

Ingénieur MLOps

Rattachement : Head of R&D

Prise de poste : dès que possible.

Vous pilotez l'ensemble du cycle de vie des modèles Physora : de la construction de l'infrastructure d'entraînement (scalable, reproductible et traçable) à leur industrialisation performante (inférence sous contraintes usine, on-premise et edge), en passant par la transformation des données brutes en jeux exploitables et l'optimisation des architectures via des expériences rigoureuses, tout en y apportant une discipline de production éprouvée, pour éviter les écueils passés et garantir des solutions fiables, différenciantes et livrables.

C'est le poste qui détermine la cadence. Il fait la différence entre une équipe qui lance trois expériences par semaine et une qui en lance trente — ce qui, dans ce domaine, constitue l'essentiel de l'écart entre réussite et échec.

Vos missions

- Construire et maintenir l'infrastructure d'entraînement : chargement de résultats volumineux, entraînement distribué multi-GPU, suivi d'expériences, registre de modèles, reproductibilité.
- Développer l'outillage qui transforme les sorties brutes de simulation en jeux de données exploitables, organisation du stockage, schéma, versionnement.
- Amener les modèles au niveau de performance d'inférence attendu : quantification, export, service sous budget de latence usine, on-premise et en périphérie (edge).
- Mener expériences et ablations, et en présenter les résultats de façon à éclairer les choix d'architecture, pas seulement à rapporter des chiffres.
- Contribuer à l'implémentation des modèles aux côtés de l'ingénieur en modélisation.
- Apporter une discipline de production venue de l'extérieur. Vous serez la personne qui a déjà vu cela échouer ailleurs, et l'on attend de vous que vous le disiez.

Profil recherché

- Master ou Doctorat en Informatique (spécialisation Machine Learning / IA) ou en Physique Numérique, complété par une expérience professionnelle significative en ingénierie des systèmes d'entraînement distribué (PyTorch, calcul haute performance, gestion de données massives).
- Expérience substantielle de l'infrastructure d'entraînement en production, hors contexte académique. Sur ce poste précisément, ce point n'est pas négociable.
- PyTorch et son écosystème en profondeur : entraînement distribué, profilage, et l'expérience du débogage qui vient d'avoir vu des modèles échouer silencieusement.
- À l'aise avec de grands volumes de données numériques et les formats de stockage qui les rendent exploitables HDF5, Zarr, Parquet ou équivalent.
- Une vraie discipline d'ingénierie logicielle : tests, gestion de versions, code lisible par d'autres.
- La curiosité pour la physique est l'essentiel. Pas besoin d'être expert en mise en forme mais vous devez avoir l'envie de la comprendre, de la décrypter, et de vous l'approprier.

Atouts appréciés

- Environnements HPC et SLURM ; entraînement distribué à grande échelle.
- Calcul scientifique (C++, Fortran, MPI).
- Service de modèles on-premise et inférence embarquée ; outillage MLOps.

Langues : nos langues de travail sont le français et l'anglais.