MoiraCorp

MOIRACORP · SPREADSHEET ANONYMISER · v1.1

Quick start for compliance officers

Three-page guide. Read this once. Refer back when you forget.

What this tool does

Takes any spreadsheet you have (.xlsx, .csv, .xls), looks at each column, replaces identifying values with pseudonyms (Acme Holdings Ltd → ClientName_047), and saves two files back to your machine:

- An **anonymised version** you can safely share with AI tools, colleagues, or external analysts. Includes an audit manifest tab.
- A **mapping file** that lets you reverse the pseudonyms back to real values. **Keep this one secure — do not share.**

Everything runs in your browser. No data is uploaded. No internet connection is needed once the tool is open.

PRIVACY MODEL (V1.1)

- The tool is fully self-contained — when you open it, the page makes zero outbound requests.
- Verify it yourself: open the tool, press F12 → Network tab → reload. The list will be empty.
- Your data, the column rules, and the mapping never leave your machine. The downloads are produced locally.

The five-step workflow

1. Open the tool.
2. Drop your spreadsheet onto the dashed area. Tool detects column types automatically.
3. Click **"Pseudonymise all PII"** (top of the column table). Then click **"Generalise all quasi-identifiers"**. Two clicks = safe defaults applied.
4. Skim the table. Each column shows a live preview (**Acme Holdings Ltd → ClientName_047**). Change anything that looks wrong via the action dropdown.
5. Click **"Export two files →"**. Two downloads appear: the anonymised file and the mapping file. You are done.

Column actions — what each does

Each column has four possible actions. Pick one per column. Defaults are sensible; change only if you disagree.

Action	What it does	When to use
Keep as-is	Column appears in output unchanged — same values, same format.	Safe columns: amounts, statuses, dates of transactions, regions. The analytical columns you want the AI to actually see.
Pseudonymise	Replaces values with consistent placeholders (Acme → ClientName_047). Same input → same placeholder.	Identifying columns: names, emails, account numbers, NI numbers, addresses. The default for anything flagged red (PII).
Generalise	Reduces detail but keeps signal. Postcode → district. DOB → year. Name → "John S."	When you need partial identifying info for analysis but full identity is too much.
Drop column	Column removed from output entirely. Doesn't appear at all.	Notes, free-text comments, internal IDs that have no analytical value.

Risk colour codes

Each column has a coloured badge based on what the tool detected. Glance left-to-right when reviewing — anything red needs your attention.

- 🔴 PII — direct personal data (names, emails, NI numbers, account numbers). Default: pseudonymise.
- 🟠 Quasi-identifier — could identify when combined (postcode, salary, ethnicity). Default: generalise.
- 🟢 Safe — not identifying (amounts, dates, statuses). Default: keep.
- ⚪ Unknown — the tool could not decide. Look at the sample values and choose.

Bulk buttons

Five buttons at the top of the column table apply changes across many columns at once. Use them to set safe defaults in one click each:

- Include all / Exclude all — toggles every column's "in output" tickbox.
- Pseudonymise all PII — sets every red-badge column to Pseudonymise. **One-click "make it safe".**
- Generalise all quasi-identifiers — sets every yellow-badge column to Generalise.
- Keep all safe columns — sets every green-badge column to Keep.

There is also a toggle below the bulk buttons:

USE GENERIC COLUMN PLACEHOLDERS

Off by default. When on, output uses C1_001, C2_001 instead of ClientName_001. Stops the output from leaking your firm's column taxonomy. Turn on when the recipient should not know what columns exist in the source.

Output files & how to handle them

You get two files back. They serve different purposes and need different handling.

File 1 — the anonymised data

SAFE TO SHARE

Filename: yourfile_anonymised_2026-06-11_1430.xlsx
Contains two tabs: Data (the pseudonymised rows) and Manifest (audit trail).
No identifying values. Safe to feed to Claude, ChatGPT, Copilot. Safe to email internally.

Manifest tab records timestamp, original filename + SHA-256, row count, per-column rules. Audit-defensible record of what was anonymised, when, and how.

File 2 — the mapping file

RESTRICTED — DO NOT SHARE

Filename: yourfile_mapping_2026-06-11_1430.xlsx
Lists every pseudonym alongside its real value. Anyone with this file reverses everything. Treat it like a password file.

Recommended handling:

- Right after download, open the file in Excel.
- **File → Info → Protect Workbook → Encrypt with Password. Set a strong password.**
- Save the encrypted version. Delete the unencrypted version. Empty Recycle Bin.
- Store the encrypted mapping file on your firm's secured drive (SharePoint with restricted access, encrypted local volume, password manager attachment).
- **Never email this file. Never paste into chat or AI tools. Never put it in a shared folder without access controls.**

Using the output with AI tools

Standard workflow once you have the two files:

1. Open the anonymised file. Skim the Data tab — confirm nothing identifying slipped through.
2. Copy the rows into the AI tool of your choice (Claude, ChatGPT, Copilot) and ask your analysis questions.
3. The AI's response refers to **ClientName_047** etc. Use the mapping file *locally* to translate findings back to real names.
4. When the analysis is complete, archive both files together (still encrypted) for audit. Delete working copies.

FCA audit notes

- Manifest tab in the anonymised file is your evidence pack: timestamp, tool version, source SHA-256 hash, per-column rules. Reproducible and auditable.
- Timestamped filenames per run, randomly shuffled pseudonym numbers, no outbound network calls. Inspect with F12 → Network in the browser to verify.

Tip: *after exporting, click "Back to columns" on the done screen — file and column rules stay loaded, change what you need, click Export again. Fresh timestamp, no re-upload.*