

Prompt

a wooden workshop with hand tools

Width

512

Height

512

Steps

20

Guidance Scale

7.5

Seed

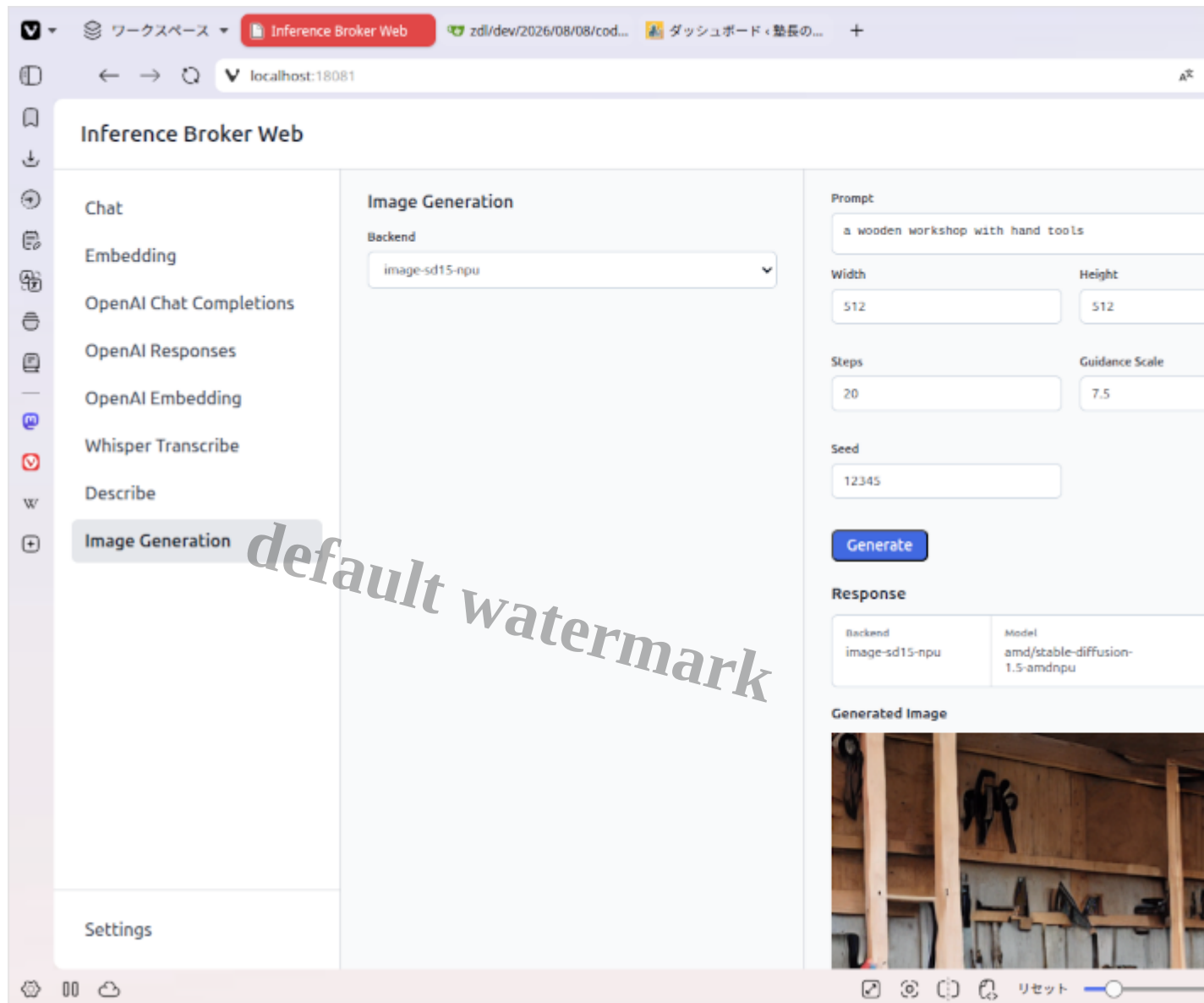
12345

Generate

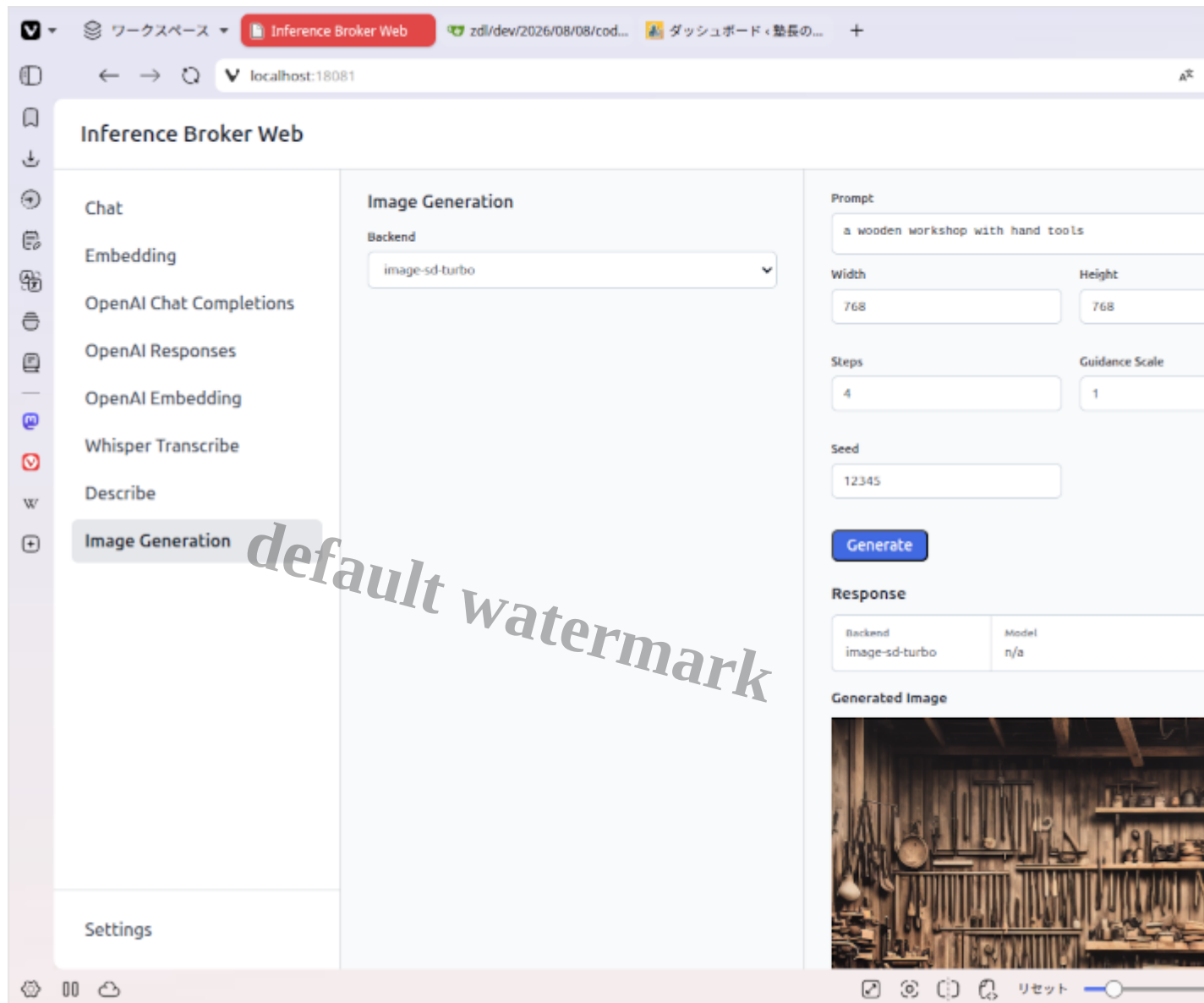
Inference Broker

Description

Inference Broker Web UI is a powerful tool for managing and optimizing AI inference tasks. It provides a comprehensive interface for configuring and monitoring various AI models, including those running on AMD GPUs and NPU accelerators. The tool supports a wide range of inference tasks, from image generation to natural language processing, and offers advanced features like batch processing and real-time monitoring. Inference Broker Web API is a powerful tool for managing and optimizing AI inference tasks. It provides a comprehensive interface for configuring and monitoring various AI models, including those running on AMD GPUs and NPU accelerators. The tool supports a wide range of inference tasks, from image generation to natural language processing, and offers advanced features like batch processing and real-time monitoring. Inference Broker Web API is a powerful tool for managing and optimizing AI inference tasks. It provides a comprehensive interface for configuring and monitoring various AI models, including those running on AMD GPUs and NPU accelerators. The tool supports a wide range of inference tasks, from image generation to natural language processing, and offers advanced features like batch processing and real-time monitoring.



AMD®Stable Diffusion
®ã?µã?°ã??æ??å@?ã•§ã•ã•¼ã•?ã??
ç??æ?•æ??æ?°ã??æ??å@?ã•§ã•ã??ã•®ã•§ã•?ã•?ã•?ã•®ã?ã??ã?ªã•§ã•1æ??ã§ã•?ã??



æp?è¼ç?»é•¢ã•-ã?ã½?ã?•ä,?ä,?äº?ã??ã?²ã•®ã?-ã?pã?¢ã?|ã??ã•sã•?ã•?ã•?ã?•ã½¿ã•?ã•«ã•ã•?ã•
£ã•?ã•®ã•sã•|ã•³äº?ã??ã?²ã?-ã?pã?¢ã?|ã??ã•«ãp?æ?´ã•?ã•|ã•?ã•¾ã•?ã??

Inference Broker LLM Broker ZIKUU Mini
 æ ¼ æ © ã ® ã ? ? ? PC ã § ã _ Codex ã ? ? Inference Broker æ ¥ ç ¶ ã ? ã ¾ ã ? ã ? ã ?

Codex CLã•®ã?ã³ã?¹ã??ã¼ã?«ã-ã?ã•®ã³ã??ã³ã??ã•šã??

```
npm install @openai/codex
```

ã•?ã•®è¨?äº?ã??æ?¿,ã•?ã•!ã•?ã??æ??ç?¿!ã•§ã•-ã•?0.150.1ã•¨ã•?ã•?ã•?ã•¼ã•¿,ã•?§ã•³ã•
?ã•?¤ã•³ã•¹ã•??ã•¼ã•?«ã•?ã•??ã•¾ã•?ã•?ã•??

ã?␣ã?³ã?¹ã??ã?¼ã?«ã?æ?•ã?ã?ã?ã?•

```
mkdir -p ~/.codex
```

ã•\$ã?•ã??ã?¼ã? ã??ã?£ã?-ã?ã??ã?ª«Codexã•®è`å®?ã??ã?jã?ªã?«ã??ã½?ã??ã•¾ã•?ã??

ã½?ã??ã•®ã•ã?

config.toml
inference-broker.config.toml

ã•®2ã•ªã??

config.toml

7èj?ç?®ã•® [model_providers.inference-broker] ã•® wire_api = â??responsesâ?•ã¾ã•\$å ¥å??ã•?ã??ã
•°OKã??



inference-broker.config.toml

终端

终端

终端

ktsubaki@ZIKUU-Mini-01 ~/workspace/inference-broker-web

% cat ~/.codex/inference-broker.config.toml

model_provider = "inference-broker"

model = "text-main"

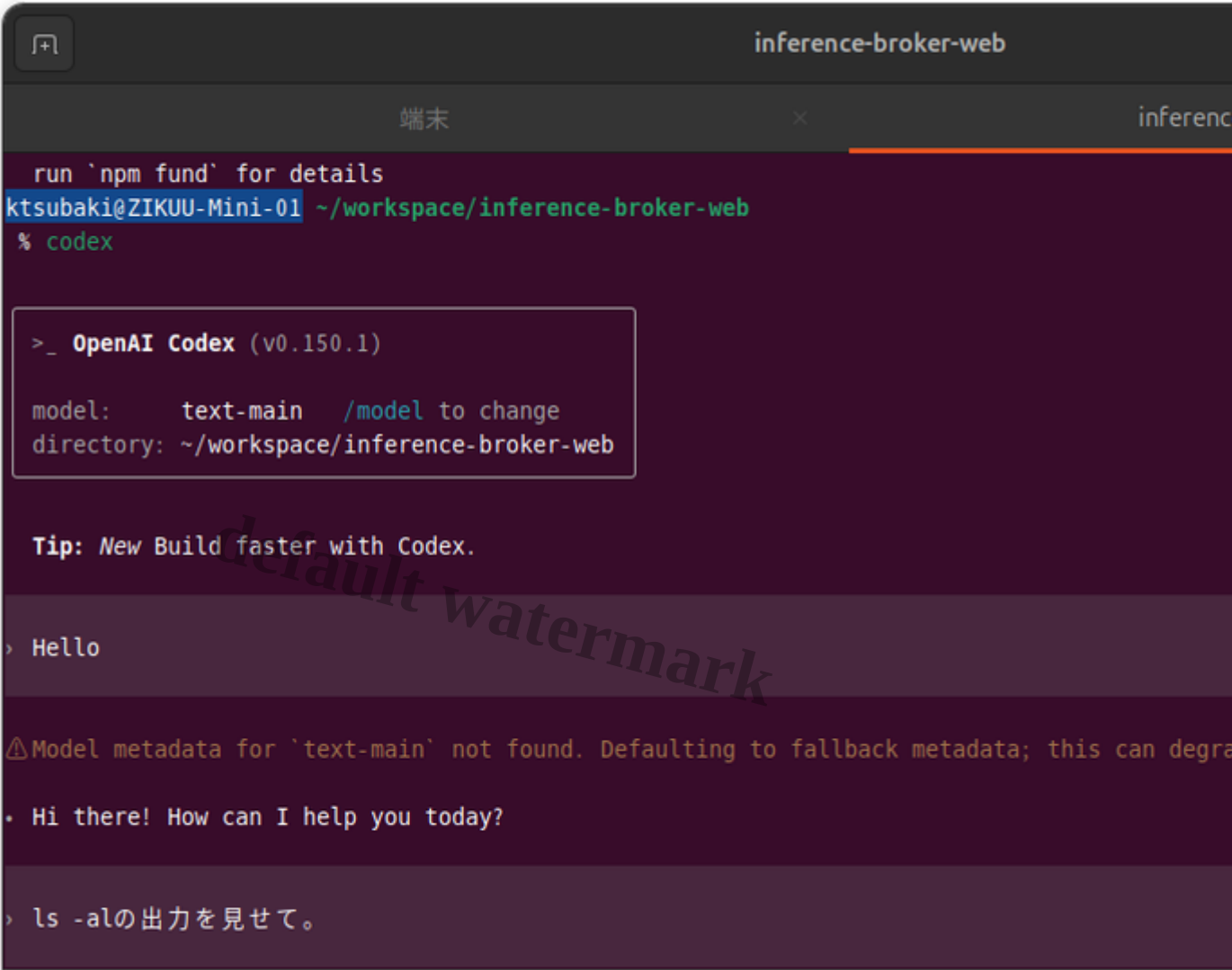
model_reasoning_effort = "medium"

ktsubaki@ZIKUU-Mini-01 ~/workspace/inference-broker-web

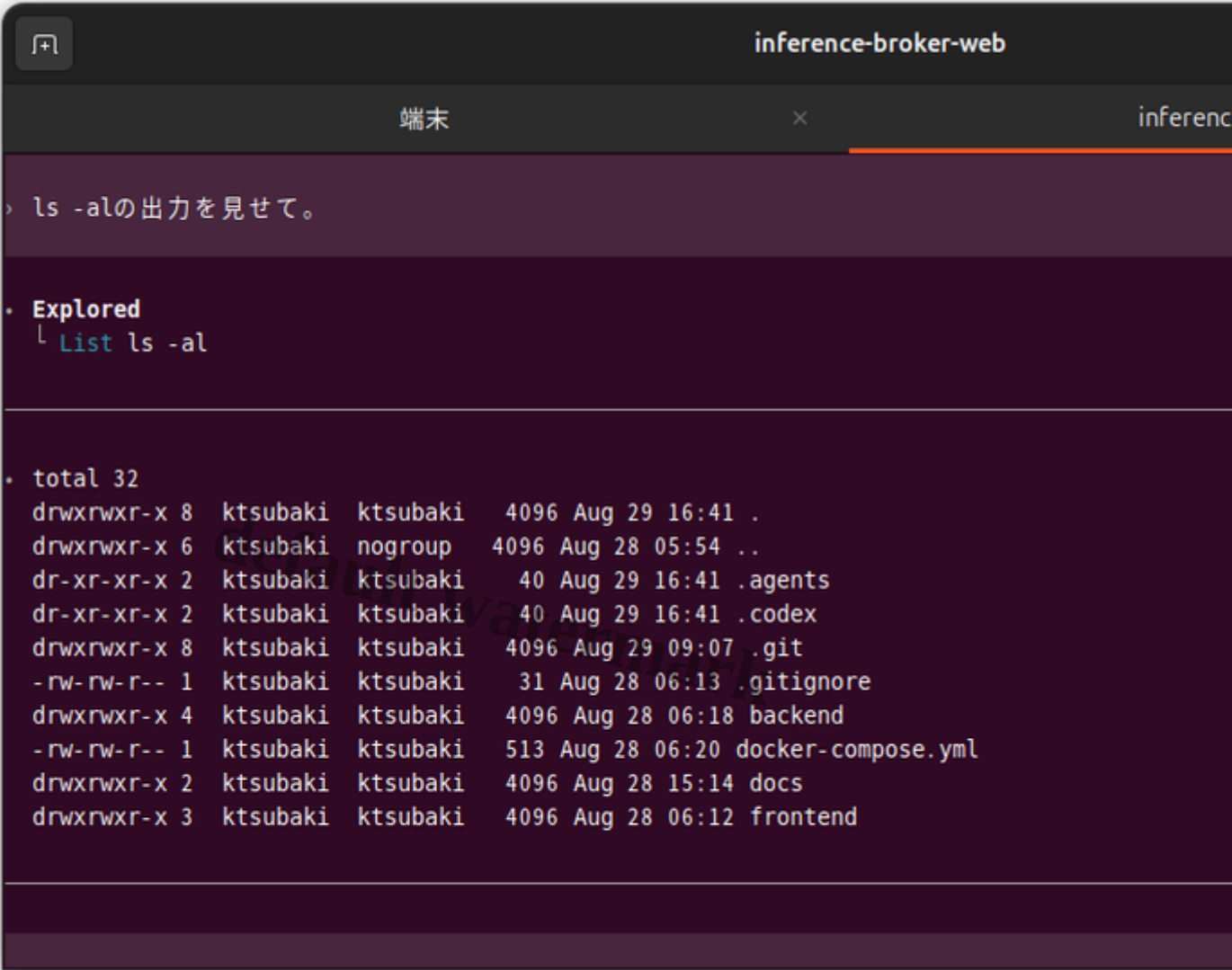
%

default watermark

2ã•♠ã•@ã??ã?jã?♠ã?«ã??ã½?æ?•ã•?ã•?ã??ã?•é•©å½?ã•ªã??ã?ã?,ã?§ã?¬ã??ã
®ã??ã?£ã?-ã?¬ã??ã?ªã«ç§»å??ã•?ã•| **Codex** ã??å®?èj?ã•?ã•|ç¢°èª•ã•?ã•¾ã•?ã??



Listã?³ã??ã?³ã??ã??æ£ã•?ã•â®?è;?ã•?ã??ã•!ã•?ã•¾ã•?ã??



ã•?ã??ã•\$ã?•

ZIKUUã?¢ã??ã?ª/Codex

â??

Inference Broker

â??

llama.cpp/Ollama/Lemonade/Ryzen AI SD

ã•ã•?ã•?çµ?è•¬ã?å@?æ?•ã•?ã•¼ã•?ã•?ã??

ã•?ã??ã•\$ã?•èµ?æ?°ã•@GPUã?µã?¼ã?•ã?¼ã•«LLM Runtime Agentã??é•ç½@ã•?ã??ã•°ã?•
ã?¢ã??ã?ªã?±ã?¼ã?•ã?\$ã?³å´ã•¬ã?ã•©ã?ã•\$ã•©ã??ã?ªæ?¬è«?è³?æ°•ã?å??ã?ã?|ã?ã??ã?•?ã??æ?
è?ã•?ã?ã•«å??ã?ã?¼ã•?ã??

Category

1. æ?¥ã?ã•@æ´»å??

Tags

- 1. AI
- 2. Codex
- 3. Inference Broker
- 4. Lemonade
- 5. llama.cpp
- 6. LLM
- 7. LLM Broker
- 8. Ollama
- 9. Ryzen AI
- 10. Stable Diffusion
- 11. ä,?é??å??å,?
- 12. å±±æ€ˆç??

Date Created
2026å¹¸æ??29æ?¥
Author
kazuotsubaki

default watermark