

Note: This English translation is provided for reference purposes only. In the event of any discrepancies between the original Japanese version and this English translation, the original Japanese version shall prevail.

September 24, 2026

Cuon to Launch On-Premises AI Solution "Private AI Appliance" in November 2026

— Combining Data Sovereignty and Secure AI Operations via a Two-Tier Architecture of Server Appliance and Cloud Management —

TerraSky Co., Ltd. ("TerraSky") announces that its group company, Cuon.Inc. ("Cuon"), an "AI & Data Integrator," will launch "Private AI Appliance," an on-premises AI solution, on November 1, 2026.

In recent years, while the deployment of AI agents for operational efficiency has accelerated, conventional cloud-based LLMs present severe security risks. These include confidential data leaks outside the corporate network, credential theft for SaaS applications caused by prompt injection attacks, and unauthorized access to host operating systems.

Private AI Appliance delivers safe, secure utilization and operation of local LLMs through a two-tier architecture: a server appliance deployed directly within the client's environment (hereafter "the Box") and a management cloud maintained by Cuon.

■ Service Architecture

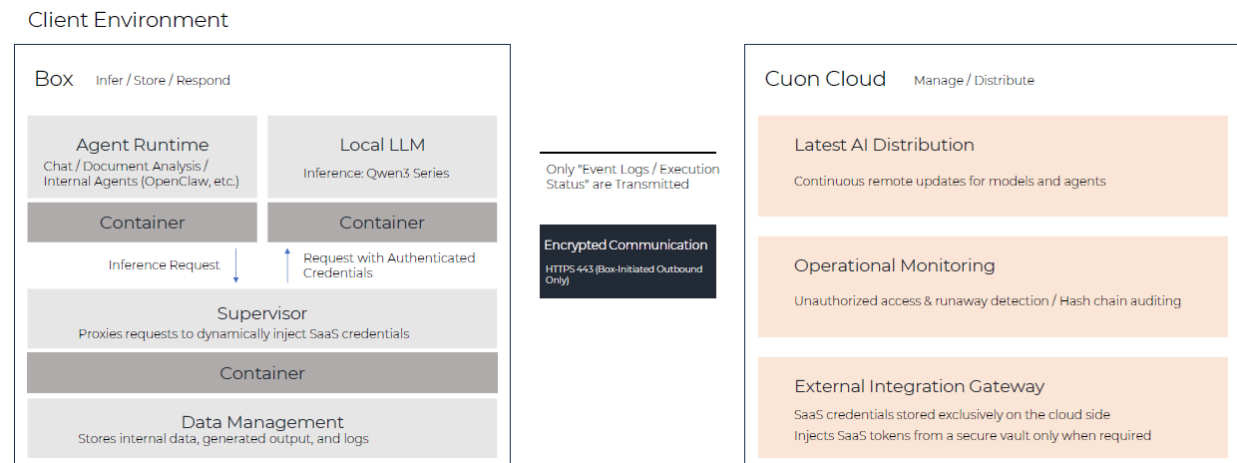
Private AI Appliance is delivered through the following two-tier structure:

The Box

Consists of pre-verified LLMs, agent runtimes, a workspace UI, and management software. By running each functional component inside isolated containers, it prevents malicious prompts embedded in files from stealing model data or compromising inference server privileges.

The Cuon Cloud

Provides centralized management functions, allowing users to monitor and manage local LLM operations and subscriptions on the Box.



(Cuon Private AI Architecture Diagram)

**All communication between both tiers is strictly limited to outbound connections initiated from the Box. There is no need to open inbound ports on the client network, nor is a VPN or dedicated line required.*

Three Strategic Approaches for a Secure Private AI Appliance

1. Building a Safe, Sandbox-Based Local LLM Environment

Isolates and operates all core functions required for local LLM and agent runtime execution within Docker containers. Even if a malicious prompt is executed, the damage is confined within the container, preventing lateral movement to the internal corporate network or host OS.

2. Credential Separation Architecture to Protect "Data Sovereignty"

When integrating with external SaaS applications, sensitive credentials such as OAuth tokens are securely managed and relayed exclusively through the cloud-side Gateway—never stored inside the Box. Keeping sensitive data and inference processing within the client's own on-premises environment ensures data sovereignty while minimizing credential leak risks.

3. Streamlined Operations via Cloud Management

Centralizes system monitoring, contract status, and local LLM health checks on the Box through the Cuon Cloud. This delivers both the robust security of an on-premises deployment and the operational convenience of the cloud.

Target Use Cases

- Secure enterprise utilization of Generative AI in data-sensitive industries such as manufacturing.
- Centralized governance, department-level SaaS authentication, and access control.
- AI agent operations within strictly restricted closed networks with limited internet access.
- Leveraging and integrating LLMs directly into system development processes.

Comprehensive Support System by the TerraSky Group

To address complex system architectures and specialized data management requirements unique to each enterprise, the TerraSky Group will combine its technical capabilities to provide full lifecycle support during solution rollout.

Support Capabilities across TerraSky Group:

- **TerraSky** : Salesforce expertise and DX consulting
- **LiberSky** : CDP/DMP and cybersecurity services
- **BeeX** : AWS and cloud infrastructure solutions

Service Availability Date

November 1, 2026 (Sun)

About Cuon.Inc.

Since its founding in 2010, Cuon has been a collective of skilled engineers specializing in the development of client-facing enterprise applications using Ruby, a programming language originating in Japan. After joining the TerraSky Group in 2019, Cuon expanded its business to develop and deliver enterprise systems that combine Ruby with cloud services such as Salesforce. As an "AI & Data Integrator" providing AI implementation support and consulting, Cuon remains committed to delivering deep expertise across system design, implementation, and operations to its clients.

Inquiries Regarding This Announcement:

Cuon.Inc.

Address: 15F Taiyo Life Nihonbashi Bldg., 2-11-2 Nihonbashi, Chuo-ku, Tokyo, Japan

TEL: +81-3-5577-4405

URL: <https://cuon.co.jp/service/paia>