

L'intelligence artificielle à l'épreuve de l'éthique : entre responsabilité et conviction

PRÉSENTATION

Cet essai s'intéresse à l'intelligence artificielle comme un outil émergent nécessitant encore d'être examiné à travers le prisme de l'éthique. Notre angle d'approche propose, en reprenant les concepts weberiens d'éthique de responsabilité et de conviction, de donner une vision relativiste de ce nouvel outil. Du point de vue éthique, notre posture est celle de la prudence et de la responsabilité, afin de préserver les valeurs, la responsabilité et la capacité de décision qui fondent l'action humaine individuelle et collective.

CONTENU

L'essai s'ouvre sur une citation désormais célèbre du président russe, qui souligne déjà la tension incarnée par ce domaine à l'échelle mondiale. Après avoir exposé l'étendue des champs dans lesquels l'IA se développe, nous posons la nécessité d'en faire un objet d'étude éthique et exposons le fil conducteur de notre essai : si l'IA présente des avantages certains, son usage ne peut être légitime qu'en vertu d'un cadre éthique strict.

Dans la première partie, et alors que nous examinons l'IA sous le prisme de l'éthique de responsabilité et exposons ses avantages, nous constatons que déjà ses limites s'imposent à nous.

Dans la seconde partie, nous passons l'IA au crible de l'éthique de conviction et mettons en exergue de nouvelles limites.

Nous concluons en reconnaissant que l'IA est l'affaire de tous et qu'elle nécessite de mêler conviction et responsabilité pour être pleinement éthique.

SOMMAIRE

I. Introduction

II. L'IA, un intérêt légitime mais relatif

1. Un outil puissant et indispensable
2. Un outil globalisé et polyvalent
3. Légitimité, éthique de responsabilité et IA

III. Les limites éthiques du recours à l'IA et son nécessaire encadrement

1. Le problème de la décision et de la responsabilité
2. Le problème de l'atteinte aux libertés publiques, aux droits fondamentaux et à la sécurité
3. Implications et conséquences d'un usage éthique de l'IA

IV. Conclusion

V. Bibliographie & sources

I. Introduction

« Celui qui maîtrisera l'intelligence artificielle sera le leader du monde ». Cette phrase prononcée par Vladimir Poutine le 1^{er} septembre 2017 souligne à elle seul l'intérêt vital que l'Intelligence Artificielle (IA) représente, et ce à toutes les échelles.

Ces dernières années, les progrès de l'informatique et le besoin d'efficacité ont conduit au développement de systèmes destinés à imiter l'intelligence humaine, tant dans sa puissance créatrice que dans ses capacités de réflexion et d'analyse. Ainsi, l'IA est née et continue d'être développée. Aujourd'hui, des outils conversationnels puissants tels que ChatGPT (OpenAI) sont désormais célèbres et permettent aux utilisateurs de dialoguer avec un interlocuteur de produire des réponses construites et argumentées. L'intelligence artificielle pénètre désormais des domaines aussi multiples que variés, allant de la finance à la médecine, en passant par l'art. Cette grande diversité de sujets entraîne l'apparition de modèles d'IA spécialisés, visant à décupler les capacités de l'homme, accroissant ainsi sa puissance de travail. Mais qu'ils soient généraux ou spécialisés, ces systèmes nécessitent d'être entraînés et nourris abondamment en données afin de pouvoir évoluer et s'améliorer. Dès lors, quelle pourrait être la meilleure façon de recourir aux IA et quelle position adopter à leur égard ? Seule l'éthique, en ce qu'elle constitue une réflexion sur la morale, c'est-à-dire l'application des valeurs socialement tenues pour bonnes, permet un éclairage pertinent sur cette question. Max Weber, célèbre sociologue du XIXe siècle, parlait d'éthique de responsabilité et d'éthique de conviction, l'une étant attachée au bien du résultat et de ses moyens, l'autre sur le bien des valeurs sur lesquelles elle repose.

Entre éthique de responsabilité et éthique de conviction, et malgré les avantages certains de l'IA, il est nécessaire d'en fixer les limites.

En effet, si le recours à l'IA se révèle comme légitime bien qu'il nécessite certaines réserves (I), ses limites éthiques imposent d'en encadrer l'usage (II).

II. L'IA, un intérêt légitime mais relatif

La globalisation de l'outil informatique et son extrême développement ont conduit à la création continue et à la mise en circulation de flux massifs de données. Pour nous aider à gérer et à exploiter ces gigantesques masses d'informations, le recours à l'IA s'est progressivement révélé comme évident et impératif.

1. Un outil puissant et indispensable

En effet, avec l'omniprésence de l'informatique et les possibilités d'internet, la quantité de flux d'informations aujourd'hui générés et mis en circulation est si importante qu'ont émergé les concepts de « tsunami numérique » et de « Big Data ». Ces deux concepts illustrent le caractère titanesque, irrésistible et inéluctable de ces vagues de données d'une part, et la menace d'un écrasement, d'une submersion par celle-ci d'autre part. Ainsi, pour éviter la submersion, l'assistance d'un outil capable d'analyser, de traiter et d'utiliser ces données se révèle indispensable.

Dans le secteur automobile par exemple, le développement de systèmes de conduite autonomes implique la génération de nombreuses données. Vitesse, météo, environnement, position dans l'espace, détection et localisation des éventuels obstacles... Mais les possibilités offertes par l'IA ne se limitent pas aux seuls domaines du recueil, du traitement et de l'analyse de données. Elles s'étendent également à la prise de décision. Reprenons l'exemple de l'automobile, le développement de systèmes de conduite autonomes impliquent nécessairement qu'ils soient en capacité de décider en tenant compte de l'intégralité des paramètres qu'ils captent et analysent, et ce en permanence. De plus, pour pouvoir générer et exploiter toutes ces données de la meilleure façon, il est nécessaire que les différents systèmes puissent communiquer ensemble et qu'ils soient de ce fait connectés entre eux à tout moment.

Ainsi, l'IA est un outil connecté, capable de collecter, de traiter, d'analyser d'importantes quantités d'informations mais également de prendre des décisions et de produire en conséquences. Par ailleurs, nous l'avons évoqué plus haut avec le concept de tsunami numérique, les masses de données à traiter peuvent s'avérer bien plus vastes lorsqu'elles relèvent d'un domaine moins spécialisé et donc démultiplier la complexité de la tâche, voire la rendre matériellement impossible, soit par manque de temps, soit par volatilité.

2. Un outil globalisé et polyvalent

Nous produisons et échangeons toujours plus de données. Qu'il s'agisse du domaine médical et de la transmission d'informations entre praticiens, du domaine judiciaire et de la procédure pénale numérique, ou encore de la sécurité avec la création de fichiers centraux nationaux et même internationaux, la quantité de données échangée de part le monde dépasse l'entendement. Par ailleurs, l'IA elle-même, par sa capacité de production, devient source de nouvelles données, constamment transmises, réutilisées, réanalysées.

De ce fait, dès lors que des données sont produites dans un domaine ou un secteur, l'IA est appelée à y trouver des applications. En outre, et la phrase du Président russe l'illustre parfaitement, dès qu'un acteur s'empare de cet outil, les autres ne peuvent que le suivre dans cette démarche s'ils ne veulent pas se voir écrasés par un concurrent aux capacités démultipliées. Catégorisation de la clientèle ou analyse des marchés ne sont que deux exemples qui déjà

permettent d'imaginer les gains de temps, de force de travail et de productivité permises par l'IA, ainsi que la réduction des coûts et l'augmentation des profits qu'elle induit.

La puissance et la variété infinie d'applications que nous continuons à trouver à l'IA permet aujourd'hui de la faire intervenir dans des domaines que nous pensions jadis propres à l'âme et à l'esprit humain tels que les arts graphiques ou encore la musique, et même la conversation.

Cette polyvalence a pour conséquence d'en faire un outil à l'interface de tous les savoirs et voué à se développer et à pénétrer tous les domaines. Ainsi, le recours à l'IA semble impérieux et donc, légitime.

3. Légitimité, éthique de responsabilité et IA

La légitimité peut se comprendre de deux manière distinctes. Elle peut désigner d'une part ce qui est fondé en droit, mais aussi ce qui est juste d'autre part. Le recours à l'IA doit donc être analysé sous ces deux aspects.

Max Weber disait de l'État qu'il a le monopole de la violence légitime. Une manière dire que le recours à la violence est fondé juridiquement et moralement dès lors qu'il est exercé par l'État. Il s'agirait là d'une forme de justification du recours à la violence par l'éthique de responsabilité. L'État assumerait de déroger aux règles qu'il défend pour protéger des valeurs supérieures. C'est ce que consacrent nombre de textes tels que la Convention Européenne des Droits de l'Homme et des Libertés Fondamentales en son article 2 lorsqu'il évoque le recours par les concepts d'un recours à la force rendu « absolument nécessaire ». Le droit positif français développe ces principes en y ajoutant le concept de proportionnalité dans le cadre des articles 122-5 et 122-7 du Code Pénal, traitant de l'irresponsabilité pénale en cas de légitime défense et d'état de nécessité.

En matière d'IA, la génération continuelle de flux toujours plus importants de donnée et le besoin de compétitivité induit par la mondialisation rendent le recours à l'IA nécessaire. Pour autant cette nécessité ne suffit pas à rendre son usage légitime. En effet, si l'IA est un outil puissant, elle n'est nécessaire que dans une logique consumériste. Depuis ses premiers pas, l'humanité a su se développer jusqu'à l'ère quantique sans recourir à l'IA. Celle-ci n'est que l'aboutissement d'un développement technologique visant à soulager l'humain d'un certain nombre de tâches, peut-être au prix de sa propre régression plus qu'à celui de son élévation. Il n'a jusqu'ici jamais été nécessaire de recourir à ChatGPT pour réfléchir ou se documenter. Beethoven n'a pas eu besoin d'une quelconque assistance numérique pour composer la V^e symphonie. De plus, le traitement de volumes toujours plus massifs engendre un coût écologique faramineux qui doit aussi être placé dans la balance de la légitimité. Le fait de pouvoir utiliser l'IA ne justifie pas nécessairement d'engendrer une dépense écologique. Ainsi, la légitimité de l'IA se révèle être en réalité, relative, et dépendantes des circonstances.

Nonobstant ces considérations, se pose la question des finalités de l'emploi de l'IA. En effet, les IA génératives permettent aujourd'hui de créer des « deep fake », documents falsifiés ou altérés, voire créés de toutes pièces, destinés à se servir sans droit de l'image d'une personne existante et même à tromper l'opinion publique. Elle peut également être utilisée pour manipuler des flux d'informations, s'introduire dans des systèmes informatiques et en altérer le fonctionnement ou en dérober les données et ainsi menacer des entreprises, la vie privée, déstabiliser des flux financiers, voire des États. Elles peuvent enfin simplement faire des erreurs sur la base de biais algorithmiques et générer des éléments inexacts, tels que des représentations

historiques erronées. Ces quelques exemples suffisent à mettre en perspective le bien fondé apparent des finalités de l'IA et à démontrer la nécessité de la mise en place de moyens de vérification et de contrôle de l'information, fussent-ils eux-mêmes recourir à cette technologie pour être efficaces au regard du raz-de-marée informationnel auquel le monde sera de plus en plus confronté.

En somme, si le recours aux intelligences artificielles peut s'avérer légitime au regard de la course technologique et d'un besoin de compétitivité, l'éthique de responsabilité ne suffit pas à le légitimer dans tous les domaines et pour toutes les applications qu'on pourrait lui trouver. Son utilisation débridée relève davantage d'une conception consumériste du monde et de ses ressources et pourrait se révéler irresponsable à certains égards. Malgré ses nombreux avantages, le recours à l'IA doit donc être maîtrisé pour se prémunir des effets néfastes d'un recours malveillant ou abusif à celle-ci. L'IA présente par ailleurs d'autres limites et nécessite d'être encadrée pour d'autres raisons.

III. Les limites éthiques du recours à l'IA et son nécessaire encadrement

Si l'éthique de responsabilité nous commanderait un usage raisonné et maîtrisé de l'IA et suffit à démontrer la nécessité de son contrôle, l'éthique de conviction nous permet d'examiner l'IA sous d'autres aspects qui lui sont intrinsèques.

1. Le problème de la décision et de la responsabilité

En effet, les IA présentent tout d'abord un problème évident lié à la prise de décision et à la responsabilité. Pour l'illustrer, reprenons une fois encore l'exemple de l'automobile. Si un véhicule autonome venait à entrer en collision, à qui la responsabilité de cet accident reviendrait-elle ? Le lancement sur les routes d'un tel véhicule nécessiterait un choix éthique sociétal. Nous pourrions décider que le conducteur, innocent de toute faute, serait tenu responsable des dégâts causés par son véhicule, comme nous le ferions pour un animal en divagation. Ceci poserait nécessairement problème en cas de collision impliquant deux véhicules autonomes. Nous pourrions aussi envisager, les assurances étant obligatoires et à supposer qu'elles l'acceptent, de dégager entièrement la responsabilité humaine et nous contenter de recourir aux sociétés d'assurances pour indemniser les éventuelles parties lésées. Dernière solution, nous pourrions rechercher systématiquement la responsabilité du constructeur, en retenant le manque de fiabilité du système. Ce dernier cas se révèle également problématique, car il exposerait injustement le constructeur aux cyberattaques.

Le sujet de la décision et de la responsabilité est d'autant plus complexe en matière de sécurité, de défense et d'armement. Plus particulièrement, ces dernières années ont vu l'apparition d'un nouveau domaine de l'armement : celui des Systèmes d'Armes Létales Autonomes (SALA). « La France ne développera pas de robots tueurs » ces mots prononcés le 5 avril 2019 par madame Florence Parly alors ministre des Armées sont sans équivoque sur la doctrine française en la matière. Toutefois d'autres pays embrassent une position doctrinale bien différente. La Russie, développe des drones antichars entièrement autonomes. Des mines intelligentes auraient également été utilisées sur le théâtre ukrainien, avec leur lot d'erreurs de ciblage et donc de civils tués... Pour la France, la décision de cet ultime recours que représente le fait de donner la mort ne peut et ne doit en aucun cas être automatisée, l'homme doit rester comptable de cette terrible prérogative pour qu'elle ne soit jamais banalisée.

Pourtant, à l'heure de la « faible empreinte au sol » ou « light footprint », un élément emprunté à la doctrine militaire états-unienne, le but est d'exposer toujours moins de personnel possible sur le terrain. Il est donc nécessaire d'envisager l'IA non pas comme un substitut mais comme un moyen d'augmentation du soldat, voire éventuellement d'aide à la décision. Cette position pourrait être étendue à tous les domaines impliquant un facteur décisionnel.

2. Le problème de l'atteinte aux libertés publiques, aux droits fondamentaux et à la sécurité

Outre les problématiques liées à la décision et à la responsabilité, les traitements de données présentent généralement des difficultés au regard des libertés publiques et des droits fondamentaux. Les traitements de données sont réglementés en France et en Europe. Comme tout traitement de données, les IA pourraient se contenter d'une subordination à ces réglementations. Néanmoins, le domaine des IA nécessite plus que la simple protection de la vie privée.

La sécurité permet à chacun de jouir de ses libertés sans nuire à l'exercice des libertés d'autrui. C'est dans un esprit sécuritaire que le projet « Bouclier Doré » a débuté en 1998. Depuis, la Chine s'est doté en 2014-2015, de plusieurs systèmes de surveillance de masse assistés par IA, notamment avec les systèmes Skynet et Sharp Eyes, et les crédits sociaux. Ces systèmes impliquent la catégorisation des citoyens selon leur comportement, une restriction d'accès à certains services tels que les transports publics, ou encore le recours à la reconnaissance faciale et donc l'enregistrement de données biométriques pour l'ensemble de la population. Ces principes contreviendraient gravement aux principes fondamentaux de nos états occidentaux, notamment l'égalité, la liberté de circulation ou encore le droit à la vie privée et à l'anonymat. De tels systèmes ne peuvent pas voir le jour en France en l'état actuel du droit et bien qu'ils présentent un avantage sécuritaire certain, ils constituent un déséquilibre manifeste entre liberté et sécurité.

Outre le domaine de la sécurité publique, la sécurité privée est également impactée par l'IA. En effet, dans le domaine de l'entreprise, l'IA peut-être utilisée comme aide à la conception. Toutefois, dans la perspective du développement d'un nouveau système, le recours à l'IA pourrait remettre en question l'attribution de la paternité d'une idée et même, représenter un risque sérieux de fuite de propriété industrielle ou intellectuelle vers l'étranger en cas d'utilisation d'un outil dont le fournisseur serait étranger. Pour palier ces éventualités, il serait nécessaire de développer des outils propriétaires ou garantis par l'État, hors réseau public et uniquement accessibles localement. Cependant ce type de systèmes se révélerait coûteux à développer et à mettre en œuvre, moins efficaces car moins nourris en données et sans garantie absolue contre la fuite de données.

Enfin, le domaine judiciaire est également impacté par l'IA. Le recours à l'IA apparaît comme une solution idéale pour examiner les dossiers, produire des synthèses, rédiger des actes et des courriers, effectuer des recherches documentaires, aider à la décision. Toutefois, les professionnels du droit alertent sur les limites et les risques associés à l'usage de l'IA. Premièrement, elle peut produire des « hallucinations », contenus vrais en apparence mais faux en réalité. Deuxièmement, elle présente des biais de conception et peut notamment conduire à des risques de discrimination. Troisièmement, elle entraîne un risque inhérent de confidentialité. Enfin, le recours à l'IA pose également un problème plus large dans le « rapport de force » judiciaire. En effet, à l'heure où les magistrats ne bénéficient pas d'outils suffisamment puissants, adaptés et sécurisés pour analyser les procédures et doivent s'acquitter manuellement de cette tâche chronophage, titanesque, des programmes d'IA permettant l'identification des vices de forme commencent à être déployés par les cabinets d'avocats. Ces systèmes permettraient de relaxer ou d'atténuer les poursuites pour des auteurs de faits pourtant accablants sur le fond, posant ainsi un

problème d'efficacité pour le système judiciaire et pouvant être considéré comme un risque pour la sécurité publique.

Ainsi, contrôle et usage raisonné restent les maîtres mots en matière d'IA et ces éléments démontrent à nouveau qu'elle ne peut en aucun cas se substituer à l'humain en matière de décision. Le recours à l'IA n'est pas nécessairement bon *per se*, ni *in fine* et nécessite d'entremêler et de concilier éthique de responsabilité avec éthique de conviction pour couvrir pleinement la vaste étendue de sujets à laquelle elle touche.

3. Implications et conséquences d'un usage éthique de l'IA

Il y a près de 200 ans, Clausewitz disait dans son traité De la Guerre : « Malheur au gouvernement qui affronte avec une politique de demi-mesures et un art de la guerre bridé un adversaire qui pareil aux éléments ne connaît d'autre loi que celle de sa propre force ! ». Bien qu'appliqué au domaine guerrier, il soulignait déjà les difficultés qu'implique le fait de s'imposer des limites face à un adversaire qui s'y refuse.

Rapportée à l'IA, cette phrase garde tout son sens. L'encadrement a un coût moral et stratégique. Celui qui se refuse à utiliser des SALA devra nécessairement exposer l'humain. Celui qui se refuse à recourir à l'IA dans son domaine, s'expose au risque de se voir distancer par la concurrence. Y recourir l'exposerait tout autant aux divers risques économique, technologiques et juridiques exposés plus haut. Ainsi, il paraît évident que les choix éthiques que nous ferons engageront les futurs professionnels.

En matière d'enseignement et de recherche, il s'agirait même d'une décision capitale dont dépendra le délitement de nos institutions, le déclassement de notre système scolaire, la baisse du niveau de nos futures élites. En effet, si l'IA peut se révéler comme un outil précieux en matière d'enseignement, un usage incontrôlé ou un recours trop systématique à son assistance risque diminuer notre capacité à penser de manière autonome, à conceptualiser, à l'abstraction et pourrait même supprimer notre capacité dépasser ce qui est déjà connu et donc à innover.

Enfin, au-delà de la balance bénéfiques/risques, les ressources naturelles et énergétiques nécessaires au déploiement des IA impliquent de s'interroger sur la balance bénéfiques/atteintes environnementales. Chacun est concerné, et à ce titre responsable, de son usage personnel de ces outils. Cette question dépasse le simple cadre professionnel ou stratégique et s'étend à tout citoyen. Ceci démontre que l'IA n'est pas une simple question d'experts mais relève de la responsabilité individuelle de chacun.

En somme, si les possibilités de l'IA en font un outil puissant pour atteindre des buts légitimes, les valeurs défendues par nos États de droit ne peuvent être sacrifiées sur l'autel de l'efficacité et de la raison d'État. Elles nécessitent d'être examinées à l'aune de leur proportionnalité vis-à-vis du but recherché, des maux à combattre. Cette question de proportionnalité s'étend à tous les domaines où l'IA peut être amenée à être mise en œuvre et nécessite d'être examinée à la fois à travers le prisme de l'éthique de responsabilité, et à travers celui de l'éthique de conviction.

IV. Conclusion

En définitive, s'il pouvait apparaître comme tout à fait légitime de recourir aux IA, la course technologique ne suffit pas à justifier à elle seule leur utilisation et la nécessité de concilier éthique de responsabilité et de conviction, d'accomplir nos objectifs en respectant nos principes moraux et ainsi de réguler leur utilisation apparaît clairement. Pour l'heure, la doctrine française est claire, notamment en matière de défense. Là où des pays comme la Russie ou la Chine cherchent à développer des systèmes d'arme ou de surveillance totalement automatisés, la France s'impose de conserver le facteur humain pour se réserver le choix et la décision, conformément aux convictions consacrées par ses textes fondamentaux.

Néanmoins, les IA dépassent largement le seul domaine des États et s'imposent à chacun dans des domaines de plus en plus nombreux. Dès lors chacun se doit d'adopter une réflexion approfondie sur ce sujet afin d'aborder le recours à l'IA de manière éclairée, responsable et citoyenne, pour en maîtriser les coûts, et les risques.

Dès lors, le réel défi des IA résidera dans la manière de se confronter à des adversaires auxquels leurs positions morales n'interdit rien ou en tout cas, permet davantage.

V. Bibliographie & sources :

Sources littéraires

- Carl von Clausewitz : « De la Guerre », 1832
- Max Weber : « Le Savant et le Politique », 1919

Sources numériques

- Vladimir Poutine : “le leader en intelligence artificielle dominera le monde”
Publié par [La Revue du Digital](#), 2 septembre 2017
<https://www.larevuedudigital.com/vladimir-poutine-le-leader-en-intelligence-artificielle-dominera-le-monde/>
- Déclaration de Mme Florence Parly, ministre des armées, sur l'intelligence artificielle et la défense, à Saclay le 5 avril 2019
<https://www.vie-publique.fr/discours/271295-florence-parly-5042019-intelligence-artificielle-et-defense>
- La France ne développera pas de "robots tueurs"- Discours de la Ministre des Armées sur l'IA, le 5 avril 2019
<https://cd-geneve.delegfrance.org/La-France-ne-developpera-pas-de-robots-tueurs-Discours-de-la-Ministre-des>
- Les systèmes d'armes létales autonomes (SALA)
<https://www.defense.gouv.fr/dems/systemes-darmes-letaes-autonomes-sala>
- La guerre « low-tech » de la Russie contre l'Ukraine a discrédité son récit de modernisation militaire
Anna Nadibaidze, 03 mars 2023
<https://lerubicon.org/la-guerre-low-tech-de-la-russie-contre-lukraine/>
- Traitement par l'intelligence artificielle des vices de procédure : rupture ou continuité ?
Billion, Arnaud ; Cassuto, Thomas, Les Cahiers de la Justice, 08 mars 2023
https://documentation.insp.gouv.fr/insp/doc/CAIRN/_b64_b2FpLWNhaXJuLmluZm8tQ0RMSl8yMzAxXzAxNDU%3D/traitement-par-l-intelligence-artificielle-des-vices-de-procedure-rupture-ou-continue?_lg=fr-FR
- L'intelligence artificielle générative pour les avocats : guide pratique
Julien Karachehayas, 09 janvier 2025 (mis à jour le 11 juillet 2025)
<https://www.francenum.gouv.fr/guides-et-conseils/intelligence-artificielle/comprendre-et-adopter-lia/lintelligence-artificielle-0>
- Grand Firewall de Chine
https://fr.wikipedia.org/wiki/Grand_Firewall_de_Chine
- « L'homme nouveau » de la Chine de Xi Jinping : surveillance et lavage des cerveaux
Pierre-Antoine Donnet, 24 Novembre 2025, (mis à jour 25 novembre 2025)
<https://asialyst.com/fr/2025/11/24/homme-nouveau-chine-xi-jinping-surveillance-lavage-cerveaux/>