



DOCUMENT DE POSITION

## “Oser penser” l’éducation à la démocratie avec l’IA

Vanessa Nurock<sup>1,2,\*</sup>

<sup>1</sup>Centre de Recherches en Histoire des Idées, UniCa et <sup>2</sup>UNESCO EVA Chair

\*[vanessa.nurock@univ-cotedazur.fr](mailto:vanessa.nurock@univ-cotedazur.fr)

### Résumé

Cet article propose de prendre au sérieux l’affirmation de Yan Le Cun selon laquelle l’IA pourrait être la cause d’une “nouvelle phase des Lumières”. Je propose l’idée qu’une éducation à la démocratie avec l’IA est possible à deux conditions, qui constituent une caractéristique de base de l’éthique by design et sont les deux faces d’une même médaille, à savoir le déploiement et l’enrichissement plutôt que la restriction et la diminution de l’humanité, à un niveau tant individuel que générationnel et culturel. Premièrement, l’IA ne devrait jamais être conçue, développée ou utilisée pour décapiter l’humanité mais plutôt pour la préservation et l’extension des capacités humaines. Deuxièmement, l’IA ne devrait pas promouvoir une quelconque forme de techno-colonialisme mais plutôt viser à préserver et étendre la diversité culturelle et intellectuelle. A condition de remplir ces deux conditions importantes, nous pouvons alors penser à l’IA non tant en termes de risques inquiétants mais plutôt en termes d’un futur constructif qui permettrait de promouvoir l’adage des Lumières “Ose penser” et de permettre ainsi une éducation à la démocratie.

*Mots-clés: Ethique by Design, Intelligence Artificielle, Démocratie, Décapitation, Techno-colonialisme*



Figure 1. Social media post by Yann LeCun on the amplification of Human intelligence with AI.

### 1. Une nouvelle “phase des Lumières”?

“En amplifiant l’intelligence humaine, l’IA pourrait causer une nouvelle Renaissance, voire une nouvelle phase des Lumières” twittait Yan Le Cun le 1er avril 2023, en précisant qu’il ne s’agissait pas d’un poisson d’avril. Cette idée a été reprise dans les principaux médias français

en ces termes : "Avec l'aide de l'IA va être amplifiée l'intelligence et la créativité de tout un chacun. Cela peut conduire à un nouveau siècle des Lumières"<sup>1</sup>.

On peut s'étonner de cette formulation de Yan Le Cun qui, non seulement insiste sur la référence aux Lumières (ainsi que j'y reviendrai amplement par la suite), mais également utilise le terme d'amplification et non d'amélioration (*enhancement*) généralement utilisé dans le contexte de l'IA, et particulièrement apprécié par les transhumanistes. Bien sûr, il importe de ne pas surinterpréter la citation de Le Cun (ou de supposer ce qu'il veut dire) mais, dans le contexte de la référence aux Lumières, on peut sans doute suggérer sans prendre trop de risque que l'idée d'une amplification suppose une forme d'élargissement non aristocratique.

Cette référence récurrente aux Lumières pourrait sembler aller de soi tant il est évident que la "révolution" de l'IA porte sur les mêmes champs que la révolution des Lumières : l'*Encyclopédie*, œuvre collective, raisonnée, intégrant à la fois les sciences et les arts et des illustrations, s'appelle maintenant *Wikipedia*. Si Rousseau vivait aujourd'hui, il ne manquerait pas de remarquer que les Confessions prennent la forme de Facebook, Tik Tok ou d'Instagram, dans la constitution d'une nouvelle histoire de soi, que L'Emile est remplacé par le e-learning, ou encore par des baby-sitters robotiques comme vecteurs d'une nouvelle éducation, et que l'on parle même d'un Contrat Social 2.0, voire de donner le pouvoir politique à des Intelligences Artificielles (IA), censées être plus "impartiales" que les humains et donc plus aptes à gouverner, sans s'interroger sur la nécessité d'être capable de se mettre à la place de l'autre pour assumer la "délégation" d'un pouvoir politique.

On peut donc assurément considérer que cette "Révolution" de l'IA ressemble fort à la Révolution des Lumières sur au moins deux plans. D'un part, elle concerne à la fois les champs technologique, scientifique, social, éthique et politique. D'autre part, elle concerne – qu'elle prolonge ou, le plus souvent, remette en question – certaines questions clés de l'époque des Lumières, telles l'identité, l'autonomie, le sens moral ou encore le Contrat Social. Toutefois, il me semble que cette similarité constitue un reflet qui risque fort d'être inversé si l'on n'y prend garde et que le risque de cette déformation ou de cette inversion des Lumières constitue un bon repère pour élaborer des garde-fous permettant de construire des IA démocratiques et de faire au mieux pour que le souhait de Yann Le Cun que l'IA facilite une nouvelle phase des Lumières ne soit pas radicalement contredit, ce qui serait bien dommage. Le but de cet article est donc, en quelque sorte, de réfléchir avec les outils de la philosophie en général, et de la philosophie morale et politique en particulier, les conditions de possibilité visant justement à permettre de réaliser ce vœu.

Le risque de cette inversion repose notamment sur un fait assez simple, au moins en apparence. Alors que la Révolution des Lumières avait permis la diffusion d'idées philosophiques prônant une pensée émancipatrice de l'autonomie au sens propre du terme, c'est-à-dire permettant de choisir ses propres normes et d'obéir à des règles librement consenties, de construire une société plus égalitaire, pour ne pas dire démocratique – malgré d'indéniables soubresauts sanguinaires et inégalités voire injustices persistantes. Au contraire, la révolution actuelle s'appuie avant tout sur une idéologie techniciste qui reproduit ou reconstruit des formes de hiérarchies en même temps qu'elle s'appuie sur des formes de connexion susceptibles d'éclipser voire d'exclure ou d'annihiler les relations. Comme l'écrivait Henry Kissinger en 2018 dans un article de *The Atlantic* intitulé *How the Enlightenment ends*: "les Lumières commencèrent essentiellement avec de nouvelles intuitions philosophiques diffusées par une nouvelle technologie. Notre époque va dans la direction opposée. Elle a généré une technologie potentiellement dominatrice, à la recherche d'une philosophie capable de la guider." Bref, souligne Kissinger, c'est comme si l'on recherchait dans l'IA des indices permettant de construire un récit, une vision directrice et comme si les IA devaient nous fournir les principes philosophiques, éthiques et politiques. Pour le dire autrement, tout se passe comme si l'éthique et la politique étaient solubles dans l'ingénierie et l'informatique et comme si l'on pouvait faire l'économie des philosophes pour confier uniquement aux ingénieurs et aux informaticiens le soin de proposer de nouveaux principes directeurs dans les champs sociaux éthiques et politiques. Au contraire du diagnostic pessimiste de Kissinger, on pourrait suggérer qu'une collaboration engageant diverses compétences serait peut-être particulièrement bienvenue dans ce type d'entreprise. Enfin, on peut suggérer que l'esprit, sinon la lettre des Lumières doit avant tout se comprendre, comme le soulignait Emmanuel Kant par l'impératif "Ose penser" et que la question n'est pas seulement celle du risque d'une inversion des Lumières mais aussi et peut-être surtout celle de leur déploiement et de leur actualisation.

## Une expérimentation technologico-sociale

Cette dilution de l'éthique et du politique dans le technologique s'opère également depuis un certain nombre d'années par le biais de l'expérimentation en amplifiant les applications et la mise en œuvre d'outils intégrant des IA, sans toutefois toujours inclure dans cette mise en œuvre une réflexion sur les implications sociales, éthiques et politiques de cette dernière. Or, il importe de souligner que la manière dont l'IA a été intégrée dans nos sociétés procède d'une forme d'expérimentation non seulement technique mais aussi, et peut-être surtout sociale, éthique et politique.

Cette expérimentation sociale, éthique et politique est voilée par l'idée très commune selon laquelle dès lors qu'une technologie est disponible, l'on ne peut forcément, que s'en servir, ainsi que le suggère ce que certains appellent communément (et sans savoir le plus souvent que Dennis Gabor critiquait précisément ce type de position techniciste) la "Loi de Gabor". Cette soi-disant "loi de Gabor" affirmerait que "ce qui peut être fait techniquement le sera nécessairement" et fait écho à la formulation (critique) du philosophe Jacques Ellul (Ellul; 1954) selon laquelle : "tout ce qui est technique, sans distinction de bien et de mal, s'utilise forcément quand on l'a en main. Telle est la loi majeure de notre époque."

Sur le plan politique, cette maxime s'inscrit en droite ligne de la formulation néo-capitaliste classique selon laquelle "il n'y a pas le choix", qui a été parfaitement formulée par Margaret Thatcher et son fameux : "En un sens, nous devons le faire car il n'y a réellement pas d'alternative" ("So in a sense we do have to do it. Because There really Is No Alternative")<sup>2</sup>, parfois abrégé en TINA. Le propre de TINA est de contraindre le champ des possibles en (im)posant une forme de dilemme et en sélectionnant une solution parce qu'on n'a pas vraiment le choix plutôt que parce qu'on la trouve vraiment désirable.

Pourtant, si l'on en revient à ce qu'écrit précisément Dennis Gabor, le père de la supposée "loi de Gabor" (Gabor; 1963): "le futur ne peut pas être prédit mais des futurs peuvent être inventés". De manière intéressante, il critique la capacité d'invention de l'homme

1 Voir par exemple (Le Monde; 4/28/2023)

2 Cette phrase fut prononcée lors d'une conférence de presse que l'on peut retrouver ici : <https://www.margarethatcher.org/document/104389>

lorsqu'elle prend les dehors d'une destruction. Il affirme ainsi que nous ne devons pas adopter l'attitude béate d'une acceptation passive d'une technologie considérée comme inexorable puisqu'il nous faut inventer nos futurs possibles, mais il affirme également que ces futurs doivent être résolument constructifs.

L'idée selon laquelle il faudrait envisager les nouvelles technologies comme une expérimentation est assez répandue aujourd'hui, notamment, aux Pays-Bas. Par exemple, Van de Poel (2017), propose l'idée selon laquelle les technologies émergentes devraient être considérées comme une expérimentation sociale. Par ce terme, il entend une triple idée : premièrement, une expérimentation dans la société, deuxièmement, une expérimentation sur la société, qui oblige à en repenser certains cadres normatifs (par exemple légaux) et troisièmement, une expérimentation par la société. Il me semble que l'IA est une expérimentation sociale dans ces trois sens à la fois et que cette triple dimension éclaire bien le fait que l'IA remodèle nos sociétés de manière complexe. Or, ainsi que le souligne de manière très pertinente Van de Poel, le problème crucial ici est de déterminer sous quelles conditions cette triple expérimentation sociale peut (ou non) être moralement acceptable.

Pour ce faire, l'une des possibilités est de situer la réflexion éthique non seulement en aval de cette triple expérimentation sociale mais tout au long du processus de développement de l'IA en axant la réflexion non seulement sur les principes ou sur les conséquences. Ce type d'approche a été proposée depuis le tournant des années 2020 sous divers angles (voir par exemple (Nurock et al.; 2021; Lachlan and Craigon; 2021; Boger; 2022; Brey and Dainow; 2023)). Dans le peu d'espace dont je dispose pour cet article, je souhaiterais proposer deux critères liés et complémentaires de l'Éthique by Design déjà évoquées dans Nurock et al. (2021) qui me semblent essentiels pour permettre de manière éthiquement acceptable l'amplification de l'intelligence (et de la créativité) humaine par l'IA, que Yann Le Cun appelle de ses vœux comme permettant "une nouvelle phase des Lumières" : la question de la décapacitation (ou "*deskilling*" en anglais) d'une part et la question du pluralisme l'IA de l'autre. Ce faisant, je souhaite souligner à quel point ces deux questions, généralement évoquées de manière distinctes, doivent être également pensées ensemble afin de permettre cette "amplification" conçue comme une consolidation des compétences que nous jugeons importantes d'un point de vue éthique et politique.

## IA et décapacitation

La question de la décapacitation n'est pas nouvelle en philosophie. A certains égards, elle pourrait être rapprochée de certains processus liés à l'automatisation susceptibles d'aboutir à une perte de compétence intellectuelle et manuelle, d'un savoir-faire, qui ont été analysés dans une veine néo-marxiste dans les années 1970. Ce processus a également été dénoncé plus récemment dans le champ de l'IA sous le terme de "deskilling" ou décapacitation.

On peut certes suggérer que cette décapacitation n'est pas forcément une (grande) perte pour l'humanité lorsqu'elle qu'elle peut concerner des tâches aliénantes, voire débilitantes, qui sont justement automatisables parce que répétitives ou ne nécessitant pas de réflexion, et donc qui ne sont pas valorisées. On peut toutefois se demander, premièrement, qui décide de valoriser ou non certaines tâches et deuxièmement, si l'automatisation concerne toujours des tâches sans valeur, voire dévalorisantes.

Au contraire, il arrive parfois, en effet, que cette décapacitation concerne la perte d'une capacité de choix pédagogique et de réflexion avec l'existence de "package" pré-programmés, ainsi que le montre (Apple; 1994). D'autres fois, cette décapacitation peut être synonyme d'une mécanisation du travail de soin et de la perte d'une attention à certains symptômes, ou l'automatisation du diagnostic par exemple, chez les infirmières ou les médecins (Rinard; 1996; Davis; 2000). C'est pourquoi, il est légitime de se demander s'il ne faut pas prendre garde au fait que nous pourrions peut-être purement et simplement perdre certaines compétences que nous n'exerçons plus.

Si l'on considère que l'IA est de plus en plus communément aujourd'hui utilisée pour prendre des décisions morales ou politiques, parfois comme aide à la décision, par exemple en permettant de proposer une variété importante de scénarii à partir desquels les citoyens ou élus peuvent choisir un certain nombre d'options à débattre, mais parfois aussi comme décisionnaire, par exemple pour programmer ces "machines morales" que sont susceptibles de devenir les véhicules sans chauffeur, on peut légitimement se demander si cette décapacitation morale et politique ne risque pas de poser un sérieux problème non seulement pour nos démocraties mais aussi, et surtout, pour l'éducation à la démocratie.

La philosophie aristotélicienne suggère que certaines de nos compétences parmi les plus importantes, comme l'éthique par exemple, sont des dispositions acquises ou hexis. Se plaçant dans cette perspective, (Vallor; 2015) suggère que c'est l'imagination morale qui doit montrer la voie aux 'techniques de soi', entendues en un sens foucauldien tout autant qu'au sens des technologies contemporaines qui sont le médium de notre rapport actuel au monde, ainsi qu'elle le souligne en note. Elle affirme ainsi que "nous conservons la capacité d'imaginer et de développer des formes nouvelles et alternatives de ces relations, même si les pouvoirs actuels de notre imagination morale continuent à être modelés par nos relations présentes aux technologies. La trajectoire future des capacités morales humaines est conditionnée par la technologie mais non déterminée par elle."

Ainsi que le souligne (Jaffro; 2007) au sujet de l'éducation morale, dans une analyse que l'on peut étendre à l'éducation à la démocratie, "les pratiques de soi ne sont pas assimilables à des techniques, mais plutôt à des dispositions et des habitudes ou, par extension, à des techniques dont l'usage essentiel et naturel est de produire de telles dispositions et habitudes." Et il poursuit : "Toutefois, alors les pratiques de soi ne relèvent pas d'un second type de techniques, mais d'un genre de choses entièrement différent – que j'appelle la 'gymnastique'". Par cette comparaison éclairante entre notre capacité morale et le jogging, (Jaffro; 2007) souligne l'importance de s'interroger sur le type de pratiques de soi susceptible de conduire non pas à n'importe quelle transformation mais à une bonne transformation de soi – c'est-à-dire un perfectionnement appuyé sur une connaissance. Il propose de dégager deux éléments essentiels de cette gymnastique. Le premier est qu'il s'agit d'un exercice, le second est que cet exercice est celui d'une volonté appuyée par la raison, elle-même orientée par une conception morale. C'est pourquoi il conclut que : "La gymnastique ne se laisse pas réduire à une technologie."

Cette gymnastique, elle va donc plus loin que l'idée d'une simple disposition acquise, et, en tout cas, est fort distincte de la routine morale qui est la base de la morale automatisée. Elle suppose également une régularité, un soin, une attention (parfois quasi-inconsciente) aux "petites choses" comme au processus, qui se perd forcément lorsque l'on cesse de la pratiquer.

Plus encore, on peut penser cette décapacitation non seule au niveau, ontogénétique, de l'individu, mais aussi phylogénétique, de l'espèce. De la même manière que certaines langues se perdent si elles ne sont plus parlées et deviennent des langues mortes, soit parce qu'elles sont remplacées par d'autres, soit parce que les populations qui les parlaient ont disparu, on peut se demander si, après un certain

nombre de générations, les compétences transférées aux IA seraient encore transmises ou si elle risquerait de disparaître dans l'évolution. Le problème alors, ne serait pas seulement la perte d'un savoir-faire mais bien d'une capacité, d'une compétence et montre bien que l'éducation à la démocratie peut devenir un problème. Faut-il alors, dans un élan technophobe, proposer, comme le craint Yann Le Cun, de promouvoir un "nouvel obscurantisme" qui bloquerait craintivement le développement des IA ? Dans l'opposition entre d'un côté, cette nouvelle phase des Lumières et de l'autre, la dénonciation de cet obscurantisme, il semblerait bien que nous soyons coincés de manière binaire, entre Charybde et Scylla. Pourtant, on pourrait peut-être sortir de cette alternative en proposant une véritable amplification de notre intelligence non pas par l'IA mais avec l'IA. Selon la formule heureuse proposée par Daniel Andler<sup>3</sup>, peut-être serait-il pertinent de promouvoir une forme de "bilinguisme" qui permettrait d'allier une double compétence : celle à penser et agir avec l'IA et celle à penser et à agir sans IA ou en étant capable d'en varier les types d'utilisations afin de ne pas perdre de compétences. Ainsi, l'amplification de notre intelligence et de notre créativité, promues par Yan Le Cun ne s'accompagnerait pas de décapitation. Ce "bilinguisme" exercé dans les champs sociaux, éthiques politiques pourrait ainsi allier IA et éducation à la démocratie.

## IA et éducation au pluralisme

La question de la perte et du gain de compétence doit être envisagée non seulement du point de vue des individus, de l'espèce, ainsi qu'on vient de le voir, mais également du point de vue des populations, sociétés et cultures. En effet, si l'on souhaite se demander à quelles conditions IA et éducation à la démocratie peuvent et doivent aller de pair, alors il est nécessaire de prendre en compte le droit des personnes, quelles que soient leurs cultures d'appartenance ou générations, à décider d'elles-mêmes. Pour le dire autrement, la question de l'éducation à la démocratie conduit à porter au premier plan la question de la dimension inclusive de l'IA.

Alors même que l'IA est souvent présentée comme un "point de vue de nulle part", scientifique, objectif et neutre, ces caractéristiques sont loin d'aller de soi, ainsi qu'en témoignent les débats sur les biais en IA qui sont allés croissant après la publication du livre de (O'Neil; 2016). Plus encore, certains vont jusqu'à affirmer que ces technologies se développent sous fond de reproduction de relations de domination économique-socio-culturelles, et dénoncent le risque d'un retour aux empires coloniaux ou le développement de nouveaux empires coloniaux, liés au développement de l'Intelligence Artificielle. Soutenant cette perspective dans un article intitulé *Géopolitique de l'Intelligence Artificielle, le retour des Empires ?* (Miaillhe; 2018) affirmait déjà sans ambages que les États-Unis et la Chine vont dominer le monde grâce à l'IA, que l'Europe est menacée de "cyber-vassalisation" et l'Afrique de "cyber-colonisation". Dans un entretien daté de début janvier 2024 (UN News; 4.1.2024), Seydina Moussa Ndiaye, maître de conférences à l'Université Numérique Cheikh Hamidou Kane du Sénégal et membre du Comité Consultatif sur l'IA de l'ONU, considère, dans la même veine que le plus grand risque que fait actuellement courir l'IA au continent africain est celui de la colonisation. Ce danger est non seulement l'un des vecteurs de méfiance de certains publics et pays envers le développement de l'IA mais pose également une difficulté de fond pour l'éducation à la démocratie, qui ne saurait se développer sur fonds de pratique techno-colonialiste à l'égard de certains pays ou de certaines communautés mais nécessite plutôt une réelle inclusivité.

Cette thématique du techno-colonialisme se développe depuis le tournant des années 2020 autour de quatre grands points d'entrée celui de l'extraction des matériaux indispensables et des conséquences environnementales de l'IA (qui dépassent la dimension techno-coloniale); celui des "ouvrier.e.s du clic", celui des données et, *last but not least*, celle des visions du monde sous-tendant le développement de l'IA, qui touchent à une forme de techno-colonialisme culturel, social, éthique et politique. Alors même que l'imaginaire et le vocabulaire attachés à l'IA donnent l'impression d'une réalité immatérielle, "en ligne", dans le nuage du "cloud", la matérialité de la production de l'IA dévoile au contraire des pratiques que Couldry et (Mejias and Couldry; 2024) résumant par une formule choc : "explorer, expander, exploiter, exterminer" (*explore, expand, exploit, exterminate*).

Les trois premiers points ayant été assez largement développés ces dernières années (voir notamment Couldry et Mejias and Couldry (2024) pour une revue récente). Dans le cadre limité de cet article, je me contenterai donc de m'attarder sur le dernier point, sur lequel la philosophie peut sans doute apporter des éléments de réflexion particulièrement pertinents à mon sens. L'une des difficultés centrales si l'on souhaite véritablement que l'IA nous permette d'étendre notre intelligence et notre créativité, pour reprendre les termes de Yan Le Cun, est de chercher à l'inscrire dans un répertoire culturel ouvert qui permette de ne pas décapiter des langues, des sociétés et des cultures – si je peux me permettre l'expression, en complément de ce qui a été développé dans la section précédente. En d'autres termes, il faudrait ici aussi avoir une approche non seulement bilingue ou plurilingue, mais aussi pluriculturelle. Or, l'une des difficultés que l'on rencontre pour ce faire réside dans la manière dont le développement de l'IA comme la critique de ce développement reposent tous deux sur des concepts et des valeurs culturellement localisés et dominants, si bien qu'il est difficile de ne pas "parler" dans les termes dans lesquels s'inscrit cette vision du monde, y compris pour la critiquer. La notion de "regard codé", proposée par Buolamwini (?) en référence à la fois au "regard colonial" et au "regard male", permet de mettre en lumière une vision du monde et des rapports de force inscrits dans un axe de domination sous-jacents au pouvoir que constitue le code ou l'IA. La spécificité de ce regard codé est qu'il passe par l'artefact d'une machine à voir qui est aussi une machine à prescrire, que ce soit explicitement (par exemple lorsque l'IA est utilisée dans le champ juridique) ou implicitement (lorsque les moteurs de recherche indiquent une figure masculine lorsque l'on tape "médecin par exemple").

Or, ainsi que je l'ai suggéré dans (Nurock; 2024, in press), ce regard codé est, la plupart du temps, au moins en Europe d'où j'écris, un regard codé États-Unien. En effet, le risque d'une exportation du "regard États-unien, s'ajoute au regard codé, que ce soit dans les biais eux-mêmes ou dans solutions proposées pour s'y soustraire.

Par exemple, d'un côté, la notion de partage ou "sharing" en anglais, qui est valorisée à l'extrême dans la culture américaine, comme l'a bien montré (John; 2017), est posée comme une valeur en soi. Cette tendance culturelle est probablement renforcée par la culture des "rideaux ouverts", qui veut que si on n'a rien à cacher, alors il faut le montrer ainsi que par la culture de la surveillance post 2001 qui veut que si l'on n'a rien à cacher, alors l'on n'a rien à craindre, comme le souligne (Solove; 2011). Pour autant, ce "partage" n'est pas celui de l'open source, mais d'un consentement à l'abandon de nos données. Or, (Varon and Peña; 2021) entre autres, ont bien montré que la question du consentement doit être située dans un rapport de force d'autant plus délicat qu'il est binaire et que ne pas consentir revient à faire un choix de type tout ou rien et qu'il n'est donc pas si facile de choisir de s'exclure d'une société mondialisée où le "partage" des données est posé comme norme.

D'un autre côté, les analyses et outils théoriques proposés – y compris ceux que j'utilise moi-même – viennent souvent de la culture américaine, ou sont diffusées via la culture américaine, par exemple s'agissant des biais, notamment raciaux, qui s'inscrivent dans un cadre culturel et historique précis aux USA. Ainsi, lorsque l'on parle d'IA "centrée sur l'humain", pour la distinguer d'une IA "centrée sur la technique", on a souvent tendance à oublier que l'humain dont il est question est ethnocentrique, centrée (en tout cas en Europe), sur la culture américaine – et ailleurs, sans doute sur la culture chinoise. C'est pourquoi il est urgent, comme le proposent déjà certains auteurs et certaines autrices, voire certaines organisations comme l'UNESCO dans sa recommandation sur l'éthique de l'IA de 2021, d'ouvrir le répertoire culturel de l'IA afin d'être capable d'y intégrer un véritable pluralisme, ferment indispensable de l'éducation à la démocratie. Cette approche ouverte est également portée par le manifeste 'Indigenous Data Sovereignty and Governance' (?) de la *Global Indigenous Data Alliance* formalisé par Carroll, Cummins et Martinez (?) qui se résume par l'acronyme CARE (*Collective benefit, Authority to control, Responsibility, and Ethics*), en complément explicite avec l'acronyme FAIR autour des principes qui promeuvent l'accès ouvert et non seulement le "partage" : *Findable, Accessible, Interoperable, Reusable*. Toutefois, la question n'est pas seulement de promouvoir certains principes d'inclusion. Comme le montrent bien (Roche et al.; 1996), il ne s'agit pas uniquement d'intégrer des éléments culturels qui permettraient de diversifier les données proposées en entrée aux IA mais bien plutôt de repenser de manière structurelle comment nous envisageons le développement de l'IA. Par exemple, (Gwagwa et al.; 2022) proposent d'intégrer dans la réflexion sur l'IA la notion d'*Ubuntu*, qui exprime l'interrelation entre les individus ainsi qu'un sens de la communauté différent de l'approche du "partage" qui pourraient permettre de renouveler la manière dont nous appréhendons les connexions en les pensant davantage sous le régime de relations enchevêtrées.

"Sapere aude", ose penser. Telle est la formule lapidaire par laquelle le philosophe Emmanuel Kant résumait la pensée des Lumières. Pour que le développement de l'IA n'entrave pas mais, au contraire facilite l'éducation à la démocratie, il est nécessaire que celui-ci ne s'accompagne pas 'une décapacitation des individus, des générations et des cultures. Bien sûr, ce critère n'est pas le seul nécessaire pour ce faire mais il permet de penser l'IA non seulement sous le régime restrictif du risque, mais aussi sous le régime constructif d'une ouverture vers de nouveaux possible.

## Remerciements

Ce travail a bénéficié d'une aide du gouvernement français, gérée par l'Agence Nationale de la Recherche au titre du projet Investissements d'Avenir UCA J.E.D.I. portant la référence n° ANR-15-IDEX-01.

## Reference

- Apple, M. W. (1994). Computers and the deskilling of teaching, *Computer Professionals for Social Responsibility Newsletter* 12(2).  
URL: <http://cpsr.org/prevsite/publications/newsletters/issues/1994/Spring1994/apple.html/>
- Boger, J. (2022). Culture change, human-centered design, and ethical by design as transactional cornerstone concepts in the development of technology for supporting aging, *Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments* pp. 556–561.
- Brey, P. and Dainow, B. (2023). Ethics by design for artificial intelligence, *AI Ethics*.  
URL: (*advance online publication*) <https://doi.org/10.1007/s43681-023-00330-4>
- Davis, A. (2000). Outsourced radiology: will doctors be deskilled?, *British Medical Journal* 337: 542.
- Ellul, J. (1954). *La technique ou l'enjeu du siècle*, Paris, Librairie Armand/The Technological Society, (J. Wilkinson, trans.), Vintage Books.
- Gabor, D. (1963). 1964, *Inventing the Future*, New York: Knopf.
- Gwagwa, A., Kazim, E. and Hilliard, A. (2022). The role of the african value of ubuntu in global AI inclusion discourse: A normative ethics perspective, *Patterns* 3(4): 100462.
- Jaffro, L. (2007). L'exercice moral est-il assimilable à une technique ? Une critique de Michel Foucault, *Revue d'Auvergne* 585: 73–96.
- John, N. (2017). *The age of sharing*, Cambridge: Polity Press.
- Lachlan, D. U. and Craigon, P. J. (2021). The moral-it deck: a tool for ethics by design., *Journal of Responsible Innovation* 8(1): 94–126.
- Le Monde (4/28/2023). Yann Le Cun, directeur à Meta : « L'idée même de vouloir ralentir la recherche sur l'IA s'apparente à un nouvel obscurantisme ».  
URL: [https://www.lemonde.fr/idees/article/2023/04/28/yann-le-cun-directeur-a-meta-l-idee-meme-de-vouloir-ralentir-la-recherche-sur-l-ia-s-apparente-a-un-nouvel-obscurantisme\\_6171338\\_3232.html](https://www.lemonde.fr/idees/article/2023/04/28/yann-le-cun-directeur-a-meta-l-idee-meme-de-vouloir-ralentir-la-recherche-sur-l-ia-s-apparente-a-un-nouvel-obscurantisme_6171338_3232.html)
- Mejias, U. A. and Couldry, N. (2024). *Data Grab: The new Colonialism of Big Tech and how to fight back*, Penguin Random House, WH Allen.
- Miaïlhe, N. (2018). Géopolitique de l'intelligence artificielle, le retour des empires., *Politique Etrangère* 3: 105–117.
- Nurock, V. (2024, in press). *What Does 'Ethical by Design' Mean?*, Enjeux éthiques et métaéthiques de l'Intelligence Artificielle « Où va la philosophie française » ed. I. Alfandary, S. Laugier, R. Zagury-Orly, Paris, Vrin.
- Nurock, V., Chatila, R. and Parizeau, M.-H. (2021). *What Does 'Ethical by Design' Mean?*, In: *Reflections on Artificial Intelligence for Humanity* ed. By B. Braunschweig and M. Khallab, Springer.
- O'Neil, C. (2016). *Weapons of Mass Destruction: How Big Data Increases Inequality and Threatens Democracy*, New York: Crown.
- Rinard, R. (1996). Technology, deskilling and nurses: the impact of the technologically changing environment, *Advances in Nursing Science* 18 (4).
- Roche, C., Wall, P. and Lewis, D. (1996). Ethics and diversity in artificial intelligence policies, strategies and initiatives, *AI Ethics* pp. 1–21.
- Solove, D. J. (2011). *Nothing to hide: the false tradeoff between privacy and security*, New Haven, Yale University Press.
- UN News (4.1.2024). Interview: Ai expert warns of 'digital colonization' in Africa.  
URL: <https://news.un.org/en/story/2024/01/1144342>
- Vallor, S. (2015). Moral deskilling and upskilling in a new machine age: Reflections on the ambiguous future of character, *Philos. Technol.* 28: 107–124.
- Varon, J. and Peña, P. (2021). Artificial intelligence and consent: a feminist anti-colonial critique, *Internet Policy Review* 10(4).