Title : Agentic Rapid Analysis of Bioinformatics Utility System for Microbial Genomics (ARABUS)

Author(s) : 1. Atsadawut Surinrangsee Student ID : 650510740
2. Student ID :
3. Student ID :

Major : Data Science

Advisor(s) : 1. Associate Professor Dr. Jeerayut Chaijaruwanich
2. Assistant Professor Dr. Prakarn Unachak
3.

Type of presentation* *(choose 1)* :
- ☑ Oral Presentation (เฉพาะ ตัวแทนนศ.ที่สาขาเลือกให้นำเสนอแบบบรรยาย)
- ☐ Poster (กรณี นำเสนอผลงานปัญหาพิเศษ/การค้นคว้าอิสระ)
- ☐ Cooperative Education (กรณี นำเสนอผลงานสหกิจศึกษา)

ABSTRACT

This cooperative education project aims to develop an agentic AI system under the project "Agentic Rapid Analysis of Bioinformatics Utility System for Microbial Genomics (ARABUS)". ARABUS is designed to provide an end-to-end microbial genome analysis workflow. Domain experts upload DNA sequence files along with associated metadata, after which ARABUS retrieves the inputs, generates an analysis plan, and orchestrates bioinformatics tools to perform key tasks including strain typing, outbreak investigation, antimicrobial resistance (AMR) gene detection, and identification of AMR gene transfer mediated by plasmids. The system then summarizes and interprets results in an accessible format, enabling scientists and public-health professionals to use genomic evidence more effectively for surveillance, outbreak tracking, and microbial genomics research.

The development process involved eliciting real-world questions and requirements from domain experts and healthcare personnel, followed by implementation of two categories of capabilities: deterministic functions, which can be executed directly without complex decision-making, and adaptive functions, which require multi-step reasoning before tool selection and execution.

From a system-engineering perspective, the project designed and implemented an agentic AI architecture that integrates Large Language Models (LLMs) with bioinformatics tools such as MLST typing (Multi-Locus Sequence Typing), SNP distance calculation (Single Nucleotide Polymorphism distance), cgMLST (core-genome MLST), and other supporting utilities. Automated workflows were developed for tool invocation, result parsing, and post-analysis interpretation. In addition, prompts and command templates were engineered to improve the reliability and correctness of LLM-driven tool use. Users interact with the system through a demonstration web

**Type of presentation must be matched with an option you choosing on student upload system.*
***The abstract can be more than one page and must be approved by project advisor before upload.*

application built with FastAPI and Gradio, allowing natural-language queries and returning outputs that are analyzed and explained in a clear, user-friendly manner.

Preliminary evaluations with domain experts and project advisors indicate that ARABUS can answer relevant questions, execute analyses, and interpret microbial genomic data accurately. The system helps reduce barriers related to specialized bioinformatics expertise and programming skills, suggesting that ARABUS has strong potential for further development and future applications in public health and bioinformatics.

**Type of presentation must be matched with an option you choosing on student upload system.*

***The abstract can be more than one page and must be approved by project advisor before upload.*