

Postdoctorat en science de données et analyse prédictive

Laboratoire CEDRIC, Conservatoire national des arts et métiers (Cnam) Paris

Projet	Projet Prima Europe CAMELAIT4MED
Type de contrat	Postdoctorat – Science de données
Date de début	Décembre 2026
Durée	16 mois
Location	Laboratoire CEDRIC, Conservatoire national des arts et métiers, 2 rue Conté 75003 Paris
Contact	Faten.atigui@lecnam.net ; vincent.audigier@lecnam.net

Contexte

La région méditerranéenne est confrontée à des défis majeurs, aggravés par la rareté de l'eau et les effets croissants du changement climatique. Le développement de la production de lait de chamelle apparaît comme une alternative durable à l'élevage laitier bovin conventionnel, très consommateur de ressources (eau/énergie, fourrages, médicaments, etc.), en offrant à la fois résilience et bénéfices environnementaux. Le projet CAMELAIT4MED vise à favoriser la transformation durable de la chaîne de valeur du lait de chamelle grâce à la mise en œuvre holistique de technologies et de pratiques innovantes à différentes étapes de la chaîne de production. Un des objectifs du projet CAMELAIT4MED est le suivi et la promotion du bien-être et de la santé des chameaux grâce à l'adoption de techniques d'élevage de précision. En s'appuyant sur les technologies des capteurs (applications IoT), le projet vise à surveiller la santé, la productivité et le bien-être du troupeau en temps réel, améliorant ainsi les normes de bien-être animal et la productivité comme l'acceptabilité par les populations tout en minimisant l'usage des antibiotiques et le recours au vétérinaire. Ces technologies, couplées à des méthodes d'analyse innovante de données massives visent à extraire des connaissances relatives à la prédiction du stress et des maladies des chameaux.

Le projet mobilise des approches interdisciplinaires associant informatique, science des données, intelligence artificielle et analyse du vivant instrumentée. Les données proviennent de capteurs portés ou embarqués, de capteurs environnementaux, de systèmes de suivi automatisés et/ou de bases de données associées aux dispositifs expérimentaux ou de terrain. La diversité des flux, la variabilité temporelle, les variations à l'échelle individuelle et au niveau de l'élevage, la présence de bruit, de valeurs manquantes et d'artefacts de mesure rendent nécessaire la proposition d'approches méthodologiques nouvelles de modélisation et d'analyse prédictive de données massives.

Objectifs

Les objectifs du poste couvrent à la fois des aspects méthodologiques et d'ingénierie. Du point de vue méthodologique, il s'agira de proposer des approches méthodologiques innovantes permettant de modéliser au mieux la structure complexe des données analysées (données fonctionnelles, multiblocs, multiniveaux, incomplètes, etc) afin de prédire au mieux des indicateurs de santé de l'animal (glycémie, taux de lipides, pression artérielle, poids corporel, etc.). Du point de vue de l'ingénierie des données, il s'agira de développer et valider une approche d'ingestion, d'intégration et d'analyse prédictive de données provenant de capteurs ainsi qu'un prototype associé à déployer et à tester dans des conditions réelles.

La démarche attendue devra couvrir l'intégration et la synchronisation de données multi-capteurs, la qualification et la préparation de données, l'extraction de caractéristiques pertinentes, la modélisation de données fonctionnelles et hétérogènes via la proposition de méthodologies à l'interface entre la statistique, l'apprentissage automatique et l'intelligence artificielle.

Responsabilités du poste

- Conduire un état de l'art sur les approches récentes d'analyse des données de capteurs.
- Proposer une approche de modélisation, d'ingestion et d'analyse adaptée aux données du projet CAMELAIT4MED et aux hypothèses scientifiques formulées par le consortium.
- Mettre en place une chaîne d'intégration et d'analyse de données : couvrant ingestion, nettoyage, qualification, synchronisation, fusion de flux, extraction de variables et visualisation.

- Produire un prototype reproductible reposant sur des environnements modernes de développement et d'expérimentation, incluant scripts, notebooks, gestion de versions et documentation technique.
- Contribuer aux livrables, publications, réunions du consortium et échanges interdisciplinaires avec les partenaires scientifiques et techniques du projet.

Mission de recherche

La mission scientifique associée au poste consiste à produire une contribution méthodologique originale en matière de modélisation appliquée aux données du vivant instrumenté. Le ou la postdoctorant(e) devra interroger les choix de représentation des signaux, les faiblesses des méthodes d'apprentissage les plus pertinentes et identifier les moyens pour lever ces faiblesses en proposant des méthodologies innovantes. Les travaux pourront mobiliser, selon la problématique scientifique, des approches supervisées ou non supervisées. Une attention particulière sera portée au caractère interprétable des résultats obtenus.

Démarche et outillage attendus

Le poste vise à la proposition d'un cadre méthodologique et technique cohérent. L'outillage pourra s'appuyer sur les environnements R/Python et devront permettre la reproductibilité des analyses et résultats. Des outils de suivi d'expériences, de gestion de pipelines analytiques et de visualisation pourront également être mobilisés afin de garantir la qualité, la maintenabilité et la réutilisabilité des travaux produits.

Profil recherché

- Doctorat en science des données, intelligence artificielle, informatique ou diplôme équivalent.
- Compétences solides en analyse prédictive, modélisation statistique, apprentissage automatique et exploitation de données hétérogènes et massives, données fonctionnelles, données de capteurs.
- Maîtrise des technologies récentes en science de données et en IA, ainsi que des bonnes pratiques de programmation scientifique et de reproductibilité.
- Capacité à travailler dans un cadre interdisciplinaire et international, avec de bonnes aptitudes de rédaction scientifique en français et en anglais.

Candidature

Le dossier de candidature devra comprendre un curriculum vitae détaillé (avec un résumé des travaux de recherche en lien avec le poste et une liste représentative de publications ou de réalisations techniques), une lettre de motivation ainsi que des lettres de recommandation.

Les candidatures se font par mail en précisant l'objet du mail : Candidature postdoc projet CAMELAIT4MED

Faten Atigui, Faten.atigui@lecnam.net et Vincent Audigier, Vincent.audigier@lecnam.net