

Après un an de ChatGPT, l'IA navigue toujours à vue

Il y a un an, OpenAI lâchait son robot conversationnel entre les mains du grand public. Après l'effet « wow », ses répercussions réelles sur le marché du travail restent floues. En revanche, l'impact systémique de l'intelligence artificielle générative semble immense. Sans qu'aucune réponse en termes de régulation n'ait encore été trouvée.

PHILIPPE LALOUX

Le mercredi 30 novembre 2022 restera gravé dans les mémoires de l'humanité comme la date du plus grand coup marketing de l'histoire des technologies. Une sorte de « moment iPhone » puissance iPhone. Ce jour-là, OpenAI, entreprise dont le nom avait à peine franchi la porte des labos de la Silicon Valley, propulsait son ovni entre les mains du grand public. ChatGPT, un agent conversationnel capable de comprendre les mots humains et de « répondre » à ses questions, ou du moins feindre de le faire, donnait chair à un concept ultra-abstrait pour le commun des mortels : l'intelligence artificielle. Pas n'importe laquelle. Pas celle dite « faible », « monotâche », qui met la pâtée à Garry Kasparov, ou reconnaît des images de tumeurs cancéreuses à la vitesse de l'éclair. Non, celle-ci pose les pierres d'une IA dite « générale », ou « forte », potentiellement capable d'apprendre et d'effectuer de multiples tâches cognitives propres à l'être humain : « parler », rédiger, répondre, analyser, résumer, suggérer, inspirer, singer...

L'effet « wow » est total. En cinq jours, ChatGPT atteint déjà le million d'utilisateurs. Du pain béni pour ce robot dont le mode d'apprentissage profond a besoin d'être nourri par des milliards de données pour s'entraîner. Tous les paradigmes semblaient exploser.

Une contribution difficile à quantifier

Alors qu'à la première ère du web, les *big tech*, singulièrement les réseaux sociaux, avaient conçu des algorithmes pour que les hommes se parlent entre eux, cette fois, on apprend à des robots à parler aux hommes. Bienvenue dans l'ère de « l'internet synthétique », comme le décrit Olivier Ertzscheid, chercheur en sciences de l'information.

Panique. Quelques semaines à peine

après la sortie de ChatGPT, Goldman Sachs dégage un rapport affolant, prédisant qu'un quart du travail actuel pourrait être automatisé avec de l'intelligence artificielle. Soit 300 millions d'emplois balayés dans le monde. Un chiffre largement tempéré depuis, notamment par l'Organisation internationale du travail selon laquelle « l'IA générative est plus susceptible d'augmenter que de détruire les emplois en automatisant certaines tâches ». Selon Bloomberg, l'IA devrait peser plus de 1.300 milliards de dollars d'ici 2032. Contribuant à une augmentation de 7 % du PIB. La Banque centrale européenne, quant à elle, estime que l'IA aura surtout un impact sur les salaires, tout en pouvant contribuer à la création d'emplois. Alors que, jusqu'ici, l'informatisation avait surtout affecté les travailleurs peu ou moyennement qualifiés, cette fois, c'est l'inverse : les postes hautement qualifiés seraient clairement plus vulnérables.

Bref, à ce stade, aucun économiste n'est en mesure de quantifier l'effet de l'arrivée de ChatGPT sur le marché du travail. Tout au plus en soupçonne-t-on ses effets systémiques, presque civilisationnels, à la mesure de sa croissance exponentielle. En filigrane se posent toujours des questions cruciales sur la sécurité et la régulation de ce marché. Mais un an après, force est de constater que personne n'a encore trouvé de mode d'emploi. L'IA navigue à vue, dans tous les sens, à la vitesse de l'éclair.

D'un côté, les pragmatiques prudents. Ils n'écartent pas l'éventualité d'une destruction de l'humanité par une IA. Et freinent des quatre fers les élans des apprentis sorciers. De l'autre, les pragmatiques entreprenants qui, au contraire, entendent appuyer à fond sur le champignon pour brider les mêmes éventuels apprentis sorciers. Réguler, puis innover. Ou innover, et puis réguler. On en revient, en somme, aux prémices d'internet où l'option B l'a clairement emporté.

En témoigne le psychodrame récent pour s'emparer de la gouvernance de la fondation OpenAI, reconquis par Sam Altman après en avoir été chassé. Tout en embarquant avec lui l'un des Gafam, Microsoft. S'il n'y a pas de lois suffisamment costaudes pour encadrer toutes les dimensions de l'IA, c'est qu'il n'y a tout simplement pas de législateur. Et quand il n'y a pas de législateur, les entreprises créent leurs lois.

La régulation au point mort

Cela étant dit, pour une fois, même les patrons de la Silicon Valley, en ce compris Sam Altman ou Elon Musk, semblent d'accord : oui, il faut fixer un cadre. Mais lequel ? Fin octobre, lors du premier sommet mondial de l'IA, à Londres, tout le monde s'est cassé les dents sur la question. Prévu pour entrer en vigueur en 2025 (soit « dans deux siècles »), le futur « AI Act » sera censé réguler les usages de l'IA dans l'Union européenne. Le tout sur la base de « niveaux de risques », les IA génératives, de type ChatGPT, étant placées en haut de l'échelle. Sauf que, à la veille de conclure un accord en trilogue Parlement-Conseil-Commission, trois pays, la France, l'Allemagne et l'Italie, demandent de renoncer à ce texte, jugé « contraire à l'approche de neutralité technologique et de prise en compte des risques de la réglementation sur l'intelligence artificielle, censée préserver à la fois l'innovation et la sécurité ». En clair, de ne pas se laisser une chance de réitérer le coup de la Silicon Valley en ouvrant la voie, sans entrave, à de futurs géants européens dans ce domaine.

Aux Etats-Unis, l'Administration Biden a plutôt opté pour la responsabilisation volontaire des entreprises (en les invitant, par exemple, à poser un filigrane sur des contenus générés par l'IA, ce que YouTube fera). Sans surprise, la Chine a choisi le contrôle gouvernemental, en dressant, par exemple, une liste noire de sources qui ne peuvent pas être utilisées

pour former des modèles d'IA. Antonio Guterres, le secrétaire général de l'ONU, a soutenu une proposition visant à créer un organisme de surveillance de l'IA. Le G7, quant à lui, a appelé à l'élaboration et à l'adoption de normes techniques pour que l'IA reste « digne de confiance ».

En attendant, les autorités de protection des données font le job. L'italienne n'a pas fait dans la dentelle en interdisant purement et simplement ChatGPT sur son territoire (le cas s'est, entre-temps, réglé par un accord avec OpenAI). Plusieurs plaintes sont en cours d'examen en France ou en Espagne.

Ce n'est pas seulement une question d'accès à la technologie, mais aussi de transparence des algorithmes, qui amplifient les biais cognitifs humains

Ce qui coince ? L'or noir de ces modèles de langage, à savoir : les données. Ou du moins, leur usage un peu sauvage. Ce qui explique le caractère hasardeux des réponses générées par ces robots, capables de créer des contenus parfaitement faux et d'assurer qu'ils sont corrects sans pouvoir fournir aucune source. Sont en particulier visés l'exercice du droit à l'exactitude des données personnelles fournies par ChatGPT, l'information quant à l'utilisation précise des données personnelles des utilisateurs de l'IA générative, le droit à la rectification des données personnelles erronées que ChatGPT utilise ou génère. En clair, en l'état, compte tenu du Règlement général sur la protection des données, ChatGPT et autres « chatbots » sont insolubles dans le droit européen.

Une nouvelle fracture numérique

Outre l'impact environnemental titanesque de ces machines ou la redéfinition de la géopolitique des composants électroniques nécessaires pour satisfaire leurs exigences en termes de puissance de calcul, un autre enjeu majeur se trame en filigrane : celui de la fracture numérique.

Selon le Baromètre de l'inclusion numérique piloté, en 2022, par des experts de l'UCLouvain et de la VUB, à la demande de la Fondation Roi Baudouin, quasiment un Belge sur deux (46 %) serait aujourd'hui en situation de fragilité numérique. Mais « le ravin qui sépare déjà les utilisateurs quotidiens d'internet des citoyens peu ou pas connectés se transforme en océan, que ces derniers ne pourront jamais traverser », prévoit, dans une tribune au *Monde*, Raphaël Maurel, secrétaire général de l'Observatoire de l'éthique publique en France. Une nouvelle fracture risque de se superposer aux autres, entre ceux qui savent utiliser un robot conversationnel. Et les autres.

Ce n'est pas seulement une question d'accès à la technologie, mais aussi de transparence des algorithmes, qui amplifient les biais cognitifs humains. En 2016, Tay, l'IA conversationnelle de Microsoft, était devenue négationniste après seulement deux mois passés sur Twitter. L'opacité des logiciels gouvernementaux, amenés à être de plus en plus utilisés pour « simplifier » la vie administrative en prenant des décisions impactant la vie des citoyens, pose quant à lui un sérieux défi démocratique. « Les règles juridiques ne sont pas faites pour être codées », écrit Elise Degrave, professeure à la Faculté de droit à l'UNamur dans le dernier *Journal des tribunaux*.

Derrière ces enjeux démocratiques de l'IA se trame donc aussi l'un de ses effets de bord : la défiance à l'égard des institutions et des géants de la tech. Une « déshumanisation » portant les germes d'une bombe sociale à retardement.

PATRIMOINE

Retour du retable d'Averbode à l'abbaye, 150 ans après l'avoir quittée



© PHOTONEWS.

Créé au début du XVI^e siècle à Anvers sur commande de l'abbaye d'Averbode, le retable fut réalisé par le sculpteur et marchand Jacob van Cothem. Le Musée des antiquités d'Anvers, l'ancêtre du MAS, l'a acheté en 1874 et il est resté à Anvers depuis lors et jusqu'à aujourd'hui. Le retable contient trois grands panneaux, deux peints et un en bois sculpté. Le panneau de gauche représente la « Crucifixion du

Christ », la sculpture centrale la « Lamentation du Christ » et la peinture de droite la « Résurrection ». En haut se trouve une scène montrant Jésus, Adam et Eve. L'œuvre d'art a été transférée d'Anvers à Averbode lundi dans une caisse spécialement conçue à cet effet. Depuis 2015, le retable d'Averbode est reconnu comme chef-d'œuvre flamand. BELGA