

IA Act : comprendre le pouvoir, les implications et le raisonnement de l'IA

Internet

Posté par : JerryG

Publié le : 13/12/2023 14:00:00

« Le contrôle et la surveillance du développement et de l'utilisation de l'IA est un sujet brûlant dans le monde entier. L'IA n'est en aucun cas une nouveauté et est en développement depuis des décennies, mais les régulateurs et les gouvernements n'y ont prêté que peu ou pas d'attention jusqu'à l'adoption rapide et généralisée de ChatGPT.

Les membres du Parlement européen ne sont pas les seuls à militer en faveur de ce qu'ils appellent la première législation régissant l'IA. En octobre, le président des États-Unis a publié un décret sur l'IA sûr et fiable.

Plusieurs pays ayant participé au sommet sur la sécurité de l'IA (au Royaume-Uni) ont publié la déclaration de Bletchley par les pays participant au sommet sur la sécurité de l'IA, les 1^{er} et 2 novembre 2023 - GOV.UK (www.gov.uk). Pour la première fois dans l'histoire, nous voyons de nombreux efforts - dans les gouvernements locaux et au niveau international - pour établir des garde-fous autour d'une technologie spécifique si tôt dans son cycle d'adoption.

Je pense que deux problèmes majeurs se cachent derrière la plupart des discussions. Premièrement, nous ne comprenons pas pleinement le pouvoir et les implications de l'IA. Elle évolue trop rapidement et dans une direction que très peu de gens peuvent prédire et prévoir.

Jusqu'à récemment, on s'attendait de manière stéréotypée à ce que les robots prennent en charge les emplois manuels et à ce que les humains s'orientent de plus en plus vers les domaines de la création et de la communication. Et puis, quelques mois seulement après le lancement de ChatGPT, nous avons assisté à des grèves à Hollywood, où l'on craignait que l'IA ne supprime des emplois aux scénaristes et même aux acteurs !

Deuxièmement, nous ne pouvons pas toujours comprendre le raisonnement de l'IA. Nous savons que dans de nombreux cas, elle peut prendre des décisions et trouver des solutions plus rapidement et mieux que des professionnels dans différents domaines.

Mais nous ne pouvons pas vraiment reproduire et comprendre sa logique, il n'y a aucun moyen d'examiner les étapes qui ont conduit un système d'IA à une conclusion. Cela signifie que nous devons faire confiance aux décisions de l'IA pour tirer parti de ses pouvoirs.

Cependant, nous savons également que dans de rares cas, l'IA peut commettre des erreurs tellement évidentes, voire drôles, que l'idée de lui confier des décisions sérieuses est tout simplement effrayante. Cependant, la promesse d'une amélioration de la qualité de vie est trop grande pour arrêter les progrès de cette technologie.

Les propositions de réglementation actuelles, que ce soit dans l'UE ou aux États-Unis, tentent essentiellement de trouver un équilibre entre la garantie de la "sécurité des produits" et la transparence, tout en imposant une utilisation éthique de l'interface utilisateur et en interdisant son application au détriment de l'intérêt public.

Trouver cet équilibre avec tant de craintes et d'inconnues est une tâche presque impossible, surtout si l'on essaie de ne pas entraver le progrès et l'innovation. Les 37 heures de négociations

au Parlement européen la semaine dernière, d'une durée sans précédent, ne sont qu'une excellente illustration des défis auxquels nous continuerons à être confrontés en ce qui concerne l'IA en tant que société.

A ce stade, il est évident que l'IA va fondamentalement changer notre vie quotidienne, mais il est presque impossible de dire ce qu'il va se passer exactement et à quelle vitesse.

Pensez à la façon dont l'accès généralisé à l'Internet mobile et les smartphones ont changé nos habitudes en matière de consommation d'informations, d'achats, d'activités commerciales et de communication avec nos amis et notre famille, le tout en l'espace de quelques années seulement. On peut s'attendre à ce que l'adoption de l'IA soit encore plus rapide et qu'elle ait un impact sur un plus grand nombre d'aspects de la vie humaine. »», dit Ilia Sotnikov, Security Strategist chez Netwrix.