

Automating document processes with PaperOffice AI

What an AI-DMS has to deliver today, how to evaluate vendors and how to roll it out step by step – for decision-makers and IT in companies and SMEs.



Contents

1	Starting point: paper, mailboxes, scattered storage	3
2	What an AI-DMS has to deliver today	4
3	Evaluation criteria: a checklist for your selection	6
4	How PaperOffice AI works	8
5	Data location and security	10
6	Rollout in steps	12
7	Measuring results	14
8	Next steps	15

About this guide

This guide is for management, business units and IT teams who want to select and introduce an AI-DMS or a solution for intelligent document processing (IDP).

It describes where document processes get stuck today, what a system has to deliver, which questions to ask vendors and how to start in manageable steps. The requirements and questions in chapters 1 to 3 apply to any selection; next to each one you will find how PaperOffice AI implements it.

All information on PaperOffice AI reflects paperoffice.ai as of October 2026. Plans and allowances are listed on the [pricing page](#).

At a glance

- An AI-DMS makes AI processing part of filing. Every capability works with the same permissions and the same audit trail.
- Start small: one workspace, one document type, one flow – usually incoming invoices or the mailroom.
- Measure success through cycle time, rework and the share of documents that pass review.
- PaperOffice AI processes and stores documents in its own data center in Navarra, Spain. With BYOS, your files are stored encrypted on your own storage.

Starting point: paper, mailboxes, scattered storage

Document problems sound trivial and still cost time, money and patience every single day.

Documents reach a company through many routes: email, scanner, upload, network drive – the inbox is scattered. They are filed on file servers, in cloud drives, in mailboxes and in scan folders, and nobody knows every place. Values from receipts, forms and orders reach the ERP, the CRM and spreadsheets through a keyboard.

Where it usually gets stuck

Document chaos Files sit everywhere and nobody knows which version counts.	Manual data entry Skilled staff retype what the document already says.	High error rate Retyped numbers and missed fields create rework.	Slow processes Approvals and filing wait for people instead of rules.
Nothing is findable Search finds file names but not the content.	Scaling More volume has meant more headcount instead of more throughput.	Compliance risk Deadlines, evidence and access cannot be proven.	Knowledge loss Knowledge sits in heads and mailboxes instead of the archive.

The time is lost between the steps

It is rarely the work that takes long. It is the waiting in between: intake, forwarding, approval, query, filing. A wrong value gets expensive the moment it sits unnoticed in accounting, in the ERP or in a contract. And whoever has to prove who saw, changed, shared or deleted which document, and when, cannot build that record after the fact.

Collect first, then sort

A clean-up project lasts a few weeks. An AI-DMS therefore starts at intake: it picks documents up where they already sit, recognizes every type and files them by your rules. You do not have to tidy up first – type, sender, date and references come from the content, not from the file name.

FOR YOUR SELECTION

Check early whether a system can connect your existing storage instead of replacing it. PaperOffice reads from email, cloud storage and business systems and hands results back through API or workflow; existing storage can be connected as your own storage mount.

What an AI-DMS has to deliver today

An AI-DMS is a complete document management system in which AI processing is part of filing. Its capabilities fall into four functional areas: Capture, Understand, Create and Agents.

What matters is that every capability is part of the same platform and works with the same permissions and the same audit trail – in the application, via API and in the AI assistant.



Capture

Scan · inbox · OCR – every input lands in the workspace, OCR included.

WHAT A SYSTEM SHOULD DO

- All channels end up in the same inbox: scanner, upload, mobile device, folders on NAS and servers, S3 storage and the API.
- Read text from skewed scans, photos and handwriting, with a position per line.
- Split scan batches into individual documents.
- Detect type, direction, language and sender without training.

HOW PAPEROFFICE AI DOES IT

- AI-OCR in more than 120 languages, bounding boxes per line
- Classification with your own document types
- Split Agent for batch scans
- Connectors for SFTP, S3, WebDAV and reverse SSH
- Mobile app as a PWA with scanner



Understand

Analyze · connect – relationships, key figures and timelines from your document base.

WHAT A SYSTEM SHOULD DO

- Search by meaning, not by the exact word – in scans too and across languages.
- Recognize people, companies, amounts, dates and reference numbers and merge spellings.
- Show related documents as one case: order, delivery note, invoice, contract.
- Build metrics that lead back to the document.

HOW PAPEROFFICE AI DOES IT

- Smart search: semantic, full text, OCR text
- Entities and knowledge graph
- Document chat with the source document
- Analytics dashboard
- Audit center with findings by severity



Create

Edit · deliver – PDF, signature, speech, translation and audit-safe delivery.

WHAT A SYSTEM SHOULD DO

- Convert files to PDF and PDF/A, and PDF back to Office formats.
- Request signatures straight from the document and keep the evidence.
- Translate without altering technical terms and product names.
- Send documents securely, with a log of every step and revocation.

HOW PAPEROFFICE AI DOES IT

- PDF Studio
- PaperOffice Sign: simple electronic signature (SES) under eIDAS
- AI translation into 35 languages
- Audit-safe email delivery with proof PDF



Agents

Autonomous AI agents – extract, check, approve and answer, automatically.

WHAT A SYSTEM SHOULD DO

- Extract fields by description instead of template, every value with its location and confidence.
- Send uncertain values to people instead of the target system.
- Remove personal data before sharing, not just cover it.
- Check documents for manipulation.
- Automate flows with triggers, conditions and approvals, and log every run.

HOW PAPEROFFICE AI DOES IT

- IDP Agents
- Human-in-the-loop
- Redact Agent
- Trust Agent
- Workflow Agents
- Chat agent that cites its sources

MODULE OR AGENT?

Modules are platform capabilities. Agents are flows that use those capabilities one after another, following your rules.

Classic DMS, cloud DMS, AI-DMS

The difference is not the filing, it is what the system does with the document.

	Classic DMS	Cloud DMS	PaperOffice AI-DMS
Text recognition	OCR as an add-on	Basic OCR	AI-OCR with bounding boxes
Data extraction	Manual tagging	Fixed templates	IDP by prompt
Search	File name and full text	Full text	Semantic with embeddings and chat
E-signature	Third-party tool	Add-on module	Included: SES under eIDAS
Redaction	Manual	Not available	Redact Agent, AI-assisted
Document verification	None	None	Trust Agent with fraud indicators
Mobile	Desktop client	Browser	PWA with scanner, offline, signature
Storage location	Own server, own maintenance	At the vendor	EU cloud or your own storage
AI assistants	Not available	Export required	MCP native: Claude, ChatGPT, Cursor

Comparison by system category, as presented on paperoffice.ai. Individual products may differ.

Evaluation criteria: a checklist for your selection

These questions make vendors comparable. Ask to see every answer on your own documents – including poor scans and exceptions.

Question for the vendor	PaperOffice AI
A · CAPTURE AND RECOGNITION	
☐ Do all input channels come together in one place? Paper, email, upload, network drive, cloud storage, API	Upload, mobile scan, email, storage mount, connector and API all end up in the same inbox.
☐ Does the system read difficult documents too? Skewed scans, photos, handwriting, tables, several languages	AI-OCR in more than 120 languages, bounding box and confidence per line, tables as structure, handwriting with confidence.
☐ Are document types detected without training? Effort for new types and new senders	Type, direction, language and sender in one pass; you describe your own types in natural language, active immediately.
B · EXTRACTION AND CHECKS	
☐ How are fields defined? Fixed template per sender or a description in your own words	IDP Agents work zero-shot by prompt; the catalog with more than 60 document types is a starting point, not a requirement.
☐ Can every extracted value be traced? Location in the original and confidence per field	Every field with value, confidence and bounding box, per page.
☐ What happens to uncertain values? Thresholds per field, review queue, four-eyes principle	Human-in-the-loop by rules per document type, field, customer, confidence or anomaly; several reviewers per case, optionally as blind review.
☐ Is content checked, not just read? Arithmetic check, format validation, master data, manipulation	Totals against line items, IBAN and VAT ID against their rules, bank details against the master record; the Trust Agent checks PDFs for hidden content and overlays.
C · PROCESSES AND HANDOVER	
☐ Can flows be built without programming? Triggers, conditions, approvals with deadline and escalation	Workflow Agents from one sentence in the AI builder or from building blocks; every run logged with steps, inputs, results and duration.
☐ How do checked values reach ERP, CRM or accounting? Handover only after checks and approval	Export, email, webhook or REST API; the workflow hands over only once checks and approvals are complete.
☐ Is e-signature built in? Signature level, evidence, filing with the document	PaperOffice Sign: simple electronic signature (SES) under eIDAS with an SMS one-time code; the signed PDF and the certificate are filed with the document.

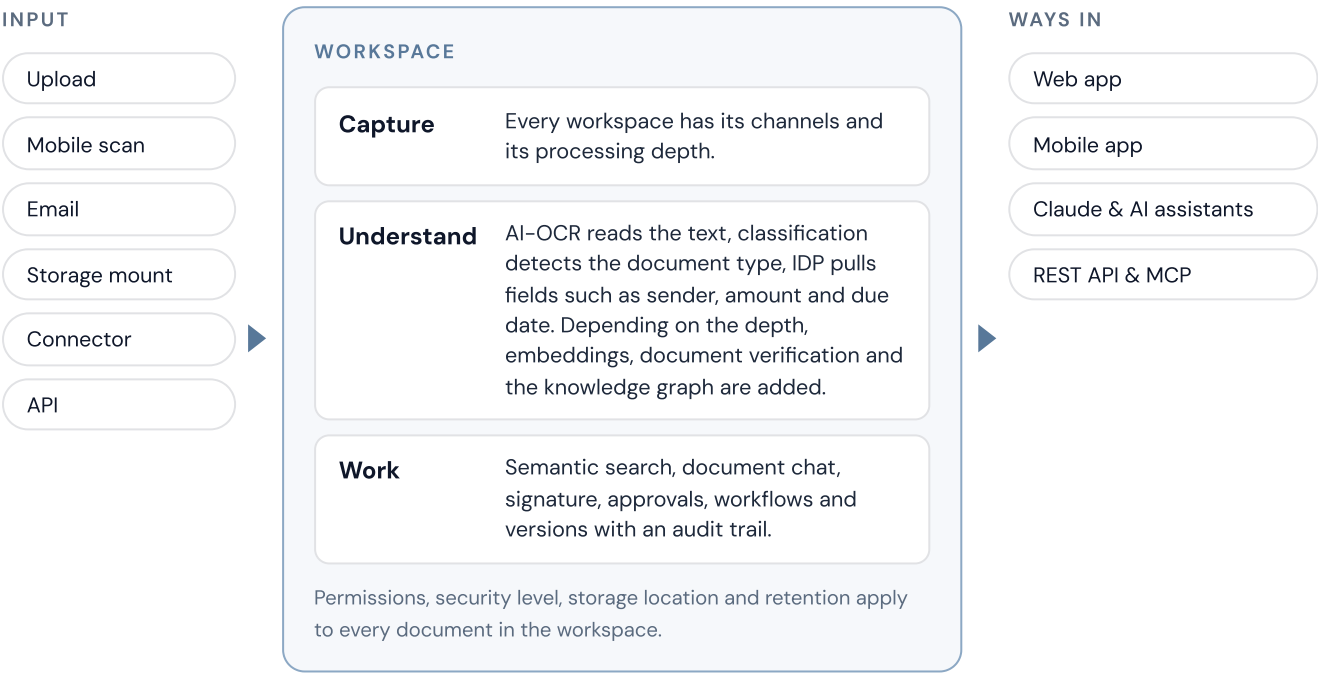
Question for the vendor	PaperOffice AI
D · SEARCH AND KNOWLEDGE	
☐ Does search find content without the right keywords? Meaning instead of word matches, scans, other languages	Semantic search, full text, OCR text and entities with one shared ranking, across languages.
☐ Do search, chat, agents and API respect permissions? No special access for AI and interfaces	Agents, REST API and MCP server work with the workspace permissions; without the right, a document appears neither in the results nor in the preview.
☐ Do answers cite their source? Document, page, relevance	The document chat names document, page and relevance for every statement.
E · SECURITY, DATA LOCATION AND EVIDENCE	
☐ Where are documents processed and stored? Location, own infrastructure, option for your own storage	In our own data center in Navarra, Spain (EU). With BYOS, the files are stored encrypted on your own storage.
☐ How finely can protection levels and permissions be set? Separation by client, department, project	Workspaces of the types Standard, Business and Compliance; read, write and admin permissions for users and groups per workspace.
☐ Is the archive revision-safe? Immutability, versions, retention, proof of deletion	Revision-safe (WORM, versions, log): audit trail with integrity hash, retention periods per document type. A basis for GoBD requirements.
☐ What evidence does the vendor provide? Certificates, DPA, TOMs, audit reports	ISO/IEC 27001 certified (certificate under NDA); SOC 2 Type II in preparation. DPA under Art. 28 and TOMs under Art. 32 GDPR from the generator in the app.
F · ACCESS, INTEGRATION AND COST	
☐ Is there a complete API and a connection to AI assistants? The same functions as in the app, webhooks, MCP	Every function via REST API, webhooks for events, MCP server with more than 300 tools; a verified connector in the Claude directory.
☐ How much effort does taking over the existing archive take? Migration, clean-up, parallel operation	No clean-up beforehand; existing archives come in as ZIP, via storage mount or from an ELO export. With a storage mount, files do not have to move.
☐ How is it billed? Seat licenses or billing for work done	A monthly plan; usage within it is paid in credits for work done. Users are included in every plan, no seat licenses.

The column "PaperOffice AI" reflects paperoffice.ai as of October 2026.

How PaperOffice AI works

PaperOffice is the DMS for companies and SMEs: OCR, IDP, signature and workflows in one platform. Four functional areas on one DMS foundation, with three ways in.

Every module works on the same document in the same workspace – no second system, no second archive. Documents arrive through any channel, are processed automatically per workspace and are ready right away: in the app, on mobile, via API and in your AI assistant.



Workspaces: the frame for every document

A workspace is more than a folder: it carries security level, processing depth, storage location, retention and permissions. Workspaces separate documents by client, department or project, and what a workspace defines applies to every document in it – in app, API and MCP.

Processing depth: four levels, chosen per workspace

Archive	Basic	Premium	Ultra
Store only	Read and organize	Understand and connect	Verify and link
File storage with your own tags and metadata, versions and audit trail; can be processed by an agent later	AI-OCR with text layer and bounding boxes, classification, language and folder suggestion, base fields and full-text search	Everything in Basic, plus extended IDP fields, parties and financial data, embeddings, semantic search, document relations	Everything in Premium, plus document verification with fraud indicators, vision, entities, knowledge graph, line items
For legacy archives and drafts	For digital documents with quick filing	For accounting and contract management	For compliance and deep analysis

The depth applies automatically to every new document in the workspace. Individual documents can be reprocessed later at a higher level.

Three ways in, one dataset



FOR TEAMS

PaperOffice App

Work directly in the app

- Web app with six views from list to Kanban, approvals and versions
- Mobile app as a PWA with scanner, offline queue and signature, no app store needed
- Review and approve on the go



FOR AI USERS

Claude connector

Work with the AI assistant

- Listed as a verified connector in the Claude directory
- Search, read, extract, approve, send for signature, archive – in plain language
- Access only with your permissions
- Via MCP also ChatGPT, Grok and Cursor



FOR DEVELOPERS

API & MCP

Integrate into your own software

- Every function headless via REST, webhooks for events such as processed, checked or approved
- MCP server with more than 300 tools, OAuth 2.1 or a user token
- Remote server, no installation

One dataset · one rights model

All three paths can be combined

Example: incoming invoices

Extraction is one step out of five. What deviates from the rule is what lands on a desk.

1

Capture

Scanner, upload, mounted storage and mobile device feed one inbox. Classification separates invoice, delivery note, reminder and quote.

2

Check

The IDP agent reads header data, line items, totals and references. Then the rules run: arithmetic check, duplicate search, VAT ID, IBAN field and order reference.

3

Approve

Routing by amount, cost center or supplier, with deadline, reminder and responsibility per role.

4

Post

Record export or handover via the REST API; webhooks report events such as processed and edited documents to your ERP.

5

File

Filing with a retention period; due dates and early payment discounts appear in financial analytics.

E-invoices in ZUGFeRD, XRechnung, Factur-X and PEPPOL format take the same path as PDF and scan.

PRICING MODEL

Billing runs through a monthly plan: usage within it is paid in credits for work done, not in seats for people – users are included in every plan. The MCP connection is included in every plan; processing is billed, not the connection. To get started there is a permanently free plan, no credit card required. Plans and prices are on the pricing page.

Data location and security

Control that can be verified: where your documents are stored, how they are protected and who accesses them.



Data location: one data center in Navarra, Spain

Processing and storage run in PaperOffice's own data center in Navarra – our own hardware, no public cloud, no shared resources. That is where your documents are processed, stored and backed up.

Your own storage (BYOS)

Connect your own storage over SFTP, such as a NAS or a server. The files are stored encrypted on your storage; PaperOffice keeps the index, metadata and AI results in its own data center in Navarra. A second storage can serve as a mirror: if one fails, the other takes over.

AI without third parties

No external AI providers, no AI subprocessors and no training on your data. Databases, storage, AI models, email servers and backup systems run on our own hardware.

Encryption and processing

AES-256 at rest, TLS 1.3 in transit, keys stored outside the database. Processing on a RAM disk, with processing data deleted after completion.

Access

Permissions per workspace for users and groups; from Business workspaces upward, password, SMS or two-person approval. Login monitoring alerts on new devices and IP addresses; "No, this was not me" locks the account and ends all sessions.

Log and retention

Every action with time, user and context, entries with an integrity hash, not editable afterwards. WORM protection and retention periods per workspace, which cannot be switched off once activated.

Three workspace types, one rights model

Standard	Business	Compliance
For everyday operations	For sensitive records	For high security
AES-256 encryption and RAM disk processing, whitelist access lists and roles, versions with audit trail, permanent deletion	Everything in Standard, plus recycle bin, WORM write protection and workspace password, external sharing, SMS or two-person approval	Everything in Business, plus hardware and chunk encryption, IP whitelist, mandatory 2FA, watermarks and protected PDF delivery
For teams, projects and daily work	For accounting, HR and contracts	For regulated industries and audits

Workspace type and processing depth are independent. Upgrading is possible any time, downgrading is not, for compliance reasons.

EVIDENCE AND DOCUMENTATION

- ISO/IEC 27001 certified (certificate under NDA); SOC 2 Type II in preparation.
- DPA under Art. 28 and TOMs under Art. 32 GDPR: the generator in the app produces both from your company data, as a PDF.
- Security whitepaper and penetration test report (under NDA) on request from the Enterprise plan.
- Processing under the GDPR; revision-safe (WORM, versions, log); a basis for GoBD requirements.

The certificates cover the operation of PaperOffice, not your processes. They make your own audit easier because processing, storage location and logging are documented. PaperOffice supplies the technical evidence – log, retention, versions and reports –, while describing your own processes stays your job.

What data protection officers ask first

Does the detection of personal data really find everything?

It finds the common categories, but do not rely on it blindly. The preview shows every location marked in the document, and you add or remove areas before anything becomes final.

Can an administrator change log entries?

No. Entries are immutable and carry an integrity hash. Even the deletion of a document remains as an entry.

How do we answer an access request?

Search collects every document about one person, you check what is still under a retention obligation and export the rest redacted. The whole process appears in the audit trail.

Is the redaction really gone?

Yes. Redacted areas are removed from the PDF, not just covered. The original is kept as its own version, the redacted document is filed as a new one.

Does the AI see all our documents?

Only what the key allows. Permissions apply per workspace, and every access is in the audit trail. Keys can be blocked at any time, which ends access immediately.

Where do we find more checkpoints?

The GDPR checklist for document processing on the downloads page of [paperoffice.ai](#) summarizes eleven checkpoints, each with the matching PaperOffice AI functions.

Rollout in steps

Start small: one workspace, one document type, one flow. Further modules come later without rebuilding what exists.

The first flow stands in a day. How fast the rest follows depends on permissions, systems and approval paths, not on recognition. You do not need your own developers to start: you describe document types, fields and rules in the app.

1

Define the pilot

- Choose the document type that eats the most time. Usually that is incoming invoices or the mailroom.
- Create a workspace, choose the processing depth, upload documents – free, with no credit card and no term.
- Record beforehand how long a case takes today and how much rework it needs (measurement sheet in chapter 7).

Result: a clearly defined flow with baseline values for comparison.

2

Document types and existing archive

- The catalog recognizes common types immediately; you describe your own types in natural language. Per type you set folder suggestion, retention period and workflow trigger.
- Bring in existing archives without tidying up first: as ZIP, via storage mount or from an ELO export. The mobile app scans paper straight into the workspace.
- Whatever cannot be assigned lands in a review list. A person decides there, and that decision can be kept as a rule for the next case of the same kind.

Result: every document has a type, workspace, folder and retention period.

3

Set up and test agents

- IDP agent: pick a template from the catalog or describe the fields in your own words. Every field gets a type, a prompt and rules – required, pattern, allowed values, page range.
- Test one to ten times, see the hit rate per field, then go to production in series: manually, in batch with up to 200 documents per run, in a workflow or automatically in the workspace.
- Describe the flow in one sentence in the prompt studio, such as “check new invoices, request approval above 5,000”; this becomes the instruction for the workflow agent.

Result: an agent whose hit rate per field you know.

4

Roles, permissions and review

- Assign read, write and admin permissions to users and groups per workspace.
- Define review rules: per document type, field, customer, below a confidence or on anomalies. Critical fields such as amount or bank details get a stricter threshold.
- Four-eyes principle and blind review where needed. After the first weeks the audit center shows which fields genuinely need review.

Result: people decide where it matters; everything else is approved automatically.

5

Handover to your systems

- Checked fields go to ERP, CRM or accounting as JSON or CSV, by export or REST API; webhooks report every processing run.
- The workflow hands over only once checks and approvals are complete. The document is kept in the workspace while your system receives the data and a reference to it.
- Then take on the next document type or workspace – further modules come later without rebuilding what exists.

Result: no more retyping between the document and the target system.

Pilot profile

Document type

e.g. incoming invoices

Volume per month

Input channels

scanner, upload, folder

Required fields

Review rules

thresholds, four eyes

Target system, handover

Responsible

business unit, IT

Storage location

EU cloud or BYOS

Measuring results

Whether it helps shows in cycle time, rework and the share of documents that pass review. Analytics show volume, deadlines and open cases per workspace.

Collect the values before the pilot and compare afterwards using the same definitions. The table shows where to find the data in PaperOffice AI and doubles as your measurement sheet.

Metric	What it shows	Where you see it in PaperOffice AI	Before	After
Cycle time	Time from intake to handover or filing	Log per workflow run with step, input, result and duration		
Rework	How many fields and documents are corrected	Audit trail with old and new value per correction, reviewer decisions		
Hit rate per field	How often a field is right without correction	Test runs and batch KPIs of the IDP Agents, confidence per field		
Review rate	Share of documents that go to people	Human-in-the-loop queue; the threshold controls the volume		
Open cases and deadlines	What sits where and until when	Task inbox with deadline and owner, analytics per workspace		
Volume	Documents and pages per period, by type and workspace	Analytics dashboard: estate, processed pages, distribution, trends		
Anomalies	Differing bank account, arithmetic errors, unusual access	Audit center: AI findings with severity from low to critical		
Cost per case	Consumption per document type and flow	Credits for work done, the same price in app, API and MCP		

Every tile and every distribution in the analytics dashboard opens the document list with the matching filter. Metrics and raw data can be collected regularly by export or through the REST API; an analysis stops at the access rights.

Next steps

Test with your own documents, try the Claude connector and get support when you need it.

TRY IT LIVE

Start for free

Create a workspace, choose the processing depth, upload documents – free, with no credit card and no term.

app.paperoffice.ai

IN CLAUDE

Try the connector

When you connect, choose the demo and try PaperOffice in a read-only demo workspace with sample documents – no account needed.

paperoffice.ai/en/claude

ADVICE

Discuss your flow

Our experts are happy to help you choose the right solution.

paperoffice.ai/en/contact

For the first conversation

- Sample documents of the pilot document type, including poor scans and exceptions
- Volume per month and today’s input channels
- Required fields and the target system for the handover
- Approval paths, review rules and owners
- Requirements for data location and storage: EU cloud or your own storage
- Baseline values from the measurement sheet in chapter 7

Further reading on paperoffice.ai

Pricing and plans	paperoffice.ai/en/pricing
AI-DMS overview	paperoffice.ai/en/ai-dms-document-management
Security and compliance	paperoffice.ai/en/company/security
Locations and data center	paperoffice.ai/en/company/locations



PAPEROFFICE ENTERPRISE OPERATIONS, S.L.U. · Calle Cataluña 8, Planta 4, Bajo · 31006 Pamplona (Navarra) · Spain
· paperoffice.ai

The contracting party for all PaperOffice services worldwide is PAPEROFFICE ENTERPRISE OPERATIONS, S.L.U., based in Pamplona. The PaperOffice brand was founded in Germany in 2002.

As of October 2026. Features, plans and allowances may change; the information on paperoffice.ai is authoritative. Third-party product names mentioned are trademarks of their respective owners.