



© Ipopba-AdobeStock

## B. Business Impact

# L'IA générative, entre menaces et opportunités : Quels défis éthiques pour les décideurs ?

ESCP Impact Paper No.2023-25-FR

Diana GARCIA QUEVEDO, Caroline VERZAT & Tristan DUPAS AMORY  
ESCP Business School

## **L'IA générative, entre menaces et opportunités : Quels défis éthiques pour les décideurs ?**

Diana Garcia Quevedo\*, Caroline Verzat\*\*,  
Tristan Dupas Amory\*

ESCP Business School

### **Abstract**

Faut-il avoir peur de l'IA générative ? Avec l'arrivée de ChatGPT et GPT4, l'IA générative - centrée sur la production de contenus textuels, visuels ou sonores à partir de données déjà existantes - bénéficie d'une couverture médiatique sans précédent. De la même manière que d'autres technologies à l'usage répandu, dont un exemple emblématique est celui d'Internet, sa généralisation entraîne toute une gamme de possibilités nouvelles et aussi potentiellement dangereuses. Les risques éthiques, écologiques et sociaux liés au volume des données et de calculs, à la création de contenus hyperréalistes ou de nouveaux *bullshit jobs*, sont réels. Pourtant, les véritables questions qui se posent, au-delà de la peur des uns et de l'enthousiasme des autres, ou même de la technique, pourraient être éthiques : Quels principes permettraient de distinguer un bon usage d'un mauvais usage de l'IA générative ? Il faut espérer que les décideurs amenés à concevoir, déployer ou contrôler l'usage des applications utilisant l'IA générative aient l'espace et le temps nécessaire pour s'interroger sur la contribution de cette technologie à trois types de finalités pas nécessairement compatibles : la performance économique de l'entreprise, le bien commun de la planète et/ou le développement des personnes. Cela nécessitera un travail important, non seulement technologique mais aussi réglementaire et organisationnel afin d'assurer son utilisation équitable, sûre et éthique.

Mots clés (keywords) : IA générative, risques, éthique

\*Ph.D. Candidates, ESCP Business School

\*\*Associate Professor, ESCP Business School

ESCP Impact Papers are in draft form. This paper is circulated for the purposes of comment and discussion only. Hence, it does not preclude simultaneous or subsequent publication elsewhere. ESCP Impact Papers are not refereed. The form and content of papers are the responsibility of individual authors. ESCP Business School does not bear any responsibility for views expressed in the articles. Copyright for the paper is held by the individual authors.

# L'IA générative, entre menaces et opportunités : Quels défis éthiques pour les décideurs ?

## L'effervescence face à la science

Avec l'arrivée de ChatGPT et GPT4, l'IA générative - centrée sur la production de contenus textuels, visuels ou sonores à partir de données déjà existantes - bénéficie d'une couverture médiatique sans précédent. L'effervescence autour de cette technologie traduit la peur limitante et l'enthousiasme qu'elle suscite dans une égale (dé)mesure.

Un véritable boom des applications utilisant l'IA générative est attendu en 2023<sup>1</sup>. Les domaines d'application possibles de l'IA inventoriés par le cabinet McKinsey en 2022 concernent aussi bien la gestion opérationnelle via des chatbots, la rédaction de contrats dans le domaine juridique, l'évaluation des CV dans le domaine des RH, le codage dans l'ingénierie, que la création de communications pour les employés<sup>2</sup>.

Dans ce contexte, faut-il avoir peur de l'IA générative ? Son usage a d'ores et déjà démontré une capacité à révolutionner les pratiques de création de contenu littéraire, vidéo, audio... mais aussi à générer des fortes controverses sur les usages qui pourraient en être faits à l'avenir. En un mot, les IA renferment une capacité colossale vis-à-vis de laquelle une armature éthique apparaît plus que jamais impérative. L'éthique peut se définir à partir de Ricoeur comme la recherche de « vivre bien avec et pour les autres dans des institutions justes<sup>3</sup> ». La véritable question qui se pose, au-delà de la justesse de la peur des uns et de l'enthousiasme des autres, peut donc être la suivante : Face aux menaces et opportunités qui s'annoncent pour l'avenir, quelle démarche éthique permettrait de distinguer un bon usage d'un mauvais usage de l'IA ?

Les controverses sont d'ores et déjà là : Alors que certaines institutions éducatives renoncent, par exemple, à proposer des devoirs à la maison de crainte que les étudiants ne les fassent rédiger par ChatGPT, d'autres proposent, au lieu d'interdire son utilisation, des moyens d'intégrer l'IA générative dans leur curriculum, car ils considèrent la technologie comme un catalyseur de l'apprentissage<sup>4</sup>. Un autre exemple troublant est celui d'un prix artistique décerné en 2022 à une œuvre créée avec le support d'une intelligence artificielle, générant une virulente controverse sur le statut d'artiste de son auteur.<sup>5</sup> Au-delà de ces polémiques récentes, et plus prosaïquement dans les entreprises, les usages marketing de l'IA générative permettant de personnaliser les contenus en fonction des historiques d'achat, de l'expérience client ou de son profil – y compris facial – dans le but de stimuler l'achat ou de fidéliser les clients après l'achat, sont en réalité déjà très avancés.

De la même manière que d'autres technologies à l'usage répandu, l'IA générative est une technologie majeure dont l'apparition implique toute une gamme de possibilités nouvelles et aussi potentiellement dangereuses. Un précédent emblématique est celui d'Internet, lequel a suscité un grand enthousiasme à la fin des années 1990. Il a permis la création de

---

<sup>1</sup> L'article en ligne du 21 mars 2023 de C. Romei, note en conclusion : « Selon un récent rapport de Vision Mobile, l'utilisation de l'IA dans les applications mobiles devrait augmenter considérablement au cours des trois prochaines années. Actuellement, seulement 4 % environ des applications utilisent l'IA, mais cela devrait atteindre 21 % d'ici 2022. » <https://www.servicesmobiles.fr/nous-sommes-entres-dans-lere-de-lia-generative-86185>

<sup>2</sup> <https://www.ictjournal.ch/etudes/2023-02-13/ia-generative-les-cas-dusages-possibles-selon-mckinsey>

<sup>3</sup> Ricoeur, P. (1990) *Soi-même comme un autre*, Paris, Seuil.

<sup>4</sup> Baidoo-Anu, David, and Leticia Anshah. "Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning," March 21, 2023.

<sup>5</sup> <https://www.lesnumeriques.com/vie-du-net/une-oeuvre-creee-par-une-ia-remporte-un-concours-d-art-aux-etats-unis-n191179.html>

nouveaux modèles économiques et démultiplié l'accès à l'information et la communication à un rythme sans précédent dans l'histoire de l'humanité. En ont découlé d'importants bénéfices pour la société comme l'accès généralisé au savoir, la facilitation des démarches administratives, les commandes et réservations en ligne, l'optimisation des flux logistiques et des voyages, ou encore la possibilité de se réunir à distance. Mais de nombreux risques, aussi divers que ses bénéfices, sont également apparus : l'« infobésité », le risque d'addiction à Internet avec un fort impact négatif sur la santé mentale, la cybercriminalité, la surexploitation commerciale des données personnelles, l'utilisation des données privées à des fins électorales ou encore la polarisation des réseaux sociaux dégradant la qualité des débats démocratiques. Après ses débuts où Internet était utilisé par un groupe d'individus restreint, cette technologie est désormais répandue dans notre vie quotidienne. Il y a, aujourd'hui encore, des mesures et des débats sur les réglementations liées à la protection des données et la cybersurveillance pour tenter de limiter les problèmes liés par exemple à la cybercriminalité<sup>6</sup>.

## **Derrière les polémiques, la technique**

L'IA générative est issue de recherches de différentes disciplines scientifiques comme la linguistique, l'informatique et les mathématiques. Des chercheurs ont travaillé de concert pour créer des modèles capables d'imiter les capacités humaines. À la fin des années 50, le pionnier de l'IA, Arthur Samuel, a développé l'idée que les ordinateurs pouvaient apprendre par eux-mêmes sans être explicitement programmés. Ce concept, connu sous le nom d'auto-apprentissage, est à la base du *Machine Learning* et de l'IA<sup>7</sup>. Au fur et à mesure que le domaine du *Machine Learning* évoluait, de nouveaux concepts tels que les modèles de « réseaux de neurones » ont émergé. L'extension de ces modèles a conduit au développement de techniques de *Deep Learning* marquant une autre étape importante dans le domaine.

Les dernières architectures de *Deep Learning* ont permis la création de grands modèles de langage – appelés *Large Language Models* en anglais (LLM) tels que GPT4. Les LLM sont capables de créer des textes, des images et des vidéos réalistes. Bien que certains puissent attribuer une intelligence à ces systèmes, il est important de rappeler que leurs résultats restent fondamentalement le produit de calculs statistiques et de prédictions<sup>8</sup>. Autrement dit, les LLM sont entraînés à utiliser des données massives afin de prédire la réponse la plus probable par rapport à une tâche spécifique. Dans le cas de ChatGPT, le modèle reçoit un *input* (une question) et prédira et sélectionnera la combinaison de mots et de phrases la plus probable pour répondre à cet *input* générant ainsi un *output* (la réponse). C'est ce fonctionnement probabilistique qui fait que ChatGPT peut produire de fausses réponses avec le même « ton » que les réponses correctes.

## **Impacts éthique, écologique et *bullshit jobs***

Comme toute innovation, ce n'est pas uniquement la maîtrise technologique qui pose question mais aussi son usage et les controverses qu'il suscite. C'est aussi le cas pour l'IA générative, dont les algorithmes sont développés par des experts depuis plusieurs années

---

<sup>6</sup> Naughton, John. "The Evolution of the Internet: From Military Experiment to General Purpose Technology." *Journal of Cyber Policy* 1, no. 1 (2016): 5–28.

<sup>7</sup> MIT Sloan. "Machine Learning, Explained." Accessed December 15, 2022. <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>.

<sup>8</sup> Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. "Power and Prediction: The Disruptive Economics of Artificial Intelligence." Harvard Business Review Press (2022).

mais dont la diffusion récente engendre des incertitudes majeures sur ses usages et leurs conséquences.

Ces inquiétudes ne se retrouvent pas seulement chez les néophytes : un groupe de personnalités éminentes de la communauté de l'IA a récemment fait part de ses inquiétudes quant à la croissance rapide des applications de l'IA et de ses effets négatifs involontaires sur la société<sup>9</sup>. À l'instar des technologies précédentes qui se déploient massivement et rapidement, il y a un risque d'oublier les risques associés à l'usage de l'IA générative. La synthèse de Koller (2023) qui s'appuie sur plusieurs rapports et articles scientifiques<sup>10</sup> alerte sur la sous-estimation de certains d'entre eux.

**Les premiers risques sont liés au volume des données et de calculs.** Ce volume exige des ressources financières conséquentes pour développer, déployer mais aussi évaluer l'impact et assurer une régulation éthique et durable des usages de l'IA générative. On pense à l'avantage compétitif des géants du numérique qui peut s'avérer problématique. Le volume des données et des calculs a aussi un impact écologique avéré, qui semble difficile à réduire. Par ailleurs, la nature du fonctionnement des IA entraînées sur des vastes volumes de données conduit naturellement à reproduire les biais et stéréotypes véhiculés dans les textes, et probablement à les amplifier si les IA sont alimentées par des textes générés par d'autres IA. En outre, les modèles d'IA peuvent dévoiler ou inférer des informations privées ou sensibles, ou encore reproduire des contenus protégés par le droit d'auteur.

**La deuxième série de risques est liée au fait que les IA produisent des contenus si convaincants que des lecteurs humains y croient quand bien même les textes contiennent des erreurs.** Ceci ouvre la porte à toutes sortes de *nudges*, malveillances et manipulations qui peuvent être extrêmement graves à grande échelle, d'autant plus qu'elles sont imprévisibles. Koller pointe du doigt cette phrase inquiétante du *Stanford AI Report 2021* : «*Les utilisateurs de demain découvriront sans doute davantage d'applications, bonnes et mauvaises, de sorte qu'il est difficile d'identifier la gamme d'utilisations possibles et de prédire leur impact sur la société*».

**Enfin, le dernier risque est de transformer de nombreux métiers en potentiels nouveaux *bullshit jobs*,** selon l'expression de l'anthropologue critique David Graeber (Graeber, 2018)<sup>11</sup>. Certains métiers dont la valeur perçue s'appuyait sur la création de contenu ou le contact humain se verraient recentrer vers des tâches plus rébarbatives de surveillance ou de validation de contenus générés automatiquement par l'IA.

## **Une technologie pour quelles finalités ? Propositions pour un discernement éthique de la part des décideurs**

Il faut espérer que les décideurs amenés à concevoir, déployer ou contrôler l'usage des applications utilisant l'IA générative aient l'espace et le temps nécessaire pour s'interroger notamment sur la contribution de cette technologie à trois types de finalités qui ne sont pas nécessairement compatibles : la performance économique de l'entreprise, le bien commun de la planète et/ou le développement des personnes quel que soit leur statut dans l'utilisation des applications

---

<sup>9</sup> Deepa Seetharaman – Elon Musk, Other AI Experts Call for Pause in Technology's Development – Wall Street Journal – March 29, 2023 <https://www.wsj.com/articles/elon-musk-other-ai-bigwigs-call-for-pause-in-technologys-development-56327f>

<sup>10</sup> <https://www.ictjournal.ch/articles/2022-04-29/modeles-de-langage-geants-des-risques-a-la-hauteur-des-capacites>

<sup>11</sup> Graeber, D. (2018) *Bullshit jobs: a theory*. Penguin.

Le discernement d'ordre éthique n'a rien d'inné, il doit s'entraîner par la réflexion critique et la confrontation au réel à travers l'action. Par essence, ce discernement n'est pas à la portée de l'IA qui n'agit pas, n'a pas d'émotions, ne considère pas les contextes, reproduit des contenus y compris leurs erreurs sans aucune conscience de leurs conséquences. Il est donc impératif d'aider les managers à discerner les conditions d'usage et les conséquences potentielles de ce nouvel outil. Les questions à se poser pourraient être les suivantes : À qui et à quoi les décideurs imaginent-ils que l'application qu'ils envisagent de concevoir, déployer ou distribuer pourrait servir dès aujourd'hui ou à l'avenir ? Les applications potentielles (ou déjà avérées) contribuent-elles à mieux vivre ensemble ? Comme dans toute réflexion éthique, il importe de construire une représentation complète des parties prenantes d'ores et déjà impliquées ou potentiellement impactées, dans l'entreprise et plus largement dans le monde<sup>12</sup> : à quelles externalités négatives (émission de CO2, perte de confidentialité des données, manipulation d'opinions, *bullshit jobs*, etc.) l'usage d'une application à base d'IA générative contribue-t-il ? Est-ce que cet outil renforce les inégalités ou bien conduit à davantage de justice ?

Les mesures d'impact chiffrées sont indispensables mais il ne faut pas négliger leurs limites pour la réflexion éthique. En effet, elles sont assises sur des indicateurs normatifs qui demandent discussion et leur essence immatérielle et impersonnelle peut engendrer une perte de vue des enjeux humains de la part des décideurs. Il est donc essentiel pour ces derniers de rencontrer différents groupes de personnes impactées, d'écouter leur vécu, leurs enjeux et leur définition de ce qui est juste afin d'identifier les tensions et questions vives reliées à l'usage d'un nouvel outil. Rencontrer par exemple des journalistes ou des codeurs dont les tâches ont profondément évolué avec l'usage de ChatGPT permettrait aux décideurs de prendre la mesure des changements positifs ou négatifs des contenus de métier liés à l'usage de l'IA générative. Puis de replacer ces changements au sein d'un système plus large d'interactions entre des donneurs d'ordre, des clients ou des régulateurs plus ou moins conscients des enjeux humains. En s'inspirant des démarches d'analyse de controverses ou d'enquête sur les questions socialement vives<sup>13</sup> qui éduquent l'esprit critique de nos étudiants, on pourrait suggérer aux décideurs de commencer par cartographier les enjeux des parties prenantes élargies avant de prendre une décision sur une IA générative. Ceci leur permettrait à minima d'étayer leurs convictions éthiques au-delà de croyances implicites dans le progrès. On peut supposer que cela permettrait aussi de limiter les risques de dissonance éthique entre valeurs personnelles ou sociétales et choix économiques dictés par la logique du profit à tout prix<sup>14</sup>.

La projection des usages futurs n'est pas aisée car elle nécessite de l'imagination et des raisonnements probabilistes. Au niveau individuel, une question éthique simple et efficace pourrait être la suivante : est-ce que je donnerais cet outil à mon enfant ? En ferait-il bon usage si cela lui évitait l'effort d'apprendre ou de créer par lui-même ? On pourrait se souvenir à cet égard que des leaders repentis des GAFAM ont reconnu dans un documentaire glaçant qu'ils limitaient fortement l'usage des réseaux sociaux par leurs enfants alors même qu'ils en organisaient la promotion marketing planétaire en

---

<sup>12</sup> Les risques environnementaux et sociaux de l'IA générative dépassent largement les parties prenantes classiques de l'entreprise. Ceci invite à souscrire à la vision de l'entreprise de Bommier et Renouard, 2018, définie comme " une action collective en vue de la production, de la répartition et de l'usage des biens au service du lien social et écologique » (p 24-25). Ce qui implique une responsabilité économique et financière, mais aussi sociale envers les employés, sociétale et environnementale et enfin politique.

<sup>13</sup> Bruno Latour "(2007) La cartographie des controverses", *Technology Review*, N. 0, pp. 82-83, Jean Simonneaux éd., *La démarche d'enquête. Une contribution à la didactique des questions socialement vives*. Educagri éditions, 2018.

<sup>14</sup> Cherré, B., Laarraf, Z. & Yanat, Z. (2014). Dissonance éthique : forme de souffrance par la perte de sens au travail. *Recherches en Sciences de Gestion*, 100, 143-172.

manipulant les clients attirés par la gratuité mais transformés en produits par le modèle économique reposant sur la publicité<sup>15</sup>.

Au niveau organisationnel, nous savons que la décision est un processus complexe, fortement sujet aux erreurs qui peut conduire à l'évitement des risques et des controverses<sup>16</sup>. Si l'on veut promouvoir une prise de décision éthique, il serait essentiel de mettre en œuvre des procédures systématiques et régulières d'évaluation des différents types de risques financiers, écologiques, sociaux ou politiques listés plus haut et d'instaurer des cercles de parole ouverts où la pensée critique est systématiquement convoquée.

Au niveau institutionnel, une réglementation serait hautement souhaitable. Mais par définition, celle-ci ne pourra s'élaborer qu'à partir des usages déjà avérés. Or cela demande du temps et du recul à la fois pour constater ces usages et pour négocier des règles acceptables pour tous les acteurs. En attendant, en l'absence de codification établie sur les bons ou les mauvais usages de l'IA générative, il semble crucial de donner aux décideurs des espaces permettant d'examiner sereinement et en profondeur les conflits et tensions de valeurs ou d'enjeux auxquels ils sont confrontés. Ceci pourrait prendre la forme de cercles de co-développement<sup>17</sup> dont la procédure permet de questionner systématiquement les différentes facettes d'un problème complexe en situation. Les différents tours de clarification, de questionnement et de conseil par les pairs constitueraient un précieux atout pour un décideur afin de prendre une décision consciente des enjeux et réfléchie sur le plan éthique.

En synthèse, l'IA générative est une technologie dont l'usage a encore besoin de temps pour mûrir. Elle nécessite un travail important, non seulement technologique mais aussi réglementaire, organisationnel et inter-individuel afin d'assurer son utilisation équitable, sûre et éthique. C'est à l'interface de l'individuel et du collectif que la question éthique doit se poser, et elle doit s'inscrire dans la mutation au long cours de notre organisation économique, politique et sociale.

## Références

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. "Power and Prediction: The Disruptive Economics of Artificial Intelligence." *Harvard Business Review Press* (2022).

Baidoo-Anu, David, and Leticia Ansah. "Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning," March 21, 2023.

Bommier S., Renouard C. (2018), *L'entreprise comme commun. Au-delà de la RSE*, Paris, Charles Léopold Mayer

Deepa Seetharaman – Elon Musk, Other AI Experts Call for Pause in Technology's Development – Wall Street Journal – March 29, 2023 <https://www.wsj.com/articles/elon-musk-other-ai-bigwigs-call-for-pause-in-technologys-development-56327f>

MIT Sloan. "Machine Learning, Explained." Accessed December 15, 2022. <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>.

---

<sup>15</sup> *The social dilemma* de Jeff Orlowski (2020).

<sup>16</sup> Voir les deux célèbres livres de Christian Morel : *Les décisions absurdes, sociologie des erreurs radicales et persistantes* (2002), *Les décisions absurdes, comment les éviter* (2012).

<sup>17</sup> Delaunay, D. et Hoffner-Lesure, A. (2018) *Le codéveloppement professionnel et managérial*, EMS Editions.

Naughton, John. "The Evolution of the Internet: From Military Experiment to General Purpose Technology." *Journal of Cyber Policy* 1, no. 1 (2016): 5–28.

Ricoeur, P. (1990) *Soi-même comme un autre*, Paris, Seuil.