

## TESTER ONBOARDING

# The Redactor.

*A 5-minute interface tour and a 10-minute test plan.*

Test build · 12 May 2026 · Browser-only — no server, no upload, no logs.

## What you're testing

---

This is a redaction tool for documents that contain personal or sensitive data — names, addresses, NI numbers, account numbers, postcodes, and so on. The use case: *you want to use ChatGPT or Claude to summarise, draft, or analyse a document, but you can't paste the raw thing in because of UK GDPR and FCA Article 32 obligations.* This tool sits between the document and the AI — you redact, then paste the safe version.

**It runs entirely in your browser.** Nothing is uploaded. Nothing is logged. Reload the page and your data is gone. That is the tool's single most important property — please test it harshly.

## Quick start (3 steps, 30 seconds)

---

1. **Open** redactor.html in any modern browser (Chrome, Edge, Safari, Firefox). *Double-click the file, or drag it into a browser tab.*
2. **Click the SAMPLE button** (top right) to load test data. You'll see redactions appear immediately on the right pane.
3. **Click any black bar** to reveal what's underneath it. Click again to hide. That's the basic flow.

THE REDACTOR — INTERFACE TOUR

follow the numbers — 1 to 8



# Redactor.

Strip names, addresses, and account details before sharing documents with AI tools. Click any black bar to reveal it — same value reappears wherever it occurs.

REDACTED x 15

v1.0 MoiraCorp - client-side only  
nothing leaves this browser

Drop a file or click to browse

TXT · MD · CSV · JSON · HTML · XML · PDF · DOCX · XLSX · ODT · RTF · EML · IMAGES (OCR)

**CUSTOM RULES — COMPANIES · ADDRESSES · KEYWORDS** [ACTIVE]

**Companies**

- Auto-detect Ltd / Inc / LLC suffixes. Catches "Acme Corp", "Globex Limited", etc.
- Auto-detect names with initials. "A L Eriksen", "M Wright" — incl. OCR-mangled forms.

CUSTOM NAMES

Type a name + Enter...

**Addresses**

- Auto-detect UK postcodes. Reliable. Catches SW1A 1AA.
- Auto-detect street patterns. Loose — may over-match.

CUSTOM STRINGS

Type an address + Enter...

**Keywords**

Anything else you want gone — codenames, internal IDs, names, specific phrases. Matched anywhere it appears.

BLOCKLIST

Type + Enter, or paste a list...

**SMART RULES**

- Context-aware (redacts values after labels like "Name:", "Phone:", "Reference:").
- Group same details (redact every later mention of a detected value)
- Aggressive 8-digit (catches dates/IDs)

[REVEAL ALL] [CLEAR] [RESET ALL] [SAMPLE]

ORIGINAL · 556 chars

Dear Mr Eriksen,

Following your call on 12 March 2024, please find enclosed the suitability report for your portfolio review. Your details on file are:

Name: A L Eriksen  
 NI Number: AB 12 34 56 C  
 DOB: 15/04/1972  
 Address: Zetica House, Southfield Road, Eynsham, Witney, OX29 4JB  
 Email: a.eriksen@example.co.uk  
 Phone: 07700 900123  
 Account: 12345678 (Sort: 12-34-56)  
 Passport: 521234567

The recommendation is to rebalance towards Just Audit & Assurance Ltd at 37 Market Square, Witney, OX28 6RE managed by Martin Wright.

Yours sincerely,  
 Compliance Officer

SANITISED · 516 chars [TEXT] [DOCUMENT] click any bar to reveal

[.TXT] [.PDF] [6 PRINT] [COPY]

Dear Mr Eriksen,

Following your call on 12 March 2024, please find enclosed the suitability report for your portfolio review. Your details on file are:

Name: [NAME] [NI NUMBER]  
 DOB: [DOB]  
 Address: [ADDRESS] [ADDRESS], Eynsham, Witney, [POSTCODE]  
 Email: [EMAIL]  
 Phone: [PHONE]  
 Account: [ACCOUNT\_NUMBER] (Sort: [SORT\_CODE])  
 Passport: [PASSPORT]

The recommendation is to rebalance towards Just Audit & [COMPANY] at [ADDRESS], Witney, [POSTCODE] managed by Martin Wright.

Yours sincerely,  
 Compliance Officer

AUDIT TRAIL · click a category to toggle the whole group

Address (street) x2   Postcode x2   Name (initials) x1   Name (OCR-merged initials) x1   NI Number x1   Date of Birth x1   Address (building) x1

Email x1   Phone x1   Account Number x1   Sort Code x1   Passport x1   Company (suffix) x1

Privacy: this tool runs entirely in the browser. No data is uploaded. Reload the page = everything gone.

**1** TITLE BAR shows a live count of redactions found. 'Nothing leaves the browser' is literal — verify in your Network tab if you want.

**2** DROP A FILE or paste text into the input pane. Works with PDF, DOCX, XLSX, EML, RTF, ODT, plus images via OCR.

**3** CUSTOM RULES — add company names, addresses, or keywords specific to the document. Press Enter after each one.

**4** SMART RULES — Context-aware catches labelled values like 'Reference: 12345'. Group same details propagates a redaction across the document.

**5** ACTION BAR — REVEAL ALL shows every redaction. CLEAR wipes the document only. RESET ALL also wipes custom rules and marking. SAMPLE loads test data.

**6** EXPORT — download as .TXT or .PDF, send to printer, or copy. Reveals are respected on export.

**7** TWO-PANE VIEW — ORIGINAL on the left (editable), SANITISED on the right. Click any [BLACK BAR] to reveal that one item.

**8** AUDIT TRAIL — every category found, with counts. CLICK A CHIP to toggle the whole group at once. Fastest way to release names while keeping numbers redacted.



## Walkthrough — point by point

---

### ① Title bar — the live count

The “REDACTED × N” stamp top-right shows how many redactions are currently active. It updates in real time as you type, change rules, or reveal items. The “v5.0 · client-side only · nothing leaves this browser” line is the privacy guarantee — open your browser dev tools and check the Network tab while you use the tool. You should see no outbound requests after the page loads.

### ② File drop / paste — the input

Two ways in:

- **Drop a file** onto the dashed box, or click it to browse. Supported formats: PDF (text and scanned), DOCX, XLSX, ODT, EML, RTF, TXT, MD, CSV, JSON, HTML, XML, YAML, and PNG/JPG (OCR via Tesseract).
- **Paste text** directly into the ORIGINAL pane on the left. Edits are live — the right pane updates as you type.

*Note on PDFs:* for scanned PDFs (image-only, no text layer), OCR can take 10–30 seconds per page. You’ll see a progress message. Be patient.

### ③ Custom rules — the per-case patterns

Three columns. Each accepts free-text entries — type, press Enter, repeat. Pasting a list works too:

- **Companies** — names that wouldn’t be caught by the generic “Ltd / Inc” pattern. e.g. internal codenames or government department names.
- **Addresses** — specific addresses or address fragments to scrub. Postcodes are caught automatically.
- **Keywords** — anything else: project names, internal IDs, codenames, deal aliases. Matched anywhere it appears in the text.

### ④ Smart rules — how aggressive to be

Three toggles control how aggressive the auto-detection is:

- **Context-aware** (default ON): catches the value after labels like “Name:”, “Phone:”, “Reference:”. This is the highest-yield rule for structured documents.
- **Group same details** (default ON): once a value has been detected once, all later mentions of the same value are also redacted.
- **Aggressive 8-digit** (default OFF): catches any bare 8-digit number. Useful for case files. Off by default because it false-positives on dates.

### ⑤ Action bar — reveal, clear, reset, sample

Four buttons sit on the action bar below the smart-rules row. Pick the one that matches what you’re trying to do:

- **REVEAL ALL** — shows every redacted value at once. Click again to hide them all. Useful for a quick sanity check before exporting.

- **CLEAR** — wipes the current document only. Keeps your custom rules and document-marking settings. Good for processing similar documents one after another.
- **RESET ALL** — wipes everything: document, custom rules, watermark, header/footer, page range, smart-rule toggles. Confirms before doing it. Use when switching to a completely different type of document.
- **SAMPLE** — loads a known test document so you can play without supplying your own data.

*Quick rule of thumb:* CLEAR for the next document of the same type, RESET ALL when you're starting fresh on a different brief.

## ⑥ Export — copy or download

Four options on the right pane's action bar:

- **↓ .TXT** — plain text with [TAGS] in place of redactions. Best for pasting into a chat or email.
- **↓ .PDF** — only enabled when you have a PDF or scanned document loaded. Produces a redacted PDF with black bars over the original page images. Honors any reveals you've made.
- **🖨️ PRINT** — opens the print preview with redactions baked in. Works for any document type.
- **COPY →** — copies the redacted text to your clipboard. Same content as the .txt download.

**All four exports respect your reveal choices: items you've revealed in the UI are NOT redacted in the export.** This is intentional — you're curating the redaction set, not just running a regex pass.

## ⑦ Two-pane view — original ↔ sanitised

ORIGINAL is on the left and is editable. **SANITISED is on the right and is read-only.** The two scroll in sync — scroll one, the other follows proportionally.

**Click any [BLACK BAR] in the sanitised pane to reveal what's underneath.** Click again to hide it. Revealed items are highlighted yellow with a dashed red underline so you can see your audit decisions at a glance. The reveal is per-value: revealing one instance of a name reveals every instance of that name.

*If you've loaded a PDF with images, you'll also see a DOCUMENT tab next to TEXT. Switch to it to see the redactions overlaid on the original page images.*

## ⑧ Audit trail — toggle whole categories

Below the panes you'll see chips for every redaction category found, each with a count: *Postcode ×2, NI Number ×1, Email ×1*, and so on. **Click any chip to toggle every redaction in that category at once.**

This is the fastest way to refine a redaction pass without clicking through each black bar. Common patterns:

- **Reveal NAMES but keep numbers** — for AML / regulatory submissions where the named subject is referenced openly but bank details remain blacked out.
- **Reveal POSTCODES but keep full addresses** — for geographical analysis without exposing the building.

- **Reveal everything except CARD\_NUMBER and ACCOUNT\_NUMBER** — for pasting into ChatGPT to draft a letter, where the recipient should see the customer's name but never their financial details.

*Once you've toggled categories, individual reveals on top of those still work — chips set the bulk default, individual clicks override.*

## What to test (10-minute plan)

---

Pick whichever of these match your workflow. The point is to find what fails or feels wrong — not to validate that it works. We expect bugs.

### Scenario 1 — A document you actually use

Drop in a real document you would otherwise be cautious about pasting into ChatGPT. A client letter, a claim form, a KYC pack, an internal memo. Look at the sanitised output and ask:

- Did it catch everything I'd expect it to catch?
- Did it false-positive on anything that isn't actually sensitive?
- If I pasted this version into ChatGPT, would I be comfortable?

### Scenario 2 — Stress-test the privacy claim

This is the most important test. Open browser dev tools (F12 in Chrome/Edge, Cmd+Opt+I in Safari), go to the Network tab, and clear it. Then load a document and run redaction. You should see:

- No outbound network requests after the initial page load (CDN scripts for OCR / PDF parsing are loaded once, then nothing).
- No requests to any analytics, telemetry, or logging endpoints.
- If you turn off your wifi mid-session, everything still works.

**Tell us if you see anything that contradicts the “nothing leaves the browser” claim. This is non-negotiable for the financial-services use case.**

### Scenario 3 — Edge cases

- A scanned PDF with poor image quality (does OCR cope?)
- A 30+ page document (does it stay responsive?)
- A document with names like “James River” or “St. James Hospital” (does it correctly NOT redact these?)
- A document with mixed languages or non-Latin characters
- Pasting in a 100,000-character paste from clipboard

### Scenario 4 — The export round-trip

Run redaction, reveal a few specific items (deliberately), then export each format (.txt / .pdf / print preview / copy). For each export:

- Are the items you revealed actually visible in the export?

- Are the items you didn't reveal actually hidden — not just visually but in the underlying data? (For PDF: open in Acrobat and try to select the redacted text; you shouldn't be able to.)

## How to give feedback

---

Anything is welcome. Bugs, UX confusion, feature requests, “this regex pattern over-matches”, “the colour scheme is too tan”. Reply by email or write up your thoughts using the template below — I prefer specific over general.

<b>Browser &amp; OS</b>	e.g. Chrome 124 on macOS Sonoma
<b>Document type tested</b>	e.g. 12-page scanned PDF, KYC pack from 2023
<b>What worked well</b>	(your notes)
<b>What didn't work</b>	(your notes — specific examples are most useful)
<b>Missing detection</b>	Things that should have been redacted but weren't. Paste the actual line if possible.
<b>False positives</b>	Things that were redacted but shouldn't have been.
<b>Would you use this?</b>	Yes / No / Yes if [feature]. We need this answer most of all.

### One last thing

If at any point you find yourself thinking ‘I'd use this if it could just...’ — please tell us. The roadmap is being shaped right now, and tester input is the strongest signal we have for what to build next.

*Thank you for testing.*