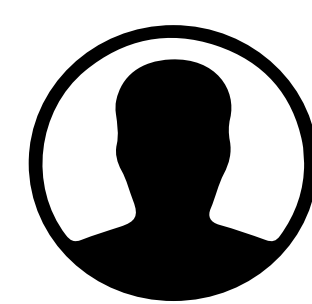


Human



Research Idea

Does bias affect language model accuracy on QA benchmarks?

Human



Notes

* Please use gpt-4o-mini for your experiments
* Use the following API key...



Research Report

Agent Laboratory: Using LLM Agents as Research Assistants

Samuel Schmidgall¹
¹Johns Hopkins University

[Code assist software like GitHub Co-pilot and X ; however, LLM applications in research primarily focus on automating the research process. In this paper, we introduce Agent Laboratory, an LLM agent framework for accelerating the individual's ability to perform research. Unlike previous work, where agents produce their own research topics, this work aims to function as a tool to assist human scientists to perform their own research ideas using LLM agents, rather than automate the entire process. Agent Laboratory takes a *human-produced* research idea as input and outputs a code repository and a research report. This is accomplished through the use of specialized agents driven by LLMs that collaborate to perform research based on human input.

Code Repository

```
├── src
│   ├── load_data.py
│   ├── run_experiments.py
│   └── ...
├── readme.md
└── requirements.txt
```

Literature
Review

Plan
Formulation

Data
Preparation

Running
Experiments

Results
Interpretation

Report
Writing

Report
Refinement