

Google, Microsoft et d'autres acceptent une action volontaire de sécurité de l'IA IT Business

Howard Solomon :



Sept grandes entreprises américaines d'intelligence artificielle, dont Google et Microsoft, ont promis que les nouveaux systèmes d'IA seraient soumis à des tests externes avant d'être rendus publics, et qu'ils étiquèteraient clairement le contenu généré par l'IA, a annoncé vendredi le président américain Joe Biden.

« Ces engagements, que les entreprises mettront en œuvre immédiatement, soulignent trois principes fondamentaux: la sûreté, la sécurité et la confiance », a déclaré M. Biden aux journalistes.

Les entreprises ont l'obligation de s'assurer que leur technologie est sûre avant de la rendre publique, a déclaré Biden.

« Cela signifie tester les capacités de leurs systèmes, évaluer leurs risques potentiels et rendre publics les résultats de ces évaluations;

Les entreprises ont promis de donner la priorité à la sécurité de leurs systèmes en protégeant leurs modèles contre les cybermenaces et en gérant les risques pour la sécurité nationale des États-Unis, ainsi qu'en partageant les meilleures pratiques et les normes de l'industrie;

— les entreprises ont convenu qu'elles avaient le devoir de gagner la confiance de la population et de donner aux utilisateurs les moyens de prendre des décisions éclairées — étiqueter le contenu qui a été modifié ou généré par l'IA, éliminer les préjugés et la discrimination, renforcer la protection de la vie privée et protéger les enfants contre les préjudices;

Les entreprises ont convenu de trouver des moyens pour que l'IA aide à relever les plus grands défis de la société – du cancer au changement climatique – et d'investir dans l'éducation et de nouveaux emplois pour aider les étudiants et les travailleurs à prospérer grâce aux énormes opportunités de l'IA.

Les entreprises qui sont d'accord incluent également Amazon, Meta, OpenAI, Anthropic et Inflection.

Ces engagements volontaires ne sont qu'une première étape dans l'élaboration et l'application d'obligations contraignantes qui seront adoptées par le Congrès.

Réaliser la promesse et minimiser le risque de l'IA nécessitera de nouvelles lois, règles, surveillance et application, [indique un document d'information de la Maison Blanche](#). L'administration continuera à prendre des mesures exécutives et à poursuivre une législation bipartite pour aider l'Amérique à montrer la voie en matière d'innovation et de protection responsables.

« Alors que nous faisons progresser ce programme au pays, nous travaillerons avec nos alliés et partenaires sur un code de conduite international solide pour régir le développement et l'utilisation de l'IA dans le monde entier », ajoute le communiqué.

L'accord indique que les entreprises qui prennent cet engagement reconnaissent que les systèmes d'IA peuvent continuer à présenter des faiblesses et des vulnérabilités même après une solide équipe rouge.

Ils s'engagent à établir des systèmes de primes, des concours ou des prix pour inciter à la divulgation responsable des faiblesses, telles que les comportements dangereux, pour les systèmes relevant du champ d'application, ou à inclure des systèmes d'IA dans leurs programmes de bug bounty existants.

Il y avait un certain scepticisme après l'annonce. [PBS cité](#) James Steyer, fondateur et PDG de l'organisation à but non lucratif Common Sense Media, qui a déclaré: « L'histoire indiquerait que de nombreuses entreprises technologiques ne joignent pas le geste à la parole en s'engageant volontairement à agir de manière responsable et à soutenir des réglementations strictes. »



[Howard Solomon](#)

Actuellement rédactrice indépendante. Ancien rédacteur en chef de ITWorldCanada.com et Informatique Canada.

Journaliste en informatique depuis 1997, Howard a écrit pour plusieurs des publications sœurs d'ITWC, dont ITBusiness.ca.

Avant d'arriver à l'ITWC, il a été journaliste au Calgary Herald et au Brampton (Ontario) Daily Times.

Recherche et mise en page:

Michel Cloutier

CIVBDL

20230725

"C'est ensemble qu'on avance"