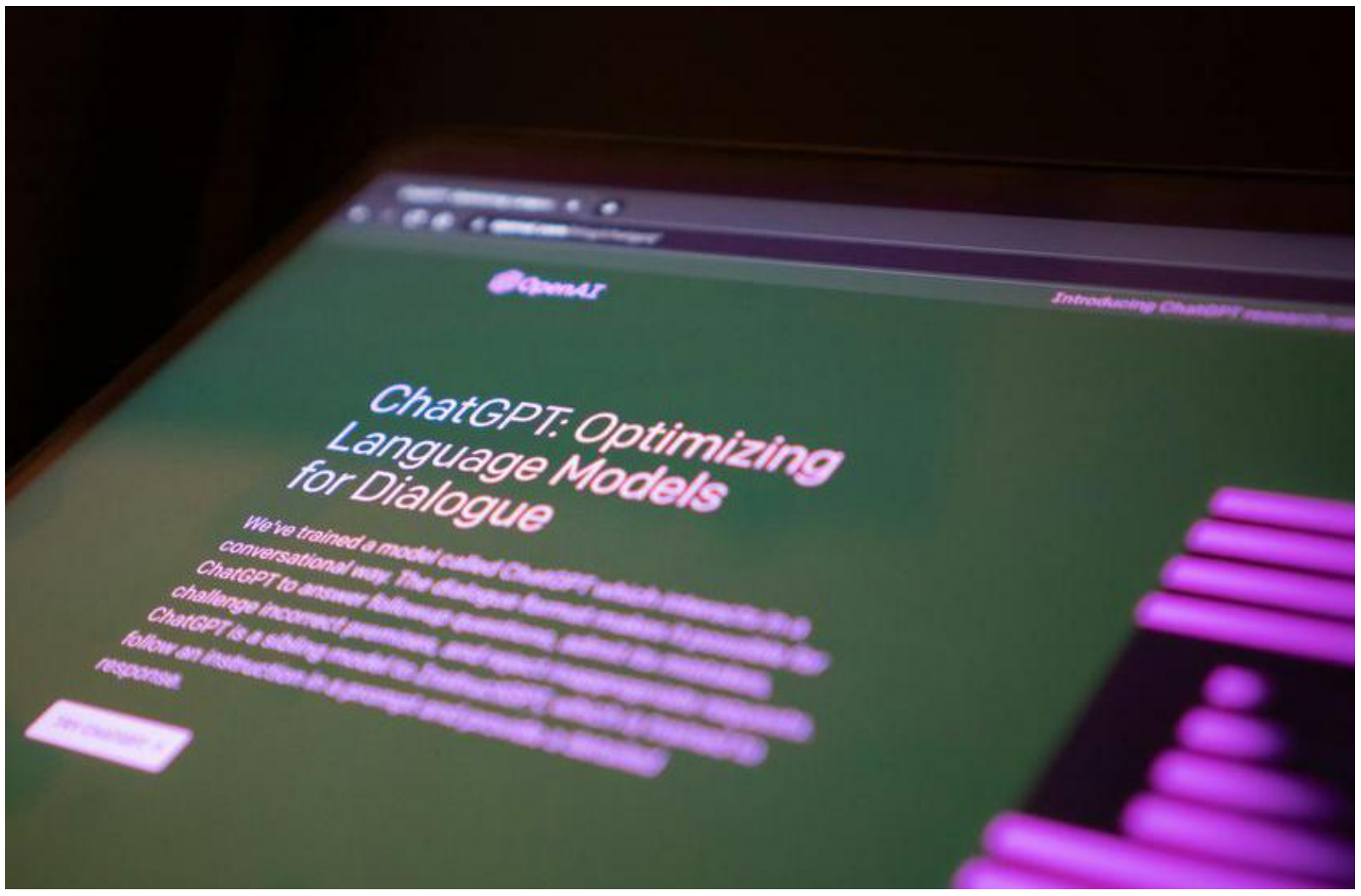
NUMÉRIQUE

Intelligence artificielle : l'Europe change d'ère

Le Parlement européen a adopté l'AI Act, soit la première législation mondiale visant à réguler les usages de l'IA, comme ChatGPT, classé « à haut risque ». Un combat des « pro-innovation » contre les « pro-régulation » qui tourne à l'avantage des seconds. Mais pour mieux grandir, promet la Commission.



DÉCODAGE

PHILIPPE LALOUX

Le projet d'Artificial Intelligence Act vient de passer un cap historique, ce mercredi, à Strasbourg. Les députés européens ont en effet donné leur feu vert, à une très large majorité, à ce qui restera dans l'histoire le texte fondateur d'un cadre légal au développement et la commercialisation de produits mettant en œuvre des technologies d'intelligence artificielle. La version finale, qui ne devrait plus évoluer, pourrait être adoptée par les 27 en avril. Pour une application effective en 2025.

Une fois de plus, l'Europe fait donc figure de pionnière en imposant une quatrième réglementation majeure dans la transition numérique, après le RGPD, le Digital Services Act et le Digital Market Act. Les entreprises vont donc devoir s'acclimater à un nouvel environnement juridique. Le pari de la Commission étant, une fois de plus, de tabler sur une sorte de « Brussels Effect », soit une onde de choc vertueuse sur l'ensemble du secteur, y compris hors de frontières européennes. Tel serait le prix de la confiance à l'égard de technologies, qui effrayent souvent autant qu'elles fascinent. Ce gage de confiance serait aussi celui de l'innovation. Et donc de la prospérité dans un secteur qui pourrait, selon une étude du cabinet McKinsey, ajouter 4.400 milliards de dollars de valeur à l'économie mondiale.

Protéger les pépites

Tel était du moins le point de vue défendu bec et ongles par la Commission et la présidence belge face à des pays plus frileux, comme la France et l'Allemagne qui auront mis tout en œuvre pour pousser sur la pédale de frein de la régulation. Et protéger leurs pépites « locales » (Mistral AI et Aleph Alpha), vues comme de futurs champions européens de l'IA, là où les 27 ont été incapables d'imposer leurs « big tech » lors de la précédente révolution, celle du web et des réseaux sociaux. Pour se faire une petite idée de la vitesse des mutations technologiques, ces deux start-up n'étaient même pas nées au moment où la première mouture de l'AI Act a été présentée, le 21 avril 2021. En trois ans, il a donc fallu intégrer à la va-vite le tsunami de l'intelligence générative (incarnée par ChatGPT), capable d'élaborer, sur la base de requêtes formulées par les utilisateurs, du texte en langage naturel, du code informatique, des images, des vidéos, de la musique…

Au terme de près de trois ans de négociations féroces, auscultées de près par les Etats-Unis (qui, du bout des lèvres, semblent plutôt opter pour l'autorégulation), voici donc l'AI Act, un texte que même Sam Altman (cofondateur d'Open AI), Sundar Pichai (Google) ou Elon Musk estiment nécessaire. Il commence par le début, soit une définition de l'IA, technologie née déjà dans les années 50. Une IA est donc « un système basé sur une machine qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, comment générer des résultats tels que des prédictions, du contenu, des recommandations ou des décisions pouvant influencer les environnements physiques ou virtuels ».

> Le choix de l'Union européenne de défendre une IA basée sur l'« open source » devrait clairement avantager les entreprises du Vieux Continent.

Pour cerner la complexité des enjeux (géopolitiques, industriels, éthiques, juridiques, sociétaux…), la Commission a choisi de réguler davantage les usages que les outils. Cela se traduit par une approche par les risques. Soit une régulation de l'IA en fonction des risques perçus de ses usages, jugés plus ou moins dangereux, de « faibles » à « inacceptables », en passant par « limités » et « élevés ». La Commission fait donc clairement la distinction entre un filtre antispams pour une boîte mail et un agent conversationnel ou un système de notation de crédit pour une banque.

Pas de score social

En creux, cela détermine donc ce qui est et sera toujours formellement banni en Europe. Typiquement, on songe aux systèmes de notation sociale à la chinoise, qui attribuent des « scores » aux citoyens en fonction de leur docilité aux règles édictées. Bannies, également : les techniques qui manipulent nos systèmes cognitifs (les *deepfakes*, par exemple) et donc les comportements individuels, la collecte aléatoire de données de reconnaissance et de vidéosurveillance (les termes *collecte aléatoire* ont évidemment tout leur sens), les systèmes de reconnaissance des émotions sur les lieux de travail et dans les milieux éducatifs, les systèmes exploitant les éventuelles vulnérabilités dues à l'âge ou au handicap physique ou mental d'un groupe de personnes. Ou encore le traitement biométrique pour la déduction de données personnelles sensibles telles que l'orientation sexuelle ou les croyances religieuses. Certaines applications de police prédictive ciblant les individus sont également exclues.

Viennent ensuite les usages classés « à risques élevés ». Ils visent des systèmes ayant une « incidence préjudiciable significative sur la santé, la sécurité et les droits fondamentaux des citoyens ». Cela comprend les « machines médicales » (chirurgie assistée par robot, par exemple), l'éducation (notations d'examens…), les ressources humaines (tri de CV…), le transport (les voitures autonomes…), la justice, les services (comme l'octroi de crédits bancaires)… En février dernier, le Parlement avait surpris la Commission en classant les robots conversationnels, comme ChatGPT ou Gemini (Google), au rayon « systèmes présentant des risques spécifiques de manipulation ». Ils seront, entre autres, tenus d'avertir que le contenu a été généré « par des moyens automatisés ». Les hypertrucages vidéos (« deepfake ») entrent dans cette catégorie. A noter en revanche que les modèles développés par Mistral et Aleph Alpha ne sont pas concernés par cette catégorie. A la différence de ChatGPT (système dit « fermé », à prendre clé sur porte), ceux-ci sont dits « ouverts », laissant la liberté à l'utilisateur de créer son propre environnement sur la base de ses données.

> Une fois de plus, l'Europe fait figure de pionnière en imposant une quatrième réglementation majeure dans la transition numérique.

Le choix de l'Union européenne de défendre une IA basée sur l'« open source » devrait clairement avantager les entreprises du Vieux Continent. L'Observatoire des multinationales ne manque de pointer du doigt le lobbying féroce de ces deux start-up, appuyées par leurs gouvernements respectifs. « Ils ont su trouver les arguments pour convaincre les gouvernements des grands pays européens de réduire les ambitions de l'AI Act », épingle le média en ligne.

Ouvrir le capot

Les fournisseurs de systèmes classés à « risques élevés » seront tenus de livrer une documentation technique aux autorités compétentes, à l'instar d'obligations de mise en conformité en vigueur dans d'autres secteurs (on pense tant aux ascenseurs qu'aux produits de cybersécurité). Cette obligation de transparence est inédite. Pour le dire plus clairement, les acteurs de l'IA « risquée » devront ouvrir le capot de leurs algorithmes. Les entreprises devront aussi livrer un résumé des données ayant été utilisées pour « entraîner » les modèles d'IA génératives. De quoi permettre aux auteurs de savoir si leurs contenus ont été exploités et le cas échéant de demander une compensation. La France est néanmoins parvenue à intégrer la mention « secret des affaires » qui atténue cette contrainte de transparence. Pour certains acteurs du marché, comme Mistral, publier le contenu de leur base de données constitue un préjudice concurrentiel. Considérée comme « bridée » par certains, l'innovation est loin d'être interdite. Les tests devront toutefois s'effectuer dans le cadre de « bacs à sable » réglementaires. Des essais dans la vie réelle sont aussi préconisés, mais les fournisseurs auront l'obligation d'obtenir l'autorisation des organes nationaux au préalable. Un rempart contre les apprentis sorciers.

L'AI Act inclut aussi des obligations d'autoévaluation et d'atténuation des risques systémiques, la notification des incidents graves, la réalisation d'évaluations de tests ainsi que la mise en place de solutions efficaces de cybersécurité. Pas question de jouer avec le feu, donc. Les systèmes « à risques limités », eux, se contenteront d'obligation de transparence, sur le fonctionnement de l'algorithme par exemple, sans autorisation préalable. Tous les objets connectés du quotidien (montres, électroménager, thermostats intelligents…) sont concernés.

Et pour superviser le tout, un Office européen de l'intelligence artificielle vient d'être mis en place en février. Il constituera le centre d'expertise dans l'ensemble de l'UE en matière d'IA. Il lui reviendra de créer des outils, des méthodologies et des critères pour évaluer les capacités des modèles d'IA à usage général comme ChatGPT ou Gemini.

A noter que l'AI Act concerne à la fois ceux qui conçoivent, fournissent ou utilisent un produit ou un service propulsé par de l'intelligence artificielle. Les entreprises, donc, mais aussi le secteur public. Dont les forces de l'ordre.

L'Artificial Intelligence Act pourrait être adopté par les 27 en avril, pour une application effective en 2025. © UNSPLASH.