BOREHOLEAI

Enterprise Data Security & Privacy Assurance

How BoreholeAI protects customer borehole logs, extracted data, outputs, and enterprise workflows

Prepared for	Prospective enterprise clients
Prepared by	BoreholeAI
Status	Rev 2
Date	17 May 2026

This document is intended for enterprise evaluation discussions.

Executive summary

Enterprise customers are right to ask whether BoreholeAI can process sensitive ground investigation records without retaining them, training on them, or exposing them to unnecessary human or third-party access. BoreholeAI implements a layered security model combining privacy-by-architecture, short retention, purge-on-download enterprise processing, strict access controls and, for the highest-security clients, customer-hosted deployment.

BoreholeAI's strongest security position is that the product is not a general-purpose document chatbot. It is a purpose-built borehole log extraction engine. The pipeline is OCR-first, spatially grounded, and largely deterministic. AI is used selectively for document understanding tasks, while the original full PDF is not sent directly to AI models.

For most enterprise clients, the preferred commercial deployment should be "Enterprise SDK" with purge-on-download: the client runs BoreholeAI from its own IDE or automation environment, files are uploaded for processing only, outputs are written back to the client's disk, and server-side files are deleted immediately after download confirmation. For organisations that cannot allow any external processing, BoreholeAI will offer a dedicated private-cloud or on-premise deployment option.

Security statement

For enterprise SDK usage, BoreholeAI processes documents only to generate the requested engineering outputs. Customer documents and result files are not used for model training, are not reviewed by BoreholeAI staff unless explicitly requested for support, and are deleted immediately after the SDK confirms successful local download under the enterprise purge-on-download mode. Minimal job metadata may be retained for billing, audit, abuse prevention, and support history. For customers requiring stronger isolation, BoreholeAI can discuss dedicated private-cloud or customer-hosted deployment.

1. Client concern to solve

For engineering firms, contractors, miners, asset owners, and public-sector bodies, borehole logs may contain project names, coordinates, investigation locations, contractor details, client names, site constraints, and commercially sensitive ground risk information. A client may accept cloud processing only if BoreholeAI can clearly answer four questions:

1. Where does the file go, and for how long?
2. Can BoreholeAI staff or third parties see the file or extracted information?
3. Is any customer data used to train or improve BoreholeAI models?
4. Can the client receive evidence that files were deleted and that access is controlled?

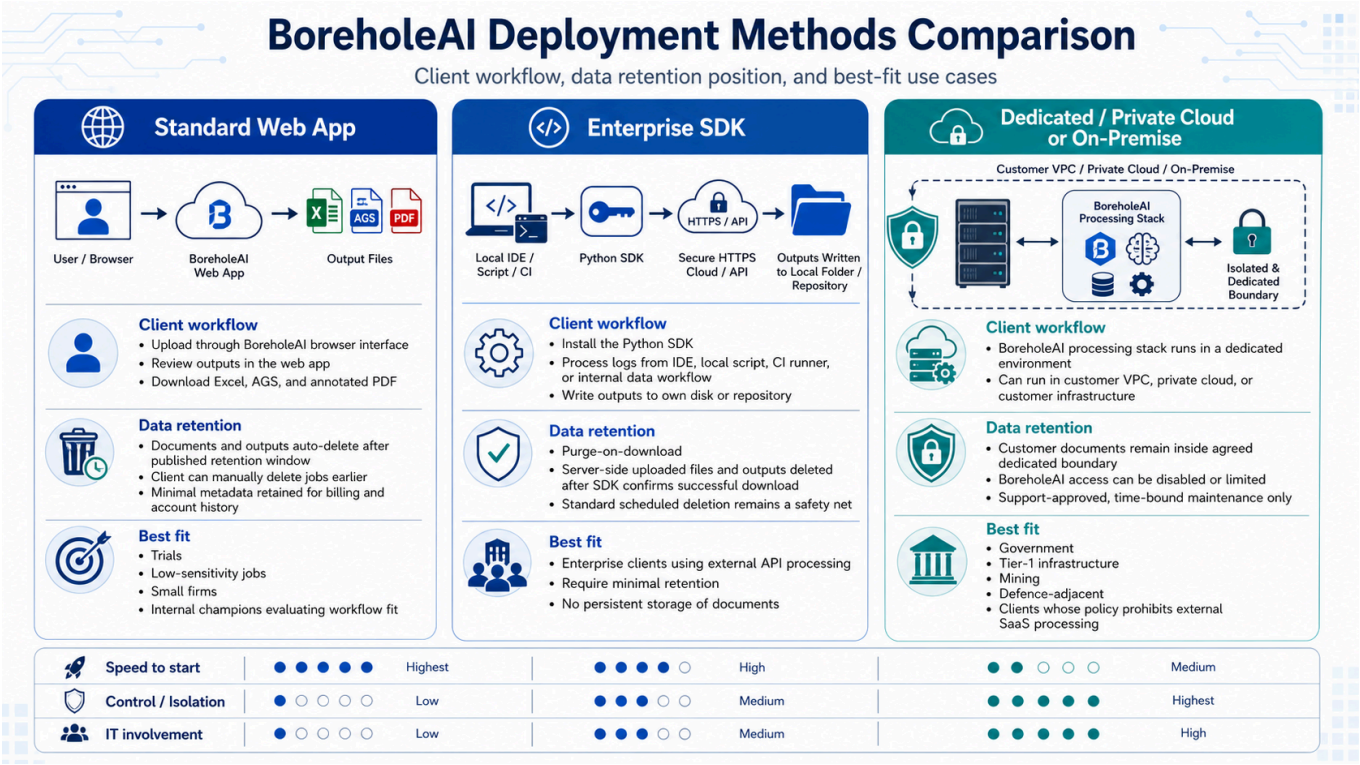
The following sections define the recommended security model.

2. Deployment and retention model

BoreholeAI comes with three deployment modes for enterprise clients.

Mode	Client workflow	Data retention position	Best fit
Standard web app	Client uploads through the BoreholeAI browser interface, reviews outputs, and downloads Excel/AGS/annotated PDF.	Documents and outputs auto-delete after the published retention window. Client can manually delete jobs earlier. Minimal metadata is retained for billing and account history.	Trials, low-sensitivity jobs, small firms, or internal champions evaluating workflow fit.
Enterprise SDK	Client installs the Python SDK, processes logs from an IDE, local script, CI runner, or internal data workflow, and writes	Purge-on-download: server-side uploaded files and outputs are deleted after SDK confirms successful download.	Enterprise clients that can use external API processing but require minimal retention and no

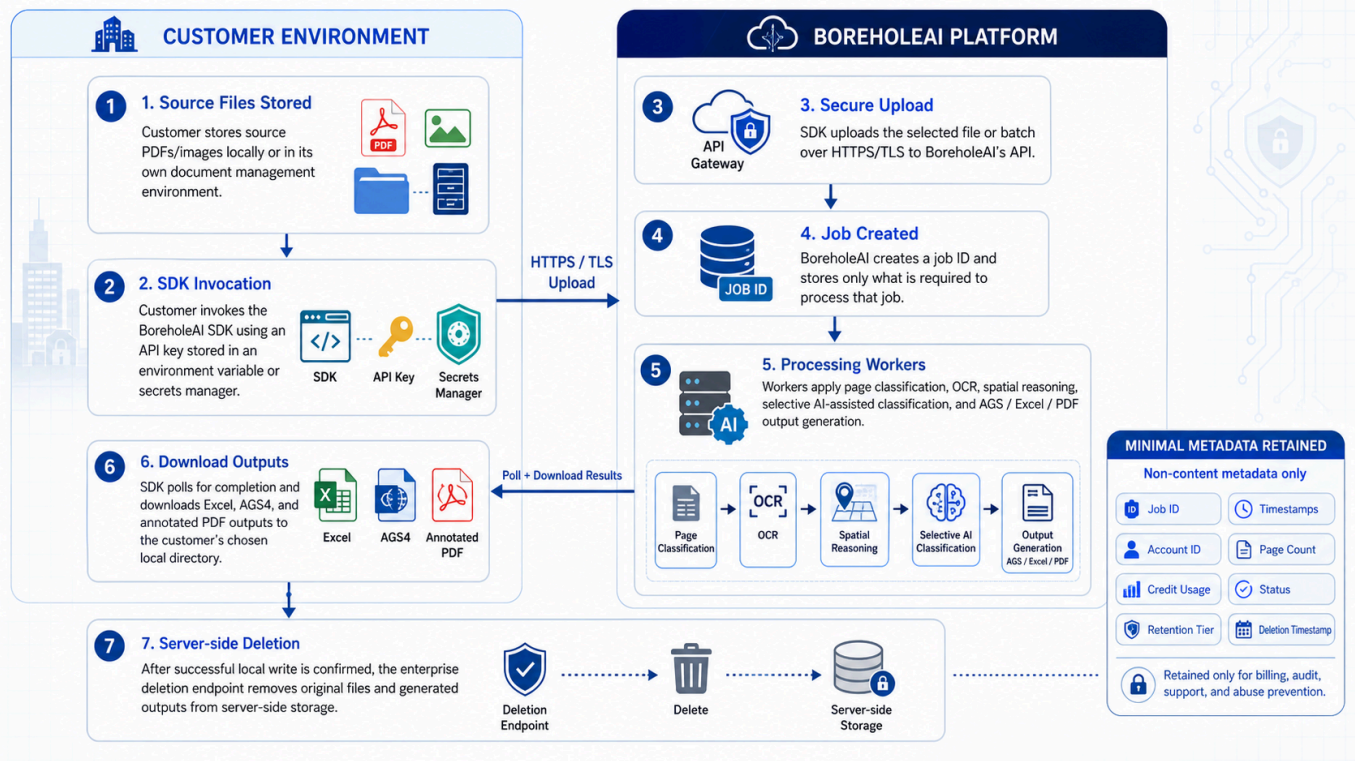
	outputs to its own disk or repository.	Standard scheduled deletion remains a safety net.	persistent storage of documents.
Dedicated/private-cloud or on-premise	BoreholeAI processing stack runs in a dedicated environment, customer VPC, private cloud, or customer infrastructure.	Customer documents remain inside the agreed dedicated boundary. BoreholeAI access can be disabled or limited to support-approved, time-bound maintenance.	Government, Tier-1 infrastructure, mining, defence-adjacent, or clients whose policy prohibits external SaaS processing.



3. Data flow for Enterprise SDK

The Enterprise SDK workflow is as follow:

1. The customer stores source PDFs/images locally or within its own document management environment.
2. The customer invokes the BoreholeAI SDK using a BoreholeAI API key.
3. The SDK uploads the selected file or batch over HTTPS/TLS to BoreholeAI's API.
4. BoreholeAI creates a job ID and stores the PDF document temporarily for processing.
5. Processing workers apply page classification, OCR, spatial reasoning, selective AI-assisted classification, and AGS/Excel/PDF output generation.
6. The SDK polls for completion and downloads Excel, AGS4, and annotated PDF outputs to the customer's chosen local directory.
7. After the SDK confirms successful local write, the enterprise deletion endpoint removes original files and generated outputs from server-side storage (Delete-on-Download Trigger).
8. BoreholeAI retains only minimal non-content metadata needed for billing, audit, support, and abuse prevention: job ID, timestamps, account ID, page count, credit usage, status, retention tier, and deletion timestamp.



4. Privacy-by-architecture

BoreholeAI achieves high standard security with the following architecture.

Control	How it works	What it means?
OCR-first processing	The original PDF/image is first processed by OCR and computer vision. Subsequent AI-assisted stages work on extracted fragments, rather than the whole original document.	Reduces unnecessary exposure of full documents to AI systems.
No third-party OCR	The OCR stage uses BoreholeAI-controlled infrastructure and an open-source OCR engine rather than sending pages to Google Vision, AWS Textract, or Azure Computer Vision.	Keeps the highest-volume document processing stage under BoreholeAI's controlled environment.
Selective AI	AI is used only where classification or reasoning is required, such as text classification. Parsing logic remains deterministic/rule-based.	Reduces black-box risk and supports engineering auditability.
No training on customer data	Customer uploads, extracted text, and outputs are not used to train, fine-tune, or benchmark models without a separate written data-sharing agreement.	Addresses the common fear that confidential project data becomes part of a vendor's model.

5. Encryption and network security

BoreholeAI commit to the following baseline controls for all cloud-hosted modes:

- TLS 1.2 or higher for all data transmitted between the browser/SDK/API client and BoreholeAI services.
- HTTPS-only access with HTTP redirected to HTTPS.
- AES-256 server-side encryption at rest for uploaded files and results stored in object storage.
- Encryption at rest for database records holding account, job, billing, and retention metadata.
- API keys stored as salted hashes rather than plaintext; full keys shown only once at creation.
- Separate development and production API keys for enterprise customers, with immediate revocation capability.
- Rate limiting and abuse prevention on public API endpoints.
- DDoS and edge protection through the web/API hosting layer where applicable.

6. Access control and internal visibility

BoreholeAI designs the system so staff do not access customer content during normal operations; access is restricted, logged, and permitted only for defined operational reasons.

Area	Control
User isolation	User-scoped storage paths and row-level security so each account can access only its own jobs, files, credits, and settings.
Support access	No browsing of customer files. Support access only after customer request, for a defined job ID, for a defined time window.
Developer/admin access	Least-privilege admin roles. Separate production access from normal development.
Back logging	Back logs must not contain source PDFs, extracted full content, or sensitive payloads. Log only job IDs, timestamps, status, error class, and performance metrics.
Deletion	Deletion endpoint removes uploaded files, result files, and job content records. Deletion event is recorded with timestamp and retention mode.

7. Data retention policy

The retention policy applies to three tiers, namely standard web app, enterprise SDK and customer hosted.

Data category	Standard web app	Enterprise SDK	Dedicated/customer-hosted
Original uploaded files	Deleted after 10 days or earlier if the user deletes the job.	Deleted immediately after successful SDK download confirmation; 10-day scheduled deletion as safety net.	Controlled by customer environment and contract.
Generated outputs: Excel, AGS4, annotated PDF	Deleted after 10 days or earlier if the user deletes the job.	Deleted immediately after successful SDK download confirmation; 10-day scheduled deletion as safety net.	Controlled by customer environment and contract.
Job metadata	Retained for billing, account history, audit, support, and abuse prevention. Exclude source content.	Retained for billing, account history, audit, support, and abuse prevention. Exclude source content.	Configured by customer policy.

8. Third-party services and subprocessor

BoreholeAI maintains high standard extraction service with the following subprocessor schedule.

Service	Purpose	Customer content exposure
Supabase	Authentication, PostgreSQL database, object storage, account/job metadata and file storage.	May hold uploaded files/results during retention window and job metadata.
Vercel	Frontend deployment/edge layer depending on architecture.	Should not store source documents unless API routes temporarily handle traffic; clarify architecture.
Google OAuth	Optional sign-in method.	Email/account identifier only, if used.
AI model provider, if applicable	Narrow classification or parsing tasks on extracted text fragments/crops.	Should not receive full PDFs, full logs, application logs, or outputs. For high-security clients, use self-hosted/private model execution.

9. Contractual commitments for enterprise clients

- Customer retains all ownership and intellectual property rights in uploaded documents and extracted outputs.
- BoreholeAI processes customer content only to provide the contracted extraction service and related support.
- BoreholeAI does not use customer content for model training, fine-tuning, benchmarking, or product analytics without separate written permission.
- BoreholeAI staff will not review customer content except where necessary for support, security, legal compliance, or incident response, and only under least-privilege controls.
- BoreholeAI will delete customer content according to the agreed retention tier and provide deletion evidence where available.
- BoreholeAI will notify the customer of material security incidents affecting customer content within a defined period.
- BoreholeAI will maintain a list of subprocessors and provide notice before material changes where commercially practical.
- For customer-hosted deployment, BoreholeAI access to the environment will be customer-controlled and time-limited.

10. Security questionnaire response pack

Common enterprise question	Recommended answer
Do you store our PDFs?	In Standard mode, yes, only during the retention window of 10 days so users can download/review outputs. In Enterprise SDK mode, customer documents and generated outputs are deleted after the SDK confirms successful local download. Minimal non-content metadata may be retained for billing, audit, support, and abuse prevention.
Can BoreholeAI staff see our documents?	Not during normal operations. Staff do not browse or use customer data. Any support access should require a customer support request, be limited to the relevant job, and be logged. For strict non-access requirements, dedicated or customer-hosted deployment can be discussed.
Do you train models on our documents?	No. BoreholeAI does not use customer documents, extracted text, or outputs to train, fine-tune, or benchmark models.

Do third-party AI providers receive our PDF?	No third-party AI provider should receive the full original PDF. When external AI is used, it should receive only minimal extracted fragments/crops required for classification. For high-security clients, BoreholeAI should offer self-hosted/private processing for those AI stages.
Can we delete data ourselves?	Yes. Standard users can delete jobs. Enterprise SDK users are configured for automatic purge-on-download. Account deletion removes account data subject to backup limitations.
What is retained after deletion?	Non-content metadata such as job ID, user/account ID, timestamps, status, page count, credit usage, deletion timestamp, and billing history may be retained.
Where is data hosted?	During the retention period and processing stage, data is hosted and stored in Supabase (SOC2 compliant).
Can you run in our environment?	Yes, for customers with strict data governance, BoreholeAI can discuss dedicated private-cloud or on-premise deployment so documents do not leave the client-controlled environment.