

# Silent AI

**Making productive use of corporate knowledge – locally, securely, without the cloud.**

Silent AI makes your company's knowledge accessible via an AI assistant – entirely on-premises, whilst fully respecting existing access rights. No prompts or tokens ever leave your network.

**100%**

local (on premises)

**100%**

secure (off-cloud)

**100%**

private (identity management)

Silent AI is a compact 2U appliance that features a high-performance GPU and fast NVMe storage as local storage.



## OVERVIEW



### Zero Cloud by design

All data and LLM inference are processed locally on the appliance. No tokens leave the network.



### Identity management

AD/LDAP integration: Employees only receive AI responses regarding content they have read-only access to.



### Seamless integration

Connectors for M365, SharePoint, Confluence, Slack, PDF, file server, SQL, ... – deployment in few days.



### Plannable low TCO

A one-off investment. No token pricing, no subscription surprises, CARE maintenance included.



### Compliance-ready

GDPR-compliant, NIS2-ready, EU AI Act-compliant. No data transfer to US infrastructure.



### Made in Europe

Development, manufacturing and support in Germany. No risk of falling foul of the US CLOUD Act.

## CHALLENGES



### Data protection risks

Cloud AI sends prompts and documents to US servers. GDPR compliance and the US CLOUD Act are incompatible.



### Authorisation gaps

Cloud-based AI requires additional rights management. Employees may view content without the appropriate access rights.



### AI without a strategy

Competitors are more productive. NIS2 and the EU AI Act are now forcing decisions on AI governance.

**Silent AI makes corporate knowledge usable and productive – today, locally and entirely under your control.**

## SOLUTION AND FEATURES

A fully local AI platform for your corporate knowledge



### Local AI assistant

LLM inferencing on a dedicated NVIDIA GPU. No cloud API calls – fully air-gapped implementation possible.



### Identity management

AD/LDAP integration: every response adheres to your organisation's existing access rights structure.



### RAG on internal data

Retrieval-augmented generation on M365, SharePoint, Confluence, file servers, SQL and others – without any glitches.



### Integrated connectors

Ready-made connectors for all common document sources. Painless deployment in days, not months.



### Compliance by design

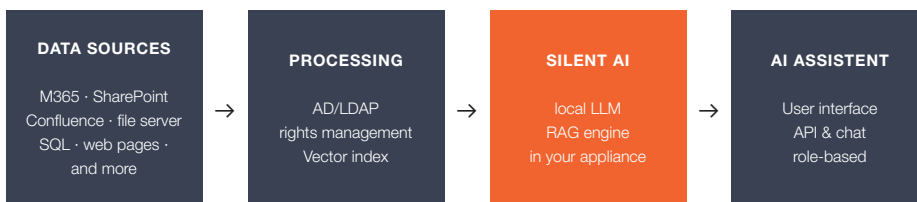
GDPR-compliant. No US CLOUD Act risk. NIS2-ready, EU AI Act-compliant. Audit-proof.



### Future-proof

Interchangeable LLMs that are becoming increasingly powerful. Up to 48 TB of Silent AI Vault storage for local vector data.

## ARCHITECTURE AT A GLANCE



## YOUR BENEFITS

- ✓ No data leaks – guaranteed by design
- ✓ More productive employees
- ✓ Full stack - reduced IT workload
- ✓ Compliance without compromise
- ✓ A future-proof, scalable platform

## SILENT ASSISTANTS

- Custom combinations of data sets, instructions and prompts, e.g.
- Service assistant with ticket system access
  - Financial assistant with a background in law
  - Medical assistant with specialist knowledge

## WHY SILENT AI

### 01 ENTERPRISE AI

**Use of AI with distributed, sensitive data. No cloud-based chatbot.**

### 02 ZERO TRUST

**Permissions inherited from the source systems. No exceptions.**

### 03 ON PREMISES

**All-in-one appliance. Made in Germany.**

### 04 FLAT RATE

**No token limits. Full cost control.**

### 05 PROOF

**Every answer with sources. Well-founded rather than made up.**

Further information and a demo session:

[www.fast-lta.de/en/silent-ai](http://www.fast-lta.de/en/silent-ai)

FAST LTA GmbH | May 2026 | Subject to change



## TECHNICAL SPECIFICATIONS

PLATFORM	
Provision	As a turnkey appliance, on-premises
GPU	Latest Nvidia GPU, 96 GB, fanless
STORAGE FOR LOCAL DATA	
Storage	Silent AI Vault storage module with 12x NVMe, triple parity
Storage slots	2 slots for Silent AI Vault
Storage capacity	12 or 24 TB (gross) per Silent AI Vault
NETWORK & INTERFACES	
Data	Dual 10GbE (RJ45) or 10GbE SFP+
Optional	Dual 25/100GbE (QSFP)
Management	1x 1GbE Admin 1x 1GbE IPMI
TECHNICAL SPECIFICATIONS	
Power	Idle: ~170 W Typical: ~636 W (1 active Silent AI Vault) Max: ~1.200 W
heat dissipation	Typical: 2170 BTU/h Max: 4094 BTU/h
Dimensions	2U x 19" x 640 mm Weight ~17 kg
AI & INTEGRATION	
Architecture	RAG (Retrieval-Augmented Generation)
Identity management	AD/LDAP integration (inherited)
User interface	Web UI · Chat API from mid 2026
Authentication	LDAP, SSO
Connectors	MS Office, MS SharePoint, MS Outlook/Exchange, Confluence, Slack, SQL, SMB/file server, PDF upload, web scraper  More in development / upon request API from mid 2026
COMPLIANCE	
Data compliance	GDPR-compliant, no transfers to third countries
Regulatory framework	NIS2-compliant, EU AI Act-compliant
Origin	German product · not subject to US CLOUD Act
SLA	CARE 1/3/5 years, fixed conditions
Availability	Now