Intro

This is a folder that contains worklogs in order to be clear what happened later when we want to look back.

Key things achieved today (10 th and 11 th Dec 2024)

- I will be using OpenHands to help me code up various aspects of the project.

    - This involved installing docker
    - Pulling the OpenHands image.

```
PS C:\Users\apagnih> docker pull docker.all-hands.dev/all-hands-ai/runtime:0.15-nikolaik
0.15-nikolaik: Pulling from all-hands-ai/runtime
5cf2c95d3f0f: Download complete
0d18fa4eae23: Download complete
11bdfd896a14: Download complete
e8a4d7981d5c: Download complete
5bd71677db44: Download complete
087caae23ec2: Download complete
8871a2ca0dba: Download complete
69f3fb8b236a: Download complete
963b287153b9: Download complete
df2904bf002b: Download complete
b6e5ddaaefa5: Download complete
53dbba7cf7ca: Download complete
4f4fb700ef54: Download complete
f0c4435bfc32: Download complete
ee2893fb6435: Download complete
551df7f94f9c: Download complete
ce82e98d553d: Download complete
b167200c1204: Download complete
4616dd779c51: Download complete
5c5f1cb4ace0: Download complete
fdf894e782a2: Download complete
7c6787aafdc2: Download complete
d4d416def475: Download complete
aad7ec0e4fe1: Download complete
```

- Running the OpenHnads image.

```
Starting OpenHands...
Running OpenHands as root
INFO:      Started server process [11]
INFO:      Waiting for application startup.
INFO:      Application startup complete.
INFO:      Uvicorn running on http://0.0.0.0:3000 (Press CTRL+C to quit)
INFO:      172.17.0.1:41782 - "GET / HTTP/1.1" 200 OK
INFO:      172.17.0.1:41782 - "GET /assets/root-BY8Piu9K.css HTTP/1.1" 200 OK
INFO:      172.17.0.1:41802 - "GET /assets/entry.client-Ch8I8ZCf.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41784 - "GET /assets/chunk-7R3XDUXW-DSKRCawN.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41782 - "GET /assets/auth-context-DKm1Ywm8.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41798 - "GET /assets/react-redux-BJt37jkE.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41796 - "GET /assets/use-config-CUlxI4fa.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41794 - "GET /assets/manifest-a7071e10.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41802 - "GET /assets/index-CdvD0Qri.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41794 - "GET /assets/store-OifvkhaZ.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41798 - "GET /assets/infiniteQueryBehavior-BtqYQEvL.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41798 - "GET /assets/with-props-CNz_g3pE.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41798 - "GET /assets/root-2NgAw5Co.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41784 - "GET /assets/index-DfJCZdWx.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41796 - "GET /assets/user-prefs-context-BSq2RHVU.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41782 - "GET /assets/i18nInstance-CnyO1Rew.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41802 - "GET /assets/initial-query-slice-CSE8L-ER.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41794 - "GET /assets/browser-slice-Cc8xkkTE.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41794 - "GET /favicon.ico HTTP/1.1" 200 OK
INFO:      172.17.0.1:41802 - "GET /locales/en/translation.json HTTP/1.1" 200 OK
INFO:      172.17.0.1:41794 - "GET /locales/en-US/translation.json HTTP/1.1" 200 OK
INFO:      172.17.0.1:41802 - "GET /assets/route-CYjuvO3r.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41794 - "GET /assets/route-B2e5AW72.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41782 - "GET /config.json HTTP/1.1" 200 OK
INFO:      172.17.0.1:41794 - "GET /assets/github-logo-RCP2sA01.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41782 - "GET /assets/chunk-VHPYXGWP-BJD6r1b6.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41796 - "GET /assets/chunk-DBLREEYE-C6ESCLQg.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41802 - "GET /assets/upload-image-input-BywfHVog.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41784 - "GET /assets/arrow-send-CBP7uKxj.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41798 - "GET /assets/declaration-B0gJQZvs.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41794 - "GET /assets/use-end-session-DaV5YqTF.js HTTP/1.1" 200 OK
INFO:      172.17.0.1:41798 - "GET /config.json HTTP/1.1" 200 OK
INFO:      172.17.0.1:41802 - "GET /api/options/models HTTP/1.1" 200 OK
INFO:      172.17.0.1:41798 - "GET /api/options/agents HTTP/1.1" 200 OK
INFO:      172.17.0.1:41798 - "GET /api/options/security-analyzers HTTP/1.1" 200 OK
```

- I got a lot of pain using the API keys on my company laptop because of self-signed certificates. Finally decided to just use my computer back in India to setup docker and run the application.
- Now I can finally use Anthropic Sonnet 3.5 to do custom tasks and just request OpenHands to do it for me. Althogh I think cost would be a big factor because just running 2 commands, to create a readme for the "project" folder and creating a branch, it costed me 7 Rupees, I expect to probably have around 100 prompts in a day, that means 700 Rupees. This is kind of expensive.
- I thought of switching to Haiku 3.5 to reduce costs because, it is said to perform similar to Sonnet 3.5 (old). And the costs are $3.00 / $15.00 | $0.80 / $4.00 | input 1 mil tokens / output 1 mil tokens. But when I tried doing it got the following errors. Which basically mean somehow Haiku can't be used, 12 retries and got a cost of 6 Rupees.
- This is a problem because I want to try out Haiku first for simple tasks and when it doesn't work, move to the bigger model because the bigger model is 3.75 times more expensive. I don't want

to pay for shit queries like make me the readme.

- One thing I can try is use Sonnet a bit, and see how effective is the usecase. If it turns out that it lets me finish the work extremely fast, it might be worth it.
- The idea is, looking at OpenRouter's leaderboard for Haiku 3.5 and Sonnet 3.5. We can easily get the companies that are using these models the most. This gave me a good idea of which company will support using Haiku 3.5 and Sonnet 3.5. *Further it is a harder proof of being useful because people won't be throwing away cash for things that don't work.* Finally, from personal experience using the web interfaces of multiple models, Anthropic is the best at the moment, therefore just going to focus using that. Things that I want to try later:
  - Aider (This used 60M tokens on Haiku 3.5)
  - Cline (This used 231M tokens on Haiku 3.5) (This seems much easier given the GUI interface along with VS Code extension)
  - SillyTavern (This is useful for me if I want to build my own LLM wrapper for the startup. It is an LLM frontend for power users.)

Apps using Claude 3.5 Haiku

Top public apps this week using this model

1.		**Cline** > Autonomous coding agent right in your IDE	231M tokens
2.		**Kuli Kuli** > new	128M tokens
3.		**Aider** > AI pair programming in your terminal	59.5M tokens
4.		**SillyTavern** > LLM frontend for power users	26.6M tokens
5.		**Chub AI** > GenAI for everyone	21.8M tokens
6.		**OpenRouter: Chatroom** > Chat with multiple LLMs at once	17.1M tokens
7.		**MosaicAI** > new	14.2M tokens
8.		**CHIM** > AI framework for Skyrim	14M tokens

Apps using Claude 3.5 Sonnet

Top public apps this week using this model

1.		**Cline** > Autonomous coding agent right in your IDE	17B tokens
2.		**websim** > Simulate the web	4.79B tokens
3.		**SillyTavern** > LLM frontend for power users	1.04B tokens
4.		**OpenRouter: Chatroom** > Chat with multiple LLMs at once	793M tokens
5.		**Chub AI** > GenAI for everyone	711M tokens
6.		**Aider** > AI pair programming in your terminal	509M tokens
7.		**Open WebUI** > new	264M tokens
8.		**YOUR_APP_NAME** > new	183M tokens

- Given this fact, I will try other options in order to see the utility, because it can get expensive if I keep using Sonnet. ~Plus, given lack of usage visible on OpenRouter, it seems that it might be better to use other technologies.~ OpenRouter only puts companies using its APIs, and OpenHand doesn't use OpenRouter so we can't directly compare.
- Additionally, I felt the OH impementation assumes the user is dumb. Because it just relies on the LLM running a command without asking if it is alright. I think if the LLM is making any changes, or running commands, it should ask the user first. It can be slow but it allows the user to learn from the LLM and be aware of what the code actually looks like. We are not there yet where we don't need to see the code at all and just let the LLM do everything for us. When things break, it won't be the the LLM's responsibility but the user's. OH design principal of

creating pull requests is smart but I think I don't want that much abstraction away from code. But all in all, the user can look at the logs and see what's happening under the hood.

```
INFO:     172.17.0.1:50424 - "GET /config.json HTTP/1.1" 200 OK
12:24:21 - openhands:INFO: socket.py:40 - sio:oh_action:fT5UfS4cnNR4DhdqAAAJ
12:24:21 - USER_ACTION
[Agent Controller fT5UfS4cnNR4DhdqAAAJ] **MessageAction** (source=EventSource.USER)
CONTENT: Please look at the README and make the following improvements, if they make sense:
* correct any typos that you find
* add missing language annotations on codeblocks
* if there are references to other files or other sections of the README, turn them into links
* make sure the readme has an h1 title towards the top
* make sure any existing sections in the readme are appropriately separated with headings

If there are no obvious ways to improve the README, make at least one small change to make the wording clearer or friendlier
12:24:21 - openhands:INFO: agent_controller.py:331 - [Agent Controller fT5UfS4cnNR4DhdqAAAJ] Setting agent(CodeActAgent) state from AgentState.INIT to AgentState.RUNNING
12:24:21 - OBSERVATION
[Agent Controller fT5UfS4cnNR4DhdqAAAJ] AgentStateChangedObservation(content='', agent_state=<AgentState.RUNNING: 'running'>, observation='agent_state_changed')
12:24:21 - OBSERVATION
[Agent Controller fT5UfS4cnNR4DhdqAAAJ] NullObservation(content='', observation='null')


==============
[Agent Controller fT5UfS4cnNR4DhdqAAAJ] LEVEL 0 LOCAL STEP 0 GLOBAL STEP 0

12:24:22 - openhands:INFO: manager.py:137 - Conversation fT5UfS4cnNR4DhdqAAAJ connected in 0.00940394401550293 seconds
INFO:     172.17.0.1:57944 - "GET /api/list-files HTTP/1.1" 200 OK
12:24:29 - ACTION
[Agent Controller fT5UfS4cnNR4DhdqAAAJ] **CmdRunAction (source=None)**
THOUGHT: I'll help improve the README. First, let me locate and examine it.
COMMAND:
find / -name "README.md" 2>/dev/null
12:24:30 - OBSERVATION
[Agent Controller fT5UfS4cnNR4DhdqAAAJ] **CmdOutputObservation (source=EventSource.AGENT, exit code=0)**
find / -name "README.md" 2>/dev/null
/usr/local/lib/python3.12/site-packages/pipenv/vendor/README.md
/usr/local/lib/python3.12/site-packages/pipenv/patched/README.md
/usr/local/lib/node_modules/npm/README.md
/usr/local/lib/node_modules/npm/node_modules/nopt/README.md
/usr/local/lib/node_modules/npm/node_modules/pacote/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmversion/README.md
/usr/local/lib/node_modules/npm/node_modules/semver/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmpublish/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmexec/README.md
/usr/local/lib/node_modules/npm/node_modules/rimraf/README.md
/usr/local/lib/node_modules/npm/node_modules/cssesc/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmteam/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmaccess/README.md
/usr/local/lib/node_modules/npm/node_modules/qrcode-terminal/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmhook/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmpack/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmorg/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmfund/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmdiff/README.md
/usr/local/lib/node_modules/npm/node_modules/glob/README.md
/usr/local/lib/node_modules/npm/node_modules/@npmcli/config/README.md
/usr/local/lib/node_modules/npm/node_modules/@npmcli/installed-package-contents/README.md
/usr/local/lib/node_modules/npm/node_modules/@npmcli/arborist/README.md
/usr/local/lib/node_modules/npm/node_modules/cross-spawn/node_modules/which/README.md
/usr/local/lib/node_modules/npm/node_modules/libnpmsearch/README.md
/usr/local/lib/node_modules/npm/node_modules/which/README.md
/usr/local/lib/node_modules/npm/node_modules/node-gyp/gyp/docs/README.md
/usr/local/lib/node_modules/npm/node_modules/node-gyp/README.md
/usr/local/lib/node_modules/npm/node_modules/node-gyp/node_modules/nopt/README.md
/usr/local/lib/node_modules/npm/node_modules/node-gyp/node_modules/which/README.md
/usr/local/lib/node_modules/npm/node_modules/corepack/README.md
/usr/local/README.md
/usr/share/doc/libtasn1-6/README.md
/usr/share/doc/sudo/README.md
/usr/share/doc/tk8.6/README.md
/usr/share/doc/libwebp-dev/README.md
/usr/share/doc/libpango-1.0-0/README.md
/usr/share/doc/shared-mime-info/README.md
/usr/share/doc/libglib2.0-0/README.md
/usr/share/doc/fonts-dejavu-core/README.md
/usr/share/doc/libgdk-pixbuf-2.0-0/README.md
/usr/share/doc/libxml2/README.md
/usr/share/doc/libzstd-dev/examples/README.md
/usr/share/doc/libcbor0.8/README.md
/usr/share/doc/git/README.md
/usr/share/doc/git/contrib/vscode/README.md
/openhands/poetry/openhands-ai-5O4_aCHf-py3.12/lib/python3.12/site-packages/modal/requirements/README.md
/openhands/poetry/openhands-ai-5O4_aCHf-py3.12/lib/python3.12/site-packages/pyarrow/tests/data/orc/README.md
/openhands/poetry/openhands-ai-5O4_aCHf-py3.12/lib/python3.12/site-packages/litellm/realtime_api/README.md
/openhands/poetry/openhands-ai-5O4_aCHf-py3.12/lib/python3.12/site-packages/litellm/llms/README.md
/openhands/poetry/openhands-ai-5O4_aCHf-py3.12/lib/python3.12/site-packages/litellm/llms/azure_ai/README.md
[12] 0:[tmux]*M
```

/

- Finally, discord popularity says that Aider and Cline might be better at the moment. OpenHands just had a viral explosion which isn't equal to better agent. And I believe in Aider more because of the continued support since 19 months.
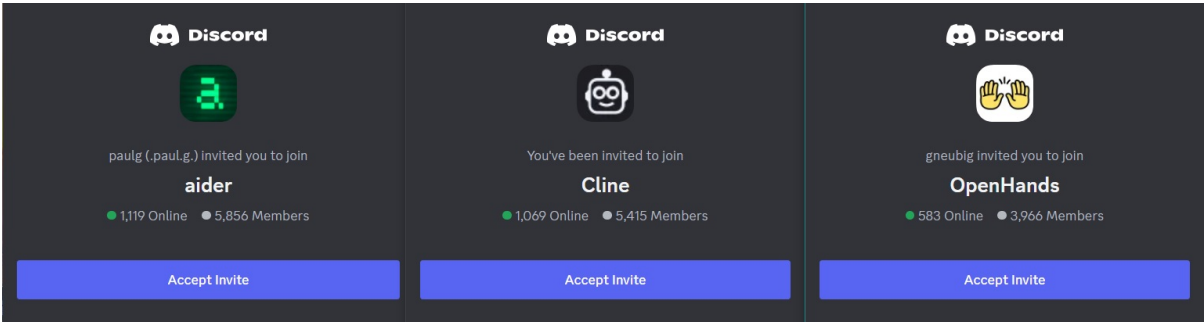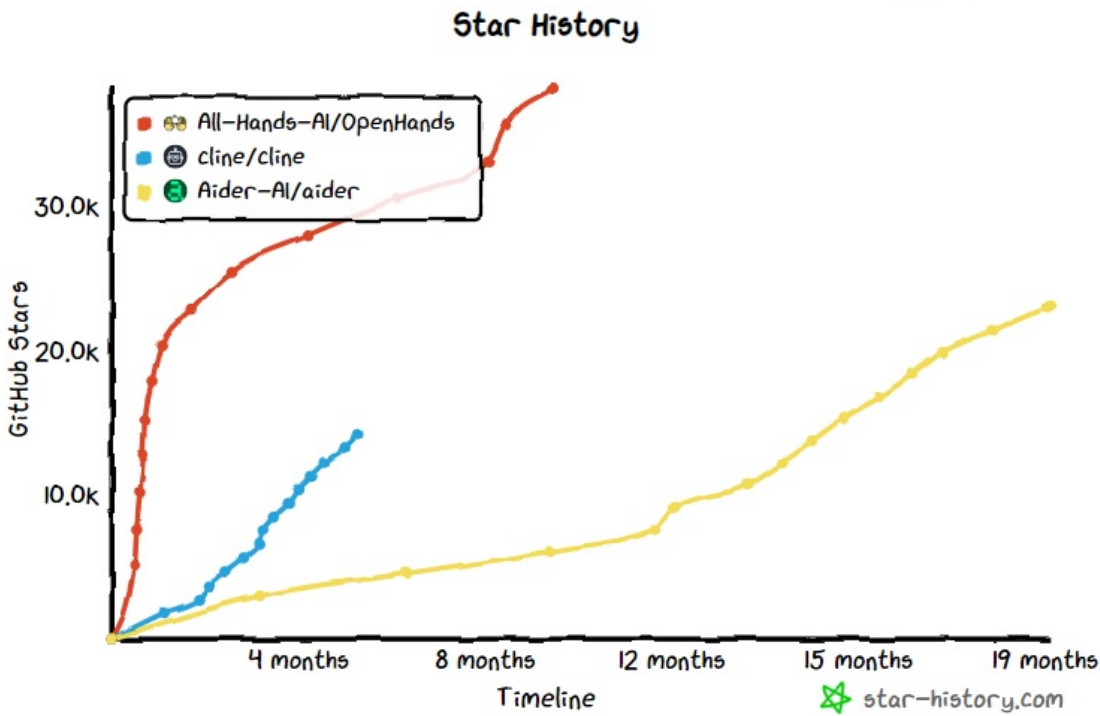
- Chat history with ChatGPT to better understand things I did and where I was stuck.
  https://chatgpt.com/share/675988ef-8720-8012-8c19-81fba550d2e6

Key things achieved today (10 th and 11 th Dec 2024)