

« Si les Européens ne développent pas d'IA, ils pourraient avoir du mal à y imprimer leur marque culturelle »

► Lire aussi PAGES 2 À 5, 28 ET 37

PROPOS RECUEILLIS PAR
Eugénie Bollait

LE FIGARO. - Les États-Unis avec OpenAI et maintenant la Chine avec DeepSeek... Le monde s'est-il lancé dans une grande course à l'IA, comme jadis il se lançait dans une course à l'espace? L'Europe est-elle à la traîne et risque-t-elle d'être marginalisée?

MURIELLE POPA-FABRE. - L'avènement de DeepSeek nous montre combien le discours sur l'absence de contraintes pour favoriser l'innovation ne tient pas l'impact avec le réel. La Chine a régulé l'IA et l'IA générative bien avant DeepSeek, et de façon extrêmement capillaire depuis 2022, algorithmes par algorithmes. L'Europe, pour sa part, n'est pas vraiment à la traîne : elle a d'ores et déjà un cadre réglementaire, et surtout des talents qui ont fait leurs preuves l'année passée avec des modèles qui ont remis en question le côté monolithique de l'intelligence artificielle américaine. D'autant que nombre d'Européens travaillent à travers le monde au développement de ces outils. Une politique de rappel de ses talents pourrait avoir un impact positif.

Ce qui manque à l'Europe, ce ne sont donc pas les compétences technologiques, mais la capacité à façonner des produits qui vont générer l'adoption. L'avancée technologique d'OpenAI est intimement associée au travail sur un produit agréable, à qui l'on a envie de poser plus de questions. Le risque de marginalisation, s'il existe, peut concerner la difficulté à façonner culturellement cette technologie : en clair, si nous ne développons pas d'IA, nous pourrions avoir du mal à y imprimer notre marque culturelle et nos valeurs. Concrètement, en demandant à l'IA de générer un air de Noël, la mélodie obtenue pourrait ne pas avoir les notes de nostalgie des chants de l'Est européen. Idem pour les environnements virtuels des jeux vidéo qui ne simulerait pas les rues de Paris ou de Cracovie. DeepSeek, de ce point de vue, est un artefact culturel : nombreux sont ceux qui l'ont testé sur des sujets politiques. En juillet dernier, l'Administration chinoise du cyberspace (CAC) a lancé un processus supplémentaire d'évaluation des modèles de langage en testant le type de réponse que l'on peut avoir sur les sujets déjà encadrés par les standards industriels chinois comme les héros nationaux ou les valeurs socialistes. Avec DeepSeek, nous sommes surtout devant un outil façonné culturellement.

De nombreux chiffres circulent concernant l'impact de l'IA sur le marché du travail, en termes de suppressions et de créations d'emplois. Les entreprises et l'ensemble des secteurs sont-ils prêts à prendre le tournant?

Depuis un an, les calculs concernant l'impact de l'IA sur le monde du travail sont de plus en plus sophistiqués. Dans le dernier rapport du FMI, chaque profession a été analysée sur la base de 56 paramètres correspondant à chaque type de tâche que les personnes ont à effectuer pendant leur travail. Cela permet de savoir où la machine peut « remplacer » et où il peut y avoir « complémentarité ». La question de fond est donc d'avoir la souplesse de jouer la complémentarité avec la machine. Et je parle ici d'une complémentarité exigeante : quelles sont les tâches dans lesquelles nous nous reconnaissons et que nous voulons effectuer nous-mêmes? Ou être aidés pour être plus efficaces? C'est cette recette de complémentarité qui va changer la donne pour les gains de productivité et aider à ce que les personnes ne se transforment pas en « grands contrôleurs ».

Une étude du MIT a par ailleurs analysé pendant deux ans la restructuration des tâches de R&D (recherche et développement, NDLR) dans une entreprise de recherche en matériaux après l'introduction d'un outil d'intelligence artificielle générative. Cette étude a montré combien la partie de créativité et d'idéation a cédé le pas à des tâches de vérification. L'idéation est passée de 39% avant l'introduction de l'outil à 16% après ; la part d'évaluation et de jugement est au contraire passée de 23% à 40%. C'est ce que j'appelle le « syndrome

de l'humain contrôleur ». Pour l'entreprise, près de 17% de produits en plus ont été développés. Beaucoup d'enquêtes insistent sur la démotivation qui peut donc émerger. Ce tournant majeur demande beaucoup d'expérimentations.

Concernant d'autres métiers, l'index de complémentarité du FMI peut être très haut, par exemple pour les chirurgiens, qui, assistés par l'IA et la technologie, exécutent un travail plus précis. La prise de risque et la responsabilité continuent toutefois de s'appliquer à une personne : là encore, il y a un enjeu essentiel. Tout cela fait que le curseur est difficile à placer. Si vous n'assumez pas une partie importante dans un processus, vous allez naturellement vous sentir moins responsable et allez progressivement perdre la capacité de vérification du résultat de la machine, et ce d'autant plus que vous êtes souvent soumis au biais d'automatisation qui porte à valider la suggestion de la machine.

Si je dois résumer, il faut une complémentarité exigeante, une vraie expérimentation de cette complémentarité pour comprendre où sont les gains de productivité tout comme les responsabilités. Le rapport du FMI introduit finalement des variables culturelles, comme, par exemple, l'acceptabilité sociale de l'automatisation du travail. Cette acceptabilité est, par exemple, très basse concernant le travail d'un juge.

L'IA va massivement investir le monde du travail, certes, mais aussi l'école et la vie privée. La société peut-elle se préparer à de tels changements? Comment ne pas y voir une vague déferlante, et ainsi nécessairement menaçante?

Les sondages qui s'intéressent à la peur de l'intelligence artificielle ou à la crainte de ce futur « augmenté » avaient commencé avant même l'arrivée de ChatGPT et ils montraient que le niveau de crainte de l'IA à travers

humaines que nous entretiendrons. Il faut donc ré-humaniser les tâches qui ont été érodées par l'administratif et tenter de continuer à répondre à ce besoin de partage humain, de compagnie, de liens et de temps avec les autres.

Concernant ces « fausses humanités » se pose la question de comment « garder l'humain dans la boucle » et comment répondre à ce besoin de partage humain qui se situe en dehors de la fonction proprement dite ou de l'efficacité du professeur, du coach ou du psychologue. En créant un outil qui répond de manière automatisée à ces rôles, on fonctionnalise de plus en plus l'action de ces professionnels : à cause de cela, tous les « côtés humains qui nous nourrissent habituellement dans ces relations » sont oubliés. C'est comme dans le télétravail : il peut arriver que la « petite information » supplémentaire vous manque. Les entreprises se sont rendu compte que la quantité d'informations qui s'échangent en dehors de la réunion est en fait essentielle. Ce sont le clin d'œil du coach sportif, le moment de respiration du psychologue ou encore le sourire en coin du professeur au moment où l'étudiant dit quelque chose de malin qui donnent de la chair aux liens sociaux. En clair, il y a une dimension essentielle qui est celle de l'intentionnalité du don, qui dépasse la fonction. Penser aux visages des professeurs qui m'ont marqué provoque parfois quelque chose de plus important que le savoir qu'ils m'ont transmis. Il y a quelque chose de lévinassien dans cette question du visage. L'air nous est essentiel, même si ces choses étaient en soi remplaçables par des machines, elles ne nous alimentent pas de la même manière. Le principe d'une IA centrée sur l'humain fait intervenir la machine autour de l'humain et non pas en remplacement.

Les jeunes générations, éduquées par ces plateformes tout au long de leur formation, posséderont-elles les mêmes capacités que leurs aînés? Y a-t-il ici un enjeu majeur?

Je pense qu'il y a un enjeu général qui concerne la perte de capital cognitif qu'implique l'automatisation de nombreuses tâches du quotidien. Dans ces conditions, on peut se poser la question de savoir comment les générations à venir vont trouver un espace de complémentarité, ainsi que la possibilité d'affirmer leur personne. Faire un devoir par soi-même, c'est aussi se rassurer sur ses propres capacités. Actuellement, des expérimentations sur le degré d'assurance des étudiants lorsqu'ils se sont habitués, de façon systématique, à utiliser à bon escient les agents conversationnels sont en cours. Cette phase de transition et d'expérimentation est fondamentale pour comprendre quelle innovation en pédagogie sera bénéfique.

Déléguer une partie de nos activités cognitives à la machine veut dire que l'on ne doit pas assurer le processus de A à Z, tout en devant vérifier le résultat. Ainsi, l'enjeu réside dans le fait de continuer à être capable de vérifier ce qui a été automatisé.

L'une des urgences concernant l'IA touche finalement à la manière dont elle inonde les réseaux sociaux : deepfakes, images de synthèse, faux audios... N'a-t-on pas créé les moyens matériels et gratuits de faire le mal à grande échelle? Comment cohabiter avec cette possibilité?

Je n'avais jamais vu les choses ainsi... Face au fait de ne plus savoir ce qui est vrai ou faux, on peut bien sûr avoir l'impression de ne pas avoir le choix. De même, au fur et à mesure que l'utilisation de systèmes génératifs ou probabilistes – selon les types d'IA dont on parle – se diffuse, on assiste à l'automatisation de larges pans de l'action humaine. Et on comprend bien que, derrière tout cela, se pose la question du libre arbitre et de la possibilité de l'exercer, pour le bien ou pour le mal. Il est donc essentiel d'éviter la perte de capacité d'intervention et en favorisant la supervision humaine sur les systèmes. En répondant à votre question, une citation me vient, de T. S. Eliot : « *They constantly try to escape/ From the darkness outside and within/ By dreaming of systems so perfect that no one will need to be good* » (The Rock). Car le risque est de rêver à des systèmes d'automatisation, de gouvernance ou de contrôle si parfaits que personne n'aurait besoin d'être bon ; des systèmes tellement parfaits qu'il n'y aurait plus besoin de choisir, ni le mal ni le bien. C'est, on peut dire, la grande tentation des idéologies.

Cela me fait revenir à la gouvernance, où l'on cherche à trouver des mécanismes pour observer l'impact de l'IA sur la société en responsabilisant tous les acteurs et en incluant la société civile. Derrière la question de la cohabitation avec la possibilité du mal à grande échelle, il y a la question de fond de la « résilience personnelle ». Cohabiter revient à trouver une manière de pouvoir exprimer sa personne et sa liberté face à l'accélération technologique, à trouver la bonne complémentarité avec la machine au travail, à se ressaisir des fondamentaux, à l'équilibre à avoir dans la recherche de liens sociaux qui donnent du sens. Tous ces éléments, pour moi, se synthétisent autour de cette possibilité de trouver dans la société une façon de gouverner ces outils sans automatiser la liberté. ■

* Ancienne chercheuse au Collège de France et à l'Irria, Murielle Popa-Fabre est actuellement experte au Conseil de l'Europe et intervient sur des questions de réglementation d'intelligence artificielle et de protection des données.

Face au fait de ne plus savoir ce qui est vrai ou faux, on peut bien sûr avoir l'impression de ne pas avoir de leviers d'action

le monde était déjà assez élevé. Les réponses à apporter sont de différentes natures. La première concerne la gouvernance de l'IA – dont un des premiers défis est d'assurer que les systèmes, ainsi que leurs effets, soient observables par des tiers de confiance. Concrètement, renforcer l'observabilité de l'IA implique de pouvoir ausculter les algorithmes, les tester, expérimenter leurs effets sur la société, les individus ou encore la santé mentale et le bien-être au travail pour construire la confiance de la société envers l'IA. La deuxième réponse se joue à l'échelle de l'individu : cette accélération technologique est l'occasion d'un retour aux fondamentaux après les changements de vie provoqués par l'arrivée d'internet, des mails et des réseaux sociaux. C'est une occasion unique de renaissance : pouvoir resituer les fondamentaux dans un nouveau contexte, avec des nouvelles possibilités et des nouveaux risques. Cela implique la sortie de la superficialité qui découle de l'accélération en repensant les priorités, c'est une occasion incroyable d'exercer son sens critique en essayant de comprendre ce qui a de la valeur.

Quels peuvent être ces « fondamentaux »?

La centralité des relations sociales, par exemple. Quelle place donner au fait d'être présent devant un ami? Quelle place donner à l'importance de l'efficacité? Beaucoup parlent d'adopter ces outils pour sortir des tâches redondantes et redonner du sens aux relations. Les professions submergées par les tâches administratives mettent ainsi beaucoup d'espérance derrière cette technologie, pour retrouver le lien : ne pas être en train de taper un rapport pendant que l'on écoute quelqu'un, par exemple pour un policier ou un médecin. Tout cela va dans le sens de cette recherche des fondamentaux : et le lien humain revient très souvent.

Intuitivement, on se dit au contraire que l'IA peut coûter cher aux liens humains. Il existe des professeurs avatés, des psychologues, des coachs sportifs... Ces fausses humanités peuvent-elles prendre la place de nos relations humaines? Comme tout outil, l'IA peut être ambivalente : elle peut être utilisée pour retrouver les liens humains, mais aussi pour les perdre. Le fait d'automatiser le langage, structurellement, à grande échelle, change tout. Auparavant, le langage était forcément lié à une interaction humaine. Aujourd'hui, il ne l'est plus, et cela va avoir un impact sur la quantité de relations

MURIELLE POPA-FABRE

À l'occasion du sommet pour l'action sur l'intelligence artificielle, qui se tient à Paris, l'ancienne chercheuse au Collège de France* décortique les enjeux philosophiques et anthropologiques que représente l'arrivée aussi inéluctable que révolutionnaire de l'IA dans tous les pans de notre vie.

