

写真・音声・PDFを、 AIに一度に任せる

マルチモーダルAIの仕組みと
中小企業での活用イメージ

2026年7月16日発行

株式会社DnD（Distill n Develop）

この資料の要点

- マルチモーダルAIは、テキストだけでなく画像・音声・PDFなど複数の種類の情報を同時に処理できるAI。
- 現場写真からの報告書生成・音声メモの整理・帳票読み取りなど、現場の「書類仕事」を大幅に短縮できる。
- 2026年には無料から使えるマルチモーダルAIが複数あり、小さな業務から試せる段階に来ている。

マルチモーダルAIとは何か

マルチモーダルAIとは、テキスト（文字）だけでなく、画像・音声・動画・PDFなど複数の種類の情報（＝モダリティ）を同時に入力・処理できるAIのことです。「マルチ（複数）＋モーダル（感覚・入力の種類）」という意味から来ており、人間が目で見、耳で聞いて、文字を読んで総合的に判断するのと近い処理をAIが担います。

従来の生成AIはテキストの入出力に特化していましたが、2024年以降、主要なAIサービスがマルチモーダル対応を標準化しました。「写真を見せて説明させる」「音声を聞かせて文字に起こす」「PDFを読み込んで要点を答えさせる」といった操作を、一つのAIシステムで行えます。

従来のAI活用と何が変わるのか

これまでのAI活用では、次のような「変換→処理」のステップが必要でした。

- OCRで画像から文字を取り出す → テキストをAIに渡す → AIが文章を生成する
- 音声認識ツールで録音を文字起こしする → テキストをAIに渡す → AIが要約する

マルチモーダルAIは、この変換ステップを内側に取り込んでいます。写真をそのまま渡すだけで「写真に基づいた報告書」を生成できる——変換の手間がなくなることが最大の変化です。

中小企業で使いやすい3つの場面

① 現場写真から報告書・点検記録を下書きする

建設・設備管理・不動産・営業など、現場で写真を撮る業務では、後から写真を見ながら報告書を書く作業が発生しがちです。マルチモーダルAIに「この写真を見て、設備点検報告書の形式で記録してください」と指示すると、写真の内容を読み取って記録文の下書きを生成します。

たとえば「外壁のひび割れ写真」を渡して「点検記録の形式で記述してください」と伝えるだけで、場所・状態・推奨対応などの項目を埋めた文章が出力されます。人が写真を見て最初から書く時間を、確認・修正の時間に置き換えられます。

② 音声メモ・録音をその場で整理する

打ち合わせ中に録音したメモや、移動中に吹き込んだ音声メモをマルチモーダルAIに渡すと、文字起こしと要点整理を同時に行えます。従来は「録音→文字起こしツール→生成AIで要約」と分かれていた3ステップが、一度の操作に集約されます。

③ 帳票・書類からデータを読み取って整理する

注文書・納品書・名刺・見積書など、紙やPDFの書類から必要な情報を取り出してスプレッドシートに整理する作業も、マルチモーダルAIが得意とする領域です。書類のレイアウトが変わっても「金額」「日付」「品名」を文脈から判断して取り出せます。

2026年時点で使えるツールの選び方

- **ChatGPT (GPT-4o)** : 画像・音声・ファイルを送信できる。無料プランでも画像入力に対応しており、入口として試しやすい。
- **Claude (Anthropic)** : 長い文書や複数ファイルを一度に処理できる。日本語精度が高く、帳票や報告書の処理に向く。
- **Gemini (Google)** : Google DriveやGmailとの連携が強み。G Suite利用企業には特に試しやすい環境。

まず「今いちばん手間のかかっている作業」を一つ選んで、無料プランで試してみることが最短の進め方です。

気をつけたい3つのポイント

- **機密情報・個人情報の取り扱い** : クラウド型サービスを使う場合、各サービスのデータポリシーを確認したうえで社内ルールを定めましょう。
- **出力の精度と確認フロー** : 「AIが下書きを作り、人が確認・修正して完成させる」という分担が、現状では最も安全で効果的です。
- **コストと使用量の管理** : 画像処理はテキストよりAPIコストが高くなります。使用量を把握しながら少しずつ対象業務を広げましょう。

まとめ

「下書きをAIに任せて、人が仕上げる」という分担から始めるだけで、現場の書類仕事は大幅に軽くなります。

まず手間のかかっている一つの作業で試してみingことをお勧めします。うまくいったら横展開する——その繰り返しが、AI活用の定着への近道です。

CONTACT

株式会社DnDでは、マルチモーダルAIを含む業務効率化の仕組みづくりを、要件定義から構築・運用まで一貫してサポートしています。

ご相談・お見積りは無料です。 <https://dnd-inc.jp/contact.html>