

SCANALYTIX.AI

Intelligence architecture for the legal and financial-accounting sectors

From the physical or digital archive to queryable evidence that holds up in court — through a modular ecosystem that digitises, processes, analyses and secures your documents, in an isolated environment, under your organisation's encryption keys.

Where the problem starts

Legal and financial professionals face the same obstacles, year after year:

- massive volumes of unstructured documents, in a variety of formats
- difficulty organising and retrieving information, with a real risk of overlooking essential evidence
- critical details missed in case files or in complex financial statements
- lengthy court proceedings, with the unavoidable re-reading and re-learning of the file before each hearing
- high costs and risks generated by manual processing

The archive remains a static repository, at exactly the moment it should be the source of the argument.

What we do

SCANALYTIX.AI turns the archive into an interactive intelligence asset.

We take in documents in whatever state they are — boxes, binders, digital files, media archives — and return them as a structured, numbered, indexed and queryable case file, with every statement traceable to the page it comes from.

The service covers the entire life cycle of the case file and is delivered in five sequential stages, each laying the groundwork for the next.

Our processing commitment:

We assess the archive, then commit to a firm deadline and volume.

The technical capacity of **SCANALYTIX.AI** is not the limiting factor; the physical condition of the material is — staples, damaged envelopes, folded or partially destroyed documents, inconsistent ordering, missing inventories, files generically named by the scanner, illegible documents, etc.

Processing takes place exclusively under a contract concluded in accordance with Article 28 of Regulation (EU) 2016/679 (GDPR), with the technical and organisational measures described in this proposal and detailed in the annexes to the contract.

Two delivery models. The choice is yours

Your documents can remain entirely on your premises, or they can be collected by us.

These are two distinct models, with different legal implications.

The model is agreed before contracting and recorded in the contract.

	Model A — Processing at the Client's premises	Model B — Full service, at SCANALYTIX.AI's premises
Where the documents are	They never leave your premises, except as specified under "Data transfer".	They are collected, transported and processed at SCANALYTIX.AI's premises.
Dedicated IDP equipment	Installed at your premises, on a rental or operating-lease basis.	The equipment and infrastructure belong to SCANALYTIX.AI .
Who handles the documents	Your staff, trained by us, or SCANALYTIX.AI staff seconded to your premises.	Vetted SCANALYTIX.AI staff, under NDA and a documented chain of custody.
Data transfer	No transfer of documents takes place. Processing is carried out within your infrastructure. Exception: if you opt for interactive dialogue with an external language model (Tier 3), the content used in the dialogue is transmitted to the provider named in the contract, on EU infrastructure. Exposure is reduced through pseudonymisation (Tier 2).	Transfer takes place. The procedure in Annex 1 applies in full.

	Model A — Processing at the Client's premises	Model B — Full service, at SCANALYTIX.AI's premises
Suitable for	Case files under strict professional secrecy, recurring volumes, organisations with strict internal document-security policies.	Organisations that do not wish to manage equipment and staff and prefer a turnkey service.
Legal framework	Data processing agreement (Annex 5), signed before the first document.	Data processing agreement (Annex 5) plus the custody procedure (Annex 1), signed before collection.

One rule, valid in both models: no document enters processing before the data processing agreement is signed.

This also applies to the archive assessment stage and to any demonstration on real documents.

Legal framework of the processing – The Data Processing Agreement

It is the entry condition for the working relationship and, in our view, the minimum condition under which a professional bound by professional secrecy can outsource the processing of a case file.

Roles of the parties

You are the data controller.

SCANALYTIX.AI is the processor and processes exclusively on the basis of your documented instructions.

We do not use the data for our own purposes, we do not use it to train AI models, and we do not make it available to any third party not authorised under the contract.

The Data Processing Agreement

It is signed before the first document is handed over, in the form set out in Annex 5 (separate document).

It covers the nature and purpose of the processing, the categories of data and data subjects, duration, security measures, sub-processors, assistance in the exercise of data subjects' rights, incident notification, audit and deletion of data on termination.

The assessment stage and pilot projects fall under the same agreement: signature precedes them, and the Client receives the handover reports and certificates for each stage.

Your relationship with your own client

The data processing agreement governs the relationship between you and **SCANALYTIX.AI**.

Your obligations towards your own client — professional secrecy and informing the client about the use of a processor and of artificial-intelligence tools — remain yours and cannot be transferred by contract.

To support you, we provide, on request, a model clause for the legal services agreement and a model client information notice, which you adapt at your own discretion. They describe what is processed, where, by whom, with what safeguards and with what options regarding external language models.

Who actually accesses the content

- Data is encrypted in transit (TLS 1.2/1.3) and at rest (AES-256). The encryption keys of the final archive are held by your organisation; without them, the stored content is unreadable to any third party, including an attacker who obtains the files.
- In Model B, for the duration of processing, the **SCANALYTIX.AI** operators assigned to the project have access to the content, strictly within the limits of their assigned task (least privilege, RBAC – Role-Based Access Control, multi-factor authentication).

This access is required for physical preparation, quality control and Human-in-the-Loop correction.

- Every access is logged by name in the audit trail, with a timestamp and the action performed.

The audit trail is available to you in full, on request, for the entire duration of the contract.

- In Model A, **SCANALYTIX.AI** staff have access only if you opt for seconded staff; otherwise, there is none.
- All **SCANALYTIX.AI** operators are vetted before hiring and sign confidentiality undertakings of unlimited duration.

Sub-processors

We use no undisclosed sub-processors. The full list, with each one's role and processing location, is annexed to the contract (Annex C of the data processing agreement) and includes, depending on the services contracted, the cloud-infrastructure and language-model providers used in the analysis tier and the equipment / equipment-servicing providers.

Using an external language model (AI) is optional and your decision

If you opt against it, Tiers 0–2 are delivered in full, without a single call to an external provider.

If you opt for it, the provider is named in the contract as a sub-processor, processing takes place on infrastructure in the European Union, and the data is not used for training.

The security conditions of processing at the provider are those of the contract with that provider, which we make available to you.

You choose the model you prefer from a list of models that cumulatively meet three conditions:

- (1) the provider is named in the contract as a sub-processor and processes under a data processing agreement;
- (2) processing takes place on infrastructure in the European Union;
- (3) the data is not used for training and is retained by the provider only within the limits set out in the contract.

Personal or free accounts of consumer AI applications do not meet these conditions and cannot be used within the service.

Models available at the date of this proposal: **Claude** (Anthropic) and **Gemini** (Google), accessed through the cloud infrastructure named in the contract, in a European Union region, under a data processing agreement and without the data being used for training. The list can be extended, at your request, with any other model that meets the same conditions.

We recommend the model best suited to each task (summarising, comparing, searching for contradictions, drafting) and support you in formulating queries, through the conversation manual established together with you.

Our recommendation: enable pseudonymisation (Tier 2) before any dialogue with an external model.

Names, personal identification numbers and all other direct identifiers are replaced with artificial identifiers before transmission and are restored only within your environment, using your organisation's keys.

Retention and deletion

Raw data is retained for the duration of the project and for a further 30 days after delivery, for remediation.

On expiry of that period or at your request, it is permanently deleted, and **SCANALYTIX.AI** issues a certificate of deletion.

Physical documents are returned against a handover report, or destroyed to DIN 66399, protection class 3, with a certificate of destruction.

Conflict of interest

Our team includes lawyers and judicial and financial experts who may be independently appointed in other cases.

The rule we commit to contractually:

- we keep a register of the cases and parties for which we have provided services;
- we check for conflicts before accepting any new project and notify you in writing of any overlap identified;
- we do not accept a case in which we have worked for the opposing party without the written consent of both parties;
- technical separation between case files is effective, not merely procedural: isolated instances, distinct keys, teams without cross-access.

Our information-security procedures are built on the ISO/IEC 27001 framework and on GDPR requirements.

Who we are

SCANALYTIX.AI is a legal-tech company built for professionals who work with complex data: lawyers, judicial experts, auditors and financial consultants, corporate legal departments, HR experts.

The platform processes the volume.

SCANALYTIX.AI's lawyers and judicial and financial experts calibrate what is being looked for — the Case-Specific Metadata Manual.

Query lines (prompts) are built for the specifics of your case.

The SCANALYTIX.AI team

Specialists from six fields in the same project team: legal, judicial, financial-accounting, information technology, cybersecurity and information security.

SCANALYTIX.AI technology

The SCANALYTIX.AI EVO ecosystem is built primarily on Canon, I.R.I.S., Alphabet, Anthropic, dtSearch and Tresorit technologies, in an isolated environment.

Providers that actually process data are named as sub-processors in the contract; the others are equipment providers or software run locally.

Tier 0 — Precision ingestion (MICAREXPUF)

Turning the raw archive, physical or digital, into an ordered and verified digital flow is the foundation of the entire process.

Physical documents

- Collection against inventory, in SCANALYTIX.AI-specific containers, from your premises.
- Staple removal, flattening, trimming of correspondence, sorting and ordering according to your requirements.

Digital documents

- Transfer exclusively through secure channels — SFTP, secure API or a dedicated client portal, encrypted with TLS 1.2/1.3. We do not use e-mail attachments or public transfer platforms.
- Full inventory on receipt.
- Media archives — video, audio, graphics — are collected, processed and inventoried separately.

The secure document-transfer procedures are set out in Annex 1.

Tier 1 — Preparation (SARAH)

This is not archiving, but digital curation. This is where the risk of human error is eliminated and the ground is prepared for precision analysis.

Double inventory	On receipt and after preparation. Full traceability for every document.
Image quality score	Numeric test per page — blur, glare, type size too small. Illegible documents are identified before processing.
Cleaning and upscaling	Removal of background noise and increase in resolution, so that every detail remains legible.
Scanning	450 - 600 dpi
Renaming and standardisation	Uniform naming to the SCANALYTIX.AI standard: date, issuer, subject, recipient.
Separate archives	The original, unprocessed version and the processed version. Both are kept, for compliance.

Tier 2 — Processing (ZAR)

Static documents become structured, secured entities ready for querying.

The stage at which the archive becomes a database.

Conversion	Transformation of documents from the formats received into the formats optimised for analysis.
BATES numbering	Unique, sequential identifier per page, the international standard in litigation.
OCR	Optical character recognition. Printed text from scans and PDFs becomes editable and searchable.
OCR validation	Identification of reading errors. Uses a dual engine.
Pseudonymisation	Direct identifiers are replaced with an artificial identifier; the process is reversible only with the organisation's encryption keys. Recommended before any dialogue with an external language model.
Anonymisation (removal of direct identifiers)	Direct identifiers are permanently removed together with the associated metadata. Irreversible. Note: removing direct identifiers does not automatically amount to anonymisation within the meaning of the GDPR if the person remains identifiable from context.
Human in the Loop – correction of misread data	Correction by an operator of data below the confidence threshold generates a new document, preserving the original and logging the changes.
Source weighting	Weighting of documents, required for the analysis in Tier 3 – “Merge, Split & File Compression”.
Case-Specific Metadata Manual	Key semantic words and phrases, established together with you, for the specific strategy of the case.

Deterministic search and identification (dtSearch)	Full-text indexing of the entire archive and reproducible search reports: every occurrence of a term, name, amount or phrase, with the document and BATES page where it appears. Exact, fuzzy, proximity or pattern search. The same query on the same index always produces the same result, independently of any AI model.
Deterministic extraction	Extraction of values (dates, amounts, parties, document numbers, issuers, etc.) based on predefined rules and validations, without a generative model. The result is reproducible and auditable; values below the confidence threshold go through Human-in-the-Loop correction.
Granular encryption	Differentiated access by department, internal team or external collaborator.
Audio transcription	Conversion to text of the audio and video recordings in the case file that you make available, provided you have the legal right to hold and process them. Verifying that right is your responsibility.

Tier 3 — Analysis (ZARIA)

From processing to strategic dialogue with the case file, in line with the EDRM (Electronic Discovery Reference Model).

Summaries	Document summary of variable length – essence, medium, detailed – focused on the topics you request.
AI-assisted extraction	From structured, semi-structured (forms, financial statements) and unstructured files, for the overall analysis of the case file. It complements the deterministic extraction in Tier 2; every value indicates the BATES page it comes from.
Identification	Establishing the relevant data sources: e-mail servers, workstations, cloud, mobile devices — at your request and within the limits of the legal basis available to you.
Legal hold	Ensuring that data is not deleted or altered; any notices to the persons involved are issued by you, on the applicable legal basis.
Collection	Transfer in a secure environment, with metadata intact — creation date, author, change history — from sources to which you have a legal right of access.
Comparative analysis	Identification of contradictions between statements and the other documents in the case file.
Review	Assessment of documents for legal relevance and for compliance with attorney-client privilege.
Interactive dialogue with the case file	Latest-generation models, accessed from your isolated environment; processing takes place at the provider named in the contract, on EU infrastructure, and only if you opt for this component.

Multimodal conversation	Specific text reports, podcast-style audio synthesis, interactive Mind Map.
Conversation manual	Established together with you — questions are calibrated to the case strategy and to eliminate ambiguous wording.
Resumption brief	Before each new court hearing: a summary of the latest hearings and the vulnerable points identified.
Correlation with legislation and case law	Through the sources your organisation already subscribes to, where the provider allows technical integration. We do not resell licensed legal content.

Anchored exclusively in the processed sources (the archive)

The system is configured to answer exclusively from the processed and uploaded documents, with a reference to the BATES page from which each statement originates.

If the information is not found in the sources, the system is instructed to say so explicitly, rather than fill in the gap.

The BATES references allow you to verify any answer quickly; verification remains necessary.

Sources admitted simultaneously in the analysis:

- the entire case file, with no contractual limit on the number of files or queries;
- external sources specifically defined by you.

Professional judgement and responsibility remain yours.

The EDRM procedures are set out in Annex 3.

Tier 4 — Storage (REBEFRE)

Data life cycle

Archive storage	Low cost between hearings, with government-grade security.
Encryption at rest	The content cannot be accessed without your organisation's keys.
Instant reactivation	At a new hearing, the case file becomes immediately queryable again.
Audit trail	Complete log: who queried what, and when. Essential for compliance and for the profession's rules of ethics.

What this service is not:

SCANALYTIX.AI is not an accredited electronic-archive administrator.

The storage we offer is secure storage, not electronic archiving with legal value within the meaning of the legislation on archiving documents in electronic form.

If you need archiving with legal value, we will refer you to an accredited provider.

The Audit Trail procedures are set out in Annex 2.

Liability, insurance and continuity

Limit of liability	Set by contract, in proportion to the value of the project.
Professional liability insurance	Set in proportion to the value of the project.
Business continuity	Should SCANALYTIX.AI cease operations, your data is returned to you in full, in the original and processed formats, within 30 days, and our copies are deleted with a certificate. As the encryption keys are yours, the archive remains accessible to you regardless of our situation.
Dependence on providers	Tiers 0–2 do not depend on any external language-model provider. If such a provider becomes unavailable, Tiers 3 and 4 are affected, while the processed archive remains fully usable.

How we work together

Signing the processing framework

The first document signed is the Data Processing Agreement in Annex 5.

We do not collect or assess any real document before the Data Processing Agreement is signed.

Archive assessment

We assess the archive on site: the physical condition of the material, the degree of standardisation, the actual volume, the types of documents and media.

The assessment is a service in its own right and is invoiced separately.

The result is a report describing what you have, what can be processed, in how long and with what risks.

It is the document on the basis of which we commit to a firm deadline and volume.

The five tiers

The tiers (0–4) can be requested separately or together.

What you get

Legal integrity	Every piece of evidence extracted is compliant, traceable and deliverable to the court, with BATES numbering and in line with the EDRM model.
Efficiency	Analysis time is reduced from weeks to hours, through deep search in any type of document and direct dialogue with the case file.
Security	Isolated environment, equipment on your premises (Model A), encryption with the organisation's keys, complete audit trail.
Expertise in the loop	Lawyers, judicial experts, financial auditors and information-security specialists in the same project team.
Complete legal framework	Data processing agreement, documented chain of custody, certificates of deletion and destruction. Outsourced processing takes place within a documented framework that allows you to demonstrate due diligence towards your client and towards the supervisory authority.

The reference offer is set out in Annex 4.

Annex 1 — SCANALYTIX.AI: Secure transfer of physical and digital documents

This annex applies exclusively to delivery Model B, in which the documents are collected by SCANALYTIX.AI.

In Model A, the documents do not leave your premises and this annex does not apply.

A secure document transfer requires establishing an unbroken chain of custody, ensuring that the data – whatever its format – is not compromised, lost or accessed without authorisation at any point in the process (collection, processing, storage and destruction/return).

1. Collection and processing of the physical archive

- **Inventory and sealing:** Before leaving the client's premises, the documents are inventoried and placed in special archive boxes, which are closed with uniquely numbered tamper-evident seals – any attempt to open them leaves visible traces.
- **Secure transport:** Transport is carried out in dedicated vehicles fitted with GPS tracking and alarm systems. Drivers and handling staff are vetted and have signed non-disclosure agreements (NDAs).
- **Chain of custody:** At every step (handover by the client → SCANALYTIX.AI carrier → receipt at SCANALYTIX.AI's premises), handover reports are signed and the integrity of the seals and unique serial numbers is checked.
- **Scanning / processing area:** At SCANALYTIX.AI's premises, the boxes are stored and unsealed only in restricted-access areas (access cards or biometrics), under 24/7 video surveillance (CCTV).
- **Wiping of equipment memory:** Physical scanning equipment temporarily retains images in its own memory. The scanners are configured to automatically clear their cache after every session.

2. Collection and management of the digital archive

For the transfer and processing of digital data, SCANALYTIX.AI uses protocols that protect information both “in transit” (while it is being sent) and “at rest” (once it reaches the servers).

- **Encrypted transfer channels:** No e-mail attachments or public platforms (such as Google Drive/WeTransfer) are used. Only secure protocols are used, such as SFTP (Secure File Transfer Protocol), secure APIs or a dedicated client portal encrypted with TLS 1.2/1.3.
- **In-transit and at-rest encryption:** Data is encrypted during transfer and also when stored on SCANALYTIX.AI's servers (e.g. using the AES-256 standard), so that even if a server is compromised, the files cannot be read without the decryption key.
- **Access control (RBAC and MFA):** SCANALYTIX.AI employees access the digital archive on the “least privilege” principle (Role-Based Access Control) – each operator sees only the data required for their task. Authentication is mandatorily via MFA (Multi-Factor Authentication).
- **Traceability (audit trails):** Every action (who uploaded a file, who viewed it, who ran the OCR/AI extraction model on it) is automatically recorded in immutable system logs.

3. Post-processing: closing the cycle

After the SCANALYTIX.AI platform extracts the data and delivers it back to the Client, again through an encrypted channel, the following takes place:

- **Return or secure (physical) destruction:** Physical documents are returned against a handover report or, at the Client's request, shredded to DIN 66399, protection class 3, with a certificate of destruction issued.
- **Digital deletion (data wiping):** Raw digital data is permanently deleted from SCANALYTIX.AI's servers, in accordance with the agreed retention policies, with a certificate of deletion issued.

Annex 2 — SCANALYTIX.AI: Audit Trail procedures

An Audit Trail is a chronological, secure and unalterable record of all the actions, decisions and changes a document goes through, from the moment it enters the system until the data is exported.

Since, depending on your options, **SCANALYTIX.AI** systems may also use external language models (AI) to “read” and interpret documents (invoices, contracts, medical forms), traceability is essential.

The **SCANALYTIX.AI** audit trail explains how the final result was reached and separates decisions made by the algorithm from those made by human users.

What exactly does an audit trail record in SCANALYTIX.AI?

For every document processed, **SCANALYTIX.AI** keeps a detailed history.

The most important elements captured include:

- **Ingestion:** The exact moment (timestamp) the document entered the system, its source (e.g. API, e-mail address, physical scanner) and who uploaded it.
- **Processing – AI/OCR extraction:** Which data fields the algorithm extracted (e.g. name, amount, date) and the artificial intelligence’s confidence score for each field.
- **Human intervention (Human-in-the-Loop):** If the confidence score was too low, the system sends the document to a human operator for verification. The audit records precisely what the operator changed (the old value extracted by the AI vs. the new value entered manually), the user’s identity and the time of the change.
- **Document-level actions:** Deletions, page additions, rotations, splits or merges of files.
- **Queries:** Who queried the case file, when, and what they asked.
- **Export:** The moment the final structured data was sent onward (e.g. to an ERP or CRM system or a database) and who/what triggered the action.
- **Storage, retention and destruction:** Entry into the archive, the retention period applied, the moment of deletion and the issuing of the certificate.

Annex 3 — SCANALYTIX.AI: EDRM-compliant procedures

The Electronic Discovery Reference Model (EDRM) is a conceptual framework that standardises the steps for managing digital evidence, known as electronically stored information (ESI), during legal proceedings and investigations.

The nine **SCANALYTIX.AI** reference stages:

- **Information Governance:** Managing the creation, storage and disposal of data before litigation begins.
- **Identification:** Locating potential sources of relevant data and custodians.
- **Preservation:** Protecting data against deletion or alteration.
- **Collection:** Gathering data in a way that maintains the chain of custody.
- **Processing:** Converting files into a reviewable format.
- **Review:** Assessing documents for relevance, confidentiality and privilege.
- **Analysis:** Evaluating content for key patterns, context and deeper insight.
- **Production:** Preparing and delivering data in the form required by the court, the judicial body or the applicable procedure.
- **Presentation:** Presenting evidence at hearings, depositions or trials.

Annex 4 — SCANALYTIX.AI: Reference offer

- **Case-file assessment** — from EUR 200 + VAT
- **Tier 0** — Precision ingestion (MICAREXPUP): from EUR 0.10/page + VAT
- **Tier 1** — Preparation (SARAH): from EUR 0.20/page + VAT
- **Tier 2** — Processing (ZAR): from EUR 0.40/page + VAT
- **Tier 3** — Analysis (ZARIA): depending on the project assessment – from EUR 500 per month of analysis + VAT; during the analysis period only
- **Tier 4** — Storage (REBEFRE) in maximum-security environments – depending on the project assessment – from EUR 100/month + VAT, per case file
- **Other processing and analysis services** — variable cost depending on complexity.
- **IDP-specific hardware** — offered under operating lease or rental — assessed according to the estimated annual volume of data to be processed.
- **SCANALYTIX.AI consumables** – containers, storage media, etc. — cost depending on the project assessment.

Note: Every project is assessed and negotiated according to its complexity and your specific requirements.

Annex 5 — SCANALYTIX.AI: The Data Processing Agreement

The personal data processing agreement, concluded under Article 28 of Regulation (EU) 2016/679, is provided as a separate document.

It is signed before the first document is handed over, including before the archive assessment and before any demonstration carried out on real documents.

Its own annexes contain: the description of the processing, the technical and organisational measures, the list of authorised sub-processors, and the templates for the handover report, certificate of deletion and certificate of destruction.