Softlight Engineering Take-Home Assignment

Task

You're building part of an AI multi-agent system. Agent A sends your agent, Agent B, different questions at runtime: like "How do I create a project in Linear?" or “How do I filter a database in Notion?”.

Your goal: Build the system that shows Agent A how to perform the requested task. It should automatically navigate the live app and capture screenshots of each UI state in the workflow. Agent B ***won’t know these tasks ahead of time***, so your system must be generalizable: it should handle any kind of request, across different web apps, and capture the right UI states ***in real time***.

The Problem

Not every UI state has a URL. Consider what needs to be captured for "creating a project in Linear":

- The project list page (has a URL)
- The "Create Project" button state
- The create modal (no URL)
- The form fields (no URL)
- Maybe the success state

Your system needs to navigate the live application and capture these states programmatically on the fly.

Testing

Pick 1–2 web apps (e.g. Linear, Notion, Asana, etc). To test your system, you can ***define a few example tasks yourself*** (like “creating a project,” “filtering issues,” or “changing settings”). But your implementation ***shouldn’t rely on these being hardcoded***. The goal is to show that your system’s approach could generalize to other apps and tasks it hasn’t seen before.

What we want to see:

- Capture 3-5 different tasks/workflows across your chosen app(s)

- Thoughtful approach to navigating and capturing non-URL states

Deliverables

1. **Code** - Your UI state capture system
2. **Loom** - A short loom where you show the agent running through a workflow and explain how it works.
3. **Dataset** - Captured UI states for 3-5 tasks across 1-2 apps, organized by task. Accompanied by a short blurb explaining the tasks. Structure this in the best way you see fit.

Practical Details

- **Time:** Up to a week.
- **Tools:** Use whatever you want - any languages, frameworks, libraries.
- **Scope**: Focus on quality. We'd rather see 3-5 tasks captured well than 20 tasks captured poorly.

Submission

To submit: github repo link, loom, and dataset.