

ChatGPT et IA générative pour les nuls et les moins nuls

19/03/2024



Source: livre blanc de l'IA, par INRIA

Qu'est-ce que ChatGPT?

ChatGPT est un prototype d'agent conversationnel utilisant des algorithmes d'Intelligence Artificielle (IA)

Il est basé sur de grands modèles de langages (*Large Language Models* en anglais, ou LLM), à savoir des modèles informatiques et mathématiques des langues, qui sont améliorés en continu: GPT3, GPT3.5, GPT4, bientôt GPT5

L'outil est disponible en plusieurs langues, dont le français, l'anglais, le chinois, ou encore le japonais.

ChatGPT a été réalisé par une entreprise américaine, *OpenAI*, spécialisée en Intelligence Artificielle, et qui développe également d'autres outils intégrant de l'IA



Le logo de ChatGPT

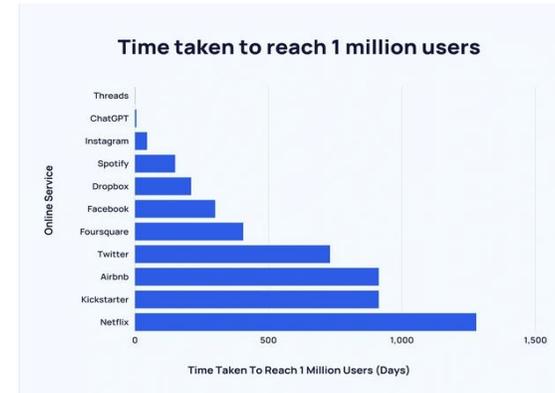


Le PDG d'OpenAI, Sam Altman

ChatGPT: un succès extraordinaire!

Lancé en novembre 2022, ChatGPT a rencontré un grand succès:

- 100 millions de comptes enregistrés en janvier 2023.
- +1561% de fréquentation du site d'*OpenAI* entre avant et après le lancement de ChatGPT
- 660 millions de visiteurs sur le site d'*OpenAI* en janvier 2023.



Source:
<https://explodingtopics.com/blog/chatgpt-users>

On note aussi des critiques, notamment sur l'exactitude des informations produites par l'outil.

Jean-Noël Barrot, lundi 20 février 2023: « à ce stade ChatGPT n'est qu'un perroquet approximatif qui restitue parfois un peu maladroitement les sommes astronomiques d'informations qu'il a compilées sur Internet. »



INDIA TODAY



Le succès de l'outil a eu d'énormes conséquences au niveau mondial:

- Les grandes sociétés sortent en urgence leurs propres outils conversationnels: *Bard*, *Gémini*, *Grok*, *Ernie*, etc.
- Certains pays repensent leur stratégie en intelligence artificielle (Japon notamment)
- Grande accélération des débats relatifs à la régulation de l'Intelligence Artificielle (comme au G7 Numérique).
- Des administrations ont adopté ChatGPT

the japan times

Tokyo Metropolitan Government to start using ChatGPT from August

BY SHINJI / Jun 18, 2023

The Tokyo Metropolitan Government will begin using the artificial intelligence chatbot ChatGPT for writing texts and carrying out other clerical work in all of its offices from August, Gov. Yuriko Koike said Tuesday.

ChatGPT "has the potential to greatly transform the way public administration is conducted," Koike said during a metropolitan assembly session. She added that "better city governance" can be achieved by assessing the positive and negative aspects of the AI service.

Koike also said the metropolitan government will use ChatGPT for tasks including preparing documents in question-and-answer format, and seek input from its employees about other practical uses for the generative AI tool.

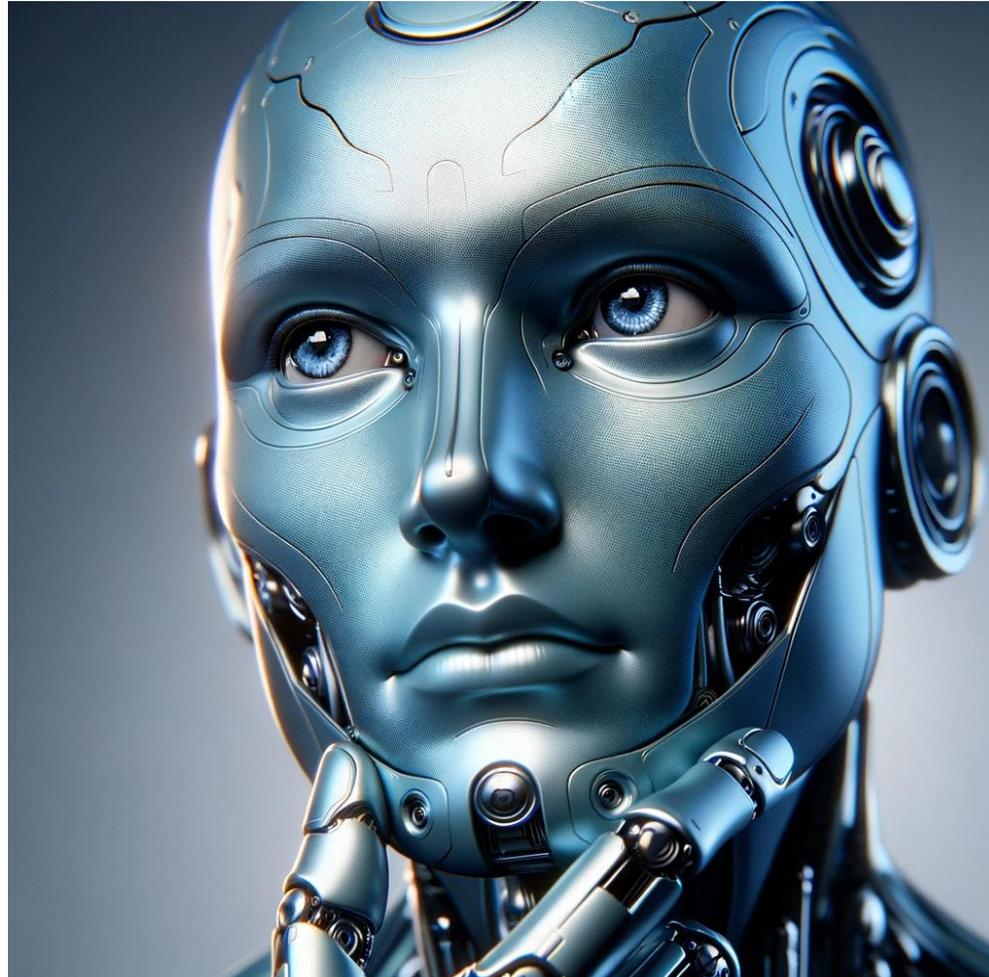
Chatbots are software applications trained using massive amounts of data from the



Plan de la présentation

1. Fonctionnement de l'IA générative
2. Utilisation pratique de ChatGPT
3. Quelques tendances et perspectives sur « l'IA générative »
4. La politique S&T et l'écosystème japonais de l'IA.

1. Fonctionnement de l'IA générative



Portrait de ChatGPT, par lui-même

D'où vient l'IA générative?

L'Intelligence Artificielle est un très vaste domaine:

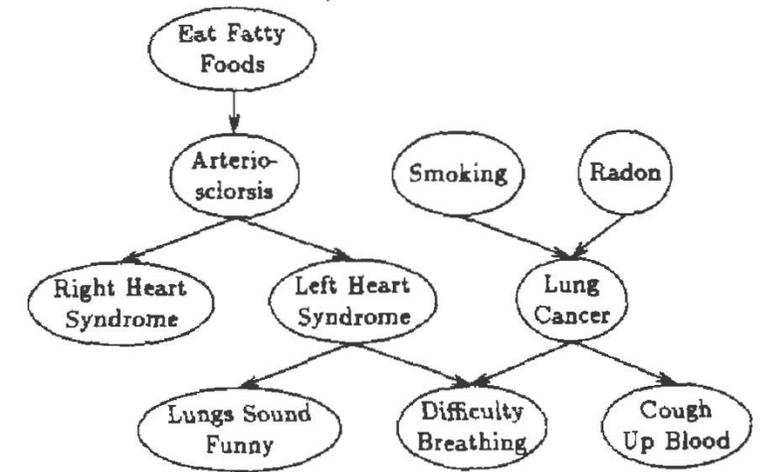
- Neurosciences
- Représentation de connaissances
- Robotique
- Etc. ...



Pour comprendre ce qu'est l'IA générative, attardons-nous sur deux approches particulières: les approches **symboliques** et les approches par **réseaux de neurones artificiels**

Approche symbolique de l'IA

- C'est une approche « descendante » de l'IA
- On interroge des experts d'un domaine, qui tentent de transmettre leurs connaissances à une machine afin que celle-ci puisse se comporter comme un expert du domaine et répondre à des questions
- Basé sur de la logique déductive.
 - Exemple: (Fièvre) et (toux) et (maux de tête) => rhume
- Approche en baisse de régime depuis vingt ans



« Système expert »

Les réseaux de neurones

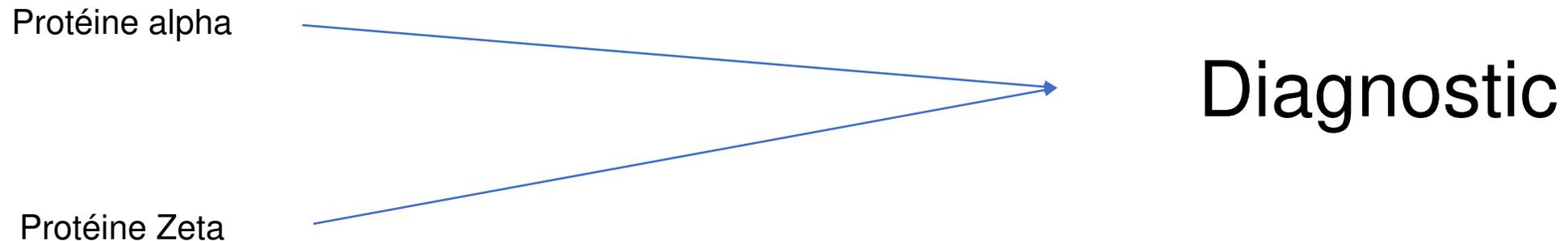
- Approche « ascendante » de l'IA
 - Généralisation à partir de données sur lesquelles « s'entraîne » un modèle
- Basée sur une représentation « formelle » du neurone, introduite en 1943
 - Lointaine ressemblance à un vrai neurone
 - Utilisation très simple et cas d'application très généraux
- Toutes les applications modernes de l'IA (dont ChatGPT), utilisent des réseaux de neurones

Neurone formel: un exemple simple

- Inspiré de notre façon de raisonner :
 - On combine plusieurs critères et ensuite on prend une décision

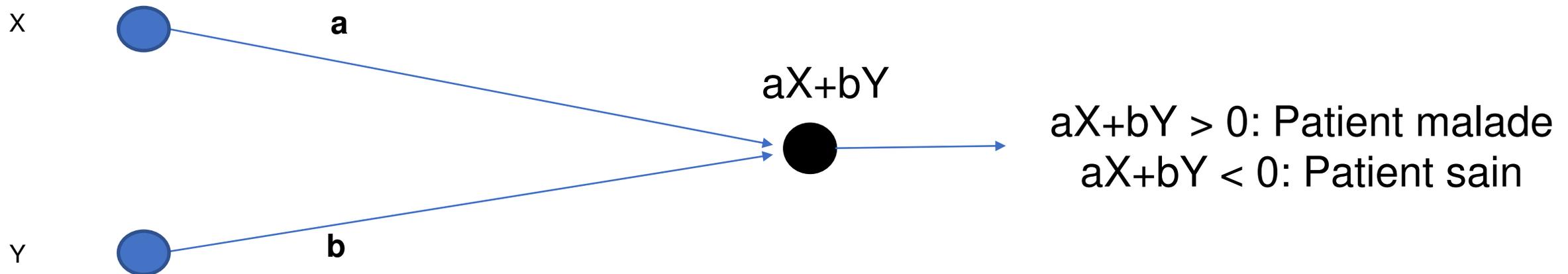
Neurone formel: un exemple simple

- Inspiré de notre façon de raisonner :
 - On combine plusieurs critères et ensuite on prend une décision
 - Par exemple: diagnostic à partir d'une prise de sang



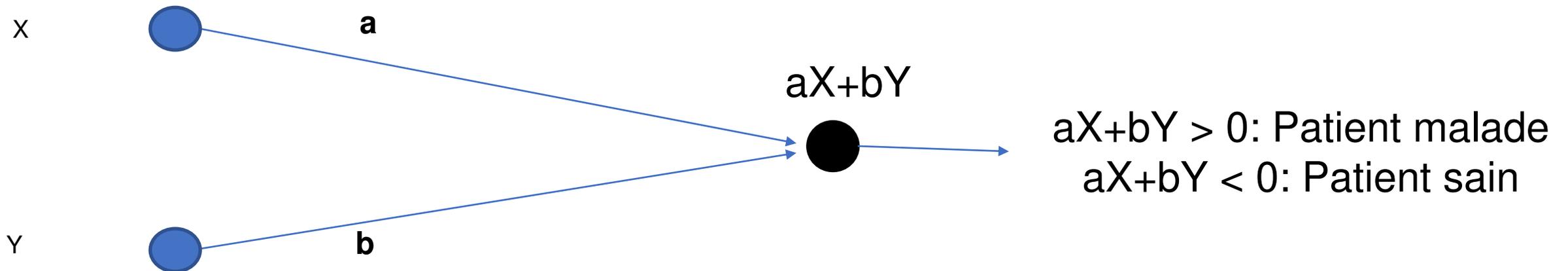
Neurone formel: un exemple simple

- Inspiré d'un vrai neurone et de notre façon de raisonner :
 - On combine plusieurs critères puis on prend une décision
 - Par exemple: diagnostic automatisé à partir d'une prise de sang



Neurone formel: un exemple simple

- Inspiré d'un vrai neurone et de notre façon de raisonner :
 - On combine plusieurs critères et ensuite on prend une décision
 - Par exemple: diagnostic automatisé à partir d'une prise de sang

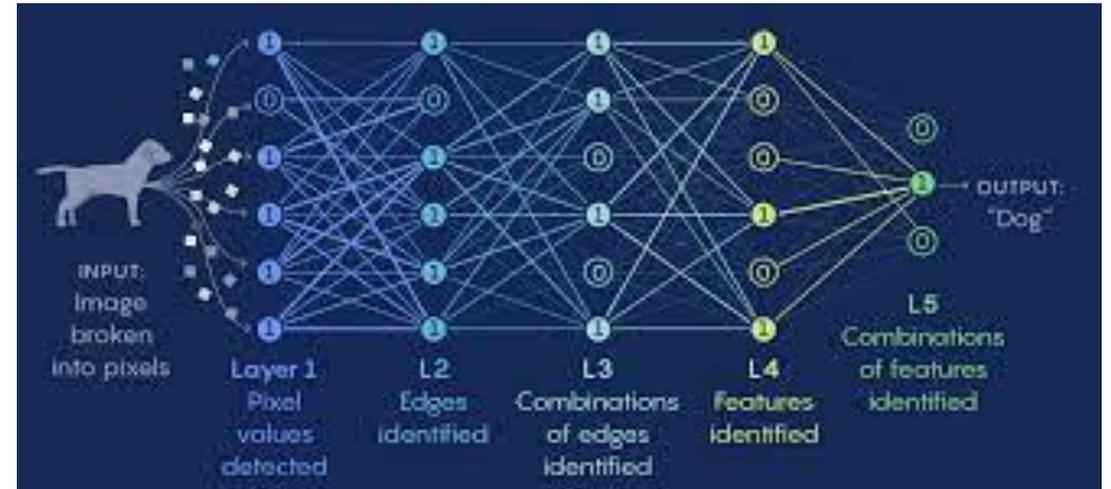


a et b sont les « paramètres » du modèle. Ils doivent être calculés par **apprentissage**

Les réseaux de neurones

Dans la réalité, les neurones sont empilés en couches

=> Nombre potentiellement très grand de paramètres



Depuis quinze ans, les grandes données accumulées grâce aux progrès de la numérisation, et l'augmentation de la puissance de calcul permettent de faire fonctionner les approches par réseaux de neurones de façon satisfaisante

Quelle que soit l'application, le principe est le même: Il s'agit de classifier en ensemble de données d'entrées en généralisant des exemples donnés:

- Des pixels de visage => reconnaissance ou non du propriétaire
- Un fichier son => une musique
- Un âge + un revenu mensuel => décision ou non d'accorder un prêt pour une banque
- Des pixels d'image => Une description en quelques mots
- Une liste de facteurs biologiques => un risque de développer une maladie cardio-vasculaire dans les cinq ans à venir, etc.
- ChatGPT n'échappe pas à la règle: à une liste de mots donnée (un prompt), il associe le mot suivant qui s'enchaîne le mieux

Forces et faiblesses des réseaux de neurones

+

- Passage à l'échelle
- Bons résultats en pratique

-

- Généralise mal sur des exemples éloignés des données d'apprentissage
- Peu transparent
- Fort coût énergétique

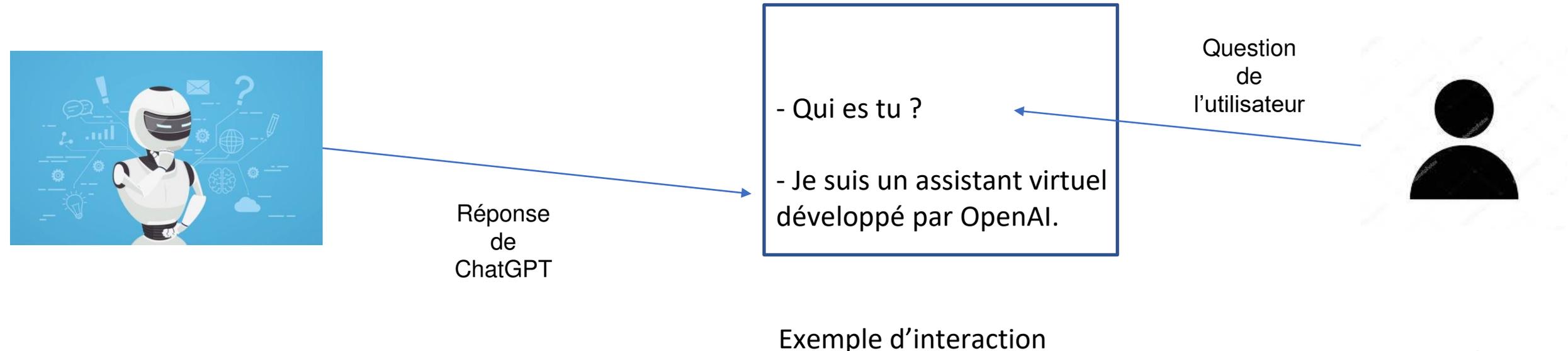
2. Comment utiliser (et aussi ne pas utiliser) ChatGPT



Source: *Cartooning for peace*

Que fait vraiment ChatGPT?

- ChatGPT est un outil de complétion de texte (survitaminé...)
- Etant donnée une requête, appelée « prompt », ChatGPT complète le texte, mot après mot.



Que fait vraiment ChatGPT?

- ChatGPT est un outil de complétion de texte (survitaminé...)
- Etant donnée une requête, appelée « prompt », ChatGPT complète le texte, mot après mot.



Point de vue de ChatGPT

Texte
complété

Texte
À compléter

- Qui es tu ?
- Je suis un assistant virtuel
développé par OpenAI.

Exemple d'interaction

Question

Réponse

Point de vue de l'utilisateur

Conditions d'utilisation: les points clés

Un mineur doit avoir l'autorisation de ses parents pour l'utiliser

Ne peut pas être utilisé pour des activités illégales

OpenAI cède les droits d'utilisation de ce qui sort de ChatGPT à quiconque utilise le service

Les interactions qu'on a avec ChatGPT sont utilisées pour améliorer le service. Donc, elles fuient...

=> Ne jamais entrer d'informations confidentielles dans ChatGPT

On déclare comprendre que ce qui sort de ChatGPT peut être inexact

=> Toujours vérifier ce qu'écrit ChatGPT

Coût de ChatGPT

Modèle de base: gratuit

ChatGPT Plus: 20 dollars par mois

Donne accès à l'interface utilisateur.

Donne un accès anticipé aux meilleurs modèles, et un accès prioritaire lors des pics d'utilisation

Possibilité d'avoir accès aux outils annexes (GPT Vision, Dall-E, etc.)

Accès à l'interface de programmation (pour développeurs):

Paiement au nombre de mots (ou « tokens »), en fonction du modèle utilisé

Cas d'utilisations intéressants

- Recherche d'idées
- Traduction de textes
- Résumé de texte / développement de texte
- Tâches créatives diverses:
 - Aide à la rédaction de mél
 - Aide à la rédaction de code informatique / html ...
- Recherche d'informations sur internet avec Bing

ChatGPT a une « **fenêtre de contexte** » limitée, qui correspond au nombre de mots qu'il est capable de prendre en compte pour faire une réponse.
=> sur de longs documents, la cohérence du texte tend à décroître
=> la taille de cette fenêtre dépend des modèles. Elle tend à augmenter à chaque nouvelle version de GPT. (GPT4 a une fenêtre de contexte de 128 000 mots)

Comment utiliser au mieux ChatGPT: le « prompt Engineering »

6 points à préciser, par ordre décroissant d'importance.

1. Tâche
2. Contexte
3. Format
4. Rôle
5. Ton
6. Exemple

Rien n'oblige à rédiger la requête dans cet ordre là.

Comment utiliser au mieux ChatGPT: le « prompt Engineering »

Tâche: préciser clairement ce qui est attendu en utilisant un verbe.

« Traduire le texte suivant:»

« Résumer le texte suivant:»

« Développer l'idée suivante: »

« Extraire les dates dans le texte suivant : »

« Trier les éléments suivants ... »

Comment utiliser au mieux ChatGPT: le « prompt Engineering »

Contexte:

- Avant de rédiger la tâche, on peut préciser dans quel contexte elle s'inscrit.
- Ce contexte peut être donné par un fichier (pdf, image, csv, etc.).

Par exemple:

« j'ai reçu le mél suivant de la part de X, qui occupe tel poste dans telle institution: *contenu du mél* . Rédiger une réponse pour accepter cette invitation ».

Comment utiliser au mieux ChatGPT: le « prompt Engineering »

Format: préciser la longueur attendue, ou si la réponse est attendue sous forme numérotée, etc.

ChatGPT peut aussi générer des fichiers spécifiques (.html, .json, code informatique) ou des images sous différents formats (jpg, png, etc.)

Comment utiliser au mieux ChatGPT: le « prompt Engineering »

Rôle: on peut préciser de qui part le message, pour aider ChatGPT à adapter la manière dont il va rédiger la réponse.

Par exemple:
rôle

Contexte

Tu es un professeur, avec en face de toi une classe de 20 étudiants de CE1.

Explique la réponse au problème suivant: « *contenu du problème* »

Tâche

en trois paragraphes numérotés correspondant aux étapes de la démonstration.

Format

Comment utiliser au mieux ChatGPT: le « prompt Engineering »

Ton: on peut préciser le style dans lequel on veut que la réponse soit rédigée: soutenue, familière, analytique, poétique, etc.

Exemple: On peut donner à la suite de la requête un exemple de réponse qu'on considère comme pertinente pour que ChatGPT s'en inspire.

3. ChatGPT et le monde du travail: Tendances et perspectives



Le monde du travail du futur par ChatGPT

L'IA: la voie vers une quatrième révolution industrielle?

- Tandis que dans les précédentes révolutions industrielles, les machines mécaniques se sont substituées à la **force physique** humaine, l'IA est sensée permettre de déployer des agents autonomes pour économiser la **force cognitive** humaine.
- Certains parlent de révolution comparable à l'invention de l'imprimerie, car l'IA générative permet de démultiplier la quantité de données pouvant être analysées et produites.
 - Documents scolaires personnalisés pour chaque étudiant ?
- Pour autant, les limitations actuelles (hallucinations, coût énergétique, etc.) poussent à un certain réalisme.
 - Mais la vitesse des progrès est, elle, très impressionnante!

Le Graal de la recherche en IA: *l'Artificial General Intelligence* (AGI)

- AGI: IA capable d'apprendre et d'exécuter n'importe quelle tâche cognitive propre aux humains
- Le temps restant avant qu'une AGI soit créée fait l'objet d'avis très divergents.
- Il n'existe pas à l'heure actuelle d'IA consensuellement considérée comme étant « générale ».
- De nombreux débats existent également sur la « conscience artificielle » des machines.
 - La conscience artificielle n'est cependant pas considérée comme un critère pour être considéré comme une AGI.



TECH • ARTIFICIAL INTELLIGENCE
Exclusive: U.S. Must Move 'Decisively' to Avert 'Extinction-Level' Threat From AI, Government-Commissioned Report Says
11 MINUTE READ

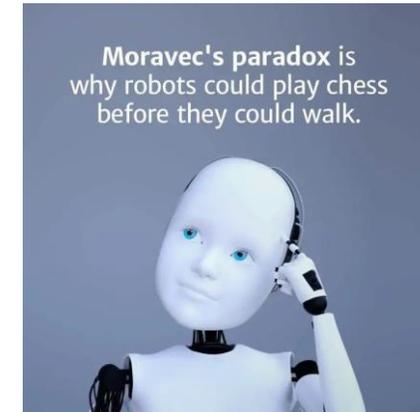


BY NIELY PEREZ FOR TIME MARCH 17, 2023 8:01 AM EDT

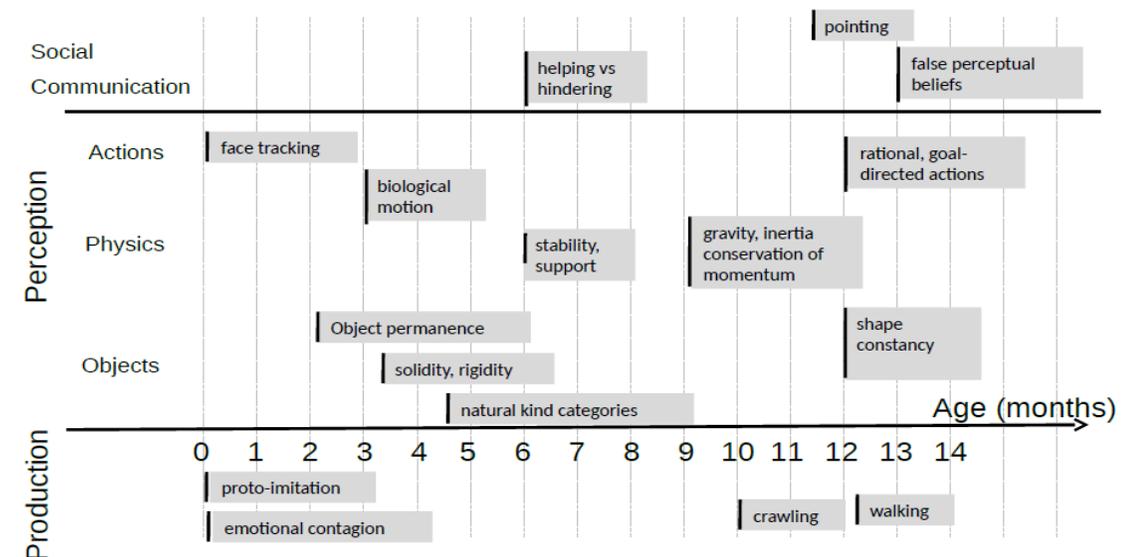
The U.S. government must move "quickly and decisively" to avert

Les défis à venir pour la recherche en IA

- **Paradoxe de Moravec** : « plus une tâche est facile, et plus elle est difficile à faire par un robot ». Toujours vrai aujourd'hui.
 - Sens commun et compréhension du monde
 - On ne sait pas vraiment ce qui est encodé dans les réseaux de neurones
 - Transfer learning: encore balbutiant
- **Fiabilité**:
 - Les « confabulations » des LLMs: imagination ou hallucination?
- **Capacité à planifier**
- **Capacité à raisonner**



Source: Medium, Février 2024



Les différents axes de recherche vers des IA au même niveau d'intelligence que l'humain

- Garder la même approche que les architectures de LLM actuels mais avec plus de données
 - Introduire du multimodal
- Trouver d'autres architectures de réseaux neuroneaux
 - Yann Lecun (JEPA)
- Faire un « système » de modules d'IA en interactions
 - Plugin Wolfram Alpha et ChatGPT
- IA « incarnée »
 - Faire découvrir le monde à un robot comme à un humain



Why next-token prediction is enough for AGI - Ilya Sutskever (OpenAI Chief Scientist)



Figure Status Update - OpenAI Speech-to-Speech Reasoning

La stratégie des grands acteurs de l'IA

- OpenAI: le leader du domaine mais des tensions visibles sur les objectifs de l'entreprise
- Microsoft veut intégrer ChatGPT dans ses suites logicielles
- Google a développé Bard et développe Gemini
- Apple a promis l'arrivée de son propre modèle, appelé Ferret.
- X développe *Grok*, mis en *open source* hier
- Amazon dit développer son propre LLM
- Meta (ex Facebook) a une approche « open source »: développement de Llama3 pour 2024



Copilot, assistant commandé par le langage naturel de Microsoft



Ferret, l'outil d'IA d'Apple



Llama, le modèle de langage de Meta³²

Au-delà de ChatGPT, une multitude de LLM

- D'autres *LLM en source ouverte* existent, dont certains ont des performances se rapprochant de ChatGPT
- Exemple: Mistral AI, nouvelle licorne française
- Ces modèles peuvent s'installer *localement* et donc ne pas faire fuiter d'informations
- La tendance est à l'entraînement de modèles plus petits (en nombre de paramètres), mais plus spécialisés.
 - Entraînés pour une tâche spécifique, sur des données métiers, pour produire un format précis



L'équipe de Mistral AI



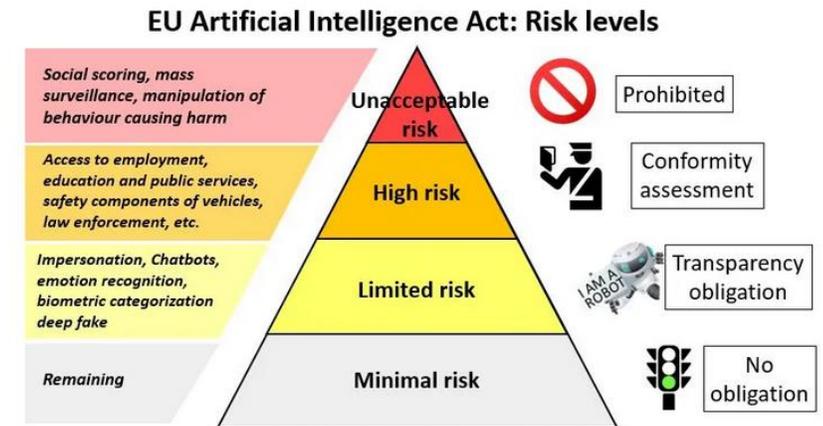
Perspectives sur l'utilisation des LLM

A l'avenir, des modèles d'IA installés « localement » pourraient être utilisés pour:

- Interroger nos données: plannings d'événements, tableaux de comptabilité, fiches institutions, compte-rendus d'entretiens et de réunions, rapports, annuaires de contacts
- Produire des brouillons de notes: « Produis moi une fiche de deux pages sur la stratégie du gouvernement japonais en Intelligence Artificielle avec le plan suivant: ...»
- On pourrait avoir des *chatbots*, entraînés comme experts sur certains sujets, qu'on pourrait solliciter à l'envie par *Tchap* lorsqu'on en a besoin ?

La régulation

- Les risques sont nombreux, en particulier:
 - Chômage de masse
 - Désinformation
 - Fuites de données
 - Vol de propriété intellectuelle
 - Cyberattaques
 - Armes létales autonomes
- Enjeux de la régulation:
 - Se prémunir des risques
 - Sans étouffer l'innovation
 - Et avec des technologies qui évoluent extrêmement rapidement.
- Approches très variables selon les pays.
 - La loi sur l'IA de l'UE est le tout premier cadre juridique sur l'IA au monde



(Source: Telefónica)

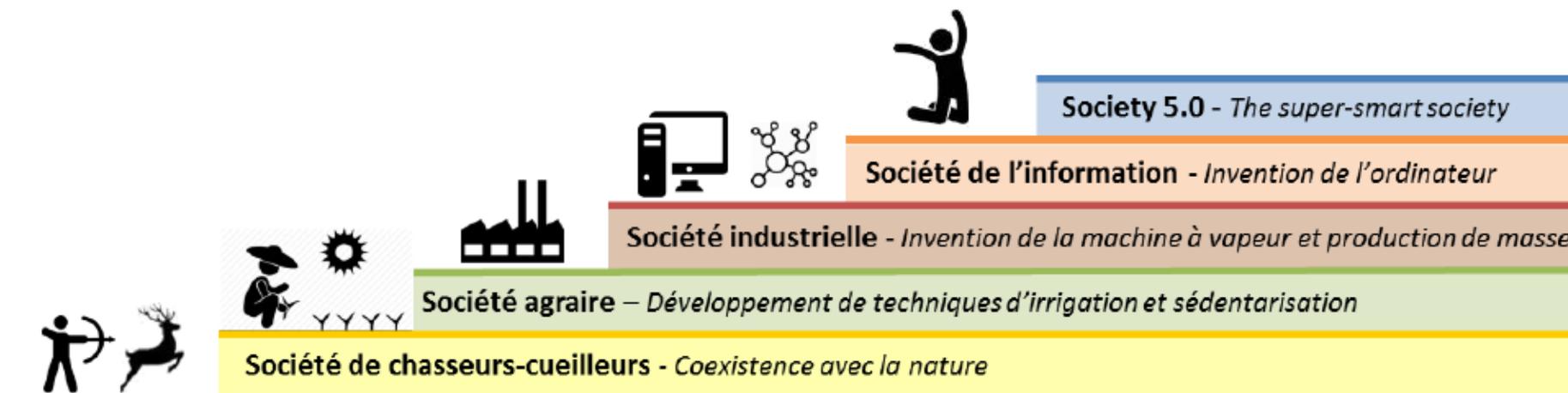
4. Politique scientifique de l'IA au Japon



Le paysage scientifique japonais, par ChatGPT

L'idéal japonais: la société 5.0

- Le Japon fait face à un grand nombre de défis: dette publique, environnement géopolitique, déclin en terme d'innovation malgré un grand investissement R&D (3,28% du PIB en 2020).
- En avril 2016, le gouvernement japonais a publié son cinquième plan quinquennal (2016-2020) pour la Science et la Technologie, qui inclut le concept « Society 5.0 »
- Le sixième plan cadre pour la *Science, Technology and Innovation* (STI) a par la suite été publié en mars 2021, et concerne la période 2021-2025. En prolongement du plan précédent, il met en avant les recherches et les technologies nécessaires pour réaliser la société 5.0.



Le pilotage National de l'IA au Japon

- Stratégie nationale en Intelligence Artificielle japonaise est définie par le *Council for Science, Technology and Innovation* (CSTI), au sein du Cabinet Office, qui est un des grands conseils gouvernementaux, présidé par le premier ministre japonais,
- Le CSTI a fixé une feuille de route intitulée « Japan's AI strategy towards Society 5.0 » en juin 2019.
- 5 volets : formation, structuration de l'écosystème R&, déploiement dans le monde réel de l'IA, développement d'infrastructures de données, et éthique de l'IA.
- Cette stratégie a ensuite été complétée et mise à jour en 2020, 2021, et 2022.



Le PM Kishida lors d'une séance du CSTI

Comités de pilotage de l'IA au gouvernement japonais

- Création en février 2023 du « Groupe de travail du Parti Libéral Démocrate sur l'évolution et l'implémentation de l'IA » (*Jimintô AI no shinka to jissô ni kansuru purojekuto chimu*), piloté par le député Masaaki Taira
- Cette équipe a rédigé un livre blanc publié le 14 avril 2023, intitulé dans sa version anglaise « The AI White Paper – Japan's National Strategy in the New Era of AI »
- Le même groupe de travail a rendu public, le 16 février dernier, un premier « projet de loi pour la promotion d'une IA responsable » (titre provisoire).
- Nouvelle stratégie nationale en IA en cours de préparation
- Création du *AI strategic Council* (8 membres) pour conseiller le premier ministre sur les grandes orientations stratégiques en IA



Le député Masaaki Taira



Une séance du AI strategic council

Engagement à l'international

- En sa qualité de président du G7, le Japon a joué un rôle actif dans les discussions internationales du groupe de travail numérique du G7 et dans le lancement du « processus d'Hiroshima » .
- Le processus d'Hiroshima a déjà permis d'aboutir à un code de conduite à destination des entreprises développant des algorithmes d'IA générative.
- Création le 14 février 2024 d'un *Japan AI Safety Institute* (AISI) dirigée par une ancienne employée d'IBM Japon
- Cet AI Safety Institute va travailler très étroitement avec ses homologues au Royaume Uni et aux Etats-Unis



Akiko Murakami, première présidente du Japan AI safety institute

Application dans les ministères et gouvernements

- Le ministre chargé de la Numérisation, Kono Taro, a ainsi déclaré en 2023 vouloir encourager l'utilisation de l'IA parmi les fonctionnaires pour améliorer la productivité et les conditions de travail.
- Des gouvernements locaux (préfectures, villes, etc.) ou des organisations diverses (universités, etc.) s'intéressent à l'IA générative
- Exemple: ville de Yokosuka, dans la préfecture de Kanagawa, qui a fait un test d'un mois de l'utilisation de ChatGPT, à compter du 20 avril 2023.
- Le *Tokyo Metropolitan Government* a déclaré officiellement, fin août 2023, que tous les bureaux pouvaient utiliser des outils d'intelligence artificielle générative.

☰ 🔍 **thejapanimes**

JAPAN

Tokyo Metropolitan Government to start using ChatGPT from August

KYODO SHARE Jun 14, 2023

The Tokyo Metropolitan Government will begin using the artificial intelligence chatbot ChatGPT for writing texts and carrying out other clerical work in all of its offices from August, Gov. Yuriko Koike said Tuesday.

ChatGPT "has the potential to greatly transform the way public administration is conducted," Koike said during a metropolitan assembly session. She added that "better city governance" can be achieved by assessing the positive and negative aspects of the AI service.

Koike also said the metropolitan government will use ChatGPT for tasks including preparing documents in question-and-answer format, and seek input from its employees about other practical uses for the generative AI tool.

Chatbots are software applications trained using massive amounts of data from the



Agences publiques

- Trois grands centres de recherche ont fondé leur centre en IA
 - Le MEXT (ministère de l'éducation, de la culture, des sports, des sciences et de la technologie) a créé le RIKEN AIP Center (RIKEN Center for Advanced Intelligence Project)
 - Le METI a créé l'AIRC (Artificial Intelligence Research Center) en mai 2015, au sein de l'AIST ;
 - Le MIC (ministère de l'intérieur et des communications) s'appuie sur un réseau de centres de recherches du NICT (National Institute of Information and Communications Technology).
- Ces trois centres forment le cœur du « AI Japan R&D Network », dont la création a été officialisée le 16 décembre 2019.
- Le NII (National Institute of Informatics) joue également un rôle actif en matière de recherche et développement, en servant notamment de Hub pour la création d'un LLM national, appelé Llm-jp.
- Les prestigieuses universités nationales (University of Tokyo, Université de Kyoto, etc.) et quelques universités privées (Waseda University, Keio University) sont également mobilisées et très en avance par rapport aux autres universités du Japon.



Acteurs privés: quelques exemples

- *Line* a développé un LLM et a annoncé avoir l'intention d'investir davantage dans cette technologie dans les années à venir ;
- La société NTT a commencé à commercialiser son LLM, *Tsuzumi* depuis 1^{er} mars 2024 ;
- Sony a monté en 2020 une branche consacrée à la R&D en IA, appelée Sony AI, dirigée par Hiroaki Kitano (cf. supra - fondateur de la *Robocup*, directeur du *Japan AI R&D Network*). Les projets de Sony AI sont essentiellement appliqués à l'industrie du divertissement.
- La société *CyberAgent* a publié en fin d'année 2023 une série de modèles en open source entraînés sur des données japonaises ;
- *Preferred Networks*: startup japonaise focalisée sur la R&D en deep learning. La société a publié *PLAMO-13B*, fonctionnant en anglais et japonais. Disponible en open source.
- *Sakana AI*: a levé récemment 30 millions de dollars auprès d'investisseurs internationaux, pour proposer des modèles petits et spécialisés
- *SB Intitutions*, startup possédée à 100% par Softbank, doit développer des LLM spécialisés en japonais.

Actions particulières pour l'IA générative

L'essentiel des efforts financiers exceptionnels consacrés à l'IA par le gouvernement japonais ces dernières années l'ont été dans le renforcement de la puissance de calcul du pays.

- 225kM\$ dépensés par le METI pour créer le *AI Bridging Cloud Infrastructure* (ABCI), pour une puissance de 55 Petaflops. D'autres actions sont en cours pour renforcer l'infrastructure de calcul du Japon.
- Projets de renforcer le supercalculateur Fugaku, mais rien d'annoncé à ce stade

OpenAI (25,000 GPU?)



Frontier (37,888 GPU)



LUMI (20,480 GPU)



Aurora (63,744 GPU)



Fugaku (158,976 CPU)



El Capitan (20,000+ GPU)



Les plus grands supercalculateurs au monde

Conclusion

- « Prompt engineering »: Tâche, Contexte, Format, Rôle, Ton, Exemple
- Vigilance sur deux points particuliers:
 - Attention aux fuites de données
 - Attention à la véracité des informations qui sont produites.
- Grosse montée en puissance de ce type de technologies
 - Des impacts possibles sur le fonctionnement des entreprises et administrations
- Le Japon a pris du retard (de son propre aveu) mais des mesures ont et vont être prises pour le combler.