

1. Project title

Insight Coding

2. Team members

Efe Erdoğmuş	22203553
Altay İlker Yiğitel	22203024
Mustafa Mert Gühan	22201895
Ece Bulut	22202662
Yasemin Altun	22202739

3. Description

The motivation for our project comes from a need in software engineering research: the systematic evaluation of diverse software artifacts created both by humans and automated systems such as Large Language Models (LLMs). Evaluating artifacts such as source code, test cases, UML diagrams, and requirements documents is often subjective, inconsistent, and time consuming. Researchers currently lack a unified platform that allows them to set up studies, recruit participants, and gather meaningful comparative feedback. Our web application will address this gap by offering an integrated environment for managing artifact based user studies.

The main goals of the project are to provide researchers with tools to easily organize and conduct empirical studies, enable participants to evaluate artifacts through structured comparison tasks with annotation and rating features and ensure data is collected, visualized, and can be exported for meaningful analysis. By doing this, the system will solve one of the important problems in software engineering: lowering the barrier to conduct reproducible, participant driven evaluations.

Key features of the application include user management across different roles such as researchers, participants, admins, and reviewers to provide with secure registration and login, artifact upload and organization, customizable study creation, participant competency assessment through quizzes, and side by side artifact comparison with annotation and rating tools. Dashboards for both researchers and participants will support monitoring progress, managing tasks, and visualizing results. Additionally, support for blinded evaluation and AI assisted quiz generation will make the system more flexible and fair.

The selling points of our application lie in its modular design, extensibility to future artifact types. It combines artifact management, evaluation, analytics, and participant engagement and puts it into a single platform. What makes this project interesting is its integration of AI-based tools into an artifact comparison platform that not only supports reproducible and replicable studies but is also easy to implement and provides valuable performance insights.

4. Extra Feature

Adaptive Task Assignment:

An extension for the Artifact Comparator platform is the addition of an adaptive task assignment mechanism. In many human subject studies, participants may have varied backgrounds, ranging from less experienced students to highly experienced professionals. Giving all participants the same evaluation tasks can lead to a number of problems: experts may find the tasks dull and basic, while beginners may find the artifacts too complicated, resulting in data that is either insufficient or of poor quality. The system can address this by dynamically adapting tasks to each participant's skill level based on the answers to the initial background questionnaire and competency quiz.