製造業における「ローカルAI」活用の最前線：機密保護とコスト最適化の両立

データ機密性の確保
社外に非開示で適用可能。

推論コストの抑制
定常処理を固定費型へ移行。

小型モデルの成熟
1GPU動作SLMで実現。



ローカルAI
（オンプレ/エッジ）

ローカルとクラウドの最適な使い分け

機密度による分離
未公開の技術情報は閾域処理。

処理量と応答性
即時処理は現場型が有利。

クラウドAI

少量・非定型・高度な推論処理



クラウドAI

- 文書検索
- 高度推論・公開情報活用

段階的な適用拡大

文書検索から → 設備保全へ拡張 → 全社的DX推進へ

オンプレミス・エッジ環境でのAI採用が本格化

業務要件に応じたローカルAIとクラウドAIの判断軸

	ローカルAI（オンプレ/エッジ）	クラウドAI
データの機密度	図面・原価・未公開発明	公開情報の調査・一般的文面作成
処理特性と応答	大量・定型処理・リアルタイム応答	少量・非定型・高度な推論処理
費用モデル	固定費型（自社設備・パッケージ）	従量課金型（利用トークン数依存）

Gemini Notebook