

ローカルLLMと「情報を外に出さない」設計

社内PCで動かすAIの向き・不向きと、クラウドのAIとの使い分け

更新日：2026-09-30（2026年9月時点の情報） | HUMAC 伊藤晴通

ローカルLLMとは

インターネットの向こうの会社のAIではなく、**自社のPC・サーバーの中で動かすAI**。入力した情報が社外に出ず、使うたびにお金がかかる仕組み（従量課金）でもない。その代わり、性能・設備・手間の見きわめが必要になる。

クラウドのAIとの違い

観点	クラウドのAI（ChatGPT・Gemini・Claude）	ローカルLLM
賢さ	最先端	一段落ちる（要約・分類・下書きは十分）
情報の行き先	社外のサーバーに送られる（契約で保護）	社内から出ない
費用	月額または使った分だけ	最初に機械代、あとは電気代
Web検索	できる	そのままではできない（決まったページを取りに行くことはできる）
手間	契約するだけ	設置・更新・故障対応が要る
向いている仕事	調査、長文の読解、難しい判断、アプリづくり	毎日の定型処理、社外に出せない情報の要約・分類、夜間の自動処理

「使い分け」が正解：講師の2台体制

人が指示するPC

クラウドのAI（Claude Code など）で、設計・調査・資料づくり・アプリ開発。賢さが要る仕事はこちら。

24時間動くPC（GPU搭載）

ローカルLLM・資料室（RAG）・定時の自動処理。夜中や休日も、PCを閉じていても動く。秘密は専用の金庫から読み、コードに書かない。

ローカルLLMに任せている定型の仕事（例）

仕事	頻度	ローカルLLMがやること	人がやること
朝のまとめ（予定・タスク）	毎朝	集めて整えてメールで届ける	読む
経理書類の仕分け	毎日	スキャンした書類を読み、種類を判定して分類・名前付け	迷った書類の最終判断（AIは消さない）
AIの現在地チェック（この資料ライブラリ）	毎週土曜	公式ページの変化とニュースを要約し、【大】【小】の目安を付けて知らせる	原文で確かめ、資料を直すかを決める（→ C3）

ローカルLLMは「知らせる係」から始める

中型のローカルLLMは、判定を読み違えることがある。実際、資料ライブラリの週次チェックでも、差分を誤読した要約が出た。**最初は「見つけて知らせる」まで**を任せ、書き換え・公開・削除は人（またはクラウドの高性能AI＋人の確認）が行う。慣れてきたら範囲を広げる。

始めるときの目安

- **機械**：GPU（画像処理用の部品）を積んだPC。中型のモデル（140億～320億パラメータ級）で、日本語の要約・分類が実用になる
- **道具**：モデルを動かすソフト（Ollama など）＋資料室（RAGエンジン）＋窓口になる小さなサーバー
- **最初の一步**：社外に出せない定型の仕事を1つ選び、「要約してメールで知らせる」だけを自動化する
- **情報の扱い**：ローカルでも、鍵の管理と権限の絞り込み（→ E1）は同じように必要