

La Maison-Blanche force Anthropic à couper ses derniers modèles aux étrangers

AUTEUR Victor Storchan

DATE 13 juin 2026

Cette décision historique est intervenue dans la nuit et concerne Fable 5 et Mythos 5.

Pour la première fois, un modèle de frontière est traité comme une technologie stratégique, dont l'accès est strictement limité, au nom de la sécurité nationale, à tout ressortissant étranger, qu'il se trouve ou non sur le sol américain, y compris aux développeurs étrangers qui le conçoivent.

L'administration américaine vient de soumettre deux modèles d'Anthropic, Fable 5 et Mythos 5, à un régime de contrôle des exportations, une décision qui, sous couvert de sécurité nationale, revient à fermer ces modèles à l'étranger.

Anthropic a dû ainsi désactiver, dans la nuit européenne du 13 juin, Fable 5 et Mythos 5 pour l'ensemble de ses clients, afin de se conformer à la directive.

- Le contrôle des exportations interdit tout accès aux deux modèles à n'importe quel ressortissant étranger, qu'il soit ou non sur le sol américain, y compris donc les employés étrangers d'Anthropic qui les développent.
- Dans un communiqué, l'entreprise dit y voir un « malentendu » et affirme travailler à rétablir l'accès au plus vite, en notant que les autres modèles Claude ne sont pas concernés.

As a result of a US government directive, we are suspending access to Claude Fable 5 for all users. You can continue to use all other Claude models.

Here's what this means for you :

Across Claude products, new sessions will run on your selected default model or Opus 4.8, and... <https://t.co/YJTiDotS2v>
– ClaudeDevs
(@ClaudeDevs) June 13, 2026

Dans les faits, un contrôle des exportations équivaut à l'arrêt du modèle.

- Une part importante des effectifs d'Anthropic, de DeepMind et d'OpenAI est composée d'employés étrangers, désormais coupés de l'outil qu'ils conçoivent.
- Si ce précédent était étendu à l'ensemble de l'industrie, il signifierait l'arrêt global du développement des modèles à la frontière.

À ce stade, pour justifier sa décision, le gouvernement a transmis à Anthropic « des éléments oraux attestant qu'elle aurait eu connaissance d'un jailbreak potentiel, limité et non universel ».

- Aucun document ne vient formellement étayer cette mesure. Pour une décision aux effets aussi radicaux que la fermeture de deux modèles, l'absence de motivation écrite prive Anthropic des moyens de contestation habituels et soulève une question relative aux garanties procédurales.
- Un jailbreak « limité et non universel » désigne une faille de contournement partielle qui n'ouvre pas un accès généralisé aux fonctionnalités dangereuses du modèle. Or, ce type de vulnérabilité touche en pratique l'ensemble des modèles de frontière. En faire le fondement d'un contrôle des exportations reviendrait à appliquer à un seul acteur un critère que presque aucun modèle du marché ne satisfait.
- Anthropic rétorque que le niveau de capacité documenté dans le rapport sur lequel s'appuie l'administration est largement atteint par d'autres modèles, comme GPT-5.5, qui ne sont pas visés par la mesure. Le périmètre de la mesure semble donc moins défini par la capacité technique que par le choix de la cible.

Cette décision s'inscrit dans une spirale escalatoire avec l'administration américaine.

- Elle fait suite à la désignation d'Anthropic, en mars, comme « supply chain risk » par l'administration américaine.
- En mai, Anthropic était déjà absente d'un accord signé entre le Pentagone et sept laboratoires d'IA.

Le ton tranche avec celui affiché un an plus tôt par l'administration.

- Au sommet de Paris en 2025, J.D. Vance déclarait : « Nous estimons qu'une réglementation excessive du secteur de l'IA pourrait étouffer une industrie transformative au moment même où elle prend son essor, et nous ferons tout notre possible pour encourager des politiques favorables à la croissance de l'IA » ^①.

Pour la France et pour tout pays souhaitant développer ou accéder à l'IA de frontière, c'est un signal majeur : la souveraineté technologique ne peut plus rester un concept.

- Plus largement, l'épisode tendrait à valider une thèse défendue dans nos pages par Anton Leicht : l'IA ne sera pas l'abondance, mais la rareté.

Dans un essai publié le 11 juin, Dario Amodei, le CEO d'Anthropic, décrivait l'« exponentielle » des capacités de l'IA et réclamait que l'État se dote du pouvoir de bloquer les modèles jugés dangereux.

- Amodei écrivait que les partisans d'une IA plus sûre, dont son entreprise Anthropic, s'étaient attachés à défendre « une législation sur la transparence, des contrôles à l'exportation des puces électroniques et des études sur les effets de l'IA sur l'emploi » – des mesures « certes insuffisantes », mais que « beaucoup pensaient » être « les seules envisageables et applicables ».
- En Argentine, Javier Milei pousse une logique inverse à l'extrême, en promettant une IA « non régulée », qu'il veut laisser se développer « sans la main mortifère d'une réglementation prématurée et mal comprise », et propose d'inscrire dans le droit une catégorie de « sociétés non humaines » pilotées par l'IA.

Si elle était confirmée, cette décision aurait deux effets directs.

- Elle porterait un coup à la rentabilité de modèles d'une valeur de plusieurs milliards de dollars, qui ne seraient alors plus en mesure de générer de revenus en dehors des États-Unis, rendant le développement de frontière non rentable.
- À ce stade, le pourcentage d'employés étrangers écartés du travail chez Anthropic pourrait atteindre deux chiffres, ce qui, étendu à OpenAI, Google ou xAI, donnerait un avantage considérable à Pékin. Comme nous l'expliquait Ng, une série d'erreurs stratégiques et la dynamique de l'écosystème chinois ont déjà considérablement réduit l'avantage des États-Unis sur la Chine dans la course à l'intelligence artificielle.

Cette décision intervient alors qu'Anthropic, à l'instar de son concurrent direct OpenAI, prépare son entrée en bourse.

- S'il pourrait s'agir d'une forme de représailles liées aux tensions entre Dario Amodei et le Pentagone, la baisse de la valorisation de l'entreprise pourrait faciliter une éventuelle prise de participation publique (une idée que Donald Trump avait déjà évoquée par le passé) à moindre coût.
- Plus largement, cette décision montre que le principal risque pour les grandes entreprises lié à l'intelligence artificielle n'est pas industriel ou technique, mais politique et réglementaire, notamment en ce qui concerne l'administration américaine.

SOURCES

- ① Remarks at the AI Action Summit, Paris, The American Presidency Project, 11 février 2025↑