KYO QA ServiceNow Knowledge Tool - Usage Instructions

Basic Usage

1. **Launch the application** using `START.bat` (Windows) or run `python run.py`

2. **Select an Excel file** with a "Meta" column to update

3. **Choose PDF files** by either:
   - Selecting a folder containing PDFs
   - Selecting individual PDF files

4. **Click "START"** to begin processing

5. **Monitor progress** with the status indicators, progress bar, and logs

PDF Processing & OCR

The application will:

1. Extract embedded text from PDFs when available

2. Automatically detect image-based PDFs and apply OCR

3. Identify model numbers using pattern matching

4. Update the Excel file with extracted information

OCR Handling

- The application detects when OCR is needed for image-based PDFs
- OCR is applied using Tesseract with image preprocessing for better results
- OCR-processed files are marked with an "(OCR)" indicator
- For best results, ensure Tesseract OCR is installed or available in the `tesseract` folder