Blaize Introduces AI Platform to Power Multi-Modal Intelligence at the Edge

August 7, 2025

In collaboration with Supermicro and a growing ecosystem of ISV partners, Blaize's AI Platform empowers faster, smarter decisions while reducing latency, cost, and complexity across critical environments

EL DORADO HILLS, Calif.--(BUSINESS WIRE)--Aug. 7, 2025-- Blaize Holdings, Inc. (NASDAQ: BZAI, NASDAQ: BZAIW) ("Blaize"), a leader in programmable, energy-efficient edge AI computing, today unveiled the Blaize AI Platform, a purpose-built, edge-native solution engineered to efficiently and flexibly deliver multimodal intelligence for mission-critical workloads across a wide range of industries. With a focus on programmability and ease of deployment, the platform is designed to accelerate AI implementation without the complexity.

With deployments already underway across Asia, the Blaize AI Platform is redefining how organizations execute and scale AI across distributed environments. Purpose-built for complex, multimodal workloads – including video, audio, and sensor fusion – the platform delivers a powerful combination of performance, energy efficiency, and total cost of ownership, while integrating seamlessly with existing GPU-based infrastructure.

"As data becomes more complex and multi-modal, the need for scalable, hybrid AI infrastructure has never been greater," said **Dinakar Munagala**, Co-founder and CEO of Blaize. "The Blaize AI Platform is already powering public safety, defense, and infrastructure projects – transforming fragmented data sources and systems into unified, responsive intelligence networks. Together with Supermicro, a global leader in energy-efficient, high-performance IT infrastructure for AI, cloud, and edge, and a growing ecosystem of ISVs and AI application providers, we're enabling customers to deploy complete, ready-to-run AI solutions at the edge and beyond."

The Blaize AI Platform was built to replace passive data collection with proactive intelligence at scale. Platform highlights include:

- **Multi-modal sensor fusion:** Consolidates video, audio, telemetry, and more into unified insights.
- **Simultaneous model execution:** Processes multiple models without batching or bottlenecks.
- **Comprehensive support of compact and scalable systems:** Enables hybrid AI deployment across diverse environments and workloads, from rugged edge devices to workstations and rack mount servers.
- **Collaborative software stack**: Pre-integrated with domain-specific AI pipelines.

ISV Ecosystem Expands Real-World Use

Blaize is launching its platform alongside a growing network of ISVs (Independent Software Vendors) whose AI software is optimized to run on Blaize's energy-efficient silicon. These companies are enabling complete AI workflows across edge environments – making the platform faster to deploy, more intelligent in operation, and easier to scale. These ISVs include:

- **AIwaysAI**: Provides tools, APIs, and model libraries for computer vision developers to accelerate time-to-solution.
- **CVEDIA**: Develops synthetic data-driven AI models that deliver accurate, real-time video analytics to enhance security operations without requiring infrastructure overhaul.
- **OrionVM**: Offers a customizable edge-to-cloud infrastructure stack that supports Blaize hardware for real-time inference.
- **Thrive Logic**: Builds custom AI workflows for operational environments—ranging from predictive maintenance to hazard detection.
- **Videonetics:** Integrates AI-driven analytics into its secure, scalable video management software for cities, airports, and critical infrastructure.
- **VSaaS.ai**: Provides a no-code interface to configure, deploy, and monitor AI video pipelines.
- **VSBLTY**: Delivers advanced audience analytics and real-time security software to retail environments, venues, and public infrastructure.

Hardware-Optimized Efficiency with Supermicro

For scalable, rack-mounted edge AI deployments, Blaize's technology is also available in **Supermicro's 1U server system** featuring **24 Blaize® GSP® Xplorer® X1600 EDSFF accelerators**. This compact form factor delivers high-density inference performance for city-wide, industrial, and defense infrastructure – at a fraction of the energy and cooling costs of traditional GPU-heavy servers. The solution enables seamless orchestration across critical sectors and is complementary to existing GPU infrastructure where appropriate.

The Blaize AI Platform is available now. To schedule a demo or request more information, please visit www.blaize.com or contact info@blaize.com. To learn more about Blaize's fast-growing partner ecosystem, please visit our blog.

About Blaize

Blaize provides a full-stack programmable processor architecture suite and low-code/no-code software platform that enables AI processing solutions for high-performance computing at the network's edge and in the data center. Blaize solutions deliver real-time insights and decision-making capabilities at low power consumption, high efficiency, minimal size and low cost. Headquartered in El Dorado Hills (CA), Blaize has more than 200 employees worldwide with teams in San Jose (CA) and Cary (NC), and subsidiaries in Hyderabad (India) and Leeds and Kings Langley (UK). To learn more, visit www.blaize.com or follow us on LinkedIn at @blaizeinc.

View source version on businesswire.com: https://www.businesswire.com/news/home/20250807227857/en/

For media inquiries:
press@blaize.com

For investor relations inquiries:
IR@blaize.com

Source: Blaize Holdings, Inc.