Internship Offer – RAG Infrastructure & Knowledge Graph Engineering (Master Level)

Position: Research & Development Internship – RAG Infrastructure & Knowledge Graph
Level: Master's Degree (BAC +5)
Duration: 4–6 months (flexible)
Location: Montpellier – HSM Hydrosciences Montpellier
Start Date: March 2026

Context & Objectives

The objective of this internship is to explore and implement an advanced Retrieval-Augmented Generation (RAG) infrastructure leveraging a Large Language Model (LLM) and a domain-specific knowledge graph. The knowledge graph will be based on the SEWERNET ontology (http://sewernet.msem.univ-montp2.fr/), commonly used for wastewater and sewer network modeling. The project will involve building the knowledge graph from existing datasets and evaluating how semantic structuring improves performance within a RAG pipeline.

The objective is to develop a system for interacting with data in natural language and knowledge graphs to obtain the best results, and to identify and propose the best methodologies for this field.

Main Missions

The intern will contribute to the following activities:

- Study and characterization of RAG architectures combining LLMs and semantic knowledge bases.
- Implementation of a prototype RAG pipeline integrating:
  - a Large Language Model
  - a vector retrieval system
  - a SEWERNET-based knowledge graph
- Construction of a knowledge graph from existing test datasets, including:
  - data ingestion
  - ontology mapping
  - validation & enrichment
- Performance benchmarking of different retrieval strategies (vector search, graph reasoning, hybrid methods).
- Proposal of improvements based on experimental results.

Expected Deliverables

- A functional prototype demonstrating the RAG + Knowledge Graph integration.
- A SEWERNET-aligned knowledge graph produced from test datasets.
- Performance analysis and experimental benchmark report.
- Presentation of findings and recommendations for future development.

Required Skills

- Master's degree level (BAC +5) in Computer Science, AI, Knowledge Engineering, Semantic Web, or related fields.
- Solid knowledge of :
  - Machine Learning or NLP
  - Ontologies / Knowledge Graphs / Semantic Web technologies (RDF, OWL, SPARQL)
  - Python development
- Familiarity with LLM tooling and RAG frameworks is a plus (LangChain, LlamaIndex, etc.).
- Strong analytical and communication skills.

Nice-to-Have Skills

- Experience with graph databases (Neo4j, GraphDB, Stardog, BlazeGraph, etc.)
- Exposure to geospatial utilities or wastewater infrastructure domains.

What We Offer

- A research-oriented internship combining semantic technologies and generative AI.
- Guidance and supervision from experts in AI and knowledge engineering.
- Opportunity to contribute to emerging industrial applications of LLMs.

Laboratoire d'accueil

HydroSciences Montpellier (HSM), HYDROPOLIS, 15 avenue Charles Flahault, F-34090 Montpellier

Duration

6 months between March and August 2026. Compensation based on the current hourly rate.

Contacts :

stephane.debard_at_ird.fr
nanee_chahinian_at_ird.fr
francoalberto.cardillo_at_cnr.it
franca.debole_at_isti.cnr.it