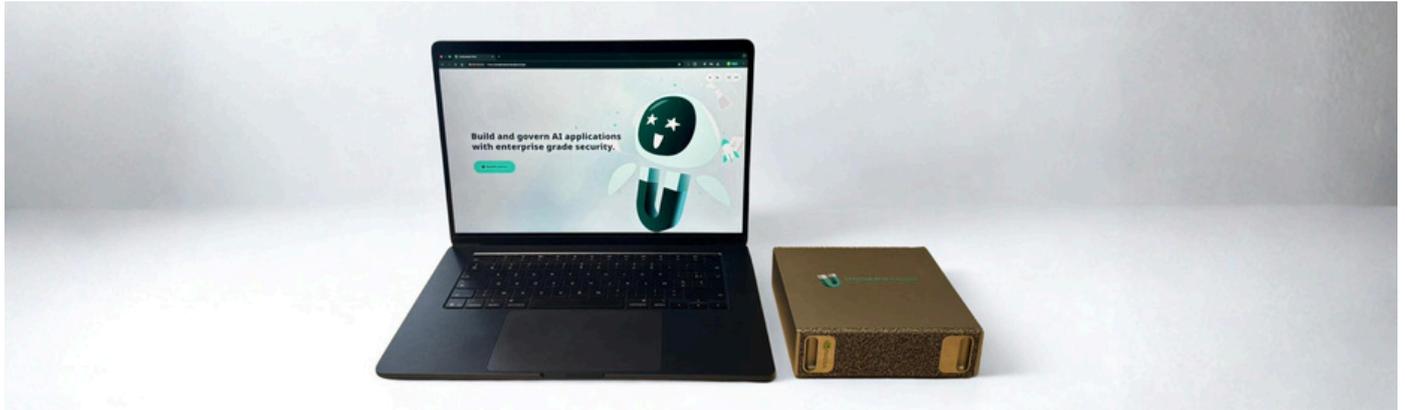# AI-In-a-Box

Fully managed, software-defined AI operating model for on-premise environments.

A compact, high-performance on-premise AI solution running the Understand Tech platform locally for secure LLM inference, RAG, and AI workflows, with no dependency on external services, networks, or vendors at runtime. Designed for regulated, air-gapped, and critical environments, it keeps data, models, and operations private, configurable, and compliant.

Key Features

- **Fully On-Prem & Air-Gapped:** Runs offline with no external dependencies.
- **Local LLMs & RAG:** On-site inference with secure local knowledge bases.
- **Multi-User & Access Control:** Concurrent users with role-based permissions.
- **APIs & Integrations:** Programmatic access for automation and custom apps.
- **Encrypted Local Storage:** All data stored and secured on-site.
- **Offline Updates:** Secure platform and model updates delivered offline.

Key Components

Integrated Reference Hardware Platform, software, and services work together seamlessly to provide built-in, always-on security.
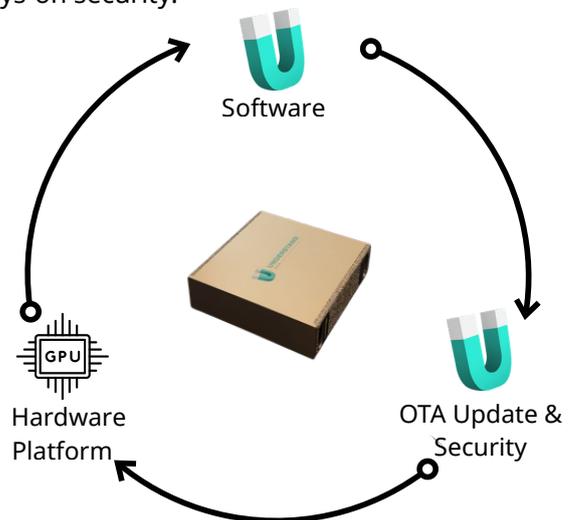
Reference Hardware Platform:
- NVIDIA Blackwell-based platform (GPU), 128 GB unified memory, 20-core Arm CPU
- 4 TB encrypted storage, local Wi-Fi/LAN connectivity

Software:
- Ready-to-run Understand Tech Platform: assistants, LLM workflows, RAG, batch jobs, and more
- Supports open-source LLMs: GPT-OSS (20B–120B), Qwen 3 (8B–32B+), and other OSS models.
- Seamless integration with your SSO, logging, and monitoring systems
- Docker-based, fully containerized with local-only access
- Encrypted, role-based, traceable access to all documents and interactions

Deployment & Security:
- Fully air-gapped or network-restricted operation
- No cloud dependencies, all computation and storage local
- Quarterly secure Docker images updates, delivered offline
- Operational continuity safeguards, enabling controlled updates, rollback, and validation prior to production use

AI-In-a-Box

is not a single-purpose device...

but a modular AI platform designed to run inside the customer's infrastructure and governance perimeter.
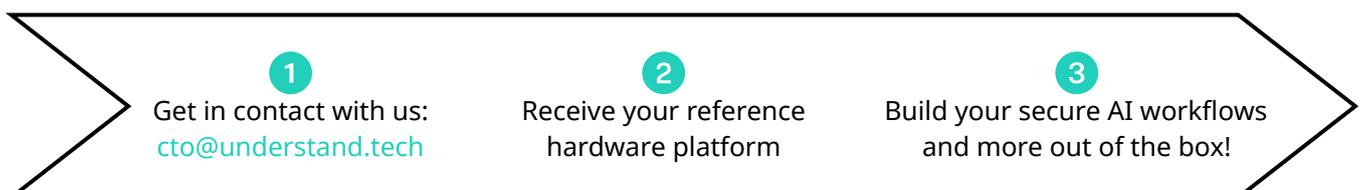

Key Use Cases

	Defense & Government	Air-gapped, classified data environments
	Critical Infrastructure	Utilities, energy, and industrial control networks
	Healthcare & Life Sciences	Hospitals & labs with PHI compliance
	Patents & R&D	Confidential invention drafting; proprietary research assistance
	Finance & HR Networks	Private payroll, compliance, and confidential audit workflows
	Secure Internal Chat/Assistants	Encrypted knowledge agents for high-trust teams

Support & Documentation

✓ **Documentation:** Admin manuals, API docs, and integration guides.
https://docs.understand.tech/documentation/ai-in-a-box

✓ **Onboarding & Best Practices:** Guidance for RAG, assistants, and governance.

✓ **Enterprise Support:** Deployment assistance, updates, and troubleshooting.

How to Get Started ?

1 Get in contact with us:
cto@understand.tech

2 Receive your reference hardware platform

3 Build your secure AI workflows and more out of the box!

Learn more:

https://understand.tech/product/ai-in-a-box