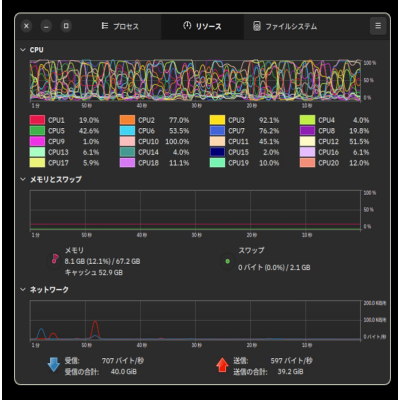




Llama3 70bã??ã?ã?¼ã?«ã?«ã•Så??ã•?ã•?ã•|ã•¿ã??

Description

ã?¢ã??ã•¥ã•ã??â|³/₄ã•®GPUæ•è¹/₄?PCã•sè©|ã•?ã•³/₄ã•?ã•?ã•??

PCă•®ă?¹ă??ă??ă?¬ă CPU Core i5 13500ă?RAM 64GBă?GPU RTX A4000ă•şă•ă??OSă¬Ubuntu
22.04 LTSă??

ăĵ¼ç??ă•?Ală•®ă??ă¼•ă•?a•Să•ă??ă??ă•?ă•ă•“ç?”æ?•ă•?ă•?PCă•Să•?ă•?ă•?ă??ă?|ă?ăă?ă•
 ă?¥ă?¼ă??ă?ă?°ă??RAGă•ă•©ă??èĵ?ă•Ēă•|ă??ă•Ēă•ă®?è“¼ă®?é“?ç??ă•ă•?ă•ă??ă•?ă•|ă•ĸă•?ă•ă
 •æ?•ă•Ēă•|ă•?ă•ă¼ă•?ă??ă¼?ă•?ă•Kubernetesă??ă½ĸă•Ēă•|GPUă??ă•©ă?¹ă?ĸă?¼ă??æſ?ç“ăăă
 |ă¼ă?ă•SLLMă??ă??ă?¹ă??ă?Ēă?ă?°ă•?ă?•ç?¹ă®?ă?é??ă•ăç?¹ă??ă•?ă•?ă°ă•ăĒă•
 ?LLMă??é??ç?°é?ç?“ă•?ă??ă•ĸă•?ă•?ă•ăă??

CPUă•?çµ•æ§?ăĹ?ă•?ă•?ă•šă•?ă??ă?~ă?¹ă•?ă³ă?¹ă??é•ă•?ă•šă•®70bă??ă??ă•?ă•ăă??ă??ă•£ă•
 ``é«?é?ă•ăCPUă??æ•è¼?ă«ă•?ă??ăĹ èĴă•?ă•?ă??ă•¾ă•?ă??ă ``ă³ă?ăšă•®Ċ¹¼ă•?ç??ă??ă¹³ă•?ă•
 ?ă??ă•³0%ă?•ă¾ăăšă•?ă•?ă??ă?ă•?ă??ă?«ă?
 ă³ă?ăšă•æ?šè½ă??ă??ă??ă?ă•ă³ă?ăšă«ă³ă?ăšă•æ?šè½ă•®é«ă?CPUă??é•ă?ă?ă•æ?¹ă•?ă•?ă•?ă•ăăă•
 ?ă•ăă??RAMă•ăăă¾ă?ăă½Ĵă•£ăĴăăă?ăăăă?ă??ă?ăšă•?ă??



GPUã•16GBã®VRAMã??ã½¿ã•?ã??ã??æ??ã•?ã•§ã•?ã??æ¶?è²»é?»å??ã•æ??ã•§ã•§50Wã»ã•©ã•
§ã•?ã•®ã•§ã•??ã?»ã??è»¢ã•ã•?ã•?ã•?ã•?ã•§ã•ã•?ã•??ã•¼ã•?ã•??ã•??

```

ktsubaki@ZIKUU-PC-02: ~
Every 1.0s: nvidia-smi
ZIKUU-PC-02: Thu Jun 6 06:47:24 2024

+-----+
| NVIDIA-SMI 545.29.06                  Driver Version: 545.29.06   CUDA Version: 12.3   |
+-----+-----+
| GPU  Name                Persistence-M | Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp   Perf          Pwr:Usage/Cap |      Memory-Usage | GPU-Util  Compute M. |
|                                           | MIG M.         |
+-----+-----+
|  0  NVIDIA RTX A4000             Off | 00000000:01:00.0 On  |           Off        |
| 41%   57C    P2              48W / 140W | 15341MiB / 16376MiB |      1%      Default |
|                                           |                     | N/A              |
+-----+-----+

Processes:
+-----+
| GPU  GI   CI           PID   Type   Process name                      GPU Memory |
|      ID   ID                                   Usage      |
+-----+-----+
|  0   N/A  N/A         8963    G     /usr/lib/xorg/Xorg                 130MiB |
|  0   N/A  N/A         9093   C+G   ...libexec/gnome-remote-desktop-daemon 161MiB |
|  0   N/A  N/A         9142    G     /usr/bin/gnome-shell               259MiB |
|  0   N/A  N/A        10754    G     ...54,262144 --variations-seed-version 130MiB |
|  0   N/A  N/A        14202    G     gnome-control-center                3MiB |
|  0   N/A  N/A        32009    C     ...unners/cuda_v11/ollama_llama_server 14634MiB |
+-----+

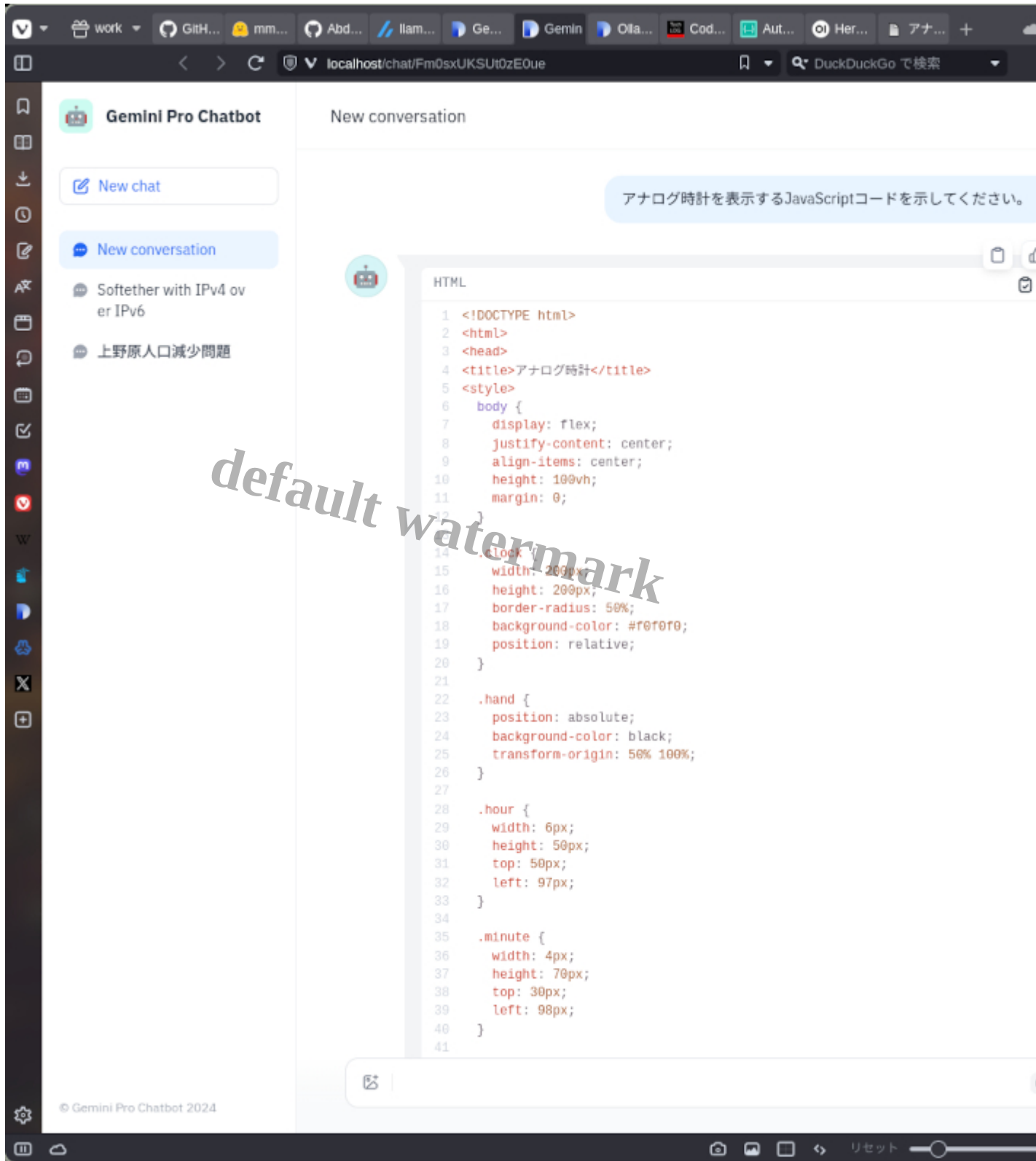
```

A4000ã¬æ¶¿è²»é?«å??ã?æ?§ã·ç?®ã·§è?·ä¹ æ?§ã·?ã·?ã??ã·¸·?ã·ã·®ã·?é å??ã·§ã·?ã·?ã·?ã·¸??ã
»ã·©é??ã·?GPUã·§ã·¬ã·ã·?ã·®ã·§ã·?ã·?ã·£ã·´??ã·?GPUã·´??ã·?CPUã·®çµ?ã·¿ã·?ã·?ã·?ã·ã·?ã·
ã·?ã·¶½¿ã·?ã·?æ??ã·?ã·«ã·ã·?ã·®ã·§ã·¬ã·ã·?ã·§ã·?ã·?ã·?ã·?ã·?ã·?ã·?ã·?ã·8bã· ã·
¬ã·¹ã·?ã·?ã·¬ã·¹ã·?ã·¾¼ã·£ã·?ã·¸·ã·?é«é?ã·ã·?ã·¬ã·¹ã·?ã·³ã·¹ã·?ã·¾¼ã·?ã·?ã·?ã·|ã·?ã·?ã·®ã·¾¼ã·¾¼ã·
§ã·?ã·½¿ã·?ã·?æ??ã·?ã·§ã·?ã·?

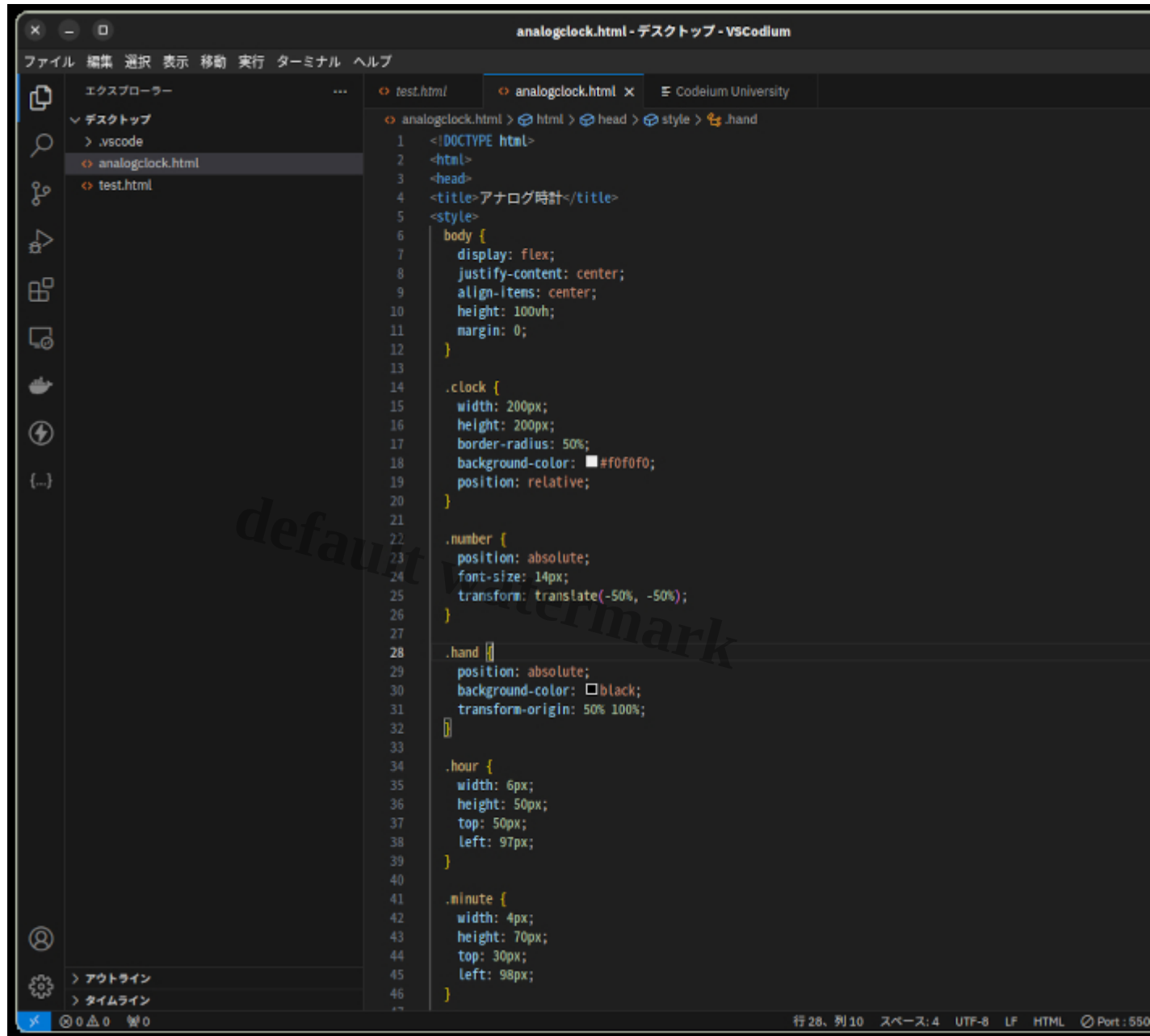
ã•jã•ªã•¿ã•«ã•?ã•??ã•-ã•??ã•¢ã•??ã•ã•°æ•??è•"?ã•??èj•çº°ã•?ã•??JavaScriptã•?³ã•?¼ã•??ã•??çº°ã•?ã•!ã••ã•
ã•?ã•?ã•??ã•?ã••ã•?ã•?è³ªã•?ã•??ã•?ã•?ã••ã••ã•®ã•??ã•®ã•§ã•?ã•??

â?ã?ã?ã?ã??8bă•Să??è©|ă?ă?¼ă?ă?ă?ă?ă?ă?ă?ă?ă?¼ă?¿ă??ă!ç¿?ă?ă|ă?ă??ă??ă?ă?ăSç??æ?
ă?ă??ă?ă?³ă?¼ă??ă-ă¼ă•£ă?ă•ă?ă?ăSă?ă?ă??

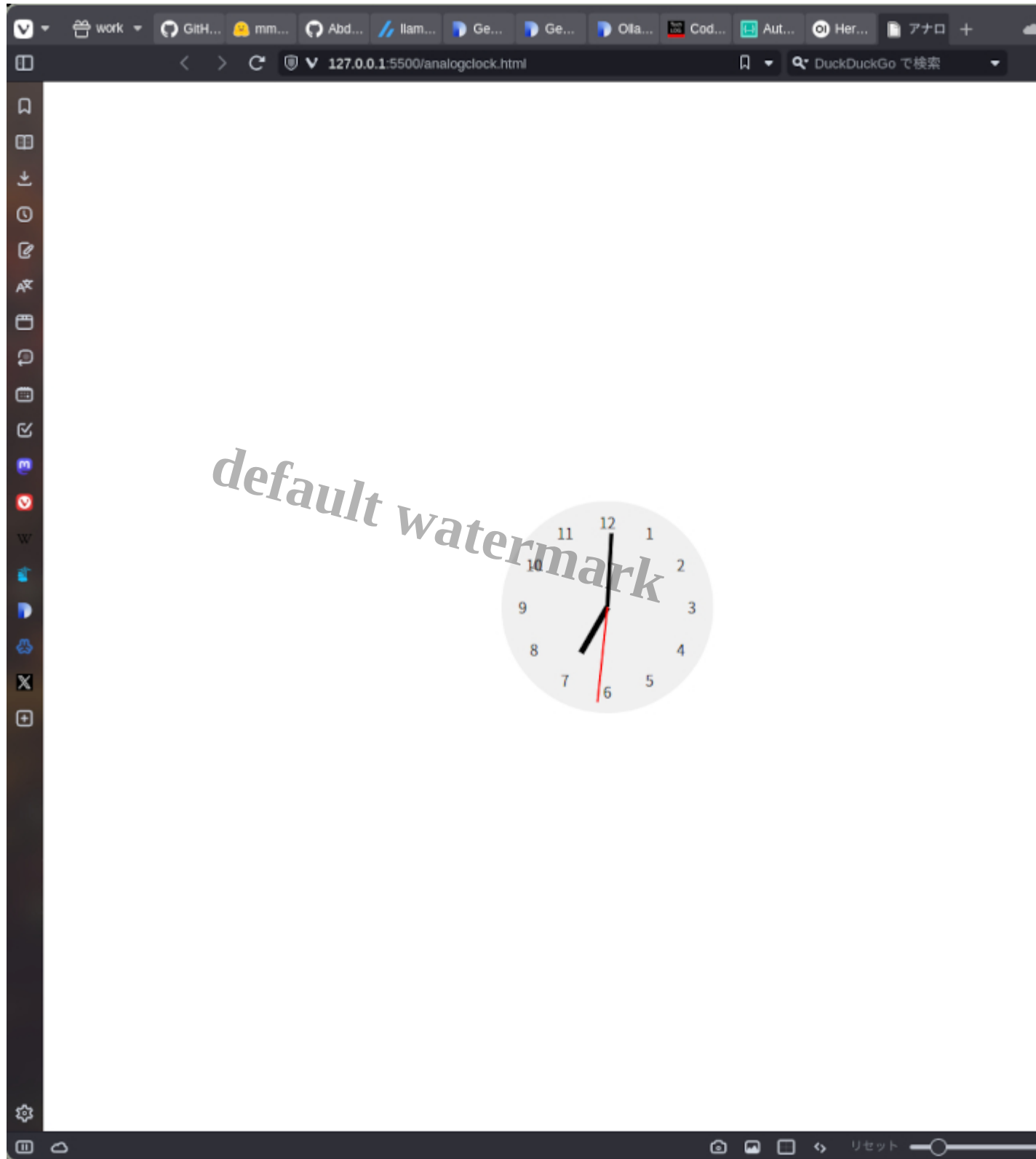
â•?â•?â•?â•?â•?Gemini 1.5 Proâ•?Sã??è©|ã•?ã•¼ã•?ã•?ã•?ã•?ã•?ã•-ã•?ã•?ã•¹ã•?ã•£ã•³ã•?ã•?ã•|ã•
?ã•?LLMã•?Difyã•?ã•?ã½¿ç?ã•?ã•ã•?ã•¾ã•?ã•?



ç??æ?•ã?•ã??ã?•ã?½ã?¼ã?¹ã?³ã?¼ã??ã??VS Codiumã«•ã?³ã??ã??ã?•ã?•lä_å-
?ã??ã?½ã?¼ã?¹ã?³ã?¼ã??ã«æ??ã-̂å ¥ã??ã?•ã?¾ã?•ã??ã??



ã•ã•?ã•|ã•@?èj?ã??æ£ã•?ã•ã•?£ã•??ã•ã•°æ??è""ã•?èj"ç°ã•?ã•??ã•|ã•??ã•ã•¼ã•?ã•??



Gemini 1.5 Proã•®å´å•?ã¬HTML+CSS+JavaScriptã?ä,?ä½?ã¬ã•£ã•?ã?ã¼ã??ã??ç??æ?•ã•?ã¬ã
•ã??ã¬Llama3ã•?ç??æ?•ã•?ã¬ã??ã•®ã??ã??èªå??ã¬Sã•?ã•?ã??

Llama3ã•?ç??æ?•ã•?ã¬ã?³ã?¼ã??ã¬ã¬CSSã?å«ã¼ã??ã¬ã•?ã¼ã•?ã??ã¬Sã•?ã•?ã¬ã•
ã??CSSã??å«ã¬ã¬ã¬ã¬ã•?ã•?ã¬ã¬ã¬ã•?ã¬ã??CSSã??å¬ã¬ã¬ã¼ã•?ã¬ã??é•ç?½ã•?ã¬ã¬ã
«æ??å?ã¬CSSã??å«ã¬ã¬æ¬²ã•?ã¬ã¬ã•?ã¬ã¬ã¬ã¬ã¬ã¬ã¬ã¬ã¬ã¬ã¬ã¬ã¬CSSç?jã•?ã¬

© 2024 Gemini 1.5 Pro. All rights reserved. Gemini 1.5 Pro is a trademark of Google LLC. Gemini 1.5 Pro is a large language model trained by Google.

Category

- æ?¥ã? ã•®æ´»å??
- ç?-ã??è´´?

Tags

- AI
- Dify
- Llama3
- LLM
- Ollama
- ä,?é??å??ä,?
- å±±æ¢ç´´ç??

Date Created

2024å¹´6æ??6æ?¥

Author

kazuo-tsubaki

default watermark