

L'IA Générative : Révolution Économique ou Simple Effet de Mode ?



Les plus anciens d'entre nous se souviennent que la puce 8086 d'Intel est à l'origine de la révolution micro-informatique de la fin des années 1980. Sans elle, Microsoft n'aurait pas pu faire fonctionner son système d'exploitation DOS. Sans elle, les ordinateurs de bureau d'IBM n'auraient pas été vendus à une société chinoise.

Une dizaine d'années plus tard, la révolution internet s'est imposée, symbolisée par l'irruption de Cisco, société spécialisée dans les routeurs télécom, et de sociétés digitales innovantes telles que Google, Amazon ou Facebook. Les opérateurs télécom, qui semblaient pouvoir bénéficier de cette révolution, ont été remisés au rôle de « commodités ».

L'irruption des smartphones dans notre quotidien a été facilitée par ARM, une société britannique qui concevait les architectures des puces ensuite développées par des spécialistes tels que Qualcomm. Samsung est devenu le premier fabricant de smartphones basés sur cette technologie. Apple a créé son propre écosystème grâce à son génie marketing. Victimes collatérales, Nokia et Blackberry ont rapidement succombé.

L'irruption de l'intelligence artificielle générative

Cela fera bientôt deux ans que OpenAI a présenté ChatGPT, son premier « agent conversationnel » organisé autour d'un Grand Modèle de Langage (LLM en anglais). Il s'agit d'un système d'apprentissage automatique des données capable de comprendre et de générer des textes en langage humain.

L'histoire s'est ensuite accélérée. Des sociétés concurrentes de OpenAI ont émergé, les performances se sont améliorées.

L'intelligence artificielle n'est pourtant pas née avec ChatGPT. Alan Turing avait construit en 1936 un modèle abstrait de fonctionnement des appareils mécaniques de calcul (la « machine de Turing »). Après la seconde guerre mondiale, les travaux de John McCarthy ont permis d'établir l'intelligence artificielle en tant que discipline à part entière. L'objectif était de combiner l'aspect « artificiel » des ordinateurs et des processus informatiques aux aspects « intelligents » d'imitation de comportements humains, notamment de raisonnement et d'apprentissage. Pour parvenir à cette intelligence « supra-humaine », les scientifiques ont essayé de créer des réseaux de neurones artificiels inspirés du fonctionnement du cerveau humain.

Cette première phase a suscité beaucoup d'espoirs ... mais également de déceptions. Les périodes euphoriques, lorsque par exemple, des « systèmes experts » permettaient d'automatiser les analyses de bilans des entreprises, ont été suivies d'hivers longs et rigoureux après avoir constaté que ces systèmes « experts » ... ne l'étaient pas vraiment compte tenu de la complexité des tâches qui leur étaient demandées. Modéliser le cerveau humain s'est révélé être beaucoup plus compliqué qu'anticipé.

La technologie de ChatGPT marque une rupture. Le cerveau humain n'est pas modélisé mais les résultats obtenus par apprentissage massif de données sont bluffants.

Les perspectives ouvertes par cette rupture technologique ne semblent pas connaître de limites, notamment dans le quotidien des entreprises : gestion des process, logistique, traduction, marketing, relations commerciales, ...

Les investisseurs en ont conclu que nous étions à l'aube d'une nouvelle révolution technologique, peut-être même plus puissante que les trois précédentes.

Les valorisations des sociétés susceptibles de bénéficier de cette révolution ont littéralement explosé.

Les concepteurs et fabricants des outils permettant de faire fonctionner les ordinateurs utilisant l'intelligence artificielle générative ont été les premiers bénéficiaires : Nvidia bien sûr (3.000 milliards de dollars) mais également TSMC, ASML, Super Micro Computer, Hynix et bien d'autres.

Les concepteurs des agents conversationnels, qui n'existaient pas il y a cinq ans, font des levées de fonds sur des bases records : 150 milliards de dollars visés par OpenAI (pas très loin de TotalEnergies), 6 milliards d'euros pour le pionnier français Mistral (22 salariés en Décembre 2023).

Les grandes sociétés technologiques US qui régnaient sur le monde avant ChatGPT sont globalement gagnantes même si, compte tenu de leur taille, la progression des valorisations est moins sensible. Elles semblent aux yeux des investisseurs les plus capables d'intégrer l'intelligence artificielle pour améliorer et rentabiliser leur propre écosystème. Elles sont également en mesure de concurrencer les concepteurs et fabricants d'outils (Amazon) ainsi que les concepteurs d'agents conversationnels (Meta).

Il est en revanche difficile d'estimer l'impact de cette « révolution » sur l'économie traditionnelle. Les gagnants n'ont pas été identifiés. Les investisseurs ont en revanche désigné par anticipation quelques perdants (Téléperformance).

La hausse des valeurs liées à l'intelligence artificielle a été tellement importante qu'elle a entraîné avec elle les indices boursiers mondiaux.

L'heure des premiers doutes sur la nouvelle révolution technologique

Depuis quelques semaines pourtant, les investisseurs commencent à avoir quelques doutes. Ils peuvent être résumés en trois questions : l'intelligence artificielle générative peut-elle réellement révolutionner le quotidien des investisseurs et des consommateurs ? Cette technologie permettra-t-elle de créer de la richesse dans l'économie réelle ? Les gagnants d'hier et d'aujourd'hui seront-ils encore leaders demain ?

Les progrès réalisés par les LLM depuis deux ans sont étonnants. Parallèlement, le prix à payer croît également exponentiellement.

Lors de la phase d'apprentissage tout d'abord, les modèles doivent intégrer un nombre de données toujours plus importants et les traiter toujours plus rapidement. Les coûts en consommation électrique explosent. Selon certaines estimations, le coût d'apprentissage des plus grands modèles s'élève aujourd'hui

à 100 millions de dollars, celui de la prochaine génération sera de 1 milliard de dollars, celui de la génération d'après de 10 milliards de dollars.

Par ailleurs, le coût informatique généré par les requêtes est lui aussi élevé. Il est généralement estimé à quelques centimes de dollar par requête. Ce coût inclut l'utilisation de la puissance de calcul nécessaire pour traiter et répondre aux demandes des utilisateurs. Cela peut paraître faible mais c'est dix fois plus élevé qu'une requête sur Google Search. Et plus la question est complexe, plus le coût informatique grimpe.

Enfin, le coût des puces explose sans que l'investisseur puisse discerner si cette augmentation provient de la complexité du process de fabrication ou du pouvoir quasi monopolistique de Nvidia. En tout état de cause, les prévisions d'investissements dans l'IA annoncées lors des dernières publications de résultats par les grandes sociétés technologiques américaines n'ont pas rassuré les marchés, surtout après le fiasco du métavers.

L'intelligence artificielle générative pourrait-elle buter sur le prix de l'investissement nécessaire à son développement ? D'autres technologies prometteuses se sont heurtées à ce type de problèmes et n'ont pu les résoudre. L'Histoire nous enseigne toutefois qu'il ne faut pas sous-estimer l'intelligence humaine.

Les obstacles rencontrés par l'intelligence artificielle générative stimulent déjà les chercheurs.

Les fabricants d'ordinateurs dédiés aux LLM (hardware) découpent les tâches induites par l'apprentissage des données et développent des puces adaptées à chacune de ces tâches. Ces dernières sont plus efficaces et moins coûteuses en énergie. Les investissements affluent depuis le début de l'année sur les start-up spécialisées sur ce type de puces.

Les concepteurs de LLM spécialisent leurs modèles pour à la fois améliorer la qualité des réponses et diminuer le coût informatique. La dernière version de OpenAI a été conçue pour être performante sur les requêtes nécessitant un raisonnement sans avoir à générer de texte. D'autres modulent la précision de la recherche en fonction de la question posée par l'utilisateur. D'autres enfin se spécialisent dans un domaine particulier (droit, sciences, ...).

Les investisseurs ont également douté ces dernières semaines de la capacité de l'intelligence artificielle générative à créer de nouveaux marchés.

Les applications pratiques de la nouvelle technologie ont bien été identifiées et recouvrent tous les domaines de l'économie. Mais leur adoption par les entreprises et les consommateurs est lente.

Les entreprises sont coupables de « IA washing ». Elles parlent toutes d'IA générative mais très peu l'ont réellement mise en place, soit parce que leur IA

n'est pas générative, soit parce que ce n'est pas de l'IA mais de l'informatique traditionnelle.

Pour innover l'ensemble de l'économie, la technologie nécessite des développements multiples et importants. Elle doit également créer ses propres besoins.

Ces débuts plutôt hésitants ont stimulé la mémoire des investisseurs. Ils se sont souvenus que la machine à vapeur, le train ou l'électricité avaient mis plusieurs dizaines d'années avant d'avoir un impact significatif sur la croissance économique.

Mais, l'IA générative n'est pas une révolution comparable aux précédentes. Elle est de nature accumulative, complétant les trois précédentes révolutions technologiques citées précédemment. Et elle intervient dans une économie mondialisée où les idées et les talents circulent librement et où l'argent finance facilement les innovations.

L'impact de l'intelligence artificielle générative ne devrait pas être immédiat et massif comme l'espéraient les investisseurs au plus fort de l'euphorie ChatGPT mais il va se diffuser rapidement dans un proche avenir.

Dans le scénario où la révolution LLM se concrétiserait rapidement, **les investisseurs s'interrogent enfin sur les gagnants à moyen terme.**

L'Histoire nous incite à être prudent. Les marchés ont parié sur IBM lors de la révolution micro ou WorldCom lors de la révolution internet. Ils n'avaient pas anticipé l'éclosion de Microsoft ou Cisco.

Aujourd'hui, les investisseurs ont adopté une philosophie basée sur l'adage « the winner takes all », se référant à l'exemple des GAFAM qui ont prospéré grâce notamment à la puissance des effets de réseaux, à la création d'écosystèmes fermés et à des comportements monopolistiques.

Selon cette philosophie, les pionniers Nvidia et OpenAI auraient pris une avance telle qu'il serait difficile de les détrôner.

Mais le contexte est différent aujourd'hui. Comme indiqué précédemment, la technologie n'est pas figée et peut permettre la confirmation ou l'émergence d'acteurs déjà existants (Google, Meta, Amazon, ...), inconnus il y a peu (Anthropic, Mistral) ou qui vont être prochainement créés.

Par ailleurs, nonobstant les spécialistes du hardware (Nvidia, TSMC, ASML), les sociétés ayant investi massivement dans l'intelligence artificielle générative ne gagnent pas d'argent et ne savent pas quand, ni souvent comment, elles pourront en gagner.

Enfin, les sanctions américaines en matière d'exportations de produits technologiques vis-à-vis des entreprises chinoises vont créer une nouvelle

concurrence. L'Empire du Milieu dispose de talents et de capitaux qu'il ne faut pas sous-estimer et qui pourraient bouleverser le paysage actuel.

*

* *

Passée l'euphorie consécutive à la présentation de ChatGPT, les marchés financiers deviennent plus prudents. Il apparaît aujourd'hui que la technologie va bien bouleverser le quotidien des entreprises et des consommateurs. Mais elle doit devenir plus frugale. Elle doit également être vulgarisée pour être plus massivement adoptée. C'est un peu comme si Apple avait lancé son iPhone sans Appstore.

Pour prendre contact avec nos conseillers :

- ✉ info@maubourg-patrimoine.fr
- F. 01.42.85.80.00