

Le comité d'éthique propose de réguler davantage les services tels que ChatGPT que les modèles "open source"

La question de la régulation des grands modèles d'intelligence artificielle (IA) sans entraver l'innovation suscite des tensions au cœur du débat sur l'AI Act, le projet de règlement européen. Certains, dont le Parlement européen, souhaitent imposer des obligations strictes aux modèles de traitement du langage qui sous-tendent des services tels que ChatGPT. D'autres, comme le président de la République Emmanuel Macron et les start-up françaises Mistral AI ou LightOn, s'opposent à des mesures qui freineraient, selon eux, l'essor des entreprises européennes dans ce domaine. Pour tenter de résoudre ce conflit, le Comité national pilote d'éthique du numérique (CNPEN) propose une solution dans un avis consultatif publié le mercredi 5 juillet.

« Nous proposons de faire une distinction entre les modèles commercialisés, soumis aux mêmes obligations strictes que les applications d'intelligence artificielle à haut risque, et les modèles publiés en open source, qui seraient simplement tenus de faire preuve de transparence et de publier des évaluations », explique Raja Chatila, professeur émérite à la Sorbonne et co-rapporteur de l'avis commandé par le ministre délégué au numérique, Jean-Noël Barrot.

Les modèles classés comme « mis sur le marché » incluraient les interfaces destinées au grand public, telles que ChatGPT, le célèbre robot conversationnel lancé en décembre 2022 par l'entreprise américaine OpenAI, ainsi que ses concurrents comme Bard de Google. Ils regrouperaient également les grands modèles de traitement du langage que les entreprises peuvent

interroger à distance via une interface appelée API moyennant quelques centimes par requête, tels que GPT-4 d'OpenAI ou leurs équivalents lancés par Google. Les modèles de génération d'images tels que Dall E2 ou Midjourney seraient également concernés.

En revanche, les modèles classés comme « en libre accès » seraient des modèles similaires publiés en open source, sans interface pour le grand public. C'est le type de diffusion que la start-up Mistral AI envisage d'adopter pour ses futurs logiciels. Ces modèles seraient simplement tenus de faire preuve de transparence, par exemple en ce qui concerne les données utilisées pour leur entraînement, et de publier des tests d'évaluation, par exemple pour détecter les biais connus comme le sexisme. Cependant, dans le cas où une entreprise utiliserait un tel modèle pour un service considéré comme à haut risque selon l'AI Act (par exemple, un diagnostic de santé, l'accès à un service public ou à un emploi), il serait soumis à des obligations plus contraignantes.

Source : www.lemonde.fr