



Blaize Announces Planned Launch of Blaize AI Services to Turn AI Infrastructure into Production-Ready APIs

April 9, 2026

New platform is designed to help cloud providers, data center operators, system integrators, and enterprises deploy application-level AI services faster, reduce cost per query, and create recurring AI service revenue

EL DORADO HILLS, Calif.--(BUSINESS WIRE)--Apr. 9, 2026-- Blaize Holdings, Inc. (Nasdaq: BZAI, Nasdaq: BZAIW) ("Blaize"), a leader in programmable, energy-efficient AI computing, today announced the planned launch of **Blaize AI Services**, a platform designed to help AI infrastructure providers and enterprises deploy production-ready, application-level AI services without building the underlying AI stack from scratch.

Most organizations have run an AI pilot. Far fewer have turned it into something that reliably delivers value at scale, because getting from experiment to production is where the real cost and complexity kicks in. Blaize AI Services solves that gap by combining modular APIs, hybrid computing, and Forward Deployed Engineering into a single platform that makes AI easier to operationalize and scale. The result is lower cost per AI interaction, faster time to value, and AI that grows as a repeatable business capability rather than an expensive one-off project.

As enterprise AI adoption expands, many organizations still face the same bottleneck: getting from working models to operational services that can scale reliably and economically. Providers want to monetize AI infrastructure, but too often are left stitching together fragmented tools, specialized inference functions, and operational workflows. Enterprises face the same challenge from the other side, often forced to assemble point solutions or build internal AI delivery capabilities from the ground up. Blaize AI Services is designed to close that gap through application-level APIs and deploy support that makes AI easier to operationalize.

Unlike approaches centered on raw compute alone, Blaize AI Services is designed around **hybrid inference economics**. The platform intelligently decomposes high-level tasks and schedules the components across Blaize accelerators and GPUs based on cost, power, and performance targets, helping customers improve utilization while supporting a broader range of workloads. Blaize describes this as an open, hybrid architecture built for heterogeneous environments, with standard APIs and tools that integrate into existing infrastructure rather than forcing platform replacement.

The platform is expected to support application-level AI services across vision, video, document processing, speech, and various extended use cases of multimodal workflows. Blaize AI Services is also designed to help customers introduce flexible commercial models, including usage-based pricing, and outcome-based services, creating a path to recurring AI revenue beyond traditional infrastructure leasing or hardware sales. This reflects Blaize's stated platform goal of helping providers compress the path from capex to recurring service revenue.

"AI adoption does not stall because of model availability. It stalls in the last mile between pilot and production," said **Dinakar Munagala, Co-Founder and CEO of Blaize**. "Blaize AI Services is designed to help our customers deploy application-level AI services faster, operate them more efficiently, and monetize infrastructure through repeatable, revenue-generating offerings."

"Providers and enterprises do not need more disconnected AI components," Munagala added. "They need a production-ready way to deliver business outcomes. With Blaize AI Services, we are packaging the core elements required to operationalize AI into modular services that can run across hybrid environments and scale with customer demand."

Key planned benefits of Blaize AI Services include:

- **Faster path from pilot to production** with modular, application-level APIs that package inference, intelligent scheduler and provisioning and business logic into production-ready services
- **Lower infrastructure complexity** through a composable platform that reduces the need to assemble fragmented point solutions
- **Improved hybrid compute efficiency** by scheduling workloads across the most appropriate processing resources based on runtime requirements
- **New recurring revenue opportunities** through application-level AI services that can be offered through usage-based or outcome-based pricing models
- **Deploy support for the last mile** through Blaize Forward Deployed Engineering, which is designed to help customers integrate, configure, and operationalize services in production environments

These benefit areas are consistent with Blaize's messaging pillars around application-level APIs, programmable silicon and hybrid

compute, and Forward Deployed Engineering.

Blaize AI Services builds on the company's broader strategy to deliver practical AI infrastructure that connects silicon, software, and deploy into a unified platform. Blaize's platform messaging emphasizes open integration, sovereign deploy options, and support for both provider and enterprise operating models, including cloud providers, data center operators, system integrators, government organizations, and large enterprises.

To learn more about **Blaize AI Services**, visit <https://www.blaize.com/ai-services-platform/> to explore the platform, watch the overview video, and request a briefing.

About Blaize

Blaize delivers a programmable AI platform, purpose-built for inference in real world environments. Its Hybrid AI architecture enables Practical AI and Physical AI workloads to run efficiently at the edge while integrating seamlessly with cloud and GPU based infrastructure. Blaize solutions support computer vision, multimodal AI, and sensor driven applications across smart cities, industrial automation, telecommunications, retail, logistics, and defense. Blaize is headquartered in El Dorado Hills, California, with a global presence across North America, Europe, the Middle East, and Asia. To learn more, visit www.blaize.com or follow us on LinkedIn @blaizeinc.

Cautionary Statement Regarding Forward Looking Statements

This press release contains forward-looking statements within the meaning of Section 27A of the U.S. Securities Act of 1933, as amended (the "Securities Act"), and Section 21E of the U.S. Securities Exchange Act of 1934, as amended (the "Exchange Act") that are based on beliefs and assumptions and on information currently available to Blaize, including what the AI Services offering will achieve; the industry in which Blaize operates, market opportunities, and product offerings. In some cases, you can identify forward-looking statements by the following words: "may," "will," "could," "would," "should," "expect," "intend," "plan," "anticipate," "believe," "estimate," "predict," "project," "potential," "continue," "ongoing," "target," "seek" or the negative or plural of these words, or other similar expressions that are predictions or indicate future events or prospects, although not all forward-looking statements contain these words. Forward-looking statements are predictions, projections and other statements about future events that are based on current expectations and assumptions and, as a result, are subject to risks and uncertainties. Many factors could cause actual future events to differ materially from the forward-looking statements in this document, including but not limited to: the uncertainties about entering a definitive agreement. These filings identify and address other important risks and uncertainties that could cause actual events and results to differ materially from those contained in the forward-looking statements. Forward-looking statements speak only as of the date they are made. Readers are cautioned not to put undue reliance on forward-looking statements, and Blaize assumes no obligation to update or revise these forward-looking statements, whether as a result of new information, future events, or otherwise, except as required by law, including the securities laws of the United States and the rules and regulations of the SEC. Blaize does not give any assurance that it will achieve its expectations.

View source version on [businesswire.com](https://www.businesswire.com/news/home/20260409740283/en/): <https://www.businesswire.com/news/home/20260409740283/en/>

Blaize Contact

press@blaize.com
www.blaize.com

Investors

ir@blaize.com
www.blaize.com

Source: Blaize Holdings, Inc.