



SHORT GUIDE · AGENT · REDACTING PERSONAL DATA

Redacting files with a marker pen? The Redact Agent marks, you decide.

Before any file goes out, names, addresses and identifiers of third parties have to disappear. This short guide shows how the Redact Agent marks personal data in PDFs and scans, how you review the preview and create a cleaned copy. For public authorities, legal departments and data protection.

KEY TAKEAWAY

13 scenarios and 16 selectable data categories, whitelist and your own terms, a preview with the reason for each area. Redacted text is removed from the copy, the original stays unchanged.

01 · THE PROBLEM

The file has to go out. Third-party names must not.

Files, expert reports and correspondence contain names, addresses, phone numbers and bank details of people who have nothing to do with the request. Before release, pages are printed, blacked out with a pen and scanned again. That takes time, and a missed spot cannot be taken back.

WHY IT HAPPENS

A black box is not a redaction

A black rectangle in a PDF editor often just sits on top of the text. Underneath, the text can still be copied, and metadata or attachments go out unnoticed.

A scan is an image first

A scanned page contains pixels, not text. A text search finds no name in it, so every page is read by hand.

Every occasion has its own rules

In a data subject request, the requester stays visible; in a publication, every name disappears. Deciding this from scratch each time leads to inconsistent redaction.

HOW YOU NOTICE IT

- Files are printed, blacked out with a pen and scanned again.
- Nobody can say for sure whether text still sits under a black box.
- Requests, courts and publications each follow their own unwritten rules.
- Redacted copies sit next to the original without a clear label.
- With large files, redaction waits for days because every page is done by hand.

02 · THE WAY WITH PAPEROFFICE

Let it mark, review, release the copy.

The Redact Agent reads the PDF, including scans, and marks the data categories of the chosen scenario. In the preview, you decide on every area. Then a cleaned copy is created as a new document.

1

Choose a scenario for the occasion

13 scenarios set the rules, for example for a data subject request, external sharing, legal proceedings or medical records. Or you choose the categories yourself.

2

Read and mark

Text recognition also reads scans. The AI marks findings in the chosen categories, from names and addresses to bank details, ID numbers and health data.

[MODULE: AI-OCR](#)

3

Decide in the preview

Click areas to redact or release them, and drag a rectangle over anything missing. Hovering shows the category and the reason. Whitelist and your own terms apply to every page.

4

Truly remove the text

Text under the boxes is deleted from the PDF and checked again afterwards. If anything remains underneath, the copy is built from redacted page images. Metadata and attachments are cleared.

5

Store and send the copy

The copy carries the suffix _redacted and is stored in the chosen workspace, the original stays unchanged.

Send by email opens the secure delivery with the copy.

[MODULE: SECURE EMAIL DELIVERY](#)

03 · IN THE APP

From the marked document to the cleaned copy.

You handle single files with a preview and whole batches with a saved Redact Agent. Both use the same scenarios, categories and lists.

1

Start redaction with preview

Select the file in the document list, open **Actions › Agents & workflows › Redact** and choose **Edit with preview**. For release to third parties, the **External Sharing** scenario is a good fit.

2

Maintain the lists

Under **Do NOT redact (Whitelist)**, enter names that stay visible, such as your authority. Under **ADDITIONALLY redact (Blacklist)**, add your own terms such as file numbers or project names.

3

Review and apply

In the preview, click areas or drag a rectangle, then **Apply redaction**. The copy goes to the chosen workspace or downloads to your device.

4

Readable instead of black

Under **Output**, choose **Pseudonymize with the Ultra AI model**. Names and identifiers are replaced by consistent fictitious values, the mapping is shown in the dialog.

5

Batches with the agent

Under **Document AI › Redact Agent**, save an agent with scenario, lists and target workspace. For many PDFs, choose **Actions › Agents & workflows › Redact** and **Automatically with Redact Agent**; processing runs server-side without a preview.

TIP FROM PRACTICE

Photos, stamps and signatures by rectangle In the preview, look closely at every page with a passport photo, stamp or signature. Drag a rectangle over such areas with the mouse, and they are removed from the copy just like marked text.

04 · AN EXAMPLE

A county office. A request for file access.



A county office receives a request for access to the file on a construction project. The file includes expert reports, correspondence and scanned objections from neighbors. The records officer has to release the file but make the neighbors' names, addresses and phone numbers unreadable.

BEFORE

- The whole file is printed and read page by page.
- The neighbors' names and addresses are blacked out with a marker pen.
- The pages are scanned again so the text underneath becomes unreadable.
- The finished copy goes to the requester as an unprotected email attachment.

WITH PAPEROFFICE

- The file is stored as a PDF in the workspace, including the scanned objections.
- With the `External Sharing` scenario, the Redact Agent marks names, addresses and phone numbers.
- The records officer reviews the preview and drags a rectangle over a signature.
- The copy goes out via secure delivery with a password, the original stays in the file.

Fictional example for illustration. Check for yourself, or with your legal or data protection advisor, which details you must redact and whether the redaction is sufficient for your purpose.

05 · CHECKLIST

Ten points for releasing files.

Go through these points with the department and data protection. Each can be answered with yes or no and shows where you need to improve.

- ☐ For each occasion, it is defined which scenario applies: request, sharing, court, publication.
- ☐ Names that must stay visible are on the agent's whitelist.
- ☐ Every single file is reviewed in the preview before release.
- ☐ Only the copy with the suffix _redacted is released.
- ☐ For batches, a saved Redact Agent with the Ultra AI model exists for each occasion.
- ☐ Before a release, someone searches for the affected names in all workspaces, including scans.
- ☐ Your own terms, such as file numbers, are on the list of terms that are always redacted.
- ☐ Pages with photos, stamps and signatures get a separate look in the preview.
- ☐ Cleaned copies are stored in a separate workspace with their own group permissions.
- ☐ Copies go out via secure delivery with a password instead of as an open attachment.

06 · FREQUENTLY ASKED QUESTIONS

What records officers ask first.

Is the redacted text really gone?

Yes. Text under the boxes is deleted from the PDF, then PaperOffice checks whether any text remains underneath. If so, the copy is built from redacted page images. Metadata, attachments, form fields and links are cleared.

Is the original changed?

No. The copy is a new document with the suffix _redacted in its name, in the chosen workspace or as a download. The original stays unchanged where it was.

What does pseudonymizing mean?

Instead of black boxes, PaperOffice inserts consistent fictitious values, and the document stays readable. This works with the Ultra AI model; the mapping is shown in the dialog, never in the PDF. Health data and access credentials stay black.

What happens with scanned pages?

Text recognition reads the scans, and the AI marks the same categories as in digital PDFs. When you apply the redaction, the image areas are painted over and the PDF is rebuilt.

Which data does the Redact Agent mark?

You can choose from 16 categories, including names, addresses, phone numbers, email addresses, bank details, tax and ID numbers, dates of birth and health data. The scenario preselects them. Drag a rectangle over signatures and photos in the preview.

Can redaction be built into our own systems?

Yes. Via REST API, `POST /documents/redact-agent-run` starts a saved Redact Agent; via MCP, the tool `po_documents_redact_agent_run` does the same. The result is the same cleaned copy as in the app.

07 · HOW TO START

Redact your next file with a preview.

Create an account on app.paperoffice.ai, upload a file and redact it via

[Actions](#) › [Agents & workflows](#) › [Redact](#) with [Edit with preview](#). Billing is pay as you go, details are on the [pricing](#) page.

Define scenarios per occasion

Agree with data protection which scenario applies to requests, sharing and publications, and which names stay visible.

Save agents

Under [Document AI](#) › [Redact Agent](#), save one Redact Agent per occasion with scenario, lists and target workspace.

Start with a test

Redact a file that contains a scan, then try to select and copy the text under a box in the copy.

DATASHEETS ON THIS TOPIC

AI-OCR

paperoffice.ai/downloads/paperoffice-ai-datasheet-ocr.pdf

Smart Search

paperoffice.ai/downloads/paperoffice-ai-datasheet-smart-search.pdf

Workspaces

paperoffice.ai/downloads/paperoffice-ai-datasheet-workspaces.pdf

Secure email delivery

paperoffice.ai/downloads/paperoffice-ai-datasheet-audit-safe-email-delivery.pdf

MORE ON THIS TOPIC

Redacting personal data

paperoffice.ai/en/ai-pii-redaction-document-anonymization-and-ai-data-anonymization



MORE ON THIS TOPIC

paperoffice.ai/en/ai-pii-redaction-document-anonymization-and-ai-data-anonymization

PRICING

paperoffice.ai/en/pricing

CONTACT

paperoffice.ai/en/contact