

# Stage M2-Data science – Retrieval Augmented Generation LLMs (H/F)

---

## Sommaire

---

<b>A.</b>	<b>Fonction .....</b>	<b>3</b>
<b>B.</b>	<b>Résumé de la fonction .....</b>	<b>3</b>
<b>C.</b>	<b>Positionnement dans l'organisation .....</b>	<b>3</b>
<b>D.</b>	<b>Missions principales .....</b>	<b>3</b>
<b>E.</b>	<b>Compétences métiers &amp; techniques .....</b>	<b>3</b>
<b>F.</b>	<b>Aptitudes professionnelles .....</b>	<b>4</b>
<b>G.</b>	<b>Signature pour acceptation .....</b>	<b>4</b>

## A. Fonction

---

Stagiaire M2 data science dans l'équipe Global AI Lab rattaché à la R&D.

## B. Résumé de la fonction

---

Le candidat, dans une démarche de recherche et d'ingénierie devra adresser la problématique de pertinence des modèles d'IA génératifs (LLMs) augmentés avec du contexte (RAG : Retrieval Augmented Generation) versus des modèles ajustés (fine tuned). Il s'agit en particulier de voir quelle approche est plus précise, plus exhaustive et plus viable dans un environnement de production. Dans un second temps, il s'agit de proposer une nouvelle approche qui améliore les résultats de l'état de l'art.

## C. Positionnement dans l'organisation

---

- Site : Grenoble
- Service : équipe Global AI Lab rattaché à la R&D
- Périmètre : France
- Responsable hiérarchique (N+1) : Ali AIT-BACHIR
- Responsabilités d'encadrement : Ali AIT-BACHIR (aait-bachir@easyvista.com)

## D. Missions principales

---

- Etat de l'art sur les LLMs et le RAG ;
- Définition d'une méthode d'évaluation pour un benchmark
- Définition d'une nouvelle méthode
- Benchmark et évaluation sur les méthodes existantes et la méthode proposée
- Rapport et interprétation des résultats
- Rédaction du rapport de fin d'étude
- Présentation et soutenance

## E. Compétences métiers & techniques

---

De formation ingénieur informatique orienté data science avec une expérience significative dans ce type d'architecture applicative distribuée à forte contrainte et volumétrie.

- Maîtrise des concepts de Machine Learning
- Compréhension des modèles statistiques et mathématiques
- Conception et implémentation de nouveaux modèles analytiques
- Familiarisé avec les modèles NLP (Natural Language Processing)
- Développement en Python des algorithmes de machine learning
- Méthode Agile et industrialisation : tests unitaires, Git, etc.

## F. Aptitudes professionnelles

---

Ces tâches s'effectueront dans un contexte Agile en place depuis plusieurs années, fortement axé sur le partage entre équipes et la collaboration étroite avec les différents intervenants sur le produit. Une bonne culture du résultat, un esprit d'initiative avec une force de proposition seront des atouts majeurs pour la réussite de ce stage.

## G. Signature pour acceptation

---

Nom/prénom	
Date et Signature	