

Les grands modèles de langage/: de l'abstraction à la réalité

Publié le 8 octobre 2024

Christopher Gannatti, CFA

Global Head of Research

- Les grands modèles de langage (LLM) ont un impact dans le monde réel, ce qui permet à des entreprises telles qu'Amazon d'économiser de nombreuses années de développement.
- Les modèles fondamentaux constituent une couche d'infrastructure essentielle pour les futures innovations en matière d'IA.
- Les LLM améliorent la productivité et l'efficacité en matière de développement logiciel et d'opérations commerciales.
- Le potentiel économique des LLM s'étendra à mesure que les grandes entreprises continueront d'investir massivement dans les infrastructures et les talents.
- L'évolution des modèles d'IA pourrait conduire à des interfaces plus spécialisées, adaptées à des tâches spécifiques.

Le problème avec les grands modèles de langage (LLM), c'est qu'il est difficile pour un ingénieur non spécialiste des logiciels de visualiser de quoi il s'agit. C'est encore plus difficile lorsque l'on découvre que le développement et la formation des plus grands LLM coûtent plusieurs centaines de millions de dollars.

Nous constatons pourtant que les PDG de certaines des plus grandes entreprises au monde indiquent que leur groupe dépensera plus de 1 000 milliards de dollars au cours des prochaines années dans la construction de nouvelles infrastructures informatiques, afin de pouvoir exécuter ces modèles plus facilement^{1&2}.

Nous sommes toujours à la recherche de cas d'utilisation ou de témoignages nous permettant de transcrire cette abstraction en quelque chose de concret, par exemple en illustrant le véritable impact d'un LLM pour une entreprise. Si nous découvrions suffisamment de cas d'utilisation, nous pourrions commencer à voir ces impacts s'intégrer dans les statistiques plus larges de la productivité économique.

C'est pourquoi nous avons été intéressés par la déclaration d'Andy Jassy, PDG d'Amazon, concernant le système Q de l'entreprise, qui est pour l'essentiel un LLM capable de générer du code logiciel³:

La durée moyenne de mise à niveau d'une application vers Java 17 est passée de 50 jours de développement à seulement quelques heures. Nous estimons que cela nous a permis d'économiser l'équivalent de 4 500 années de travail de développement (oui, ce chiffre est incroyable mais vrai).

En moins de six mois, nous avons été en mesure de mettre à niveau plus de 50 % de nos systèmes Java de production vers des versions Java modernisées, en une fraction du temps et des efforts habituellement nécessaires. Par ailleurs, nos développeurs ont expédié 79 % des révisions de code générées automatiquement sans modifications supplémentaires.

Les avantages s'étendent au-delà des efforts économisés pour les développeurs. Les mises à niveau ont en effet amélioré la sécurité et réduit les coûts d'infrastructure, ce qui a permis d'estimer les gains d'efficience annualisés à 260 millions de dollars.

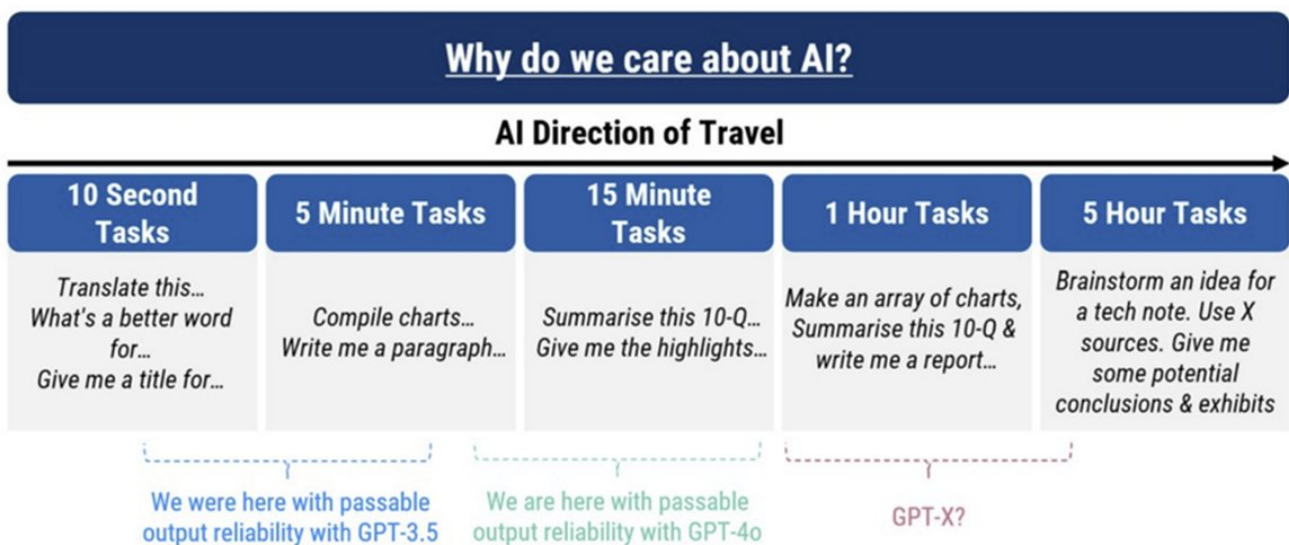
Une économie de 4 500 années de développement ??? Des gains d'efficience annualisés de 260 millions de dollars ??? Ces chiffres sont absolument vertigineux. Nous avons conscience que la révolution de l'IA ne fait que commencer, mais la conceptualisation formulée par Jassy concernant l'impact de l'IA sur les développeurs de son équipe conduit d'autres personnes à publier en détail des témoignages similaires.

Une feuille de route conceptuelle pour l'évolution de l'IA4

L'illustration 1 est selon nous essentielle pour permettre à chacun de rapidement visualiser d'une autre manière l'intérêt de tous ces LLM. Tous les professionnels de la connaissance peuvent comprendre que leur travail se compose de différents ensembles de tâches, et que chacun de ces ensembles peut impliquer des quantités de temps très différentes.

Les versions actuelles des LLM sont capables de répondre à des questions ou à des e-mails simples, mais il est beaucoup moins certain que ces systèmes puissent créer à partir de zéro des rapports entièrement nouveaux ou des idées originales. Cela ne signifie pas qu'ils ne sont pas en capacité de le faire, mais simplement qu'à un certain stade le niveau d'analyse nécessaire pour des réponses ou des questions simples est très différent du niveau d'analyse nécessaire pour un tout nouveau diaporama susceptible de contenir 60 diapositives originales, toutes créées par l'IA.

Illustration 1 : Une feuille de route conceptuelle pour l'évolution de l'IA



Source : Stanley, Edward et al. « Mapping AI's Rate of Change. » Morgan Stanley Research. 4 juin 2024.

Les LLM sont-ils des logiciels, des infrastructures, ou les deux ?

Certains qualifient les LLM de « modèles fondamentaux ». Le terme « fondamental » suscite la réflexion dans la mesure où, dans de nombreux contextes, un fondement est une base sur laquelle vous pouvez bâtir quelque chose. Si nous prenons l'exemple de la création de valeur dans plusieurs écosystèmes :

- Qu'est-ce qui confère de la valeur aux ordinateurs ? On pourrait considérer tous les composants matériels et additionner la valeur de chacun d'entre eux pour déterminer la valeur d'un ordinateur. Or, si l'on songe à la manière dont les ordinateurs ont été adoptés au fil de leur histoire, il fut un temps où l'on pensait que les individus pourraient construire leurs propres ordinateurs. La facilité d'utilisation que Microsoft a apportée à l'expérience d'exécution de tâches précises grâce aux ordinateurs a été extrêmement précieuse, tout comme le fait que de nombreux utilisateurs aient adopté ces progiciels et appris à les utiliser tous simultanément.
- Qu'est-ce qui confère de la valeur à Internet ? Beaucoup répondront qu'Internet permet aux êtres humains d'échanger des informations à l'échelle mondiale beaucoup plus facilement. Or, dans ce cas, que penser de Google, un exemple d'entreprise qui cherchait à organiser les informations disponibles sur Internet afin que vous puissiez obtenir ce dont vous aviez besoin beaucoup plus rapidement ? Comment pensez-vous que nous nous en sortirions si nous devions utiliser Internet sans moteurs de recherche ?
- Si l'on prend l'exemple de l'iPhone, la valeur de l'appareil réside-t-elle dans les composants du téléphone et dans ce que la technologie peut spécifiquement accomplir, ou dans l'ensemble des applications intégrées dans le logiciel iOS ?

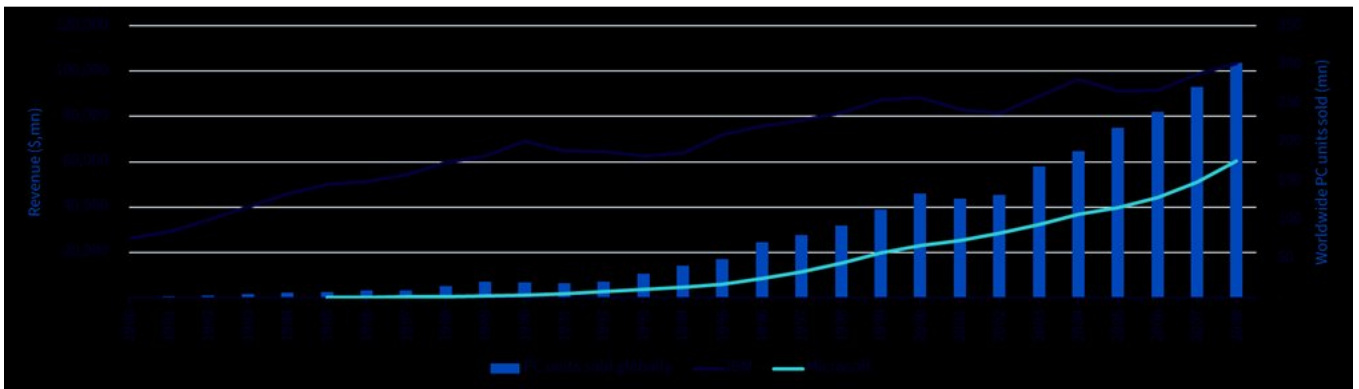
Dans chacun de ces cas, la réponse consiste probablement à considérer que tout revêt une certaine valeur, mais la raison pour laquelle nous faisons souvent référence aux produits et services proposés par des entreprises dont la capitalisation boursière dépasse les 1 000 milliards de dollars réside dans le fait que différents effets se répercutent et se multiplient de manière exponentielle. Des effets de réseau considérables entrent également en jeu : rien ne génère plus d'utilisateurs et de croissance qu'une immense base initiale d'utilisateurs.

Les illustrations 2a et 2b présentent des tendances que nous avons déjà observées :

- L'illustration 2a indique la relation entre les revenus d'IBM, de Microsoft, et le nombre d'unités de PC vendus à l'échelle mondiale. Dans un premier temps, IBM vend logiquement des PC, ou permet à d'autres entreprises de concevoir et commercialiser ce que l'on appelait les clones IBM. Le nombre de PC augmente et, après un certain temps, une fois que la base d'utilisateurs de PC atteint un certain niveau, les revenus de Microsoft décollent dans la mesure où tous ces utilisateurs de PC en entendent parler, et recherchent les avantages de l'utilisation des systèmes d'exploitation de Microsoft. Il aurait été absurde que le système de Microsoft prenne de la valeur avant que les PC ne se retrouvent sur la quasi-totalité des bureaux.

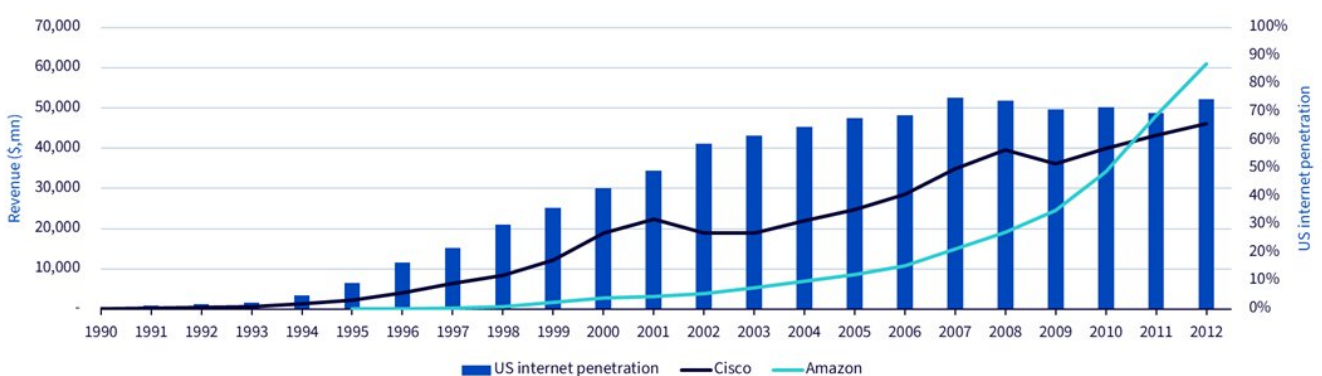
- L'illustration 2b établit la relation entre l'adoption d'Internet et les revenus de Cisco et d'Amazon.com. Les appareils Cisco permettent à un plus grand nombre d'entreprises et de personnes d'utiliser Internet. À mesure que de plus en plus de personnes profitent de l'Internet, la valeur de l'offre d'Amazon.com - le commerce électronique - devient de plus en plus évidente. Il n'aurait pas été logique que les efforts de commerce électronique d'Amazon.com revêtent de la valeur avant l'adoption d'Internet par une masse critique d'utilisateurs.

Illustration 2a : Relation entre le nombre d'unités de PC vendus à l'échelle mondiale et les revenus d'IBM et de Microsoft au fil du temps



Source : WisdomTree, Bloomberg. Les données relatives au nombre d'unités de PC vendus proviennent de « Total Share: Personal Computer Market Share 1975-2010, Jeremy Reimer » et de Gartner. **Les performances historiques ne garantissent pas les résultats futurs, et tout investissement est susceptible de perdre de la valeur.**

Illustration 2b : Relation entre l'adoption d'Internet aux États-Unis et les revenus de Cisco et d'Amazon.com au fil du temps



Source : WisdomTree, Bloomberg, Banque mondiale. **Les performances historiques ne garantissent pas les résultats futurs, et tout investissement est susceptible de perdre de la valeur.**

Conclusion : Que développerons-nous sur la base des modèles fondamentaux ?

Ce qui est intéressant avec les modèles fondamentaux, du moins au deuxième semestre 2024, c'est que seules les entreprises mondiales les plus grandes et les plus rentables disposent des ressources pour continuer de les développer et de les faire progresser. Même s'il semble que certains de ces modèles fassent partie d'entreprises indépendantes, les plus grandes sociétés au monde ont tendance à acquérir d'importantes participations financières permettant les investissements croissants nécessaires dans les talents et l'infrastructure informatique.

Nous ignorons de quoi l'avenir sera fait, et il s'agit sans doute d'une question à mille milliards de dollars. Ce que nous savons, en revanche, c'est que les individus ont tendance à privilégier des tâches personnelles discrètes, et qu'ils n'auront pas nécessairement besoin d'accéder à un modèle capable de passer n'importe quel examen de haut niveau. Nous aimons cependant l'idée d'interfaces plus spécialisées, axées sur des tâches plus spécifiques, et susceptibles ensuite d'utiliser certaines parties des modèles plus larges pour accomplir le travail.

Quelle que soit l'évolution de la situation, nous pensons que les plus grandes entreprises mondiales connectées à ces modèles fondamentaux auront un rôle important à jouer dans les années à venir.

1 Goldman Sachs, <https://www.datacenterdynamics.com/en/news/goldman-sachs-1tn-to-be-spent-on-ai-data-centers-chips-and-utility-upgrades-with-little-to-show-for-it-so-far/>

2 Nvidia, https://www.linkedin.com/posts/leeps_nvvidia-ceo-predicts-1-trillion-will-be-spent-activity-7101349410281836544-hB72/

3 Source : Extrait d'une publication LinkedIn du PDG d'Amazon, Andy Jassy, à laquelle il est fait référence sur le lien <https://nextbigteng.substack.com/p/hello-ai-world-evolution-of-developer-economy-in-the-age-of-ai>.

4 Source : Stanley, Edward et al. « Mapping AI's Rate of Change. » Morgan Stanley Research. 4 juin 2024.

Important Risks Related to this Article

Informations importantes

Communications commerciales publiées dans l'EEE Ce document est publié et approuvé par WisdomTree Ireland Limited, une société autorisée et réglementée par la Central Bank of Ireland.

Communications commerciales émises dans des juridictions en dehors de l'EEE Ce document est publié et approuvé par WisdomTree UK Limited, une société autorisée et réglementée par la Financial Conduct Authority du Royaume-Uni.

WisdomTree Ireland Limited et WisdomTree UK Limited sont toutes les deux désignées comme « WisdomTree » (le cas échéant). Notre Politique sur les conflits d'intérêts et notre Inventaire sont disponibles sur demande.

Réservé aux clients professionnels uniquement. Les informations figurant dans ce document sont fournies à titre informatif et ne constituent pas une ore de vente, ou une sollicitation d'ore d'achat de titres ou d'actions. Ce document ne doit pas être utilisé comme fondement d'une décision d'investissement. La valeur des investissements peut fluctuer et vous êtes susceptible de perte tout ou partie du montant investi. La performance passée ne constitue pas nécessairement une indication des performances futures. Toute décision d'investissement doit être fondée sur les informations figurant dans le prospectus approprié et sur des conseils indépendants en matière d'investissement, fiscaux et juridiques.

L'application des réglementations et lois fiscales peut souvent conduire à des interprétations diérentes. Tous les points de vue ou opinions exprimés dans cette communication représentent les points de vue de WisdomTree et ne doivent pas être interprétés comme des conseils réglementaires, fiscaux ou juridiques. WisdomTree ne donne aucune garantie ou représentation quant à l'exactitude des vues ou opinions exprimées dans cette communication. Toute décision d'investissement doit être fondée sur les informations contenues dans le prospectus approprié et après avoir sollicité des conseils indépendants en matière d'investissement, fiscaux et juridiques. Ce document n'est pas et ne doit en aucun cas être interprété comme une publicité ou une ore publique d'actions ou de titres aux États-Unis ou dans toute province ou tout territoire des États-Unis. L'introduction, la transmission et la distribution (directes ou indirectes) de l'original ou d'une copie de ce document sont interdites aux États-Unis.

Bien que WisdomTree s'efforce d'assurer l'exactitude du contenu de ce document, WisdomTree ne peut en garantir l'exactitude. Les fournisseurs de données tiers sollicités pour obtenir les informations contenues dans le présent document ne donnent aucune garantie ou représentation de quelque sorte en rapport avec ces données. Lorsque WisdomTree exprime ses propres opinions concernant le produit ou l'activité du marché, ces opinions sont susceptibles de changer. WisdomTree, ses alliés et leurs dirigeants, directeurs, partenaires ou employés respectifs déclinent toute responsabilité pour toute perte directe ou indirecte découlant de l'utilisation de ce document ou de son contenu.