Offensive chinoise en Europe : l’open source, nouveau cheval de Troie numérique

7.9.2026 - François Godement | Institut Montaigne

Le 11 août, Mistral a annoncé qu’il proposerait un modèle développé par l’entreprise chinoise Z.ai. Objectif : permettre à ses clients de bénéficier d'un rapport qualité-prix attractif sans qu'ils ne partent à la concurrence. Cette décision privée interroge : la logique de concurrence commerciale doit-elle prévaloir sur la sécurité nationale ? Quels sont les risques et où placer le curseur entre opportunité économique et risque de manipulation des données ? Quelle est la stratégie du pouvoir chinois pour mettre la technologie au service de ses objectifs de puissance ?

L’usage des normes, un atout majeur au service de la Chine

Quels sont les liens entre le **destin de l’industrie automobile** et **l’avenir des modèles ouverts** (dits “*open source*”) ou plus précisément, des modèles à poids ouverts (ou “*open weight*”, modèle qui rend ses paramètres d’entraînement accessibles, sans pour autant donner accès à son code, à la différence de l’*open source*) dans le domaine de l’intelligence artificielle ?

Procédons d’abord par une analogie : l**a Chine a parfaitement compris le jeu des normes dans la compétition industrielle.** C’est ainsi qu’elle a fait main basse sur la **quasi-totalité de l’industrie des télécommunications, avec une stratégie d’adoption, puis d’influence**, sur l’élaboration et la diffusion de normes communes dont l’Europe s’était réjouie au départ, notamment avec la 3G. Elle est ensuite devenue une force majeure dans la 5G - aussi bien dans les équipements de réseau que dans son déploiement au sein des chaînes industrielles et logistiques -, ainsi qu’un **acteur majeur de la définition des normes futures.**

Cette expérience se répète dans le domaine de l’IA avec la diffusion de modèles ouverts ou à poids ouverts, désormais "offerts" au monde par Xi Jinping lui-même. **Nombre de pays en développement, de pays émergents et même de pays européens avaient préféré négliger le risque sécuritaire dans les télécommunications afin de bénéficier des bas coûts chinois.** Les modèles de langage chinois à poids ouverts connaissent une adoption croissante, en raison de leur coût apparent très faible.

Ils ont aussi permis à la Chine d’avancer en réutilisant légalement le code, les architectures, les poids et, lorsqu’ils étaient effectivement publics, les jeux de données mis à disposition par des projets internationaux - au nom de la communauté d’échange libre, l’un des grands mythes de l’ère digitale depuis la fondation de l’Internet. À cette (ré)utilisation des ressources ouvertes s’est ajoutée la possibilité de **reproduire certaines capacités et certains comportements observables de modèles propriétaires de pointe** en utilisant leurs sorties comme données d’entraînement ou de comparaison : c’est le processus dit de "distillation", ou plus précisément de **distillation de connaissances.**

Selon les propos d’Elon Musk lui-même, lors du procès qui a opposé au printemps 2026 xAI Corp., l’entreprise à l’origine du modèle xAI, à son concurrent OpenAI , la distillation serait une pratique fort répandue. À une question de la Cour qui lui demandait si xAI avait distillé la technologie d’IA, il a répondu que la plupart du temps, les entreprises d’IA distillaient des modèles d’IA d’autres entreprises et, alors qu’on lui demandait si cette réponse valait pour l'affirmative dans le cas d’xAI, Musk a répondu : "*En partie.*"

Il est vrai que l'entreprise xAI avait elle-même lancé son premier modèle Grok-1 sous licence permissive à poids ouverts. L'entreprise a donc vraisemblablement adopté la stratégie commerciale logique, qui fut aussi celle de la Chine : celle d'un concurrent débutant avec un retard. Le jugement est difficile, tant cette situation peut ressembler au cas des retardataires latecomers dans les révolutions industrielles précédentes. **Mais la logique de concurrence commerciale doit-elle prévaloir sur le risque sécuritaire ?**

Il faut bien sûr préciser ce que recouvre ici l'expression "*open weight*". Dans la plupart des cas évoqués, il s'agit moins de systèmes d'IA intégralement ouverts que de grands modèles de langage (LLM) dont les poids sont publiquement accessibles. Ces modèles à poids ouverts peuvent être téléchargés, exécutés et parfois modifiés sous une licence permissive. Mais **l'ouverture des poids n'implique automatiquement ni la publication des données d'entraînement, ni celle du code complet** de pré- et de post-entraînement, ni l'ouverture des applications dérivées telles les API.

Des restrictions peuvent en outre être appliquées à différentes couches du système : au niveau du modèle et de son post-entraînement, dans les instructions système, au sein de la chaîne logicielle qui l'entoure ou, lorsque le modèle est exploité comme un service, au niveau de l'API contrôlée par son fournisseur. Cela peut aller du refus de fournir des instructions relatives à la fabrication d'une arme biologique aux jugements politiques.

La sécurité : un risque limité, pour l'instant ?

La possibilité de vulnérabilités intentionnelles - portes dérobées, comportements activés par un déclencheur ou mécanismes d'exfiltration - est souvent citée. Ce risque doit toutefois être distingué de celui que présentent les bibliothèques logicielles, les appels à des serveurs distants ou l'infrastructure d'hébergement. Il est un risque moins évident lorsque les modèles sont exécutés localement, ou sur une infrastructure indépendante de celle de leur concepteur, sans appels distants ni transmission de télémétrie ou de données non maîtrisées.

Un autre argument récurrent doit être nuancé. **Les concepteurs des modèles propriétaires seraient les seuls à connaître et à contrôler entièrement leur fonctionnement**, puisqu'ils conservent l'accès aux poids, au code d'entraînement, aux données internes d'évaluation et aux informations relatives au processus de développement. On pourrait contre-argumenter en soulignant que leurs détenteurs doivent être tenus de communiquer ces éléments, notamment aux chercheurs, aux autorités de contrôle ou à des auditeurs indépendants. Et malgré le climat actuel de défiance envers telle ou telle administration - américaine, israélienne, ou, pourquoi pas, française -, il n'y a pas de comparaison possible avec **l'opacité et l'arsenalisation, c'est-à-dire la mise systématique des technologies et des données au service d'objectifs de puissance, de contrôle ou de coercition pratiquée par le pouvoir chinois.**

"Les risques de cybersécurité sont tels dans les domaines numériques que l'Union européenne a déjà fortement découragé le recours aux produits chinois dans certains secteurs."

Les risques de cybersécurité sont tels dans les domaines numériques que l'Union européenne a déjà fortement découragé le recours aux produits chinois dans certains secteurs, tels que les onduleurs de forte puissance. Les États-Unis les interdisent et viennent aussi d'interdire les robots humanoïdes chinois. Certains pays, comme Israël, interdisent les véhicules électriques chinois pour les mêmes raisons sécuritaires, tout comme la Chine interdit la circulation des Tesla – pourtant fabriquées en

Chine – dans certaines régions. Que ce soit par le code, par le matériel ou par l'infrastructure, la stratégie et même la législation chinoises amplifient ces risques.

Les modèles à poids ouverts utilisés exclusivement en environnement exclusivement local diminuent bien certains risques : la dépendance à un fournisseur extérieur, puisque les données ne sortent plus de l'organisation, l'interruption à distance (sauf *backdoor*, notamment dans les logiciels et matériels associés) et l'exposition aux législations extra-territoriales. L'absence de licence ouvre des choix d'utilisation indépendante. Mais les données d'entraînement et les choix qui ont présidé à la construction du modèle restent opaques. Les biais, les erreurs et les comportements imprévus sont donc bien là, tout comme les risques provenant des nombreux logiciels (API) qui entourent généralement le modèle. Le gain de souveraineté suppose de disposer des compétences permettant réellement de contrôler l'outil. À une dépendance extérieure, on substitue des risques internes plus difficiles à identifier et donc à réduire. On remplace une dépendance extérieure difficilement maîtrisable par des risques internes, à commencer par les biais et failles de fonctionnement laissées par le fournisseur initial.

La méfiance reste donc de mise en ce qui concerne les modèles chinois à poids ouverts, les services d'accès par API, les bibliothèques logicielles et les composants matériels susceptibles de leur être commercialement associés. Mais où placer le curseur entre opportunité ou coût économique, d'une part, et risques dûs à l'opacité et l'absence de responsabilité des fournisseurs, y compris pour l'ensemble de la "pile logicielle" (technologies utilisées pour faire fonctionner une application logicielle), les bibliothèques, matériels et infrastructures qui leur seront souvent associés par des utilisateurs publics insuffisamment protégés ou des entreprises privées dépourvues de ces moyens ? **La pente européenne**, au nom de la réduction des risques (*de-risking*) préférée au découplage, **a longtemps été de placer ce curseur au plus près de la sécurité nationale ou de la sécurité publique, entendues au sens le plus restreint.** Qu'en est-il aussi du risque de quasi-monopole obtenu à partir d'une pression réussie sur les prix ?

"La logique sécuritaire commanderait de se priver des ressources logicielles et des infrastructures chinoises qui leur sont associées. L'opportunisme économique, lui, place le curseur dans l'autre sens."

Le dilemme européen : coûts immédiats versus sécurité future

La logique sécuritaire commanderait de se priver des ressources logicielles et des infrastructures chinoises qui leur sont associées. L'opportunisme économique, lui, place le curseur dans l'autre sens, et cela plus encore pour les plateformes numériques ou les entreprises achetant des services d'accès par API. **Le coût est le facteur essentiel de décisions privées qui ne prennent pas en compte les risques publics,** ou qui minimisent le simple risque d'accès non autorisé aux données, de transfert vers une juridiction tierce ou de réutilisation à des fins d'entraînement et de profilage.

Car deux vérités s'imposent. **Ce que Xi Jinping annonce au monde, c'est bien une politique de diffusion subventionnée ou quasi gratuite des modèles de langage chinois** - une forme de "*dumping technologique*" au sens stratégique du terme, et non nécessairement au sens juridique du droit commercial -, et, moins ouvertement, de leurs applications dérivées. À l'abondance du capital-risque américain, la Chine oppose ses ressources financières étatiques et la capacité de l'État à orienter le financement, le crédit et l'épargne vers les secteurs prioritaires.

Deuxième vérité : l**e niveau de sophistication des modèles de langage chinois, s'il a été beaucoup exagéré par l'emballement médiatique** autour du cas DeepSeek, a franchi un cap décisif au sein de la concurrence entre modèles à poids ouverts. Les limites chinoises portent plutôt aujourd'hui sur les ressources de calcul effectivement disponibles, qui dépendent des performances et du nombre des accélérateurs, de leur interconnexion, de l'efficacité de la pile logicielle ainsi que de la capacité des centres de données. Certes, **la Chine est en avance sur l'Europe dans ces différents domaines, mais elle reste derrière les États-Unis.**

Rien n'interdit, en principe, de déployer un modèle à poids ouverts sur une infrastructure indépendante de celle de son concepteur et de **conserver localement les données** qu'il traite, à condition bien sûr de disposer des capacités nécessaires et de maîtriser l'ensemble de la chaîne d'inférence, des bibliothèques utilisées et des communications sortantes. C'est une barre très haute, que très peu d'organisations sont capables d'atteindre - et souvent en ajoutant une dépendance américaine.

En théorie, **l'utilisation de modèles chinois à poids ouverts, couplée à une infrastructure de calcul et à des composants matériels contrôlés de façon autonome, peut constituer une combinaison gagnante, alliant moindre coût et moindre risque.** C'est bien ce que font désormais de grands fournisseurs américains de services cloud, tels que Microsoft Azure et Amazon Web Services (AWS), lorsqu'ils référencent, hébergent ou permettent de déployer des modèles chinois tels que DeepSeek, Moonshot ou Qwen - pour l'instant sans obstacles de la part de l'administration américaine. Cette activité d'hébergement, de mise à disposition et, selon les cas, de distribution repose sur une infrastructure opérée par le fournisseur cloud et distincte de celle du concepteur chinois. Elle répond aussi à une logique concurrentielle, puisqu'elle exerce une pression à la baisse sur les tarifs facturés par les fournisseurs de modèles propriétaires de pointe, tels que les services ChatGPT d'OpenAI ou Claude d'Anthropic.

Un symbole : le tournant Mistral

Venons-en à l'Europe et au champ des possibles pour les Européens. Il existe désormais un test et même un symbole du choix décrit ci-dessus : la décision de Mistral, annoncée le 11 août 2026, de proposer dans son catalogue, parallèlement à ses propres modèles Mixtral, un service d'inférence fondé sur GLM-5.2, modèle à poids ouverts développé par l'entreprise chinoise Z.ai. Ses propres modèles sont en effet fondés sur une architecture à mélange d'experts et conçus pour réduire le coût d'exécution, une approche qui avait d'ailleurs inspiré certains concurrents chinois.

Ce choix fait par la start-up française qui a symbolisé l'autonomie stratégique dans le domaine numérique a de quoi interpeller. Au-delà du gain de légitimité et de la **reconnaissance symbolique** que cette décision apporte aux modèles de langage chinois en France et en Europe, tentons d'en décomposer les implications.

Tout d'abord, c'est bien **une reconnaissance par Mistral du niveau atteint par son concurrent,** avec un rapport qualité-prix sans doute attractif pour les clients de Mistral lui-même. Mieux vaut conserver ces clients au moins pour une part de la valeur (hébergement, mise à disposition, services, etc.) que de les perdre. C'est aussi une convergence potentielle d'intérêts concurrentiels face aux fournisseurs américains de modèles propriétaires de pointe et payants. En théorie, la sécurité peut être assurée par l'usage d'infrastructures distinctes et par des mesures d'isolation logicielle et réseau destinées à empêcher toute transmission de données, de journaux d'usage ou de télémétrie au concepteur chinois.

L'argument cité plus haut concernant l'impossibilité d'accéder aux données d'entraînement des

modèles à poids ouverts a également ses limites : on peut dans une certaine mesure évaluer directement son comportement et modifier ses paramètres par réentraînement partiel. Dans le cas d'un modèle à poids fermés, seul le détenteur de ces poids - ou un prestataire auquel il les confie - peut les modifier directement. Les autres utilisateurs peuvent tout au plus recourir aux mécanismes d'adaptation proposés par le fournisseur.

Il reste qu'en cas d'incident - typiquement les cas récents et remarqués de modèles expérimentaux manifestant, lors de tests, des comportements non conformes aux contraintes ou aux objectifs fixés par leurs concepteurs - l'identification de l'origine du comportement indésirable peut exiger une analyse rétrospective longue et coûteuse des données, des différentes étapes d'entraînement et des consignes du système. Il faut alors retrouver le moment où des instructions insuffisantes ou des objectifs non hiérarchisés, une anomalie dans les données ou tout simplement une erreur de code ou une négligence ont pu engendrer un **désalignement entre le comportement effectif du modèle et les objectifs et consignes qui lui avaient été assignés.** Si l'on est bien obligé de faire confiance aux concepteurs des modèles propriétaires et des modèles de frontière pour effectuer cette recherche, qu'en sera-t-il pour des concepteurs qui n'assument pas de responsabilité durable sur les systèmes diffusés, ou pour des acteurs malveillants ?

Il y a d'ailleurs une raison pour laquelle des dirigeants et chercheurs appartenant aux principales entreprises qui développent des modèles propriétaires appellent à une pause dans leur développement. Cette prise de position va contre leurs intérêts commerciaux à court terme. Elle illustre la difficulté désormais reconnue du contrôle de systèmes d'IA dont les capacités générales et le degré d'automatisation augmentent rapidement, et qui pourraient se rapprocher de ce que l'on désigne comme une intelligence artificielle générale (IAG). Pour les sceptiques, une démonstration illustrée à partir de rapports d'incidents récemment publiés est éloquente…

"La concurrence ouverte est bénéfique au consommateur, mais pour les producteurs, le risque de cannibalisation de leurs propres voitures ou de leurs propres modèles de langage persiste."

Le choix effectué par Mistral présente aussi une **analogie avec ceux de certains constructeurs automobiles**, moins bien placés en Europe même que leurs concurrents locaux : faute du savoir-faire, des capitaux ou de l'échelle nécessaires pour s'adapter à la quatrième révolution industrielle, nous en venons à ouvrir les portes aux firmes chinoises. La concurrence ouverte est bénéfique au consommateur, mais pour les producteurs, le risque de cannibalisation de leurs propres voitures ou de leurs propres modèles de langage persiste.

Une caractéristique des licences ouvertes est que la mise à disposition d'un modèle sous une telle licence n'accorde par elle-même aucune exclusivité à un hébergeur ou à un distributeur. Le concepteur (Z.ai et Alibaba pour Qwen ou Moonshot AI pour Kimi) reste parfaitement libre de distribuer directement ses produits ou de les proposer par l'intermédiaire d'autres plateformes. Sauf à voir Mistral pratiquer avec succès une analyse approfondie du modèle et une ingénierie inverse lui permettant d'en reproduire certaines innovations, démarche qui a tant servi ses concurrents chinois, **l'entreprise peut se trouver progressivement reléguée à ce rôle d'hébergeur et de distributeur non exclusif.**

La double dépendance américaine et chinoise : peut-elle

s’équilibrer ?

La dépendance américaine est forte : dépendance à l’ensemble de la pile de calcul, depuis les accélérateurs et autres semi-conducteurs jusqu’aux serveurs, aux logiciels d’orchestration et aux infrastructures cloud à très grande échelle. À un horizon prévisible, les capacités européennes de calcul consacrées à l’entraînement et à l’inférence des grands modèles devraient continuer de croître moins rapidement qu’aux États-Unis - un écart de capitaux qu’aucune décision publique nationale ou même communautaire ne peut combler en Europe.

"Il faut donc parler d’une double dépendance acceptée : dépendance à un modèle conçu en Chine et dépendance à une infrastructure de calcul largement fondée sur des technologies américaines."

Il faut donc parler d’une double dépendance acceptée : dépendance à un modèle conçu en Chine et dépendance à une infrastructure de calcul largement fondée sur des technologies américaines. Le paradoxe est que chacune peut, dans certaines conditions, limiter une partie des risques créés par l’autre. **L’infrastructure généralement occidentale peut empêcher le concepteur chinois d’accéder aux données traitées, tandis que l’ouverture des poids réduit la dépendance envers les fournisseurs américain de modèles propriétaires.** C’est en ce sens qu’il est absurde de parler, comme c’est devenu la mode en France, de "vassalisation heureuse" à l’égard des États-Unis tout en présentant le recours aux outils chinois comme un facteur de liberté.

Le choix effectué vaut tant que les États-Unis ne durcissent pas leur régime de contrôle des exportations applicable à Z.ai, ou les conditions juridiques applicables à l’utilisation du modèle GLM-5.2, dont les poids publiquement accessibles sur Hugging Face sont proposés sous licence MIT. L’absence chez Mistral de lien commercial direct avéré, de prestation fournie à Z.ai, de participation capitalistique ou de transmission de données à l’entreprise réduit l’exposition juridique et sécuritaire de l’hébergeur - tout comme elle la réduit pour Microsoft Azure et AWS.

Au stade actuel, un opérateur européen ou américain peut exécuter sur une infrastructure américaine ou européenne les poids publiquement disponibles de GLM-5.2, dès lors que la puissance de calcul ou une technologie soumise au contrôle américain des exportations ne sont pas mises à la disposition de Z.ai elle-même. Il est possible que l’utilisation de GLM-5.2 pour des applications militaires déclenche d’autres mesures, y compris pour l’importation aux États-Unis de produits dont le développement ou le fonctionnement intégreraient l’usage du modèle. GLM-5.2 n’est par exemple disponible que dans son service Fireworks AI, qui ne figure pas au catalogue Azure Government fourni aux administrations publiques.

Nous ne savons pas quelle part de la stratégie de Mistral passera par le tournant de GLM 5.2, ni si l’adoption quasi-simultanée par les grands hyperscalers américains accélérera la diffusion des modèles chinois au détriment de leurs homologues européens - ou des modèles langagiers américains fermés. À l’échelle de l’entreprise, **la décision de Mistral a sa logique économique, en particulier pour ne pas verser complètement dans une dépendance américaine de facto.**

Signe-t-elle une étape de construction patiente d’une autonomie logicielle européenne, dont la première condition, après le talent des concepteurs, est la disponibilité de capitaux ? Nous ne le savons pas non plus. Est-ce que ceci accroît à terme les risques européens, tout comme a pu l’être le recours fréquent et persistant de certains états-membres à Huawei dans la 5G ? Avant que le signal du tournant Mistral ne devienne un exemple à suivre et ne fasse boule de neige, il faut s’interroger. **Une Europe de l’IA dont le porteur serait de plus en plus chinois, notamment pour les**

applications aux entreprises, est-elle souhaitable à seule raison de son coût immédiat ?

https://www.institutmontaigne.org/expressions/offensive-chinoise-en-europe-lopen-source-nouveau-cheval-de-troie-numerique