## Stage de fin d’étude : Exploitation des données de maintenance du parc hydraulique à l’aide d’Intelligence Artificielle textuelle

Contexte :

Créée en 1946 pour le contrôle et la surveillance des ouvrages hydrauliques français, la  
Division Technique Générale (DTG) est **une unité d’ingénierie spécialisée dans les activités de**  
**mesure et d’expertise** et intervient sur tout le parc de production d’EDF.   
Les activités principales de DTG sont :

* **La Surveillance, le diagnostic et le pronostic** pour les sites de production d’électricité du parc EDF (nucléaire, hydraulique, thermique et renouvelables) pour conseiller les exploitants en matière de sûreté et de performance
* **Le développement de techniques innovantes et des connaissances nécessaires à l’exploitation des parcs de production,** pour apporter les réponses au besoin de développement énergétique et aux enjeux environnementaux
* La contribution au développement d’EDF en assurant ses missions pour le compte de filiales du Groupe ou de tiers, en France et à l’international

Elles sont organisées autour de 4 départements

* Le département mécanique dont les domaines d’expertises sont le diagnostic et les contrôles mécaniques, la mécanique vibratoire, l’acoustique, la performance thermodynamique et la performance hydraulique
* Le département électricité dont les domaines d’expertises sont les matériels électriques (alternateurs, transformateurs, régulateurs)
* Le département ouvrage et génie civil dont les domaines d’expertises sont l’auscultation, la surveillance et la topographie
* Le département eau et environnement dont les domaines d’expertises sont l’hydrométéorologie et les mesures environnementales

Au sein de DTG l’équipe du **LabData regroupe des datascientists ayant pour mission de développer des solutions d’Intelligence Artificielle (IA) en appui aux départements pour leurs activités de surveillance et de diagnostic.**

Description du stage :

Au sein du LabData, le stage a pour objectif d’exploiter les bases de données d’évènements d’exploitations (maintenances et avaries sur les matériels) du parc hydraulique. Il s’agit de données textuelles, renseignées par les exploitants. Des modèles d’IA de Traitement Automatique du Langage naturel (TAL ou NLP en anglais) seront utilisés pour déterminer l’historique des maintenances, les avaries, et les modes de défaillances des matériels. Ces résultats seront ensuite utilisés pour les études de fiabilité des matériels du parc hydraulique.

Le stage comportera plusieurs phases :

* Prise en main des données et rencontre avec le métier
* État de l’art des méthodes envisagées et montée en compétence
* Amélioration des modèles existants
* Développement de modèles sur de nouveaux matériels (annotation avec le métier, choix du modèle, apprentissage, déploiement)
* Rédaction d’une note recensant les travaux effectués

Profil souhaité :

Vous êtes actuellement étudiant.e en d’école d’ingénieur ou en master 2 avec une spécialité en data science, mathématiques appliquées, IA ou deep learning.

Vous avez un intérêt pour le secteur de l’énergie et pour les solutions d’IA textuelles (NLP).

Une bonne maîtrise de Python, ainsi que des frameworks usuels tels que pytorch ou scikit-learn seront nécessaires.

Vous serez amené.e à mettre en œuvre et/ou acquérir des compétences en Machine Learning, Deep Learning (Architecture Transformers), TAL(Traitement Automatique du Langage)/NLP (Natural Langage Processing), Python, développement logiciel et en lecture et synthèse d’article de recherche

Dates : à partir du premier trimestre 2024  
Durée du stage : 5 à 6 mois  
Lieu du stage : EDF DTG Grenoble, 134 chemin de l’étang, Saint Martin le Vinoux  
Horaires : 35 h / semaine.  
Indemnité : variable selon la formation, suivant les grilles EDF