

Supporting the Autonomous Operation of Generative AI Optimized for Your Business in a Dedicated Environment

Fujitsu leverages various business-specific AI technologies, including generative AI, action recognition AI, and demand forecasting, to support the creation of new services and businesses that improve daily convenience and enhance people's abilities and knowledge.

As a foundation for using these AIs safely and continuously, we provide, in an integrated manner, a suite of functions that realize the development, operation, and advancement of AI in the customer's dedicated environment, together with accompanying support services that underpin them.

Overview

This service provides a dedicated AI platform that can also be utilized on-premises as a means to realize highly sovereign AI, and supports the autonomous operation of generative AI optimized for the customer's business.

Service Features

In the customer's dedicated environment, we provide, in an integrated manner, a suite of functions for continuously improving models and agents, together with accompanying support services that underpin them, so that the customer themselves can autonomously drive the entire cycle of optimal generative AI model development, operation, and additional training.

Realization of "AI that continues to be used"
Break away from being stuck in the PoC stage, toward practical application

End-to-end support from planning and design to operation
Support for seamless operation of the AI infrastructure

Customized configuration dedicated to the customer
Meet customer needs by utilizing research technologies at the forefront

Robust AI infrastructure linked with infrastructure
Provision of a secure and safe foundation

Provided Functions

1. Infrastructure foundation that realizes a dedicated environment • Supports on-premises and realizes AI utilization that does not send data externally • Provides consistent support from introduction to operation and advancement • Adopts a high-performance GPU infrastructure based on PRIMERGY*	2. Generative AI trust technologies • An LLM vulnerability scanner that addresses over 9,000 types of vulnerabilities • Suppression of unauthorized input and leakage through guardrails • Automated operation that ensures security even for non-experts	3. High-accuracy generative AI models and customization / lightweighting technologies • Adopts high-accuracy Japanese-language LLM "Takane" as the core • Creates and improves business-specific models through in-house fine-tuning • Realizes low-cost and high-efficiency AI operation through LLM quantization technologies	4. AI agent development and efficiency technologies • Rapid construction of business AI with low-code / no-code • Integration with existing systems and cloud through MCP support • Advanced utilization through coordinated operation of multiple AI agents
---	---	---	--

*Japan: Private AI Platform on PRIMERGY, Europe: Private GPT

For details of this service and inquiries, please check ["Fujitsu Kozuchi Enterprise AI Factory."](#)