

NVIDIA RTX Sparkが拓く「完全ローカル知財AI」の実践ガイド

RTX Sparkが知財業務に変革をもたらす理由

128GB統合メモリで70Bモデルをローカル完全稼働



未公開特許や契約書など外部送信NGな超機密データを端末内で安全に処理。

1台約90万円(US\$5,999)から導入できる現実的コスト



数千万円規模のデータセンターサーバーと異なり、部門単位で手軽に調達可能。

日常作業は8~14B、高難度案件は70Bの階層運用が最適



メモリ帯域を考慮し、高速な小規模モデルと高品質な大型モデルを使い分け。



知財実務における最適ユースケースとアーキテクチャ

文章生成より「クレーム・明細書の整合性検査」で高ROI



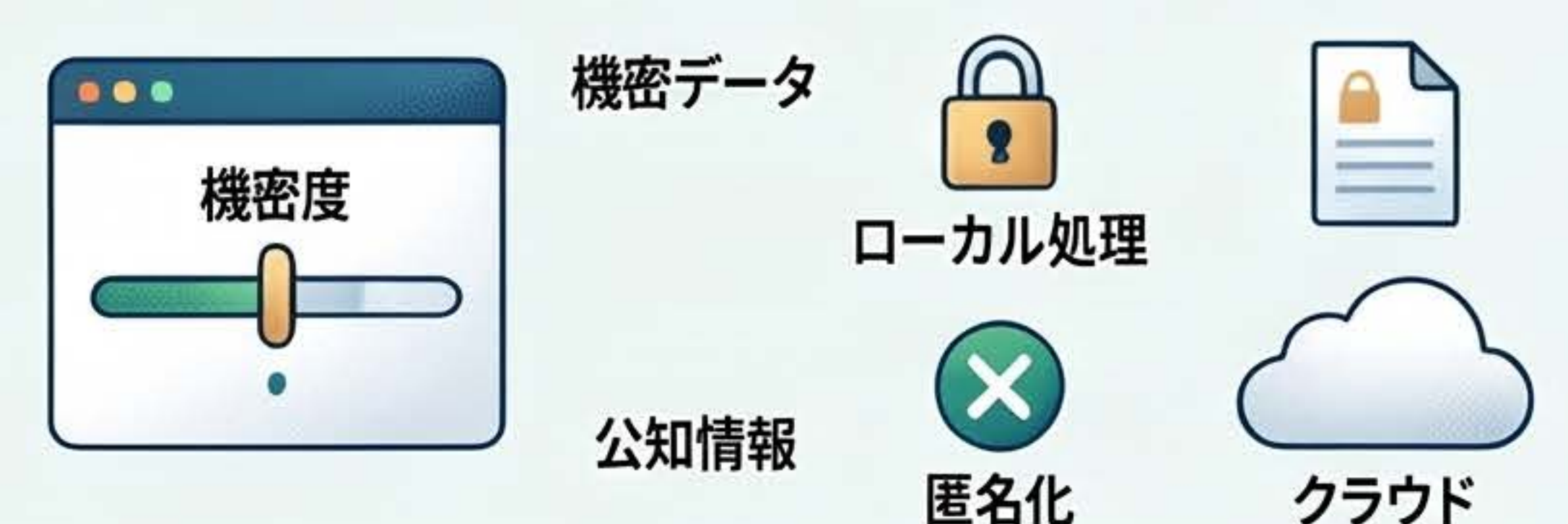
符号誤りや用語不一致など、正解が文書内にあるメカニカルな検証で真偽を発掘。

「LLM単体検索」を避け、Local RAG+引用検証を徹底



ハルシネーションによる存在しない特許番号の生成を防ぐため、原文照合を必須化。

AI Gatewayによる機密度に応じた自動ルーティング



機密データはローカルで処理し、公開情報のみ匿名化してクラウドへ安全に振り分け。