User Manual

Damarrion Morgan-Harper   Stephany Portillo   Nadia Hernandez

Samantha Moncada   Tommy Works

December 2025

Jira Project Link

Jira Project URL: https://[YOUR-JIRA-LINK].atlassian.net/

Project Overview

The **Box.com Discovery Bates Namer** is a cloud-native serverless application designed to automate the ingestion, parsing, and renaming of legal discovery files. By leveraging **AWS Lambda** and the **Box Skills Framework**, this tool replaces manual desktop processing with an automated workflow that scales to handle terabytes of data monthly.

Features & Capabilities

- **Automated Bates Extraction:** Scans uploaded PDFs to identify and extract sequential Bates stamp identifiers.

- **Metadata Cascade:** automatically captures metadata (PD Case Number, Disc Number) from the parent folder structure and assigns it to the file.

- **Intelligent Renaming:** Renames files according to the standardized convention:

  `{BatesNumber}_Disc{XX}_{OriginalFileName}.pdf`

- **Error Notifications:** Detects "soft" errors (e.g., inconsistent stamps) and "critical" errors (e.g., file size limit), notifying the user via email.

Usage Instructions

1. Folder Naming Convention (Critical)

For the automation to trigger correctly, you must upload files into a folder that strictly follows this naming convention:

$$pd[CaseNumber]\_disc[Number]$$

Examples:

- `pd12345_02` (Correct)

- `Case12345` (Incorrect - Automation will fail)

2. Uploading Files

1. Log in to **Box.com** and navigate to the designated "Skill Applied Folder".

2. Create a new folder using the naming convention above (e.g., `pd55921_01`).

3. Drag and drop your PDF discovery files into this new folder.

4. **Wait:** The system will automatically detect the upload. Within seconds, the file will be renamed and moved to the target directory.

Troubleshooting & Error Messages

If a file fails to process, you will receive an automated email notification. Common errors include:

- **"Invalid folder name format"**: You uploaded a file to a folder that does not look like `pdXXXXX_XX`. The system may default the disc number to '00' or fail entirely.

- **"File size exceeds limit"**: The uploaded PDF is larger than 64 MB. AWS Lambda cannot process this file in one pass. Please split the document.

- **"No Bates stamps found"**: The system could not detect a valid stamp pattern on the first page. Ensure the scan quality is high.

- **"Duplicate PD case folders"**: A destination folder with this name already exists. The system prevents overwriting data to ensure safety.

Technical Requirements

Users do not need to install local software. However, the system relies on the following constraints:

- **File Type:** PDF Only (Text-readable preferred).

- **Browser:** Google Chrome, Safari, or Microsoft Edge (Latest Versions).

- **Internet:** Stable HTTPS connection required for Box API interactions.