



Avril 2025

La technologie open source à l'ère de l'IA

Alors que de plus en plus d'organisations déploient l'IA générative dans toutes les fonctions de l'entreprise, une nouvelle enquête révèle que les dirigeants se tournent de plus en plus vers des solutions open source pour développer leurs piles technologiques.



Ce rapport fait partie d'une collaboration de recherche parmi McKinsey, la Fondation Mozilla et Patrick J. McGovern Fondation.

En un coup d'oeil

Utilisation et tendances de l'IA open source

- L'utilisation des technologies d'IA open source est répandue. Plus de 50 % des répondants ont déclaré exploiter l'open source dans chacun des domaines de données, de modèles et d'outils de la pile technologique.
- La maturité technique et l'expérience des développeurs influencent l'utilisation de l'open source. L'IA open source L'utilisation est la plus élevée dans les secteurs de la technologie, des médias et des télécommunications (70 %), et les développeurs d'IA expérimentés sont 40 % plus susceptibles d'utiliser l'open source.
- Les organisations utilisent des outils open source d'acteurs connus. En janvier 2025, les outils open source les plus couramment utilisés par les entreprises sont ceux développés par de grands acteurs technologiques, tels que Llama de Meta et Gemma de Google.

Valeur pour les organisations et les développeurs

- La plupart des répondants sont satisfaits de leurs modèles d'IA open source. Environ dix fois plus Les répondants se sont déclarés satisfaits plutôt qu'insatisfaits de leur utilisation des technologies open source. Les principales raisons de satisfaction sont les performances et la facilité d'utilisation.
- L'open source est en tête en termes de coûts, tandis que les outils propriétaires ont un délai de rentabilisation plus rapide. Les répondants affirment que l'IA open source présente des coûts de mise en œuvre et de maintenance inférieurs (60 % des répondants) et (46 %). En revanche, les répondants estiment que les outils propriétaires permettent une rentabilisation plus rapide (48 %).
- Les développeurs apprécient les outils open source. La plupart (81 %) déclarent que l'expérience avec les outils open source est très valorisée dans leur domaine et que travailler avec ces outils est important pour leur satisfaction professionnelle (66 %).

Perspectives d'avenir sur l'IA open source

- L'utilisation de l'open source est susceptible de croître. Les trois quarts des répondants prévoient d'accroître leur utilisation des technologies d'IA open source au cours des prochaines années.
- Les organisations sont ouvertes à un mélange de solutions ouvertes et propriétaires. Près des trois quarts des répondants (plus de 70 %) se disent ouverts aux technologies open source ou propriétaires dans tous les domaines de la pile technologique.

Risques et stratégies d'atténuation

- Les outils d'IA open source présentent des défis potentiels. Les répondants expriment des inquiétudes concernant la cybersécurité (62 %), la conformité réglementaire (54 %) et la propriété intellectuelle (50 %) lors de l'utilisation d'outils d'IA.
- Les organisations mettent en œuvre des mesures de protection pour gérer les risques liés à l'IA open source. Stratégies Cela comprend le renforcement des cadres de sécurité de l'information et des contrôles de la chaîne d'approvisionnement en logiciels, l'évaluation des modèles par des tiers et des garde-fous pour limiter le comportement des modèles.

Les logiciels libres sont depuis longtemps un élément essentiel de l'écosystème technologique. Contrairement à la plupart des logiciels commerciaux, qui nécessitent généralement une licence commerciale ou un abonnement et restreignent l'accès à leur technologie de base, les outils libres sont développés de manière collaborative et mis à la disposition du public pour utilisation, modification et distribution avec beaucoup moins de restrictions, offrant aux développeurs la possibilité d'adapter et de concevoir des solutions sur mesure aux besoins spécifiques de leurs organisations.

L'ère actuelle de l'intelligence artificielle ne fait pas exception. Alors que de plus en plus d'entreprises développent et déploient des solutions basées sur l'IA, elles se tournent vers un éventail croissant de technologies open source, comme la famille Llama de Meta, la famille Gemma de Google, la famille OLMO de l'Allen Institute for Artificial Intelligence et, plus récemment, la famille NeMo de Nvidia, DeepSeek-R1 et Qwen 2.5-Max d'Alibaba. Nombre d'entre elles comblent rapidement l'écart de performance par rapport aux modèles d'IA propriétaires.

Une nouvelle enquête, inédite, menée auprès de plus de 700 responsables technologiques et développeurs seniors dans 41 pays par McKinsey, la Fondation Mozilla et la Fondation Patrick J. McGovern fournit l'analyse la plus complète et la plus détaillée de la manière dont les entreprises envisagent et déploient l'IA open source. Les résultats suggèrent que les dirigeants adoptent les outils open source comme un composant essentiel de leurs infrastructures technologiques, citant des avantages tels que des performances élevées, une facilité d'utilisation et des coûts de mise en œuvre et de maintenance réduits par rapport aux outils propriétaires. Les développeurs, quant à eux, considèrent de plus en plus l'expérience avec l'IA open source comme un élément important de leur satisfaction professionnelle globale. Si les solutions ouvertes suscitent des inquiétudes quant à la sécurité des données et au délai de rentabilisation, plus des trois quarts des personnes interrogées prévoient d'accroître leur utilisation de l'IA open source dans les années à venir.

En d'autres termes, l'open source devrait être un élément clé de la stratégie de solutions technologiques d'IA de chaque entreprise. Alors que l'IA s'infiltré dans presque toutes les fonctions de l'entreprise, dirigeants et experts en technologies doivent exploiter pleinement le potentiel des technologies ouvertes, sous peine de perdre un avantage concurrentiel potentiel.

Vilas S. Dhar, président de la Fondation Patrick J. McGovern :

En démocratisant l'accès aux écosystèmes d'innovation, l'open source met les outils de création à la portée de tous, et pas seulement des plus fortunés. Les utilisateurs deviennent ainsi des bâtisseurs et les consommateurs des créateurs. Nous le constatons clairement aujourd'hui : les défis les plus urgents de l'humanité exigent une intelligence collaborative qui transcende les frontières et les disciplines. L'innovation s'accélère lorsque l'expertise circule librement, lorsque les idées se rencontrent et se combinent sans autorisation. Être visionnaire, c'est reconnaître que notre imagination collective dépassera toujours ce qu'un laboratoire peut produire seul. L'avenir de l'IA appartient aux écosystèmes, et non aux empires.

Comment l'open source est utilisé dans la pile technologique de l'IA

L'enquête a été menée auprès de plus de 700 personnes réparties dans 41 pays, avec une attention particulière portée au Brésil, aux États-Unis, en France, en Inde et au Royaume-Uni. Les répondants étaient classés comme décideurs technologiques ou développeurs seniors et avaient tous une expérience des systèmes d'IA. (Voir l'annexe pour plus de détails sur les répondants.) À l'aide d'une pile technologique d'IA adaptée de précédentes publications de Mozilla¹ (Pièce 1), les répondants ont été interrogés sur leurs expériences avec les technologies d'IA open source et propriétaires dans sept domaines : données ; modèles ; hébergement/inférence ; modifications ; API et gestion des invites ; outils ; et expérience utilisateur/applications. (Voir les définitions et exemples de technologies ouvertes et « partiellement ouvertes » dans le tableau.) Les répondants ont sélectionné parmi une longue liste d'exemples dans l'enquête :

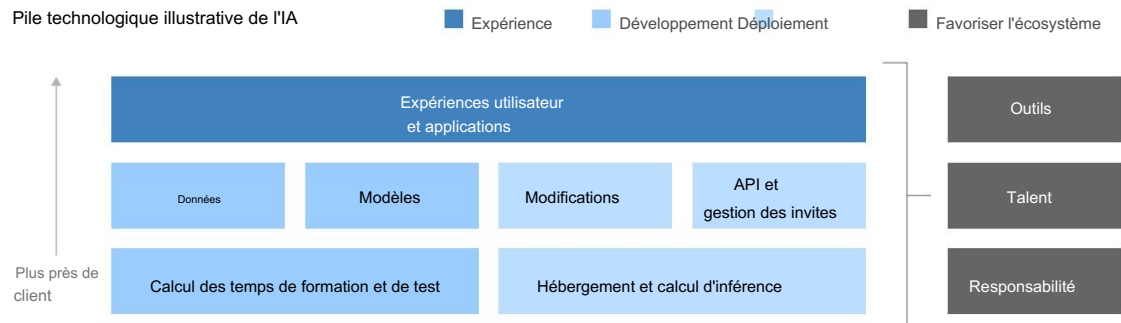
- Données : Il s'agit de données qui peuvent être utilisées pour former ou adapter des modèles d'IA ; elles incluent la préformation, données d'évaluation, de réglage fin et de préférence (telles que Common Voice de Mozilla, Pile d'EleutherAI et Dolma de l'Allen Institute).
- Modèles : Ceux-ci incluent les pondérations du modèle (par exemple, pré-entraînement, point de contrôle, adaptation) ainsi que le code utilisé pour entraîner le modèle (par exemple, Mistral, Gemma, Llama, GPT-J d'EleutherAI, Stable Diffusion et Aya de Cohere).
- Hébergement/calcul d'inférence : cela implique des serveurs ou une infrastructure cloud qui hébergent des modèles d'IA, afin que les utilisateurs puissent exécuter des inférences (comme Llamafire, NomicAI, Ollama, Llama.cpp).
- Modifications : Cela implique des réglages fins, des adaptateurs, des intégrateurs, etc., qui aident à adapter les modèles de base à des cas d'utilisation spécifiques (tels que PEFT [réglage fin efficace des paramètres] et LoRa [adaptation de bas rang]).
- API et gestion des invites : les API et la gestion des invites rendent l'IA utilisable dans des contextes et des domaines spécifiques (par exemple, l'API d'inférence sans serveur Hugging Face).
- Outils : il s'agit d'outils qui améliorent le processus de développement et de déploiement de l'IA, notamment l'orchestration, la sécurité, l'observabilité et l'évaluation de l'IA (tels que PyTorch, Tensorflow, LangChain et Llama Guard).
- Expériences/applications utilisateur : il s'agit d'expériences utilisateur final rendues possibles par l'IA (telles que Chat Câlines).

Web <2025>

Enquête sur l'IA open source>
Pièce <1> de <12>

Une pile technologique d'IA peut être construite avec des composants open source et propriétaires.

Pile technologique illustrative de l'IA



Source : IA publique : faire en sorte que l'IA fonctionne pour tous, par tous, Mozilla, 30 septembre 2024

¹ IA publique : faire en sorte que l'IA fonctionne pour tous, par tous, Mozilla, septembre 2024.

Pour les éléments non modélisés de cette pile technologique (outils logiciels, packages et ensembles de données), les technologies sont considérées comme ouvertes si elles correspondent aux définitions standard des logiciels open source, comme la liberté d'utilisation, d'étude et de modification. Cependant, la complexité des modèles d'IA, et en particulier des grands modèles de langage (LLM), a soulevé des questions quant aux critères des modèles d'IA open source. Cela dit, les définitions et les critères d'« ouverture » en IA font l'objet de nombreux débats. Mozilla soutient la définition de l'IA open source proposée par l'Open Source Initiative (OSI),² qui comprend des exigences d'ouverture sur les informations de données, le code et les paramètres. Les technologies conformes à la norme OSI pour « open source » ou « IA open source » ont été considérées comme open source dans l'enquête. Les outils populaires comportant des composants ouverts mais ne respectant pas la norme OSI (par exemple, les modèles de pondération ouverts ou les logiciels dont les licences open source ne sont pas approuvées par l'OSI) ont été considérés comme partiellement ouverts dans l'enquête. Dans les graphiques où cette distinction n'est pas faite, nous incluons les outils partiellement ouverts et open source dans notre catégorie open source (tableau).

Comment les organisations intègrent l'open source dans leurs piles technologiques

Notre enquête révèle que l'utilisation de l'IA open source est courante chez les répondants, en particulier dans les domaines de la pile technologique où elle peut être intégrée de manière transparente aux systèmes d'entreprise et aux protocoles de sécurité existants. Plus de 50 % des répondants déclarent utiliser l'IA open source dans les domaines des données, des modèles et des outils de la pile technologique (Figure 2). L'open source est moins répandu dans les modifications (telles que les réglages fins et les adaptateurs) et l'hébergement/le calcul d'inférence ; cela peut s'expliquer par le fait que les projets d'inférence open source sont encore relativement récents (par exemple, l'un des nouveaux outils les plus remarquables, vLLM, a été développé au Sky Computing Lab de Berkeley en avril 2024³). Cela peut également indiquer que les utilisateurs choisissent de modifier leurs modèles avec des packages internes et des données propriétaires pour des cas d'utilisation personnalisés.

Tableau

Les modèles d'IA se situent sur un spectre allant de totalement fermé à complètement ouvert.

	Description	Exemple
Fermé	Niveau 1 : Interne/étroitement contrôlé ; accès externe limité ou inexistant	Gemini Ultra (initialement)
	Niveau 2 : Accès externe (par exemple, via l'API) ; absence de poids complets et d'informations sur la formation (les poids ne sont pas open source)	GPT-4 (API), Claude
Partiellement ouvert	Niveau 3 : Modèles de pondération ouverts ; code et données limités ou non accessibles au public	Lama 3 et 4
	Niveau 4 : Modèles de pondération ouverts avec matériel d'appui (par exemple, code) ; raisons pour lesquelles la définition de l'IA open source établie par l'Open Source n'est pas respectée L'initiative (OSI) pourrait inclure des licences non approuvées par l'OSI ou des données insuffisantes	Commande R+ (Cohere), Diffusion stable 2
Ouvrir	Niveau 5 : Répond à la définition OSI de l'IA open source (par exemple, disposer d'un code source, d'aucune restriction de redistribution et d'une licence autorisant les modifications et neutre sur le plan technologique)	Bloom (BigScience)
	Niveau 6 : dépasse la norme OSI ; les données sont accessibles au public	Pythia (EleutherAI), T5 (Google), Dolly 2.0 (Databricks)

² Blog Mozilla, « Célébration d'une étape importante pour l'IA open source », article de blog d'Ayah Bdeir, Imo Udom et Nik Marda, 22 août 2024.

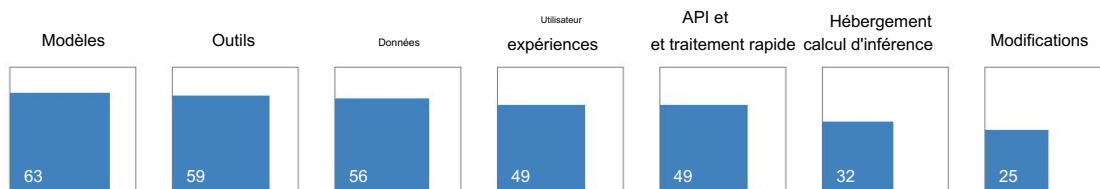
³ Ivan Ortega, « Un moteur d'inférence et de service à haut débit et à faible consommation de mémoire pour les LLM », UC Berkeley Sky Computing Lab, 25 avril 2024.

Web <2025>

Enquête sur l'IA open source
Pièce <2> de <12>

L'utilisation de solutions d'IA open source est importante dans l'ensemble de la pile technologique, la plus grande part des répondants la signalant dans les modèles et les outils.

Utilisation régulière de solutions d'IA open source par les organisations, par domaine de la pile technologique, % des répondants



Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

Bien qu'il existe de nombreuses technologies d'IA open source disponibles dans l'ensemble de la pile technologique, les organisations utilisent le plus souvent celles créées par de grandes entreprises telles que Google, Meta et Microsoft. Dans les modèles, 61 % des répondants ayant déclaré utiliser des modèles open source avaient utilisé Llama de Meta, 40 % Gemma de Google, 32 % Mistral et 28 % Phi de Microsoft. Parmi les répondants utilisant des outils d'IA open source, les plus populaires étaient PyTorch (58 %), TensorFlow (57 %), PostgreSQL (45 %) et LangChain (33 %). Avec la transition vers le développement d'agents, les bibliothèques fournissant des frameworks et des bibliothèques d'agents open source, ainsi que des services d'évaluation, pourraient gagner en popularité.

Les secteurs des technologies, des médias, des télécommunications et des industries de pointe sont en tête en matière d'utilisation de l'IA open source. Parmi les secteurs comptant un grand nombre de répondants (plus de 70 %), le secteur technologique est en tête, avec 72 % des répondants utilisant des modèles open source et 39 % d'entre eux utilisant l'open source depuis plus de trois ans. Géographiquement, l'Inde, le Royaume-Uni et les États-Unis sont en tête en matière d'utilisation de modèles open source, ce qui reflète probablement la maturité relative de leurs secteurs technologiques (graphique 3).

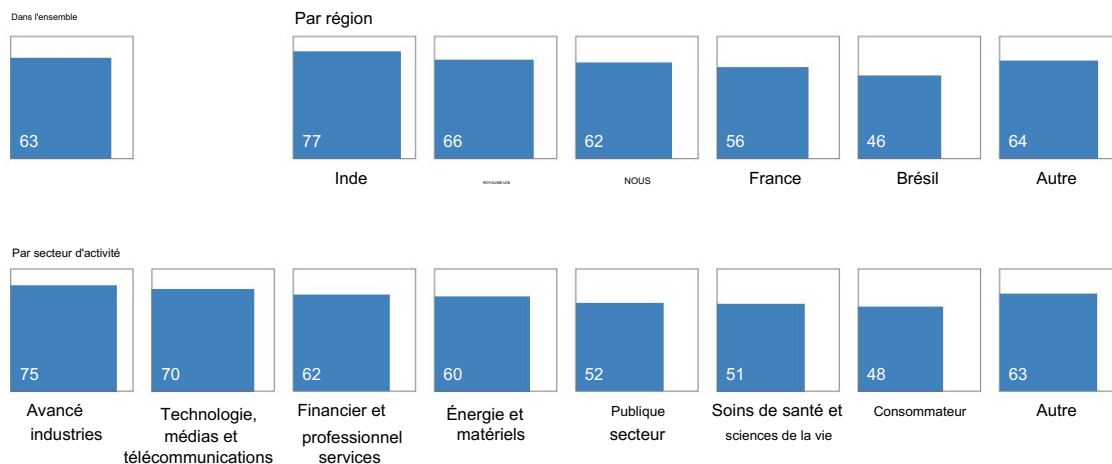
Quel que soit le secteur d'activité ou le lieu, les répondants qui qualifient l'IA d'importante ou de très importante pour l'avantage concurrentiel de leur organisation (81 % des organisations) étaient plus de 40 % plus susceptibles de déclarer utiliser des modèles, des outils et des données d'IA open source, ce qui suggère que les organisations ayant une plus grande maturité en matière d'IA et des développeurs plus expérimentés sont plus susceptibles d'utiliser des technologies d'IA open source.

Mark Surman, président de la Fondation Mozilla : « L'élan en faveur de l'IA open source est indéniable. Rien que l'année dernière, nous avons vu d'innombrables exemples prouvant que l'innovation communautaire peut non seulement concurrencer, mais même surpasser les modèles propriétaires. Le prochain grand pari consiste à créer des outils et une pile ouverts qui rendent l'IA véritablement accessible, comme une boîte de Lego IA utilisable par tous. Si nous y parvenons, l'IA open source ne sera pas seulement une alternative aux systèmes fermés. Elle constituera le fondement d'un avenir plus compétitif, créatif et innovant. »

Web <2025>
Enquête sur l'IA open source>
Pièce <3> de <12>

L'Inde et les secteurs technologiques sont des points chauds pour l'utilisation régulière de modèles d'IA open source.

Utilisation régulière de modèles d'IA open source par les organisations, par région et par secteur,¹ % des répondants



¹ Industries de pointe (y compris les semi-conducteurs), n = 52 ; technologie, médias et télécommunications, n = 260 ; services financiers et professionnels, n = 157 ; énergie et matériaux, n = 43 ; secteur public, n = 59 ; soins de santé et sciences de la vie, n = 53 ; consommation, n = 45 ; autres, n = 34.
Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

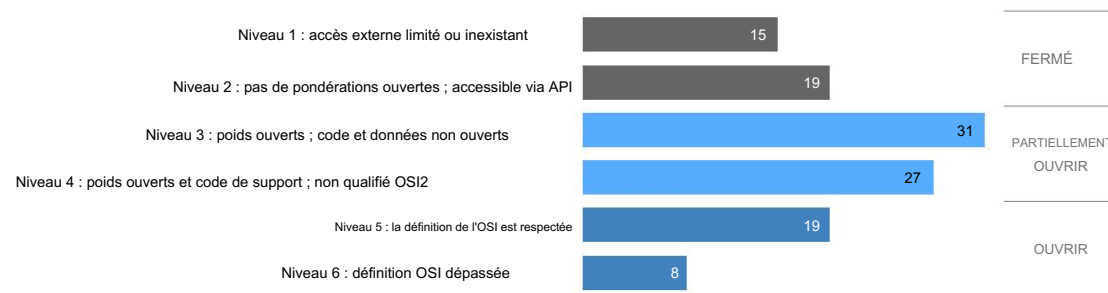
Pour les modèles d'IA, les répondants ont été interrogés sur leurs préférences sur l'échelle d'ouverture. Les répondants ont le plus souvent sélectionné des modèles partiellement ouverts, c'est-à-dire des modèles avec des pondérations ouvertes, mais qui ne respectent pas nécessairement la norme OSI d'ouverture aux données publiques (graphique 4). Ce résultat reflète probablement le contexte concurrentiel actuel, où les modèles partiellement ouverts incluent bon nombre des modèles les plus connus et les mieux dotés en ressources, tels que les modèles à pondérations ouvertes des familles Llama 3 et 4, et les modèles soumis à des restrictions d'utilisation commerciale (comme Stable Diffusion).

Web <2025>

Pièce 4 de 12 sur l'IA open source>
Pièce <4> de <12>

La plupart des organisations préfèrent utiliser des modèles d'IA classés comme partiellement ouverts.

Intérêt pour l'utilisation de modèles d'IA en production, par niveau d'ouverture d'accès, % des répondants¹



¹Les répondants pouvaient sélectionner >1 réponse.

²Initiative Open Source.

Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

L'attrait des modèles de pondération ouverts pourrait également s'expliquer par leur capacité à être hébergés sur l'infrastructure interne de l'organisation, une préférence citée par 40 % des dirigeants d'entreprise. Les dirigeants interrogés ont indiqué que les LLM auto-hébergés offrent un meilleur contrôle des données.

confidentialité, réduisant le risque de violations de données impliquant des fournisseurs tiers et permettant la mise en œuvre de protocoles de sécurité personnalisés - une approche qui minimise l'accès externe aux données sensibles et donne un contrôle total sur la gestion des données et les protocoles de sécurité.

Pour les modèles hébergés sur l'infrastructure de l'entreprise, 43 % des répondants estiment que les offres open source sont plus faciles d'accès et d'exécution (Figure 5). Cependant, lorsque les modèles sont hébergés par un hébergeur ou utilisés via une API (comme l'API Claude, l'API Gemini ou l'API OpenAI), davantage de répondants déclarent que les modèles fermés sont plus faciles d'accès et d'exécution (respectivement 45 % et 39 %), ce qui reflète probablement la disponibilité et l'utilisation répandue des API populaires des principales entreprises d'IA. Une distinction peut être établie entre les organisations axées sur l'IA qui disposent des compétences et des ressources nécessaires pour gérer leurs propres déploiements de modèles et celles qui souhaitent utiliser une API externe pour bénéficier d'une suite de développement et d'une rentabilisation plus rapide. L'adoption de l'une de ces deux approches par une organisation est souvent déterminée par plusieurs facteurs, tels que la profondeur des talents technologiques, les cas d'utilisation prioritaires et les préoccupations en matière de sécurité.

Lareina Yee, directrice et associée principale du McKinsey Global Institute : « De nombreuses organisations sont intéressées par une approche multimodèle (mélangeant des modèles plus ouverts avec des solutions propriétaires).

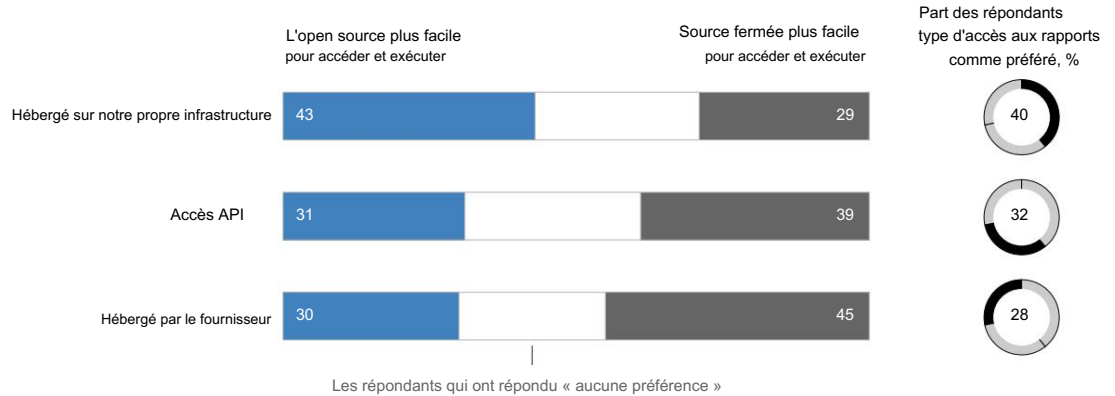
Nous constatons une évolution similaire à celle de nombreuses grandes organisations qui utilisent des solutions provenant de plusieurs fournisseurs de services cloud. »

Web <2025>

Enquête sur l'IA open source
Pièce <3> de <12>

Avec une infrastructure interne, les organisations trouvent plus facile d'exécuter des modèles d'IA open source que d'exécuter ceux qui sont fermés.

Facilité d'utilisation perçue et préférence des modèles d'IA, par type d'accès, % des répondants



Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

Les organisations tirent une valeur significative des technologies open source

De nombreuses organisations utilisant l'IA open source en constatent déjà la valeur, et la plupart des répondants utilisant à la fois l'IA open source et l'IA propriétaire se disent satisfaits de leurs outils d'IA.

Les principaux facteurs de satisfaction sont la performance et la facilité d'utilisation. Nous avons constaté que les personnes satisfaites de leur utilisation des technologies open source sont bien plus nombreuses que les personnes insatisfaites : plus de dix fois plus de répondants se sont déclarés « assez

satisfaits », « satisfaits » ou « très satisfaits » de l'open source que de répondants ayant choisi l'une des options « insatisfait

Les utilisateurs de technologies d'IA open source signalent des coûts de mise en œuvre inférieurs (60 %

des décideurs) et des coûts de maintenance inférieurs (46 % des décideurs) par rapport aux outils propriétaires.

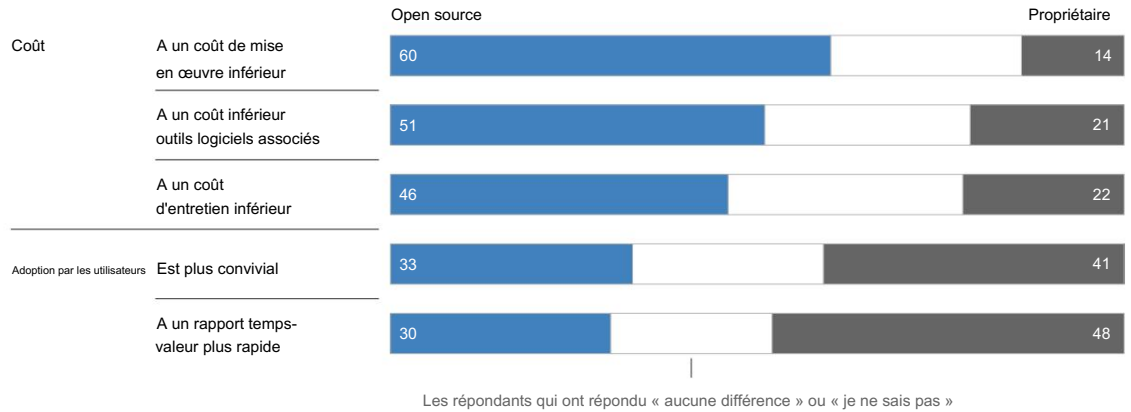
Cinquante et un pour cent des répondants estiment que l'utilisation de technologies open source réduit les coûts d'utilisation des outils logiciels associés dans leurs organisations (graphique 6).

Web <2025>

Enquête sur l'IA open source
Pièce <6> de <12>

Les leaders technologiques affirment que les modèles d'IA open source sont moins coûteux à déployer que les modèles d'IA propriétaires, mais qu'ils sont en retard en termes de délai de rentabilisation.

Coût et adoption par les utilisateurs des modèles d'IA open source par rapport aux modèles propriétaires, % des répondants



Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

L'utilisation de l'IA open source est plus courante dans les grandes organisations et les industries qui investissent déjà dans l'IA, car elles sont plus susceptibles de disposer des ressources nécessaires pour investir dans les talents et le soutien interne. De même, les fournisseurs de modèles d'IA propriétaires ont investi massivement dans la simplicité d'utilisation et les outils de développement, afin d'offrir une expérience fluide. Cela est particulièrement évident pour les produits destinés au grand public, dont l'objectif principal est d'élargir l'adoption en proposant des API bien documentées, des extraits de code prêts à l'emploi et une interface de développement intuitive.

Les utilisateurs d'outils open source et propriétaires font état de niveaux similaires de création de valeur (en termes d'augmentation des revenus ou d'économies) dans toutes les fonctions de l'entreprise. Les fonctions qui déclarent le plus souvent utiliser l'IA open source sont l'informatique, l'ingénierie logicielle et la gestion des produits.

Développement de services, identique à celui de l'IA propriétaire. Cependant, l'IA open source a affiché un léger avantage en termes de réduction des coûts, avec une moyenne de 4 % supérieure à celle des solutions propriétaires, avec une amélioration typique des coûts de 26 %. Le marketing a réalisé les plus importantes économies grâce à l'IA open source, surpassant l'IA propriétaire de 5 %. Les opérations de services ont enregistré les plus fortes augmentations de revenus grâce à l'IA open source, avec une croissance de 6 % supérieure à celle des solutions propriétaires.

Les développeurs déclarent être particulièrement intéressés par les outils open source, même si leur utilisation reflète souvent le niveau d'expérience du développeur, puisque les outils propriétaires ont tendance à nécessiter moins d'expertise.

L'utilisation régulière de modèles d'IA open source était plus courante chez les développeurs qui avaient plus d'expérience de travail avec des systèmes d'IA (plus de 40 % de chances en plus) que chez ceux qui n'en avaient pas (graphique 7).

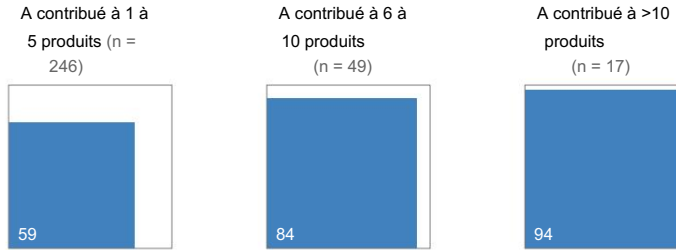
Interrogés sur leur familiarité et leur aisance avec les outils open source, 81 % des développeurs déclarent que l'expérience avec ces outils est très valorisée dans leur domaine, et 42 % d'entre eux affirment que les outils open source sont plus valorisés que les outils propriétaires. De plus, une majorité de développeurs (66 %) déclarent que l'expérience avec les outils open source est importante, voire très importante, pour leur satisfaction professionnelle (Figure 8).

Web <2025>

Enquête sur l'IA open source
Pièce <7> de <12>

Plus le développeur est expérimenté, plus il est susceptible d'utiliser régulièrement des modèles d'IA open source.

Utilisation régulière de modèles d'IA open source, par niveau d'expérience de travail avec des systèmes d'IA, % de développeurs répondants (n = 312)



Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

Web <2025>

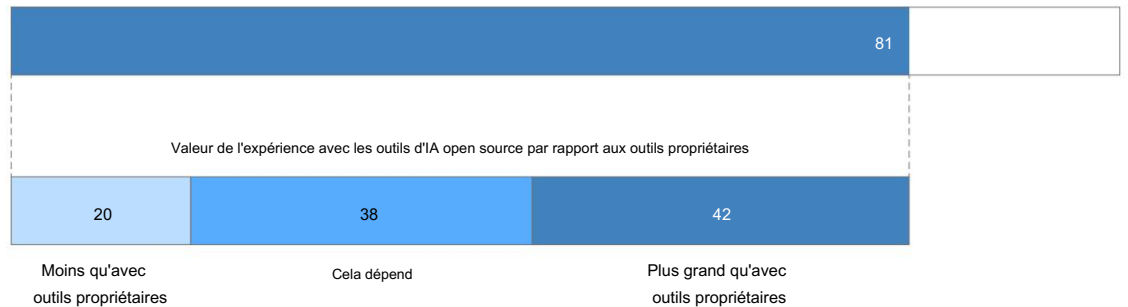
Enquête sur l'IA open source
Pièce <8> de <12>

Les développeurs affirment que travailler avec des outils d'IA open source est très apprécié et important pour la satisfaction au travail.

Importance des outils d'IA open source pour la satisfaction au travail, % des développeurs interrogés (n = 312)



Expérience avec des outils d'IA open source perçue comme valorisée sur le terrain, % de développeurs répondants (n = 312)



Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

Harrison Chase, cofondateur et PDG de LangChain : « Les outils open source créent un écosystème d'intégrations qui accélère l'innovation grâce aux contributions partagées, simplifie la prise en main d'un large éventail d'outils et assure l'interopérabilité avec les magasins de vecteurs et les fournisseurs de modèles. L'adoption croissante des frameworks open source par les développeurs s'accompagne naturellement d'une augmentation de l'utilisation d'outils commerciaux. Cette tendance se poursuit, les outils open source favorisant le développement et l'adoption de services payants simplifiant le déploiement open source. »

La route à suivre : plus d'expérimentation, plus d'utilisation

Le paysage des modèles d'IA open source a connu une croissance significative en 2024, marquée par une augmentation des versions et une meilleure parité des performances avec leurs homologues propriétaires. Si la tendance se poursuit, davantage de modèles seront probablement rendus open source. De plus, des initiatives open source telles qu'EleutherAI, Hugging Face et OpenMined ont connu un essor considérable, alimentant davantage la dynamique du développement d'IA accessible et communautaire. Parmi les étapes clés de 2024, citons Llama 3 de Meta (paramètres 8B/70B), qui a surpassé des modèles fermés tels que Claude 3 Sonnet et Gemini Pro 1.5 lors des benchmarks, et DeepSeek-V3, un modèle open source rivalisant avec les meilleurs systèmes propriétaires en termes de vitesse d'inférence.⁴ Des entreprises comme Apple (OpenELM) et Microsoft (Phi-3-mini) ont élargi leurs offres open source, tandis que des start-ups comme Reka AI ont introduit des modèles multimodaux compatibles avec les fonctionnalités de GPT-4. Ces deux dernières années ont également été marquées par des collaborations, telles que SmoLLM2 et SmoVLM de Hugging Face, mettant l'accent sur l'accessibilité et l'efficacité. Bien que les modèles open source soient à la pointe de la transparence et de l'innovation communautaire, des défis subsistent pour faire évoluer l'infrastructure de formation afin de correspondre aux capacités de calcul de formation des systèmes propriétaires, associées à des cycles d'itération rapides.⁵

Pourtant, en moyenne, 75 % des personnes interrogées s'attendent à ce que leur organisation utilise davantage les technologies d'IA open source au cours des prochaines années. Les nuances varient bien sûr d'une organisation à l'autre, et les perceptions des dirigeants quant aux risques et aux cas d'utilisation les plus adaptés à l'open source sont diverses. Si environ 70 % des utilisateurs privilégient les technologies open source ou propriétaires, la plupart indiquent qu'ils envisageraient les deux (Figure 9). Très peu se sont déclarés « puristes » des technologies propriétaires ou open source.

⁴ Chronologie de l'intelligence artificielle, 2022 – présent, projet AI Timeline, GitHub, consulté en mars 2025 ; Weblog de Simon Willison, « Chronologie des versions de modèles d'IA en 2024 », 31 décembre 2024 ; Open LLMs for code, GitHub, consulté en mars 2025.

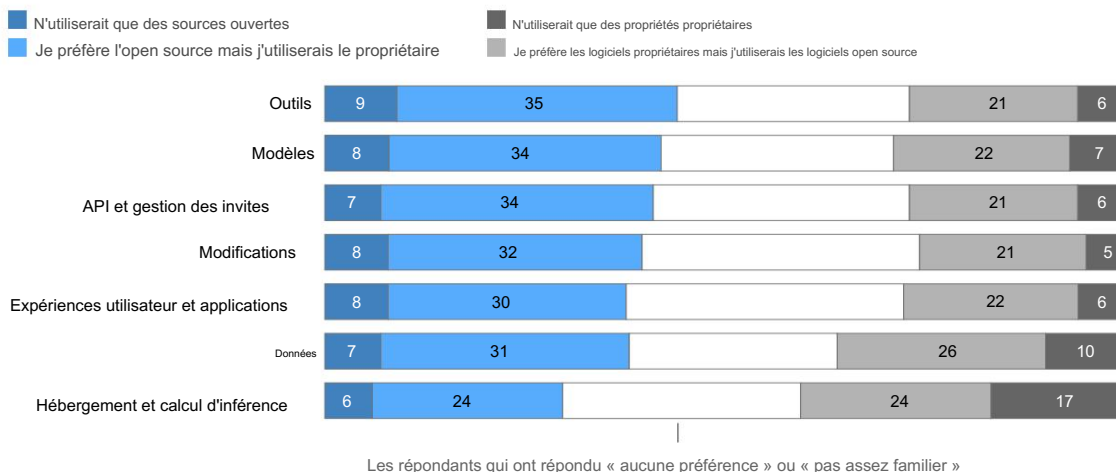
⁵ Ben Cottier et al., À quel point les modèles ouverts sont-ils en retard ?, Epoch AI, 4 novembre 2024.

Web <2025>

Pièce 9 de l'IA open source
Pièce <9> de <12>

Les organisations sont plus susceptibles de préférer les outils d'IA open source aux outils propriétaires dans leurs piles technologiques.

Préférence pour les outils d'IA open source par rapport aux outils propriétaires, % de répondants familiarisés avec l'IA open source¹



¹Outils, n = 633 ; modèles, n = 679 ; API et gestion des invites, n = 644 ; modifications, n = 558 ; expériences utilisateur et applications, n = 670 ; données, n = 624 ; hébergement et calcul d'inférence, n = 645.

Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

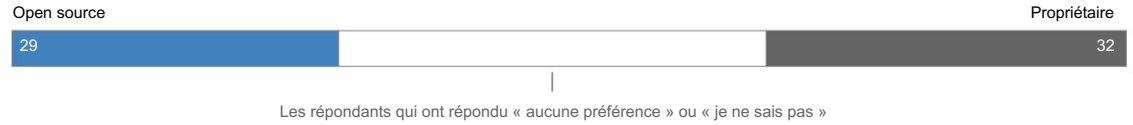
En ce qui concerne l'avenir, les préférences des dirigeants sont partagées quant à leur recours à l'IA open source ou propriétaire : 32 % des répondants indiquent que leurs dirigeants ont exprimé une préférence pour les technologies d'IA propriétaires, 32 % n'ont aucune préférence et 29 % privilégient l'IA open source. Ceux qui optent pour l'IA open source citent le coût comme principale raison dans 63 % des cas. Les dirigeants ayant une préférence stratégique pour l'IA propriétaire citent la sécurité, le risque et le contrôle du système comme principale raison dans 72 % des cas (graphique 10).

Lorsqu'ils comparent des outils open source et propriétaires similaires, les répondants prennent principalement en compte la sécurité, le risque et le contrôle du système (56 % des répondants) ; le coût (47 %) et la qualité (46 %) sont également des facteurs à prendre en compte.

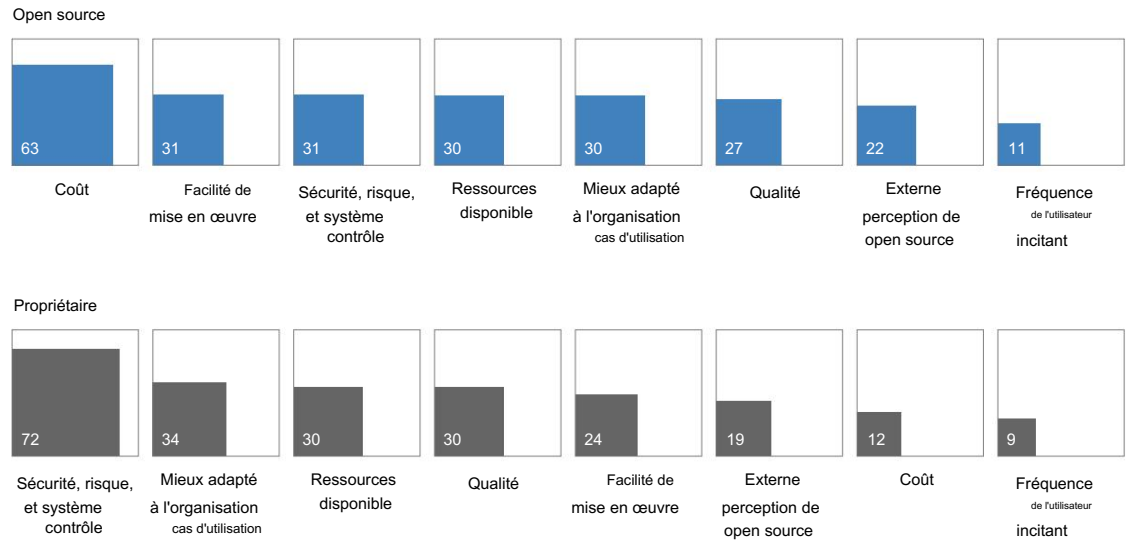
Malgré ces avantages avérés, les différences de coûts perçues entre l'IA open source et l'IA propriétaire pourraient être plus marquées que l'impact réel. À l'avenir, à mesure que l'écosystème de services prenant en charge les logiciels open source se mondialise et deviendra plus abordable, nous anticipons que l'écart de coûts entre les technologies open source et les technologies fermées pourrait encore se creuser.

Les dirigeants sont attirés par le coût des technologies d'IA open source et par les protocoles de sécurité des technologies d'IA propriétaires.

Préférence déclarée de la direction de l'organisation pour l'utilisation de logiciels open source plutôt que propriétaires
 Technologies de l'IA, % des répondants



Raisons de la préférence, % des répondants sélectionnant comme 3 premières réponses



Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

Si les répondants sont enthousiastes à l'idée d'accroître leur utilisation des technologies open source, rares sont ceux qui contribuent à l'écosystème open source. Seuls 13 % des répondants indiquent avoir contribué à des projets open source, et 50 % se disent incertains quant à leur participation à de futurs projets d'IA open source. Parmi ceux qui y ont participé, les raisons les plus fréquemment invoquées étaient l'attraction de talents et une image de marque positive. Ce faible taux de contribution pourrait s'expliquer par le fait que les contributions à l'open source sont souvent motivées par un sentiment d'altruisme et une croyance dans les idéaux du mouvement open source, plutôt que par un bénéfice financier direct. Les contributions de multiples parties prenantes, notamment les entreprises, les institutions universitaires et autres fondations, sont essentielles pour stimuler le progrès technique dans chaque domaine de la pile technologique de l'IA et favoriser un déploiement sûr et la réduction des risques (voir l'encadré « L'impact de l'open source sur l'avenir du développement de modèles »)

L'impact de l'open source sur l'avenir du développement de modèles

Les modèles open source peuvent accélérer l'innovation grâce à la collaboration, en réduisant les développements redondants et en favorisant le progrès collectif. Les innovations en matière d'IA open source sont susceptibles d'avoir des répercussions en aval sur deux tendances technologiques clés : les applications de périphérie axées sur la confidentialité et alimentées par des modèles de langage compacts (SLM) et l'émergence de modèles de raisonnement avec des temps de calcul d'inférence plus élevés.

Pour les applications utilisant des SLM, on peut s'attendre à l'émergence de modèles de langage sectoriels économiques, adaptés à des tâches spécifiques et déclinés en outils métier pour alimenter les applications. Des acteurs à très grande échelle comme Amazon Web Services (AWS), Google Cloud et Microsoft Azure proposent déjà de tels modèles pour des secteurs tels que l'industrie manufacturière et la finance. Les développeurs open source jouent également un rôle important dans la création de ces SLM, permettant la création de grands modèles de langage (LLM) polyvalents, dont les performances peuvent égaler, voire dépasser, celles des modèles plus importants.¹

Les petits modèles permettent également des applications de pointe et une intelligence sur appareil pour les organisations qui accordent la priorité à la latence et/ou à la confidentialité.² Certains

Parmi les exemples de plateformes de petits modèles distribuant des modèles open source (et autres), on peut citer le Qualcomm AI Hub, qui répond aux besoins des fabricants de produits d'IA de pointe, et Ollama, qui propose un cadre et des outils pour déployer des modèles ouverts sur les PC des utilisateurs avancés. Nous nous attendons à ce que ces plateformes intègrent des évaluations de tiers de confiance, des outils de certification (par exemple, AILuminat de MLCommons), renforçant la confiance des clients lors de la sélection des modèles.

La deuxième tendance clé est l'émergence de modèles de raisonnement, qui utilisent des capacités de calcul plus élevées pendant la phase d'inférence (plutôt que pendant la phase de pré-apprentissage) pour exceller dans des tâches spécifiques. Si la première vague de modèles de raisonnement était propriétaire (comme le modèle de raisonnement o1 d'OpenAI), des alternatives open source, notamment DeepSeek-R1 et un modèle similaire d'Alibaba, ont rapidement suivi. D'autres acteurs les développent et les adaptent. Perplexity a modifié une version de DeepSeek3 pour fournir des informations plus objectives et plus précises. Smolagents de Hugging Face a également créé un modèle alternatif de recherche approfondie⁴, proposant des offres plus compétitives.

d'OpenAI et de Google DeepMind.

D'autres technologies ouvertes émergent pour aider les constructeurs à optimiser et à améliorer leurs pipelines et processus de formation de modèles.

DeepSeek, par exemple, a continué à proposer des référentiels open source, incluant des capacités de parallélisme et d'intégration, pour ses modèles de

Si les capacités des modèles open source étaient autrefois inférieures à celles des modèles propriétaires, les modèles de base se sont considérablement améliorés. Si les entreprises peuvent rencontrer des difficultés pour adapter certains composants des modèles de raisonnement, les offres open source permettent désormais aux fournisseurs de services de modélisation de regrouper une gamme complète de technologies offrant une expérience développeur efficace, favorisant la modularité et tirant parti des avantages du développement communautaire. De plus, elles offrent aux entreprises une plus grande flexibilité et un plus grand choix pour déployer l'IA en périphérie ou dans le cloud, en fonction de leurs besoins en matière de confidentialité, de latence et de performances. Ce modèle opérationnel et cette flexibilité architecturale peuvent contribuer à la création de systèmes d'IA plus résilients, une qualité particulièrement précieuse dans un monde en constante évolution.

¹ « Comment l'open source façonne l'avenir de l'innovation », DevOps Online, consulté en mars 2025.

² Shreyas Subramanian, Vikram Elango et Mecit Gungor, Les petits modèles de langage (SLM) peuvent encore avoir du punch : une enquête, document de travail de l'Université Cornell, 3 janvier 2025.

³ Blog Perplexity, « Open-source RL 1776 », 18 février 2025.

⁴ Hugging Face Smolagents, Open Deep Research, GitHub, consulté en mars 2025.

Naviguer dans les risques présentés par l'IA open source

Malgré les avantages et la valeur de l'IA open source, de nombreux risques, principalement liés à la sécurité et à la confidentialité, associés aux outils open source pourraient affecter leur adoption. Les personnes interrogées considèrent que les outils d'IA open source sont plus risqués que l'IA propriétaire pour la plupart des types de risques liés à l'IA. Les risques les plus pertinents cités en matière d'IA sont la cybersécurité (62 % des personnes interrogées), la conformité réglementaire (54 %) et la propriété intellectuelle (50 %).

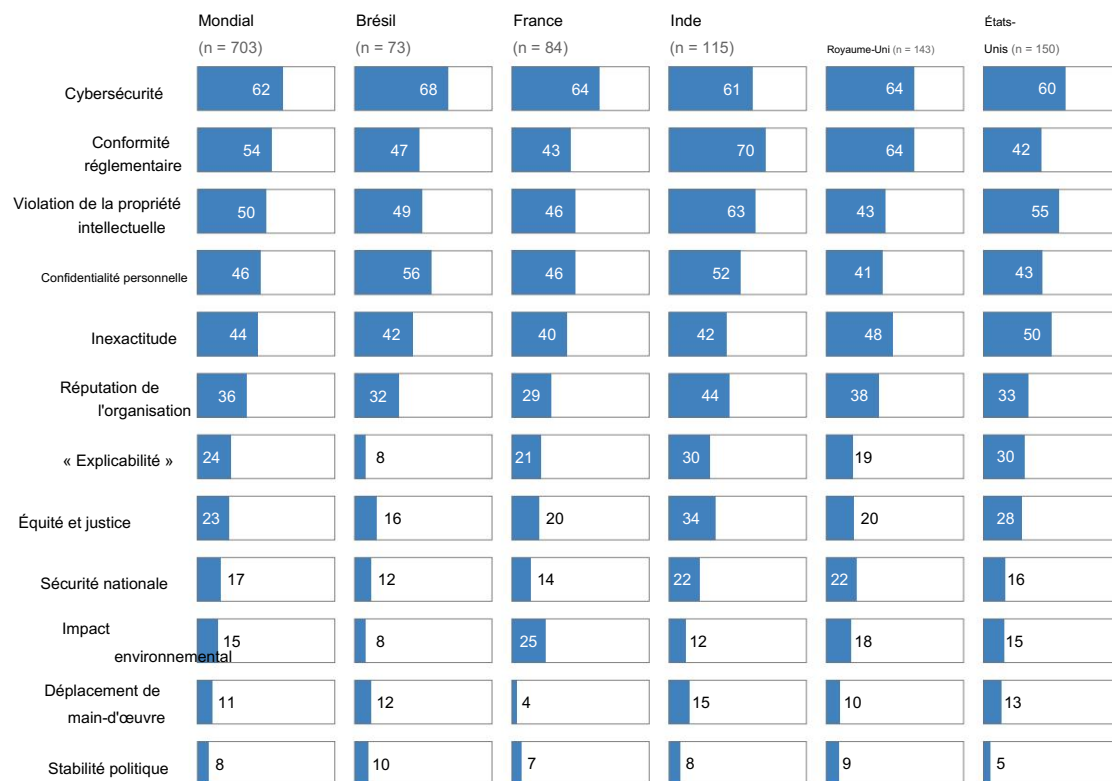
Cependant, la perception du risque varie selon les pays, probablement en fonction de leur environnement réglementaire, de leur tolérance au risque et de la maturité de l'IA (Figure 11). Les personnes interrogées au Brésil sont 10 % plus susceptibles d'identifier la cybersécurité comme un risque significatif par rapport à la moyenne mondiale. En Inde, 70 % des personnes interrogées expriment des inquiétudes quant à la conformité réglementaire, soit 31 % de plus que la moyenne. De plus, l'Inde affiche une inquiétude 26 % supérieure à celle des autres pays concernant les atteintes à la propriété intellectuelle.

Web <2025>

Enquête sur l'IA open source
Pièce <11> de <12>

La perception du risque lié à l'IA varie selon les pays.

Pertinence perçue du risque lié à l'IA, par région, % des répondants



Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

Bien que la plupart des développeurs que nous avons interrogés considèrent l'IA open source comme risquée, la perception du risque varie également en fonction de l'expérience du développeur (voir l'encadré « Prendre des mesures contre les risques potentiels »). Les développeurs plus expérimentés (définis comme ceux ayant contribué à six systèmes d'IA ou plus en production) sont bien plus à l'aise avec l'IA ouverte. Ils sont environ 11 % moins susceptibles de considérer l'open source comme plus risqué en termes de violation de propriété intellectuelle, 15 % moins susceptibles de considérer l'IA open source comme plus risquée en termes de cybersécurité et environ 11 % moins susceptibles de considérer l'IA open source comme plus risquée en termes de conformité réglementaire.

Prendre des mesures contre les risques potentiels

Notre enquête montre que les entreprises perçoivent des risques plus importants liés à l'open source. Alors, quelles mesures devraient-ils prendre ? Nous voyons quatre points clés que les dirigeants doivent prendre en compte lors de la mise en œuvre d'un système basé sur des modèles, qu'il soit open source ou propriétaire :

— Garde-corps : La mise en place de

des garde-fous robustes, tels que le filtrage automatisé du contenu, la validation des entrées/sorties et la surveillance humaine, peuvent contribuer à garantir une utilisation responsable et des sorties sécurisées.

Les exemples open source incluent NeMo Guardrails, Llama Guard et Guardrails AI de Nvidia, qui peuvent

aider en conformité avec les réglementations et normes éthiques.¹

— Évaluations par des tiers : une autre façon de renforcer la confiance dans le fait que les modèles ouverts

ne pas causer de préjudice involontaire, c'est mener des évaluations régulières avec des repères standardisés qui permettent la certification. Au cours de cette analyse comparative, les entreprises privées les évaluations garantissent que les ensembles de données de test sont gardé privé du modèle.²

— Documentation et suivi :

Sur le plan opérationnel, une nomenclature logicielle peut aider à suivre les divergences de version et les vulnérabilités en conservant des inventaires détaillés des composants open source.

Les évaluations quantitatives des risques, telles que le Common Vulnerability Scoring System Calculator (CVSS) v3.0, peuvent évaluer la gravité des vulnérabilités dans les systèmes open source.³

— Pratiques de cybersécurité : Pour garantir la confidentialité des données et l'intégrité du système,

L'exécution de modèles dans des environnements d'exécution sécurisés (TEE) peut contribuer à garantir le chiffrement des données sensibles pendant le traitement. L'intégration de techniques de confidentialité différentielle et d'apprentissage fédéré lors de l'entraînement peut empêcher les modèles de mémoriser des informations confidentielles. Des contrôles d'accès stricts au sein des référentiels de modèles, une segmentation du réseau entre les serveurs d'entraînement et d'inférence, une surveillance continue des incidents de sécurité et la vérification du hachage cryptographique garantissent l'origine fiable des modèles. Les référentiels sont quelques exemples qui peuvent répondre davantage aux défis de sécurité du contenu et de cybersécurité dans les environnements d'IA de production.

¹ Yi Dong et al., « Sauvegarde des grands modèles de langage : une enquête », IEEE Transactions on Pattern Analysis and Machine Intelligence, mai 2024.

² Tanmay Rajore et al., TRUCE : Analyse comparative privée pour prévenir la contamination et améliorer l'évaluation comparative des LLM, document de travail de l'Université Cornell, juin 2024.

³ Zoë Brammer et al., Châteaux construits sur du sable : Vers la sécurisation du système logiciel open source, Institute for Security and Technology, avril 2023 ; Lucie-Aimée Kaffee, « Rapports sur le hub : un premier regard sur l'autogouvernance dans le développement de l'IA open source », Hugging Face, 12 juin 2024.

Philip Reiner, PDG de l'Institute for Security and Technology :

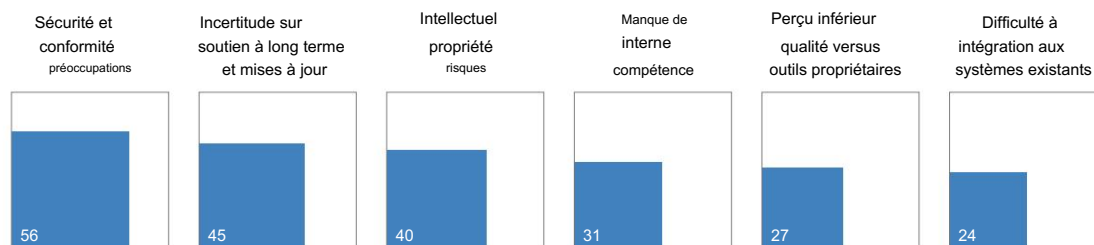
« Les organisations doivent revenir aux principes fondamentaux et se concentrer sur les pratiques de cybersécurité de base en matière d'IA. En réalité, les outils open source souffrent souvent d'une maintenance insuffisante et de logiciels obsolètes. Pour gérer ce risque, les entreprises doivent consulter régulièrement les rapports de non-conformité et désigner une équipe dédiée pour suivre les risques et les mises à jour des outils open source qu'elles utilisent. Nombre de cadres dirigeants ne comprennent pas pleinement les risques encourus, et nous verrons probablement de plus en plus de rôles évoluer vers l'assurance qualité. »

Pour les répondants qui n'utilisent pas actuellement l'IA open source, le principal obstacle à l'adoption réside dans les problèmes de sécurité et de conformité (56 % des répondants), parmi d'autres répertoriés dans la figure 12. Cependant, les organisations des répondants qui utilisent l'IA open source ont commencé à mettre en place des mesures de protection techniques lors de l'utilisation afin d'atténuer certains de ces risques potentiels. Plus d'un cinquième (21 %) des répondants utilisant l'IA open source mettent en œuvre des pondérations alignées, 35 % utilisent des garde-fous programmables (tels que NeMo Guardrails de Nvidia), 47 % utilisent des modèles de protection (par exemple, Llama Guard) et 49 % utilisent des ajustements rapides. Parmi ceux qui sont préoccupés ou très préoccupés par les risques associés à l'entraînement des modèles d'IA, 49 % consultent un conseiller juridique, 43 % recherchent des versions nettoyées des modèles, 57 % disposent de mesures de protection/tests techniques et 36 % évitent complètement les modèles open source. Parmi les répondants qui abordent les questions de droits d'auteur lorsqu'ils travaillent avec des ensembles de données à grande échelle, 53 % consultent un conseiller juridique et 46 % mettent en œuvre des mesures de protection et de test techniques. De plus, 29 % achètent des ensembles de données nettoyés et 33 % évitent complètement les ensembles de données open source.

Web <2025>
 Enquête sur l'IA open source>
 Pièce <12> de <12>

Les principaux obstacles à l'adoption d'outils d'IA open source sont les préoccupations liées à la sécurité et à la conformité, au support à long terme et à la propriété intellectuelle.

Obstacles signalés à l'adoption d'outils d'IA open source, % des répondants



Source : Enquête McKinsey sur l'IA Open Source, 703 participants ayant une expérience de travail avec des systèmes technologiques d'IA, du 9 décembre 2024 au 24 janvier 2025

L'avenir de l'IA open source

L'IA open source devient un élément clé du paysage émergent de l'IA. Notre enquête révèle une demande forte et croissante pour les technologies open source dans tous les secteurs, toutes les zones géographiques et toutes les technologies. Avec l'expansion de l'écosystème open source et l'augmentation de son utilisation, les entreprises auront une opportunité significative de réduire leurs coûts de développement et de déployer des systèmes d'IA personnalisés sur les appareils périphériques. Cependant, pour exploiter pleinement les avantages de l'open source, les entreprises devront peut-être renforcer leur confiance dans leur capacité à gérer les risques techniques et juridiques qu'elles perçoivent comme plus importants pour certaines utilisations de l'IA open source.

De plus, les développeurs expérimentés étant les principaux contributeurs et utilisateurs des projets open source, la montée en compétences des équipes est essentielle. Une collaboration accrue dans ce domaine favorisera probablement l'émergence de nouveaux services visant à améliorer l'expérience des développeurs. À l'instar de ce que nous avons observé dans les secteurs du cloud et des logiciels, une approche hybride deviendra probablement la norme, avec des technologies open source et propriétaires coexistant sur plusieurs couches de la pile technologique d'IA pour répondre à la diversité des besoins organisationnels.

À propos des auteurs et remerciements

[Ankit Bisht](#) est associé au bureau de McKinsey à Dubaï, et [Lareina Yee](#) est associé principal au bureau de la région de la baie de San Francisco, où [Roger Roberts](#) est associé, Brittany Presten est associée associée et Katherine Ottenbreit est consultante.

Les auteurs souhaitent remercier leurs partenaires de recherche de la Fondation Mozilla et de la Fondation Patrick J. Foundation McGovern ; leurs collègues de QuantumBlack Labs, la branche de développement logiciel et de R&D de QuantumBlack, AI by McKinsey, qui apporte des innovations en matière d'IA à ses clients et a apporté de nombreuses contributions à l'écosystème des logiciels open source dans l'IA et l'apprentissage automatique ; Cayla Volandes du bureau de McKinsey à New York ; Natasha Maniar du bureau de McKinsey dans la région de la baie de San Francisco ; et leurs collaborateurs universitaires externes pour leurs idées et leurs points de vue sur le projet d'enquête et son analyse, notamment Knut Blind de Fraunhofer ISI, Luca Vendraminelli de l'Université de Stanford, Sayash Kapoor de l'Université de Princeton et Genevieve Smith de l'Université de Californie à Berkeley.

Appendice

Données démographiques des répondants

Notre enquête s'est déroulée de décembre 2024 à janvier 2025, auprès de 703 répondants dans 41 pays, avec un accent particulier sur le Brésil, la France, l'Inde, le Royaume-Uni et les États-Unis.

Voir un résumé des données démographiques des répondants ci-dessous.

— Géographies :

- États-Unis : 150
- Royaume-Uni : 143
- Inde : 115
- France : 84
- Brésil : 73
- Europe (hors France) : 70
- Autre : 68

— Industries :

- Industries de pointe (y compris électronique avancée ; aérospatiale et défense ; automobile et assemblage ; semi-conducteurs) : 52
- Consommation (comprend les biens de consommation et les biens emballés ; vente au détail) : 45
- Énergie et matériaux (y compris l'énergie électrique et le gaz naturel ; l'ingénierie, la construction et les matériaux de construction ; le pétrole et le gaz ; les métaux et l'exploitation minière) : 43

- Services financiers et professionnels (y compris les services commerciaux, juridiques et professionnels ; les services financiers ; les investisseurs privés et les principaux investisseurs) : 157
- Santé et sciences de la vie (comprend les soins de santé, les produits pharmaceutiques et les produits médicaux) : 53
- Secteur public (comprend le secteur public ; le secteur social) : 59
- Technologie, médias et télécommunications (comprend les médias et le divertissement ; la technologie ; les télécommunications) : 260
- Autre (y compris l'agriculture ; produits chimiques ; immobilier ; voyages, logistique et infrastructures) : 34

Tout au long de l'enquête, nous catégorisons les répondants en deux archétypes principaux : les décideurs et les développeurs. Les décideurs sont des cadres, des vice-présidents et des cadres supérieurs. Les développeurs sont des développeurs/technologues seniors ou occupant des postes équivalents. Les répondants devaient posséder une expérience dans au moins un domaine de la pile technologique de l'IA. Les développeurs ont également été éliminés s'ils n'avaient pas contribué aux systèmes d'IA.

en production. Voir la répartition des titres de poste des répondants ci-dessous.

— Décideurs : 391 répondants

- Titres avec au moins cinq répondants : directeur de l'analyse/des sciences, directeur des données, directeur du numérique, directeur général, directeur financier, directeur de l'information, directeur de l'exploitation, directeur de la stratégie, directeur de la technologie, département/division/ chef d'unité commerciale, directeur général, directeur général, autre Cadre supérieur, autre cadre (niveau de vice-président ou supérieur), propriétaire/associé/directeur, vice-président principal, vice-président

— Technologues : 312

- Titres comportant au moins cinq répondants : ingénieur principal (ou équivalent technologue), autre technologue/ingénieur senior (niveau développeur/architecte senior et supérieur), ingénieur principal (ou équivalent technologue), ingénieur senior (ou équivalent technologue), ingénieur senior (ou équivalent technologue), ingénieur d'état-major (ou équivalent technologue)

Avril 2025

Copyright © McKinsey & Company

Conçu par McKinsey Global Publishing

[McKinsey.com](https://www.mckinsey.com)

