, cette intervention de la société civile suppose que s'opère une vraie bascule idéologique. Comme l'a écrit Michaël Foessel, la référence permanente à l'Apocalypse est une fermeture du temps, alors même que la fonction même du politique est de rouvrir le temps^④. Au catastrophisme éclairé qui aime aujourd'hui tous les débats, devrait succéder un progressisme lucide et éclairé. Pas le progressisme aveugle du XIXe ou du XXe siècle, bien sûr ! Trop de violences et trop de souffrances nous ont appris qu'il fallait être lucides. Mais il nous faut apprendre à nous méfier des prophètes de malheur qui ne sont parfois que des grenouilles en train de gonfler leurs poumons... Pour La Fontaine, l'avenir ne repose ni sur une menace obscure, ni sur un trésor caché. Dans une de ses fables les plus connues, le laboureur mourant dit à ses enfants : « Travaillez, prenez de la peine / C'est le fonds qui manque le moins ». L'avenir repose sur l'action, le travail et la confiance dans ce qu'on peut construire.

Notes

- ① Jean-Pierre Dupuy, *Pour un catastrophisme éclairé*, Paris, Seuil, 2004.
- ② Damien Leloup, « IA : Le spectre du Terminator », *Le Monde*, 20-21 septembre 2026
- ③ Léon XIV, *Magnifica Humanitas. Sur la protection de la personne à l'ère de l'intelligence artificielle*, Cité du

Vatican, 15 mai 2026, publiée le 25 mai 2026.

- ④ Michaël Foessel, *Après la fin du monde : critique de la raison apocalyptique*, Seuil, 2023.