M2 funded internship position

Structured Extraction of Biomedical Data for AI-Guided Gene Signature Interpretation in Precision Oncology

Context:

This M2 internship is the first step in a project developing an automated AI-driven biomedical knowledge system for context-specific interpretation of gene lists leveraging generic biological ontologies and disease-specific scientific literature. This AI-based system will be powered by an innovative Retrieval-Augmented Generation (RAG) architecture with hierarchical associative memories (Hopfield models) developed at CEA.

The specific objective of this internship is to develop a robust pipeline for extracting and structuring knowledge from scientific publications in PDF format. We will leverage advanced parsing tools (MistralOCR), structured extraction techniques based on Large Language Models (Claude by Anthropic), and AI agent frameworks (LangChain) to generate structured data linking genes, biological ontologies, and treatment responses while preserving traceability to sources.

The system will be validated on a concrete use case: analyzing transcriptomic response signatures in clear cell renal carcinoma (ccRCC) patients, focusing on literature concerning biological mechanisms of resistance to anti-PD-1 immunotherapies.

Profile:

- Master's level in computer science, bioinformatics, or related field
- Strong Python skills (experience with LangChain, API integration preferred)
- Interest in NLP and large language models
- Background in biology/bioinformatics is a plus

Supervision and working environment:

The intern will work at CEA-Grenoble (BGE laboratory, UA13 INSERM/CEA/UGA) under triple supervision: **Guido Uguzzoni** (AI methods for molecular biology), **Christophe Battail** (pharmacogenomics), and **Yoann Curé** (RAG-Hopfield architecture). This interdisciplinary environment offers unique training at the intersection of machine learning, systems biology, and precision medicine.

Location: Genetics & Chemogenomics team, BGE lab, IRIG, CEA-Grenoble, Building C3, 17 avenue des Martyrs, 38054 Grenoble Cedex 9

Duration: 6 months (March - August 2026)
Salary: ~650-700€ net/month (EFELIA/MIAI funding)

How to apply:

Send CV, cover letter, transcripts, and the name of one referee to Guido Uguzzoni (guido.uguzzoni@cea.fr) and Christophe Battail (christophe.battail@cea.fr) before **January 15, 2025**