

## Offre de POST-DOCTORAT en INFORMATIQUE – Intelligence artificielle

(english version below)

**Domaine** : IA / santé

**Titre** : Analyse de données de santé datées et incomplètes sur des patients âgés en vue d'évaluer un ensemble de facteurs de risques pour la santé

**Durée** : 17 mois

**Début** : septembre 2024

### Contexte

Ce projet est financé par le CPER Cornelia qui vise à poser les bases d'une IA responsable et durable. Les porteurs du projet sont Véronique Delcroix (UPHF, UMR CNRS 8201 LAMIH, Valenciennes) et Emmanuel Chazard (ULR 2694 METRICS, Université de Lille et CHU de Lille, CPER Tec'Santé).

Le/la postdoc intègre une équipe composée de membres du département informatique du LAMIH et de l'URL 2694 METRICS. Le LAMIH est un laboratoire pluridisciplinaire reconnu, traitant de questions centrées sur l'Humain, liées à l'informatique, l'automatique et la mécanique industrielles et Humaines. Des travaux sur un sujet connexe à celui proposé ont déjà été réalisés dans le département informatique du LAMIH. V. Delcroix est enseignante chercheuse en informatique dans le domaine de l'IA, sur la modélisation des connaissances et le raisonnement avec incertitude à l'aide de réseaux Bayésiens. L'URL 2694 METRICS est une équipe de recherche pluridisciplinaire en Santé Publique qui réunit des cliniciens et des méthodologistes. E. Chazard est professeur des universités et praticien hospitalier à l'Université de Lille et au CHU de Lille. Il est coordonnateur de l'axe 3 de METRICS, directeur adjoint de la SFR-TSM, et co-responsable de la plateforme e-santé du CPER Tec'Santé, pour la génération de données structurées de santé fictives.

En collaboration avec l'équipe projet, le/la postdoc travaillera sur un projet en IA pour la santé à partir de données de santé.

### Sujet

Vous participerez à la proposition de méthodes et modèles pour exploiter des données de santé provenant de cabinets de médecine générale dans le but d'améliorer le suivi de santé des personnes âgées (prévention, diagnostic, anticipation des risques, ...). Ces données, recueillies sur plusieurs années ou dizaines d'années pour un même patient, présentent des caractéristiques particulières qui doivent être prises en compte pour être efficacement exploitées.

Vos missions seront les suivantes :

- Etat de l'art
  - o Dynamique temporelle des variables
  - o Données essentiellement manquantes, hétérogènes, déséquilibrées
  - o Modèles graphiques probabilistes avec une dimension temporelle,
  - o Appropriation des travaux antérieurs
- Compréhension des données médicales des cabinets de groupes fournies par l'URL METRICS
- Proposition de modèles et méthodes
  - o Définition de classes de variables suivant leur persistance dans le temps.
  - o Identification des variables cibles et sélection des données
  - o Proposition de modèles et algorithmes pour exploiter des données de santé datées
  - o Développement de programmes pour tester et évaluer les solutions proposées
- Rédaction d'au moins un article scientifique pour un journal international.
- Présentation d'au moins un séminaire en interne sur les travaux menés

**Profil du candidat**

Le candidat doit avoir une thèse récente dans le domaine de l'intelligence artificielle, avec si possible une connaissance des modèles graphiques probabilistes, et une expérience de travail sur des données de santé. Un attrait pour la modélisation des connaissances, une bonne capacité d'abstraction et des compétences en développement sont nécessaires. Le candidat devra être autonome.

Il s'agit d'un travail multidisciplinaire dans lequel de bonnes capacités de communication sont indispensables, en particulier pour acquérir une connaissance suffisante du domaine d'application et des variables. Le Français n'est pas exigé mais facilite les échanges concernant la compréhension des données de santé. L'Anglais est requis (lecture, écriture essentiellement).

**Adresse d'emploi :** LAMIH, UPHF, Le Mont Houy, Valenciennes, 59 300, France

**Salaire net :** 4300 euros brut mensuel

**Contacts :**

- Véronique Delcroix : [Veronique.Delcroix@uphf.fr](mailto:Veronique.Delcroix@uphf.fr), tel : 03 27 51 14 65 ou 06 03 62 61 53
- Emmanuel Chazard : [Emmanuel.Chazard@univ-lille.fr](mailto:Emmanuel.Chazard@univ-lille.fr)

## POST-DOCTORATE position in Artificial intelligence

**Field:** AI / health

**Title:** Analysis of dated and incomplete health data on elderly patients to assess a set of health risk factors.

**Duration:** 17 months

**Start date:** September 2024

### Context

This project is funded by the CPER Cornelia program, which aims to lay the foundations for responsible, sustainable AI. The project leaders are Véronique Delcroix (UPHF, UMR CNRS 8201 LAMIH, Valenciennes) and Emmanuel Chazard (ULR 2694 METRICS, Université de Lille and Lille's hospital, CPER Tec'Santé).

The postdoc will join a team regrouping members of the LAMIH Computer Science department and the ULR 2694 METRICS. The LAMIH is a recognized multidisciplinary laboratory, dealing with human-centered issues related to computer science, automation and industrial and human mechanics. Work on a subject related to that proposed has already been carried out in the LAMIH computer science department. V. Delcroix is a researcher in computer science in the field of AI, on knowledge modeling and reasoning with uncertainty using Bayesian networks. The ULR 2694 METRICS is a multidisciplinary public health research team that brings together clinicians and methodologists. E. Chazard is a university professor and hospital practitioner at the University of Lille and Lille University Hospital. He is coordinator of Axis 3 of METRICS, co-director of SFR-TSM, and co-leader of the e-health platform of CPER Tec'Santé, for structured health data generation.

In collaboration with the project team, the postdoc will work on an AI project based on health data.

### Subject

You will be involved in proposing methods and models for exploiting health data from general practitioners, with the aim of improving health monitoring of the elderly (prevention, diagnosis, risk anticipation, etc.). These data, collected over several years or decades for the same patient, present specific characteristics that need to be taken into account to be effectively exploited.

Your tasks will include

- State of the art about
  - o Temporal dynamics of variables
  - o Essentially missing, heterogeneous, imbalanced data
  - o Probabilistic graphical models with a temporal dimension,
  - o Appropriation of previous work
- Understanding group practice medical data provided by the METRICS URL
- Proposition of models and methods
  - o Definition of variable classes according to their persistence over time,
  - o Identification of target variables and data selection,
  - o Proposal of models and algorithms to exploit dated health data,
  - o Development of programs to test and evaluate proposed solutions.
- Writing of at least one scientific article for an international journal.
- Presentation of at least one in-house seminar on the work carried out.

### **Candidate profile**

The candidate should have a recent thesis in the field of artificial intelligence, if possible with knowledge of probabilistic graphical models, and experience of working with healthcare data. An interest in knowledge modeling, a good capacity for abstraction and development skills are required. The candidate must be autonomous.

This is a multidisciplinary job in which good communication skills are essential, in particular to acquire sufficient knowledge of the application domain and variables. French is not required, but will facilitate exchanges concerning the understanding of health data. English is required (mainly reading and writing).

**Job address:** LAMIH, UPHF, Le Mont Houy, Valenciennes, 59 300, France

**Net salary:** 4300 euros gross per month

### **Contacts :**

- Véronique Delcroix : [Veronique.Delcroix@uphf.fr](mailto:Veronique.Delcroix@uphf.fr) tel : 03 27 51 14 65 or 06 03 62 61 53
- Emmanuel Chazard : [Emmanuel.Chazard@univ-lille.fr](mailto:Emmanuel.Chazard@univ-lille.fr)