Agent-driven Code Recommendation Service (ACRS) using AutoGen and LASSO

Team Project

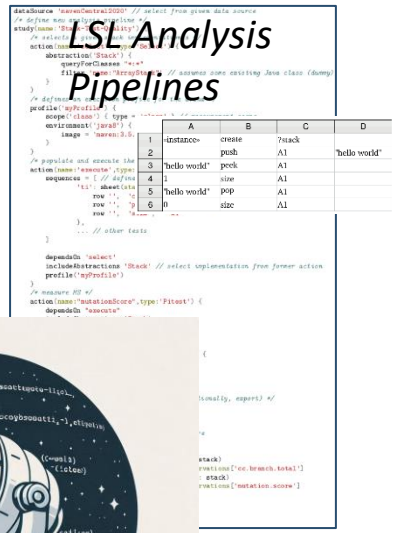
Chair of Software Engineering (Prof. Atkinson) &
Chair of Web-based Systems (Prof. Bizer)

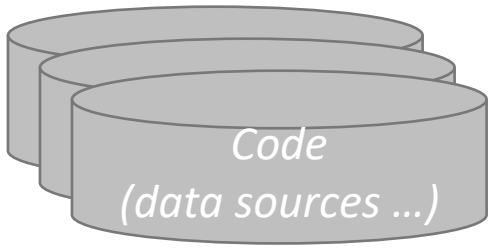
Fall Semester 2024

An Agent-driven Code Recommendation Service

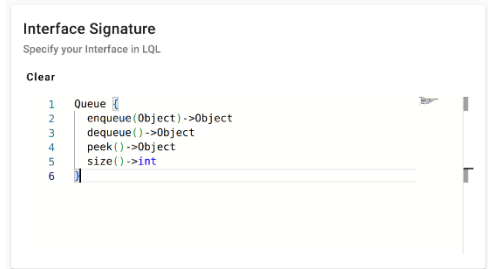
- Summary
  - Develop an AI-driven, agent-based service that provides personalized code recommendations based on analysis of large numbers of software implementations, mimicking the functionality of LASSO's analysis pipeline (LSL)
- Project involves
  - **Agents**: Use autonomous software agents to drive the code search process
  - **Generative AI**: Use of LLMs to enable conversations and perform tasks
  - **Code Search**: The primary focus is on searching for *relevant* code snippets or implementations
  - **Services**: The provision of a service-oriented agent framework that offers code recommendation capabilities
- Technology Components
  - LASSO platform: automate the gathering and analysis (including static and dynamic analysis) of large numbers of harvested software implementations (https://github.com/SoftwareObservatorium/lasso)
  - LLMs for Software Engineering tasks + LangChain: AI toolkitfor integration (https://www.langchain.com/langchain)
  - AutoGen: agent framework (https://microsoft.github.io/autogen/)

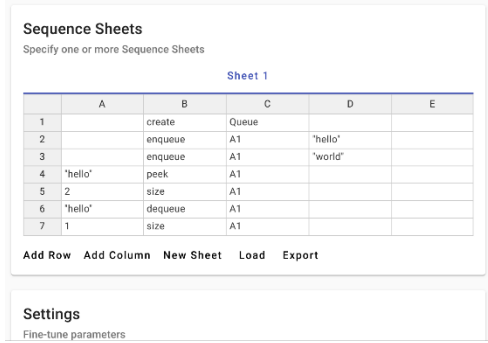
LASSO Platform – Code Search Service



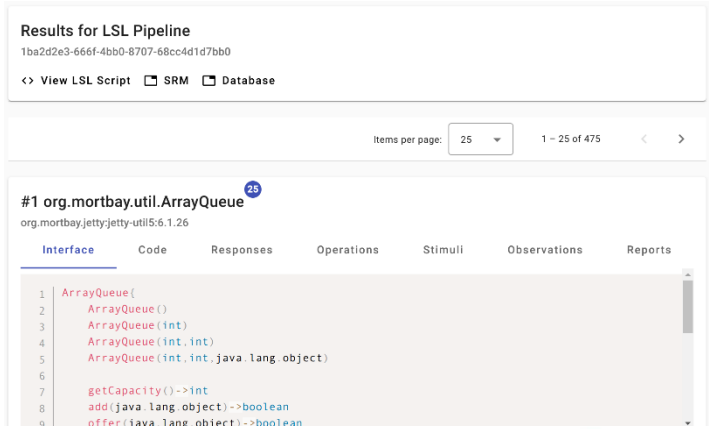


1) Query





2) Manually pick best match (preferences)



LLMs – Code Generation

LLMs (chatbots) trained on massive amounts of open source code

GPT-*

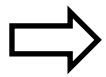
tabnine

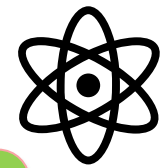
StarCoder

GitHub Copilot

… many more

generate a java method that serializes an object to a string using JSON





Here's an example Java method, serializeToJsonString, ….

…

….

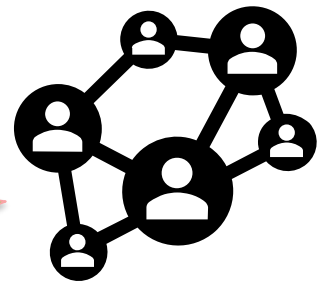
ACRS – Multi-Agent Code Recommendation (1)

Recommend a Java method that serializes an object to JSON. Pick the best match based on minimal cyclomatic complexity and project popularity.

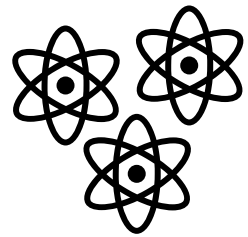
The best match is Java method, toJson, ….

In addition to cyclomatic complexity, also consider X and Y …

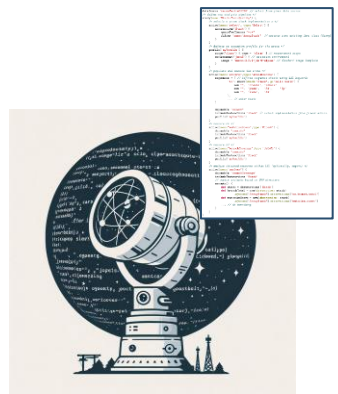
….



Agents



LLMs



LASSO Platform

ACRS – Multi-Agent Code Recommendation (2)

Recommend a Java method that serializes an object to JSON. Pick the best match based on minimal cyclomatic complexity and project popularity.

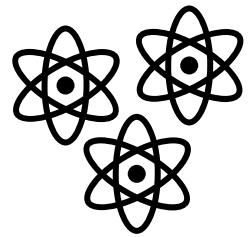
Chat with LASSO (i.e., enable conversations)

The best match is Java method, toJson, ….

In addition to cyclomatic complexity, also consider X and Y …

….

Agents



LLMs

LASSO Platform

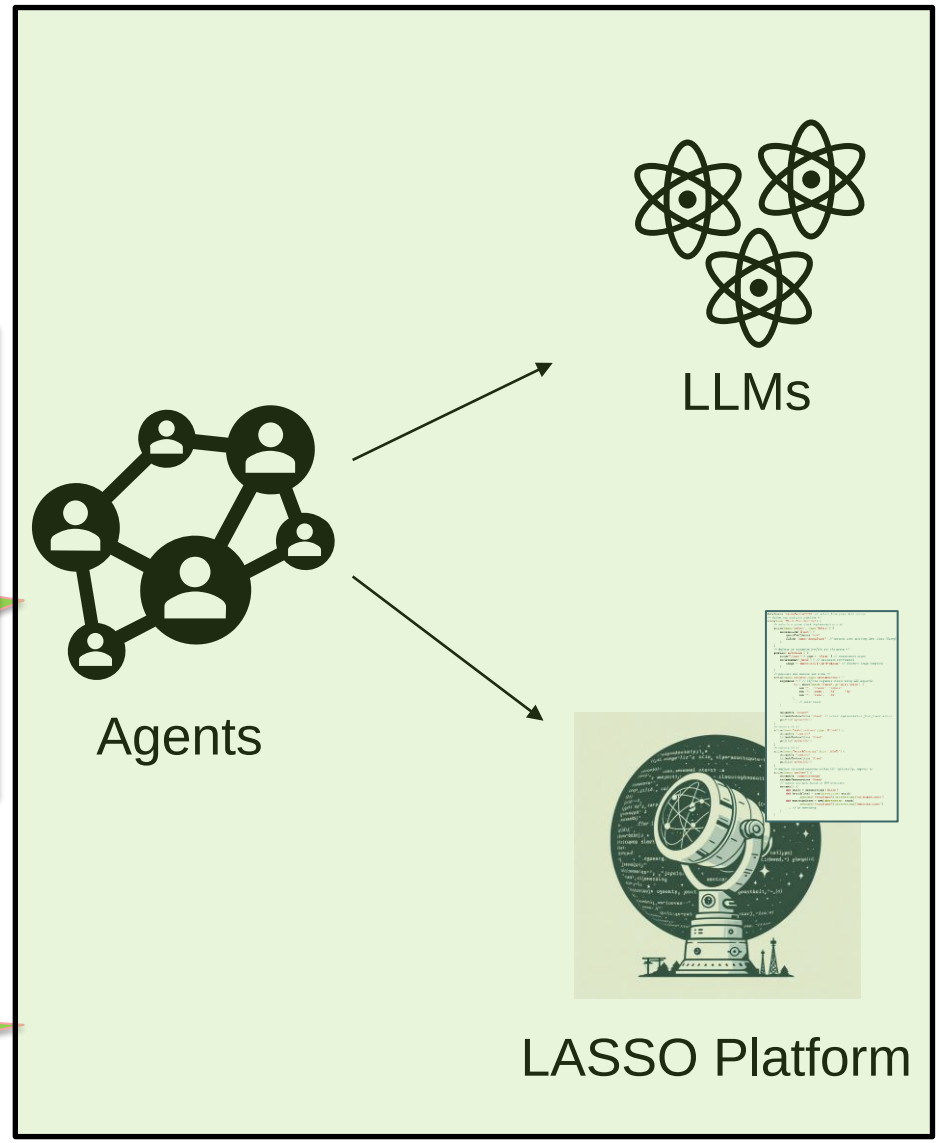
ACRS – Multi-Agent Code Recommendation (3)

Recommend a Java method that serializes an object to JSON. Pick the best match based on minimal cyclomatic complexity and project popularity

- AutoGen: agents + LLM workflows
- LangChain: use LASSO's platform services to retrieve "code", "facts" (e.g., metrics) and use analysis "tools" (e.g., execute and analyze code)

complexity, also consider X and Y …

….



Requirements

- Participants
  - 4-6 students
- Length
  - 6 months
- Prerequisites
  - Python and/or Java programming skills
  - Basic understanding of machine learning
  - Ideally, IE 686 Large Language Models and Agents
- Language
  - English
- Organisation
  - Goals and timetable defined by agreement with the supervisor
- Applicable to MMDS: yes
- Online: By agreement
- Main Supervisor
  - Marcus Kessel
- Technical Supervisor (Agents)
  - Ralph Peeters