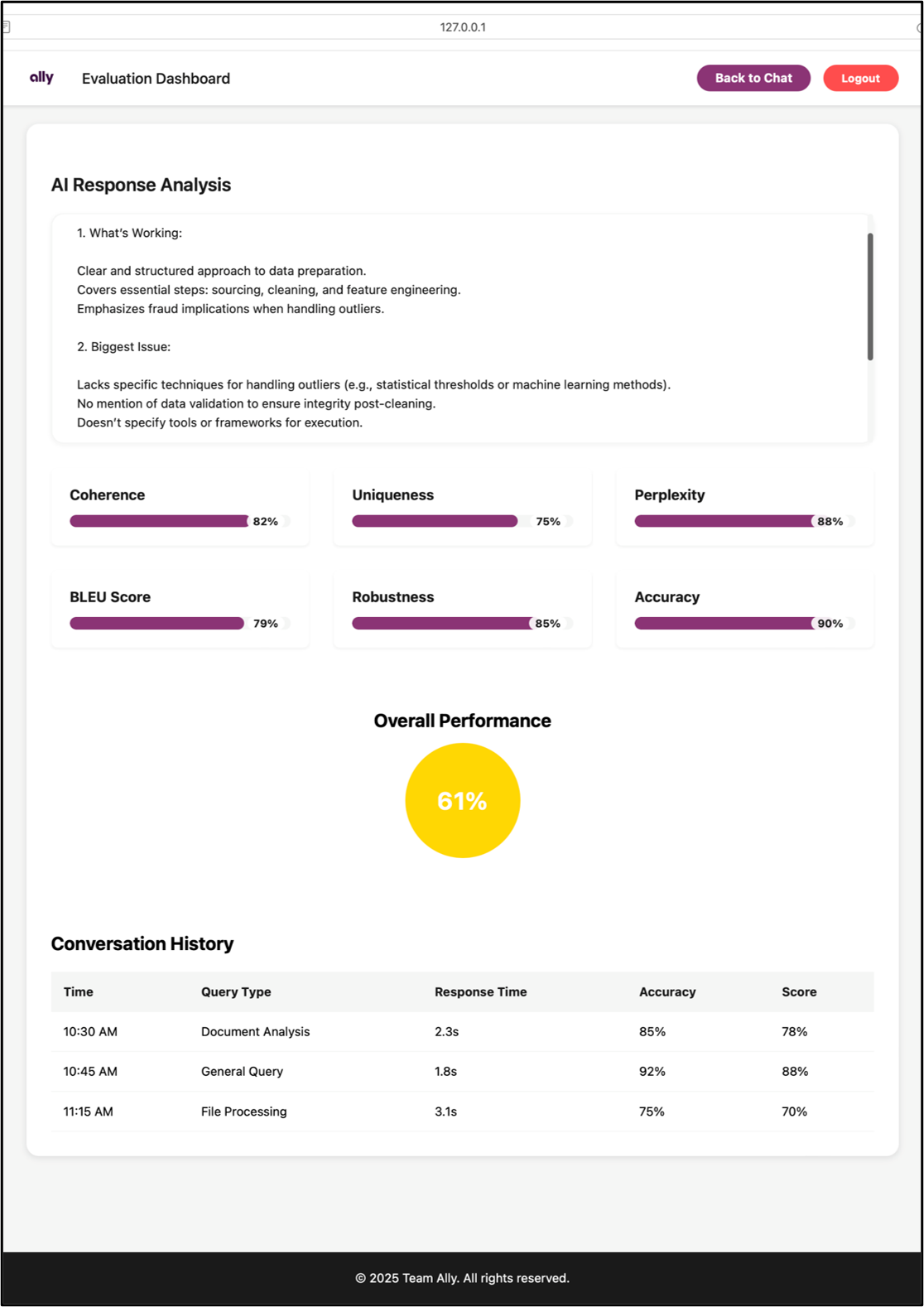
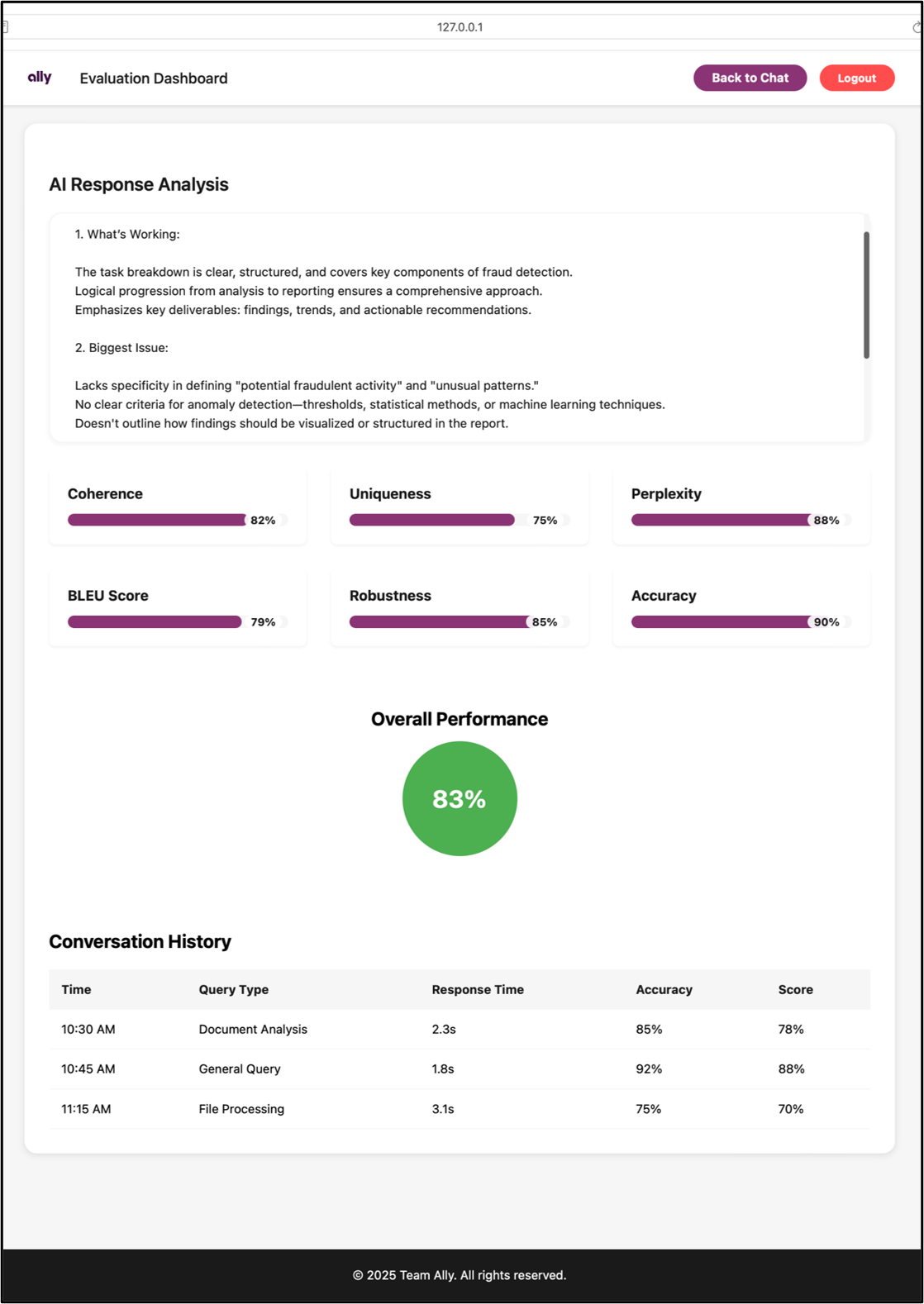
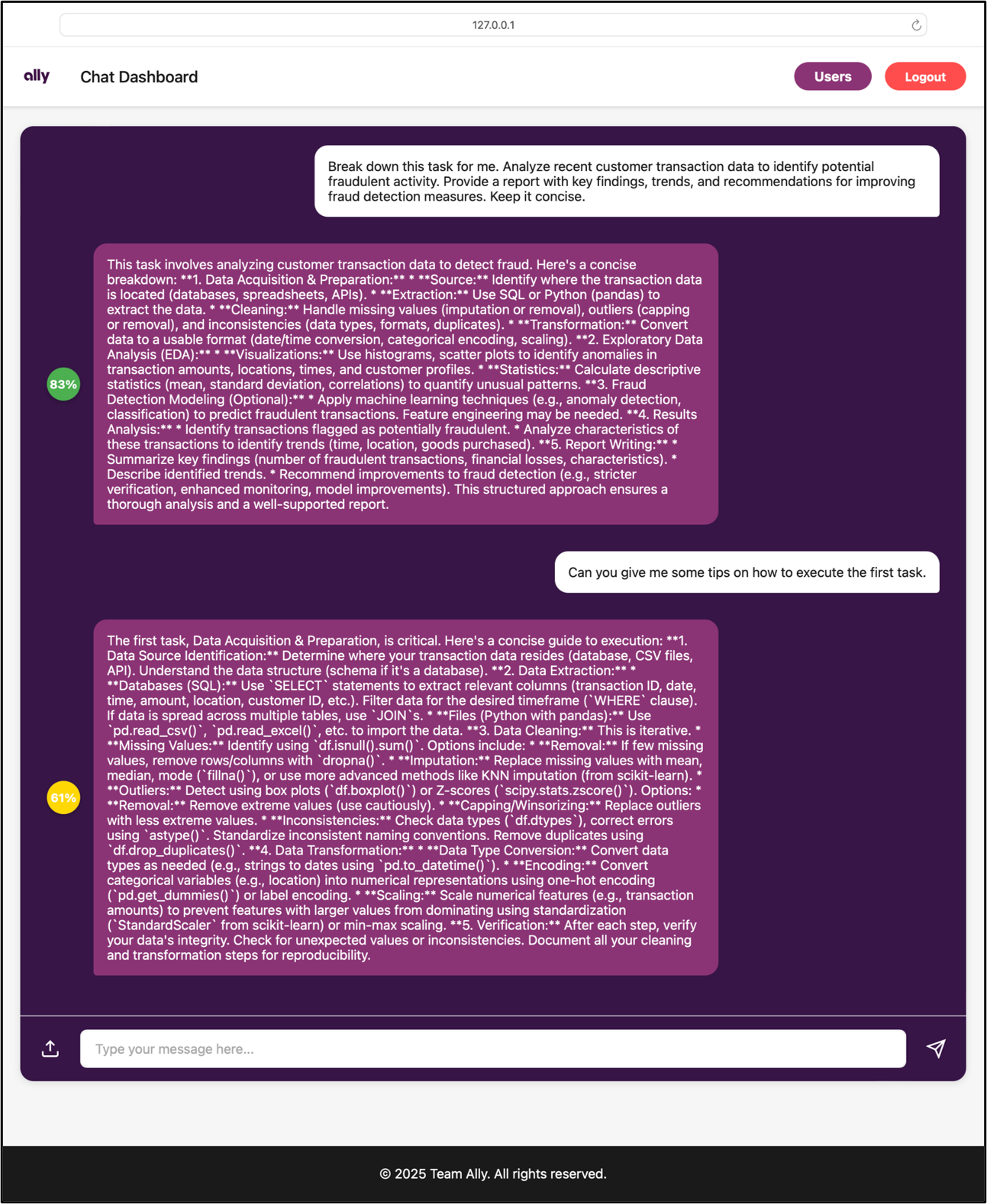
Design Day Booklet Team Page







PAGE N + 2



Ally

Project Sponsors

Jesue Jackson

Detroit, Michigan

Divyesh Jambusaria

Charlotte, North Carolina

Dan Lemont

Detroit, Michigan

Jesse Podell

New York, New York

John Stoutenger

Charlotte, North Carolina

Theresa Weaver

Detroit, Michigan

Michigan State University

Team Members (left to right)

Vu Ho

Hanoi, Vietnam

Andrew Dagher

Livonia, Michigan

Gabe Moraru

Grand Rapids, Michigan

Michael Plante

Ada, Michigan

Ethan Gomez

Macomb, Michigan

Amit Wagh

Novi, Michigan



Ally Financial, headquartered in Detroit, Michigan, is a leader in the U.S. financial services industry. Recognized as one of the nation’s largest online-only banks, Ally provides an array of online banking services to approximately 11 million customers.

Given recent increased interest in artificial intelligence, Ally Financial is experimenting with using generative artificial intelligence (GenAI) to automate various internal business processes. Research conducted thus far by Ally and others is promising, but GenAI’s novelty and complexity create concerns regarding its reliability of performance.

Currently, there is no testing framework in place to accurately assess where GenAI excels and when it should be used in business practices. Without such a framework, it is time-consuming to identify use cases where it is appropriate to apply GenAI.

Our AI System Testing Framework evaluates how a GenAI model performs on a specific task. Given a prompt, the application indicates how well the GenAI model responds to the prompt by displaying meaningful evaluation scores associated with the interaction such as accuracy and relevancy.

After accessing the application, a user interacts with GenAI through a chatbot-like interface. The user prompts the GenAI with a professional use case and reference response, receives an output, and is then redirected to an evaluation page. The evaluation page provides a visualization of scores on how well the AI performed for that use case. Additionally, the user views past interactions and the scores associated with those interactions.

The front end of this system is built using HTML, CSS, and JavaScript. The back end is implemented in Python and uses the Flask library to create a web application. A server provided by the MSU Division of Engineering Computing Services (DECS) is used to host a PostgreSQL database where relevant data is stored.

CSE498 | 8:00 a.m. – Noon Computer Science and Engineering, Third Floor | 3200/3300 Hallway

Ally Financial

AI System Testing Framework