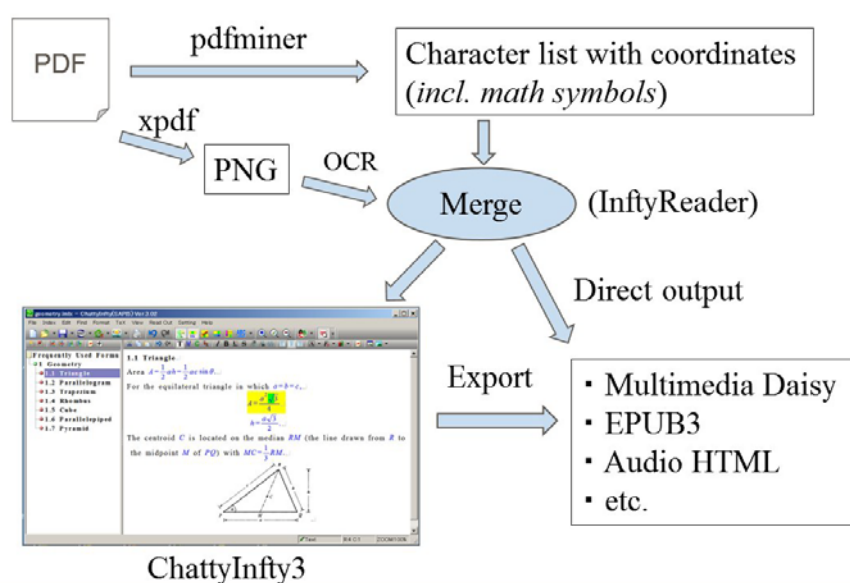


# InftyReader Ver. 3.1 の新しい機能

## 1. PDF の認識率が向上

InftyReader Ver.3.1 では [pdftminer](#) というソフトを用いて、PDF から直接文字や記号のコードと位置座標を抽出して認識に利用します。PdfminerとOCR, InftyReader による構造解析の統合により、PDF の認識率が、通常のOCR ソフトや従来のバージョンの InftyReader に比べて飛躍的に認識率が向上しています。



**(注 1)** 上図は英文の数学文書ですが、日本語の PDF でも同様に認識率が向上しています。

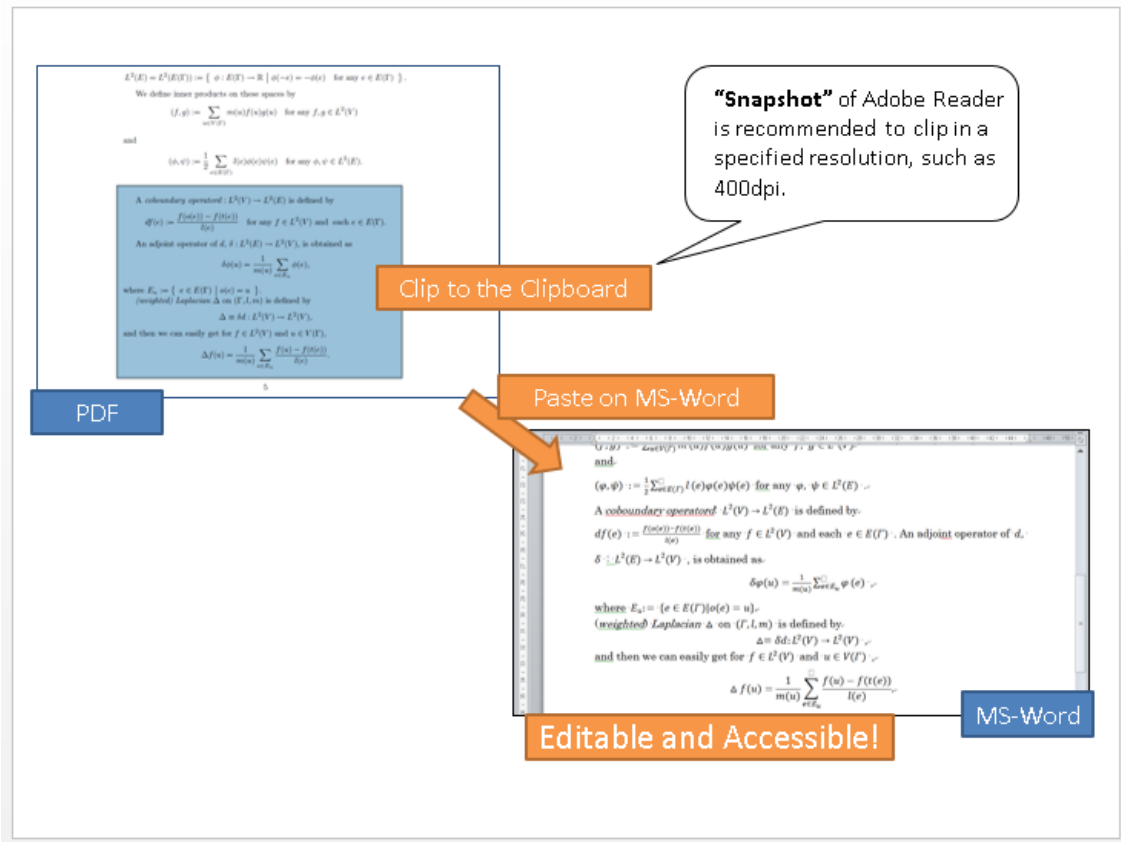
**(注 2)** 上記の方法で認識率が向上するのは LaTeX や Word、In Design 等々により電子的に生成された PDF であり、印刷文書のスキャン画像を PDF で保存した場合など、文字や記号が埋め込まれていない PDF については、認識率は従来バージョン (Ver.3.0 以前) の InftyReader と同レベルです。

**(注 3)** Pdfminer は優れたソフトウェアですが、特殊なフォントや記号については文字コードが取得できない場合があり、また座標についても抽出できるのはフォントの矩形情報であり、実際の幾何学的矩形とは大きくずれる場合などがあります。そのため特殊な記号や数式部分の認識には従来の OCR による結果と照合して最終結果を確定します。

## 2. クリップボードの画像を認識して貼り付け

InftyReader Ver. 3.1 をインストールすると、クリップボードにコピーされた画像を認識して MS Word や InftyEditor, ChattyInfty の文書に、右クリックのメニューから直接認識結果を貼り付けることが出来るようになります。

## Illustration of “Clip and Paste as Math Text” to Word



(注) InftyReader は低解像度の画像は認識できないため、クリップボードには 400DPI 程度の高解像度でコピーする必要があります。PDF から文章や数式などを範囲指定して「指定した解像度」の画像としてコピーするためには、をクリップするには、Adobe Reader の “スナップショット” が便利です。詳しくは、InftyReader 3.1 のパッケージに含まれている [CopyAndPasteAsMathText.pdf](#) をご覧下さい。