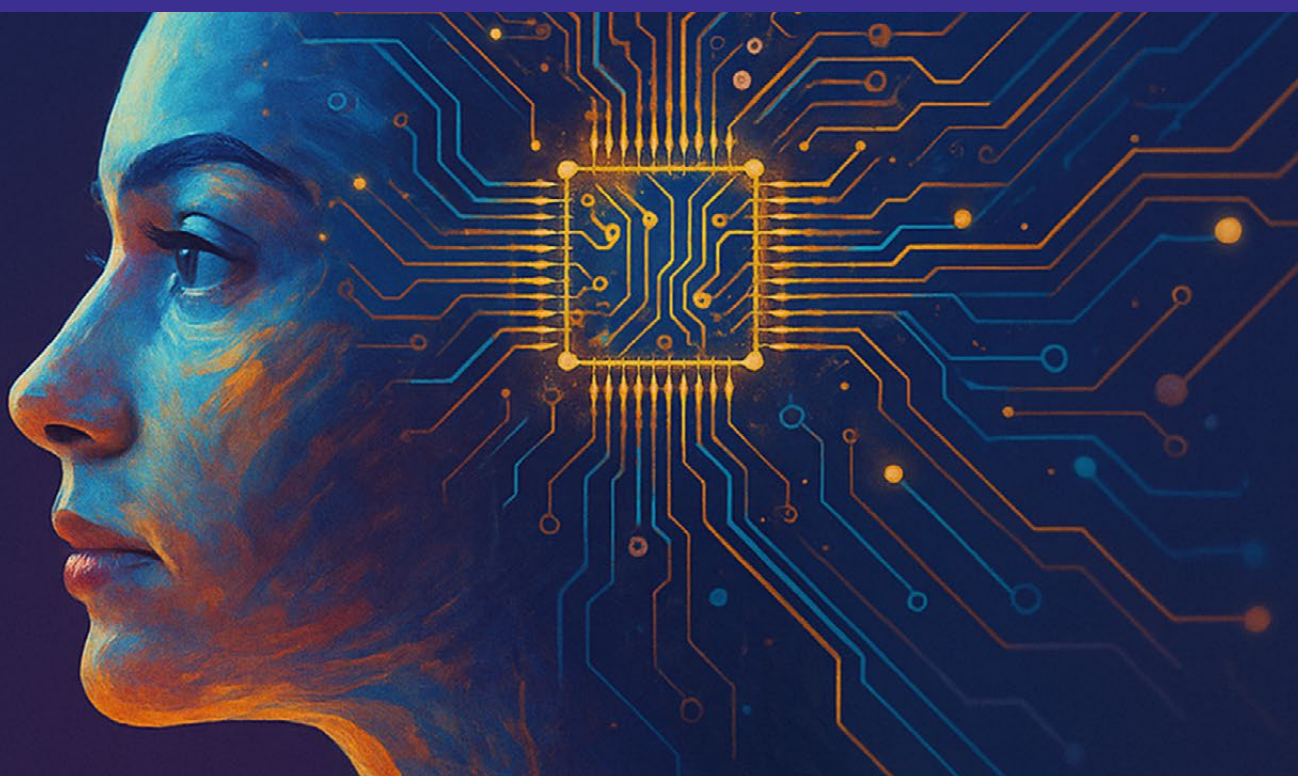




LIGHTS - Technology

L'intelligence artificielle et la paréidolie

ESCP Impact Paper No.2025-48-FR

Frédéric FRÉRY
ESCP Business School

[**L'intelligence artificielle et la paréidolie**

Frédéric Fréry*

ESCP Business School

Résumé

La paréidolie est le biais perceptif par lequel l'esprit attribue du sens à des motifs qui n'en ont pas. C'est par exemple ce qui nous fait voir un visage dans des rochers sur Mars ou des animaux dans la forme des nuages. La recherche montre que les IA, tout comme les humains, sont victimes de ce phénomène, ce qui les pousse à surinterpréter des informations aléatoires. Cependant, la notion de paréidolie peut aussi s'appliquer à notre propre perception des résultats fournis par les IA. En effet, lorsque nous discutons avec une IA, nous avons parfois tendance à lui attribuer une véritable intelligence, alors que ce n'est qu'un programme informatique, limité à ses algorithmes et aux données qui ont permis de l'entraîner. Il est important de rappeler que l'intelligence artificielle n'est pas intelligente, et que nous ferions mieux de parler d'intelligence simulée. En revanche, l'IA nous donne une occasion unique de mieux comprendre ce qu'est véritablement l'intelligence humaine, avec ses limites et ses spécificités.

Mots clés : paréidolie, IA, biais, illusion, psychologie

*Professeur, ESCP Business School

L'intelligence artificielle et la paréidolie

En 1976, la sonde américaine Viking 1, en orbite autour de la planète Mars, prend une photo du massif montagneux Cydonia Mensae dans laquelle on croit percevoir un visage, ce qui déclenche de nombreuses spéculations plus ou moins fantaisistes, certains étant convaincus que cela prouve l'existence d'une civilisation martienne. Pourtant, en 1998, une autre sonde, Mars Global Surveyor, photographie le même lieu, sous un autre éclairage et avec une résolution dix fois supérieure, et les clichés montrent une simple colline, sans aucun visage. De même, dans une scène du film *Le Fabuleux Destin d'Amélie Poulain*, de Jean-Pierre Jeunet (2001), la toute jeune Amélie, pour tromper son ennui, photographie des nuages en forme d'ours en peluche ou de lapins.

Dans les deux cas, on parle de paréidolie, du grec *para* qui signifie « à côté » et *eidôlon*, diminutif de *eidos*, qui signifie « apparence, forme ». Il s'agit d'une illusion d'optique, qui donne du sens là où il n'y en a pas, en nous faisant voir des formes familières (un visage, un animal, etc.) dans des images vagues ou ambiguës (nuages, rochers, taches d'encre, etc.). La paréidolie peut également se manifester sur le plan auditif, lorsqu'un son familier est perçu dans un bruit indistinct, ou lorsqu'une structure mélodique ou rythmique semble émerger, par exemple, du bruit d'une machine.

La paréidolie est un biais cognitif qui a vraisemblablement été sélectionné par l'évolution naturelle : pour nos lointains ancêtres, être capable de repérer immédiatement un prédateur dangereux ou au contraire un visage familier était une question de survie, et ce phénomène est la plupart du temps inoffensif.

Cependant, il peut mener à des interprétations erronées, et il contribue notamment à alimenter des superstitions. En 1994, une habitante de Floride a ainsi affirmé avoir identifié l'image de la Vierge Marie dans un sandwich au fromage grillé qu'elle s'apprêtait à manger. Après avoir conservé ce sandwich pendant dix ans, elle a fini par le mettre en vente sur eBay, où il a atteint le prix astronomique de 28 000 dollars. De même, en 2009, la forme d'une tache d'humidité sur le mur d'une gare ferroviaire désaffectée de Santiago du Chili a été interprétée par plusieurs passants comme une apparition du Christ. Les médias chiliens ont largement relayé cet événement, et le mur est devenu un lieu de pèlerinage et de prières.

En psychologie, c'est sur la paréidolie que reposent les tests projectifs, tels que le test de Rorschach, conçus comme des outils de révélation de la personnalité. En effet, la paréidolie est influencée par le capital culturel et les expériences propres à chaque individu. La paréidolie est aussi référencée en psychiatrie comme une forme d'apophénie, c'est-à-dire un trouble de la perception qui fait relier des événements ou des phénomènes n'ayant aucun lien entre eux, ce qui provoque des illusions sensorielles particulièrement intenses et persistantes. Pry (2021) rappelle ainsi que certaines lésions du lobe temporal peuvent occasionner des paréidolies, tandis que les personnes atteintes de névroses semblent y être plus exposées que les autres.

Enfin, certains artistes, comme Salvador Dalí ou Giuseppe Arcimboldo, ont largement utilisé la paréidolie dans leurs œuvres. Pour Dalí, on peut penser notamment à la toile « Marché d'esclaves avec apparition du buste invisible de Voltaire » ou à l'agencement de la pièce du musée-théâtre de Figueres « Marilyn Monroe » (dans laquelle deux tableaux figurent les yeux de l'actrice et des rideaux ses cheveux). Il en est de même pour les célèbres représentations des saisons et des éléments par Arcimboldo, où des personnages sont composés de fruits, de légumes et de plantes.

Or, si la paréidolie est un phénomène largement documenté chez l'être humain, mais aussi chez d'autres primates (Taubert *et al.*, 2017), il semble connaître une actualité nouvelle avec l'extraordinaire succès des systèmes d'intelligence artificielle (IA) à base de *machine learning* et de *deep learning*.

L'intelligence artificielle, elle aussi victime de paréidolie

En effet, les IA sont elles aussi sujettes à la paréidolie. Du fait même de leur entraînement, les systèmes de *machine learning* ont parfois tendance à identifier des visages humains dans des objets inanimés (comme des panneaux ou des bâtiments) ou à confondre des animaux avec des fruits ou des légumes. Ces hallucinations conduisent à des « faux positifs », qui peuvent être particulièrement problématique, par exemple pour les véhicules autonomes. C'est la raison pour laquelle des techniques telles que le ReCAPTCHA, utilisé notamment pour se connecter à certains sites Internet, demandent aux utilisateurs humains d'identifier le sens d'images ambiguës (par exemple des feux de signalisation ou des motos), afin de parfaire – gratuitement – l'entraînement des IA et de leur éviter de fausses interprétations.

Réciproquement, l'IA peut être utilisée pour créer des paréidolies. On trouve ainsi sur Internet des requêtes qui permettent de demander à des IA génératives comme Dall-E, Midjourney ou Stable Diffusion de produire des paréidolies, généralement en mélangeant la photo d'un visage avec celle d'un paysage. Les résultats sont probants, et les êtres humains reconnaissent généralement le visage dans l'image obtenue.

Une recherche du MIT Computer Science and Artificial Intelligence Laboratory (Hamilton *et al.*, 2024), a exploré de manière plus systématique le lien entre IA et paréidolie. En s'appuyant sur un jeu de données inédit comprenant 5 000 images paréidoliques, cette étude montre que les modèles d'IA ne détectent pas aussi bien que les humains les paréidolies faciales, c'est-à-dire la tendance à percevoir des visages dans des objets inanimés, comme un rocher sur Mars ou un sandwich au fromage grillé. Contrairement aux humains, capables d'identifier des visages aussi bien dans des images très rudimentaires (ou très bruitées), que dans des scènes visuellement surchargées (ou confuses), les algorithmes échouent dans les deux cas. Leur performance maximale – que Hamilton *et al.* qualifient de *pic paréidolique* – s'observe dans des images de complexité modérée, comportant peu d'objets bien séparés et distincts. Même en entraînant les IA avec encore plus de visages humains, elles restent bien moins sujettes que nous à la paréidolie faciale. En revanche – et c'est une des conclusions les plus surprenantes de cette recherche – lorsqu'on entraîne les IA à reconnaître des visages d'animaux, elles deviennent beaucoup plus sujettes à la paréidolie de visages humains. Cela semble confirmer que la paréidolie est une capacité évolutive, enchâssée dans notre système au sens de Kahneman (2011), c'est-à-dire notre système de pensée rapide, implicite, automatique et non conscient.

L'intelligence artificielle, une vaste paréidolie ?

Pour autant, il existe une autre manière de considérer le lien entre IA et paréidolie, beaucoup plus générale : et si notre perception de l'IA était elle-même une vaste paréidolie ?

En effet, lorsque nous discutons avec une IA, nous avons parfois tendance à lui attribuer une véritable intelligence, alors que ce n'est qu'un programme informatique, limité à ses algorithmes et aux données qui ont permis de l'entraîner. Tout comme les animaux que

nous croyons voir dans les nuages, nous croyons voir une forme d'intelligence dans les réponses des IA.

Pour rappel, l'expression « *Artificial Intelligence* » est apparue pour la première fois en 1955, dans le titre du document *Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, rédigé par quatre chercheurs, John McCarthy (Dartmouth College), Marvin Minsky (Harvard University), Nathaniel Rochester (IBM) et Claude Shannon (Bell Telephone Laboratories). Ce document proposait l'organisation d'un atelier de recherche à l'été 1956, atelier qui est généralement considéré comme l'acte de naissance de l'IA. Or, le choix de l'expression « *Artificial Intelligence* » n'était pas anodin. Il visait à se démarquer des domaines existants, et notamment de la cybernétique. Centrée sur l'adaptation, la régulation et le contrôle des systèmes, cette discipline était fortement associée à son fondateur, le mathématicien Norbert Wiener. Or, McCarthy souhaitait s'écarter de la cybernétique – mais aussi de la figure tutélaire de Wiener – pour se consacrer spécifiquement à la simulation de l'intelligence humaine : « *The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.* » (McCarthy et al., 1955).

Cependant, parler dans le titre du document d'intelligence *artificielle* plutôt que d'intelligence *simulée* avait un autre objectif : trouver une dénomination originale, intrigante, voire paradoxale (avec l'idée d'une intelligence non humaine), afin attirer l'attention et donc les financements. Le document sollicitait d'ailleurs le soutien de la fondation Rockefeller pour l'organisation de l'atelier de recherche de l'été 1956, pour un montant de 13 500 dollars (soit plus de 160 000 dollars actuels). « Intelligence Artificielle » était un nom porteur de promesses spectaculaires, ce qui le rendait séduisant pour les financeurs, notamment gouvernementaux, en particulier dans le contexte de la guerre froide et de la compétition technologique avec l'URSS.

Au total, nous utilisons donc une terminologie conçue dans un but de communication, qui dès l'origine a cherché à enjoliver la réalité. Comme le soulignent certains experts actuels (Seifert, 2025 ; Venkatesh, 2025), parler d'Intelligence Artificielle est un abus de langage : l'IA est bien artificielle, mais elle n'est certainement pas intelligente. C'est une simulation algorithmique programmée pour faire ce qu'elle fait, et il serait bien plus pertinent de parler d'intelligence *simulée* ou d'intelligence *synthétique*.

Pour autant, notre tendance à la paréidolie nous incite à percevoir de l'intelligence dans l'IA, et parfois de manière excessive. Selon une recherche du MIT Media Lab (Phang et al., 2025), si la plupart des utilisateurs de l'agent conversationnel ChatGPT le considèrent comme un simple outil, une minorité développe un véritable rapport émotionnel avec lui, au point de lui confier des informations personnelles qu'ils hésiteraient à partager avec des êtres humains. Encore plus étonnant, 10 % des utilisateurs de ChatGPT le considèrent comme un ami et reconnaissent éprouver un sentiment de détresse lorsqu'ils ne peuvent plus y accéder. Cette recherche révèle également que les échanges vocaux avec ChatGPT suscitent jusqu'à 10 fois plus d'attachement émotionnel que les échanges textuels : la paréidolie est ainsi beaucoup plus intense lorsque la voix est utilisée.

Les individus qui déclarent un niveau élevé de solitude ont plus tendance à s'attacher affectivement à ChatGPT, ce qui renforce davantage leur isolement social et leur dépendance émotionnelle à l'IA. Dans certains cas extrêmes, cela peut conduire au pire. Un adolescent américain de 14 ans s'est ainsi suicidé en 2024 après avoir développé une relation toxique avec un agent conversationnel proposé sur abonnement par l'entreprise character.ai, qui vante sa capacité à tenir des conversations réalistes.

Certains auteurs – comme Geneste et Galindo (2025) ou Lardellier et Carré (2025) – soulignent que le discours autour de l'IA mobilise de plus en plus des métaphores religieuses. Des figures du transhumanisme comme Ray Kurzweil sont parfois considérées comme des prophètes, des oracles ou des évangélistes, tandis que le Dataïsme érige les algorithmes en divinités séculières, promettant des réponses absolues grâce à l'analyse des données (Harari, 2016). Cette sacralisation de la technologie s'inscrit dans un contexte où l'IA, en offrant un nouveau récit existentiel, tend à combler le vide laissé par le recul des religions traditionnelles. Comme un écho à la tache d'humidité sur le mur de la gare de Santiago du Chili.

Enfin, la paréidolie est aussi encouragée par les lois du marché. D'un point de vue managérial, face au tsunami annoncé de la substitution de l'humain par la machine dans la plupart des entreprises, les écoles et universités se sont empressées d'inclure l'IA dans leurs programmes, au risque de remplacer par un succédané synthétique ce qui définit notre humanité (voire, selon le *cogito* de Descartes, ce qui légitime notre existence). À tout le moins, ces formations devraient inclure une dimension critique : un clone numérique n'est que la simulation d'un client, d'un investisseur ou d'un électeur, et il doit toujours être considéré comme tel, ni plus, ni moins. Or, pour susciter l'intérêt des utilisateurs et capter leur attention, c'est plutôt l'anthropomorphisme qui est recherché par les développeurs, notamment au travers d'interfaces vocales (et bientôt vidéo) toujours plus convaincantes, capables de simuler non seulement la réflexion, mais aussi les émotions. Malheureusement, cela accroît les risques de détournements, de tromperies et de manipulations. Si la frontière entre l'humain et l'IA se brouille, les entreprises, dans leur volonté de séduire leurs clients, n'y sont certainement pas étrangères. C'est une des responsabilités des formations au management que d'alerter sur cette dérive et de faire émerger une approche plus éthique de l'IA.

Et si la paréidolie était le propre de l'intelligence humaine ?

Au total, le concept de paréidolie permet de mieux comprendre non seulement notre relation à l'IA, mais peut-être aussi ce qui la différencie de l'intelligence humaine. Simulation algorithmique exagérément présentée dès sa naissance comme une véritable intelligence, l'IA fait désormais partie de notre quotidien, au point que certains lui attribuent une authentique intelligence, égale, voire supérieure, à l'intelligence humaine. Un outil informatique, rapide, pratique et ubiquitaire, voit ainsi ses lignes de codes assimilées à une véritable pensée.

Or, notre propension à percevoir de l'intelligence là où il n'y en a pas, à chercher du sens, de l'émotion et une narration dans la machine, est profondément humaine. Comme le rappelait Herbert Simon (1955), ce ne sont pas une logique parfaite ou une rationalité absolue qui nous définissent, mais nos biais cognitifs, nos imperfections, nos impulsions et nos émotions. Bien sûr, les IA sont capables d'hallucinations, mais elles ne font en cela que suivre mécaniquement ce que nous avons inclus dans leurs algorithmes et leurs données. Bien sûr, l'utilisation de l'IA peut nous aider à combler les lacunes de notre rationalité limitée pour nous rapprocher d'une rationalité procédurale, certainement plus absolue, mais au prix de ce qui fait notre singularité humaine.

Être exposé à la paréidolie, imaginer que la machine pense, n'est pas une faiblesse. C'est au contraire la preuve de ce qui manque à l'intelligence artificielle : une conscience façonnée par l'émotion, une perception à la recherche de sens, une intelligence contrainte par ses failles.

Et c'est peut-être cela, plus que tout, qui fait l'intelligence humaine.

Références

Geneste S., Galindo G. (2025), "'I believe in AI': An exploration of employee acceptance of AI through religious lenses", *Academy of Management Annual Meeting*, Copenhagen.

Hamilton M., Stent S., DuTell V., Harrington A., Corbett J., Rosenholtz R., Freeman W.T. (2024), "Seeing Faces in Things: A Model and Dataset for Pareidolia", Working Paper, *MIT Computer Science and Artificial Intelligence Laboratory*.

Harari Y.N. (2016), "Homo Deus: A brief history of tomorrow". Harper.

Kahneman D. (2011), "Thinking, Fast and Slow", Farrar, Straus and Giroux.

Lardellier P., Carré E. (2025), « L'IA réactive des croyances mystiques », *The Conversation France*.

McCarthy J., Minsky M., Rochester N., Shannon C. (1955), "Proposal for the Dartmouth Summer Research Project on Artificial Intelligence", Dartmouth College.

Phang J., Lampe M., Ahmad L., Agarwal S., Mengying Fang C., Liu A.R., Danry V., Lee E., Pataranutaporn P., Maes P. (2025), "Investigating Affective Use and Emotional Well-being on ChatGPT", Working Paper, *MIT Media Lab*.

McCarthy J., Minsky M., Rochester N., Shannon C. (1955), "Proposal for the Dartmouth Summer Research Project on Artificial Intelligence", Dartmouth College.

Pry R. (2021), « L'enfant et les paréidolies », *Enfance*, 2(2), pp. 205-2011.

Seifert J. (2025), "The Meaning of Intelligence: A Realist Phenomenological Groundwork for a Philosophy of AI, Transhumanism, and Posthumanism", *Journal of East-West Thought*, 15(2), pp. 1-20.

Simon H.A. (1955), "A behavioral model of rational choice", *The Quarterly Journal of Economics*, 69(1), pp. 99-118.

Taubert J., Wardle S.G., Flessert M., Leopold D.A., Ungerleider L.G. (2017), "Face pareidolia in the rhesus monkey", *Current Biology*, 27(16), pp. 2505-2509.

Venkatesh P. (2025), "On artificial intelligence in healthcare", *Indian Journal of Ophthalmology*, 73(6); pp. 777-778.