

Grounded AIの進化と 視覚的知識創造

Google NotebookLMにおける「Nano Banana Pro」の実装と実務適用性の評価

Report Context: 2026年 AIトレンド分析

Category: Academic / Authority

Key Technology: Gemini / Nano Banana Pro

情報の抽出から、 信頼できる 「知識の創造」へ



2026年現在、生成AIの役割は根本的なシフトを迎えています。従来の「要約・抽出」機能を超え、信頼性を担保した「視覚的な知識創造（Knowledge Creation）」のフェーズへと移行しました。

The Challenge

従来のLLMが抱えていた「ハルシネーション（幻覚）」のリスク。

The Solution

NotebookLMによる「Grounded AI」アプローチ。外部データではなく、ユーザーが提供した特定のソース（グラウンドデータ）のみを根拠に回答を生成することで、教育や高度な業務現場での信頼性を獲得しています。

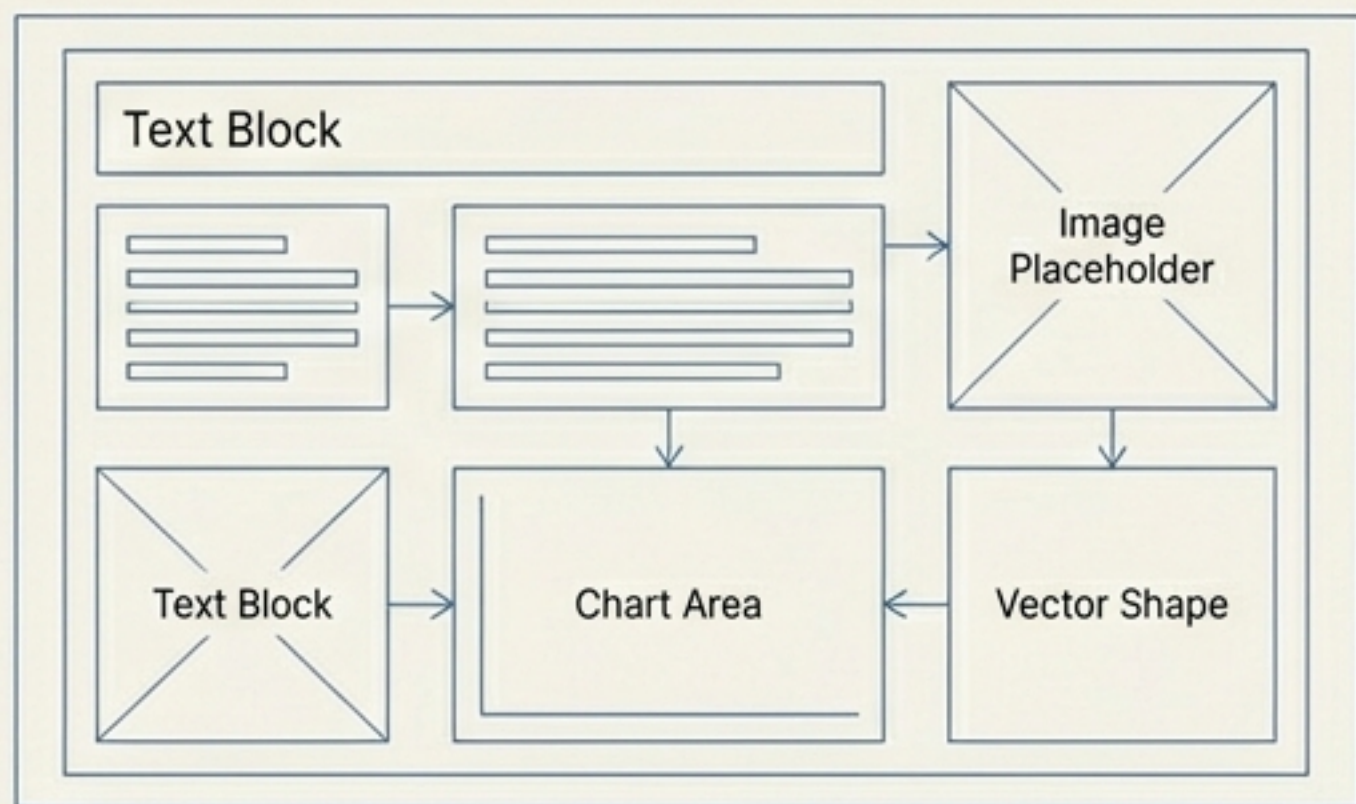
The Milestone

2025年11月、大規模言語モデルと画像生成モデルの統合により、単なるテキスト出力ではない、完全な「スライドデッキ」および「インフォグラフィック」の自動生成が実現しました。

方法論：画家型アプローチとNano Banana Pro

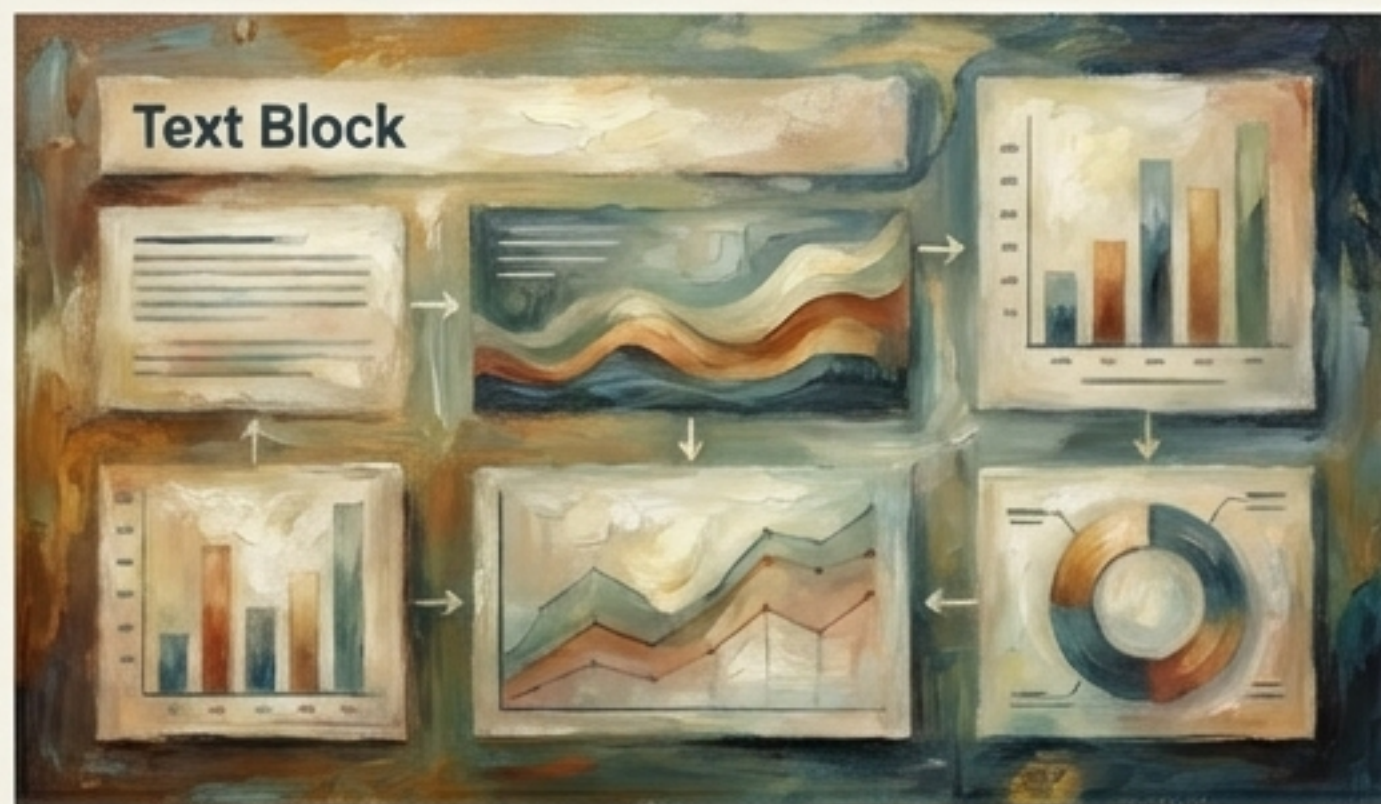
Geminiによる論理構成と、画像生成モデルによる視覚表現のパイプライン処理

The Architect (Classic Gemini)



編集容易 / オブジェクト配置 / 標準的デザイン

The Painter (NotebookLM)



編集困難 / 一枚の絵として描画 / 高度なデザイン

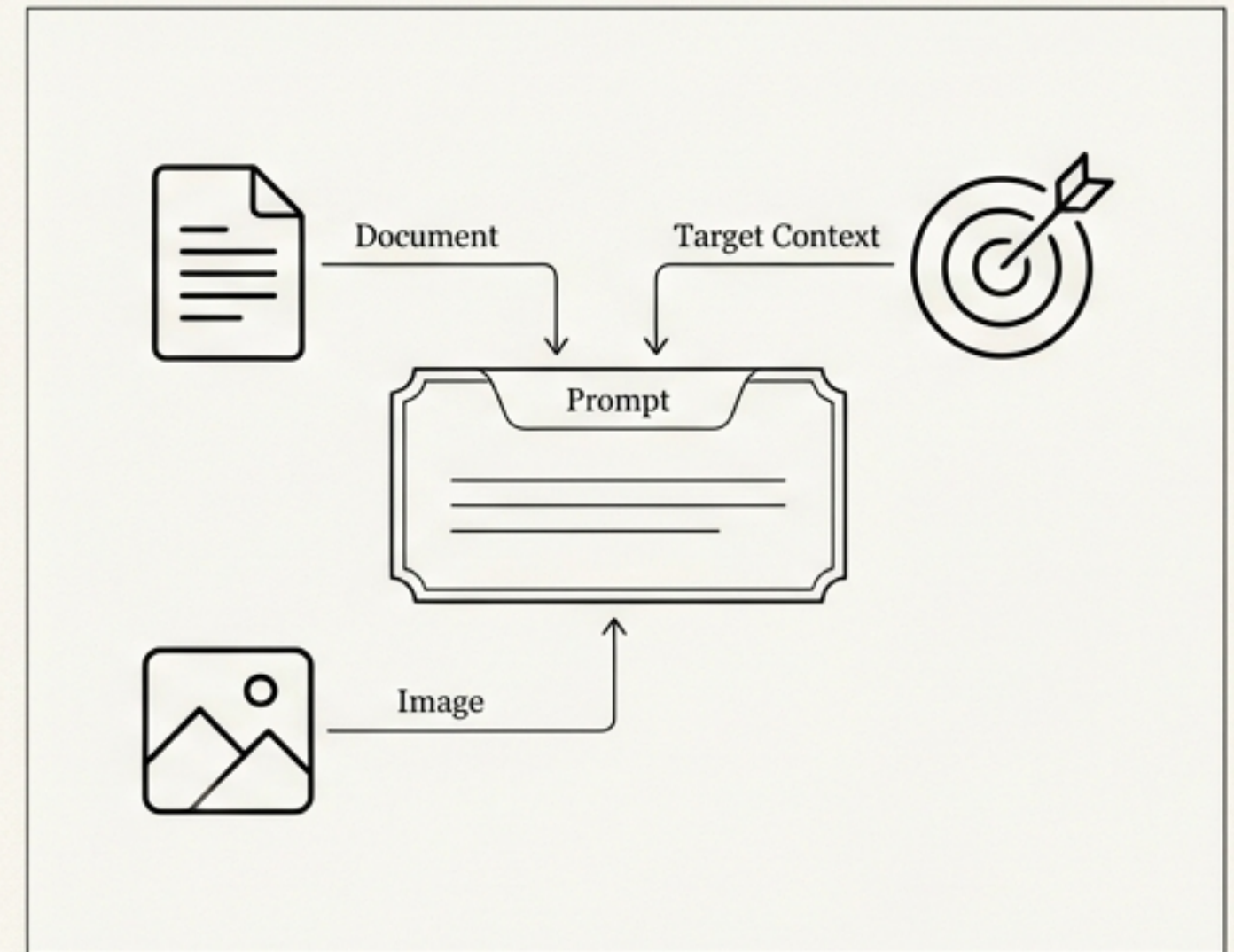
NotebookLMのスライド生成は、テキストボックスを積み上げる「建築家 (Architect)」型ではなく、Nano Banana Proを用いて一枚の絵として情報を描く「画家 (Painter)」型のアプローチを採用しています。

制御：コンテキストの明示とプロンプト設計

「画家」型の生成モデルを制御するためには、曖昧さを排除した明確な指示（ディレクション）が不可欠です。

01 / Contextual Definition

「誰に向けて」「何のために」というターゲットとゴールを明示する。「部署別・シーン別プロンプト」を適用することで、生成される構成の精度が劇的に向上します。



02 / Input Modality

テキストやPDFに加え、2025年12月より画像入力に対応。コンテキストウィンドウは100万トークンまで拡張されており、膨大な資料背景をAIに理解させることが可能です。

一貫性の維持：キャラクターシートの活用



The Method

単一の画像ではなく、複数の表情や角度を含む「キャラクターシート（Character Sheet）」を入力ソースとして使用する。

The Command

「exact same appearance」や「lock reference」といった強力な固定プロンプトを使用し、AIに対して参照元を厳密に守らせる。

Constraints

Googleの安全ポリシーにより、実写の人物（顔写真）はブロックされる可能性があるため、非人間キャラクターやシルエット化などの回避策が推奨されます。

実証結果：文脈維持とストーリーテリング

14 / 15

Slides with Consistent Character Design

自作キャラクターを用いた生成実験において、全15スライド中14枚でキャラクターのデザイン（形状、パーツ配置）の一貫性が維持されました。

- Speed: スライドデッキ生成時間は約5～10分
- Comparison: インフォグラフィック機能と比較して、連続した文脈維持において明確な優位性を確認
- Deep Research: Web調査機能と連携し、情報の収集から構成案作成までを自動化

課題：「編集不可」のパラドックス



The Dilemma

NotebookLMの出力は、Nano Banana Proによって描かれた「一枚の絵（PDF）」です。これは美的な完成度が高い反面、テキストの修正やレイアウトの微調整ができないというジレンマを抱えています。

Implication

「綺麗な絵画には手を入れることができない」。ビジネス現場で頻頻発する修正要求に対して、ネイティブな機能だけでは対応できないという壁が存在します。

解決策：ポストプロセス・エコシステム

PDF出力の制約を突破する3つの技術的アプローチ



01 Kirigami

画像化されたスライドから文字認識（OCR）とオブジェクト分離を行い、編集可能なPPTXファイルを再構築する技術。画像の文字潰れ補正も行います。

02 Codia (NoteSlide)

レイアウト保持精度の高い変換を行う専門ツール（クレジット制モデル）。

03 Gemini Canvas

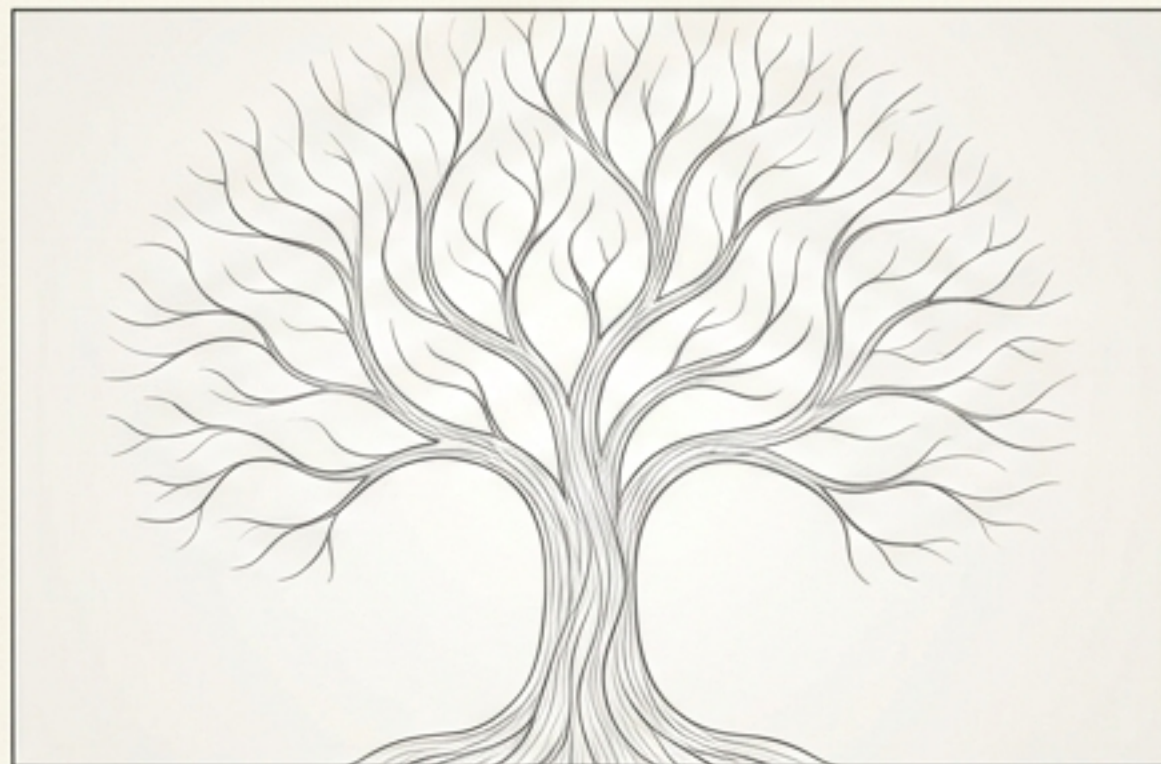
生成されたPDFをYAML形式で構造化し、Google Slidesとしてコードベースで再構成させる手法。

展望：2026年の技術的地平

Gemini 2.5 Pro & Flash Thinking

AIが回答に至るまでの推論過程（論理展開、反証検証）をユーザーに表示する

「Flash Thinking」により、ブラックボックス化が解消されます。



Expanded Inputs

会議音声からの議事録生成や、動画コンテンツのタイムスタンプ付き解析など、入力ソースの多様化が予測されます。

Collaboration

個人の知的生産ツールから、Google Workspace（Drive, Calendar等）と完全統合されたチームコラボレーション機能への進化が期待されます。

結論：オペレーターから 「監督（Director）」へ

NotebookLMは単なる「資料を読むツール」から、企業の「知的生産インフラ」へと進化しました。人間は「0からの作成」を行う作業者ではなく、AI生成物を導く「監督」に役割をシフトすべきです。正確性（Grounding）と視覚能力（Nano Banana Pro）の結合は工数を劇的に削減しますが、その品質を決定づけるのは人間の「プロンプトエンジニアリング」と「情報キュレーション」能力です。