Iowa Initiative for Artificial Intelligence
Final Report

Project title:	Development of Artificial Intelligence Chatbot as Expert System for Antibiotic Resistance Interpretation and Stewardship	
Principal Investigator:	Lemuel Non, MD	
Prepared by (IIAI):	Avinash Mudireddy	
Other investigators:	Kunatum Prasidthrathsint, MD Meredith Parsons, MD Bradley Ford, MD	
Date:	May 20, 2025	
Were specific aims fulfilled:	Yes	
Readiness for extramural proposal?		
If yes … Planned submission date		
Funding agency		
Grant mechanism		
If no … Why not? What went wrong?		

Brief summary of accomplished results:

We developed a web-based chatbot application hosted on Microsoft Azure, leveraging Azure OpenAI and Promptflow. The chatbot interprets MIC (Minimum Inhibitory Concentration) data into susceptibility categories using CLSI guidelines and identifies phenotypic resistance patterns. The system integrates local antibiotic stewardship rules to assist with antibiotic recommendations. The platform provides fast, consistent, and accurate responses, with evaluations indicating high agreement with clinical expert assessments.

Research report:

Aims (provided by PI):

- Interpret MIC values into Susceptible, Intermediate, or Resistant using CLSI guidelines.
- Identify phenotypic resistance patterns (e.g., ESBL, carbapenemase producers).
- Incorporate local antibiotic treatment rules to guide therapy.
- Validate chatbot responses using bacterial isolates and ID expert review.

Data for Aims:

We used two core datasets for driving and validating the expert system:

1. MIC and Zone Diameter Breakpoints:
   This dataset includes interpretive criteria for a wide array of genus-level organisms and antibiotics, detailing susceptibility breakpoints in terms of MIC (µg/mL) and zone diameters (mm). It follows EUCAST standards and provides category thresholds (Susceptible, Intermediate, Resistant) based on disk diffusion and MIC data.
2. EUCAST Expert Rules for Enterobacterales:
   A set of rule-based interpretations, encoding resistance phenotypes such as AmpC, ESBL, and carbapenemase production based on observed susceptibility to certain indicator antibiotics. These rules are applied to taxa like E. coli, Klebsiella, Enterobacter, etc., using if-then logic chains that guide inferred resistance mechanisms and antibiotic reporting constraints.

These structured data sources were embedded into the Promptflow logic through carefully crafted prompts, enabling the AI to apply microbiological rules consistently without direct training. They also enabled phenotype recognition by pattern-matching antibiotic susceptibilities against codified expert logic. Human experts used these same datasets as reference standards for reviewing and validating model outputs.

AI/ML Approach:

This project implements a cloud-based chatbot application using a combination of Flask (Python backend), HTML frontend, and Azure OpenAI's GPT-4o model, orchestrated via Microsoft Promptflow. The application is designed to interpret antimicrobial susceptibility test results and support clinician queries related to antibiotic resistance and treatment decisions. At its core, it leverages structured promptflow chains deployed via an Azure Machine Learning scoring endpoint, embedding clinical microbiology logic derived from expert-validated sources such as CLSI guidelines and EUCAST rules.

The backend logic, defined in app.py, runs a Flask server that handles POST requests from the web frontend. When a user submits MIC data or a clinical question, the backend constructs a JSON payload containing the query and prior chat context, which is then sent to the Azure-hosted Promptflow endpoint. This endpoint uses GPT-4o to execute logic encoded in the form of modular prompts, each responsible for a specific reasoning step—such as MIC-to-category translation, resistance pattern detection, or therapy recommendation.

A key architectural feature of the application is its use of Azure AI Search indices to store and retrieve tabular and rule-based references in real time. Specifically, the MIC and Zone Diameter Breakpoints dataset and the EUCAST Expert Rules for Enterobacterales are indexed and queried during inference. This allows the model to ground its outputs with exact matches to susceptibility thresholds and resistance inference rules, enhancing consistency and interpretability. During runtime, Promptflow logic performs dynamic lookups in these indices, incorporating relevant context into the model prompt, rather than relying on static or hardcoded rules.

The system modularizes its decision logic across different Promptflow steps, which include:

1. MIC classification based on EUCAST/CLSI thresholds,
2. Pattern recognition using multi-antibiotic resistance logic (e.g., detecting ESBL phenotypes), and
3. Contextual therapy guidance based on localized antibiotic policies.

The GPT-4o model operates strictly under rule-constrained prompts, without any fine-tuning, ensuring that all interpretations remain transparent and auditable. The architecture allows rapid updating of rules and thresholds simply by modifying the AI search index content or prompts—without needing to retrain the model.

Validation was performed via a human-in-the-loop setup, where expert microbiologists and infectious disease physicians assessed the chatbot's interpretations and responses for clinical accuracy. The application is fully modular, scalable, and future-ready for EMR integration or extension to additional resistance phenotypes and pathogens.

Experimental methods, validation approach:

No formal experimental methods were conducted due to the development-focused nature of this pilot. Validation was done through human-in-the-loop expert review of chatbot outputs. MIC interpretation and antibiotic guidance were assessed by microbiologists and ID physicians, with results deemed satisfactory.

Results:

- MIC interpretation accuracy was consistently high
- Resistance pattern recognition was satisfactory based on expert review
- Antibiotic recommendation responses aligned with local guidelines