

**Type d'offre :** Offre en laboratoire

**Date de publication :** 04.03.25

**INRAE-LBE**

# **PhD offer - Multi-Criteria Optimization of Anaerobic Digestion Processes Using Artificial Intelligence**

## **Informations générales**

**Type de contrat :** CDD

**Durée du contrat :** 3 ans

**Contact :**

[Jean-Philippe Steyer](#) / [Rémi Servien](#)

**Date de prise de poste :** mer 01/10/2025 - 12:00

**Métier :** PhD

**Thématique :** Autre

## **INRAE-LBE :**

Le **Laboratoire de Biotechnologie de l'Environnement** (LBE) de l'INRAE est un institut de recherche de premier plan qui se consacre aux technologies environnementales innovantes. Nous nous concentrons sur le développement de solutions durables pour répondre aux défis environnementaux mondiaux les plus pressants. Notre équipe interdisciplinaire de scientifiques et d'ingénieurs travaille en collaboration pour faire progresser le domaine des biotechnologies environnementales, avec un accent particulier sur la récupération des ressources, le traitement des eaux usées et la production de bioénergie. Plus d'informations ici.

## **Détail de l'offre (poste, mission, profil) :**

### **Contexte général**

Les grands modèles de langage (LLM), illustrés par exemple par le GPT-4, ont trouvé diverses applications dans des domaines tels que la robotique, l'éducation et la médecine. Dans le domaine de la robotique, les robots intelligents utilisent les LLM pour comprendre leur environnement, facilitant ainsi l'interaction et la collaboration avec les humains (Open X, 2023). Actuellement, les LLM sont principalement utilisés pour contrôler les robots. En nous appuyant sur ce paradigme, nous proposons d'explorer une nouvelle application : l'utilisation d'un LLM pour optimiser l'alimentation des substrats dans la digestion anaérobie (DA) selon plusieurs critères qui sont difficiles à intégrer dans une approche classique (c.-à-d. en utilisant uniquement des modèles mécanistes). Les modèles LLM présentent plusieurs avantages par rapport aux méthodes de contrôle traditionnelles, ce qui justifie une exploration et un examen plus approfondis. Tout d'abord, ils permettent une certaine explicabilité : l'algorithme de contrôle résout la question derrière les sélections de substrat, permettant ainsi à l'opérateur de l'usine d'évaluer et de participer activement au processus de prise de décision. Ce dialogue ouvert entre l'opérateur et l'algorithme facilite ensuite l'amélioration itérative basée sur la compréhension mutuelle et la collaboration. Il permet également d'utiliser toutes les connaissances : en tirant parti de la richesse des informations documentées dans toute la littérature scientifique et technique sur la digestion anaérobie, les MFR peuvent être formés de manière efficace. Par conséquent, l'exploitant d'une seule installation est susceptible de bénéficier d'une expertise globale. Par exemple, les

LLM Open Source, illustrés par Llama 2 (Touvron et al., 2023), offrent une adaptabilité en permettant une formation locale sur les nouvelles données ou l'accès à la littérature sans nécessiter de formation supplémentaire. Enfin, en ce qui concerne l'interaction homme-machine, l'un des principaux avantages des LLM est leur capacité à dialoguer avec leurs opérateurs. Ce dialogue permet aux opérateurs d'ajuster dynamiquement les points de consigne, de signaler la disponibilité ou l'épuisement des substrats, de demander des éclaircissements sur les résultats des mesures et d'évaluer la stabilité de l'exploitation, ce qui est particulièrement important pour les agriculteurs qui n'ont pas forcément de compétences avancées dans ces domaines complexes aux objectifs divers et parfois contradictoires. Les premiers résultats obtenus par Gaida (2024) sur le sujet montrent que l'utilisation des LLM à cette fin est prometteuse, mais qu'un travail important d'étalonnage et de stabilisation des modèles est encore nécessaire avant de pouvoir les rendre utilisables à un niveau industriel.

---

## **Mission et objectifs du doctorat**

Les objectifs de ce doctorat sont de différentes natures. L'objectif général est de démontrer l'utilité des modèles d'IA, plus particulièrement des LLM, comme contrôleurs pour optimiser les procédés de digestion anaérobie. En utilisant en plus des modèles mécanistes de type bilan matière pour garantir une certaine robustesse à l'approche, le LLM pourra formuler des recommandations pour la gestion du processus, par exemple pour l'alimentation en substrat, en offrant des justifications transparentes et compréhensibles de ses décisions. Cela permettra finalement d'exploiter les modèles selon une approche multicritères (énergétique, économique et environnementale).

Cet objectif global sera divisé en différents sous-objectifs :

1. Une bibliographie sur les différentes implémentations des modèles LLM comparant leurs différents avantages et inconvénients ;
2. Une comparaison méthodologique sur les différents tests des modèles sélectionnés ;

3. Une application pour la gestion en temps réel des méthaniseurs au travers d'un algorithme implémenté.
- 

## **Exigences**

Nous recherchons un(e) boursier(e) très motivé(e) pour ce poste. Le candidat idéal doit posséder les qualifications suivantes :

- Un master ou un diplôme d'ingénieur à l'interface entre l'agriculture et la science des données ;
  - Une bonne connaissance de la programmation ;
  - Une maîtrise de la langue anglaise ;
  - Un intérêt pour les problèmes environnementaux serait un avantage.
- 

## **Conditions**

- Durée du contrat : 3 ans ;
  - Localisation : Unité LBE, Narbonne INRAE, France ;
  - Salaire brut : 2100€ (incluant la sécurité sociale) ;
  - Avantages scientifiques : Financement pour aller à différentes conférences ;
  - Autres avantages : 45 jours de congés par an, accès à la salle de sport et autres activités sportives ;
  - Encadrement : Jean-Philippe Steyer et Rémi Servien ;
  - Financement : La moitié du financement de la thèse a déjà été obtenue. La moitié du financement est demandée et une autre demande sera faite avec le candidat.
- 

## **Références**

1. *Open X-Embodiment Collaboration, Open X-Embodiment : Robotic Learning Datasets and RT-X Models, 2023, <https://doi.org/10.48550/arXiv.2310.08864>.*

2. Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., and Scialom, T. *Llama 2 : Openfoundation and fine-tuned chat models*, 2023, <https://doi.org/10.48550/arXiv.2307.09288>.
3. Gaida, D. *Synergizing language models and biogas plant control : A GPT-4 approach*, 2024. In : *Proceedings of the 18th IWA World Conference on Anaerobic Digestion*, <https://iwa-ad18.org/proceedings-book/>.

**URL de l'offre :** <https://www.dataia.eu/sites/default/files/phd2025.pdf>

**Lien vers l'offre sur le site dataia.eu :** <https://da-cor-dev.peppercube.org/node/1251>