MÉTIERS DE L'IA

# MLOps : ce qu'il faut pour qu'un modèle d'IA tienne en production



Par Pierre-Louis Gaudreau, Président d'IA.consulting · Mis à jour le 20 octobre 2026

## La question que vous vous posez

Un prestataire vous a livré un modèle d'IA, ou vous en propose un. Dans le devis, une ligne parle de « MLOps ». Vous voulez savoir ce que c'est, et ce que cela coûtera quand le projet sera fini.

Un modèle est un programme qui a appris à partir d'exemples (vos factures, vos devis, vos tickets clients) et qui produit ensuite des prévisions ou des réponses.

Nous pensons que le MLOps est le coût caché de l'après-projet, parce que la livraison d'un modèle n'est que le début de son exploitation. Les pages sur le sujet parlent à des ingénieurs. Celle-ci parle à celui qui signe. Le métier se situe sur la carte des métiers de l'IA (ia.consulting/fr/metiers-ia/).

## Ce que recouvre le MLOps

MLOps contracte « machine learning » (apprentissage automatique : des programmes qui apprennent sur des exemples) et « operations » (exploitation). Google Cloud le définit comme une culture et une pratique d'ingénierie qui unit le développement et l'exploitation de ces systèmes [S1]. C'est une définition d'éditeur de cloud, orientée vers ses outils, sans coût ni taux de réussite [S1].

Deux cousins éclairent le mot. Le DevOps rapproche, pour le logiciel ordinaire, ceux qui écrivent le programme et ceux qui le font tourner. Le DataOps fait de même pour les données : les préparer, les contrôler, les livrer propres. Un modèle dépend des deux à la fois.

Dans un système réel, le code du modèle n'est qu'une petite fraction de l'ensemble [S2]. Le reste : les données, l'infrastructure (serveurs et réseaux) et la supervision (le suivi de ce que fait le modèle). Les chercheurs de Google le montraient déjà en 2015, par un schéma où le code tient dans une petite boîte au milieu [S4]. Aucune ne chiffre la proportion, et elles décrivent de très grands systèmes, pas des PME.

### Le cycle de vie d'un modèle

Cinq étapes, et la dernière ramène à la première : contrôler les données, entraîner et tester le modèle, passer en production (en service réel, après l'essai), surveiller ses résultats, le ré-entraîner quand ils se dégradent.

Les devis oublient la dernière. Un modèle vieillit sans que personne y touche : vos clients, vos produits, vos factures changent, et la qualité baisse. On parle de dérive. Sans surveillance, vous ne la voyez que lorsqu'un client se plaint.

### Les outils

Les plateformes MLOps abondent, chez les grands fournisseurs de cloud (serveurs loués à distance) et chez des éditeurs spécialisés. Nous ne les comparons pas : le choix revient à votre prestataire. Demandez-lui lesquelles il utilise et qui paie les licences.

## MLOps, AIOps, LLMOps : trois sigles à distinguer

La requête « ai ops » mène le plus souvent à l'AIOps, qui est autre chose. Ces définitions sont notre lecture de l'usage courant, sans source citée ici.

| Sigle | De quoi il s'agit |
|---|---|
| MLOps | Exploiter dans la durée des modèles d'IA construits ou adaptés pour votre activité |
| AIOps | Utiliser de l'IA pour exploiter l'informatique : repérer une panne, trier les alertes des serveurs |
| LLMOps | MLOps appliqué aux grands modèles de langage (LLM), ceux qui rédigent du texte comme les assistants de conversation |

En AIOps, l'IA sert à faire tourner l'informatique. En MLOps, l'informatique sert à faire tourner l'IA. Avec un abonnement à un assistant conversationnel, le LLMOps est l'affaire de l'éditeur. Si l'on adapte un modèle à vos données, vous entrez en MLOps.

## Ce que fait un ingénieur MLOps

Il construit et entretient la chaîne qui mène un modèle des données à la production, puis veille à ce qu'elle continue de tourner. Une journée type : lire les alertes, vérifier qu'une mise à jour des données n'a rien cassé, corriger.

Ses livrables se vérifient : un pipeline (enchaînement automatique des étapes, des données à la mise en service) documenté, un tableau de suivi de la qualité, une procédure pour

revenir à la version précédente. Exigez ce dernier point par écrit. Le métier voisin, qui construit le modèle plutôt que de l'exploiter, est décrit dans le métier d'ingénieur IA (ia.consulting/fr/metiers-ia/ingenieur-ia/).

Le profil mêle programmation (souvent Python), cloud, conteneurs (paquets qui embarquent un programme avec ce qu'il faut pour tourner à l'identique partout), statistiques de base et culture DevOps. Nous n'avons pas de source vérifiée sur les cursus, les salaires ni le nombre d'offres. Nous n'en donnons donc aucun.

## Ce que les autres pages ne disent pas : le coût de l'après-projet

Aucune page lue ne chiffre l'exploitation d'un modèle dans la durée. Nous non plus : aucun montant sourcé pour une PME. Nous avons trois indices.

Le premier est ancien et qualitatif. Les chercheurs de Google constataient en 2015 qu'il est courant d'encourir des coûts de maintenance continus très élevés dans les systèmes réels [S3]. Aucun montant, et un article antérieur aux grands modèles de langage.

Le deuxième est un témoignage. Un responsable de Deloitte, cité par la presse spécialisée, rapporte qu'une enquête auprès de ses clients montrait que 20 % seulement des modèles d'IA parviennent en production [S6]. L'enquête n'est ni nommée ni datée, l'article date de 2023.

Le troisième est une prévision. En août 2024, Gartner annonçait qu'au moins 30 % des projets d'IA générative seraient abandonnés après la preuve de concept (l'essai initial) fin 2025, entre autres à cause de coûts croissants [S5]. Elle porte sur l'IA générative et ne parle pas du MLOps. À revérifier contre le bilan de Gartner.

Aucun ne donne un budget. Ils justifient une habitude : demander le prix de l'après-projet avant de signer l'avant-projet.

## Pour le dirigeant : recruter, louer ou former

### Trois façons de porter le MLOps

- **L'éditeur.** Vous achetez un logiciel d'IA en abonnement, l'exploitation est comprise. C'est le cas le plus courant en PME, notre lecture sans source. Vérifiez au contrat qui surveille, qui corrige, sous quel délai.
- **Le prestataire.** Une agence ou un indépendant a construit le modèle. Il faut un contrat de maintenance après livraison, au tarif écrit. Choisir un expert, un consultant ou une agence (ia.consulting/fr/ia-pme/budget-choix/choisir-consultant/) détaille la comparaison.
- **L'interne.** Vous recrutez un ingénieur MLOps, ou vous formez des salariés (ia.consulting/fr/ia-pme/equipes-controle/formation-equipes/) qui connaissent déjà vos données. Notre lecture : cela se justifie avec plusieurs modèles, rarement avec un seul.

### Les lignes de coût après la livraison

Aucun montant ici. Demandez ces lignes une par une : hébergement et licences, supervision (qui regarde, à quelle fréquence), ré-entraînement et son prix, temps de vos équipes pour valider les résultats, mise à jour quand l'éditeur change ses conditions. Les lignes du budget IA après le projet (ia.consulting/fr/ia-pme/budget-choix/) donnent le cadre.

### Cinq questions à poser avant de signer

1. Qui surveille le modèle, et comment suis-je prévenu d'une baisse de qualité ?
2. À quelle fréquence est-il ré-entraîné, et à quel prix ?
3. Puis-je revenir à la version précédente, et en combien de temps ?
4. Que récupère mon équipe si nous changeons de prestataire : code, documentation, accès ?
5. Quelles lignes de coût s'ajoutent après la livraison ?

### Cas construit

Cas construit : une entreprise de négoce fait prévoir ses délais de livraison par un modèle. Six mois plus tard, de nouveaux fournisseurs faussent les prévisions. Sans contrat de suivi, personne ne le voit avant les plaintes de clients. Avec une supervision prévue au devis, l'écart apparaît vite et le ré-entraînement est déjà payé.

---

## Questions fréquentes

**AIOps et MLOps, est-ce la même chose ?** Non. L'AIOps utilise l'IA pour exploiter l'informatique. Le MLOps exploite des modèles d'IA.

**Quelle différence avec le DevOps ?** Le DevOps traite le logiciel ordinaire. Le MLOps y ajoute les données et le modèle, qui peuvent se dégrader sans que le code change.

**Une PME a-t-elle besoin d'un ingénieur MLOps ?** Rarement à temps plein. Le plus souvent, l'éditeur ou le prestataire tient ce rôle. Il faut que le contrat dise qui.

**Combien coûte le MLOps ?** Nous n'avons aucun chiffre sourcé pour une PME. Il dépend du nombre de modèles, de la fréquence de mise à jour et de l'hébergement. Demandez-le par écrit.

---

## La question suivante

Une fois l'après-projet compris, reste à l'inscrire dans vos comptes. Les lignes du budget IA après le projet (ia.consulting/fr/ia-pme/budget-choix/) montrent où placer hébergement, supervision et ré-entraînement dans l'enveloppe totale.

## Sources


- **S1** · Google Cloud (Architecture Center), MLOps : pipelines de livraison continue et d'automatisation en apprentissage automatique, 2024-08. https://docs.cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning (https://docs.cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning)
- **S2** · Google Cloud (Architecture Center), même page, part du code du modèle dans un système réel, 2024-08. https://docs.cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning (https://docs.cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning)
- **S3** · NeurIPS (Sculley et al., Google), Hidden Technical Debt in Machine Learning Systems, 2015-12. https://proceedings.neurips.cc/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf (https://proceedings.neurips.cc/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf)
- **S4** · NeurIPS (Sculley et al., Google), même article, figure 1, 2015-12. https://proceedings.neurips.cc/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf (https://proceedings.neurips.cc/paper/2015/file/86df7dcfd896fcaf2674f757a2463eba-Paper.pdf)
- **S5** · Gartner (communiqué du 5 août 2024, repris par IntelligentCIO), 30 % des projets d'IA générative abandonnés après la preuve de concept, 2024-08. https://www.intelligentcio.com/eu/2024/08/05/gartner-predicts-30-of-generative-ai-projects-will-be-abandoned-after-proof-of-concept-by-end-of-2025/ (https://www.intelligentcio.com/eu/2024/08/05/gartner-predicts-30-of-generative-ai-projects-will-be-abandoned-after-proof-of-concept-by-end-of-2025/)
- **S6** · Solutions Numériques (Alain Clapaud), propos de Subhadip Roy (Deloitte), 2023-09. https://www.solutions-numeriques.com/?p=110371 (https://www.solutions-numeriques.com/?p=110371)





**PUBLIÉ SOUS LA RESPONSABILITÉ DE**

**Pierre-Louis Gaudreau**, Président d'IA.consulting. Sources citées et datées. Une erreur ? Écrivez-nous.