

Hiding names with tape?

From now on, the text under the box is truly gone.

Files for courts, data subject requests, documents for service providers. This guide shows how the Redact Agent marks personal data in PDFs and scans, how your team reviews every area in the preview, and how you release a copy from which the redacted text has been removed.



TRY IT YOURSELF IN 45 MINUTES

One claim file in five documents: claim report, scanned witness statement, medical report, correspondence on a data subject request and a power of attorney with a passport photo. Includes a test package.

WHY THIS GUIDE

A black box is drawn in seconds. What sits underneath, you cannot see.

Before any file goes out, names, addresses and identifiers of third parties have to disappear. Printing, taping over and scanning again costs hours. And a missed spot cannot be taken back.



In legal departments and data protection teams, releasing a file often starts with tape, a marker pen and a scanner. In the end, nobody can say for sure whether text still sits under a black box in the PDF.

This guide shows how the Redact Agent takes over this part: one scenario per occasion, marks by the AI, a preview in which your team decides, and a copy from which the text is truly removed. With the test package, you can try the whole process yourself in 45 minutes.

Daniel Schönland

Founder and CEO, PaperOffice

Legal department

You send files to courts, opposing counsel and experts while protecting uninvolved third parties.

Data protection

You answer data subject requests and review documents before they leave the company.

Case handling

You prepare file access and sharing, often under time pressure and with scanned pages.

WHAT YOU CAN DO AFTERWARD

Choose a scenario per occasion, maintain the whitelist and your own terms, review the preview, redact signatures and photos with a rectangle, send the copy securely and process batches with a saved Redact Agent.

OVERVIEW

Contents

01	The file for the regional court	04
02	What holds up the release	05
03	The process at a glance	06
04	Prepare the file and start Redact	07
05	The scenario for the occasion	08
06	Lists and AI model	09
07	Reviewing the preview	10
08	Signatures, photos and stamps	11
09	Apply: truly remove the text	12
10	Releasing securely	13
11	Readable instead of black: pseudonymizing	14
12	Batches with the Redact Agent	15
13	What teams ask when they start	16
14	Try it yourself in 45 minutes	17
15	Further reading	18
16	Your first release in five days	19
17	Checklist and start	20

AT A GLANCE

- 13 scenarios set the rules per occasion, for example “Legal Proceedings”, “External Sharing” or “Medical Records”.
- Text recognition also reads scans. The AI marks findings in the chosen categories.
- In the preview, your team decides on every area. Drag a rectangle over signatures and photos.
- Text under the boxes is deleted and checked again. The original stays unchanged, the copy is named ...
_redacted.pdf.
- The copy goes out via secure delivery. A saved Redact Agent handles batches, also via API and MCP.

01 · THE FILE FOR THE REGIONAL COURT

One file, three recipients. And each may see something else.

Tuesday, 9:40 a.m. Malik Osei, a paralegal in the legal department of Ostkap Versicherung AG, has to send the file on claim SA-2026-04417 to the regional court by Friday.



The file holds the claim report of the policyholder Jonas Reimer with address, date of birth and IBAN, the scanned statement of a neighbor with her signature, a medical report on the other party to the accident, and internal notes with direct dial numbers.

At the same time, Jonas Reimer has asked for access to his data, and a service provider needs a copy to settle the claim. Three recipients, three sets of rules, one file. Until now, that meant print, tape over, scan again.

THE REAL HURDLE

Malik Osei may only release the file once every area has been reviewed. A black rectangle in a PDF editor is not enough: the text underneath often stays in the file.

02 · WHAT HOLDS UP THE RELEASE

Four time sinks. And one of them is a risk.

Hardly any spot is hard to find. The time goes into reading every page, covering every spot and creating every copy again.

Tape and scanner

Pages are printed, taped over and scanned again so the text underneath becomes unreadable.

Text under the box

A rectangle in a PDF editor often just sits on top of the text. Copying, metadata and attachments still give it away.

Scans without text

A scanned page is an image. A text search finds no name in it, so someone reads every page.

Every occasion has its own rules

For a court, the parties stay visible; for a data subject request, the requester does. Deciding this from scratch each time leads to inconsistent redaction.

TIME SINK	HOW PAPEROFFICE HELPS	IN THE TEST PACKAGE
Tape and scanner	The AI marks, your team reviews in the preview, the copy is created as a new document	Claim file, 3 pages
Text under the box	Text is deleted and checked again, metadata and attachments are cleared	Copy test in the copy
Scans without text	Text recognition with positions, a rectangle for signature and passport photo	Witness statement and power of attorney as scans
Own rules	13 scenarios, whitelist and your own terms	three occasions, three scenarios

All documents in the test package are fictional. Chapter 14 shows what to try with them.

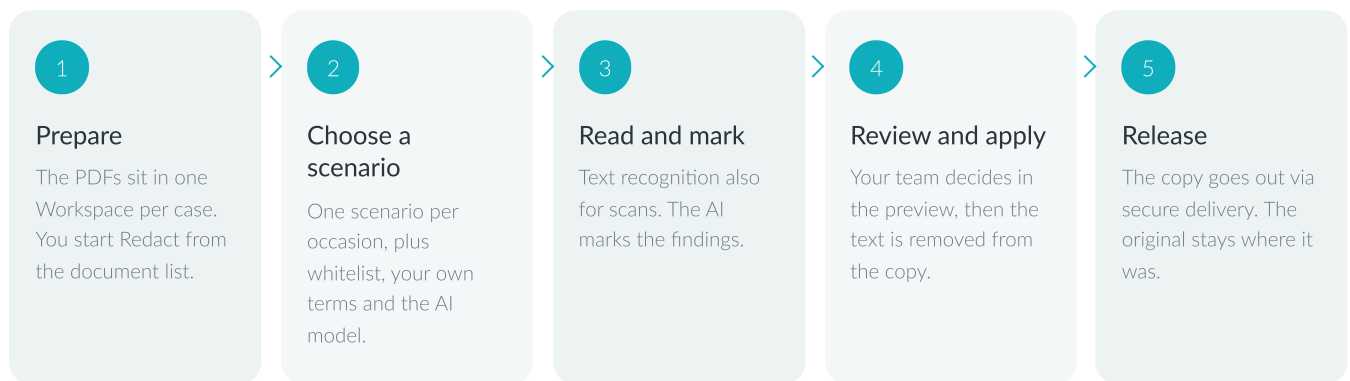
WHAT REMAINS IN YOUR HANDS

The decision what a recipient may see, and whether the redaction is sufficient for your purpose. PaperOffice marks, removes and stores. Your team reviews and releases.

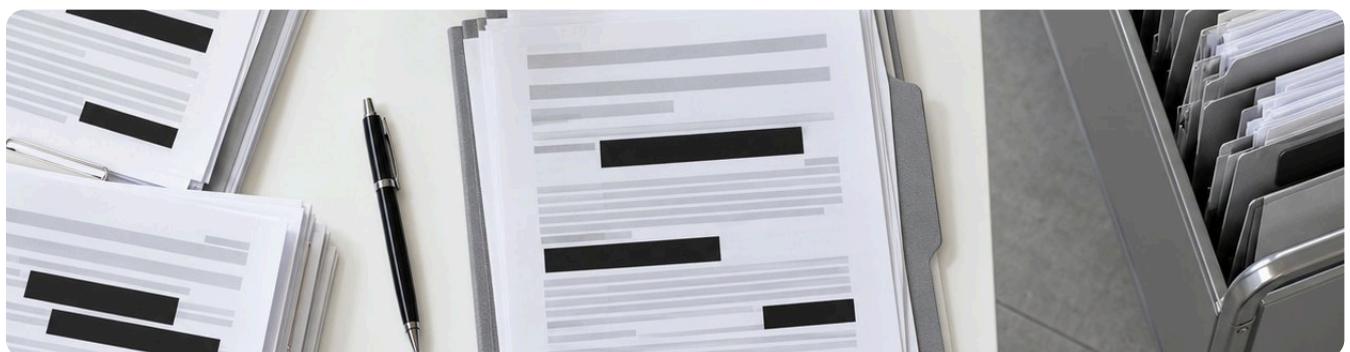
03 · THE PROCESS AT A GLANCE

Five stations. From the file to the release.

Every file follows the same path, whether created digitally or scanned, whether one at a time with a preview or in a batch.



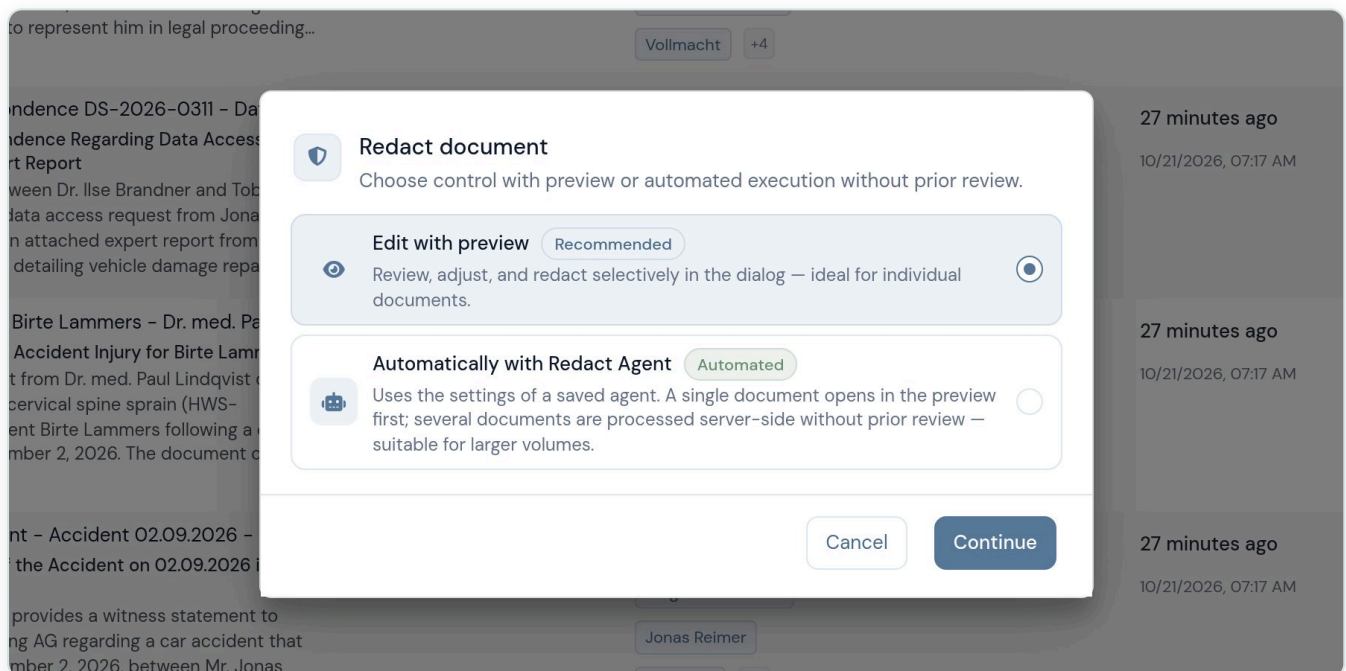
STATION	WHAT PAPEROFFICE DOES	WHAT YOUR TEAM DECIDES
Prepare	accepts digital PDFs and scans in the Workspace	one Workspace per case and who has access
Choose a scenario	sets the data categories per occasion	which scenario applies to which occasion
Read and mark	reads text and positions, marks findings	whitelist and your own terms
Review and apply	shows every marked area, removes the text	who reviews the preview
Release	copy as a new document, delivery with a password	who approves and who receives it



04 · PREPARE THE FILE AND START REDACT

One Workspace per case. Redact right from the list.

Under **Document AI** > **Workspaces**, click "Create New Workspace": name "Claim SA-2026-04417", type "File / Case". Upload the PDFs of the file via "Add document".



1 **Redact document.** Select a PDF, then **Actions** > **Agents & workflows** > **Redact**. "Edit with preview" is meant for individual files, "Automatically with Redact Agent" for batches. Then "Continue".

1 **Select:** the PDFs of the file in the document list. Redaction works on PDF documents, created digitally or scanned.

2 **Choose the path:** "Edit with preview" opens the "Document anonymization" dialog (chapters 05 to 09).

3 **Series:** With a preview, several PDFs run one after another. After each copy, "Next document" moves on.

Search first

The search with "All Workspaces" also finds a name in the text of scans processed by AI-DMS. This shows you where else a person appears.

Permissions per Workspace

Group permissions apply per Workspace: view and download, plus upload or full access. Only those who work on the file can see it.

05 · THE SCENARIO FOR THE OCCASION

13 scenarios. One for every occasion.

In the **Scenario** tab, you choose a scenario. It defines which data categories the AI marks. The default is "GDPR Request (Art. 15)".


Document anonymization
 Helps you redact personal data


Scenario


Fine-tuning


GDPR Request (Art. 15)
 Document for disclosure to data subject


External Sharing
 Document for business partners/service provider


Legal Proceedings
 Evidence with data protection for third parties


Whistleblower (HinSchG)
 Identity protection under Whistleblower Protection

1 "Legal Proceedings" card ("Evidence with data protection for third parties") for the file sent to the regional court. The hint at the top of the "Fine-tuning" tab says what the scenario redacts: third-party contact data, while the whitelist keeps the parties to the proceedings visible.

OCCASION	SCENARIO	WHAT THE SCENARIO SETS
File for a court	Legal Proceedings	third-party contact data, parties visible via whitelist
Request from a person	GDPR Request (Art. 15)	third-party data, requester data stays visible
Documents for service providers	External Sharing	personal data, business data stays visible
Publication	Public Publication	all personal data, including highly sensitive data
Medical letters, reports	Medical Records	personal and medical data
Bank or contact data	Redact Financial Data, Redact Contact Data	IBAN, credit cards, tax numbers, or names, addresses, phone, email
Your own selection	Custom	you choose from 16 data types

There are also "Whistleblower (HinSchG)", "Internal Audit", "Anonymous Application", "Personnel File / HR" and "Scientific Research". The names describe the occasion. Check for yourself, or with your legal and data protection advisors, whether the result meets the needs of your procedure.

06 · LISTS AND AI MODEL

What stays visible is set. What must always go, too.

In the **Fine-tuning** tab, you maintain two lists and optional instructions for the AI. In the **Execution** tab, you choose the AI model and the output.

The screenshot shows the 'Document anonymization' interface. The 'Fine-tuning' tab is active, displaying a 'Scenario' section with a 'Third-party contact data redacted, use whitelist to keep parties to proceedings visible' instruction. Below this, there are two input fields: 'Do NOT redact (Whitelist)' containing 'Ostkap Versicherung AG, Jonas Reimer, Birte Lammers' and 'ADDITIONALLY redact (Blacklist)' containing 'KH-88 301 552; Projekt Nordlicht'. There are also sections for 'Custom AI Instructions (optional)' and 'Custom instructions for the AI...'. The 'Execution' tab is also visible, showing 'Saved to current Workspace', 'Start priority' set to 'Best Effort', 'AI model' set to 'Ultra', and 'Output' set to 'Black boxes (standard)'.

1 **Fine-tuning.** Whitelist with the parties, blacklist with the policy number and the internal project name.

2 **Execution.** The copy is saved to the current Workspace. AI model "Ultra", output "Black boxes (standard)", then "Start".

Whitelist

"Do NOT redact (Whitelist)": names that stay visible, such as the parties and your own company. Comma-separated or one entry per line.

Your own terms

"ADDITIONALLY redact (Blacklist)": terms that are always redacted, such as internal project names or claim numbers.

Box color

At the top of the dialog under "Redaction color": Black, White, Red, Blue or Gray. Black shows every recipient that something was removed.

AI MODEL	HOW IT WORKS	INTENDED FOR
Basic	compact model, redacts whole lines	simple, well-readable letters
Premium	large model, word-precise redaction	scans, tables, forms
Ultra (default)	large model, word-precise, pseudonymizing available	sensitive or complex cases
Manual	text recognition without AI detection, you choose the areas yourself	when you know exactly what has to go

07 · REVIEWING THE PREVIEW

Every area in view. Your team decides.

After “Start”, PaperOffice reads the pages with text recognition, including scans, and looks for the data of the scenario. Then the “AI redaction review” opens.

- 1 The claim file in the preview. Page thumbnails on the left, the page with the marked areas in the middle, at the top the counter “areas will be redacted” and “Page 2/3”.

Name	Anschrift, Kontakt
Jonas Reimer	Möwenweg 7, 24000 Mustematen
Birte Lammers	Deichweg 3, 24000 Mustematen
Helga Sommerfeld	Möwenweg 9, 24000 Mustematen Tel. 040 66969 193

Uhr fuhr Jonas Reimer mit seinem Fahrzeug zurück
Möwenweg 7. Dabei stieß er gegen die Seite des Fahrzeugs
ing Deichweg befuhr. Beide Fahrzeuge waren langsam unter

- 1 **Look:** Black areas with a red border will be redacted, areas with a blue outline stay visible.
- 2 **Click:** A marked area is released, an unmarked one is redacted. The counter keeps track. In the image, a click blacked out the witness's phone number, which the AI had not marked.
- 3 **Page through:** page by page, including the scans. Then “Apply redaction”.

GOOD TO KNOW

The AI can miss a spot or mark too much. That is why a person goes through the preview page by page before every release.

08 · SIGNATURES, PHOTOS AND STAMPS

Whatever is an image, you cover with a rectangle.

The AI reads text. It does not detect signatures, passport photos, stamps and logos as images. You mark such areas yourself in the preview.

The screenshot shows a document titled "Vollmacht". Below the title, there is a line "Hiermit erteile ich". Then, there are several fields with redacted content: "Name, Vorname" (redacted as "Reimer, Jonas"), "Geburtsdatum" (redacted as "14.03.1988"), "Anschrift" (redacted as "Möwenweg 7, 24000 Musterhafen"), and "Personalausweis-Nr." (redacted as "C4K7N2R91"). To the right of these fields is a passport photo, which is also redacted with a black rectangle. A small red 'x' icon is visible in the top right corner of the photo's redaction box.

- 1 The power of attorney from the test package. A rectangle is drawn over the passport photo, and another one over the signature further down. The AI had not marked the ID number and the date of birth; a click blacked them out. The cross on an area removes it again.
- 1 Open the page: look closely at every page with a passport photo, stamp, logo or signature.
- 2 Drag a rectangle: with the mouse across the area. When you apply, it is redacted like marked text.
- 3 Correct: The cross on the area removes it, then drag it again.

"Signatures" category

It is called "Signatures (mark by hand)": drag a rectangle over the signature itself in the review step.

"Manual" mode

Text recognition shows all areas, and you click what should be redacted.

Files with images

A Redact Agent in a batch works without a preview. Handle files with photos or signatures one at a time with a preview.

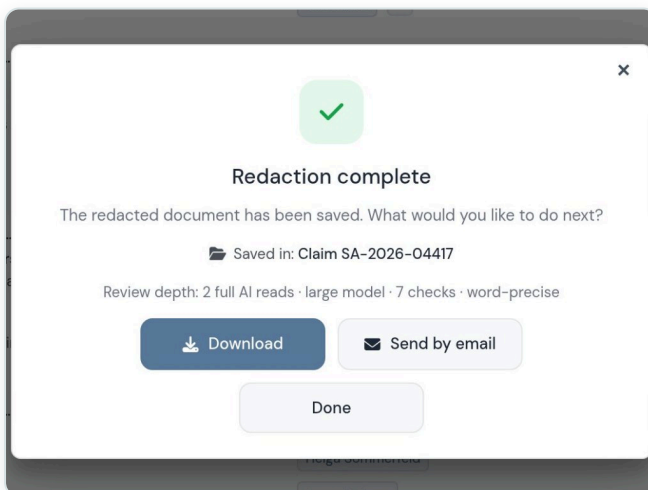


09 · APPLY: TRULY REMOVE THE TEXT

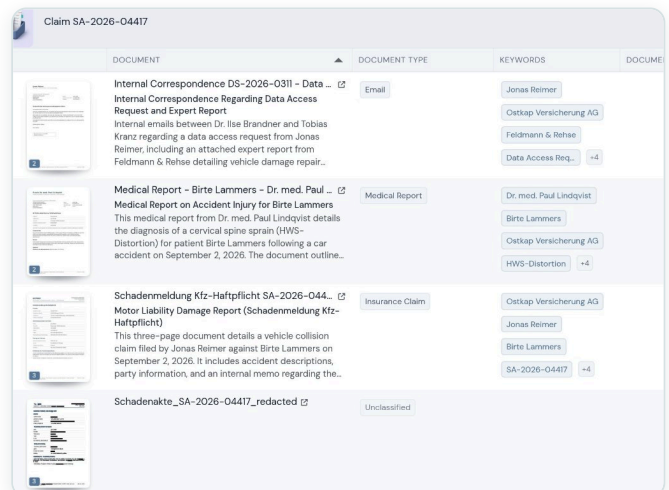
No box on top of the text. The text is gone.

"Apply redaction" creates a copy as a new document in the current Workspace. The original stays unchanged.

STEP	WHAT PAPEROFFICE DOES
Digital PDF	Text under every box is deleted from the file, and so are the image pixels underneath.
Verification	If text still sits under a box, the digital version is discarded and the copy is built from redacted page images.
Scans	The image areas are painted over, and the PDF is assembled again.
Side channels	Metadata, XMP, attachments, bookmarks, form fields, annotations, links and JavaScript are cleared.
File name	<name>_redacted.pdf, when pseudonymizing <name>_pseudonymized.pdf



1 **Redaction complete.** "Saved in" names the Workspace, the "Review depth" line shows AI reads, model and whether the redaction is word-precise.



2 **Original and copy** sit side by side in the Workspace. The list shows the original under the name AI-DMS assigned, and the copy under its file name with _redacted. What goes out is always the file with _redacted.

THE COPY TEST

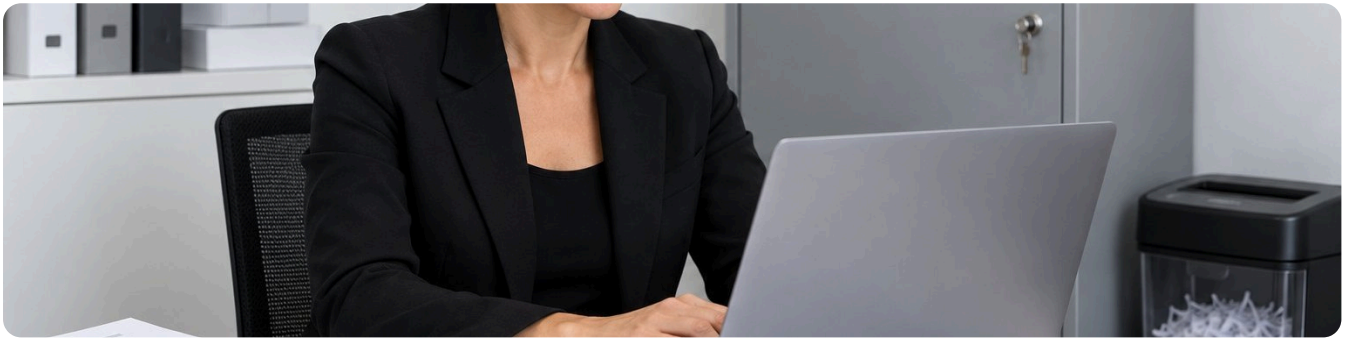
Open the copy, select the spot under a box and copy it into a text editor. This way you can see for yourself that no text remains there.

10 · RELEASING SECURELY

The copy goes out. With a password and proof.

In the completion dialog, “Send by email” opens secure delivery with the copy. Later, you reach it via

Actions > Share & export > Send by email .



Send by email
 Send selected original files by email. PDFs receive a password (email or SMS); the recipient downloads via a secure link — no additional portal code.

Schadenakte_SA-2026-04417_redacted

RECIPIENT & MESSAGE

Recipient email *

Recipient name

Subject

Personal message

SECURE ACCESS

PDF password by email DEFAULT
 The recipient receives the link and the PDF password in the same email. Opening the link and every download are logged.

PDF password by SMS HIGHER SECURITY
 The recipient receives the link by email; the PDF password is sent separately by SMS. Opening the link and every download are logged.

Delivery events are logged per document and in your account under Delivery history.

- 1 **Delivery to the regional court.** The dialog lists the copy, followed by recipient, subject and message. The password is sent via “PDF password by email” or separately via “PDF password by SMS”.

BUILDING BLOCK	HOW IT WORKS
Link instead of attachment	The recipient downloads via a secure link and does not need an account.
Validity	In the app, the link is valid for 7 days, and each file can be downloaded up to 5 times.
Proof	The “Delivery history” shows the events, the “Proof PDF” is created after the first download.
Revocation	“Revoke” blocks the link immediately, and the revocation is logged.

11 · READABLE INSTEAD OF BLACK: PSEUDONYMIZING

Names replaced, meaning kept. Pseudonymizing with Ultra.

In the **Execution** tab under “Output”, choose “Pseudonymize”. Names, addresses, dates and identifiers are replaced by consistent fictitious values. The document stays readable, for example as a training case for new colleagues.

Pseudonymization complete

The pseudonymized document has been saved. The table shows which values were replaced.

Saved in: Claim SA-2026-04417

Review depth: 2 full AI reads · large model · 7 checks · word-precise · pseudonymized

Original → in the document Copy table

Only visible here, never inside the PDF. Use the right column when you talk about the document in a chat.

Original	In the document
24000 Musterhafen	40849 Düsseldorf
Amtsgericht Musterhafen	Hannover
Deichweg 3	Feldweg 147
Hafenring 1	Am Anger 67
Landgericht Musterhafen	München
Möwenweg 7	Gartenweg 165
Märzenweg 9	Gartenweg 3

10 areas were blacked out because no safe replacement was possible.

Download Send by email Done

1 **Pseudonymization complete.** The “Original → in the document” table shows which value was replaced by what. It appears in the dialog, never in the PDF. “Copy table” copies it.

When it is available

With the Ultra AI model. The scenarios Whistleblower, Legal Proceedings, Public Publication and Anonymous Application always use black boxes.

What stays black

Health data, access credentials, signatures and special details such as religion or ethnic origin. Without a safe replacement, the area is blacked out as well.

For the chat

Select the copy, **Actions** →
Agents & workflows →
Ask Agent . In the chat, use the values of the right column.

IMPORTANT

A pseudonymized copy is not anonymous: with the table, every person can be identified again. Treat the copy and the table like the original.

12 · BATCHES WITH THE REDACT AGENT

Set it up once. Redact every batch the same way.

A Redact Agent stores the scenario, lists, AI model and target. Under **Document AI** > **Redact Agent**, click "Create Redact Agent".

- 1 "Output" tab. "New Workspace" with the template "Herausgabe_Gericht". Each run appends the agent ID, date and time, and the "Workspace name preview" shows the result.
- 1 **General:** name, "Active", AI model and output, that is, black boxes or pseudonymizing.
- 2 **Scenario and Fine-tuning:** scenario, whitelist, your own terms and instructions, as in the dialog.
- 3 **Output:** "Existing Workspace" or "New Workspace", plus the redaction color. Then "Save".
- 4 **Apply:** select PDFs, **Redact**, "Automatically with Redact Agent", choose the agent. A single PDF opens the preview first; several run server-side without a preview.

PATH	CALL
REST API	POST /documents/redact-agent-run with redact_agent_id and a pofid or file, wait for the result with client_wait
MCP	tool po_documents_redact_agent_run
Agent ID	in the overview under "Copy Redact Agent ID"

Seven questions before your first release.



Is the redacted text really gone?

Yes. Text under the boxes is deleted from the PDF and then checked again. If anything remains, the copy is built from redacted page images. Metadata, attachments, form fields and links are cleared.

Does PaperOffice detect signatures and photos?

The AI reads text. Drag a rectangle over signatures, passport photos, stamps and logos in the preview, and they are redacted like marked text.

Is the original changed?

No. The copy is a new document with _redacted in its name. The original stays unchanged where it was.

Which AI model should we use?

The default is Ultra. Premium is also word-precise, Basic redacts whole lines. With "Manual", you choose the areas yourself without AI.

What does pseudonymizing mean?

Instead of black boxes, PaperOffice inserts consistent fictitious values. This works with Ultra; the mapping is shown in the dialog, never in the PDF. Such a copy is not anonymous.

Can we redact whole batches?

Yes, with a saved Redact Agent, server-side without a preview. Via REST API and MCP, you start it from your own systems.

What does it cost?

Billing is pay as you go. The details are on the pricing page: paperoffice.ai/en/pricing.

14 · TRY IT YOURSELF IN 45 MINUTES

Download the test package. Redact a claim file.

Five fictional files from Ostkap Versicherung, in German, show the whole path: from the scenario to the rectangle and the copy.

REDACTION TEST PACKAGE

Five PDFs and a guide (README.txt)

Schadenakte_SA-2026-04417.pdf digital, 3 pages
 Zeugenaussage_Sommerfeld_Scan.pdf scan with signature and stamp
 Gutachten_Lammers.pdf medical report, 2 pages
 Auskunft_Reimer_Schriftverkehr.pdf letters on a data subject request
 Vollmacht_Reimer_Scan.pdf scan with passport photo and signature
paperoffice.ai/downloads/paperoffice-ai-test-package-redact-agent.zip



- 1 **Create a Workspace:** at app.paperoffice.ai, type "File / Case", upload all five PDFs.
- 2 **Court:** the claim file with "Legal Proceedings", whitelist "Ostkap Versicherung AG, Jonas Reimer, Birte Lammers".
- 3 **Scans:** witness statement and power of attorney in the preview, rectangles over signature, stamp and passport photo.
- 4 **Request:** the correspondence with "GDPR Request (Art. 15)", whitelist "Jonas Reimer".
- 5 **Test:** try to copy the text under a box in a copy, then the claim file with "External Sharing", Ultra and "Pseudonymize".

FILE	WHAT TO LOOK FOR IN THE PREVIEW
Claim file	Do the parties on the whitelist stay visible? Are the witness's address and phone number marked?
Witness statement (scan)	Is the text from the scan marked? Drag the signature and stamp yourself.
Medical report	With "Medical Records": are the diagnosis and medications marked?
Correspondence	Does Jonas Reimer stay visible? Are the names of third parties marked?
Power of attorney (scan)	Is the ID number marked? Drag a rectangle over the passport photo.

All companies, names, addresses, license plates and bank details in the test package are fictional, and the IBAN is a test IBAN. The cost of the test is shown on the pricing page: paperoffice.ai/en/pricing.

15 · FURTHER READING

The short version. And every module in detail.

The short guide sums up redaction in eight pages. The datasheets describe the modules that support you when you release a file.

SHORT GUIDE · 8 PAGES

Redacting personal data

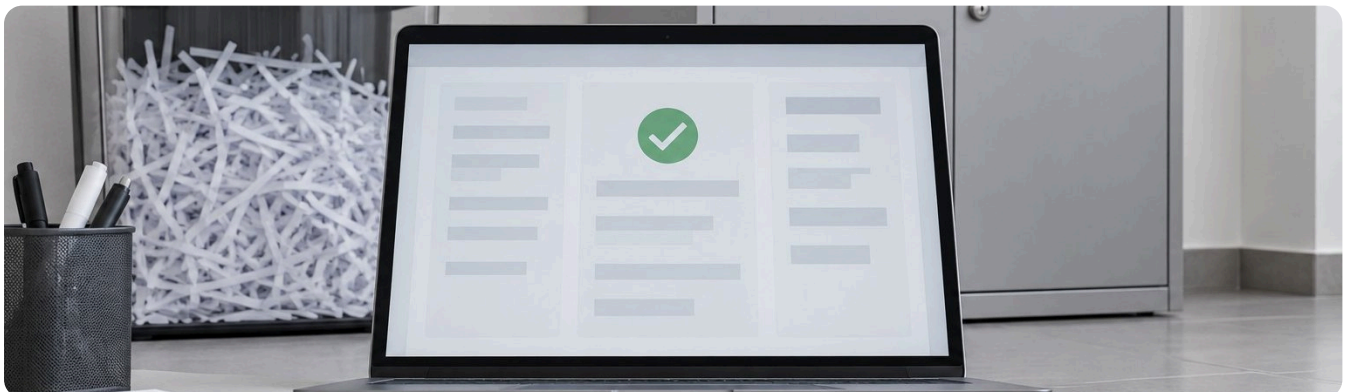
Redacting files with a marker pen? The Redact Agent marks, you decide. With a checklist and the most common questions.

paperoffice.ai/downloads/paperoffice-ai-short-guide-pii-redaction.pdf



DATASHEETS ON THIS TOPIC

DATASHEET	WHY IT MATTERS FOR REDACTION	DOWNLOAD
AI-OCR	text recognition with a position for each line, also for scans and photos	paperoffice.ai/downloads/paperoffice-ai-datasheet-ocr.pdf
Smart Search	find a name in all Workspaces before the release	paperoffice.ai/downloads/paperoffice-ai-datasheet-smart-search.pdf
Workspaces	one Workspace per case, group permissions and access protection	paperoffice.ai/downloads/paperoffice-ai-datasheet-workspaces.pdf
Secure email delivery	send the copy with a password, event log and proof PDF	paperoffice.ai/downloads/paperoffice-ai-datasheet-audit-safe-email-delivery.pdf



16 · YOUR FIRST RELEASE IN FIVE DAYS

Start with one occasion. Then add the next.

To get started, you need one occasion that comes up often, such as files for courts, and a team that reviews the preview.

Day 1 · Occasions and scenarios

- 1 Agree with the legal department and data protection which scenario applies to courts, requests, service providers and publications. One Workspace per case.

Day 2 · Lists

- 2 Whitelist with your own company and the usual parties, blacklist with internal project names and number ranges.

Day 3 · Test

- 3 Work through the test package, drag rectangles over signatures and do the copy test in a copy.

Day 4 · Redact Agent

- 4 Save one Redact Agent per occasion with scenario, lists and target Workspace, and apply it to a batch without images.

Day 5 · Release routine

- 5 Define who reviews the preview, that exclusively files with _redacted go out, and that they go via secure delivery. For the due date, add a task with a deadline to the document; "Deadline reminder (paid)" reminds you beforehand and when the deadline is reached.

MEASURE THESE FOUR VALUES

Files

Files per week released in redacted form

Added by hand

Areas your team adds or releases in the preview

Time to release

Time from the request to sending the copy

Follow-ups

Questions or corrections after a release



Measure the first weeks carefully. Only the comparison with the weeks after shows how much taping and rescanning you save.

17 · CHECKLIST AND START

Before the first file goes out.

- ☐ For each occasion, it is defined which scenario applies: court, request, service provider, publication.
- ☐ Names that must stay visible are on the whitelist, your own terms on the blacklist.
- ☐ Before a release, someone searches for the affected names in all Workspaces, including processed scans.
- ☐ Every single file is reviewed in the preview before release, page by page.
- ☐ Signatures, passport photos, stamps and logos are covered with a rectangle.
- ☐ Exclusively the copy with _redacted goes out, via secure delivery with a password.
- ☐ For batches without images, a saved Redact Agent exists for each occasion.
- ☐ A run with the test package and the copy test are done.

GET STARTED NOW

Create an account, set up a Workspace for your first case, upload the test package and redact the claim file with a preview. Billing is pay as you go.

app.paperooffice.ai**MORE ABOUT THE REDACT AGENT**

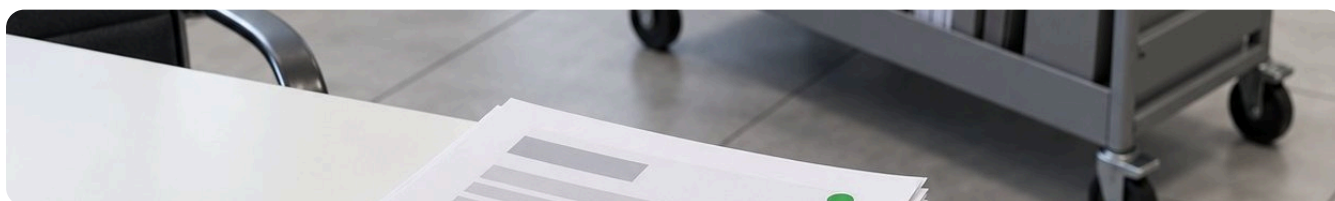
Scenarios, preview, pseudonymizing and the path via API and MCP on one page.

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

**QUESTIONS ABOUT THIS GUIDE?**

Contact us if you have questions about the setup or your occasions.

paperooffice.ai/en/contact



As of October 2026. Screenshots are from a test account, and the photos are AI-generated scenes. All examples, companies, names, addresses, license plates, bank details and values in this guide are fictional. PaperOffice, paperooffice.ai.