



SHORT GUIDE · RISK · DATA PROTECTION

Personal data in every folder and attachment?

Redact with a preview instead of a marker.

Names, addresses, IBANs and dates of birth hide in personnel files, attachments and old scans. This short guide shows how to find them with PaperOffice, redact them with a preview before sharing and store them separately. For data protection officers, HR teams and anyone who hands out files.

WHAT YOU TAKE AWAY

Search across all workspaces, preview-based redaction in 16 data categories, a whitelist and your own terms, plus a cleaned copy next to the unchanged original.

01 · THE PROBLEM

The data is there. But where exactly?

Personal data spreads across personnel files, email attachments, contracts and scanned archives. Most teams only search when an access request arrives or a file goes to a third party. Then people search by hand, copy and black out with a marker. In the end, the question remains whether something was missed.

WHY IT HAPPENS

Data travels with every document

An application arrives by email, gets forwarded, printed and scanned again. Every stop leaves another copy behind, often in a different folder or mailbox.

A scan starts out as an image

A scanned page holds pixels instead of text. Only text recognition lets search find names, addresses or an IBAN in it; until then the file slips through every search.

A black box merely hides the text

A black box drawn in a PDF editor often just sits on top of the text. The text underneath stays copyable, and file metadata or attachments leave the house unnoticed.

HOW YOU NOTICE IT

- An access request ties up a colleague for days because every storage location is searched one by one.
- Files for lawyers, auditors or service providers are printed, blacked out and scanned again.
- Nobody can say which archives contain health data or bank details.
- Personnel files, applications and invoices sit in the same folder with the same permissions.
- Before sending, someone asks: is everything on page twelve really redacted?

02 · THE WAY WITH PAPEROFFICE

Find first, then redact with purpose.

PaperOffice reads every document, including scans, and makes it searchable. Search, redaction and separation into workspaces build on that. You decide in the preview what gets redacted.

1

Find every document about a person

Search covers full text, the text from scans and recognized entities, across all workspaces you have rights to if you like. Each match is highlighted in the document.

[MODULE: SMART SEARCH](#)

2

Make sensitive data types visible

After processing with AI-DMS Ultra, **Audit › Detected data types** shows whether a document contains personal data, bank, tax or health data.

[MODULE: AUDIT CENTER](#)

3

Redact with a preview

The Redact Agent marks findings in 16 categories, from names to health data. You click areas on or off and drag a rectangle over anything it missed.

4

Cleaned copy, original stays

Redacted areas are removed from the PDF, and the copy's metadata, attachments and form fields are cleared. The copy goes to the workspace you choose; the original stays unchanged.

5

Store separately, send securely

Workspaces separate HR, clients and projects with their own group permissions. The copy goes out via secure delivery with a password and an event log.

[MODULE: SECURE EMAIL DELIVERY](#)

03 · IN THE APP

From search result to redacted copy.

Day to day there are two ways: single documents with a preview, whole batches with a saved Redact Agent. Both use the same categories, scenarios and lists.

1

Search for the person

Type a name or employee number into the search bar and choose **All Workspaces**. Select the matching results in the list.

2

Start redaction

Open **Actions > Agents & workflows > Redact** and choose **Edit with preview**. Pick a scenario such as **GDPR Request (Art. 15)**: third-party data is marked, the requester's own data stays visible.

3

Maintain the lists

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

4

Review and apply

In the preview, click areas or drag a rectangle, then **Apply redaction**. The copy lands in the workspace you chose or comes to you as a download.

5

Batches automatically

Under **Document AI > Redact Agent**, save an agent with scenario, lists and target workspace. For many documents choose **Actions > Agents & workflows > Redact** and **Automatically with Redact Agent**; processing runs on the server.

TIP FROM PRACTICE

Add photos and stamps in the preview Go through every page with a passport photo, stamp or signature in the preview. Drag a rectangle over such areas with the mouse; they are then removed from the copy just like text.

04 · AN EXAMPLE

An access request in the HR department.



A machine builder with 180 employees receives an access request from a former employee. His data sits in the personnel file, payroll records, emails and scanned application documents. The data protection officer has to compile everything and redact the data of other people.

BEFORE

- The personnel file, mailboxes and network drive are searched one by one by hand.
- Scanned documents are read page by page because search skips them.
- Colleagues' names are blacked out on printouts and scanned again.
- The documents go out as an unprotected email attachment.

WITH PAPEROFFICE

- One search across all workspaces finds the documents with his name, scans included.
- With GDPR Request (Art. 15), third-party data is marked and his own stays visible.
- In the preview she checks each finding and adds a photo with a rectangle.
- The cleaned copies go out via secure delivery with a password.

A fictional example for illustration. Which documents you hand over and whether the redaction is enough for your purpose is for you or your data protection advisor to assess.

05 · CHECKLIST

Ten points for your data protection.

Go through the points for your own organization. Each one can be answered with yes or no and shows where to improve.

- | | |
|--|---|
| ☐ We know where personnel files, applications and customer data are stored. | ☐ Scanned archives are searchable thanks to text recognition. |
| ☐ A search for a name covers all workspaces we have rights to. | ☐ For each occasion we have defined which data categories are redacted before sharing. |
| ☐ Names that must stay visible are on a whitelist. | ☐ Redacted areas are removed from the file, and metadata and attachments are cleared. |
| ☐ A person reviews every redaction of a single document in the preview. | ☐ HR documents sit in their own workspace with their own group permissions. |
| ☐ Retention years and legal basis are set for each document type. | ☐ Confidential copies go out as a password-protected link instead of an open attachment. |

06 · FREQUENTLY ASKED QUESTIONS

What data protection officers ask first.

Does detection find everything?

It marks the categories you choose, from names and addresses to ID numbers and health data. Still rely on the preview: it shows every finding, and you add or remove areas before anything becomes final.

How do we prepare an access request?

Search for the person across all workspaces, select the results and redact with the **GDPR Request (Art. 15)** scenario. Third-party data is marked, the requester's own data stays visible. Whether the response is complete is for you to assess.

What about names that must stay visible?

Enter them under **Do NOT redact (Whitelist)**, such as your company name or the responsible clerk. Under **ADDITIONALLY redact (Blacklist)** you add your own terms that are always redacted, for example project names.

Is the redacted text really gone?

Yes. Redacted areas are removed from the PDF; on scans the image areas are painted over. The copy's metadata, attachments, form fields and links are cleared. The original remains as a separate document.

Can we redact entire archives?

Yes. A saved Redact Agent processes selected PDF documents in a batch on the server and stores the copies in the target workspace. For sensitive files, use the preview, one document at a time.

Can a document be anonymized and stay readable?

Yes, with the Ultra model. Pseudonymization replaces names, addresses, dates and identifiers with consistent fictitious values. You see the mapping in the dialog; it never appears in the PDF itself.

07 · HOW TO START

Redact your first document with a preview

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

[Actions](#) › [Agents & workflows](#) › [Redact](#) with [Edit with preview](#). Billing is based on usage; details are on the pricing page.

Separate your storage

Create dedicated workspaces for HR, clients and projects and assign group permissions per workspace.

Agents for batches

Under [Document AI](#) › [Redact Agent](#), save one Redact Agent per occasion, for example for access requests or external sharing.

Connect your systems

Via REST API, POST `/documents/redact-agent-run` starts a saved agent; via MCP, use the tool `po_documents_redact_agent_run`.

DATASHEETS ON THIS TOPIC

Smart Search

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

Audit Center

paperoffice.ai/downloads/paperoffice-ai-datasheet-audit-center.pdf

Workspaces

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

Classification

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

Secure email delivery

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

MORE ON THIS TOPIC

Data protection

paperoffice.ai/en/solutions/risk/ai-data-protection-gdpr-compliance-pii-detection



MORE ON THIS TOPIC

paperoffice.ai/en/solutions/risk/ai-data-protection-gdpr-compliance-pii-detection

PRICING

paperoffice.ai/en/pricing

CONTACT

paperoffice.ai/en/contact