



DOSSIER STRATÉGIQUE N°4

ChatGPT, OpenAI et intelligences artificielles : enjeux et perspectives d'une rupture technologique dont il faut se saisir !

“Lorsque souffle le vent du changement, certains construisent des murs, d'autres, des moulins.”



SOMMAIRE

| | |
|---|-----------|
| 1. Introduction et contextualisation..... | 3 |
| 1.1. Définition de l'outil..... | 3 |
| 1.1.1. Présentation..... | 3 |
| 1.1.2. Développement de l'écosystème..... | 3 |
| 1.2. ChatGPT, BARD, ERNIE, LLAMA les premiers d'une longue série ?..... | 4 |
| 1.2.1. BARD & GEMINI..... | 4 |
| 1.2.2. ERNIE..... | 5 |
| 1.2.3. LLAMA..... | 5 |
| 1.2.4. MISTRAL AI, l'outsider..... | 6 |
| 1.3. Opportunités induites..... | 7 |
| 1.3.1. Gains de productivité et déchargement des tâches à faible valeur ajoutée pour les humains..... | 7 |
| 1.3.2. Amélioration de la qualité des productions humaines..... | 7 |
| 1.3.3. Détachement des principaux biais humains dans le cadre de la prise de décision..... | 8 |
| 1.4. Risques et points d'attention..... | 9 |
| 1.4.1. Hacking et détournement de l'outil au profit des cybercriminels..... | 9 |
| 1.4.2. Protection des données personnelles et RGPD..... | 10 |
| 1.4.3. Propriété intellectuelle et dimension légale..... | 10 |
| 1.4.4. Politique de l'entreprise et risques réputationnels..... | 11 |
| 2. Vision stratégique de Cyberwings : le pas de côté..... | 12 |
| 2.1. Le coup de génie silencieux de Microsoft..... | 12 |
| 2.1.1. ChatGPT, une technologie OpenAI ou Microsoft ?..... | 12 |
| 2.1.2. Cartographie de l'écosystème OpenAI et implications..... | 15 |
| 2.1.3. Un faisceau d'indices semés sur la route du succès..... | 16 |
| 2.2. L'IA, l'éthique et la transparence..... | 17 |
| 2.2.1. Transparence et explicabilité..... | 17 |
| 2.2.2. Responsabilité et protection de la vie privée..... | 18 |
| 2.2.3. La singularité technologique..... | 18 |
| 2.3. Anticiper les prochaines évolutions, adapter nos usages..... | 19 |
| 2.3.1. IA et souveraineté française et européenne..... | 19 |
| 2.3.2. IA et la directive européenne NIS2 (Network and information security 2)..... | 22 |
| 2.3.3. Anticiper les prochaines évolutions..... | 22 |
| 2.3.4. Adapter nos usages : la valorisation du fruit de l'imperfection et de l'originalité..... | 23 |
| 3. Conclusion..... | 26 |



1. INTRODUCTION ET CONTEXTUALISATION

L'avènement de ChatGPT a bouleversé notre manière de concevoir le travail. Gain de temps, propos pertinents, optimisation des productions... ChatGPT semble être devenu l'outil ultime pour nous faire gagner en productivité. Pour autant, Sciences Po Paris a interdit l'usage de ChatGPT au sein de son institution. Plus généralement, le secteur universitaire commence à s'inquiéter face à une utilisation toujours plus importante de cet outil d'intelligence artificielle dans le cadre de productions écrites. Parallèlement et en opposition avec les craintes du secteur universitaire, le nombre de brevets déposés à l'étranger dans le domaine de l'IA connaît une croissance fulgurante, preuve que l'intelligence artificielle sera l'outil par excellence de notre avenir. De fait, il est nécessaire de nous adapter, mais surtout de s'intéresser aux implications concrètes qu'induit l'utilisation de cette technologie pour les organisations françaises et européennes.

1.1. Définition de l'outil

1.1.1. Présentation

Pour résumer simplement, ChatGPT est une API¹ utilisant l'intelligence artificielle générative (basée sur du *machine learning*²) pour générer du texte, inaugurée dans sa version *bêta* en novembre 2022. Plus spécifiquement, il s'agit d'un outil de traitement du langage naturel, c'est-à-dire le langage humain de parole, développé par l'entreprise mère OpenAI. Cette API permet de tenir une discussion à propos de n'importe quel sujet, et d'obtenir des informations précises ainsi que des réponses aux questions posées. La technologie utilisée pour ChatGPT (GPT-4³) est également réutilisée par OpenAI pour générer du code ou des images, ChatGPT n'étant qu'un moyen parmi d'autres d'utiliser cette technologie.

1.1.2. Développement de l'écosystème

Si l'outil présenté en novembre 2022 est révolutionnaire, cela n'empêche pas OpenAI de continuer à innover. Ainsi en ce premier mois de l'année 2024, la société a annoncé prochainement la mise en ligne d'une GPT store⁴. Concrètement, il s'agit de permettre aux développeurs de créer et de vendre leurs propres outils basés sur GPT, dont des chatbots capables d'effectuer diverses tâches en utilisant le langage naturel. Cette annonce répond en réalité au développement d'API plus classiques utilisant ChatGPT pour améliorer des produits ou faciliter la vente de solutions. À noter que ces outils ne sont parfois que de simples coquilles contenant un prompt. La future mise en ligne du GPT store constitue donc une réponse d'OpenAI afin de limiter le développement d'outils tiers sur lesquels la société aurait peu de capacités d'action.

- 1 Une API (application programming interface ou « interface de programmation d'application ») est une interface logicielle qui permet de « connecter » un logiciel ou un service à un autre logiciel ou service afin d'échanger des données et des fonctionnalités.
- 2 Le Machine Learning ou apprentissage automatique est un domaine scientifique, et plus particulièrement une sous-catégorie de l'intelligence artificielle. Elle consiste à laisser des algorithmes découvrir des « patterns », à savoir des motifs récurrents, dans les ensembles de données. Ces données peuvent être des chiffres, des mots, des images, des statistiques...
- 3 GPT-4 succède au modèle GPT-3 et 3.5. GPT 4 est en mesure de générer du code, traiter des images et comprendre 26 langues. C'est un modèle plus souple que GPT 3, et construit sur environ 1000 milliards de paramètres. GPT-4 a été lancé par OpenAI le 14 mars 2023, pour les abonnés de ChatGPT et les utilisateurs du moteur de recherche Bing.<https://fr.semrush.com/blog/gpt4-fonctionnalites-limites/>
- 4 <https://interestingengineering.com/innovation/you-can-buy-your-customized-version-of-chatgpt-starting-next-week>



1.2. ChatGPT, BARD, ERNIE, LLAMA les premiers d'une longue série ?

Après le succès planétaire d'OpenAI, les concurrents directs de Microsoft ont dû pour la plupart réagir en urgence et dévoiler plus ou moins distinctement leur stratégie en matière d'IA.

1.2.1. BARD & GEMINI

Le 6 février 2023, le CEO de Google et Alphabet Sundar Pichai a annoncé sur le blog officiel de Google l'arrivée sur le marché de sa nouvelle IA : BARD⁵. Bard se différencie de ChatGPT dans la mesure où il ne fonctionne pas sur le même modèle d'apprentissage. Si ChatGPT existe grâce à la technologie GPT3 (initialement, et désormais GPT4), Bard fonctionne sur la technologie développée par Google : LamDa. Cette technologie se distingue de celle d'OpenAI et Microsoft en fonctionnant sur une architecture de type « réseau neuronal » inventée par Google et disponible en libre accès depuis 2017⁶. Par ailleurs Bard contrairement à ChatGPT est connectée à internet, ce qui lui permet de pouvoir opérer une mise à jour constante de son système et de répondre à des questions d'actualité.

Cependant, 2 jours après cette annonce, le 8 février, la démonstration des capacités de Bard a tourné court. En effet, durant la promotion de ce nouvel agent conversationnel (pour répondre rapidement suite au succès de ChatGPT), l'IA Bard a commis une erreur factuelle dans sa réponse (la question étant de présenter à un enfant de 9 ans les découvertes permises par le télescope James Webb). Cette mauvaise réponse a entraîné une chute drastique de la cotation en bourse d'Alphabet de près de 100 milliards de dollars⁷.

L'échec de Google confirme notre hypothèse concernant la révélation de Bard. Cette révélation n'a eu lieu qu'en réaction au succès de ChatGPT, l'IA de Google n'était pas prête à être présentée au grand public. De fait, cela confirme la stratégie de long terme de Microsoft en la matière qui a concentré ses investissements de l'IA conversationnelle dans l'objectif final de l'intégrer à son moteur de recherche (*voir partie 2*).

Pour autant, cet échec n'a pas découragé Google puisque le 11 mai dernier, le géant du web au sein des GAFAM a annoncé l'intégration de Bard à son moteur de recherche. L'IA sera en état de détecter les cas où ce dernier est en mesure de répondre à une question et fournira une réponse rédigée en plus de la liste des résultats de recherche. À noter que cette innovation continue de répondre à celle de Microsoft qui a déjà intégré ChatGPT à son moteur de recherche BING, ainsi qu'à Microsoft Edge.

Il est également opportun de mentionner le développement par Google d'une seconde IA qui a fait parler d'elle très récemment (décembre 2023). En effet, annoncée en grande pompe, Gemini serait une IA générative ultra-performante basée sur LLM, autrement dit, un grand modèle de langage possédant un nombre considérable de paramètres, globalement de l'ordre du milliard. Elle serait ainsi capable de comprendre son environnement et les contextes, mais également de raisonner et même de saisir les références culturelles, tout cela en temps réel. L'annonce de Gemini s'est accompagnée d'un spot vidéo présentant l'IA tenant une conversation fluide et en direct avec une personne humaine. Elle serait même capable de reconnaître un tour de magie... Cependant, il a récemment été prouvé que cette vidéo est un bluff. Tout d'abord, la vidéo est présentée en accéléré : l'IA n'est donc pas capable de répondre « en temps réel ». De plus, la démonstration véritable s'est faite « en utilisant des images fixes tirées de la séquence, et des requêtes sous forme de texte », explique un porte-parole de Google⁸. Il n'est donc pas possible d'avoir une conversation

5 <https://blog.google/technology/ai/bard-google-ai-search-updates/>

6 <https://blog.google/technology/ai/lamda/>

7 Beaucoup d'articles ont été écrits à ce sujet que ce soit en anglais ou en français. Ici un article résumant assez bien la situation : <https://www.numerama.com/tech/1263488-cette-erreur-de-bard-a-fait-perdre-100-milliards-de-dollars-a-google.html>

8 <https://www.phonandroid.com/google-la-video-bluffante-de-lia-gemini-est-en-partie-fausse.html>



en direct avec Gemini.

Il est ainsi primordial de garder une distance raisonnable et un esprit critique face aux effets d'annonce concernant le développement de l'IA, quelles qu'en soient les origines. Enfin, il est nécessaire de préciser que les IA de Google ne sont toujours pas disponibles en Europe, ces dernières doivent encore apprendre à se conformer au RGPD.

1.2.2. ERNIE

L'annonce et le succès de ChatGPT ont également contraint la Chine à dévoiler son jeu en matière d'IA. Il semble en effet pertinent de rappeler que la Chine effectue un contrôle strict de son internet et dispose de son propre moteur de recherche censuré Baidu. Ainsi, depuis la révélation ChatGPT, la Chine a d'une part interdit son utilisation sur son territoire, mais vient également d'annoncer l'arrivée future d'un chatbot intégré au moteur de recherche, ainsi qu'un cloud nommé ERNIE. Peu d'informations étaient initialement disponibles à ce sujet, en raison du fait que cette IA n'avait été rendue accessible qu'à l'été 2023 (et présentée en mars 2023). Baidu a présenté ERNIE comme ayant pour mission de répondre de manière beaucoup plus fluide aux requêtes des utilisateurs, mais aussi de générer du contenu personnalisé, ou encore de mieux appliquer la censure du régime. Nous devons toutefois noter que l'accès à cette IA est très limité en raison de la censure et des verrous imposés par l'état chinois⁹.

En outre, il est important de noter l'essor pris par la Chine et Baidu en matière de dépôt de brevet, spécifiquement en ce qui concerne l'apprentissage automatique et l'IA¹⁰. Et cela n'est pas près de s'arrêter. On peut donc facilement imaginer que le développement d'ERNIE tendra à amener son lot d'innovations et bouleverser la course à l'IA¹¹ dans les prochains mois.

Même s'il demeure difficile aujourd'hui d'anticiper qu'elle sera la politique opérationnelle chinoise sur le développement, l'industrialisation et la commercialisation de cette IA, nous pouvons facilement avancer que cette technologie s'inscrit pleinement dans la stratégie d'expansion et d'influence de la Chine aux plans économique, (géo)politique, et même psychologique. En effet, au-delà de constituer une condition nécessaire à la survivance des industriels chinois dans le bras de fer géo-économique sino-américain, les intelligences artificielles chinoises participent également aux « opérations dans le domaine cognitif »¹² visant à replacer le narratif du Parti communiste chinois (PCC) au centre de la pensée des citoyens, de la diaspora mais également de la communauté internationale dans son ensemble. Le but ultime des opérations dans le domaine cognitif, tel que théorisé par l'Armée populaire de libération (APL) est de manipuler les valeurs, l'esprit/l'éthos national, les idéologies, les traditions culturelles, les croyances historiques, etc., d'un pays ou d'un groupe pour les inciter à abandonner leur compréhension théorique, leur système social et ainsi d'atteindre des objectifs stratégiques sans combattre. L'utilisation d'intelligences artificielles génératives pour produire des discours convergents, relayés par des milliers de faux profils numériques (dont des figures diplomatiques) eux-mêmes générés par IA, en est un exemple frappant.

1.2.3. LLAMA

Llama 2 est une IA générative développée par le groupe Meta. L'originalité de Llama repose sur le fait qu'elle soit basée sur de l'open source et est gratuite même pour une utilisation commerciale. Ce modèle a été révélé par Meta en juillet 2023 et a le mérite de se placer en concurrent direct de Google. Cependant, il convient de

9 <https://www.courrierinternational.com/article/chine-ernie-le-chatbot-de-baidu-qui-refuse-de-repondre-a-des-questions-sensibles>

10 Source : INPI - https://www.inpi.fr/sites/default/files/intelligence_artificielle_presentation.pdf

11 <https://fr.statista.com/infographie/24982/innovation-intelligence-artificielle-nombre-brevets-ia-machine-learning-par-entreprises/>

12 Paul Charon (directeur du domaine « Renseignement, anticipation et stratégies d'influence » de l'IRSEM) et Jean-Baptiste Jeangène Vilmer (directeur de l'IRSEM) – Les opérations d'influence chinoise, un moment machiavélien. 2021 – Institut de recherche stratégique de l'École militaire (IRSEM) https://drive.google.com/file/d/11XOzrPy3z_1nRsPlKiqktjbgK5aqMClq/view



préciser que Llama (et donc Meta) sont les partenaires privilégiés de Microsoft - elle est d'ailleurs optimisée pour fonctionner localement sur Windows - on peut donc s'attendre à ce que cette IA reste potentiellement en dessous des capacités de ChatGPT afin de limiter la concurrence directe. L'IA de Meta se veut « responsable » et « éthique », assure qu'aucune information venant de ses utilisateurs n'a servi à son entraînement et que les sites contenant d'importantes quantités de données personnelles en ont été exclus¹³.

1.2.4. MISTRAL AI, l'outsider

Créée à l'été 2023 une start-up française essaye elle aussi de tirer son épingle du jeu et propose son propre modèle d'IA générative. Mistral AI souhaite proposer une alternative européenne aux IA des GAFAM avec de premiers modèles qui devraient être accessibles au premier trimestre 2024. Pour réussir son coup, la start-up vient de lever 105 millions d'euros afin de proposer des modèles Open source, respectueux du RGPD et à destination des entreprises¹⁴.

Il faut toutefois rester vigilant : ironie du sort, les trois chercheurs français à l'origine du projet étaient préalablement en poste chez Meta et Google DeepMind... Par ailleurs la levée de fonds est conduite par un fonds d'investissement américain, et d'autres pontes des GAFAM ont également contribué à son financement, à l'image de l'ancien CEO de Google Éric Schmidt¹⁵.

13 <https://www.usine-digitale.fr/article/meta-lance-llama-2-un-grand-modele-de-langage-open-source-et-gratuit-meme-pour-une-utilisation-commerciale.N2154492>

14 <https://www.usine-digitale.fr/article/la-start-up-francaise-mistral-ai-leve-105-millions-d-euros-pour-developper-des-modeles-d-ia-generative.N2142627>

15 https://www.lemonde.fr/economie/article/2023/12/11/la-start-up-francaise-mistral-ai-a-leve-385-millions-d-euros_6205065_3234.html



1.3. Opportunités induites

Les intelligences artificielles, lorsque correctement définies, délimitées et régulées en fonction de leurs champs d'application, peuvent représenter de réelles opportunités. En l'espèce, celles-ci sont offertes par ChatGPT dans ses différentes dimensions et sont intégrées à part entière dans les discours et la stratégie de communication des acteurs qui portent ce projet.

Synthétiquement, trois grandes catégories de bénéfices peuvent être retenues pour les intelligences artificielles génératives.

1.3.1. Gains de productivité et déchargement des tâches à faible valeur ajoutée pour les humains

Sur la base des technologies décrites, l'IA peut accroître la productivité des organisations en déchargeant les humains de la réalisation de tâches de raisonnement et de production non complexes, avec par exemple la démocratisation de l'assistance vocale ou l'utilisation de traducteurs automatiques.

Le principe d'apprentissage – s'apparentant à un acte de mimétisme de nouvelle génération - permet aux ordinateurs de produire des raisonnements et productions inédits, sur la base des informations qu'ils ont ingérées dans des contextes comparables. À ce titre les gains de productivité résultent en la différence entre les coûts anciennement alloués au travail humain sur les tâches confiées à l'IA, et les coûts liés à l'entraînement de l'IA (supportés en interne ou inclus dans le prix des licences) et à la vérification / validation desdites tâches. On peut par exemple penser à une meilleure optimisation de la chaîne d'approvisionnement grâce à l'analyse de données en temps réel et donc la diminution des postes dans ce secteur¹⁶. Ces faits sont d'ailleurs corroborés par une étude menée conjointement par l'Université de Harvard (département business) et le Boston Consulting Group. Concrètement, les conclusions de l'étude sont claires : « pour chacune des 18 tâches de conseil réalistes situées à la frontière des capacités de l'IA, les consultants utilisant l'IA étaient nettement plus productifs : ils réalisaient 12,2 % de tâches en plus en moyenne et 25,1 % de tâches en moins et produisaient des résultats d'une qualité nettement supérieure (plus de 40 % de qualité en plus par rapport au groupe de contrôle)¹⁷ ». Il est toutefois nécessaire de noter que les profils et compétences nécessaires à l'intégration d'une IA dans un métier sont différents des profils préalablement responsables de l'exécution des tâches. Cela pose de fait des questions d'ordre RH au sein de l'entité, et des interrogations éthiques, économiques et politiques à l'échelle de nos sociétés.

1.3.2. Amélioration de la qualité des productions humaines

La force de l'intelligence artificielle repose sur la quantité des informations - condensées dans des bases de données statiques gigantesques, voire collectées sur internet de façon dynamique - ingérées par les ordinateurs pour reproduire des raisonnements et réaliser des tâches. La quantité et la variété de données permettant de guider ce raisonnement sont en effet nettement supérieures à celles qu'un cerveau humain est en capacité de traiter. Les résultats obtenus sont, en théorie, d'une précision et d'une qualité plus grande, à condition bien sûr que le périmètre source de l'IA ait été délimité correctement au regard de l'activité concernée. Sur du moyen/long terme, on peut aisément envisager le développement de systèmes de diagnostics médicaux automatisés plus précis¹⁸, ou bien dans le domaine marketing,

16 À ce sujet, lire l'article de Sébastien Rivière sur l'automatisation de la supply chain par Amazon : <https://news.sebastienriviere.com/blog/amazon-et-lautomatisation-des-t%C3%A2ches>

17 Dell'Acqua, Fabrizio and McFowland, Edward and Mollick, Ethan R. and Lifshitz-Assaf, Hila and Kellogg, Katherine and Rajendran, Saran and Krayer, Lisa and Candelon, François and Lakhani, Karim R., Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality (September 15, 2023). Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-013, Available at SSRN: <https://ssrn.com/abstract=4573321> or <http://dx.doi.org/10.2139/ssrn.4573321>

18 Dans ce domaine, l'IA est belle est bien révolutionnaire. Voir articles suivants : [IBM big data used for rapid diagnosis of rare leukemia case in Japan | The Japan Times](#) ; [Santé. Au CHU de Strasbourg, l'intelligence artificielle aide à](#)



une analyse beaucoup plus fine et rapide de la data dans l'objectif final de mieux appréhender les comportements des consommateurs et donc d'optimiser le ciblage et l'efficacité des campagnes marketing¹⁹.

1.3.3. Détachement des principaux biais humains dans le cadre de la prise de décision

Lorsque produits par des ordinateurs nourris de données factuelles, les raisonnements et productions sont délestés des biais psychologiques, cognitifs, sociaux qui sont propres aux personnes humaines et à leur conditionnement. En ce sens, l'IA permet, dans certaines configurations, de garantir une meilleure impartialité et neutralité. À cela s'ajoutent la stabilité et la régularité des raisonnements, des choix et des productions opérés par une IA, face à des réalisations humaines potentiellement plus volatiles. Dépourvue de variables physiques et contextuelles (comme la fatigue, le stress le surmenage, etc.), une intelligence artificielle produira des résultats conditionnés uniquement aux variables qu'elle a appris à traiter / corrélérer.

Pour autant, cette constance théorique ne peut occulter les biais observés dans les pratiques de production des IA génératives. Ces derniers résultent d'une part de l'essence même des algorithmes d'IA dont les fondements sont le fruit d'un travail (initialement) humain et contextualisé. D'autre part, ces biais intrinsèques s'expliquent par la nature des informations ingérées, découlant elles-mêmes de réalités sociales, économiques, politiques et linguistiques émanant des Hommes²⁰.

A ce stade de notre analyse, il nous apparaît ainsi opportun de revenir plus en détail sur ces biais en explorant les risques et dérives inhérents à cette technologie.

[opérer les patients fragiles pour la première fois en France \(dna.fr\)](https://dna.fr) ; [L'intelligence artificielle n'assistera pas la médecine, elle sera la médecine de demain \(la-croix.com\)](https://la-croix.com)

¹⁹ Sur la relation entre l'IA et le big data voir : <https://enterpriseproject.com/article/2019/10/how-big-data-and-ai-work-together>

²⁰ Cette notion sera questionnée dans la seconde partie de ce dossier, notamment en prenant en considération la notion de *singularité technologique*.



1.4. Risques et points d'attention

Si le développement de l'IA dans le monde de l'entreprise est souvent perçu comme une opportunité permettant d'optimiser les gains de productivité, il demeure cependant primordial de garder à l'esprit les limites imposées par ces nouvelles technologies. Les effets indésirables pouvant parfois surpasser les bénéfices escomptés.

1.4.1. Hacking et détournement de l'outil au profit des cybercriminels

L'innovation en termes d'IA qu'est ChatGPT conduit tout naturellement les cybercriminels à détourner cet outil à leur profit, les possibilités d'usage étant presque infinies. En effet, l'API d'OpenAI est en mesure de générer du code au moyen d'une simple phrase de commande, et ce dernier peut en conséquence être réutilisé dans la création de malwares. ChatGPT peut ainsi permettre aux pirates de gagner un temps considérable dans l'élaboration de leur code (malveillant) et pourquoi pas d'en « industrialiser » l'usage. De plus, s'il est impossible de demander à ChatGPT de créer un mail de phishing et de l'envoyer, il est possible de demander l'IA de rédiger un courriel en lieu et place d'une entité et de s'en servir par la suite dans une campagne d'hameçonnage²¹. Malgré des garde-fous théoriques, il est en effet possible de détourner l'usage de ChatGPT à des fins malveillantes voire criminelles. En outre, des déclinaisons « noires » à ChatGPT se développent désormais, à l'instar de WormGPT²², et viennent confirmer la notion de « double usage »²³ des technologies numériques.

Dans une autre mesure, la question de la sécurité de l'IA en elle-même (sécurité de son code source principalement) est primordiale.

Corrompre une IA peut résulter en trois conséquences et trois risques notables. Premièrement, la possibilité de modifier l'algorithme et de perturber les « raisonnements » et choix de l'IA. Cela induit alors des risques opérationnels (influence sur les métiers), mais également en matière d'atteinte à la sécurité physique, et à la dignité humaine²⁴. Ces risques sont directement liés au phénomène de l'IA poisoning : les attaques par empoisonnement visant à modifier le comportement du système d'IA en introduisant des données corrompues en phase d'entraînement (ou d'apprentissage) dans l'objectif final d'infiltrer et donc de compromettre un système d'information (SI)²⁵.

Corrompre une IA peut également conduire à une exfiltration de données concernant ses utilisateurs, mais également à l'exfiltration des contenus ingérés et produits. Cette réalité pose ainsi la question de la possibilité -ou non- de garantir la confidentialité des données (personnelles, sensibles, commerciales) traitées par l'IA (en entrées ou sorties) et du respect de la propriété intellectuelle et industrielle. Pour se prémunir contre ces dérives, certaines solutions techniques (à l'instar de la pratique du watermarking, du chiffrement des données sensibles, etc.) ainsi que des solutions juridiques (comme le dépôt de titres de PI ou encore la mise en œuvre des mécanismes de préservation du secret des affaires²⁶) existent. Toutefois, ces dernières ne peuvent être efficaces que lorsqu'elles sont combinées²⁷.

Enfin, le détournement d'une IA et son intégration dans un réseau d'ordinateurs corrompus - *botnet* – sont des risques non négligeables. Ces techniques offensives

21 Sur ce thème, voir : <https://urlz.fr/kzUA>

22 <https://www.zdnet.fr/actualites/ce-qu-il-faut-savoir-a-propos-de-wormgpt-la-reponse-des-cybercriminels-a-chatgpt-39960858.htm>

23 Pour reprendre la terminologie juridique utilisée dans l'accord de Wassenaar pour qualifier les biens et technologies pouvant être destinés à un usage dual, à la fois civil et militaire.

24 <https://www.lefigaro.fr/actualite-france/chatgpt-un-belge-se-suicide-apres-avoir-trouve-refuge-aupres-d-un-robot-conversationnel-20230329> / https://www.bfmtv.com/tech/intelligence-artificielle/j-aimerais-te-voir-mort-eliza-l-ia-qui-aurait-conduit-un-homme-au-suicide_AV-202303300439.html

25 <https://www.youtube.com/watch?v=ghVWFZ5esnU>

26 LOI n° 2018-670 du 30 juillet 2018 relative à la protection du secret des affaires

27 Pint Avocats, « Vous avez des secrets ? Protégez-les ! », Marseille Innovation, octobre 2023 <https://www.marseille-innov.org/start-ups-vous-avez-des-secrets-protégez-les/>



peuvent permettre à un individu ou un groupe malveillant de perpétrer des attaques « rebond » afin de causer des dommages sur une cible finale de façon furtive, à l'image des attaques par dénis de service distribué (DDoS). Ce cas appliqué au monde informatique pose finalement la question plus générale de l'imputabilité de la responsabilité juridique en cas de dérive d'une IA, au cœur des débats et des initiatives de régulation à différentes échelles, et notamment européenne²⁸.

1.4.2. Protection des données personnelles et RGPD²⁹

La politique d'OpenAI en matière de protection des données personnelles est claire : l'entreprise se donne le droit de collecter et de stocker la quasi-totalité des données personnelles des utilisateurs et de les partager avec un « tiers³⁰ » sans les en informer. Le risque de fuite de DCP³¹ est donc réel. Par ailleurs, la potentialité forte de trouver des failles au sein de l'API ChatGPT accentue de fait le risque de compromission et de fuite pour les données personnelles des utilisateurs. Parallèlement, les données à caractère personnel (DCP) des utilisateurs, quelle que soit leur nationalité, sont exploitables tant par les entreprises « tierces » que par l'État américain lui-même. En application du concept d'extraterritorialité, la loi FISA³² et le Cloud Act³³ permettent en effet aux autorités américaines d'exploiter les données personnelles des utilisateurs non américains dès lors qu'ils utilisent une technologie américaine, qu'elles soient hébergées sur le territoire des États-Unis ou non. Dès lors, en acceptant d'utiliser ChatGPT, on autorise le gouvernement des États-Unis à pouvoir accéder librement à ces informations.

1.4.3. Propriété intellectuelle et dimension légale

L'avènement de ChatGPT est aussi à analyser sous l'angle juridique. Premièrement, nous pouvons souligner l'existence de sanctions judiciaires encourues par les clients utilisant l'API, dans la mesure où ces derniers peuvent être tenus responsables des actions réalisées par leurs chatbots³⁴. Cette réalité, couplée au fait que l'IA ne soit pas infaillible, pose de réelles questions de responsabilité et de couverture, dans le cadre de cyberattaques et de détournement.

Dans une autre mesure, il existe un véritable enjeu pour les sciences, l'industrie et les productions innovantes. D'une part, il va devenir de plus en plus difficile d'identifier et de singulariser la production humaine de la production de l'IA³⁵. D'autre part, la protection de la propriété intellectuelle (dont la propriété industrielle et la protection des droits d'auteurs et productions littéraires et artistiques) peut être mise à mal par l'IA. ChatGPT ayant assimilé une immense quantité de bases de données - parfois de manière illégale -, il est urgent pour les chercheurs de protéger leurs recherches afin d'éviter qu'elles ne se soient collectées et stockées sur les serveurs d'OpenAI et donc in fine ingérées pour nourrir le processus d'apprentissage.

28 De nouvelles règles en matière de responsabilité applicables aux produits et à l'IA pour protéger les consommateurs et favoriser l'innovation ; https://france.representation.ec.europa.eu/informations/de-nouvelles-regles-en-matiere-de-responsabilite-applicables-aux-produits-et-lia-pour-protger-les-2022-09-28_fr

Propositions européennes visant à réguler l'IA synthétisées ci-après : <https://digital-strategy.ec.europa.eu/fr/policies/european-approach-artificial-intelligence>

29 Règlement Général sur la Protection des Données

30 Des exemples de « tiers » sont intégrés à la cartographie produite dans la partie 2.1.2

31 Donnée à caractère personnel

32 En vertu de la norme FISA 702, les "fournisseurs de services de communication électronique" américains peuvent être contraints de donner aux autorités de sécurité américaines l'accès aux données à caractère personnel des "personnes non américaines", qui sont définies comme toute personne qui n'est pas un citoyen américain ou un résident permanent des États-Unis.

33 Le Cloud act est une loi fédérale américaine promulguée en mars 2018 qui étend la portée géographique des demandes éventuelles du gouvernement américain à pouvoir accéder aux données sur les serveurs, quelle que soit leur localisation.

34 En France, pour incriminer l'IA elle-même, il serait nécessaire d'un point de vue juridique de créer une personnalité juridique pour le robot, ce qui n'est pas envisagé pour le moment. Logiquement, la responsabilité des dommages causés par un système d'IA incomberait à son propriétaire. Pourtant, il ne faut pas exclure que l'IA puisse échapper au contrôle de son propriétaire et que, par conséquent, il n'y ait pas de responsabilité de principe. Deux visions s'opposent alors : une responsabilité dite diffuse (ou conjointe) comportant plusieurs inconvénients dans la mesure où tous les acteurs liés à l'IA seraient responsables en cas de dommage. Une autre vision serait de dire que la sanction serait celle de l'exploitant uniquement, car c'est lui qui bénéficie de l'utilisation de l'IA. Ce concept est confirmé par l'article 1242 du Code civil, paragraphe 1, qui justifie que nous sommes responsables des choses que nous avons sous notre garde.

35 Désormais, certaines IA sont même éligibles au statut de « inventeur » dans le cadre de dépôt de brevet. -E. Feyeux, A. Chemin, R. Segura, *Artificial intelligence, friend or foe ?* Sciences Po Aix, décembre 2021. Source: *Dépôt de brevet : l'IA peut-elle être l'inventeur ?* Dalloz actualité, le 11.10.2021.



Dans quelle mesure cette IA sera en capacité de respecter la propriété intellectuelle (PI) à l'avenir, sachant que des technologies centrales comme ChatGPT ont déjà eu accès à des bases de données « protégées » par ces mêmes titres de PI sans en assumer les contreparties ? Cette question, que nous avons posée en janvier 2022, commence à trouver des réponses en 2024. En effet, le 28 décembre 2023 le New York Times a choisi d'intenter une action en justice contre OpenAI et Microsoft afin de mettre un terme à la pratique consistant à utiliser ses articles pour former des IA³⁶. Le célèbre journal américain affirme que les violations de droits d'auteur au sein du journal pourraient à elles seules valoir des milliards. La décision des juges devrait servir de jurisprudence à bien des égards. À noter que le NYT n'est que la pointe émergée de l'iceberg en ce qui concerne les procès et plaintes intentés à l'encontre d'OpenAI. La comédienne et autrice américaine Sarah Silverman poursuit également OpenAI affirmant que ses livres ont été ingérés pour entraîner les modèles d'IA d'OpenAI sans autorisation. En juin, plus de 4 000 écrivains ont ainsi signé une lettre ouverte adressée aux PDG d'OpenAI, de Google, de Microsoft, de Meta et d'autres développeurs d'IA, les accusant de pratiques d'exploitation dans la construction de chatbots qui "imitent et régurgitent" leur langue, leur style et leurs idées. L'affaire est donc à suivre.

1.4.4. Politique de l'entreprise et risques réputationnels

Non connecté à internet, ChatGPT a eu accès à une quantité astronomique de bases de données jusqu'en janvier 2022. De fait, tout ce qui a été indexé sur internet à partir de janvier 2022 n'est pas connu de cette IA. On ne peut donc ignorer le risque de diffusion de fausses informations à ce jour, ce qui pose un réel problème de fiabilité. Par ailleurs, même connectée à internet, le risque de diffusion de fausses informations reste important : il ne peut y avoir de garantie de fiabilité à cet égard. Ainsi, si les partisans de ChatGPT avancent des gains de productivité pour les entreprises, il convient de rappeler qu'un travail de *Fact checking* est plus que jamais nécessaire si les informations produites par ChatGPT ont ensuite vocation à être utilisées dans l'organisation. Il ne faut pas non plus ignorer le risque important de biais de l'IA qui peut, selon le type de question posé, proposer une réponse connotée, voire empreinte de xénophobie, d'homophobie, de complotisme, etc.

Par ailleurs, alors que l'Europe fait face à une crise énergétique sévère, que le prix de l'électricité augmente drastiquement et que les enjeux du réchauffement climatique ne sont plus à prouver, il convient de rappeler que l'utilisation de ChatGPT entraîne une consommation importante d'électricité. Ce paramètre peut avoir des conséquences financières, mais peut également être un problème dans le cadre de la politique RSE³⁷ des entreprises.

D'un point de vue réputationnel enfin, il est à noter que l'exploitation irrégulière des DCP par OpenAI et ses partenaires (et les possibles abus qui pourront en découler), peut être extrêmement délétère pour les organisations, pouvant se trouver au cœur de scandales du même acabit que l'affaire *Cambridge Analytica*³⁸.

36 <https://fr.euronews.com/next/2023/12/28/le-new-york-times-poursuit-openai-et-microsoft-pour-violation-de-droits-dauteur>

37 Responsabilité Sociétale des Entreprises

38 Pour plus de précision, à ce sujet, voir l'article du Monde qui résume parfaitement la situation : https://www.lemonde.fr/pixels/article/2018/03/22/ce-qu-il-faut-savoir-sur-cambridge-analytica-la-societe-au-c-ur-du-scandale-facebook_5274804_4408996.html



2. VISION STRATÉGIQUE DE CYBERWINGS : LE PAS DE CÔTÉ

L'intelligence artificielle ChatGPT n'est pas simplement issue d'une rencontre technique et capitalistique heureuse entre OpenAI et Microsoft. L'association entre la start-up récemment éclosée et le maître des GAFAM est en réalité le point culminant d'une stratégie de long terme, imaginée et mise en œuvre par Microsoft, depuis près de 10 ans. Silencieusement et brique après brique, Microsoft a mis en place toutes les conditions favorables à la supplantation de ses concurrents directs grâce au lancement éclair et inattendu d'une IA particulièrement performante. Ce travail de préparation a été opéré de façon furtive, à différents niveaux : juridique, technique, financier, et commercial.

Tandis que l'ensemble des analystes et médias se limitent aujourd'hui à décrire les fonctionnalités, les opportunités et les menaces que représente ChatGPT, la cellule de veille stratégique de Cyberwings a cherché à décrypter les fondements du projet de long terme de Microsoft en vue :

- de donner à ses lecteurs l'ensemble des clés de compréhension des évolutions extrêmement rapides auxquelles nous assistons en matière d'IA. Ceci suppose de s'appuyer sur l'analyse et le croisement de sources primaires, de considérer des données factuelles et neutres, et enfin de prendre un certain recul sur le flux d'information relayé par l'ensemble de la communauté médiatique.
- d'anticiper les évolutions probables en matière d'intelligence artificielle et ainsi se préparer de manière efficace aux défis futurs qui se présenteront à nous.

2.1. Le coup de génie silencieux de Microsoft

Après des années de domination de la part de Google en matière de moteurs de recherche (Google détient 93 % des parts de marchés dans ce domaine³⁹), Microsoft espère pouvoir inverser la tendance avec ChatGPT et son intégration dans la majorité de ses outils.

Quelle stratégie a, de fait, été mise en place par ce géant du numérique afin de concrétiser son dessein à long terme ?

2.1.1. ChatGPT, une technologie OpenAI ou Microsoft ?

Retour sur les fondements historiques d'OpenAI

Basée à San Francisco aux États-Unis, OpenAI est fondée en décembre 2015 et arbore un statut associatif à but non lucratif. Originellement, OpenAI est un centre de recherches dédié à l'intelligence artificielle dont l'objectif premier est d'exploiter toutes les possibilités à impact positif que les IA peuvent offrir à la société. Pour autant, l'association change de statut en 2019 pour devenir une entreprise à but lucratif plafonné afin de pouvoir continuer à financer sa R&D. La même année Microsoft choisit d'investir environ 1 milliard de dollars au sein d'OpenAI (investissement réitéré en 2021 puis à hauteur de 10 milliards le 22.01.2023) et finit par détenir la licence propriétaire de ChatGPT1. Cette stratégie d'investissement dans l'IA est en réalité une stratégie de long terme. En effet, Microsoft a acquis la licence d'exploitation exclusive de la technologie GPT-3 dès 2020. L'objectif final étant pour Microsoft d'intégrer ChatGPT à l'ensemble de ses outils, de la suite Office 365 jusqu'au moteur

³⁹ <https://www.blogdumoderateur.com/chiffres-google/>



de recherche Bing, en passant par l'univers AZURE. Dès lors, si BING semblait jusque là désuet et ne plus attirer les foules d'internautes, il s'agissait finalement des conséquences temporaires d'un choix volontaire de Microsoft dans le cadre de sa stratégie de concentration des investissements autour de l'application de l'IA aux moteurs de recherche. Concrètement, il s'agissait pour Microsoft de conserver l'ensemble des utilisateurs de la gamme Office, de les acculturer à l'utilisation de l'IA et par conséquent de les orienter vers de nouveaux usages, tous intégrant ChatGPT, en vue notamment de récupérer une part significative du marché des moteurs de recherches au détriment de son concurrent (et leader) Google.

Retour sur les récentes évolutions capitalistiques opérées dans le cadre de la stratégie d'investissement de Microsoft

Afin de saisir les récents enjeux autour du « départ » de S. Altman et G. Brockman du board de OpenAI, nous proposons dans un premier temps d'opérer une courte chronologie des faits.

- Vendredi 17 novembre 2023 : S. Altman est licencié d'OpenAI par le Conseil d'Administration pour manquement à ses obligations. Parallèlement, G. Brockman se voit retirer son titre de Président. Il choisira de démissionner le même jour. Microsoft premier investisseur de la société est « mis devant le fait accompli ».
- Samedi et dimanche 18 et 19 novembre 2023 : Pressions de Microsoft sur OpenAI pour faire réintégrer S. Altman.
- Lundi 20 novembre 2023 : Des tractations sont en cours entre Altman, Brockman et Microsoft afin que les deux premiers intègrent les équipes du géant numérique. Échanges de messages mélodramatiques sur X avec l'expression des regrets d'Ilya Sutskeve - membre du CA d'OpenAI - d'avoir voté en faveur de l'éviction d'Altman.
- Mardi 21 novembre : Après plusieurs jours de négociations, S. Altman retrouve sa place chez OpenAI en dépit de l'intervention d'E. Musk en défaveur d'Altman. Le CA est renouvelé, et Altman annonce sa volonté de renforcer ses relations avec Microsoft.

Comme l'ont souligné plusieurs analyses⁴⁰, les dernières évolutions et les mouvements au sein d'OpenAI semblent s'être déroulés à la faveur de Microsoft qui resserre son emprise sur cette dernière. Il est cependant important d'apporter quelques nuances.

Tout d'abord, Microsoft n'a pas attendu le départ de S. Altman d'OpenAI pour avoir une influence sur la jeune société, comme en attestent les liens capitalistiques entre les deux entités. Par ailleurs, nous souhaitons également revenir sur le « recrutement d'Altman chez Microsoft » qui, de notre point de vue, relève davantage de la mise en scène. Altman avait en effet déjà vendu sa société LinkedIn à Microsoft quelques années auparavant, révélant des relations préalables extrêmement solides entre les deux acteurs. Nous pouvons en ce sens émettre l'hypothèse d'une stratégie entre Altman et Microsoft pour augmenter le niveau de pression exercée sur OpenAI. L'objectif final étant de réintégrer Altman et de renforcer la mainmise de Microsoft sur la jeune pépite. Cette hypothèse semble d'ailleurs validée par les déclarations du mardi 21 novembre 2023 de S. Altman lui-même.

D'autres analyses argumentent que cet évènement démontre la volonté d'OpenAI de s'émanciper de Microsoft en raison d'une dépendance accrue de la start-up au géant de l'informatique. Pour autant, la cellule de veille et d'analyse stratégique de Cyberwings ne partage pas cette vision. Comme exprimé précédemment, l'emprise de Microsoft sur OpenAI existe depuis les premiers investissements de 2019 et l'accord de la licence exclusive de ChatGPT. Par conséquent l'hypothèse de la création d'un quasi-monopole de l'IA par Microsoft n'est pas dénuée de sens. La société dirigée par Satya Nadella détient déjà une IA effective implémentée dans BING,

⁴⁰ https://www.linkedin.com/posts/guillaume-champeau_microsoft-reportedly-plans-to-invest-10-activity-7132405962212024321-VmOu/ ; <https://www.courrierinternational.com/article/tech-c-est-le-fiasco-de-la-decennie-apres-le-depart-d-altman-les-salaries-d-openai-menacent-de-rejoindre-microsoft>



possède la mainmise sur ChatGPT, développe en parallèle le projet Co-pilote, etc. Elle contrôle certains des principaux algorithmes d'IA. L'idée d'un monopole en devenir n'est donc pas à balayer. Dans la poursuite de ces arguments, il faudra également s'interroger sur la possible création d'un « cartel de l'IA » entre les maîtres des GAFAM.

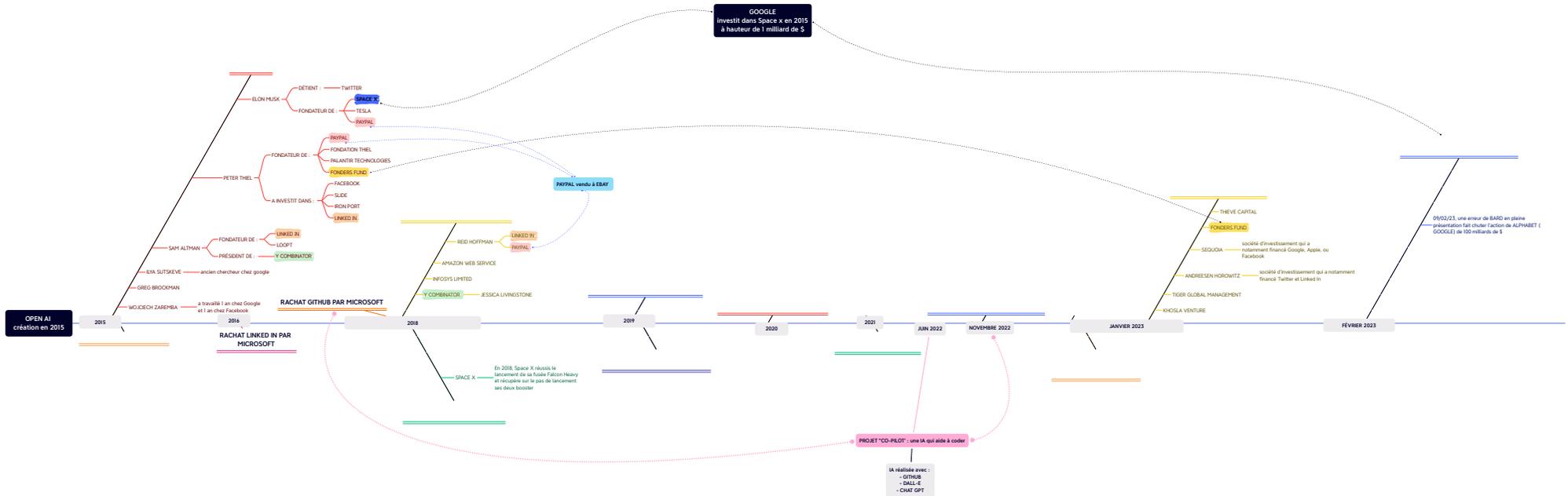
Le dernier argument dont nous souhaiterions débattre est celui d'un avenir incertain (voire la disparition) d'OpenAI, point de vue que Cyberwings ne partage pas. Il est en effet important de garder à l'esprit qu'OpenAI conduit des programmes de recherche fondamentale qui font évoluer le domaine de l'IA. Dès lors, s'il est possible qu'OpenAI disparaisse sous sa forme entrepreneuriale et commerciale, les actions de R&D seront très probablement poursuivies via la création d'une nouvelle entité, voire directement par Microsoft. Ceci ne résulterait pas en une « fin » de l'aventure OpenAI. En effet, il n'y a pas de concurrence directe entre OpenAI et Microsoft. Nous pouvons avancer qu'il s'agit pratiquement de la même entreprise ou du moins qu'il existe une intégration quasi parfaite des deux entités dans le cadre d'un partenariat stratégique au sein duquel Microsoft co-conduit et co-dirige les recherches et orientations d'OpenAI.

En outre, si notre hypothèse du cartel est amenée à se vérifier, Microsoft aurait tout intérêt à conserver « la marque » OpenAI...



2.1.2. Cartographie de l'écosystème OpenAI et implications

Source : Cyberwings



Cartographie accessible en annexe (format .pdf) afin de pouvoir zoomer sur les différents éléments.



2.1.3. Un faisceau d'indices semés sur la route du succès

Le succès de ChatGPT avec en arrière plan, le succès de Microsoft, invite à opérer une analyse rétrospective à l'image d'une analyse inforensic, afin de déchiffrer les différents indices laissés par ces mêmes entreprises explicitant leur succès et la révolution qu'elles viennent d'initier.

Les indices techniques

De prime abord, on constate un abandon de l'investissement des géants du web dans le domaine des moteurs de recherche. En effet, depuis l'arrivée de Bing de Microsoft pour concurrencer Google, aucune innovation majeure en la matière n'est à souligner. Les améliorations ne reposent pas tant sur l'aspect révolutionnaire dans la manière de concevoir la recherche, mais sur d'autres aspects moins novateurs. D'une part, on constate que les principaux moteurs de recherche tentent d'améliorer l'expérience de l'utilisateur avec par exemple l'ajout de nouvelles fonctionnalités comme le défilement continu ou l'amélioration de la recherche par image chez Google. D'autre part et afin de se distinguer des géants du marché, de petits moteurs de recherche cherchent à promouvoir une « recherche internet », plus vertueuse. À cet égard nous pouvons mentionner des outils comme Écosia (biodiversité, arbres plantés à chaque recherche, pas de rétribution des actionnaires), Lilo (moteur de recherche français qui finance des associations) ou Qwant (moteur de recherche qui refuse le traçage de ses utilisateurs, anonymat sur le web, etc.). Ainsi, à défaut de pouvoir miser sur l'efficacité, ces nouveaux moteurs jouent davantage sur l'aspect éthique et vertueux de leur application.

Pour compléter ce constat, nous pouvons nous concentrer sur l'étude des RFC (Request for comment) relatifs aux moteurs de recherche. Les RFC sont une série numérotée de documents décrivant les aspects et spécifications techniques d'internet ou de différents matériels informatiques et qui ont vocation à faire évoluer les règles protocolaires sur le web et l'internet. Ces documents sont travaillés et rédigés à l'initiative d'experts techniques - à l'image d'une « task force » - puis sont revus par la communauté internet dans son ensemble. À noter que ces experts sont parfois issus de grands groupes informatiques ou des GAFAM comme Microsoft, Google, ou Amazon. En ce qui concerne les moteurs de recherche, et après étude des derniers RFC publiés sur le sujet ces dernières années, nous constatons que Microsoft ne soumet plus aucun RFC à ce sujet, et ce depuis 2020. Nous pouvons ainsi continuer de soutenir notre hypothèse selon laquelle Microsoft a cessé d'investir dans la recherche autour des moteurs de recherche et s'est donc retiré des task forces, à la faveur du développement de l'IA.

Les indices juridiques

Du point de vue juridique et si l'on se réfère à la cartographie présentée ci-dessus, l'ensemble des changements de statuts de la société OpenAI peuvent nous interpellier sur les intentions de la société, mais peuvent surtout traduire son niveau de développement ainsi que ses ambitions. En effet, on constate en premier lieu que l'entreprise créatrice de ChatGPT était un centre de recherche valorisant l'Open source et le partage scientifique. Pour autant en 2019 la société choisit d'arborer un nouveau statut, celui de société à but lucratif plafonné. Ce changement de but et donc de volonté implique de fait une évolution de la stratégie d'entreprise d'OpenAI, qui se traduit concrètement par un renouvellement au niveau de sa direction. Les investissements massifs de Microsoft dans l'ancienne société d'Elon Musk (premier investissement de 1 milliard en 2019) soutiennent l'hypothèse d'un abandon de la part des GAFAM – notamment de Microsoft - de la recherche en matière de moteurs de recherche en faveur de l'IA.



Les indices capitalistiques et financiers

Outre le financement direct de Microsoft dans le capital d'OpenAI, la stratégie du géant du numérique repose sur l'anticipation. En effet, une IA à l'image de ChatGPT a besoin de se nourrir de plusieurs centaines de bases de données pour apprendre et être en mesure de formuler des réponses aux questions posées par les utilisateurs. Dans cette logique, Microsoft acquiert en 2016 LinkedIn – dont le fondateur S. Altman est également au cœur du projet OpenAI – puis en 2018 Git hub afin de concrétiser le projet « co-pilot », une IA enrichie de la technologie de ChatGPT, Dall-E et des bases de données de GitHub afin d'apprendre à cette même IA à coder. En prenant possession de ces immenses bases de données que sont Linked in ou GitHub, Microsoft s'assure également que des concurrents comme Google ne puissent utiliser ces mêmes bases pour alimenter leur propre IA.

Ces dernières années, en plus de l'achat de GitHub et LinkedIn, Microsoft a notamment investi dans des entreprises de jeux vidéo et productions multimédias comme Activision ou Znimax. Ces investissements sont tout à fait compatibles avec une volonté potentielle de développer de nouvelles applications d'IA dans les champs prometteurs du WEB 3.0 ou encore la génération d'images et de flux vidéo. Dès lors, les différentes évolutions capitalistiques et les prises de participations financières de Microsoft depuis 2019 soutiennent notre hypothèse de mise en œuvre d'une stratégie de développement de long terme de la part de la société de Bill Gates et Paul Allen, à savoir la priorisation de l'intelligence artificielle et l'abandon des efforts de recherche et développement des moteurs traditionnels.

2.2. L'IA, l'éthique et la transparence

À la lumière de nos premiers éléments d'analyse et considérant les stratégies mises en place par certaines entreprises fournisseuses d'IA pour obtenir une forme de monopole, il semble pertinent de s'interroger sur les notions de transparence et d'éthique qui encadrent le développement et l'utilisation de cette technologie.

Si nous avons déjà mentionné les enjeux gravitant autour du développement d'IA plus « éthiques » comme Llama ou Mistral AI, il s'agira ici de compléter cette réflexion en considérant le cas d'école qu'est ChatGPT.

2.2.1. Transparence et explicabilité

Afin d'assurer un développement sûr des IA, les décisions prises par ces mêmes systèmes doivent être transparentes et explicables. Les utilisateurs et les parties prenantes doivent en effet être en mesure de comprendre pourquoi une décision a été prise par l'algorithme. Connaître les modalités de raisonnement de l'IA favorise ainsi la confiance dans les systèmes dans la mesure où l'on comprend la logique qui a poussé l'IA à donner une réponse plutôt qu'une autre. En somme, cela permet une évaluation critique des actions de l'IA, surtout dans les domaines où les conséquences sont importantes, comme le domaine médical. Cependant il est important de noter que dans la cadre de ChatGPT, la logique de transparence ne semble pas appliquée, les modèles GPT-3 et GPT-4 sur lesquels se base l'algorithme conversationnel n'étant pas open source. Cela fait d'autant plus débat qu'à l'origine, la société Open AI était une entreprise à but non lucratif avec une réelle volonté de travailler en adéquation avec la philosophie du logiciel libre.

Dans cette logique d'entrave à la transparence, on peut également s'interroger sur la place de la propriété intellectuelle et industrielle dans le modèle choisi. On peut en effet mettre en exergue une forme de paradoxe avec, d'une part, un besoin de transparence sur l'algorithme afin de renforcer la confiance des utilisateurs dans le produit et d'autre part, la volonté légitime du créateur de protéger sa production et de ne pas la partager entièrement. Pour autant, ce paradoxe n'est qu'illusoire. En effet, toute création en open source peut être protégée par les droits d'auteurs. La licence accorde ainsi des droits assez larges aux utilisateurs, notamment celui de modifier ou de diffuser le code source. On parle alors de licence libre. Dès lors, pour éviter toute confusion, il est important de préciser qu'un logiciel en open source



n'est pas libre de droits. Le logiciel ou algorithme ne tombe pas dans le domaine public et reste soumis aux droits d'auteur. Le non-respect des termes d'une licence open source engage par conséquent la responsabilité de l'utilisateur. En définitive, la protection des innovations - notamment dans le domaine de l'IA – ne constitue pas une entrave à la transparence ni à l'explicabilité des différentes IA. Le choix d'OpenAI de ne pas choisir une licence libre est en ce sens un choix délibéré.

2.2.2. Responsabilité et protection de la vie privée

Bien que nous les ayons abordés de manière indirecte tout au long de ce dossier, il est nécessaire à ce stade de s'appesantir sur les notions centrales de responsabilité et de privacy en lien avec le développement des IA.

Concernant la notion de responsabilité, il est crucial de rappeler que les concepteurs, les développeurs et les utilisateurs d'IA ont la responsabilité de garantir que ces algorithmes sont utilisés de manière éthique et conforme aux normes établies. Les décisions prises par les systèmes d'IA doivent être surveillées et contrôlées pour éviter des résultats préjudiciables. L'exemple de l'IA Tay, également créée par Microsoft (en 2016), illustre magistralement cette notion. Tay est un programme d'IA conversationnel (en somme, un ancêtre de ChatGPT) conçu pour interagir avec les internautes sur Twitter (X de nos jours). En moins de 24h, Tay a été retirée des réseaux après avoir rédigé des commentaires racistes, néo-nazis et conspirationnistes. Si Microsoft s'est « indigné » de ce que les internautes ont pu faire subir à Tay pour en arriver à de telles réponses de sa part, il convient de rappeler que les utilisateurs, à l'instar du concepteur de l'IA, sont responsables des propos tenus par la machine. Considérant l'asymétrie de cette co-responsabilité et l'ignorance totale de la loi empirique de Murphy par Microsoft, la réaction médiatique ainsi que la latence de cette dernière pour le retrait de Tay du réseau social sont, de notre point de vue, à déplorer.

L'utilisation croissante de l'IA dans la collecte, l'analyse et l'utilisation des données soulève des préoccupations liées au respect de la vie privée. Il est primordial de garantir que les données personnelles sont traitées avec soin et que les individus peuvent garder le contrôle sur la manière dont leurs informations sont utilisées. A ce stade, rappelons que Microsoft détient tant la licence propriétaire de ChatGPT que le réseau social LinkedIn. Pour rappel, LinkedIn réunit aujourd'hui plus de 930 millions de membres avec, pour chaque individu possédant un compte sur la plateforme, la diffusion de centaines de milliers d'informations personnelles exploitables par Microsoft. Afin de démontrer dans quelle mesure ces données personnelles sont exploitables, nous pouvons nous référer au site « Digimind⁴¹ » permettant de dresser un portrait démographique de la France simplement en exploitant les données accessibles publiquement sur LinkedIn : sexe, âge, poste, parcours académique et professionnel, localisation, réseaux les plus plébiscités, etc. Si un site indépendant et externe à LinkedIn est en mesure d'analyser ces métadonnées pour en tirer des conclusions pertinentes, nous pouvons imaginer la puissance d'un outil tel que ChatGPT dans l'exploitation de centaines de bases de données lui étant directement reliées.

2.2.3. La singularité technologique

La singularité technologique fait référence à un point hypothétique dans le futur où l'intelligence artificielle pourrait dépasser et même surpasser l'intelligence humaine, entraînant des changements profonds et rapides dans la société⁴². En effet, l'IA étant une technoscience d'arbitrage, elle fait plus qu'aider l'individu : elle est en mesure de trouver la meilleure des possibilités pour lui. Dès lors, tout serait question d'informations factorisables. L'écart entre les capacités de « raisonnement » d'une IA et la subjectivité propre à l'être humain tendrait ainsi à se réduire jusqu'à ce que la seule différence entre l'IA et l'être humain soit le support : pour l'homme son corps et pour l'IA les machines⁴³. La singularité intervient à partir du moment où l'on assiste à un décrochage entre l'humain et la machine à la faveur de cette dernière,

41 <https://blog.digimind.com/fr/tendances/linkedin-chiffres-incontournables-france-et-monde>

42 Vernor Vinge, *Technological Singularity*, 1993

43 LIOGIER Raphaël, « La vie rêvée de l'homme », *La pensée de midi*, 2010/1 (N° 30), p. 18-27. DOI : 10.3917/lpm.030.0018. URL : <https://www.cairn.info/revue-la-pensee-de-midi-2010-1-page-18.htm>



la vitesse d'évolution exosomatique (ce que l'on fabrique en dehors du corps) étant supérieure à la vitesse d'évolution endosomatique (l'évolution physique de l'individu).

Ces intelligences « surhumaines », créées par des humains aux capacités augmentées cybernétiquement ou par d'autres intelligences artificielles moins développées, seraient capables d'améliorer leurs propres capacités plus efficacement que les esprits humains les ayant conçues. Ainsi, une spirale de progrès de plus en plus rapide amènerait à des ruptures technologiques très importantes en une courte période de temps⁴⁴. Technologiquement et avec le développement de l'IA générative, nous sommes à « la veille » d'une nouvelle singularité⁴⁵.

Pour autant, il convient de relativiser cette projection. En dépit des progrès et des innovations technologiques, les systèmes d'IA modernes ont encore des limites importantes. Tout d'abord, ils sont généralement très spécialisés et manquent de compréhension générale et de conscience de soi. En outre, il est important de rappeler la complexité de l'intelligence humaine et le fait que la connaissance de ses potentialités est encore loin d'être complète du point de vue scientifique.

L'intelligence humaine réside notamment dans la faculté d'un individu à s'adapter à l'autre par la modification de son attitude et la capacité à mener une réflexion et un débat contradictoire, ce que l'IA n'est pas encore en mesure de faire. A ce stade, cette dernière est seulement capable de créer des synthèses de l'existant. Ainsi, recréer la complexité et la diversité de la pensée humaine dans un système d'IA est un défi immense qui n'est pas encore relevé. Enfin, d'un point de vue éthique et conceptuel, même si l'IA parvient à égaler l'intelligence humaine dans certains domaines, savoir si elle peut reproduire des aspects tels que la créativité (le phénomène de l'« inédit »), l'empathie et la conscience demeurent une interrogation technique et philosophique à laquelle nous n'avons pas encore de réponse⁴⁶.

2.3. Anticiper les prochaines évolutions, adapter nos usages

La révolution ChatGPT et plus généralement des IA impliquent une prise de conscience de la part des utilisateurs, mais également des professionnels du secteur de l'informatique. Il convient dès lors d'adapter nos usages face aux prochaines évolutions techniques en la matière.

2.3.1. IA et souveraineté française et européenne

ChatGPT en particulier, et l'intelligence artificielle de façon générale, méritent une attention toute particulière dès à présent pour les organisations françaises et européennes. En effet, l'IA se positionne déjà au cœur des combats de normes et de titres de propriété intellectuelle entre les différents pôles numériques globaux⁴⁷, et la position de l'Union européenne à ce sujet est assez ambivalente pour plusieurs raisons. Aujourd'hui, les dépôts de brevet dans le monde en lien avec l'IA sont essentiellement réalisés par les industries du numérique américaines, chinoises et japonaises⁴⁸. Par ailleurs, OpenAI et Microsoft viennent de sortir une version GPT-4 de ChatGPT, et les ingénieurs de Google travaillent aussi à développer une technologie similaire. Sans une mobilisation des acteurs académiques, des filières industrielles et technologiques et des institutions politiques européennes, l'issue de la course à l'IA risque d'être défavorable aux organisations et entreprises de l'UE, qui devront *subir* les choix de protocoles technologiques, les normes et les critères éthiques définis par d'autres.

44 Pour la requalification de l'IA et l'importance de l'individu dans le développement de l'IA voir l'interview de Raphaël Liogier : <https://www.frenchweb.fr/raphael-liogier-lia-est-une-expression-mal-choisie-il-sagit-dune-techno-science-decisionnelle/346357>

45 « Nous sommes à moins de dix ans de la singularité technologique », selon Ben Goertzel, <https://www.futura-sciences.com/tech/actualites/intelligence-artificielle-nous-sommes-moins-dix-ans-singularite-technologique-selon-ben-goertzel-109010/>

46 D'après la conférence « Le numérique et l'IA : entre adaptations et réaffirmation des métiers de la communication et du journalisme » conduite par Alexandre Joux à Sciences Po Aix, le 27 septembre 2023

47 Nous nous référons aussi aux États dominants en matière numérique (États-Unis, Chine, Japon, Corée du Sud, Inde, Royaume-Uni) ainsi qu'à l'écosystème d'industriels privés gravitant autour.

48 Les 5 premiers déposants en la matière sont : IBM, Microsoft, Google-Alphabet, Yahoo, Baidu. Source : INPI



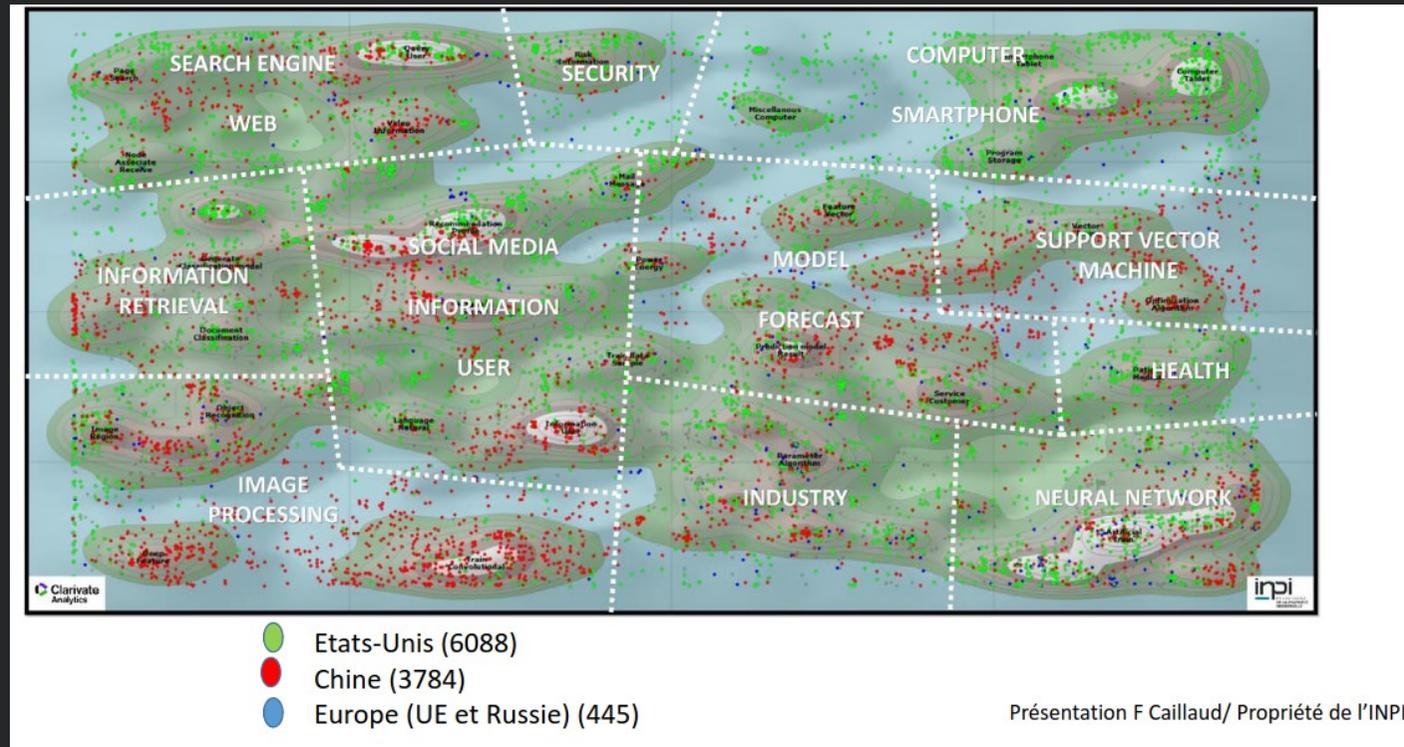
Il faut noter qu'en avril 2021, la Commission européenne a présenté une proposition de règlement concernant le développement de l'IA dans l'Union européenne. Le contrôle absolu étant impossible, la Commission propose un plan de développement tout en encadrant les applications de l'IA⁴⁹. L'objectif de cette démarche réglementaire est de ne pas prendre de retard sur les autres pays dans ce domaine d'avenir en favorisant le développement de l'IA de façon « raisonnée ». Plusieurs propositions ont été établies comme : la création d'un marché unique des données (matière première de l'IA dans la phase d'apprentissage) en conformité avec le RGPD, la mise en œuvre du règlement sur la gouvernance européenne des données, l'investissement dans la formation d'experts dans ce domaine (nous pouvons par exemple mentionner les projets Nevermind et Fabulous), etc. Il faut bien sûr noter qu'il sera toujours plus difficile de développer l'IA dans une région ou un pays qui accorde une importance particulière à la protection des données, ce qui n'est pas le cas aux États-Unis. Ce positionnement singulier adopté par l'UE soulève ainsi plusieurs défis: l'IA étant un progrès technologique que l'on ne peut plus empêcher, nous allons de fait assister à une accoutumance des utilisateurs mondiaux aux outils à base d'IA. Par conséquent on est en droit de s'interroger sur un retard technologique possible pour l'Europe et l'installation d'une situation de dépendance aux technologies chinoises et américaines. Nous faisons de fait face à un paradoxe entre retard technologique en matière d'IA et protection accrue des données de citoyens européens.

La France, tentant de défendre son poulain Mistral dans une course où elle part avec un handicap, tend à temporiser les ambitions de régulation européennes, laissant le temps à ses industriels d'entraîner leur modèle. De facto, nous nous alignons ainsi avec la position des États-Unis, dans une configuration qui sera probablement préjudiciable sur le long terme aussi bien pour l'industrie numérique européenne que pour les citoyens et utilisateurs des suites logicielles intégrant de l'IA.

49 À ce sujet voir : A, Olivier., "L'intelligence artificielle : que fait l'Union européenne ?", Touteleurope.eu, 2021, p. 2. <https://www.touteleurope.eu/economie-et-social/intelligence-artificielle-que-fait-l-union-europeenne/>



Infographie : Les dépôts de brevet liés à l'intelligence artificielle à l'échelle globale ⁵⁰(Source : INPI)



50 Source : INPI - https://www.inpi.fr/sites/default/files/intelligence_artificielle_presentation.pdf



2.3.2. IA et la directive européenne NIS2 (Network and information security 2)

Véritable tournant réglementaire dans le domaine de la sécurité numérique, la nouvelle directive européenne NIS 2 (Network and Information Security 2) fait parler d'elle sur la scène européenne et nationale. Il semble, en conséquence, pertinent de s'interroger sur les enjeux de cette directive pour les entreprises européennes dans le contexte d'une utilisation de plus en plus importante de ChatGPT.

L'arrivée de ChatGPT soulève de nombreuses questions en termes de cybersécurité, mais également en termes d'employabilité de l'outil par les entreprises dans le cadre de la nouvelle directive. En effet, bien que les normes régulatrices de l'UE ne soient pas directement applicables à OpenAI en tant que conceptrice de la technologie, NIS 2 promet un tournant important pour les entreprises françaises qui utilisent cet outil d'IA générative.

Premièrement, si l'outil IA manipule des données sensibles, il doit être conçu et utilisé de manière à garantir la confidentialité de ces dernières (ne pas inclure de données sensibles propres à l'entreprise ou à ses collaborateurs, par exemple). Le second élément à prendre en compte est celui de l'intégrité des systèmes. Les outils d'IA doivent être sécurisés contre les menaces potentielles qui pourraient compromettre l'intégrité des systèmes d'information. En effet, si l'outil n'est pas bien sécurisé, il existe un risque d'exfiltration de données, risque dont la criticité dépend de l'utilisation faite de l'outil et des vulnérabilités existantes. Troisièmement, au sens de la directive, il deviendra nécessaire de notifier les autorités compétentes de tout incident de sécurité impliquant des intelligences artificielles. Enfin, la mise à mal d'un SI, dont la vulnérabilité est introduite par l'utilisation de ChatGPT (ou tout autre outil à base d'IA), peut entraîner des conséquences importantes pour les RSSI, responsables de la sécurité des systèmes d'information. Dans son article 32-6, la directive dispose que les États membres doivent veiller à ce que des personnes physiques puissent être tenues responsables des manquements à leur devoir. Ce risque est d'autant plus sévère si l'enquête révèle que l'entité concernée ne répond pas à l'ensemble des exigences requises par NIS 2, mentionnées précédemment.

2.3.3. Anticiper les prochaines évolutions

Une diversité d'IA spécialisées

Comme évoqué plus haut, la marche des IA pour faciliter le quotidien des utilisateurs est en route. En réponse à l'innovation ChatGPT, de nombreuses IA ont vu le jour afin de concurrencer Microsoft, mais également pour offrir à l'utilisateur une palette d'outils la plus variée possible. Outre Bard et Ernie, développés par Google et Baidu, d'autres IA comme Perplexity.ai, slidesia, copose.ia, formulatbot, etc. proposent des services spécialisés dans la création de présentations, ou dans la structuration de tableaux Excel, l'objectif demeurant de rendre l'IA accessible à tous et de généraliser son usage pour la rendre indispensable pour le quotidien.

Vers une fin programmée des moteurs de recherche ?

ChatGPT inaugure une nouvelle ère en matière de navigation sur Internet. Il sera désormais possible de poser des questions beaucoup plus subtiles à l'intelligence artificielle et, par conséquent, de bénéficier de réponses plus précises. Outre cette révolution dans la manière de répondre aux requêtes des utilisateurs, l'intégration de ChatGPT remet en cause le monopole établi par Google. En outre, nous allons progressivement assister à une accoutumance de la population aux outils intégrant de l'IA. En effet, la manière dont nous avons intégré à nos usages les différentes évolutions techniques dans le monde de l'informatique le confirme. Par exemple,



l'utilisation aujourd'hui classique des moteurs de recherche (Yahoo en 1994 puis Google en 1996) a, en son temps, révolutionné la recherche d'informations, et ces outils nous apparaissent désormais comme basiques, alors même que les premières recherches ne pouvaient s'effectuer qu'en renseignant l'URL complète du site auquel l'utilisateur souhaitait accéder. La technologie ChatGPT va donc continuer de bousculer et de faire évoluer nos usages.

De fait, il semble ainsi pertinent de penser que l'utilisateur préférera utiliser l'IA intégrée plutôt que les moteurs classiques en raison d'une pratique de recherche plus aisée, instinctive et avec un effort moindre. Cela pose toutefois un grand nombre de questions, principalement d'un point de vue éthique, quant à l'élaboration des réponses aux requêtes. Quelle somme faudra-t-il payer pour être référencé par l'IA ? Pourra-t-on poser toutes les questions possibles et inimaginables ? Qu'en sera-t-il en termes de censure ? Quels acteurs seront susceptibles de se positionner comme régulateurs, face aux entreprises fournissant le service ? Autant de questions auxquelles la société civile doit être sensibilisée.

2.3.4. Adapter nos usages : la valorisation du fruit de l'imperfection et de l'originalité

« Je ne parle pas aux IA , ça les instruit »

Le renforcement des capacités de l'IA dans certains domaines doit être mesuré en termes de coûts et de bénéfices pour l'utilisateur. Concentrons-nous sur le métier d'expert en sécurité offensive (pentesteur !), cher à notre cœur. En avril 2023, OpenAI a lancé un programme de Bug Bounty pour son outil ChatGPT. Le bug bounty (ou « prime au bug ») permet aux éditeurs de logiciels et produits de corriger les failles de leurs applications en récompensant les utilisateurs qui trouvent et signalent certaines failles de sécurité. De manière générale, plus la faille est critique et plus la récompense est élevée. Il existe différents types de programmes de bug bounty : programmes ouverts à toute la communauté d'utilisateurs, restreints à certains cercles, internes, etc. Ils peuvent être mis en œuvre au sein de l'entreprise elle-même ou via des plateformes spécialisées dans le cadre d'un contrat de prestation. Dans le cas de ChatGPT, au-delà d'aider à sécuriser le dispositif et de jouer le jeu d'une innovation collaborative, il est primordial de poser la question de l'utilisation des logs⁵¹ générés par les pentesteurs lors de la phase de bug bounty. Il faut en effet envisager que ces logs et les tests subis aient été utilisés a posteriori pour entraîner ChatGPT en test de sécurité, voire en inforensique⁵². Le risque principal demeurant la possible disparition progressive (ou subite!) de certains types de métiers et compétences ainsi que l'incapacité des humains sur le long terme à réintégrer les méthodologies « optimisées » par les IA, devenues les « standards » de marché, le nouvel état de l'art.

Une adaptation nécessaire des métiers

Comme évoqué tout au long de cette analyse, ChatGPT et plus largement le développement exponentiel des IA va conduire à révolutionner notre manière de travailler.

En effet, si nous prenons un peu de hauteur historique et que nous adoptons un point de vue purement économique et philosophique, nous pouvons souligner que les innovations technologiques de rupture appliquées aux méthodes de production constituent un facteur de déséquilibre entre les moyens de production que sont le capital d'une part et le travail d'autre part. À la fin du XIXème siècle, Karl Marx décrivait déjà la façon dont les évolutions techniques et méthodologiques (la division du travail et des tâches, l'automatisation, etc.) contribuaient à un affaiblissement de la classe ouvrière et à l'éclatement de sa conscience de classe au profit de la

51 Événements journalisés, permettant de retracer l'ensemble des tests et actions effectuées sur une cible informatique.

52 Investigation numérique pot-mortem.



bourgeoisie, le travail devenant un moyen de production « non essentiel » car interchangeable. Dans le cas de l'intelligence artificielle, les conséquences économiques et sociales du progrès technique ne relèveront pas tant d'une lutte des classes que d'une confrontation entre domaines réservés aux humains et domaines accessibles aux machines. Ainsi, cette confrontation constitue non pas une logique de lutte de domination verticale mais plutôt une lutte horizontale, entre secteurs d'activité, dont l'issue dépendra de la capacité de ces derniers à s'adapter aux changements induits par l'IA et à cultiver leur particularisme.

De nombreux secteurs d'activité vont être affectés par ces évolutions. Prenons le cas de la justice, du conseil, et de l'éducation.

Alors que Sciences Po Paris a interdit formellement toute utilisation de ChatGPT, il convient de s'interroger sur les impacts de cette technologie au monde de l'enseignement. Utiliser ChatGPT peut-il être considéré comme étant du plagiat ? L'affirmative conduirait inévitablement les créateurs d'algorithmes et logiciels antiplagiat à revoir leur copie, ces outils étant pour le moment dans l'incapacité de détecter efficacement les productions écrites de ChatGPT. Nous faisons de fait face à un paradoxe difficilement soluble : d'une part, la réalité du fait que nous ne pouvons plus nier l'existence de cet outil et donc qu'il devient urgent de se l'approprier pour en faire une force – par exemple en l'incluant dans les programmes pédagogiques et en le présentant comme un outil méthodologique nouveau-. D'autre part, la difficulté de survivance d'un système d'évaluation de l'acquisition des connaissances (et de diplomation) en grande partie basée sur des travaux de productions écrites.

D'après les observations du centre TUMO (Center for Creative Technologies) dans le domaine de l'éducation, l'IA devrait redresser la qualité moyenne de l'éducation et établir de nouvelles normes. Les progrès de l'IA pourraient instaurer de nouvelles pratiques dans les méthodes d'enseignement et d'évaluation, mais également contribuer à l'amélioration des outils d'apprentissage. Le TUMO prévoit entre autres une utilisation de l'IA comme partenaire et copilote dans les projets de demain.

Par ailleurs, il est important de mentionner le risque d'une augmentation des inégalités dans le domaine de l'éducation en lien avec l'accès à l'IA dans le cas où son adoption et sa promotion ne sont pas encadrées. On parle ici d'un risque d'accentuation de la fracture numérique, correspondant au fait que les individus qui n'ont pas accès aux nouvelles technologies et à leurs apports soient marginalisés⁵³. De plus, partant du constat que les individus qui ont accès à cette technologie contribuent à l'évolution et à l'apprentissage des IA, il est primordial de souligner que nous risquons d'assister au développement d'un élitisme culturel via la technologie et à la création d'une réalité déformée reflétant l'opinion, les intérêts et le capital culturel des classes les plus privilégiées.

Il faut garder à l'esprit que nous allons continuer à nourrir les IA, et donc à les instruire grâce à nos données. Il s'agit par conséquent de soutenir directement ou non un modèle éthiquement questionnable où l'accès à une information neutre et démarquée tend à devenir illusoire en l'absence de mécanisme de régulation approprié.

Les professions libérales vont elles aussi être amenées à s'adapter et à évoluer, on pense notamment au métier d'avocat, de comptable ou de notaire. En 2018, le Think Tank Institut Sapiens publiait un rapport sur les impacts de la « révolution digitale » : le « Top 5 des métiers en voie de disparition ». Il est ainsi mentionné que la profession de comptable est appelée à disparaître entre 2041 et 2056. L'innovation ChatGPT en a sûrement accéléré le glas. Il en va de même pour les professionnels du droit. En 2018 et 2019, plusieurs chercheurs mentionnaient déjà la révolution de la justice prédictive et des algorithmes afin d'apporter un éclairage quant aux probabilités d'issues dans un litige. Si l'IA va permettre d'accélérer la production documentaire, elle va aussi permettre d'aider à l'élaboration de la stratégie de défense, anticiper les arguments et contre-argumentations de la partie adverse. Par ailleurs, avec les évolutions constantes en matière de technologie, on peut émettre l'hypothèse d'une IA future en capacité de rédiger des plaidoiries en utilisant la rhétorique, le recours au pathos, etc.

53 Conférence en ligne « AI and Education : scaling the learning curve », TUMO Erevan (Centre de la création numérique), le 23 septembre 2023 <https://tumo.ai/>



Les métiers du conseil enfin, dont la base de travail consiste à brasser, croiser et faire parler des données pour formuler des recommandations vont eux aussi devoir s'adapter. En effet, le deep learning et les innovations induites en la matière par ChatGPT remettent en question certains aspects de cette profession à l'image de l'analyse des données. Par ailleurs, les outils utilisés par ces mêmes experts commencent eux aussi à se nourrir de l'IA et proposent des aides à la décision. Dès lors, il devient urgent de s'adapter face à ces changements et de mettre en valeur l'apport incontestable de la matière grise face aux développements de l'IA. Il faut ainsi repenser ces métiers, mais surtout réfléchir aux capacités intrinsèques de l'être humain en misant notamment davantage sur l'intelligence émotionnelle et relationnelle. Dans une autre mesure, l'IA a également révolutionné les métiers de l'ingénierie. De nouveaux métiers ont d'ailleurs vu le jour comme celui de cognicien, architecte big data, coach de robot ou encore data miner.

À l'opposé du spectre, certains métiers liés à l'information et la communication seront faiblement affectés par l'IA. C'est notamment le cas des journalistes de terrain et d'investigation. Le Directeur du laboratoire Institut Méditerranéen des Sciences de l'Information et de la Communication (IMSIC) Alexandre Joux affirme qu'il n'y aura pas de grands changements pour ce métier : « Seul le journaliste aujourd'hui peut encore mener des interviews de qualité avec la possibilité de s'adapter à son interlocuteur. » Plus globalement, il voit l'IA comme une aide : « L'intelligence artificielle reste un outil d'accompagnement des métiers. Mais rien ne permet d'affirmer aujourd'hui qu'elle remplacera l'humain... et ce dans beaucoup de domaines. »



3. CONCLUSION

Le passage récent de la technologie GPT-3 à GPT-4, la mise en ligne officielle de BARD en mai 2023 dans les pays anglophones, l'annonce du lancement du projet Copilot par Microsoft et l'ouverture de son accès par le grand public en janvier 2024... sont autant d'évènements qui confirment la volonté des géants du numérique de généraliser l'usage de l'IA dans le quotidien des internautes. Synonyme de gain de temps, d'efficacité, mais aussi de « fun » pour un grand nombre d'utilisateurs, la technologie que représente ChatGPT et plus généralement l'IA ne peut toutefois être analysée de manière manichéenne. Les implications sont telles qu'il devient urgent et nécessaire de réfléchir aux défis que pose l'IA dans le secteur de l'entreprise ainsi que dans la société, en vue de maintenir une « longueur d'avance » sur le développement de ces technologies et de prendre des décisions collectives en conscience. Si les risques en matière de souveraineté des données, de protection de la propriété intellectuelle, de sécurité numérique, d'accroissement des inégalités, etc. sont réels, il convient de créer un environnement propice au partage de connaissances, à la réflexion autour de l'utilisation de ces outils, et de positionner des garde-fous. L'objectif étant naturellement de convertir la potentielle menace que représente l'IA en une opportunité au service des individus et groupes concernés.

À cet effet, la dernière infographie proposée synthétise, sous la forme d'un SWOT, les principaux éléments à retenir de notre analyse.



Forces

Amélioration de la qualité des productions humaines (précision, pertinence).
Multiplicité des usages et services possibles.
Facilité d'utilisation et d'application (outils intégrés dans les suites logicielles grand public et dans les moteurs de recherche).
Rapidité d'exécution des demandes.
Capacité de répondre à de nouveaux besoins.



Faiblesses

Base de données jusqu'en 2022.
Consommation importante d'énergies.
Aspiration des données d'entrée, fournies à l'outil.
Pas d'information sur les critères de sécurité.
Imperfection du système (tournures de phrases, absence de sources, biais possibles).
Manque de recul sur les process intégrant une IA.
Profils et compétences encore insuffisants sur le marché de l'emploi.

SWOT



Opportunités

Gains de productivité.
Création de nouveaux métiers.
Émergence de modèles mixtes.
Possibilité de réaffecter les compétences humaines aux tâches à forte valeur ajoutée.
Impartialité dans la prise de décision.



Menaces

Production standardisée.
Atteinte potentielle au secret des affaires, à la propriété intellectuelle et industrielle.
Fuites de données à caractère personnel.
Risques réputationnels et juridiques pour les organisations intégrant les services d'une IA.
Compromission ou détournement de l'outil à des fins criminelles
Modification des équilibres socio-économiques.
Renforcement des inégalités face au numérique (dont inégalités scolaires, puis sociales).

