

# [Crustdata Build Challenge : AI Agent for Browser Automation]

## Goal:

Build an AI agent to automate browser workflows on your machine's native browser

**Example:** [OpenAI's Operator](#) but on your browser

## Grammar:

- **Interact** API
  - An API that accepts natural language commands and translates them into browser automation actions implemented using browser automation frameworks like Playwright, Puppeteer, or Selenium.
- **Extract** API
  - An API that allows the automation agent to retrieve and parse structured data from web pages that the browser has loaded.

## Milestones:

### Level 1 - Basic Browser Control

#### Minimum requirement:

- Implement an **Interact** API that accepts natural language commands to control browser actions
- Properly handle common error scenarios with clear error messages
- Successfully demonstrate one complete flow with the **Interact** API:
  - Log into a popular website
  - Perform a search with user-specified keywords
  - Navigate through search results and interact with a specific result item

#### How to submit:

- Send private github repo via email to buildchallenge AT crustdata.co with loom link of the working demonstration
- NOTE: WILL ONLY EVALUATE WITH A RECORDING URL

#### On successful completion:

- Technical interview call scheduled with Crustdata team
  - Note: You would only be judged on your performance in interviews and your work up to Level 1

## Level 2 - Advanced Browser Integration

### Minimum requirement:

- All Level 1 requirements, plus:
- Enhance the interact API to directly control native browsers on your machine (Chrome/Firefox) through their respective browser-specific interfaces, rather than relying on third-party browser automation frameworks
- Implement an **extract** API to retrieve **structured** data from web pages
- Demonstrate at least one complete automation flow using both APIs:
  - Log into a popular website
  - Perform a search with user-specified keywords
  - Navigate through search results
  - Parse necessary information in a structured format
- Support for the following:
  - Proxy configuration
  - Browser extension integration

### How to submit:

- Successful submission of level 1
- Send private github repo via email to buildchallenge AT crustdata.co and loom link of the working demonstration
- NOTE: WILL ONLY EVALUATE WITH A RECORDING URL

### On successful completion:

- 3,000 USD (total)

## Level 3 - Contextual Intelligence & Advanced Workflows

### Minimum requirement:

- All Level 1 and Level 2 requirements, plus:
- Cross-platform compatibility with identical functionality across Windows, macOS, and Linux
- Agent can spawn automatically at periodic intervals and do specific tasks as laid out in Level 2
- Implement a conversational interface where the agent can maintain context across multiple commands

### How to submit:

- Successful submission of level 1 and 2
- Send private github repo via email to buildchallenge AT crustdata.co and loom link of the working demonstration
- NOTE: WILL ONLY EVALUATE WITH A RECORDING URL

### On successful completion:

- 15,000 USD (total)
- Full-time offer (after 15 min call with CEO)

## General Rules & Tips:

- Loom recordings demonstrating successful automation across multiple websites with complex workflows will be scored significantly higher
- Bonus points for implementing solutions to common automation challenges:
  - CAPTCHA detection and handling
  - Managing login sessions effectively
  - Dealing with dynamic page content and loading states
  - Graceful recovery from unexpected page states
- Solutions must be built from scratch - not extending existing tools (feel free to take inspiration from projects from such)
  - <https://github.com/lavague-ai/LaVague>
  - <https://github.com/Skyvern-AI/skyvern>
  - <https://github.com/steel-dev/steel-browser>
  - <https://github.com/browserbase/stagehand>
  - <https://github.com/browser-use/browser-use>
  - Etc