

PDFの種類

PDFの基礎知識

PDFとは？

Portable Document Format（ポータブル・ドキュメント・フォーマット）

 どんな環境でも同じように表示できる文書形式

- ✓ OSや端末が違って同じレイアウトで表示
- ✓ 印刷したときも同じ見た目
- ✓ ビジネスシーンで広く使われている

実は、PDFには3つの種類がある！



テキストベース

デジタルPDF



画像ベース

スキャンPDF



ハイブリッド

OCR処理済み

① テキストベースPDF（デジタルPDF）




最も扱いやすいPDF


作成方法

- Word、Excel、PowerPointから「PDFとして保存」
- Webページを「PDFで印刷」
- デザインソフトから直接出力

特徴

 テキスト選択・コピー可能

 ファイルサイズ小

 文字検索が可能

 変換・編集が容易

② 画像ベースPDF（スキャンPDF）



実は「画像」のPDF

作成方法

- 紙の書類をスキャナーで読み取り
- スマホで撮影してPDF化
- FAXを電子化

特徴

- ✗ テキスト選択・コピー不可
- ✗ 文字検索できない
- ⚠ ファイルサイズが大きい
- ⚠ Word等への変換にOCRが必要

💡 見分け方：PDFを開いて文字を選択してみる

→ 選択できない場合は画像ベースPDF

③ ハイブリッドPDF（OCR処理済み）



いいとこ取りのPDF

仕組み

スキャン画像 + OCR（光学文字認識）で抽出したテキストレイヤー

構造イメージ

上層：見えないテキストレイヤー（検索・選択可能）

下層：スキャン画像（見た目）

メリット

✓ 元の見た目を保持

✓ テキスト選択・コピー可能

✓ テキスト検索可能

✓ 変換が容易

3種類の比較表

項目	テキストベース	画像ベース	ハイブリッド
テキスト選択	✓ 可能	× 不可	✓ 可能
文字検索	✓ 可能	× 不可	✓ 可能
ファイルサイズ	小	大	中～大
変換の容易さ	◎ 簡単	△ OCR必要	◎ 簡単
作成方法	アプリから直接保存	スキャン・撮影	スキャン+OCR

OCR（光学文字認識）とは？

Optical Character Recognition

画像の中の文字を認識して、テキストデータに変換する技術



OCRツールの例

- Adobe Acrobat Pro（有料）
- Google Drive（無料・精度高い）
- Microsoft Word（無料・Office利用者）
- オンラインサービス（Smallpdf、ILovePDFなど）

実務での使い分け

テキストベースPDFを使うべき場面

- 報告書、提案書、契約書などの文書作成
- データの検索・抽出が必要な場合
- ファイルサイズを抑えたい場合

画像ベースPDFになるケース

- 手書き書類のアーカイブ
- 署名・押印入りの原本保存
- 図面や設計図の保存

→ 可能ならOCR処理してハイブリッド化を推奨

ハイブリッドPDFの活用

- スキャンした書類を検索可能にする
- アーカイブの利便性向上
- 既存の画像PDFを活用しやすくする

画像PDFをWord等に変換する方法

🌟 おすすめ：Google Driveを使う方法（無料）

1. PDFをGoogle Driveにアップロード
2. 右クリック → 「アプリで開く」 → 「Googleドキュメント」
3. 自動的にOCR処理される
4. 「ファイル」 → 「ダウンロード」 → 「Microsoft Word」

その他の方法

- Adobe Acrobat Pro
- Microsoft Word（挿入機能）
- Smallpdf、iLovePDF

⚠️ 注意点

- OCR精度は100%ではない
- 複雑なレイアウトは崩れる
- 変換後は必ず確認を

まとめ

覚えておきたいポイント

- 1 PDFには**テキストベース**、**画像ベース**、**ハイブリッド**の3種類がある
- 2 見分け方：**文字が選択できるか試してみる**
- 3 画像PDFは**OCR処理**でテキスト化・検索可能にできる
- 4 可能な限り**テキストベースPDF**で作成するのがベスト
- 5 スキャンPDFは**Google Drive**で簡単にOCR処理できる