RDTI General Approval Application (Draft) - v1.61

Project Details
Project Identifier: PRJ-002

Objective (limit 1000 chars): Develop a secure, decentralized platform utilizing Federated Learning (FL) to train robust diagnostic AI models across multiple hospital data silos without sharing raw patient data. This aims to overcome data privacy limitations and improve the accuracy of early disease detection. [Intentional brevity to respect the 1000 character limit.]

Start Date: 2026-03-01

End Date: 2027-11-30

ANZSRC: 080399

Project Owner (technical contact)
Name: Anna Rodriguez

Role: Director of R&D

Email: anna.rodriguez@company.com

Phone: +64 021-999-0000 (Mobile)

Core Activity 1
Name: Design and Implementation of Secure Aggregation Protocol

Activity dates: 2026-04-01 -> 2026-10-31

Uncertainty (~500 words): The scientific uncertainty is whether a homomorphic encryption scheme can be practically applied to aggregate high-dimensional model parameters (like those from a deep neural network) without introducing unacceptable latency or computational load that would render the Federated Learning process infeasible in a real-time clinical environment. Existing aggregation methods often compromise on either privacy or efficiency.

Systematic approach (~250 words): A systematic approach involves: 1) Benchmarking three different homomorphic encryption libraries. 2) Developing a tailored, optimized aggregation function for the selected library. 3) Implementing a prototype on a simulated network of ten nodes. 4) Rigorously measuring key metrics: aggregation time, bandwidth usage, and proof-of-correctness error rate across various network conditions.

Intentions (~200 words): This activity intends to create New Knowledge in cryptographic engineering by proving the feasibility and efficiency of homomorphic encryption in high-scale FL for medical imaging. It will also create a New Process—the highly secure and optimized aggregation method—for distributed AI model training.

Core Activity 2
Name: Decentralized Fault Tolerance and Model Drift Mitigation

Activity dates: 2026-11-01 -> 2027-05-31

Uncertainty (~500 words): The technological uncertainty lies in effectively designing a decentralized mechanism to identify and isolate malicious or significantly inaccurate local model updates (e.g., from 'poisoned' data) without relying on a central authority or trusted party, a challenge amplified by the high data heterogeneity common across different hospital systems.

Supporting Activity 1
Name: Data Preparation and Pre-processing Standards Definition

Activity dates: 2026-03-01 -> 2026-04-30

Description (~250 words): Establishing standardized protocols and tools for cleaning, normalizing, and anonymizing medical records (DICOM/EHR data) to ensure uniform data quality across all participating hospital nodes before model training can commence.

Definition (~250 words): This satisfies the Supporting R&D activity definition because it involves the systematic preparation and standardization of materials (data) which is essential for conducting the Core R&D activities. Without this uniform pre-processing, the model updates would be inconsistent, invalidating the entire federated training experiment.

Supporting Activity 2
Name: Acquisition and Configuration of Cloud Computing Resources

Activity dates: 2026-03-15 -> 2026-05-15

Description (~250 words): Procurement and setup of necessary specialized computing resources (GPUs, secure cloud infrastructure) and configuring the virtual network environment required to simulate the high-scale decentralized architecture for development and testing.

Definition (~250 words): This satisfies the Supporting R&D activity definition as it is an activity in support of R&D that involves the necessary technical and logistical preparation of the specialized tools (hardware/software infrastructure) required to execute the intensive model development and testing defined in the Core R&D activities.

Supporting Activity 3
Name: Ethical and Regulatory Compliance Review

Activity dates: 2026-03-01 -> 2027-11-30

Description (~250 words): Continuous consultation with legal and ethical experts to ensure the platform design and aggregation protocols comply with international data privacy regulations (\text{GDPR}, \text{HIPAA}$) specific to the target deployment countries.

Definition (~250 words): This satisfies the Supporting R&D activity definition because it is work undertaken for the purpose of, or in support of, R&D. Navigating the legal and ethical constraints (which influence cryptographic choices and data handling) is fundamental to designing a technologically and commercially viable solution for the healthcare sector.