blaize®







Blaize AI Platform: Purpose-built to Enable AI in the Real World



blaize

Actionable Intelligence for critical industry workloads

From traffic intersections to tactical command posts, Blaize delivers purpose-built AI with edge-first performance and flexible cloud integration ready to scale across any environment.

Real-time Intelligence

The world is moving from passive data collection to real-time operational intelligence and Blaize is helping customers make that leap.

Every day, organizations deploy cameras, sensors, and edge devices that see the world but lack the ability to assimilate it. Video Management Systems (VMSes) collect massive amounts of footage, but without leveraging AI, critical insights are delayed, need to be manually reviewed, or missed entirely. The explosion of multi-modal data such as video, audio, sensor fusion, system telemetry running in parallel has quickly outpaced traditional architectures designed for siloed workflows or cloud-only processing.

Blaize AI Platform delivers real Transformation

A hybrid AI platfrom engineered to support **scalable inference** and edge-to-cloud intelligence. Purpose-built for the real-world serving multiple industry verticals, the Blaize AI Platform combines energy-efficient AI hardware, integrated SDK, and vertical-ready applications with an ecosystem of partners, enabling real-time inference directly at the point of action. Blaize helps to empower organizations with a streamlined approach to deploying real-time, multi-modal AI across varied environments.

Blaize provides

Real-time AI inference where operations happen on-site, at the edge, in-real time.

Leverage our best-in-class Blaize AI Platform for VMS application or implement for Audio, LiDAR, Radar, Thermal and Sensor Fusion for a true multi-modal approach to AI enabled data management.

Support autonomous safety, logistics, vision and operational decision-making at scale

Real-time Intelligence

The Blaize AI Platform moves beyond theoretical AI to deliver real-time intelligence directly into operational environments. Built for physical world deployments, Blaize enables true multi-modal AI to process complex data where it's generated, and seamlessly connect to the cloud for centralized orchestration and deeper analytics making intelligent decisions in the moment and informing long-term strategies.

Blaize Advantages

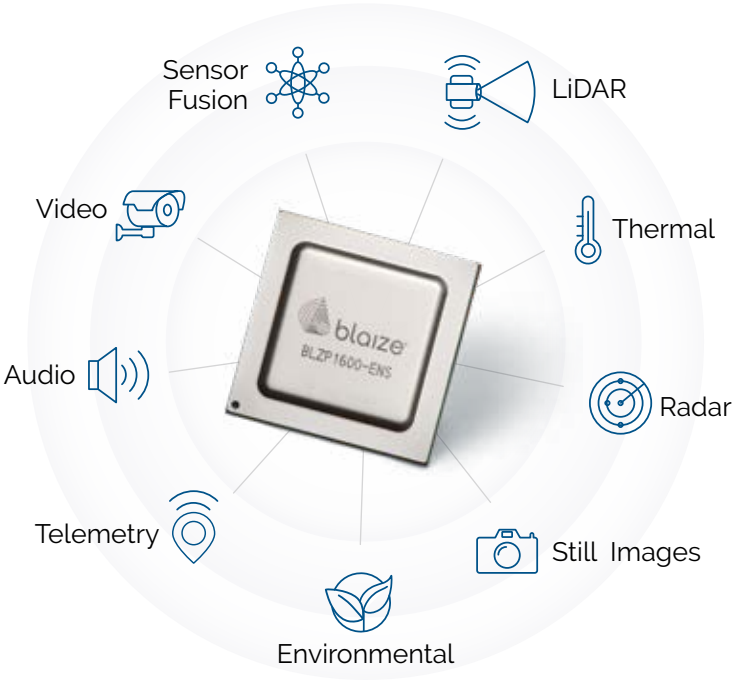
Autonomous, real-time inference across video, audio, environmental, LiDAR, radar, thermal, sensor and telemetry data streams

Integrated edge-to-cloud orchestration that synchronizes models, data, and alerts across distributed locations and assets

Multi-modal sensor fusion that combines disparate data types into unified, context-aware event detection

One unified platform that supports safety, security, logistics, maintenance, and mission operations simultaneously eliminating fragmented AI silos

Multi-Modal Capability



Sensor + data streams visual flow diagram

Unlocking next-generation operational capabilities

- **PPE and safety compliance monitoring** using thermal imaging, video analytics, and access control integration
- **Proactive equipment health diagnostics** by analyzing vibration patterns, sound anomalies, and telemetry deviations in real time
- **Public safety threat detection** that fuses synchronized video feeds and directional audio to assist with situational awareness, helping with quicker response times and decision making.

24/7 automated monitoring

Faster incident detection = Reduced congestion

Resource-optimized patrols

Public satisfaction
via safer, smoother commutes

Today's Reality

Today's physical world infrastructure was not built for real-time AI. Across industries, organizations continue to deploy sensors, cameras, and systems that collect vast volumes of data, but these deployments may lack the necessary intelligence to process and act on that data with the speed and efficiency needed to deliver desired outcomes.

Many organizations are working with

Outdated VMS and SCADA systems that capture data but offer no native analytics or automated decision-making

Cloud-dependent AI pipelines that introduce unacceptable latency, recurring bandwidth costs, and operational risks especially in remote, disconnected, or sovereign environments

GPU-centric architectures that are costly, power-intensive, oversized, and increasingly impractical for field deployments and distributed operations

Fragmented multi-sensor deployments where data is trapped in independent silos, preventing unified visibility and operational response eliminating fragmented AI silos

These translate directly into business bottlenecks

Delayed event detection resulting in missed safety incidents, security gaps, and operational downtime

Failure to correlate cross-modal signals like video + audio + telemetry, undermining situational awareness

High cost of ownership due to expensive infrastructure, manual oversight, and complex integration workstreams

Rigid architectures that cannot scale across diverse environments or adapt as new AI models and sensor types emerge

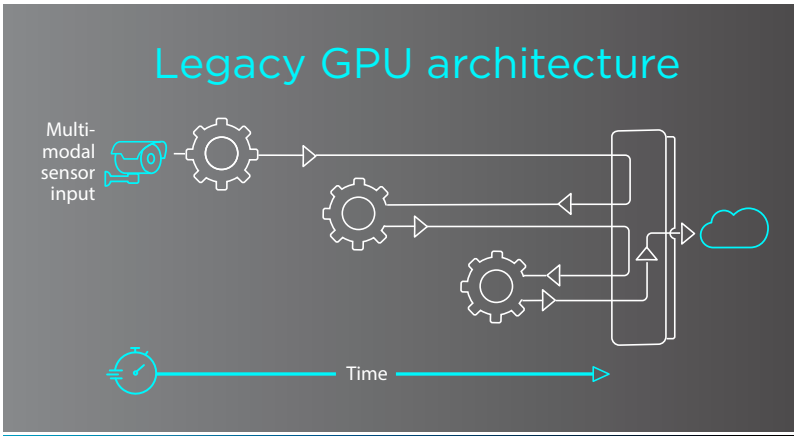
The Blaize Differentiation

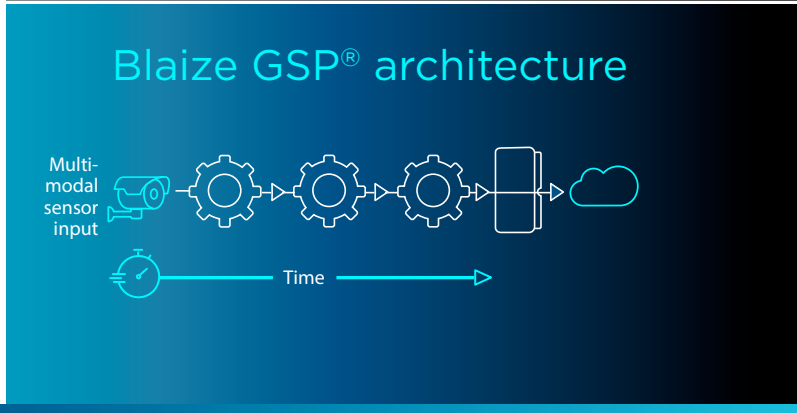
Legacy GPU architecture

Traditional solutions suffer from high latency and require a significant power draw from constant DRAM transfers.

Blaize AI Platform: purpose-built to eliminate these barriers

The Blaize GSP® (Graph Streaming Processor) replaces reactive, disconnected systems with integrated, real-time intelligence across modalities, data types, and deployments providing efficient task-level parallelism without redundant memory transfers





Where Blaize Fits into the Future

Blaize AI Platform supports advanced, cross-modal deployments across key verticals, some examples include:

Smart City

Monitor traffic, loitering, crowd anomalies, and public safety threats using combined visual and acoustic inference.

Smart Energy

Monitor pipelines, detect leaks, and analyze safety and compliance events across remote infrastructure using integrated sensor and video streams.

Smart Defense

Integrate ISR, drone detection, audio cues, and field telemetry to power decisions in contested environments.

Smart Healthcare

Enable patient and staff safety through vision-based fall detection, occupancy monitoring, and real-time anomaly alerts.

Smart Retail

Gain insights into shopper behavior, shelf activity, and inventory flow through real-time visual and environmental analytics.

Smart Manufacturing

Monitor equipment health, worker safety compliance, and production flow using integrated AI across visual and sensor inputs.

Smart Restaurant

Monitor drive- thru, kitchen performance, staff presence, and occupancy trends in real time.

Smart Mining

Combine camera, thermal, sensor, and equipment data to monitor compliance, productivity, and safety.

Smart Airport

Combine visual and sensor streams to monitor flow, baggage safety, and restricted zone activity.

Smart School

Use access control, behavioral monitoring, and audio signals to detect and respond to anomalies.

Smart Automotive

Power in-cabin monitoring, object detection, and predictive maintenance for multi-modal sensor fusion, helping with vehicle autonomy and safety applications.

Each deployment unlocks real-time AI across domains
not just what you see, but what you sense, hear, and measure.

The Ecosystem

Blaize partners with leading innovators to ensure Blaize AI Platform remains open, extensible, and vertical-ready.

The Blaize Solution

The Blaize AI Platform is a purpose-built, fully integrated system designed to deploy and operate AI directly within physical environments where latency, power, and infrastructure constraints have historically limited adoption. Hybrid AI systems that include a GPU and Blaize GSP® can help to optimize cost, compute, and power for edge-to-cloud inference applications. Blaize natively supports real-time, multi-modal inference at the edge while seamlessly extending into cloud-based orchestration as needed.

At the core is Blaize's Graph Streaming Processor (Blaize GSP®), a proprietary architecture that executes multiple AI models concurrently, without batching, while dramatically reducing power and compute overhead. This enables AI to operate in real-world conditions: inside factories, vehicles, critical infrastructure, campuses, and field-deployed environments, not just inside the data center.

Core Capabilities

- **Supporting 1,000+ cameras at scale** in VMS applications with 6X Blaize® Xplorer® X1600p-Q PCIe Card, providing Up to 2,160 video streams at 720p (12,000 fps total) and up to 960 video streams at 1080p (5,320 fps total) to help high-throughput performance.
- **True multi-modal concurrency** running simultaneous video, audio, telemetry, and environmental models without compromising speed or accuracy.
- **Deployable, rugged hardware form factors** designed for edge sites, disconnected operations, and constrained spaces.
- **Blaize's full software stack** works in concert with partner solutions to help streamline sensor onboarding, model deployment, and multi-modal orchestration across edge and cloud environments.
- **Hybrid edge-to-cloud integration** for centralized orchestration, multi-site management, and long-term analytics without sacrificing on-site response time or data control.

Blaize AI Platform is built to help support AI in diverse deployment scenarios, including air-gapped edge locations and networked campus environments.

Key benefits

The Blaize AI Platform transforms how customers deploy and scale AI across physical environments, helping to deliver immediate operational, financial, and technical advantages.

- **Real-time decision-making at the edge**
  Process multi-modal data directly at the point of capture for instant alerts, faster response, and higher safety and security outcomes.
- **Hybrid AI System Efficiency (GPU and Blaize GSP®)**
  Blaize's Graph Streaming Processor works alongside existing GPU infrastructure in a hybrid AI approach, assisting with optimizing compute allocation, reducing power consumption, and helping to lower overall operational costs.
- **Eliminate cross-sensor blind spots**
  Fuse video, audio, telemetry, and environmental data for unified, context-rich event detection and situational awareness.
- **Faster time-to-deployment**
  Pre-integrated hardware paired with partner-driven orchestration solutions helps enable rapid field activation without complex custom integration cycles.
- **Sovereign, secure, and compliant AI operations**
  Localized processing keeps sensitive data on-prem where necessary, supporting privacy, security, and industry-specific compliance requirements.
- **Scalable across industries and sites**
  From single-site deployments to city-wide networks, Blaize's modular platform adapts to growing infrastructure and evolving use cases.

The result is accelerated insight, improved safety, helping to lower TCO, and operational intelligence that scales while reducing complexity at every level.

Blaize AI Platform: Building Blocks

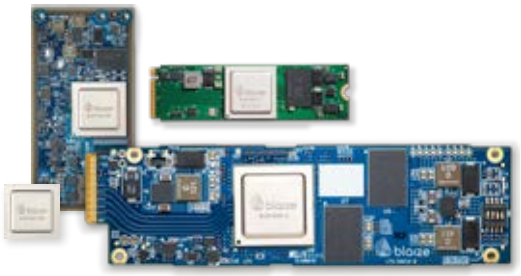
Full stack, Fully integrated Infrastructure

- Plug-and-play application onboarding.
- Real-time incident visibility and response.
- Self-managed topology (campus, city, facility).
- Rapid deployment with no-code AI application testing and rollout

Blaize can scale from a single site to regional or national infrastructure—**smart cities, campuses, facilities, and beyond.**

Blaize Graph-Native Architecture Processor

Blaize revolutionary Graph Streaming Processor (GSP®) architecture built to be fully programmable and efficient for AI applications.



Blaize Accelerators

SOC, embedded and accelerator platforms built for maximum system level performance and efficiency.

Blaize Software – SDK, AI Studio & Picasso

Comprehensive software portfolio – easily build, optimize and productize complete AI applications.

Blaize Systems*

A range of systems from deep edge devices to rack mount servers.

A unified AI platform delivering intelligence across Edge, Cloud, and Hybrid environments

*System configurations shown are representative examples. Actual implementations may vary based on specific requirements and deployment environments.

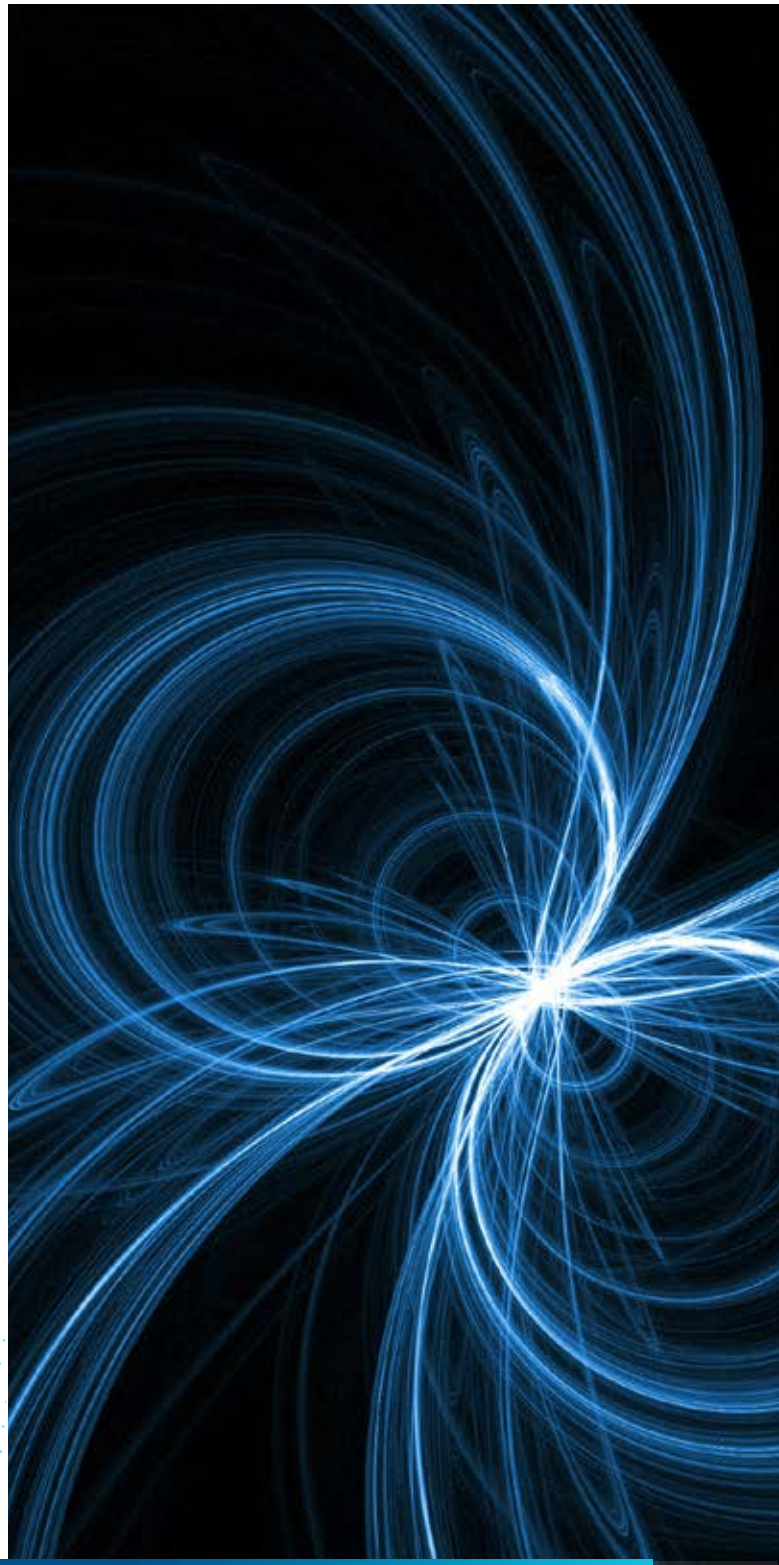
Why Blaize

The Blaize AI Platform delivers real-world AI where it matters most, directly at the point of data capture while eliminating the cost, power, and complexity barriers of legacy GPU and cloud-first approaches.

Powered by Blaize's proprietary Graph Streaming Processor (Blaize GSP®), and supporting SDK, the platform executes multiple AI models simultaneously across video, audio, sensor, and telemetry inputs with lower power consumption and a smaller hardware footprint.

- **Deploy anywhere AI is needed** even in disconnected or resource-constrained environments
- **Run more concurrent models** using less compute and energy than traditional architectures
- **Leverage partner-delivered orchestration** and deployment tools for rapid activation
- **Scale from edge to cloud** with unified management and data control

Blaize enables multi-modal AI that's operational, scalable, and built for the real world.



Take the Next Step to Enable AI in your Environment

With Blaize, you can build on what you have and create a foundation for what's next.

Schedule a Consultation

Assess your goals and create a tailored deployment plan.

Request a Demo

Experience real-time detections and intelligent orchestration in action.

Launch a Pilot Program

Validate performance at priority sites and prove ROI.

Plan for Scale

Develop a roadmap for phased expansion.

Contact us at
info@blaize.com

Visit us at
www.blaize.com

MSB05007
v1.0 Aug 2025