



AGIR #4

ACTIONS DE LA GOUVERNANCE INITIÉES PAR LA RECHERCHE



**ChatGPT, Bard... Quels impacts
de l'IA générative sur l'enseignement
et la recherche ?**

Synthèse du séminaire du 30.03.2023

CHATGPT, BARD... QUELS IMPACTS DE L'IA GÉNÉRATIVE SUR L'ENSEIGNEMENT ET LA RECHERCHE ?

INTRODUCTION

El Mouhoub MOUHOUD

Président de l'Université Paris Dauphine - PSL

INTERVENANTS

Alexandre ALLAUZEN

Professeur en informatique à l'ESPCI-PSL et chercheur au LAMSADE de l'Université Paris Dauphine - PSL

Jamal ATIF

Professeur et chercheur au LAMSADE de l'Université Paris Dauphine - PSL

Pierre LANIRAY

Maître de conférences en sciences de gestion, délégué à l'innovation et à l'accompagnement pédagogiques de l'Université Paris Dauphine - PSL

Mattias MANO

Directeur du Centre d'Innovation Pédagogique de l'Université PSL

Sophie MERITET

Maître de conférences en économie, Vice-Présidente Affaires Internationales de l'Université Paris Dauphine - PSL

MODÉRATRICE

Florence GELIN

Directrice Générale des Services de l'Université Paris Dauphine - PSL

Les intelligences artificielles génératives sont sur le devant de la scène médiatique depuis le lancement de ChatGPT en novembre 2022. Comment fonctionnent-elles ? Qu'en attendre ?

C'est la problématique du quatrième séminaire AGIR « Actions de la Gouvernance Initiées par la Recherche » proposé par El Mouhoub MOUHOUD, président de l'Université Paris Dauphine - PSL, et présenté le 30 mars 2023.

Après un éclairage sur les technologies des IA génératives par Alexandre ALLAUZEN, le débat s'est ouvert sur les enjeux soulevés par ces nouvelles technologies au sein du monde académique.

« Notre rôle en tant que grand acteur académique n'est pas d'interdire l'utilisation de ces intelligences artificielles au sein de notre établissement, mais de réfléchir sur l'émergence de ces technologies pour mieux agir. »

E.M. MOUHOUD,
président de l'Université Paris Dauphine - PSL

Synthèse de la présentation

Alexandre ALLAUZEN

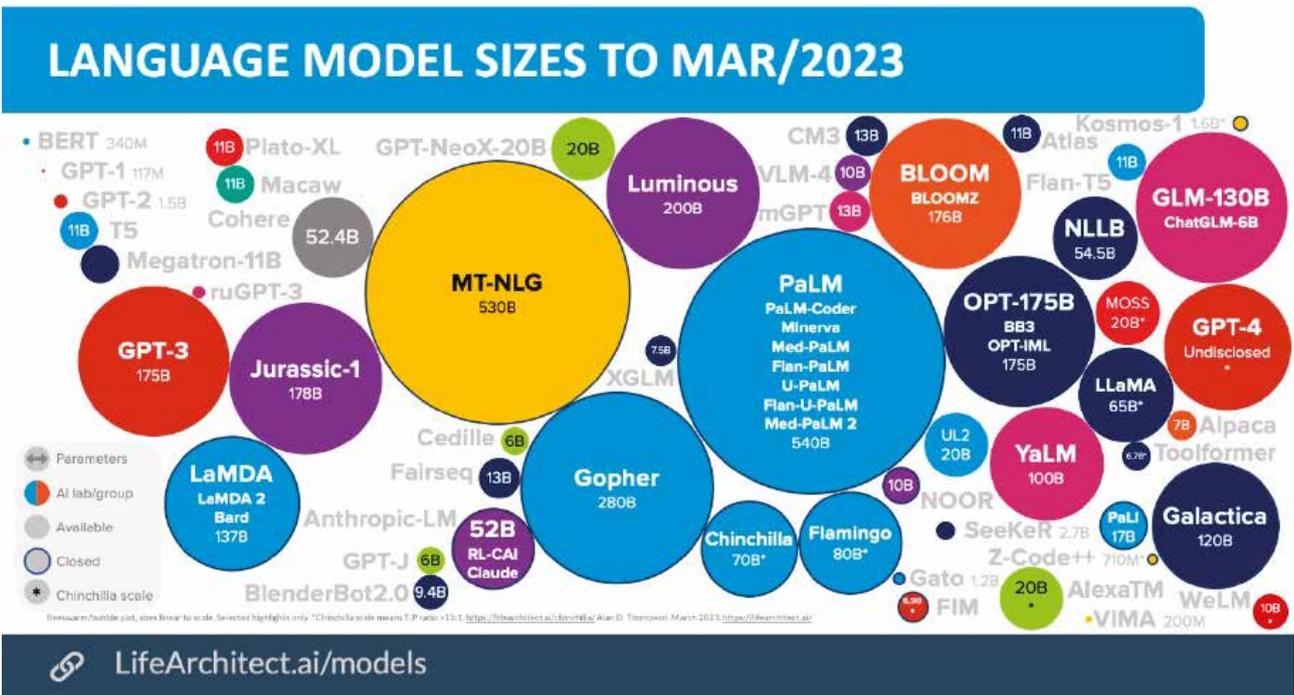
Derrière ChatGPT, la révolution des « large language models »

Le 30 novembre 2022, l'entreprise OpenAI lançait officiellement ChatGPT, un agent conversationnel — ou chatbot — nouvelle génération. En une semaine, la plateforme de mise à disposition de l'outil recensait plus d'un million d'utilisateurs. En deux mois, plus de cent millions. **Phénomène immédiat, et mondial.** Un succès dû en grande partie à la versatilité de l'outil et au réalisme de ses réponses. L'écriture de l'agent conversationnel bluffe aussi bien lorsqu'il s'agit de présenter des idées de start-up à des investisseurs, que de proposer un plan de dissertation.

L'outil base ses performances sur un « large language model », ou en français : très grand modèle de langue. Un modèle de langue est avant tout une représentation statistique d'un langage. Il permet de prédire, à partir d'une suite de mots, quel est le mot suivant le plus probable. La phrase « Le temps passe si... » sera plus probablement complétée par le mot « vite » que par le mot « tractopelle ». À chaque mot, le modèle de langue associe donc un poids statistique en fonction des mots précédents, c'est-à-dire en fonction du contexte. C'est ainsi que le modèle peut générer, mot après mot, des phrases ayant du sens pour un être humain.

« À chaque mot, le modèle de langue associe donc un poids statistique en fonction des mots précédents, c'est-à-dire en fonction du contexte. C'est ainsi que le modèle peut générer, mot après mot, des phrases ayant du sens pour un être humain. »

Il existe aujourd'hui plus d'une centaine de modèles de langue. GPT-3, celui utilisé par ChatGPT, n'en est qu'un. Des géants de l'industrie numérique comme Google, Amazon ou Microsoft travaillent sur leurs propres modèles de langue, tout comme des acteurs académiques tels que les universités de Berkeley, Stanford ou Fudan. Si les modèles de langue sont si nombreux, c'est notamment parce qu'ils ne sont pas nouveaux et que les acteurs académiques et industriels de l'intelligence artificielle se positionnent dessus depuis plus de 20 ans.



Représentation des modèles de langue existant en mars 2023.

« Grâce à l'essor des réseaux de neurones dans les années 2000, il devient possible à des modèles d'apprendre automatiquement ces fonctions de mots, avec une approche statistique et non pas humaine. »

Les « large language models » en 3 étapes

Au début des années 2000, le domaine du traitement automatique des langues connaît un tournant significatif avec les « *word embeddings* ». Cette technique consiste à représenter un mot en fonction de son comportement. « Je », « tu » et « nous » ont une fonction similaire. « Montélimar », « Brest » et « Londres » en ont une autre. Grâce à l'essor des réseaux de neurones dans les années 2000, il devient possible à des modèles d'apprendre automatiquement ces fonctions de mots, avec une approche statistique et non pas humaine. Et pour entraîner ces réseaux de neurones, le domaine peut compter sur une autre technologie en plein boom à cette période : **internet, et son nombre colossal de textes désormais accessibles en ligne qui vont nourrir l'apprentissage.** C'est la première étape vers les outils tels que ChatGPT.

Deuxième étape : le milieu des années 2010, et le **concept-clé de l'attention**, issu des outils de traduction automatique. Dans la phrase « je livre une pizza en boîte », comment traduire le mot « livre » en anglais ? *Book ? Pound ? Deliver ?* Pour cela, il faut canaliser l'attention portée sur la phrase à traduire. « Livre » suit le pronom « je ». Le pronom devient important pour le modèle de traduction car il conditionne la présence d'un verbe derrière, et permet de sélectionner « *deliver* » dans le champ des possibles.

Et enfin troisième étape : les *transformers*, qui capitalisent sur les deux étapes précédentes. Non seulement les modèles peuvent identifier correctement la fonction d'un mot, mais les *transformers* permettent à présent d'en déterminer le sens en analysant les mots précédents et suivants. Autrement dit : **les modèles de langues tiennent compte à présent du contexte des mots et des phrases.**

De la théorie à ChatGPT

GPT-3 se base sur ces *transformers*. L'acronyme signifie d'ailleurs *generative pretrained transformers*, ou *transformers* génératifs pré-entraînés. **L'essence de GPT est d'empiler en cascade ces briques technologiques** que sont les *transformers* pour donner à chaque mot, à chaque phrase, un sens compréhensible par l'humain. Ce modèle de langue est qualifié de « très grand » (*large*) car **il fait appel à 175 milliards de paramètres pour représenter un contexte.** Une taille de paramètres qui a subi une croissance exponentielle récemment — les plus grands modèles de langue datant d'il y a 2 ans ne comptaient que quelques centaines de millions de paramètres. Cette croissance rapide est ce qui explique en grande partie les performances subites des outils récents tels que ChatGPT. Et la croissance continue à un rythme très soutenu.

Une fois le modèle de langue GPT suffisamment performant, reste à lui **adosser la fonction de chat pour en faire l'outil conversationnel qu'il est aujourd'hui.** Pour cela, il faut trois étapes. D'abord, montrer manuellement au modèle la réponse humaine attendue à une question pour l'entraîner de manière supervisée. Ensuite, lui demander de formuler des réponses et

les classer de la meilleure à la plus mauvaise. Enfin, apprendre au modèle à évaluer lui-même ses réponses pour qu'il puisse interagir de manière autonome.

Applications et limites des grands modèles de langue

Cette capacité d'interaction implique une nouvelle relation à l'information, non plus basée sur la recherche, mais sur la génération. ChatGPT peut traiter de l'information et la trier pour la raffiner. Il peut **générer du code, des synthèses de sujets complexes, transcrire une image...** Autant d'applications qui posent des questions quant à **l'impact de cette technologie sur nos quotidiens**. Pour l'expérience, des chercheurs ont fait passer des examens à ChatGPT. Droit, art, biologie : le programme réussit ces tests basés sur l'acquisition de connaissances.

De manière plus pragmatique, l'application d'apprentissage des langues **Duolingo** a prévu d'utiliser ChatGPT dans son système pour enrichir le retour de l'apprenant. La startup **Be My Eyes** l'utilise pour décrire l'environnement des personnes malvoyantes. Et des applications de travail collaboratif l'intègrent également pour faciliter la prise de note ou les interactions entre participants.

Pour autant quelques limitations subsistent. Les modèles et les données ne sont pas ouverts à la communauté, le coût de l'outil pose des questions d'accessibilité, et **ChatGPT n'est pas encore adapté à des usages de niche** tels que les applications juridiques. Comme toutes les technologies numériques basées sur des jeux de données, **la technologie est aussi tributaire de la qualité des données utilisées pour l'entraînement** : biais sociaux, ethniques, ou de genre. Toutefois, ces limites sont périssables. Celles mentionnées hier sur le niveau de performance n'ont déjà plus lieu d'être, et celles citées précédemment pourraient bien ne plus être pertinentes d'ici quelques mois.

« Les modèles et les données ne sont pas ouverts à la communauté, le coût de l'outil pose des questions d'accessibilité, et ChatGPT n'est pas encore adapté à des usages de niche tels que les applications juridiques. »



Discussion autour de l'impact sur la société et l'enseignement

Synthèse des interventions

Jamal ATIF, Pierre LANIRAY, Mattias MANO, Sophie MERITET

Avec l'IA générative, le paradigme change. À présent nous parlons d'outils performatifs, utilisés pour générer des contenus et plus seulement les traiter. Il faut s'attendre à des contenus créés par des IA, diffusés largement, qui serviront eux-mêmes à entraîner de nouveaux modèles d'IA. En étant confrontés plus fréquemment à ces contenus, **notre perception du monde peut s'en trouver modifiée**. Ce qui s'opère est donc fondamentalement nouveau.

ChatGPT arrive à une période où nous mettons en données le monde, et **il se positionne sur la notion d'interaction entre l'humain et le web**. Cela pose la question des futurs contenus, produits de manière massive et automatique, sur la promotion de produits, les opinions politiques, tous similaires et **difficiles à différencier de contenus rédigés par des humains**. ChatGPT présente ce problème dual, que l'on retrouve dans les sujets sociotechniques controversés, sur **l'utilisation positive ou négative de la technologie**. En management des systèmes d'information, le **concept d'affordance** résume cette dualité : les technologies sont une suite de potentialité d'actions et chaque utilisateur a son usage propre. Intrinsèquement, il est donc imprévisible de connaître à l'avance les effets d'un outil comme ChatGPT, d'autant plus à la vitesse à laquelle il évolue.

Comment alors former les étudiants ? Quel sera leur cœur des compétences ? S'il s'agit de leur permettre de développer une pensée libre, alors effectivement l'utilisation de technologies d'IA génératives interroge sur la qualité d'une dissertation rédigée sans supervision enseignante. Cependant, le problème n'est pas nouveau. Depuis la fin des années 2010, les attentes des étudiants ne sont plus au cours descendant, et les enseignants perçoivent bien qu'il **faut renouveler l'interaction entre enseignants et apprenants**.

C'est la démarche qui a poussé à la **création de l'outil Tuteur virtuel** par l'Université Paris Dauphine-PSL, testé actuellement auprès de 600 étudiants. Il s'agit d'un assistant en ligne proposant aux étudiants des activités personnalisées selon leur progression dans l'apprentissage. L'accent est mis sur la pratique et l'accompagnement de l'étudiant dans son acquisition de compétences. L'outil, basé sur un chatbot, **augure des nouvelles formes d'interaction pédagogique rendues possibles par la technologie**.

Une chose est sûre : il faut accompagner les acteurs de l'enseignement dans cette transition numérique, dont ChatGPT est une composante. **Les compétences numériques des enseignants doivent être renouvelées** pour permettre de se saisir des nouveaux outils, et pour **sensibiliser les étudiants au fonctionnement des IA génératives** afin d'éviter les écueils et faire preuve d'esprit critique.

C'est notamment la mission du **centre d'innovation pédagogique (CIP)** de l'Université PSL, qui effectue un travail de veille poussé sur les innovations qui touchent à la pratique pédagogique. **ChatGPT permet de retravailler des activités d'enseignement**, en aidant par exemple à revisiter des plans de cours. Mais pour utiliser l'outil de cette manière, **il faut d'abord apprendre aux enseignants à interagir avec lui**. Quelles sont les capacités de ChatGPT ? Comment s'adresser à lui pour obtenir une réponse de qualité ? Quelles sont ses limites ? Tout cela doit faire l'objet de formations en amont.

Maîtriser et former, voilà le nerf de la guerre pour une utilisation raisonnée des nouvelles technologies d'IA. Il faut s'y atteler à la fois sur le volet informatique et mathématique, mais aussi sur les impacts en les intégrant dès les étapes de développement et d'appropriation de l'outil. Sans maîtrise de la technologie, et sans formation attenante, les utilisateurs ne pourront pas s'en faire une représentation correcte.



« En s'appuyant sur une approche pluridisciplinaire, le séminaire AGIR a pour objectif d'éclairer par la recherche les actions de la gouvernance de notre université. Grâce à la contribution d'enseignantes-chercheuses et d'enseignants-chercheurs dauphinois, ainsi que la participation d'invités externes, ces séminaires ont pour ambition de nourrir nos stratégies de pilotage et visent à impulser de nouvelles démarches innovantes. »

El Mouhoub MOUHOUD

Président de l'Université Paris Dauphine - PSL

À propos d'AGIR

Organisés autour de différentes thématiques clés liées au programme de la présidence de l'université pour la période 2021-24, les Séminaires AGIR (Actions de la Gouvernance Initiées par la Recherche) invitent les acteurs de la communauté dauphinoise à partager les analyses et enjeux permettant de mieux fonder les actions de la gouvernance de l'université. Les conférences et débats sont ouverts au grand public.

Séminaire AGIR #1 / 21.05.2021

Quels enjeux pour les programmes d'égalité des chances en France ?

Séminaire AGIR #2 / 01.10.2021

Quels enjeux en matière de responsabilité environnementale dans l'ESR ?

Séminaire AGIR #3 / 21.01.2022

Enseignement supérieur, pour un investissement plus juste et plus efficace

Séminaire AGIR #4 / 30.03.2023

ChatGPT, Bard,... Quels impacts de l'IA générative sur l'enseignement et la recherche ?



Retrouvez en ligne, l'intégralité du replay du séminaire AGIR du 30.03.2023 ainsi que tous les replays des précédents séminaires sur le site dauphine.psl.eu



UNIVERSITÉ PARIS DAUPHINE-PSL
Place du Maréchal de Lattre de Tassigny
75775 Paris cedex 16
dauphine.psl.eu