



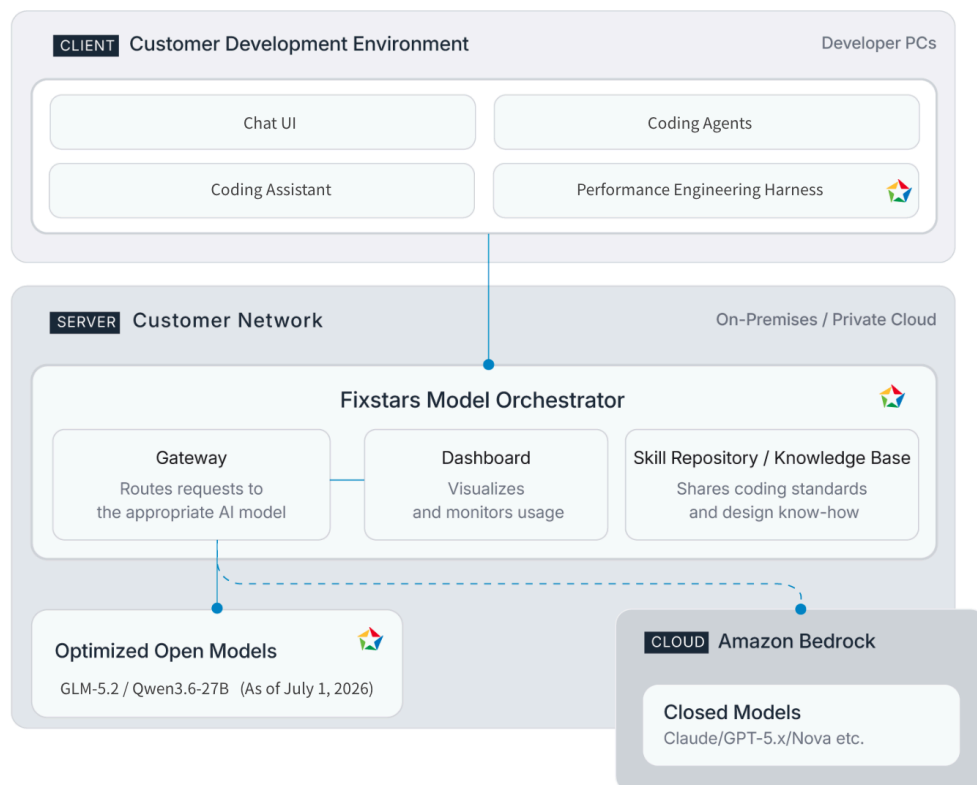
FOR IMMEDIATE RELEASE

Fixstars Launches Fixstars Vega, a Secure AI Appliance for Embedded Development Companies

Delivering high-security, low-cost AI operations to dramatically improve development efficiency through AI-driven development

IRVINE, California – July 06, 2026 – Fixstars Corporation (TSE Prime: 3687, US Headquarters: Irvine, CA), a leading company in performance engineering technology, today announced the launch of Fixstars Vega, an AI-driven development appliance that delivers high security and low operating costs for companies engaged in embedded software development.

Fixstars Vega is an appliance that enables customers to securely operate verified and optimized state-of-the-art open-weight large language models [1] (open-weight LLMs) in dedicated environments, whether on-premises or in a private cloud. Furthermore, by incorporating the embedded software development expertise Fixstars has cultivated over 20 years in the form of a performance engineering harness [2], Fixstars Vega supports the use of AI in embedded product development.



The Fixstars Vega architecture

Features of Fixstars Vega, a Secure AI Appliance for Embedded Development Companies

Embedded development sites—where systems involve strict security requirements as well as a high degree of specialization and uniqueness—face the challenge that typical cloud-based AI services are difficult to use.

Fixstars Vega is an appliance product that addresses these challenges with advanced security and engineering capabilities.

Feature 1: Unlimited AI usage with no token-based charges

Because open-weight LLMs run in the customer's dedicated local environment (on-premises, private cloud, etc.), there are no pay-per-token charges as with cloud AI services.

Organizations can make full use of AI without worrying about usage volume, keeping costs low even when rolling out AI-driven development across the entire company.

The latest open-weight LLMs have reached performance on par with cloud AI services.

Fixstars will also continuously verify newly released models and deliver them as an appliance tuned to draw maximum performance from the hardware.

Feature 2: High security through AI built in a company-dedicated environment

Because the AI environment is built to match the confidentiality of the information handled and internal policies—on-premises, hosted private cloud, or cloud (VPC)—customers can use AI without sending confidential information such as source code and design documents outside the company network. There is also no risk of input data being used to train external AI services.

Feature 3: A proprietary harness that improves efficiency and quality in embedded development

A harness is a set of software components—providing functions such as context management and tool operation—that AI models need to execute tasks appropriately and efficiently.

Combined with an LLM, it forms an AI agent.

Fixstars Vega incorporates the performance engineering know-how Fixstars has accumulated over 20 years in the form of a harness.

By using AI agents that combine this proprietary harness with verified state-of-the-art open-weight LLMs, users can generate performance-optimized code that general-purpose AI cannot produce.

Acceleration results achieved with Fixstars Vega

- Image segmentation foundation model inference: 4.7x speedup
- SparseDrive inference: 4.0x speedup
- ResNet-50 training: 2.1x speedup

Fixstars Vega Configuration Overview

Fixstars Vega is an appliance that builds and provides a dedicated AI development environment for each customer, on-premises or in a private cloud.

To help each customer identify and deploy the optimal AI development environment, Fixstars also offers careful assessments, consultation, and training programs.

Main Components and Functions

AI infrastructure:

Based on an initial assessment, we prepare the environment best suited to each customer, such as on-premises or private cloud.

Open-weight LLMs:

Rather than simply providing available LLMs as they are, Fixstars verifies and optimizes the LLMs best suited to embedded developers.

Fixstars Vega standard LLMs (as of July 2026): GLM-5.2, Qwen3.6-27B

Gateway for a secure AI environment:

Receives AI queries from each user and routes them to approved AI models based on the customer's pre-configured security policy.

Performance engineering harness:

Fixstars' proprietary harness that optimizes AI output for embedded development, including profiling, bottleneck identification, computation optimization, and optimization for specific devices.

AI coding environment:

A set of peripheral software required for AI-driven development, such as coding assistants and chat UIs, composed of open-source applications.

Associated Services

Secure AI infrastructure design and construction:

We design and build environments that meet your security requirements, such as on-premises, hosted private cloud, or cloud (VPC).

Deployment of AI coding environments:

We deploy tools useful for development work—such as AI chat, AI coding agents, and IDE integration—and support their utilization.

AI model selection, evaluation, and operational design:

AI models are constantly evolving. The Fixstars Vega team will verify newly released models as well, delivering state-of-the-art models optimized for the hardware.

Design for utilizing internal knowledge:

We build mechanisms that allow AI to reference specifications, design documents, internal rules, and past materials.

Training and adoption support:

We support AI utilization through developer seminars, establishment of usage rules, visualization of usage, and more.

For details on Fixstars Vega: <https://www.fixstars.com/en/products-services/vega>

[1] Open-weight LLM

An open-weight LLM is a large language model whose trained weights (parameters) are publicly available, allowing anyone to download and run it in a local environment. In recent years, running such models in dedicated environments—such as local PCs, on-premises servers, or private clouds—as “local LLMs” has been attracting attention.

[2] Harness

A harness is a set of software components—providing functions such as context management and tool operation—that AI models need to execute tasks appropriately and efficiently. Combined with an LLM, it forms an AI agent.

###

About Fixstars Corporation

Fixstars is a technology company dedicated to accelerating AI inference and training through advanced software optimization solutions. It supports innovation in healthcare, manufacturing, finance, mobility, and other industries. For more information, visit: <https://www.fixstars.com/>

Media Contact

Public Relations, Fixstars Corporation

Email: press@fixstars.com

Tel: +81-3-6420-0751