Thursday, April 30, 2026

TReC 2026 Project Proposal Submission Form

Submit your project proposal for the 7th TRAIL Research Camp (August 24th - September 4th, 2026, Lausanne, Switzerland). Please complete all required sections and submit your proposal before April 30th, 01:00 PM (CET).

Administrative Data

Full Name of Team Leader Xavier Lessage

Contact Email xavier.lessage@cetic.be

Project Information

Project Title FED-SORAG: Federated Sovereign Retrieval-Augmented Generation for Secure Knowledge Systems

Profile of the Team Leader(s) & Expected Team Composition

Xavier is an expert research engineer in CETIC's Data Science department. His work focuses mainly on artificial intelligence, cloud computing and distributed data processing. He is particularly interested in the synergies between AI and cybersecurity, notably through advanced approaches such as federated learning and homomorphic encryption. He actively collaborates with several universities and research centers, contributing to both academic research and applied innovation in distributed and secure AI systems. Several publications about Federated learning are available in collaboration with universities and researchers.
Xavier has shown his ability to supervise projects in the 5 past TRAIL Summer Workshops. Xavier federated a very efficient approach to foster the teams of researchers involved in publishing scientific publications and technical reports as well as producing results released in the TRAIL Factory. Furthermore, results from these past workshops raised the interest of Walloon businesses and various industry-driven use cases in Health and Industry domains are explored in on-going collaborative projects.

Looking for researchers with the following profiles:

- AI / ML PhD student with foundations in NLP, LLMs, or RAG systems, ideally with interest in federated learning
- Distributed systems or cybersecurity PhD student with interest in privacy-preserving architectures and secure communications
- Medical informatics or digital health PhD student familiar with interoperability standards (FHIR) or clinical datasets
- SSH PhD student (ethics, law, social sciences) sensitive to data sovereignty and trustworthy AI issues

Have you already identified potential team members for your project? Yes

List the team members you have identified and briefly describe their profiles/roles (e.g., expertise, affiliation, expected contribution).

Researchers from ULIEGE, UMONS, CETIC, ...

Domain of Application Healthcare

Scientific Theme Privacy-preserving AI

Proposal Content

Abstract

This work investigates a Federated Sovereign Retrieval-Augmented Generation (FED-SORAG) architecture that combines federated learning principles with RAG systems to enable secure access to distributed and sensitive knowledge bases without centralizing data. The goal is to allow large language models to query decentralized sources (e.g., healthcare, industry, public institutions) while ensuring that raw data always remains within local infrastructures.

The project is motivated by recurring challenges observed across several federated and data space initiatives: enabling effective querying over distributed systems while preserving sovereignty, privacy, and interoperability. A central difficulty lies in handling heterogeneous knowledge sources, including unstructured text, structured records, and multimodal data. This problem can be broken down into sub-problems such as data modality alignment, storage heterogeneity, indexing strategies, and federated query optimization.

FED-SORAG is structured as a federation of local RAG systems coordinated through a global orchestration layer that aggregates results across nodes. Internal knowledge is prioritized, while external sources are used only when they add clear value. This raises additional requirements for secure external access, including query sanitization, validation of retrieved content, and strict provenance tracking to distinguish internal from external information.

Key technical challenges include scalable distributed indexing, query orchestration, latency management, and robustness against adversarial threats such as prompt injection. The architecture also explores confidentiality mechanisms and emphasizes transparency through traceable outputs and confidence indicators.

For evaluation, the project uses a federated simulation of clinical data transformed into an interoperable format, enabling realistic distributed healthcare scenarios with strict data locality. External biomedical knowledge can be selectively integrated in a controlled manner.

The expected outcomes are a proof-of-concept, a formal architecture, and an initial evaluation of performance and security trade-offs. Overall, FED-SORAG aims to provide a sovereign, trustworthy alternative to centralized AI systems for sensitive domains.

Background Information & Problem Statement

Retrieval-Augmented Generation (RAG) systems have significantly improved the reliability and domain relevance of large language models by grounding their outputs in external data. However, most existing RAG approaches rely on centralized architectures, requiring data to be aggregated or accessed through external services. This creates major limitations when dealing with sensitive, regulated, or proprietary data, particularly in sectors such as healthcare, finance, and public administration, where data cannot leave local infrastructures due to privacy, legal, or sovereignty constraints. Beyond data locality, real-world deployments also face the difficulty of leveraging knowledge bases that are large and heterogeneous in nature, combining unstructured documents, structured records, and multimodal content, which complicates indexing, retrieval, and consistent reasoning.

Federated learning enables collaborative model training without sharing raw data, but it does not address the challenge of real-time knowledge retrieval and reasoning across distributed sources. As a result, there is currently a gap between privacy-preserving learning approaches and practical, secure knowledge access systems.It also considers how external knowledge sources, such as public repositories or content from the internet, can be integrated in a controlled and secure manner, so that they enrich generation when relevant without compromising the confidentiality of internal data or the integrity of retrieved content.

This project addresses this gap by proposing a Federated Sovereign Retrieval-Augmented Generation (FED-SORAG) framework. The goal is to enable language models to query and reason over decentralized knowledge bases through a federation of local RAG systems, each operating within its own secure environment. This approach preserves data locality while enabling cross-organizational knowledge access. To bridge this gap in a realistic setting, we leverage healthcare as a representative high-stakes domain. Specifically, we build upon an open-source conversion of the MIMIC Database into FHIR, enabling the creation of interoperable yet distributed clinical data silos. This setup reflects real-world constraints where institutions maintain control over sensitive patient data while still requiring cross-organizational knowledge access. However, this paradigm introduces key challenges, including distributed orchestration, latency, scalability, the management of large and heterogeneous knowledge sources, and security risks such as prompt injection or untrusted coordination layers. It also raises important questions about reliability, explainability, and trust.

Positioned at the intersection of federated learning, information retrieval, and trustworthy AI, this project aims to design a secure, scalable, and transparent architecture for accessing sensitive distributed knowledge, enabling the next generation of sovereign AI systems.

Project Objectives & Concrete Implementation

The goal of this 2-week camp is to build a **working proof-of-concept of FED-SORAG**, a federated and sovereign Retrieval-Augmented Generation system enabling secure cross-organization knowledge access without data centralization.

We will implement a **multi-node federated RAG architecture** composed of local RAG systems (each hosting its own data and retriever) and a central orchestration layer that routes queries, aggregates responses, and ensures that raw data never leaves local environments. Each node will independently retrieve and generate answers from its private corpus, while the orchestrator combines results into a unified response. The architecture will also account for the fact that local corpora may be large and heterogeneous, and that external knowledge sources (e.g., public repositories or web content) may need to be consulted in a controlled way when internal knowledge is insufficient.

On the implementation side, based on the framework flower (https://flower.ai/, https://flower.ai/docs/examples/fedrag.html), we will develop:

- A set of **local RAG nodes** with document ingestion, embedding-based retrieval, and LLM-based generation
- A **federated orchestration layer** for query routing and response fusion
- A **secure communication interface** between nodes and orchestrator
- A basic **response aggregation mechanism** (ranking or synthesis of multi-node outputs)
- A controlled gateway for accessing selected external knowledge sources, including basic egress filtering and provenance tagging of retrieved content

- An **integrated federated model distribution update and monitoring system** (Fed MLOps)

We will also integrate initial **robustness features**, including simple defenses against prompt injection and mechanisms to filter or validate external inputs.

The system will be evaluated on a realistic distributed healthcare dataset derived from an open-source conversion of the MIMIC Database into FHIR. This dataset will be partitioned into multiple simulated institutional silos (e.g., hospitals or departments), each hosting its own local RAG node. This setup enables controlled experimentation on cross-silo query answering, interoperability, and privacy-preserving access to structured clinical data.

By the end of the workshop, we will deliver:

- An open source conversion of a MIMIC-based dataset in FHIR
- A working **FED-SORAG prototype**
- A toolchain **for the deployment, update and monitoring of federated models**
- A **demonstration of federated question answering across data silos**
- A short **technical report with architecture and evaluation results**

This work will provide a first practical validation of sovereign, federated RAG systems for secure and distributed knowledge access.

Do you plan to deliver, as an outcome of your project, a reusable "brick" for the TRAIL Factory (https://factory.trail.ac/en/home_page) that could later be transferred and converted into a company process?

Yes

Briefly describe what the brick would be and its intended users.

The brick is a generic federated AI component enhanced with Retrieval-Augmented Generation (RAG), designed to enable privacy-preserving data analysis and knowledge extraction across distributed datasets. It is developed and validated using the MIMIC dataset (converted into the FHIR format), while remaining adaptable to other datasets and domains. By leveraging advanced federated AI architectures, it enforces ethical principles, complies with legal privacy frameworks, and preserves organizational data sovereignty. The intended users include researchers, data scientists, healthcare professionals, and organizations across various sectors such as energy, education, and legal.

Project Dataset

The project will use a realistic, domain-specific dataset based on an open-source conversion of the MIMIC Database into the FHIR format. The dataset will be:

- Structured as interoperable healthcare records (patients, encounters, observations, medications)
- Partitioned into multiple independent silos to simulate distinct healthcare institutions
- Composed of heterogeneous content where possible, combining structured FHIR resources with

unstructured clinical notes and, where relevant, other modalities, so as to reflect the diversity typically found in real-world institutional knowledge bases
- Extended with synthetic or derived annotations where needed for retrieval and evaluation tasks
- Optionally complemented by a limited set of external biomedical references (e.g., excerpts from public guidelines or open biomedical literature) to support controlled experimentation on the integration of outside knowledge alongside sovereign internal data

This approach enables the development and testing of federated RAG in a realistic healthcare setting while remaining fully compliant with open-data usage constraints.

This project is designed to be generic and applicable across a wide range of use cases, including industries such as energy, education, legal and more.

Detailed Work Plan

Week 1: Architecture & Core System

- Prepare and deploy simulated healthcare data silos using FHIR-formatted data derived from the MIMIC dataset
- Define FED-SORAG architecture and interfaces
- Set up environments, repositories, and datasets
- Implement local RAG nodes (retrieval + generation)
- Add text-to-structured and structured-to-test conversion (text→ FHIR queries and FHIR resources →usable prompts)
- Deploy multiple simulated data silos
- Build federated orchestration layer (query routing)
- Set up a basic external-knowledge gateway allowing controlled retrieval from selected outside sources (e.g., a limited set of public biomedical references), with provenance tagging of retrieved content
- MLOps/DevOps integration

Goal: Working baseline system with multi-node federated RAG.

Week 2: Security, Evaluation & Optimization

- Integrate security layer (prompt injection defenses, sanitization)
- Test robustness against adversarial and noisy inputs
- Optimize latency and retrieval performance
- Implement retrieval strategies tailored to FHIR resources (Patient, Observation, MedicationRequest, etc.), and complementary strategies for unstructured or mixed content within the same node
- Implement response aggregation and explainability (sources, confidence)
- Run evaluation (quality, latency, robustness), including a basic scenario where external knowledge is consulted alongside sovereign internal data

Goal: A secure, optimized, and evaluated FED-SORAG prototype enabling federated question answering over distributed FHIR-based clinical data silos.

Bibliographic References

- Stripelis et al.,(2026), Supercharging Federated Intelligence Retrieval, **Flower Labs**, https://arxiv.org/abs/2603.25374
- Jung et al., Federated Learning and RAG Integration: A Scalable Approach for Medical Large Language Models. arXiv preprint arXiv:2412.13720 (2024).
- Addison et al., C-FedRAG: A Confidential Federated Retrieval-Augmented Generation System. arXiv preprint arXiv:2412.13163 (2024).
- Zhao, Dongfang. FRAG: Toward Federated Vector Database Management for Collaborative and Secure Retrieval-Augmented Generation. arXiv preprint arXiv:2410.13272 (2024).
- Xiong et al., “Benchmarking retrieval-augmented generation for medicine.” In Findings of the Association for Computational Linguistics ACL 2024, pp. 6233-6251. 2024.
- Cormack, Gordon V., Charles LA Clarke, and Stefan Buettcher. “Reciprocal rank fusion outperforms condorcet and individual rank learning methods.” In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pp. 758-759. 2009.
- Mao, Q. et al. (2025), FedE4RAG: Privacy-Preserving Federated Embedding Learning for Localized Retrieval-Augmented Generation. https://arxiv.org/abs/2504.19101
- P. et al. (2024), C-FedRAG: A Confidential Federated Retrieval-Augmented Generation System. NVIDIA/Deloitte,Addison, https://arxiv.org/pdf/2504.19101
- Chakraborty, A. et al. (2025), Federated Retrieval-Augmented Generation: A Systematic Mapping Study. Arizona State University, https://arxiv.org/abs/2505.18906
- Gao, Y. et al. (2023), Retrieval-Augmented Generation for Large Language Models: A Survey, https://arxiv.org/abs/2312.10997
- Chen Zhang et al. (2022), A survey on federated learning: Knowledge-Based Systems. https://www.sciencedirect.com/science/article/abs/pii/S0950705121000381
- Yifan Yao et al. (2024), A survey on large language model (LLM) security and privacy. https://www.sciencedirect.com/science/article/pii/S266729522400014X

Eligibility & Evaluation

Does the project include multidisciplinarity between STEM & SSH?

Yes

How?

Yes, the project is clearly multidisciplinary, combining both STEM and SSH approaches.

From the STEM perspective, it relies on advanced AI techniques, including federated learning and Retrieval-Augmented Generation (RAG), to enable analysis and exploitation of distributed data while preserving privacy. The project is developed and validated using the MIMIC dataset (converted into the FHIR format), while remaining adaptable to other datasets and domains.

From the SSH perspective, it integrates ethical, legal, and societal considerations such as data protection, data sovereignty, GDPR compliance, and responsible AI principles (transparency, fairness, and accountability).

The link between both domains is ensured by the federated architecture and RAG, which embed these SSH constraints directly into the technical system design. The project therefore combines AI engineering and governance challenges for applications across sectors such as healthcare, energy, education, and legal domains.

We confirm that the Team Leader will be present for the full duration of TReC'26 if the project is selected (August 24th - September 4th, 2026, Lausanne, Switzerland)

I/We agree and confirm