OFFRE DE STAGE MASTER 2 / ÉCOLE D'INGÉNIEURS

Apprentissage géospatial et modèles de fondation pour la cartographie du risque de rage animal–humain

Durée	6 mois	**Lieu**	UMR ESPACE-DEV, Montpellier
Niveau	Master 2 / dernière année ingénieur	**Partenaire**	IOC/FIOCRUZ, Brésil

Contexte et problématique

La rage est une zoonose mortelle dont le risque de transmission dépend à la fois de la circulation du virus chez les animaux, de la distribution des espèces hôtes, des caractéristiques environnementales et de la présence humaine. **GeoRabIA (Geospatial and AI-Driven Modeling of Animal-to-Human Rabies Transmission Risk)** vise à développer une approche d'intelligence artificielle capable de produire des cartes spatiales du risque de rage animale et d'exposition humaine, en combinant données épidémiologiques géolocalisées, contexte environnemental, informations sur les espèces et représentations géospatiales apprises [1–5].

Le projet est porté par l'UMR ESPACE-DEV (IRD / Université de Montpellier) en collaboration avec l'**Instituto Oswaldo Cruz (IOC/FIOCRUZ), Brésil**, qui apporte son expertise en éco-épidémiologie, surveillance des maladies infectieuses et analyse de données de rage humaines et animales au Brésil. Les données épidémiologiques nécessaires au projet ont déjà été collectées par **l'IOC/FIOCRUZ** et seront mises à disposition dans le cadre de cette collaboration, selon les conditions d'accès et d'utilisation applicables.

Objectifs et méthodologie

Le stage consistera à construire et évaluer une chaîne reproductible de modélisation du risque de rage à partir des données disponibles. Les travaux s'organiseront autour des axes suivants :

- **Préparation et harmonisation des données :** contrôle qualité, harmonisation spatiale et temporelle, gestion des coordonnées, espèces, attributs épidémiologiques et variables environnementales.
- **Extraction de représentations pré-entraînées :** extraction des représentations avec SatCLIP [2] pour les informations de localisation et TaxaBind [3] pour les informations écologiques et multimodales.
- **Modélisation :** comparaison de méthodes de machine learning classiques et de modèles de deep learning intégrant les embeddings géospatiaux, écologiques et les variables épidémiologiques.
- **Transfer learning et few-shot learning :** comparaison de représentations gelées, fine-tuning et stratégies adaptées à la rareté et au déséquilibre des cas de rage.
- **Cartographie du risque :** production de cartes du risque relatif de présence de la rage animale à partir des données épidémiologiques et environnementales disponibles. Ces cartes seront ensuite croisées avec les données de répartition de la population humaine issues du Global Human Settlement Layer (GHSL) [5], afin d'identifier des zones potentielles d'exposition humaine. Les résultats seront accompagnés d'une analyse des incertitudes et des limites associées aux données de surveillance.

Profil recherché et compétences requises

Étudiant·e en Master 2 ou dernière année d'école d'ingénieurs en informatique, intelligence artificielle, data science, géomatique, télédétection ou domaine proche.

Compétences requises	Compétences appréciées
• Bonne maîtrise de Python et des outils scientifiques (NumPy, Pandas, scikit-learn). • Bases solides en machine learning et deep learning. • Expérience avec PyTorch ou un framework équivalent. • Capacité à conduire des expérimentations, évaluer les modèles et analyser les résultats. • Lecture de publications scientifiques en anglais, autonomie et rigueur.	• Analyse de données géospatiales : GeoPandas, Rasterio/GDAL, QGIS. • Notions de télédétection, données raster/vectorielles et systèmes de coordonnées. • Intérêt pour les foundation models, representation learning, multimodal learning ou self-supervised learning. • Connaissances en transfer learning, few-shot learning ou apprentissage sur données rares/déséquilibrées.

Encadrement et environnement scientifique

Le stage sera réalisé au sein de l'UMR ESPACE-DEV à Montpellier, dans le cadre du projet GeoRabIA, avec une collaboration scientifique avec l'IOC/FIOCRUZ au Brésil. L'encadrement réunira des compétences en deep learning, géospatial AI, télédétection, apprentissage frugal / à partir de peu de données et éco-épidémiologie. Des échanges réguliers avec les partenaires brésiliens permettront de confronter les choix méthodologiques aux connaissances épidémiologiques et aux données de terrain.

Candidature

Les candidatures sont à envoyer aux adresses suivantes : sara.mobsite@ird.fr, celia.blondin@ird.fr et thomas.hallopeau@ird.fr, en précisant « STAGE GeoRabIA » dans l'objet du mail.

Le dossier de candidature devra comprendre : un CV, une courte lettre de motivation, les relevés de notes récents et, si disponible, un lien vers un dépôt GitHub/GitLab ou des projets antérieurs.

Gratification et financement

La gratification du stage sera conforme à la réglementation en vigueur.

Ce travail a bénéficié d'une aide : de l'Etat gérée par l'Agence Nationale de la Recherche (ANR) au titre de France 2030 et par l'Union européenne NextGenerationEU portant la référence « ANR-21-EXES-0005 » ; de la Région Occitanie ; et de l'Institut ExposUM de l'Université de Montpellier.

Références

[1] Catozo, R. G., et al. "Two Decades of Human Rabies in Brazil: Epidemiological Trends, Emerging Risks and Treatment Challenges." Revista da Sociedade Brasileira de Medicina Tropical, 58, e0110-2025, 2025.

[2] Klemmer, K., et al. "SatCLIP: Global, General-Purpose Location Embeddings with Satellite Imagery." Proceedings of the AAAI Conference on Artificial Intelligence, 39(4), 2025.

[3] Sastry, S., et al. "TaxaBind: A Unified Embedding Space for Ecological Applications." IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2025.

[4] Food and Agriculture Organization of the United Nations. EMPRES-i: Global Animal Disease Information System.

[5] European Commission, Joint Research Centre. Global Human Settlement Layer (GHSL), 2025.