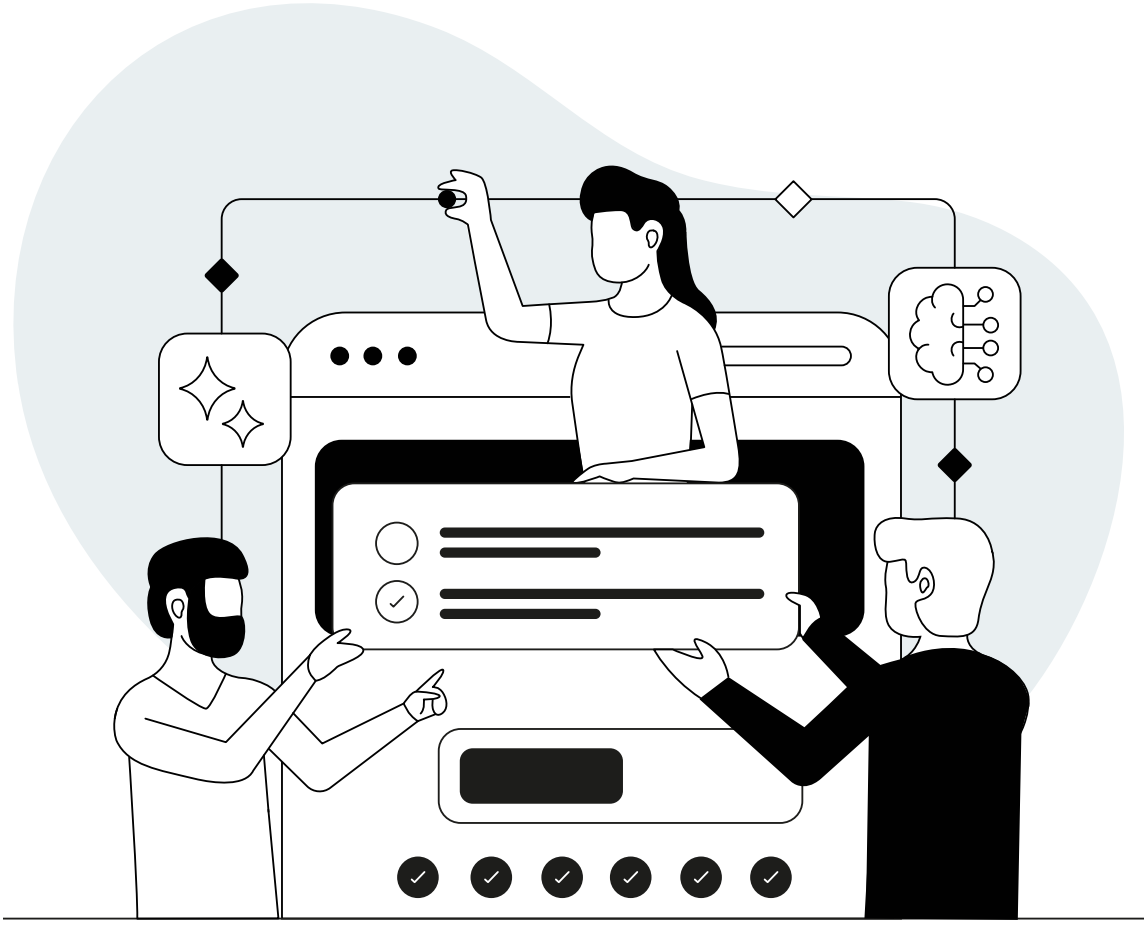
AVALLAIN

Avallain AI Trust Model



A Personal Data Protection Framework for Generative AI Workflows

At Avallain, we design and continuously evolve educational technologies for leading publishers and educators worldwide. In line with this purpose, our products incorporate generative AI and large language models (LLMs). While we recognise the potential of AI and explore the new opportunities it brings to the educational community, we are also aware of the privacy risks these tools pose to the protection of personal data.

For this reason, we are committed to creating secure environments for our users' data by applying the technical measures necessary to reduce data vulnerability. These safeguards are structured within the **Avallain AI Trust Model**: a multi-layer framework designed to enforce strict data privacy strategies at every stage of the data workflow.

This framework sits within a broader commitment to responsible AI use in education. For a full account of Avallain's position on AI ethics, human oversight, risk classification and data governance, see our companion publication, '**Avallain Position Paper on AI Ethics and Security in Education**'.

Layer 1: User Application

This layer focuses on the tool's interface where data entry begins, as well as on the practices and enablement of the users interacting with the system.

- **Secure Interfaces (Data Containment):** AI interactions are isolated entirely within our applications. Unlike public web-based AI tools, our interfaces act as closed sandboxes that prevent external data scraping, session crossover or unauthorised third-party tracking, ensuring that inputs are securely contained from the moment of entry.

- **Data Minimisation and Active Prompt Hygiene:** In our products, the interface uses active, contextual UI cues and automated reminders that nudge users to remove unnecessary personal details before submitting a prompt. This enforces data minimisation from the very start.

- **AI Literacy Training:** We provide practical training courses to help users understand how AI tools work and where the privacy risks lie. This training gives users the technical knowledge needed to understand how best to manage personal data.

Layer 2: The Sovereignty Layer

This layer defines the core of our data governance framework. Data sovereignty means that Avallain ensures that user data always remains within the appropriate regulatory jurisdiction and geographic boundaries. We achieve this by establishing a strict compliance boundary, the **Data Sovereignty Firewall (DSF)**, which confines data entirely within the appropriate region, usually either the EU/UK or the USA.

However, our architecture is open to other regions and higher geographic specificity if required, for example retaining all data within a specific country. This means your data is restricted to authorised server nodes, preventing unauthorised cross-border transfers and ensuring strict compliance with relevant legislation such as the EU GDPR and UK GDPR.

Within this secure architectural boundary, we deploy two specific subsets of controls to protect data and maintain oversight:

1. Data Protection Strategies

Once personal data enters the Sovereignty Layer, it undergoes automated stripping and cleaning before it ever reaches the core AI engine:

- **PII Masking (Pseudonymisation):** The infrastructure supports redacting Personally Identifiable Information (such as addresses, ID numbers, email addresses) from the text payload while the query is processed.

- **Identity Decoupling (Anonymisation):** The system strips away underlying network and user metadata, ensuring the AI model has no way to trace a query back to a specific individual or session.

2. Compliance Oversight

To provide transparency and support security auditing, a robust record of system activity is put in place:

- **Audit Trail & Logging:** The system maintains continuous event logs, providing centralised traceability for security checks and compliance verification.
- **Data Minimisation in Logs:** To protect user privacy, Avallain logs only prompt payloads insofar as necessary for provided functionality and features, with controllable retention periods to ensure data is kept only for the required period before being removed.

Layer 3: Secure Transport

This layer secures the data while it is in transit and strictly controls its physical destination. It ensures that data is encrypted dynamically and routed exclusively to verified, compliant endpoints.

Within this infrastructure layer, we deploy two core network controls:

1. End-to-End Encryption:

To secure data against interception while moving between systems, all communications are fully protected:

- **Encryption in Transit:** All data passing between the user interface, our applications and the processing nodes is encrypted over secure HTTPS connections using industry-standard TLS protocols. This ensures that the data payload cannot be intercepted or read by unauthorised third parties while moving across the network.

2. Regional Endpoint Pinning:

This is the technical mechanism that enforces our Data Sovereignty Firewall policies at the network level:

- **Hard-Coded Routing Restrictions:** Our transport infrastructure is explicitly configured to communicate solely with enterprise API endpoints that are contractually and physically anchored within the desired cloud regions.

Layer 4: AI Inference Engine

This layer governs the actual processing of data by the AI, ensuring that inputs are handled transiently and are never permanently stored or reused. By establishing a stateless computation environment, this layer guarantees that your data is treated purely as a temporary operational calculation, remaining completely insulated from the underlying model's long-term memory.

Within this processing layer, we enforce two absolute data-handling standards:

1. Stateless Processing

When a prompt is processed, it exists strictly within volatile runtime memory (RAM):

- **Temporary Calculation Only:** Data is handled entirely as a transient calculation and is completely purged immediately after the AI generates its response.
- **No Disk Writes:** No user data or prompt history is ever written to a permanent storage disk or database during the inference phase.

2. Zero-Retention and Training Ban

This is our legal and technical guarantee that your intellectual property and user inputs remain exclusively yours:

- **Contractual Zero-Retention:** Backed by strict, enterprise-tier contractual agreements with our infrastructure providers, no user data or prompt inputs are ever retained by the vendor post-session.
- **Absolute Training Ban:** Our agreements strictly prohibit the use of our data to train, retrain or improve any underlying public or proprietary AI models.

Our Commitment: Innovation Without Compromise

We believe that the future of education should be shaped by intelligent, innovative technologies, but never at the expense of the privacy, autonomy and rights of students and educators. Whether you are a global publisher or a school community, the **Avallain AI Trust Model** ensures our products are designed to respect and protect your data.

Our strict, multi-layered safeguards enable administrators, teachers and students to use advanced generative tools with confidence. Teach, learn and create knowing that your data remains protected, managed in line with relevant legislation and not used to train external AI models. With Avallain, you can bring AI into the classroom safely.

AVALLAIN

Avallain AI Trust Model

A Multi-Layer Personal Data Protection Framework for Generative AI Workflows

LAYER 1
User Application

Interface Security & User Empowerment

- Secure Interfaces
- Active Prompt Hygiene
- AI Literacy Training

LAYER 2
The Sovereignty Layer

Regional Governance & Privacy Controls

- Data Sovereignty Firewall
- PII Masking & Identity Decoupling
- Audit Trail & Data Minimisation

LAYER 3
Secure Transport

Encrypted & Routed Data Transit

- End-to-End Encryption
- Regional Endpoint Pinning

LAYER 4
AI Inference Engine

Stateless Processing, Zero Retention and Training Ban

- Stateless RAM Processing
- Contractual Zero-Retention
- Absolute Training Ban

Teach, learn and create with confidence. Avallain ensures your data remains protected, securely managed and insulated from external AI training.