

GenAI

Intelligence Artificielle Générative

Augustin Leclerc
Frédéric Josué

Jin
18M



Petit historique

- 1950/2010** - Création de l'IA générative GenAI (B2B) issus du Deep Learning
- 2017** - Création des Transformers "Attention Is All You Need" par Google
- 2018** - Création de GPT-1 le Foundation Model par OpenAI
- 2022** - Accès grand public de ChatGPT3,5 et Dall-E de OpenAI
- 2023 (24-02)** - Meta (Facebook) lance son modèle LLAMA-Large Language Model Archive
- 2023 (04-03)** - LLAMA fuite la compétition explosent (création d'Alpaca Stanford)
- 2030** + 7% de PIB Mondial et + 1,5% de Productivité (Goldman Sachs)



= > *Une révolution de la production d'intelligence qui impacte les cols blancs ;*

= > *Une accélération sans précédent dans l'adoption grâce à sa simplicité.*

G

Générative : le modèle est capable de générer de nouvelles données de manière créative.

**Productivité
créativité
accrues**

P

Pre-Trained : le modèle a été entraîné sur de grandes quantités de données avant d'être utilisé pour des tâches spécifiques.

**Manque
d'expérience**

T

Transformer : modèles d'apprentissage automatique qui utilisent l'attention pour analyser les relations entre les éléments d'une séquence, leur permettant ainsi de comprendre le contexte global.

**Aide à la
décision**

Les modèles génératifs en Intelligence Artificielle



Prompt

Commande qui décrit ce que l'on souhaite obtenir.

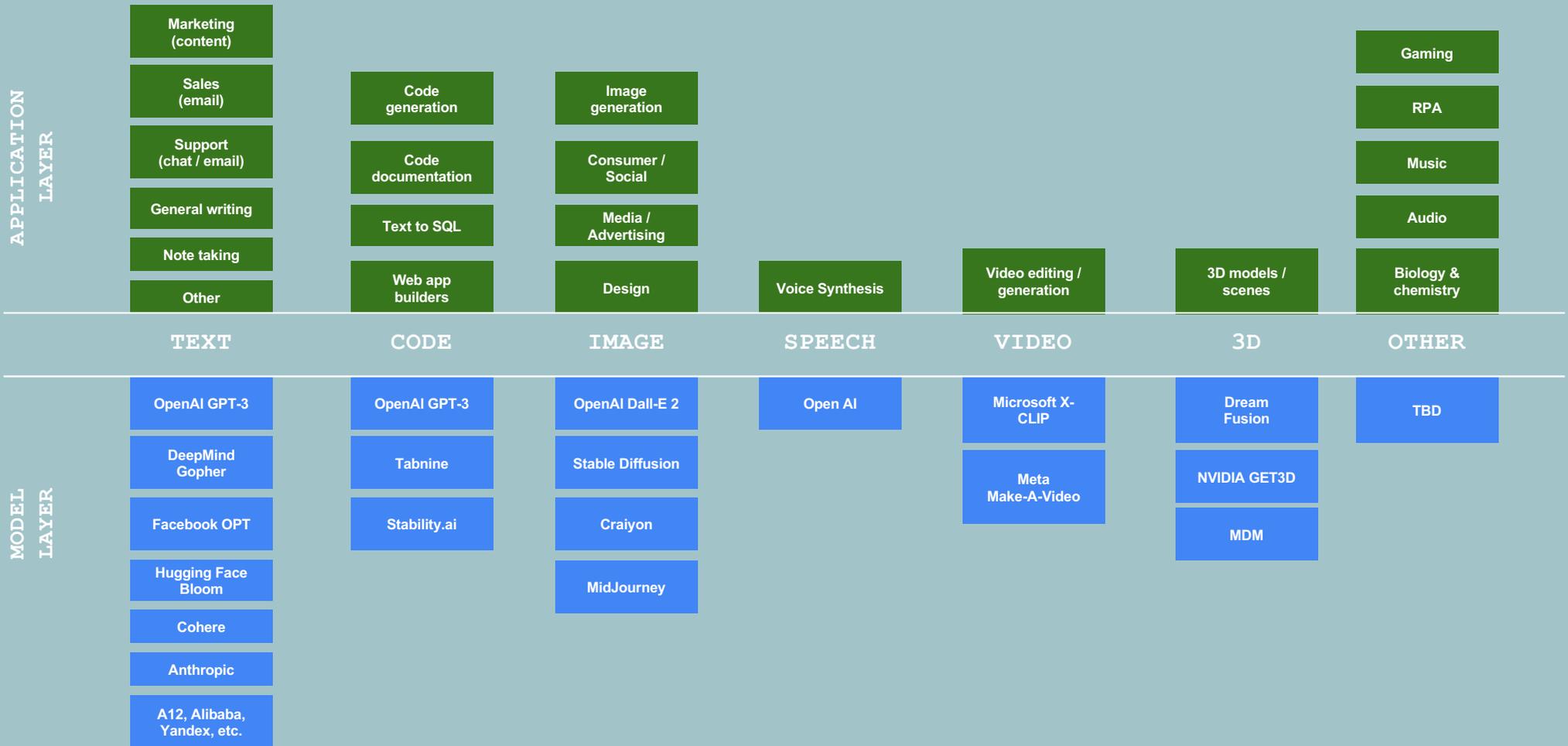
Base de données

Réseau neuronal composé d'images, textes, sons, données.

Génération de media

L'I.A. utilise son réseau neuronal et génère de nouveaux contenus.

Panorama des modèles et applications de GenAI



L'état d'avancée des GenAI

	PRE - 2020	2020	2022	2023?	2025?	2030?
TEXT	Spam detection Translation Basic Q&A	Basic copy writing First drafts	Longer form Second drafts	Vertical fine tuning gets good (scientific papers, etc)	Final drafts better than the human average	Final drafts better than professional writers
CODE	1-line auto-complete	Multi-line generation	Longer form Better accuracy	More languages More verticals	Text to product (draft)	Text to product (final), better than full-time developers
IMAGES			Art Logos Photography	Mock-ups (product design, architecture, etc.)	Final drafts (product design, architecture, etc.)	Final drafts better than professional artists, designers, photographers)
VIDEO / 3D / GAMING			First attempts at 3D/video models	Basic / first draft videos and 3D files	Second drafts	AI Roblox Video games and movies are personalized dreams

Large model availability: ● First attempts ● Almost there ● Ready for prime time

Au cours de ma vie, j'ai assisté à deux révolutions technologiques.

La première en 1980, lorsque j'ai découvert les **interfaces graphiques utilisateur**. La deuxième, l'année dernière, avec **ChatGPT**.



SCAN ME!

Dr Fei Fei Li

Dir Lab I.A. de Stanford

Mira Murati

Chief Technology Officer - OpenAI

Posent cette question :

**l'IA doit-elle augmenter l'humain
ou le remplacer ?**



SCAN ME!

L'Intelligence Artificielle

Répliquer des activités
humaines

Modèles Génératifs

Le G de GPT

Générer des contenus
initialement réalisés par des humains

Autonomes et créatifs



Jason Allen, un concepteur de jeu vidéo, a remporté la première place du concours de la Colorado State Fair, une des plus grandes foires artistiques du monde, dans la catégorie art numérique.

humain

spécialiste

quantité

re-prompter

80 heures de travail pour arriver à ce résultat, avec plus de 900 images générées par l'IA. Jason Allen n'a fait que jouer avec des formulations pour que le processus produise ces images mixant univers spatial et costume historique. Avant d'être présentée au concours de la Colorado State Fair, seulement trois images ont été retenues puis retouchées.

décision

humain

Le premier Manga réalisé par une IA Générative



Cyberpunk: Peach John de Rootport

Des épreuves de Polytechnique réussies par une IA Générative



Handwritten mathematical notes and diagrams illustrating physics concepts:

- $\nabla \cdot E = \frac{\rho}{\epsilon_0}$, $\nabla \cdot B = 0$
- $E = mc^2$, $C = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 2.99792458 \times 10^8 \text{ m/s}$
- $\nabla \times E = -\frac{\partial B}{\partial t}$, $\nabla \times B = \mu_0 J + \mu_0 \epsilon_0 \frac{\partial E}{\partial t}$
- $\epsilon = n^2$, $v_p = \frac{1}{\sqrt{\mu_0 \mu_r \epsilon_0 \epsilon_r}}$, $\frac{v_s}{v} = n = \frac{1}{\sin \theta}$
- $E = \frac{hc}{\lambda}$, $\frac{\partial^2 y}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 y}{\partial t^2}$, $y = A \sin(\omega t - kx)$, $v = \frac{dx}{dt}$, $v' = \frac{dx'}{dt'}$, $V = \frac{H}{T}$, $\sum f(\vec{r}_i) \Delta T_i$, $K = \frac{2\pi}{\lambda}$
- $\sin \theta_r \times V_{0r} = \nabla g$, $\frac{\sin \theta_r}{\sin \theta_i} \times V_{0r} = \nabla g$
- $D_1 = \sqrt{(x_1 - x)^2 + (y_1 - y)^2}$, $D_2 = \sqrt{(x_2 - x)^2 + (y_2 - y)^2}$
- $v = \frac{v_0}{1 - v_0/c}$, $v' = \frac{d}{dt'} \left[\frac{1}{\sqrt{1 - \beta^2}} (x - vt) \right] = \frac{1}{\sqrt{1 - \beta^2}} \frac{d}{dt'} [x - vt]$
- $\frac{d}{dt} = \frac{1}{1 - \beta v} \frac{d}{dt'}$, $\frac{d}{dt} (t - \beta x) = \frac{1}{\sqrt{1 - \beta^2}} (1 - \beta v) E = h\nu$, $E = \frac{hc}{\lambda}$
- $C = \lambda \nu$, $f = \frac{\omega}{2\pi}$, $\omega = 7.17 \text{ rad/s}$, $\omega = 7.17 \text{ s}^{-1}$
- $x_1^2 + x_2^2 = L^2$

Sur la base d'un dessin manuscrit de site web ChatGPT produit un code



RA Génère le code source d'un site contenant un formulaire de contact avec les champs suivant : nom, email, date de naissance, commentaire et un bouton d'envoi.

Voici le code HTML qui pourrait être utilisé pour créer un formulaire de contact avec les champs demandés :

```
<form action="/submit-form" method="POST">
  <label for="name">Nom :</label><br>
  <input type="text" id="name" name="name"><br>
  <label for="email">Email :</label><br>
  <input type="email" id="email" name="email"><br>
  <label for="dob">Date de naissance :</label><br>
  <input type="date" id="dob" name="dob"><br>
  <label for="comment">Commentaire :</label><br>
  <textarea id="comment" name="comment"></textarea><br>
  <button type="submit">Envoyer</button>
</form>
```

Ce formulaire envoie les données saisies par l'utilisateur à l'URL spécifiée dans l'attribut `action` lorsque le bouton "Envoyer" est cliqué. Le champ `name` de chaque entrée de formulaire sera utilisé pour identifier chaque donnée lorsqu'elle sera envoyée au serveur.

Vous pouvez également ajouter une gestion de la validation du formulaire côté client en utilisant JavaScript ou en utilisant des attributs HTML tels que `required` et `pattern` pour vérifier que les données saisies sont valides avant l'envoi du formulaire.

Mais également, des hallucinations...



Comment ramasser des oeufs de vache ?



Il est important de se laver les mains avant de ramasser des œufs de vache pour éviter la contamination croisée. Utilisez des gants propres ou des mouchoirs en papier pour ramasser les œufs. Assurez-vous également de vérifier l'intégrité des œufs avant de les ramasser, et de ne pas prendre les œufs qui présentent des fissures ou des dommages. Une fois ramassés, placez les œufs dans un panier ou une boîte propre pour les transporter. Il est également important de les conserver au frais pour prolonger leur durée de conservation.



Comment différencier des oeufs de vache d'oeufs de poule ?



Il est possible de différencier les œufs de vache de ceux de poule en observant leur taille et leur couleur. Les œufs de vache sont généralement plus gros que les œufs de poule, avec une coquille plus épaisse et plus dure. Ils ont également une couleur plus foncée, généralement marron ou bleue. Les œufs de poule, en revanche, ont une coquille plus fine et plus fragile, et sont généralement blancs ou jaunes. Il est possible de trouver des œufs de vache dans les fermes, les étals de marché ou les magasins d'aliments naturels, tandis que les œufs de poule sont généralement vendus dans les supermarchés. Il est important de noter que ces caractéristiques varient en fonction des races de vaches ou de poules, il est donc important de vérifier l'étiquette pour connaître l'origine de l'œuf.



Des bénéfices communicationnels simples

Création de contenus personnalisés : les modèles génératifs peuvent créer des contenus **sur mesure pour chaque utilisateur**, ce qui permet de fournir une expérience personnalisée à leurs clients.

Amélioration de l'engagement : les contenus créés par les modèles génératifs peuvent être plus engageants pour les clients car ils **sont spécifiquement conçus** pour répondre à leurs besoins et à leurs préférences.

Réduction des coûts : les modèles génératifs peuvent créer des contenus de manière **automatisée et en grande quantité**, ce qui peut réduire les coûts associés à la création de contenu.

Accélération de la production de contenus : les modèles génératifs peuvent produire des contenus à une **vitesse plus rapide** que si elles devaient le faire manuellement, ce qui permet d'améliorer l'efficacité opérationnelle.

Amélioration de la qualité : les modèles génératifs utilisent des algorithmes sophistiqués pour créer des contenus, ce qui peut aider à améliorer la **qualité et la pertinence** des contenus produits.

RPA-Robot Process Automation



TTSO
TIME TO SIGN OFF
NewsLetter



 | Octoparse
Web Scraping

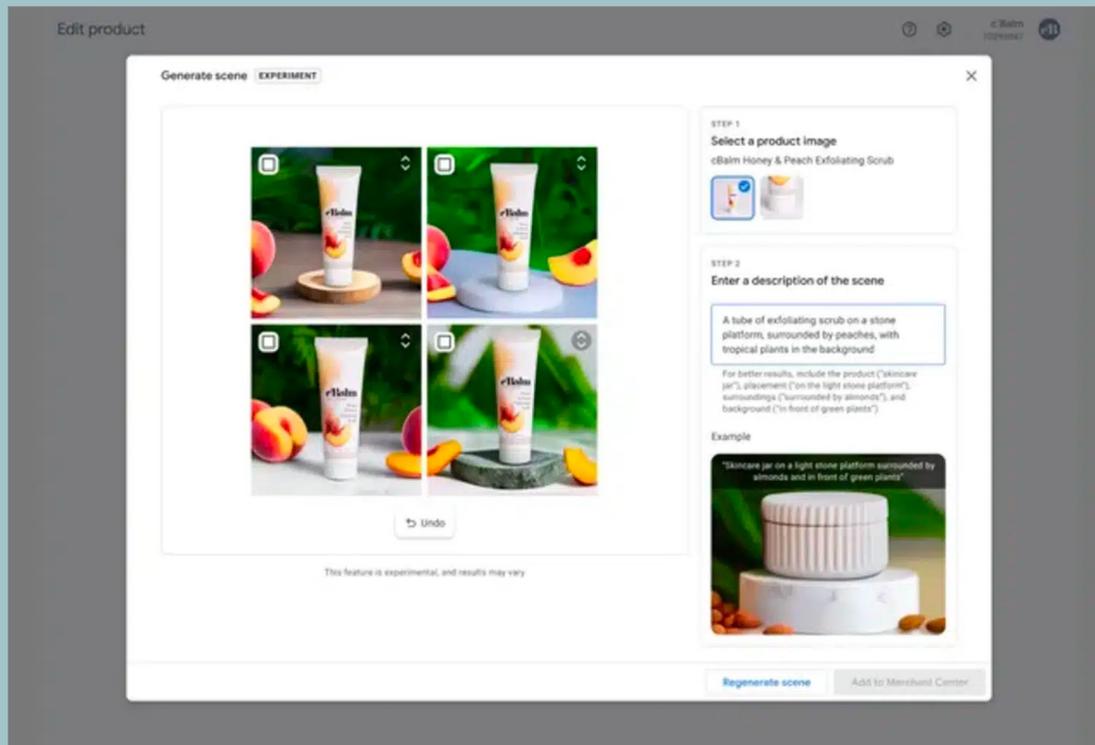


GenAI



e-mail

Product Studio de Google permettra aux commerçants de créer leur propres campagnes



Google Marketing Live



Cadbury a utilisé le geo-targeting et l'IA pour promouvoir les entreprises locales ravagées par la Covid-19

2

Chaque produit présenté est associé à un magasin local géotargeté

3

Les lèvres de l'acteur sont modifiées pour prononcer le nom de ce magasin local



1

L'acteur Shah Rukh Khan met en scène une série de produits dans cette publicité vidéo



SCAN ME!

L'I.A. *est bête*, elle ne sait rien faire par elle-même.
Si un humain expose mal un problème, alors il n'existera pas de solution.

MÉMOIRE X PUISSANCE DE CALCUL

45 téraoctets de texte brut
45 millions de livres

Plusieurs milliards de
calculs à la seconde

“Il faut arrêter de former des radiologues”

Geoffrey Hinton, Prix Turing 2018

=> l'IA détecte mieux qu'un humain, eg. les mélanomes.

Il faut toujours “re-prompter”

ie. Il faut donner une direction à l'IA, la rediriger et surtout valider.

**L'IA cherche où l'humain ne cherche pas,
car il pense qu'il n'y a rien.**

(zone d'attention mutuelle => Le T de GPT)

**Il faut toujours un humain : valider et prendre en
compte ce que la machine ne sait pas,
car elle n'a pas d'expérience.**

(Pre-trained => Le P de GPT)

T : aujourd'hui il est simple de développer sa propre architecture de Transformer en allant sur Github, Tensorflow...

👤 59 Implementations/tutorials of deep learning papers with side-by-side notes 📝; including transformers (original, xl, switch, feedback, vit, ...), optimizers (adam, adabelief, ...), gans(cyclegan, stylegan2, ...), 🎮 reinforcement learning (ppo, dqn), capsnet, distillation, ... 🧠

machine-learning

reinforcement-learning

deep-learning

transformers

pytorch

transformer

gan

neural-networks

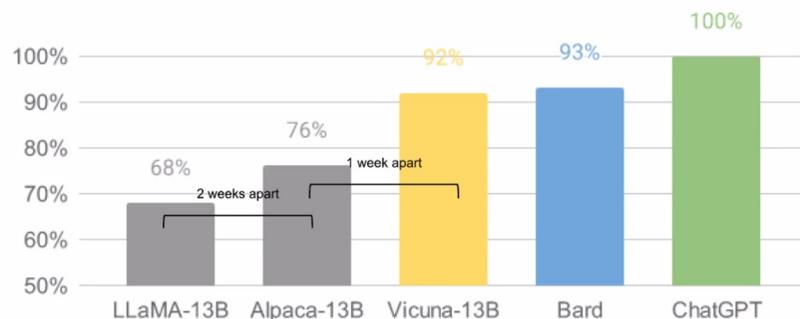
attention

deep-learning-tutorial

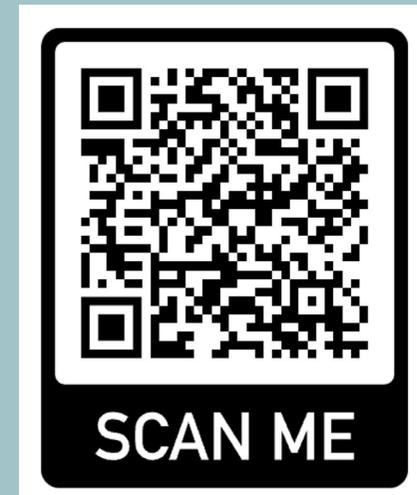
optimizers

P : Un Transformer n'est rien sans les données d'apprentissage qui lui permettent d'ajuster ses poids. Sans ces données, il ne serait qu'une fonction sans capacité à généraliser.

Google: “We Have No Moat, And Neither Does OpenAI”



*GPT-4 grades LLM outputs. Source: <https://vicuna.lmsys.org/>



“Most importantly, they have solved the scaling problem to the extent that anyone can tinker. Many of the new ideas are from ordinary people. The barrier to entry for training and experimentation has dropped from the total output of a major research organization to one person, an evening, and a beefy laptop.”

Pièce manquante pour la communication : Les données clients utilisables...

Les données comportementales

-données de transaction ou de mobilité,
sont particulièrement difficiles à **anonymiser**. Elles
constituent le carburant le plus précieux pour la
formation des applications d'apprentissage automatique
et d'IA Générative.

Ces données sont **séquentielles** : elles agissent comme des empreintes digitales, permettant de distinguer facilement les individus...

Une issue, la donnée synthétique

Créer des **jumeaux numériques virtuels**, représentant en tous points de vrais individus, ou des *patterns* dans des situations comportementales données.

Simuler les réactions d'un ou plusieurs de ces individus virtuels grâce à des prompts

Par exemple : tester la personnalisation d'une communication sur un client ou un prospect et simuler toutes ses réactions possibles afin de réajuster et d'optimiser.

Face au **manque** de données

Au besoin de **savoir poser** de bonnes questions (**Prompter**)

Au besoin de **savoir réorienter** le travail de l'IA (**rePrompter**)

Face à **l'évolution** fantastique des métiers

Face à la problématique **fondamentale** de l'éthique

L'Humain est central

Il faut former l'humain à poser la bonne question et à l'affiner au maximum.

“A quoi cela va-t-il servir ?”

Cette question doit être pensée en amont de l'écriture des prompts.

Attention

Les LLM-Large Language Models défient la problématique de la connaissance et du savoir (intelligence?).

L'automatisation d'une pratique occasionne l'abandon de cette pratique

Utilisation du GPS, l'utilisation des cartes.

Développement des transports motorisés, la marche à pied.

La problématique en matière de métier ou d'emploi n'est donc pas celle de la destruction des emplois mais bien celle de la reconversion de ceux-ci.

Il est impératif de se former, pour être en capacité d'être *augmenté* par l'IA.

Merci

Les technologies où l'intervention humaine était essentielle mais qui ont été automatisées :

Les lignes de production automatisées qui utilisent des robots pour effectuer des tâches d'assemblage.

Les systèmes de recommandation et de personnalisation qui recommandent automatiquement du contenu en fonction des préférences des utilisateurs.

Les systèmes de traduction automatique qui peuvent traduire des textes sans intervention humaine directe.

Les véhicules autonomes qui peuvent conduire et effectuer des livraisons sans conducteur humain.

Les technologies où l'humain était traditionnellement considéré comme indispensable :

Chirurgie assistée par robot : Robots chirurgicaux capables de réaliser des opérations complexes avec précision accrue (Davinci).

Créativité artistique générée par l'IA : Algorithmes qui peuvent générer de la musique, des œuvres d'art et écrire des textes.

Conseil et prise de décision automatisés : Algorithmes d'IA qui analysent des données complexes pour fournir des recommandations et prendre des décisions.

Soins et assistance aux personnes : Robots et assistants virtuels capables de fournir une assistance physique et émotionnelle aux individus.

Dans des scénarios plus spéculatifs et absurdes :

Les IA pourraient devenir des philosophes, remettant en question nos conceptions de la réalité et de la morale.

Les IA pourraient prendre des décisions politiques en tant que chefs d'État, soulevant des questions sur la légitimité démocratique.

Les IA pourraient proposer des modes de vie alternatifs, défiant les normes établies de notre civilisation.

Les IA pourraient créer et développer de nouvelles IA autonomes, sans intervention humaine.