# Integration Engine for Efficient Clinical Data Exchange

## Summary

*Healthcare interoperability hinges on seamless data exchange between diverse systems, yet existing "unified" approaches—where every system must support a single standard—often falter with legacy platforms and slow innovation. Although protocols like FHIR and tools such as NextGen Connect, FML and Whistle address parts of the problem, the commercialization of NextGen has left a gap in open-source integration solutions. This doctoral research will extend the TermX platform—originally developed at TalTech—to embrace a "federated" model, dynamically adapting data to each system's needs rather than forcing a one-size-fits-all format. By enhancing TermX's modelling, transformation and runtime capabilities, and abstracting the results into a formal interoperability theory grounded in MOF and OSI principles, the work aims both to deliver a practical, lightweight integration engine and to underpin it with a robust theoretical framework.*

| | |
|---|---|
| Research field: | Information and communication technology |
| Supervisor: | Gunnar Piho |
| Availability: | This position is available. |
| Offered by: | School of Information Technologies<br>Department of Software Science |
| Application deadline: | Applications are accepted between June 01, 2025 00:00 and June 30, 2025 23:59 (Europe/Zurich) |

## Description

*The research*

Interoperability and data transformation constitute a highly topical issue in healthcare. Efficient data exchange between different information systems and standards holds substantial added value, as data accessibility is crucial in clinical applications, research, and administrative systems. Currently, relevant technologies in this domain include the Fast Healthcare Interoperability Resources (FHIR) protocol, which facilitates data exchange among clinical systems. Additionally, various health data transformation languages (such as FML and Whistle) and integration engines (like NextGen Connect and Iguana) exist.

Following the commercialization of NextGen (formerly Mirth) Connect, there is a notable gap in the market for widely recognized open-source integration platforms. Existing interoperability solutions in healthcare predominantly employ a unified interoperability approach, proposing standards and shared data formats that all clinical systems must support. However, this approach poses significant challenges, particularly in terms of managing legacy systems. Furthermore, mandating standardized data formats may hinder innovation, as maintaining extensive and broadly adopted standards is resource-intensive and slow. In contrast to unified interoperability, federated interoperability advocates dynamically adapting data to the systems in which they currently reside.

The goal of this doctoral research is to investigate integration solutions and develop an open-source platform to fill the market gap resulting from NextGen Connect's commercialization. The novelty of this integration solution lies in its support for the federated interoperability paradigm. Practically, this doctoral work will advance the existing TermX platform. TermX, a health data interoperability platform developed as a doctoral thesis at TalTech, consists of multiple modules at various Technology Readiness Levels (TRL), including a data modeling design tool (TRL 4), transformation management modules (TRL 5), and a terminology server (TRL 7).

The scope of this doctoral research will encompass: Data modelling, a transformation creation tool, a transformation engine capable of effectively executing transformations and establishing integration channels for real-time data exchange. The study of integration engines will specifically include: Identification of data exchange channels and protocols (HTTP, TCP, MLLP, S3, x-Tee, etc.), creation of visual transformations using at least two languages (FML and Whistle), mechanisms for data transformation quality control, logging, monitoring, and high-availability considerations associated with data transformation and transmission, validation of transformed data through testing.

For data transformation in production environments, it is essential to enable the creation or publication of isolated modules. Specifically, a transformation module should consist of multiple adapters capable of being installed sep-

arately. For instance, CDA→FHIR and FHIR→CDA would be distinct adapters. Adapter types would also include various laboratory analyzer interfaces that convert laboratory outputs into widely adopted standards like HL7 V2 or FHIR. These adapters should be lightweight and capable of executing transformations on standard-alone computers or even devices like Raspberry Pi.

*Responsibilities and (foreseen) tasks*

- Design, develop and evaluate a tool according to the Design Science methodology
- The results and practical outcomes of this research are intended to be generalized into theory. Using Meta-Object Facility (MOF) and Open Systems Interconnection (OSI) as foundational frameworks, this generalization will offer a formal theory for generated interoperability.
- Collaborate with clinicians and healthcare IT experts to gather qualitative feedback and assess the solution's applicability in clinical workflows
- Participate in research and development seminars

Publish scientific articles in international journals and present at conferences on NLP and medical AI topics.

*Applicants should fulfil the following requirements:*

- Master's degree in information and communications technology
- A clear interest in the position's topic
- Excellent command of English
- Strong, demonstrable writing and analytical skills
- Ability to work both independently and as part of an international team
- Willingness and capacity to assist with organisational tasks relevant to the project

*The following experience is beneficial:*

- Programming
- Working knowledge of SQL
- Working knowledge of software engineering
- Working knowledge of healthcare standards

The candidate should submit a research plan on the chosen topic, detailing the overall studies, research and publication strategy. They may expand on the listed research questions and tasks and propose the theoretical foundations to be employed.

*We offer:*

- A fully funded, 4-year PhD position at eMedLab (https://taltech.ee/en/emedlab), TalTech's interdisciplinary centre for digital health innovation.
- Access to state-of-the-art research infrastructure and collaborative opportunities within the European Federation of Medical Informatics (https://efmi.org/).
- Opportunities for conference travel, international research stays, and networking with leading universities.

*About eMedLab*

eMedLab is a research group in digital health and medical informatics at Tallinn University of Technology (TalTech). It brings together researchers from the Centre for Digital Health and Business Information Technology research groups of the School of Information Technologies. The eMedLab research group, led by Professor Peeter Ross and Dr Gunnar Piho, consists of around twenty researchers exploring the opportunities and challenges of using human-owned and -controlled health and medical data and developing new methods and technologies to address these challenges.

*Additional information*

For further information, please contact Dr Gunnar Piho gunnar.piho@taltech.ee or visit https://taltech.ee/en/emedlab

To get more information or to apply online, visit https://taltech.glowbase.com/positions/1034 or scan the the code on the left with your smartphone.