Mosquito Data Transcription Web Application

A web-based tool for digitizing handwritten mosquito surveillance data from field forms. This application streamlines the process of converting scanned paper forms into structured digital data through document identification, image correction, and assisted transcription.

Overview

This tool is designed to help transcribe handwritten data from mosquito surveillance forms collected in Mozambique. The forms contain tabular data about mosquito captures, including information about locations, collection dates, house characteristics (walls, floors, ceilings), and mosquito counts by species.

The application guides you through three main stages: 1. **Document Identification** - Identify the document type and enter metadata 2. **Dewarping** - Correct image distortion to create a clean, rectangular table view 3. **Transcription** - Efficiently transcribe the data using keyboard shortcuts

Getting Started

Uploading Files

1. Click the **"Carregar Ficheiros"** (Upload Files) button in the sidebar
2. Select individual image files or an entire folder of scanned forms
3. Supported formats: PNG, JPG, JPEG
4. Files will be uploaded to the server and appear in the source files list

File Status Indicators

The sidebar shows the processing status for each document: - **Doc** - Document metadata has been entered - **Seg.** - Image has been dewarped/segmented - **Dados** - Data has been transcribed - **Yellow** - Partial completion - **Green** - Fully complete

Click on any file in the sidebar to begin working on it.

Tab 1: Document Information (Documento)

The Document tab is where you identify the form type and enter key metadata about the document.

Purpose

Before transcribing the data, you need to tell the system what type of form it is. Different form types have different column structures, which the system uses to guide transcription.

Steps

1. **Select Document Type**
    - Choose from the dropdown menu (e.g., "Anopheles Gambiae s.1", "Procopack", etc.)
    - The system will load the appropriate template for that form type
    - You can click on header examples to help identify the correct type
2. **Enter Location Information**

- **Província** (Province): Select from dropdown
- **Distrito/Município** (District/Municipality): Type to autocomplete
- **Localidade/Povoação** (Locality/Town): Up to 2 locations
- **Bairro** (Neighborhood): Up to 2 neighborhoods
3. **Enter Collection Details**
   - **Data de colheita** (Collection Date): Date the mosquitoes were captured
   - **Ano/mes de pulverização** (Spray Year/Month): When the area was last sprayed
4. **Submit**
   - Click **"Submeter informação do documento"** to save the metadata
   - The system will automatically advance to the Dewarping tab

Tips

- Use the autocomplete features for locations to maintain consistency
- The system will auto-fill the Province when you select a District
- You can come back and edit this information later if needed

---

Tab 2: Image Correction (Corrigir Distorção)

The Dewarping tab corrects perspective distortion and skewing in the scanned images to create a clean, rectangular view of the data table.

Purpose

Field forms are often photographed at an angle or with the paper curved/warped. This tab lets you mark the corners of the data table so the system can digitally "flatten" it into a perfect rectangle, making transcription easier and more accurate.

Steps

1. **Mark the Table Corners**
   - Click along the **top edge** of the data table, from left to right
   - Start at the top-left corner
   - Click several points along the top (corners are easiest, but more points = better accuracy)
   - Then click along the **bottom edge**, going from right to left (clockwise)
   - End at the bottom-left corner
2. **Adjust Points**
   - You can drag and reposition any point by clicking near it
   - Make sure you have the same number of points on top and bottom
3. **Apply Correction**
   - Click **"Aplicar correcção"** (Apply Correction)
   - The system will process the image and create a dewarped version
   - You'll automatically move to the Transcription tab
4. **Try Again if Needed**
   - If the result isn't square enough, return to this tab
   - Click **"Redefinir pontos"** (Reset Points) to start over
   - Add more points or adjust their positions for better results

Tips

- More points = more accurate correction (aim for 4-6 points per edge)

- Click on the actual corners/edges of the data cells, not the paper edge
- The table outline should turn blue when you have matching points on top and bottom

---

Tab 3: Data Transcription (Transcrição)

The Transcription tab is where you enter the actual data from the form into the system using keyboard shortcuts for maximum efficiency.

Purpose

Transcribe the tabular data cell-by-cell. The system highlights the current cell, shows a zoomed view, and uses keyboard shortcuts to speed up data entry.

Getting Started

1. **Define the Data Area**
   - Click on the **top-left corner** of the data table (usually row 1, column 1)
   - Click on the **bottom-right corner** of the data table
   - The data area will be outlined in blue
2. **Begin Transcribing**
   - The first data cell will be highlighted in red (row 1, column 1)
   - A zoomed view of the cell appears above
   - Type the value you see in the cell

Transcription Workflow: Column-by-Column

Important: The application is designed for **column-by-column transcription**, not row-by-row. This means you fill in all values in one column before moving to the next column.

How it works: - When you enter a value (especially single digits in number columns), the system automatically moves **down** to the next row in the same column - When you reach the bottom of a column, the system automatically wraps to the **top of the next column** - This vertical workflow is much faster than moving horizontally across rows

Example workflow: 1. Start in column 1 (house numbers) - transcribe all house numbers from top to bottom 2. System automatically moves to column 2 - transcribe all values in column 2 3. Continue column-by-column until you reach the last column 4. The form is complete!

Keyboard Shortcuts

Navigation: - **Arrow Keys (↑↓←→)** - Move between cells - **Enter** - Save and move to next row - **Space** - Toggle overlay (show/hide all transcribed values)

Fast Data Entry (Column-by-Column): - **0-9** - For number columns, automatically advances to next row **in the same column** - For numbers ≥10: press the first digit, press ←, then type the second digit - **Letters (a-z)** - For dropdown columns (wall/floor/ceiling types), autocomplete by typing - Example: Type "c" then "a" to autocomplete "Caniço" - System advances automatically to next row when only one match remains

Special Functions: - **X** (capital) - Clear entire current column and return to top of that column - **Shift+Symbol** - Type the underlying number (Shift+! = 1, Shift+@ = 2, etc.)

Action Buttons

- **? Ajuda** - Open the keyboard shortcuts help modal
- **Submeter transcrição** - Save your work (do this periodically!)
- **Marcar transcrição como completa** - Mark this document as finished
- **Tentar corrigir colunas** - Auto-detect column boundaries (useful if columns are misaligned)
- **Tentar corrigir linhas** - Auto-detect row boundaries (useful if rows are misaligned)
- **Tentar transcrição automática** - Use OCR to automatically transcribe (review results carefully!)

Tips

- **Work column-by-column** - Don't jump around; complete each column from top to bottom
- **Save frequently** - Click "Submeter transcrição" every few columns
- **Toggle overlay** - Press Space to check your work against the image
- **Pattern detection** - For latitude/longitude columns, the system will auto-fill prefixes after you enter the first cell
- **Column wrapping** - When you finish a column, the system automatically moves to the top of the next column
- **Focus on speed** - The shortcuts are designed for rapid vertical entry; trust the system to advance automatically

Workflow Optimization

1. **Start with column 1** (house numbers) - transcribe all rows from top to bottom
2. **Move to column 2** - the system wraps automatically when you reach the bottom
3. **Continue column-by-column** - complete each entire column before moving to the next
4. For number columns: let single-digit auto-advance work for you (just type and it moves down)
5. For wall/floor/ceiling columns: learn the letter prefixes for common values (e.g., "ca" for Caniço)
6. Toggle overlay (Space) periodically to verify accuracy of the entire column
7. Save after completing every 2-3 columns

Common Issues

- **Wrong column?** Use arrow keys to navigate, not mouse clicks
- **Misaligned columns?** Try "Tentar corrigir colunas" button
- **OCR made mistakes?** Auto-transcribe is a starting point; always review
- **Lost your place?** Press Space to show the overlay with all entered values

---

Questions or Issues?

- Click the **? Ajuda** button in the Transcription tab for quick keyboard reference
- The application auto-saves periodically, but manual saves are recommended
- Each document must go through all three tabs to be marked complete