

Stage Ingénieur Développement IA

Titre : Benchmark d'un système agentique expert pour le diagnostic et le chiffrage technique

Lieu : Bordeaux – Village by CA, 11 cours du 30 Juillet

Début : Janvier ou février 2026 – **Durée :** 5 à 6 mois

Organisation : Télétravail partiel possible

Rémunération : Selon profil

Candidature : CV + court mail de motivation à envoyer à rh@daryl-ai.com

Contexte de l'entreprise

DARYL révolutionne la vente de produits techniques dans l'industrie grâce à un assistant IA capable de générer en quelques secondes des diagnostics techniques et des devis commerciaux connectés aux outils CRM, CPQ et ERP.

Notre solution est déjà utilisée par plusieurs PME industrielles qui l'emploient pour transformer leur cycle de vente, mieux qualifier leurs leads et réduire drastiquement les délais de réponse aux clients.

Installée au Village by CA Bordeaux, DARYL connaît une croissance soutenue, portée par la montée en puissance des outils d'IA appliqués au CPQ industriel.

Objectif du stage

L'objectif du stage est de concevoir et de mettre en place des métriques robustes pour évaluer un système agentique expert s'appuyant sur des technologies avancées d'orchestration d'agents et de gestion de connaissance (RAG, KAG).

En partant du graphe agentique existant, il faudra comparer différentes configurations basées sur différents modèles de fondation, afin de mesurer leur impact sur la qualité du diagnostic, la pertinence du chiffrage et la robustesse globale du système.

Missions principales

Vous réaliserez un état de l'art des approches d'évaluation de systèmes agentiques et des métriques utilisées pour mesurer la performance de systèmes d'IA complexes orientés production.

Définition et formalisation d'un ensemble de métriques quantitatives et qualitatives pour évaluer le système (exactitude des réponses, stabilité, temps de réponse, coût, robustesse, expérience utilisateur, etc.). Instrumentation de la stack pour remonter ces signaux.

Vous mettrez en œuvre des campagnes structurées de benchmarking en faisant varier les architectures agentiques, les modèles de fondation et les configurations de RAG / KAG, afin d'identifier les compromis performance / coût / complexité les plus pertinents.

Vous produirez un rapport de synthèse opérationnel, incluant recommandations, visualisations de résultats et guidelines pour l'évolution du système en environnement produit.

Profil recherché

Vous êtes étudiant(e) en école d'ingénieur ou en Master (IA, data, génie logiciel, mathématiques appliquées) et vous souhaitez travailler au cœur de systèmes d'IA déployés en conditions réelles.

Vous disposez de solides bases en machine learning, NLP et Python, et vous êtes à l'aise avec l'évaluation de modèles et la manipulation de données expérimentales.

Vous avez idéalement une expérience (ou un intérêt marqué) pour des outils comme LangFuse, MLflow, LangChain ou LangGraph, ainsi que pour l'observabilité et le suivi des performances de systèmes d'IA.

Vous appréciez les systèmes complexes et industriels, vous êtes sensible aux enjeux d'IA explicable et vous aimez concevoir des méthodologies d'évaluation rigoureuses.

Encadrement & environnement

Vous serez encadré(e) directement par le CTO de DARYL et intégré(e) à une équipe pluridisciplinaire mêlant IA, produit, tech et business.

Vous évoluerez dans un environnement dynamique au cœur de l'écosystème French Tech Bordeaux, avec des cas d'usage concrets, visibles et rapidement impactants chez des clients industriels.