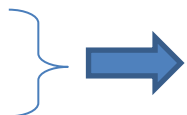


はじめに

生成AI(LLM：大規模言語モデル)を利用する上での主な留意点

- ・機密情報の扱い
 - ー クラウド等にデータを上げてよいか
 - ー 学習に利用されてしまわないか
- ・ハルシネーション
- ・特定情報を反映させること
- ・他者の権利を侵害してしまわないか



ローカル生成AI(LLM や SLM)

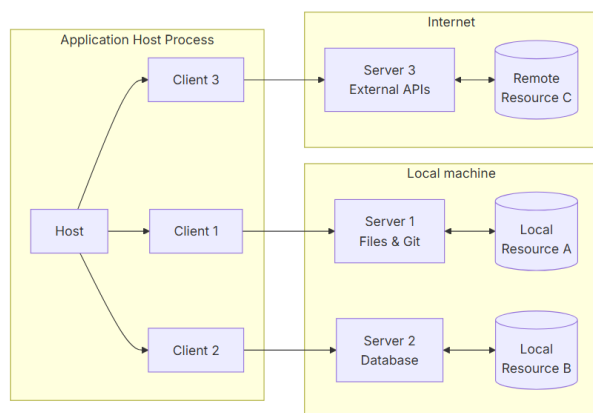
- ・生成AIモデルのチューニング
- ・生成AI(LLM)と外部データとの連携
RAG(拡張検索生成)や MCP

ローカル生成AIとSLM(小規模言語モデル)

- ・Ollama や LM Studio などの無償ソフトもある
上記の無償ソフトはチャット用GUIを備える
- ・3B程度の phi4-mini や Llama, 7B 程度の Granite4-H-tiny などの SLM も選択可

MCPによる外部データとの連携

- ・MCP(Model Context Protocol)
Anthropic社が提案したオープンプロトコル
- ・一方、RAG(拡張検索生成)は、DBの検索結果を加味したプロンプトを生成AIに入力する



MCPのアーキテクチャ

<https://modelcontextprotocol.io/specification/2025-06-18/architecture>

All in OneツールによるローカルSLMとMCP

- ・AIモデルの選択・切替, MCPによる連携設定, チャット用GUIなどを統合したツール
- ・LM Studio, AnythingLLM, Dive などが候補
- ・MCPを使うには, MCP 対応の生成AIモデルが必要
SLM では phi4-mini(3.8B), Llama3.2(3B), granite4-H-tiny(7B)など
- ・mcp-server-time を使い、「現在の時刻を教えてください」という内容のプロンプトで時刻の取得を試みる

➡ PC(Core i7/メモリ32GB/Intel UHD Graphics)でローカルSLMとMCPの実行ができた

現在の時刻を教えてください。日本語をお願いします。

📄 コピー 📝 編集

▶ ✖ get_current_time ツール呼び出し結果

現在の日本時間は **2025年10月9日 (木) 6:15** です。

Diveでの実行結果

(granite4-H-tiny を Ollama経由で使用)

現在の時刻を教えてください。日本語をお願いします。

llama-3.2-3b-instruct

🔗 get_current_time

mcp/time >

The current time in Japan is 2025-10-10T12:36:19+09:00, which is not daylight saving time.

LM Studioでの実行結果(Lama3.2使用)

問い合わせ先

事業化支援部橋渡し支援課

TEL 046-236-1500