

# Demo and Status Report: Autonomous Web Agent

- 05.05.2023



# Overview

## Recent progress

- Got a basic agent of UI-tars to run on the cloud
- Divided ourselves into subgroups

## Working on:

- Comparing UI-Tars with the previous cohort's agent

## Goals for next week

---



---

# Progress

This week we two key tasks, the first being that a basic web agent is running on the cloud. Second being that we divided ourselves into subgroups to tackle different tasks simultaneously. These subgroups and their focuses are:

1. Get the UI-Tars to run on the cloud and experiment with it.
  2. Improve the basic version of the agent on Colab.
  3. Compare the 2B to the 7B model.
-

---

# Our Current Progress

We are currently increasing our focus on getting the desktop version of UI-Tars to run as it might simplify our project significantly.

We are also looking into where the model is lacking or if any bugs arise while running the model to further work on them and improve the overall model.

Recently, we added an image analysis feature to the agent. This allows it not only to accept images as input but also to interpret and understand the content of those images

---

---

# Issues we are facing

UI-Tars has such a highly intensive code that makes running it locally really hard, we are still working on getting the full desktop version to run locally. When we tried to run the original source code we still got some errors but many of them have been resolved.

The agent may repeat the same answer multiple times or get in a loop while giving the answer so that is another thing that we will need to work on.

---

---

# Next Week's Goals

- We would still try to get the UI-Tars desktop version to run locally/on the cloud.
  - Fine tuning: we are going to work on fine tuning the model to enhance the model's performance and accuracy. This should help eliminate any redundancies like the problem of repeated answers.
-

# Goals for next meeting

1. A fine tuned model that is capable of returning accurate answers.

---