



WE
DON'T
NEED
NO
LIMITATIONS

Déconstruire les murs de l'IA

We don't need no limitations

Déconstruire les murs de l'IA

Le 30 novembre 1979, Pink Floyd dévoilait *The Wall*, un album conceptuel devenu une icône culturelle, explorant les thèmes de l'isolement, de la souffrance et de la construction de barrières intérieures. À l'occasion du 45e anniversaire de sa sortie, cet eBook revisite la métaphore des murs, mais cette fois dans le contexte de l'intelligence artificielle (IA). L'album raconte l'histoire de Pink, un personnage qui, pour se protéger, érige progressivement un mur autour de lui. Ce mur, censé le préserver de la douleur, se transforme en une prison, alimentant défiance et désespoir.

Aujourd'hui, la théorie de Bertrand Braunschweig sur les «murs» de l'IA offrent un éclairage fascinant sur les obstacles, tant techniques que conceptuels, qui séparent l'IA de la société humaine. Dans ce contexte, les murs ne sont pas toujours des barrières intentionnelles comme celle de Pink, mais ils érigent des frontières invisibles entre l'IA et ses utilisateurs. Ce cadre théorique identifie cinq grands murs qui entravent le développement de l'IA : la confiance, la consommation énergétique, la sécurité des systèmes, les interactions homme-machine et l'inhumanité des machines. Enrichi d'un sixième mur, celui de la limitation des données d'apprentissage disponibles, cet eBook analyse comment ces barrières affectent l'évolution de l'IA et son acceptation.

Nous explorerons dans cet eBook les dynamiques de ces six murs : certains sont érigés consciemment – au moins partiellement - pour protéger, comme le mur de Pink dans *The Wall*, tandis que d'autres se dressent involontairement, fruits des limites techniques et des lacunes de compréhension. En abordant chaque mur, nous soulignerons la distinction entre les obstacles activement érigés et ceux subis, pour montrer comment ils enferment l'IA dans des contraintes spécifiques et parfois aliénantes. Avec, bien entendu, une réflexion sur la meilleure façon de salutairement les contourner ou démolir.



Images générées par Midjourney avec les consignes : « <https://s.mj.run/zyMlw-D7AIM> the wall theory of AI --chaos 10 --ar 3:4 --style raw --sref <https://s.mj.run/zyMlw-D7AIM> --cref <https://s.mj.run/zyMlw-D7AIM> --weird 300 --v 6.1 »

et : « Human and machine (artificial intelligence) separated by five walls. »

Comprendre le concept des murs de l'IA

L'intelligence artificielle, en dépit de ses avancées prodigieuses, se heurte à des obstacles qui freinent son développement et limitent son intégration dans la société. Bertrand Braunschweig, chercheur et coordinateur scientifique du programme *Confiance.ai*, a théorisé cinq de ces obstacles majeurs, qu'il nomme les «murs» de l'IA. Ces murs représentent des barrières structurelles qui empêchent l'IA d'atteindre son potentiel idéal. Loin d'être des obstacles simples ou ponctuels, ces murs soulèvent des questions profondes d'ordre technique, éthique, écologique et social. Ils sont à la fois des défis et des garde-fous, reflétant la complexité des rapports entre l'homme et la machine.

Les cinq murs théorisés par Braunschweig incluent : le **mur de la confiance**, le **mur de la consommation énergétique**, le **mur de la sécurité des systèmes**, le **mur des interactions homme-machine** et le **mur de l'inhumanité des machines**. Chacun de ces murs a ses particularités, mais ils ont en commun de poser des limites à l'IA dans son rapport aux humains et à l'environnement. Ces obstacles ne sont pas tous intentionnels : certains sont construits consciemment (partiellement) comme des mesures de protection ou de régulation, tandis que d'autres émergent passivement des limitations techniques ou des conceptions imparfaites. Par conséquent, ces murs sont souvent perçus comme des éléments qui alourdissent et isolent l'IA, la rendant difficile à adopter et à intégrer de manière fluide dans la société.

À ces cinq murs de Braunschweig, nous ajoutons un **sixième mur** : celui de la **limitation des données d'apprentissage disponibles**. Ce dernier obstacle représente une barrière potentielle pour les modèles de langage et les systèmes d'IA qui reposent sur de vastes volumes de données.



Il est estimé que d'ici 2026 à 2032, les données textuelles humaines accessibles et exploitables pourraient être intégralement utilisées, créant une impasse pour le développement de ces systèmes. Ce mur de la limitation des données est particulièrement inquiétant, car il pourrait imposer un plafond à l'évolution de l'IA, affectant directement sa capacité à s'adapter et à s'améliorer.

Ainsi, ce cadre théorique des «murs de l'IA» nous invite à une analyse approfondie des défis que ces barrières posent.

Six murs de l'IA

Pour mieux comprendre les obstacles que représentent les murs de l'IA, on peut les mettre en parallèle avec le mur métaphorique de *The Wall*. Dans cet album conceptuel, le protagoniste, Pink, construit progressivement un mur autour de lui pour se protéger des traumatismes et des déceptions. Ce mur, qu'il croyait protecteur, devient une prison qui l'isole, alimente sa défiance et le coupe du monde extérieur.

De manière similaire, les murs de l'IA isolent cette technologie de la société, l'enferment dans des limites et génèrent des perceptions de méfiance, de complexité ou de danger.

En examinant chacun des murs de l'IA définis par Bertrand Braunschweig, et en ajoutant celui de la limitation des données, on découvre comment ces barrières créent, à leur tour, une prison invisible pour l'IA, limitant son potentiel et son intégration sociale.

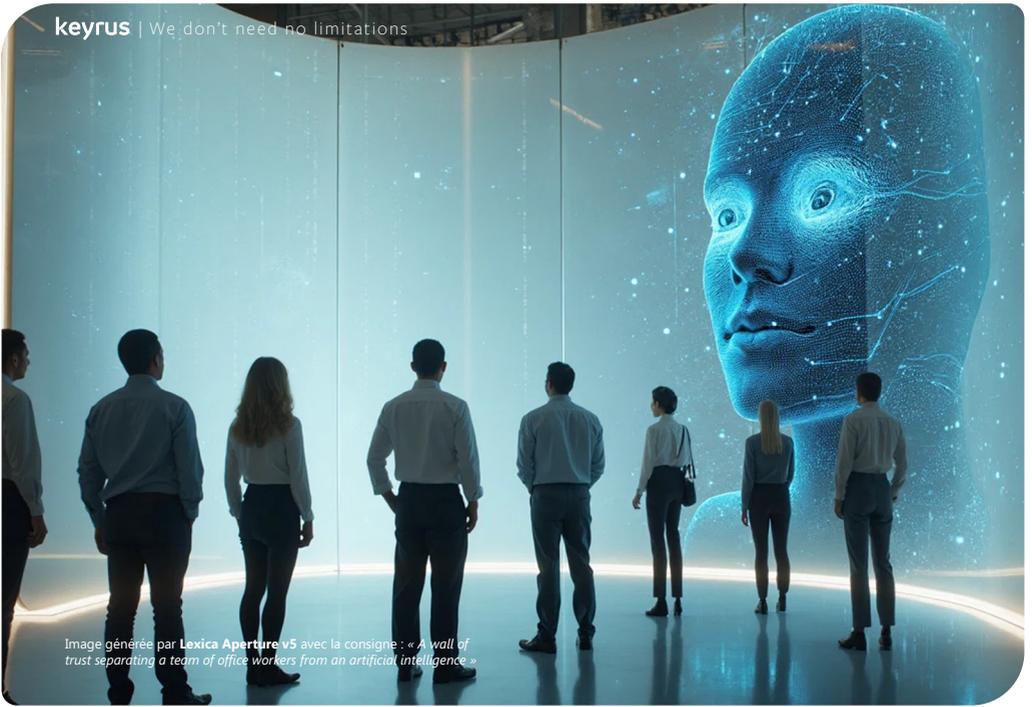


Image générée par Lexica Aperture v5 avec la consigne : « A wall of trust separating a team of office workers from an artificial intelligence »

Le mur de la confiance envers l'IA

Dans *The Wall*, Pink construit un mur pour se protéger, mais cette barrière érode progressivement sa confiance en autrui, le rendant méfiant et isolé. De la même manière, le mur de la confiance envers l'IA représente une méfiance croissante des utilisateurs face aux systèmes IA, perçus comme opaques et difficiles à comprendre. Ce manque de transparence, souvent lié à la nature de boîte noire des modèles d'IA, crée une barrière entre l'IA et la société humaine. Comme Pink qui, par son isolement, devient inaccessible aux autres, l'IA, sans efforts de transparence et d'explicabilité, reste perçue comme un « autre » incompréhensible et potentiellement dangereux.

Le mur de la consommation énergétique

Dans *The Wall*, l'entretien de ce mur mental et émotionnel exige de Pink une énergie colossale, le poussant dans une spirale de surcharge et de fatigue psychologique. Ce concept d'épuisement trouve un écho dans le mur de la consommation énergétique de l'IA.

Les modèles d'IA, notamment ceux qui s'appuient sur l'apprentissage profond, demandent des quantités massives d'énergie pour fonctionner. Ce mur est un obstacle concret à la durabilité des technologies IA, posant des enjeux écologiques essentiels. Sans optimisation de la consommation, l'IA risque de se retrouver dans une impasse où ses besoins énergétiques croissants deviendront insoutenables, créant ainsi une forme de « prison énergétique ».

Le mur de la sécurité des systèmes

Dans *The Wall*, le mur de Pink est conçu pour le protéger de la souffrance, mais il devient un piège, augmentant sa vulnérabilité. Le mur de la sécurité des systèmes IA reflète ce dilemme : les mesures de sécurité sont censées protéger les utilisateurs, mais elles peuvent créer de nouvelles failles. Les systèmes d'IA, en particulier ceux qui régissent des infrastructures critiques, peuvent eux-mêmes devenir des cibles d'attaques, transformant les mécanismes de sécurité en zones de vulnérabilité. Ce mur expose l'IA à des dilemmes de fiabilité et d'éthique, où l'isolement voulu pour des raisons de sécurité peut paradoxalement augmenter les risques.

“ Les systèmes d’IA, en particulier ceux qui régissent des infrastructures critiques, peuvent eux-mêmes devenir **des cibles d’attaques**, transformant les mécanismes de sécurité en zones de vulnérabilité. ”



Image générée par **Lexica Aperture v5** avec la consigne : « A wall separating a team of office workers from non-human artificial intelligences »

Le mur des interactions homme-machine

L'isolement de Pink symbolise une incapacité à interagir sainement avec autrui, le laissant prisonnier d'une solitude aliénante. Le mur des interactions homme-machine en IA évoque une situation similaire : bien que les technologies IA évoluent, elles peinent encore à interagir de manière fluide et intuitive avec les utilisateurs. Ce mur de l'incompréhension technologique empêche l'IA de s'intégrer harmonieusement dans la vie humaine, la rendant distante et impersonnelle. Pour surmonter ce mur, il est essentiel de repenser les interfaces et les modes d'interaction pour créer des échanges plus naturels et compréhensibles, et ainsi abattre cette barrière de déconnexion.

Le mur de l'inhumanité des machines

Le mur que Pink érige finit par l'aliéner totalement, le transformant en une figure froide, déshumanisée, incapable d'établir des liens authentiques. Ce mur de l'inhumanité est aussi un obstacle majeur pour l'IA : malgré tous les progrès, elle reste fondamentalement une machine, sans empathie ni véritable compréhension des émotions humaines. Ce manque de dimension humaine crée une barrière émotionnelle entre l'IA et ses utilisateurs.

Ce mur pourrait être partiellement surmonté par une recherche interdisciplinaire visant à introduire des cadres éthiques et à améliorer le contexte dans lequel l'IA est utilisée. Toutefois, cette inhumanité de la machine reste un obstacle difficilement franchissable, limitant le lien de confiance et de familiarité que les humains pourraient avoir avec ces systèmes.

Le mur de la limitation des données d'apprentissage

Enfin, Pink craint de manquer des ressources émotionnelles nécessaires pour se protéger, une peur qui alimente sa construction de murs. Ce thème trouve un écho dans le mur de la limitation des données d'apprentissage en IA. Les modèles de langage et les systèmes IA exploitent de vastes volumes de données humaines pour s'améliorer, mais il est estimé que d'ici 2026-2032, l'ensemble des données publiques disponibles pourrait être épuisé. Ce mur représente une limite potentielle à la croissance de l'IA, créant une «stagnation cognitive» où le manque de nouvelles données ralentirait ou stopperait le progrès. Sans solutions alternatives — comme l'optimisation de l'apprentissage ou la création de données synthétiques — ce mur pourrait freiner considérablement l'évolution de l'IA.

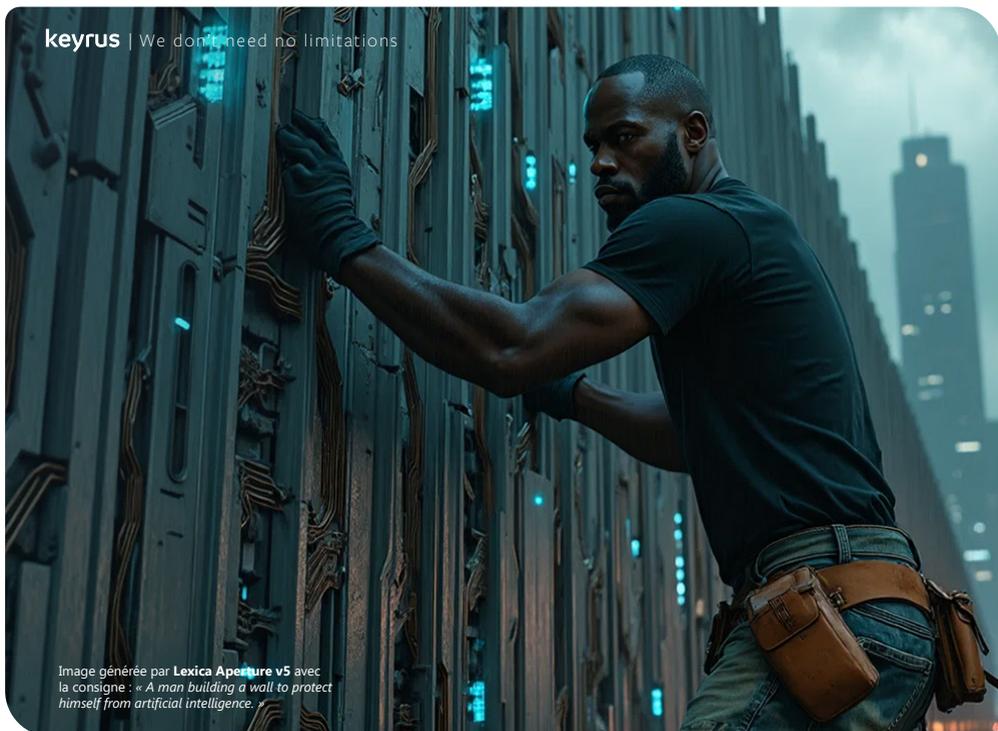


Image générée par **Lexica Aperture v5** avec la consigne : « *A man building a wall to protect himself from artificial intelligence.* »

Actif ou passif : une distinction essentielle

Pink, construit volontairement son mur pour se protéger des blessures émotionnelles et des souffrances de sa vie. Ce mur, bien qu'invisible aux autres, est le fruit d'une intention consciente, même si elle est destructrice. Les murs de l'IA, en revanche, ne sont pas tous le résultat d'une décision consciente. Certains d'entre eux, comme la sécurité ou l'interaction homme-machine, sont érigés de manière volontaire pour éviter des dérives ou pour protéger les utilisateurs. D'autres, cependant, apparaissent passivement, fruits de limitations techniques, de contraintes contextuelles ou de lacunes dans notre compréhension. Cette distinction entre barrières «érigées» et barrières «subies» est essentielle pour appréhender les défis que ces murs posent et la manière de les surmonter.

Les murs érigés de l'IA

Certains murs de l'IA sont érigés délibérément pour répondre aux préoccupations de sécurité, de confiance ou d'éthique. Par exemple, le **mur de la sécurité des systèmes** est mis en place pour protéger les utilisateurs des vulnérabilités de l'IA et pour limiter les risques liés à l'autonomie des systèmes. De même, le **mur de la confiance** repose sur des efforts délibérés pour instaurer la transparence et renforcer la crédibilité des systèmes d'IA. Enfin, le **mur des interactions homme-machine** s'élève en partie du fait de choix conscients visant à éviter des comportements de l'IA qui seraient trop intrusifs ou socialement perturbants.

Ces murs «construits» représentent une forme de contrôle que nous exerçons sur l'IA pour préserver des valeurs et protéger les individus. Mais, à l'instar du mur de Pink, ils peuvent aussi devenir des barrières oppressives et auto-limitantes, suscitant méfiance et isolement plutôt que sécurité et acceptation. En voulant «protéger» l'utilisateur, ces murs créent parfois un fossé entre l'IA et la société, un paradoxe où les mesures de sécurité peuvent finir par alimenter la défiance et isoler davantage l'IA.

Les murs *subis* de l'IA

D'autres murs, en revanche, se dressent sans qu'ils aient été volontairement érigés. Ils résultent de contraintes techniques, comme **la consommation énergétique** nécessaire pour faire fonctionner les modèles d'IA à grande échelle, ou de limitations contextuelles, telles que **le manque de données d'apprentissage disponibles** pour nourrir des modèles de plus en plus complexes. Ces murs *subis* sont souvent le reflet de défis que l'on n'avait pas anticipés ou que l'on n'a pas encore su surmonter. Contrairement aux murs construits, ils n'ont pas pour but de restreindre l'IA mais résultent des limites mêmes de la technologie et de ses besoins.

Ces barrières passives posent des problèmes de durabilité et de croissance. Par exemple, le mur de la consommation énergétique questionne la viabilité à long terme de l'IA et de ses impacts écologiques, tandis que le mur de la limitation des données d'apprentissage pourrait freiner l'amélioration des modèles IA, limitant leur capacité à évoluer. Contrairement à un mur érigé consciemment, ces barrières techniques apparaissent comme des freins inévitables auxquels la technologie devra s'adapter pour progresser.

Implications de la distinction actif/passif

Cette différence entre les murs actifs et passifs de l'IA a des implications majeures pour les stratégies à adopter. Les murs construits consciemment sont plus facilement ajustables, car ils résultent de décisions humaines et peuvent être reconfigurés en fonction des besoins et des contextes. Des choix éclairés, basés sur des discussions éthiques et sociales, pourraient réduire leur impact négatif en ajustant les exigences de sécurité ou en créant des interfaces homme-machine plus intuitives.

Image générée par Midjourney avec la consigne : « Vintage poster of a computerized wall consuming tremendous amounts of energy. »



En revanche, les murs *subis* exigent une innovation technique et des recherches poussées pour être surmontés. Ils posent des défis de taille, nécessitant de nouvelles architectures, des technologies de pointe en matière de gestion de l'énergie et une exploration continue des données disponibles ou des données de synthèse.

En somme, comprendre cette distinction est essentiel pour élaborer des solutions adaptées et équilibrées. Les murs actifs, comme ceux de la sécurité et de la confiance, relèvent d'une volonté de maîtrise ; leur démantèlement partiel dépend d'un choix collectif visant à alléger les restrictions sans compromettre les principes de sûreté et de fiabilité. Les murs passifs, quant à eux, nous rappellent que la technologie est soumise à des contraintes qu'il faudra vaincre par l'innovation et la collaboration interdisciplinaire.

Ainsi, comme le mur de Pink dans *The Wall*, les murs de l'IA peuvent être à la fois protecteurs et oppressifs. Démanteler certains de ces murs ou en atténuer l'impact nécessitera non seulement des avancées techniques mais aussi un questionnement profond de la place que nous voulons accorder à l'IA dans notre société.

“ Les **murs construits** sont plus facilement ajustables, car ils résultent de décisions humaines et peuvent être reconfigurés en fonction des besoins et des contextes. Les **murs subis** exigent une innovation technique et des recherches poussées pour être surmontés. ”

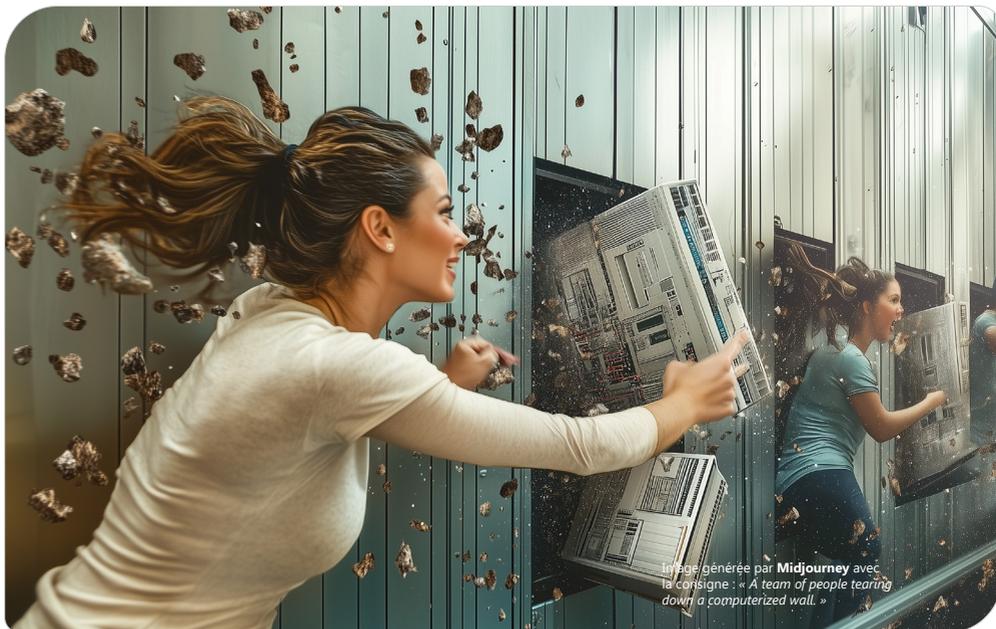


Image générée par Midjourney avec la consigne : « A team of people tearing down a computerized wall. »

Vers une démolition consciente des murs de l'IA

On l'aura compris, l'intelligence artificielle, malgré ses promesses et ses prouesses, reste encerclée de murs qui limitent son potentiel et entravent sa pleine intégration dans la société. Ces murs — qu'ils soient de confiance, de sécurité, d'interaction, d'inhumanité, de consommation énergétique, ou de disponibilité des données — forment une structure complexe de barrières, parfois érigées volontairement pour protéger, parfois subies en raison des limites de la technologie.

Et cette distinction entre les murs «construits» consciemment et les murs «subis» est essentielle pour élaborer des solutions adaptées à chaque type de barrière.

Bien évidemment, démolir ces murs ne signifie pas ouvrir toutes les portes de l'IA sans réflexion. Il s'agit plutôt de détruire les aspects oppressifs de ces barrières tout en préservant les garde-fous nécessaires.

La question n'est pas seulement de rendre l'IA plus puissante ou plus présente dans notre quotidien, mais de l'intégrer de manière harmonieuse et équilibrée dans la société. Cela implique une démarche proactive et consciente, où l'IA est modelée non seulement par des impératifs techniques, mais aussi par des valeurs humaines qui garantiront son acceptation sociale et sa viabilité à long terme.

Finalement, cette «démolition consciente» des murs de l'IA invite à une réflexion profonde sur la place de cette technologie dans notre monde. À l'image de Pink, il est possible de dépasser ces murs et d'éviter l'isolement — en transformant les frontières en points de connexion entre l'humain et l'intelligence artificielle. La clé réside dans un équilibre subtil entre innovation, régulation et éthique, pour que l'IA devienne un outil au service de l'humanité, sans jamais s'en détacher ni s'enfermer dans des limites inutiles.

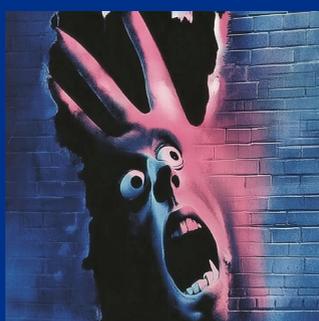
Article co-écrit par keyrus, Chat-GPT^{4.0} et Mistral, Claude, Copilote, Perplexity et Gemini

Les unes auxquelles vous avez échappé...

Lors de la création de nos ebooks, chaque image naît d'un processus itératif avec le moteur d'IA. Des allers-retours répétés permettent d'affiner chaque visuel, jusqu'à atteindre l'expression la plus fidèle de l'idée souhaitée. Mais le travail ne s'arrête pas là : parmi les images générées, il faut ensuite opérer des choix. Ce processus de sélection repose sur un effort d'objectivation, où il s'agit de mettre de côté sa propre subjectivité pour respecter au mieux le « cahier des charges ». Une démarche délicate, tant les images peuvent inspirer, surprendre, voire fasciner, mais essentielle pour rester fidèle à l'intention initiale.

L'atelier du prompteur...

Images générées par **Midjourney** avec la consigne : « <https://s.mj.run/zyMlw-D7AIM> the wall theory of AI --chaos 10 --ar 3:4 --style raw --sref <https://s.mj.run/zyMlw-D7AIM> --cref <https://s.mj.run/zyMlw-D7AIM> --weird 300 --v 6.1. »



Vous avez trouvé cette lecture utile ?

Vous aimerez sûrement aussi :

Chat-Geppetto

Façonner la Conscience Artificielle

Chat-Geppetto explore la notion de conscience artificielle (CA) et ses implications éthiques et technologiques. Contrairement à l'intelligence artificielle traditionnelle, la CA vise à doter les machines d'une perception de soi et d'une capacité d'adaptation contextuelle, permettant une meilleure efficacité et flexibilité. Toutefois, Keyrus met en garde contre les risques d'une autonomie décisionnelle complète et prône une approche de «conscience assistée», où les systèmes restent sous contrôle humain. En utilisant la métaphore de Pinocchio, l'ouvrage illustre la vision d'une IA qui aide et guide sans jamais s'émanciper, garantissant un usage éthique et sécurisé de ces technologies avancées.

Quelles sont les 3 idées principales ?

- 1. Conscience assistée** : la conscience artificielle doit améliorer l'adaptabilité et la performance des machines tout en restant sous supervision humaine, sans atteindre l'autonomie décisionnelle complète.
- 2. Éthique et régulation** : encadrer la CA avec des garde-fous est essentiel pour éviter les dérives, garantir un usage sécurisé et éviter l'illusion de la fiabilité totale.
- 3. Métaphore de Pinocchio** : comme Geppetto façonnant un pantin, les créateurs de CA développent des systèmes conscients, mais qui ne doivent jamais devenir autonomes, soulignant l'importance d'une approche mesurée et contrôlée.



keyrus TAB
MOBILE CLOUD TRUSTED Tech Advisory Board

Chat-Geppetto

Façonner la Conscience Artificielle

www.keyrus.com



keyrus

make data matter

Acteur international du conseil et des technologies, Keyrus a pour mission de donner du sens aux données, en révélant toute leur portée, notamment sous un angle humain.

Parce que ce ne sont pas tant les données elles-mêmes qui importent, mais les opportunités que nous pouvons développer en les apprivoisant vraiment, nous nous efforçons constamment de comprendre les objectifs que nos clients souhaitent atteindre. Nous explorons et mesurons les comportements, nous les comprenons et les traduisons en un résultat concret. Nous donnons un sens aux réalités que les données portent afin d'aider nos clients à prendre des décisions plus efficaces.

Les données, qu'elles soient grandes, petites, humaines, complexes, historiques ou prospectives, n'ont de sens que lorsqu'elles sont utilisées pour développer les expériences, affiner la compréhension du quotidien et prendre les meilleures décisions.

Notre proposition de valeur est fondée sur cinq grands groupes de services, chacun comprenant des offres multiples :

- **Automatisation et intelligence artificielle** : nous fournissons à nos clients les moyens d'améliorer leur productivité et leur précision sur l'ensemble de leurs processus, afin de se concentrer sur le travail à plus forte valeur ajoutée.
- **Expérience numérique centrée sur l'humain** : la relation avec les clients et l'engagement des collaborateurs constituent deux des plus grands contributeurs au succès global des entreprises. Nous aidons les entreprises à imaginer et à créer des expériences numériques multimodales et fluides pour atteindre leurs objectifs.
- **Mise en œuvre des données et des analyses** : les données sont une clé incontestable du succès pour les entreprises. Lorsqu'elles sont utilisées intelligemment, elles ouvrent des opportunités uniques pour faire face aux défis actuels et futurs. Nous permettons aux organisations de déployer tout le potentiel de leurs données : nous mettons la science des données au profit du développement de l'entreprise.
- **Cloud et sécurité** : le Cloud et les plateformes numériques ont le potentiel de révolutionner la façon dont les données sont transformées en valeur, tout en portant l'extensibilité et la flexibilité à un niveau supérieur. Nous sécurisons l'ensemble de vos données et veillons à ce qu'elles soient protégées et confidentielles.
- **Transformation et innovation** : pour prospérer dans l'écosystème actuel, chaque entreprise doit non seulement accélérer sa transformation numérique, mais aussi acquérir des compétences pour stimuler son adaptabilité, sa résilience et sa compétitivité. Nous aidons nos clients à se transformer avec succès pour développer un meilleur futur.

S'appuyant sur l'expérience cumulée de plus de 3 500 collaborateurs et présent dans 27 pays sur 4 continents, Keyrus est l'un des principaux experts internationaux en matière de données, de conseil et de technologie.

Pour en savoir plus : www.keyrus.fr

Jean-Philippe CLAIR
Directeur Marketing, Communication & Expérience client
jean-philippe.clair@keyrus.com