# Document OCR & Extraction Tool - User Guide

This guide will walk you through using the Document OCR & Extraction Tool to process your scanned PDF or image documents, extract text, identify and edit tables, and export the data.

## 1. Getting Started: Uploading Your Document

The application interface is divided into three main panels:

1. **Left Panel:** Document upload and page navigation.
2. **Center Panel:** Displays the currently selected page for viewing and area selection.
3. **Right Panel:** Shows extraction results and action buttons.

**Steps to Upload:**

1. **Locate the Upload Area:** In the top section of the **Left Panel**, you'll find an area labeled "Click or Drag & Drop (PDF/Image)".
2. (Screenshot: Show the initial UI with the upload area highlighted)
3. **Upload Your File:**
4. **Option A (Click):** Click anywhere inside the dashed border of the upload area. A file selection dialog will open. Navigate to your PDF or image file (e.g., .png, .jpg, .tiff) and select it.
5. **Option B (Drag & Drop):** Drag your PDF or image file from your computer directly onto the upload area.
6. **Processing:**
7. Once a file is selected or dropped, the application will automatically start processing it.
8. A global loading spinner with the message "Processing Document..." will appear, indicating that the backend is:
9. Converting images to PDF (if an image was uploaded).
10. Running ocrmypdf to perform OCR, deskew the pages, and clean the document.
11. This step might take some time depending on the document size and number of pages.
12. **Document Loaded:**
13. When processing is complete, the loading spinner will disappear.
14. Page thumbnails will appear in the **Left Panel** below the upload area.
15. The first page of the document will automatically be displayed in the **Center Panel (Page Viewer)**.
16. (Screenshot: Show the UI after a document is successfully loaded, with thumbnails and the first page in the viewer)

## 2. Navigating and Viewing Pages

1. **Page Thumbnails (Left Panel):**
2. Scroll through the thumbnails to see all pages of your document.
3. The currently selected page thumbnail will be highlighted.
4. **Page Viewer (Center Panel):**
5. Click on any thumbnail in the Left Panel to display that page at a larger size in the Center Panel.

## 3. Extracting and Editing Tables (Manual VLM Method)

This is the primary method for extracting and correcting table data.

**Steps:**

1. **Select a Page:** Click on a page thumbnail in the Left Panel to display it in the Center Panel.
2. **Identify a Table:** Visually locate the table you want to extract on the page displayed in the Center Panel.
3. **Select the Table Area:**
4. In the **Center Panel (Page Viewer)**, click and drag your mouse to draw a rectangular bounding box around the table you want to extract. Ensure the box covers all the table content, including headers.
5. (Screenshot: Show an image in the Page Viewer with a red dashed selection box around a table)
6. Once you release the mouse, the coordinates of your selection will appear in the **Right Panel** under "Manual Table Extraction (VLM)" next to "Area Selected:".
7. **Extract the Table:**
8. In the **Right Panel**, under "Manual Table Extraction (VLM)", click the blue button labeled "**Extract Table from Area**".
9. A loading spinner will appear while the VLM processes the selected image area with contextual information.
10. **View and Edit the Extracted Table:**
11. Once processing is complete, the extracted table will appear in the **Right Panel** under the "All Extracted Tables" section. It will be labeled (e.g., "Table from Page X (Extraction Y)").
12. The table is displayed in an editable format:
13. **Cell Editing:** Click inside any cell (including header cells) to edit its text content.
14. (Screenshot: Show an editable table in the "All Extracted Tables" section, perhaps with one cell focused for editing)
15. **Add Row:** Click the "**+ Add Row**" button below the table to add a new empty row at the end.
16. **Delete Row:** Click the red "**×**" button at the right end of any data row to delete that row (a confirmation prompt will appear).
17. **Move Row:** Click the "**↑**" (Up Arrow) or "**↓**" (Down Arrow) buttons in the first column of each data row to reorder the rows. Buttons will be disabled if a row cannot be moved further (e.g., up for the first row).
18. **Add Column:** Click the green "**+**" button in the last header cell (far right of the header row) to add a new column. You will be prompted to enter a name for the new column.
19. **Delete Column:** Click the small gray "**×**" button next to any header cell's input field to delete that entire column (a confirmation prompt will appear).
20. **Move Column:** Click the "**←**" (Left Arrow) or "**→**" (Right Arrow) buttons within each header cell to reorder columns. Buttons will be disabled if a column cannot be moved further.
21. (Screenshot: Show an editable table with the row/column action buttons highlighted)
22. **Repeat for Other Tables/Pages:**
23. You can navigate to other pages or select other areas on the same page and repeat steps 2-5 to extract more tables. Each newly extracted table will be added to the "All ExtractedTables" list in the Right Panel.
24. **Delete an Entire Extracted Table:**
25. If you want to remove an entire extracted table from the list, click the "**Delete Table**" button found next to that table's title in the "All Extracted Tables" section.

**Troubleshooting Table Extraction:**

1. If the extracted table data is incorrect or parsing fails, the "All Extracted Tables" section will display the "Raw VLM Output" in a text area for that entry. This can help diagnose issues with the VLM's interpretation or the parsing logic.

## 4. Viewing Full Page Text (ocrmypdf Output)

This section allows you to see the raw text extracted by ocrmypdf for the entire current page, which is generally very accurate for plain text.

**Steps:**

1. **Navigate to a Page:** Select the desired page in the Left Panel.
2. **Show OCR Layer Text:** In the **Right Panel**, under "Full Page Information", click the button "**Show OCR Layer Text (ocrmypdf)**".
3. **View Text:** The extracted plain text for the entire page will appear in a text area below the button.
4. (Screenshot: Show the Right Panel with the OCR Layer Text displayed in the text area)
5. **Export OCR Text:** Click the "**Export OCR Text**" button to download the displayed plain text as a .txt file.

## 5. Exporting Extracted Tables to Excel

Once you have extracted and edited all desired tables:

1. **Locate Export Button:** In the **Right Panel**, within the "All Extracted Tables" section (if at least one table has been extracted), you will see a green button: "**Export All Tables to Excel**".
2. **Click to Export:**
3. Click this button.
4. A loading spinner might appear while the Excel file is generated.
5. Your browser will then download an .xlsx file.
6. Each table you extracted (and potentially edited) will be on a separate sheet within this single Excel file. Sheets are typically named based on the page number and extraction order (e.g., "Page1\_Table1", "Page5\_Table1").
7. (Screenshot: Show a file download dialog or an example of the Excel file structure with multiple sheets)

This guide should cover the main functionalities. Remember that the accuracy of table extraction can depend on the quality of the scan and the complexity of the table structure. The editing features allow you to refine the results to ensure data integrity.