INGÉNIEUR DATA & MODÉLISATION

Profil recherché

Diplôme d'ingénieur ou doctorat (Bac+5 à Bac+8) en intelligence artificielle ou apprentissage automatique, avec de solides compétences en traitement du signal et en biostatistique. Une expérience en bioacoustique ou en données multimodales serait un vrai atout.

Le projet

Le projet vise à caractériser l'état affectif des orques (Orcinus orca) du détroit de Gibraltar en alliant trois disciplines : la bioacoustique (répertoire vocal, diversité et fréquence des appels), l'éthologie (postures, synchronie de groupe) et la physiologie (cortisol et ocytocine comme marqueurs d'activation négative et positive), croisées dans un modèle d'intelligence artificielle pour identifier les patterns acoustiques collectifs associés à ces états.

Où vous travaillerez

Vous travaillerez au sein du laboratoire CETRA, rattaché à Aix-Marseille Université, principalement basé(e) à Marseille/Aix. Vous intégrerez l'équipe du programme ECHO, avec un déplacement sur le terrain du détroit de Gibraltar en début de projet.

Quotité de travail : 100 %. Vous débuterez la semaine du 9 novembre.

Contraintes

- <> Volume et complétude des données dépendants des conditions de collecte terrain.
- <> Nécessité de concevoir un modèle robuste malgré des données potentiellement incomplètes ou bruitées.
- <> Exigence de transparence méthodologique (le modèle devant être explicable, pas seulement performant, pour un usage scientifique).

Compétences et savoir-être requis

- <> Maîtrise des architectures de deep learning multimodal et des méthodes d'apprentissage supervisé.
- <> Compétences en traitement du signal acoustique.
- <> Maîtrise des outils de biostatistique et de gestion de bases de données.
- <> Capacité à vulgariser des choix techniques complexes auprès d'une équipe non spécialiste.
- <> Maitriser l'anglais scientifique.

Vos Missions

- <> Calibration initiale du matériel de collecte (hydrophones, drones, D-Tags) avec les systèmes de traitement.
- <> Conception de l'architecture de deep learning multimodal (alignement temporel par fenêtres glissantes entre acoustique, comportement et physiologie).
- <> Entraînement et validation du modèle, visant à identifier les patterns acoustiques collectifs associés à l'état affectif des orques.
- <> Gestion des données manquantes en mode dégradé.
- <> Évaluation de l'opportunité d'intégrer des outils existants (ex. orcAI) comme brique de classification acoustique de base.
- <> Analyse statistique des corrélations entre signatures acoustiques et états hormonaux, en lien avec le data analyste/biostatisticien.

Hygiène et sécurité

- <> Poste majoritairement en laboratoire ; déplacements ponctuels en mer en début de projet pour la calibration des outils de collecte.