

# Build AI Foundation Models and APIs for Breast Cancer Early Detection

## Objectives

The main goal of the project is to build a set of AI foundation models that are trained on large data and can be adapted to many downstream tasks in healthcare. The specific objectives include 1) pretraining AI foundation models (i.e., visual and language models), 2) building application programming interfaces to enable open access, and 3) demonstrating the effectiveness by applying to breast cancer early detection.

## Background

An AI foundation model in healthcare is a machine learning model that is trained using big healthcare datasets, and could be adapted to many downstream tasks. Early examples of foundation models consist of pre-trained language models and image processing models, e.g., OpenAI's "GPT-n" series, Meta AI's Segment Anything, and Google DeepMind's GATO. These models can be used for domain-specific tasks using different targeted datasets, such as cancer detection.

Breast cancer has become one of the most common cancers worldwide, accounting for approximately 12% of all new cancer cases. In the U.S., it is estimated that breast cancer affects 30% of all new female cancer cases every year. Early detection of breast cancer can significantly reduce mortality and expand treatment options. Recently, deep learning-based approaches for breast cancer detection have met or exceeded human expert performance levels.

In the last six years, the U of I's [Machine Intelligence and Data Analytics \(MIDA\)](#) Lab has been focused on developing groundbreaking AI/ML methodologies that enable future dependable and trustworthy technologies for healthcare. The Lab published the world's largest publicly available breast ultrasound benchmark, developed a series of deep multitask learning approaches, built novel robust self-training approaches to improve the robustness of AI models, proposed new semantic and instance segmentation approaches, and developed generative AI models for image synthesis.

## Hardware and Environment

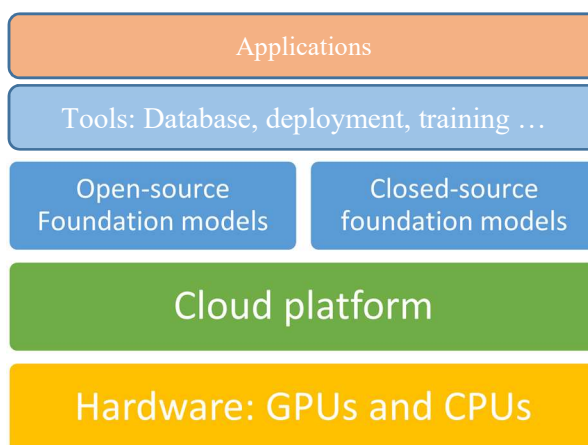
The team will get access to MIDA Lab's two high-performance GPU servers and INL's Falcon supercomputer.

## Customer

U of I's Machine Intelligence and Data analytics (MIDA) Lab, Dr. Min Xian

## Expected team

Three students.



AI Foundation Model Stack