

Excel Automation Toolkit (v2.5.0)

User Guide & Technical Reference

1. Overview

Excel Automation Toolkit is a standalone Windows desktop utility built for high-volume spreadsheet operations. It handles regional number normalization, table consolidation, file splitting, export data cleansing, PII sanitization, workbook structure repair, and document conversion without relying on cloud services.

The application executes entirely on local CPU and memory. It makes no external network requests, collects no telemetry, and never uploads spreadsheet contents to remote servers. Financial, operational, and customer data remains strictly on your workstation.

Core Capabilities

- **100% Offline Processing:** Zero network transmission; compliance-ready for confidential banking and corporate records.
 - **Zero-Installation Portable:** Runs directly from any local directory or USB drive without administrative privileges or registry writes.
 - **Broad File Support:** Reads and writes standard OpenXML (`.xlsx`), Excel Binary (`.xlsb`), legacy Excel 97-2003 (`.xls`), XML Spreadsheet 2003, and raw HTML table exports from enterprise ERP systems.
 - **Safe COM Isolation:** Background automation tracks and manages dedicated Excel COM process IDs. Halting or force-stopping a task targets only toolkit workers, leaving your personal open Excel workbooks untouched.
-

2. Requirements & Getting Started

System Requirements

- **Operating System:** Windows 10 or Windows 11 (64-bit).
- **Hardware:** 4 GB RAM minimum (8 GB recommended for workbooks over 250,000 rows); 200 MB free disk space.

- **Office Integration:** Microsoft Excel (2016, 2019, 2021, or Microsoft 365) recommended for COM-based formatting and formula preservation. Core text parsing, splitting, and merging operate independently.

Quick Start

1. Extract the downloaded archive (`ExcelAutomationToolkit_Latest.zip`) into your preferred local folder (e.g., `C:\Tools\ExcelAutomationToolkit`).
2. Double-click `ExcelAutomationToolkit.exe` to launch the application.
3. Drag and drop your target Excel or PDF files into the file queue area.
4. Select the desired module from the tool panel, configure your parameters, and click **START PROCESS**.

3. Module Reference

The toolkit provides nine dedicated processing modules accessible from the main window.

3.1 Change Regional Format

Normalizes thousand and decimal separators across mismatched international standards.

- **The Problem:** ERP exports, core banking extracts, and cross-border accounting reports frequently use commas for decimals and periods for thousands (e.g., `1.250.500,75`). When opened on computers configured with US/UK regional settings, Excel classifies these values as uncalculable text strings, causing `#VALUE!` errors in formulas.
- **Target Regional Formats:**
- **Standard / US Format (`1,234.56`):** Decimal point (`.`) and thousands comma (`,`). Standard for the US, UK, Canada, Australia, Japan, and international financial systems.
- **European Format (`1.234,56`):** Decimal comma (`,`) and thousands dot (`.`). Standard for Germany, Spain, Italy, Latin America, and Indonesian banking standards.
- **French / ISO Format (`1 234,56`):** Decimal comma (`,`) and non-breaking space () for thousands. Standard for France, Nordic regions, and BIPM / ISO-31 metric accounting.
- **Processing Engines:**
- **Safe Mode (MS Excel COM):** Operates through Microsoft Excel COM automation. Retains 100% of cell formatting, font families, fill colors, borders, and column layouts. Supports both `.xlsx` and compact `.xlsb` output formats.
- **Fast Mode (Pandas / OpenPyXL):** Direct binary stream processor engineered for massive tabular dumps where speed is paramount and visual formatting is secondary. Generates `.xlsx` output.
- **Advanced Accounting & Number Cleaning (Enterprise & Lifetime):**

- **Clean Invisible Characters & Spaces:** Strips non-breaking spaces (`CHAR 160`), horizontal tabs (`CHAR 9`), carriage returns, line breaks, and leading apostrophes (`'`) that force numeric entries into text.
 - **Normalize Accounting Negatives:** Converts SAP-style trailing minus values (e.g., `1,500,000-`) and financial brackets (e.g., `(1,500,000)`) into standard negative numbers (`-1500000`).
 - **Strip Currency Symbols:** Removes currency prefixes and suffixes (`Rp` , `IDR` , `$` , `USD` , `EUR` , `€`), converting currency text into raw calculable numbers.
 - **Force Text-to-Number Conversion:** Clears green triangle "Number Stored as Text" warning flags across cells, enabling `=SUM()` and arithmetic evaluation.
 - **Identifier Protection:** The parser automatically detects alphanumeric codes with leading zeros (such as National IDs, phone numbers, postal codes, and product SKUs) and maintains them as text, preventing Excel from stripping leading zeros.
-

3.2 Merge Files

Consolidates multiple Excel files or worksheets into a unified master workbook.

- **Processing Modes:**
 - **Normal / Standard:** Sequential row stacking. Reads data across all supplied workbooks and worksheets, automatically mapping and aligning matching column headers. Columns missing in individual files receive blank cells, and extra columns append dynamically to the master grid.
 - **Multi-Table Sheet:** Specialized layout detector that scans worksheets containing multiple discrete tables separated by section headers or blank intervals, consolidating them into organized data blocks.
 - **Data Lineage Tracking:** Automatically appends `File Name` and `Sheet Name` audit columns to each row, allowing you to trace any consolidated record back to its exact source file and worksheet.
 - **Output Configuration:**
 - Output formats: `.xlsb` (Fast Binary Format) or `.xlsx` (Standard OpenXML).
 - Custom master file naming (defaults to `MERGE_DATA`).
 - Flexible storage location: Default subfolder (`RESULT_MERGE`), source file directory, or a custom target folder.
-

3.3 Convert PDF to Excel

Extracts tabular data and text structures from PDF documents directly into editable Excel workbooks.

- **Conversion Modes:**
- **Separate File per PDF:** Extracts each PDF document into an independent Excel workbook (`1 PDF → 1 Excel File`).

- **Merge All PDFs:** Combines extracted tables from all queued PDF documents into a single consolidated workbook.
 - **Extraction Controls:**
 - **Password Support:** Handles encrypted or password-protected PDF documents.
 - **Page Range Selection:** Target specific document pages (e.g., `1-3` , `5` , `8-10` , or leave blank for `All`).
 - **Output Formats:** Saves extracted data as `.xlsx` (Standard OpenXML) or `.xlsb` (Fast Binary).
-

3.4 Split Excel Files

Divides massive multi-sheet workbooks or large data tables into smaller, focused distribution files.

- **Split Modes:**
 - **Split by Worksheet:** Extracts every individual worksheet within a multi-tab workbook into an independent single-sheet file.
 - **Split by Column Value:** Filters and partitions a master table based on distinct values in a designated column (for example, generating separate workbooks for each Branch, Department, Sales Representative, or Fiscal Quarter). You can select detected headers from a dropdown or enter a column name, column letter (e.g., `B`), or column index.
 - **Formatting Options:**
 - **Keep Formatting:** Retains all cell styles, fonts, number formatting, colors, and column widths using COM automation.
 - **Data Only:** Lightweight export focused purely on raw values for rapid batch generation.
 - **Output Compression:** Optional checkbox to automatically package all generated split workbooks into a compressed `.zip` archive upon completion.
-

3.5 Bulk Data Cleansing

Resolves common data quality issues, unprintable artifacts, and layout defects across tabular datasets.

- **Content Cleansing Controls:**
- **Remove Empty Rows and Columns:** Eliminates completely blank rows and trailing empty columns.
- **Trim Extra Spaces:** Trims leading and trailing spaces while collapsing multiple consecutive internal spaces into single spaces.
- **Remove Duplicate Rows:** Identifies and removes duplicate records across the dataset.
- **Strip Non-Printable Characters:** Removes hidden ASCII control codes, zero-width spaces (`\u200B-\u200F`), Byte Order Marks (`\uFEFF`), and irregular line breaks.
- **Convert Text Numbers to Numbers:** Coerces text numbers into native numeric data types while protecting leading zeros on identifiers.

- **Fix Formula Error Cells:** Replaces formula error values (`#N/A` , `#REF!` , `#VALUE!` , `#DIV/0!` , `#NAME?` , `#NUM!` , `#NULL!`) with either a clean `Blank` cell or numeric `0` .
 - **Format Text Casing:** Standardizes text capitalization across textual fields to `UPPERCASE` , `Title Case` , or `lowercase` .
 - **Layout & Grid Standardization:**
 - **Unmerge All Cells:** Unmerges merged ranges across the sheet so tables can be sorted, filtered, and queried via PivotTables.
 - **Disable Wrap Text:** Disables text wrapping to restore uniform row heights.
 - **AutoFit Columns and Rows:** Adjusts column widths and row heights to fit cell contents cleanly.
 - **Clear Formatting / Reset Grid:** Strips fills, custom fonts, borders, and irregular styles to restore a clean, default spreadsheet grid.
 - **Formula Injection Defense:** Automatically neutralizes dangerous spreadsheet formula prefixes (`=` , `@` , `+` , `-`) in text fields by escaping them with a leading apostrophe (`'`), preventing malicious Dynamic Data Exchange (DDE) code execution when opened downstream.
-

3.6 Excel to PDF

Batch converts Excel workbooks into vector PDF documents for client presentation, compliance, and archiving.

- **Export Modes:**
 - **Separate PDF:** Converts each Excel file into an individual PDF document matching the source workbook name.
 - **Combined PDF:** Merges all converted Excel workbooks and their worksheets into a single sequential PDF document (e.g., `COMBINED_REPORT.pdf`).
 - **Security & Fidelity:**
 - **Password Protection:** Encrypts generated PDF documents with optional password security.
 - **Vector Layout Preservation:** Employs Excel's native PDF rendering engine to preserve exact print areas, margins, page breaks, chart objects, and typography.
-

3.7 Workbook Manager

Performs structural maintenance, security scrubbing, and hygiene actions across workbooks in bulk.

- **Management Actions:**
- **Remove VBA Macros:** Strips VBA macro code, user forms, and ActiveX modules, re-saving the workbook cleanly as a standard macro-free `.xlsx` file.

- **Break External Links:** Identifies external workbook references and sever links, converting referenced formula results into static values to eliminate broken link update prompts.
 - **Hidden Sheets and Rows:**
 - **Keep Hidden Content (Default)** : Leaves hidden structures unchanged.
 - **Unhide Hidden Sheets and Rows** : Reveals all hidden worksheets, including deep **xlSheetVeryHidden** tabs, and unhides all hidden rows and columns across every worksheet.
 - **Delete Hidden Sheets and Rows** : Permanently removes hidden and very-hidden worksheets, along with hidden rows and columns, purging unseen data before distribution.
 - **Unprotect Worksheets:** Batch-removes worksheet password protection across multiple files using a supplied password.
-

3.8 Audit and Security Suite

A specialized three-part module for financial auditing, data protection compliance, and spreadsheet locking.

Sub-Module 1: Data Reconciliation (Audit Diff)

Compares two datasets to detect discrepancies, missing rows, and modified values.

- **Comparison Modes:**
 - **Compare Two Separate Files:** Reconciles Base Dataset (File A) against Comparison Dataset (File B).
 - **Compare Two Sheets in Single File:** Compares two worksheets (Sheet A vs Sheet B) within the same workbook.
- **Matching Criteria:**
 - Matches records using a designated **Unique Key Column** (e.g., **ID** , **Invoice_Number** , **National_ID** , **Tax_Invoice_Number** , **Account_Number** , or any custom header).
 - Auto-detects or manually targets the header row index.
 - Optional flags to ignore whitespace differences and toggle case sensitivity.
- **Audit Report:** Produces a structured reconciliation summary identifying matched records, unmatched records in File A, unmatched records in File B, and cell-level value variances.

Sub-Module 2: Data Redaction (Sensitive Data Masking)

Automates PII, banking, and confidential data masking in compliance with PCI-DSS, GDPR, and regional privacy frameworks.

- **Supported Redaction Patterns:**
 - **Credit and Debit Cards (PCI-DSS):** Scans 13 to 19-digit account numbers with Luhn checksum validation (Visa, Mastercard, Amex, Discover).
 - **US Social Security Numbers (SSN):** Masks US SSN patterns (***--1234**).

- **IBAN and SWIFT/BIC:** Masks international bank account numbers and BIC routing codes.
- **Email Addresses:** Masks usernames while preserving domain structure (e.g., `je@company.com`).
- **Phone Numbers:** Masks international (E.164) and domestic telephone numbers.
- **Bank Account Numbers:** Masks internal bank account numbers (`123**890`).
- **National ID Numbers:** Masks national identification numbers across global formats.
- **Tax Identification Numbers (TIN):** Masks corporate and individual tax identifiers (EIN, TIN, NPWP).
- **Street Addresses:** Scans and masks postal street addresses.
- **Custom Regex Pattern Masking:** Define custom regular expression patterns (e.g., `\bCONFIDENTIAL-\d+\b`) to redact proprietary project codes or internal references.
- **Column Targeting:** Apply redaction across all columns or limit masking to specific selected columns.

Sub-Module 3: Protect / Lock

Enforces spreadsheet security policies before external distribution.

- **Protection Modes:**
- **Lock Formulas Only:** Automatically locks all formula cells and protects the worksheet while leaving input data cells unlocked for end-user entry.
- **Full Sheet Protection:** Locks all cells and makes worksheets read-only.
- **Bulk Sheet Unprotect:** Removes sheet-level passwords across multiple workbooks.
- **Freeze Formulas to Values:** Converts dynamic formulas into static values ("Paste Values") across all sheets while preserving 100% of visual styling, fonts, and borders.
- **Password Control:** Optional password assignment for protection and unprotection actions.

3.9 Workflow Pipeline Engine (Enterprise & Lifetime)

Chains multiple processing operations into an automated, single-session pipeline executed directly in memory without repeatedly saving and re-opening intermediate files.

- **Built-in Presets:**
- **Cleanse and Merge Master:** `Format → Cleanse → Merge → Deduplicate`
- **Merge, Cleanse and Freeze:** `Merge → Format → Cleanse → Freeze`
- **Audit and Sanitizer:** `Cleanse → Redact → Lock`
- **Client Distribution Report:** `Cleanse → Format → Freeze → PDF`
- **Data Cleaning and Formatting:** `Cleanse → Format`
- **Formula Freeze and Protection:** `Freeze → Full Protect`

- **Custom Workflow Pipeline:** Build a customized sequence using step controls (**Add Step** , **Remove** , **Clear All** , and reordering).
 - **Preset File Management:** Save customized workflows as **.json** configuration files to share standardized operating procedures across your organization.
 - **Output Format:** Final compiled results can be written directly to **.xlsx** (Standard Excel) or **.xlsb** (Fast Binary).
-

4. Licensing & Activation

14-Day Evaluation Period

When downloaded, Excel Automation Toolkit runs with **all features enabled for 14 days**. There are no disabled tools, file size caps, row limits, or watermark restrictions during this evaluation. You can process production workloads immediately while your permanent license is being configured.

Commercial License Activation

Commercial licenses are bound to workstation hardware to enable 100% offline activation without an ongoing internet connection.

Instant Self-Service Activation:

1. Open **Excel Automation Toolkit** on the target computer.
2. Click the **License Badge** in the top-right header (or select **About → Activate License** from the top menu).
3. Copy your unique **Machine ID** (formatted as **MCH-XXXX-XXXX-XXXX-XXXX**).
4. Visit our official **License Claim Portal**:

<https://excel-automation-toolkit.pages.dev/claim>

1. Enter your **Polar.sh Order ID / Receipt ID** and paste your **Machine ID** (or list of IDs for Enterprise).
2. Click **Generate & Claim License Key** to receive your permanent cryptographic key instantly.
3. Return to the activation window in the toolkit, paste the key into the input field, and click **ACTIVATE LICENSE**. The status indicator will switch to **Licensed**.

(Manual support is also available via email at viztumdev@gmail.com if you require assistance).

License Tiers

- **Pro:** Single-workstation commercial license bound to 1 designated PC.
- **Lifetime:** Permanent single-workstation commercial license including all future v2.x updates, Advanced Accounting formatters, and the Workflow Pipeline engine.

- **Enterprise:** Multi-seat organization license (minimum 10 seats) with hardware-locked workstation keys, custom corporate watermarks on generated reports, and Advanced Accounting formatters.
 - **Universal:** Hardware-unbound commercial license that activates instantly on any machine without hardware fingerprint locking.
-

5. Windows Explorer Context Menu

The application provides an optional Windows Explorer right-click shortcut for rapid file processing:

1. In the application header menu, open **Settings / Regional Format**.
 2. Click **Register Windows Explorer Context Menu**.
 3. Right-clicking any Excel file or directory in Windows Explorer will now display the **Excel Automation Toolkit** menu, providing one-click shortcuts to launch conversions and merges.
 4. To remove these shortcuts at any time, click **Unregister Windows Explorer Context Menu**. This integration writes strictly to `HKEY_CURRENT_USER\Software\Classes` and requires no administrative privileges.
-

6. Safety & Process Isolation

Job Cancellation

- **Single Click (Cancel):** Requests a graceful termination. The background worker finishes writing the current row or file to prevent file corruption, then halts cleanly.
- **Double Click (FORCE STOP):** Immediately terminates background automation workers. The toolkit tracks the exact Windows Process IDs (PIDs) of Excel instances it created and closes only those processes—**personal workbooks open in your own Excel sessions outside the toolkit are never closed or lost**.

Overwrite & File Safeguards

- **Conflict Resolution:** If an output file with the same name already exists in the destination folder, the application pauses and prompts whether to overwrite, auto-rename, or cancel.
 - **Read-Only Source Handling:** Source files are opened strictly in read-only mode, guaranteeing that original data is never modified or overwritten in place unless you explicitly select an in-place output mode.
-

7. Support & Resources

For license inquiries, workstation transfers, or volume licensing:

- **Self-Service Claim Portal:** <https://excel-automation-toolkit.pages.dev/claim>
- **Support Email:** viztumdev@gmail.com
- **Product Store:** https://buy.polar.sh/polar_c1_bVH2wmBY9DqUMeTl6TAkXxW5hXYHh9gzufmqM3j3xMa
- **Release Version:** v2.5.0 Commercial Release
- **Platform:** Windows 10 & 11 (64-bit)