

Why aion?

Move from endless AI strategizing to immediate production value with aion's full-stack platform.

We don't just provide infrastructure, we build the complete stack. Our FDEs act as your AI R&D team to deliver the outcomes and innovation your business needs- without ever hiring an internal AI team.

GPUs available:

 **40,000 chips** **800 MW**

<> SIMPLE

Write code, not infrastructure. Smart routing and automated resource management provide a seamless developer experience with transparent pricing.

🚀 FAST

Directly access frontier models on aion-owned high-bandwidth clusters with sub-second inference speed.

⚙️ OPTIMIZED

Built-in performance optimizations automatically reduce cost and maximize throughput across training and inference.

The Enterprise AI Paradox

88%
of organizations use AI
in at least one function

...but only 7%
have scaled AI across
the enterprise

Our Research Team



From pilot to production by combining research, infrastructure, and execution into one forward-deployed AI platform

FDEs

- Engineers & Integration Deployment
- Research Partnerships

Forge

- Bare Metal
- Kubernetes
- Slurm
- Identity & RBAC



Nexus

- Inference
- Models
- Optimize
- Data
- Agents
- Public Cloud Integration

Factory

Owned infrastructure of 800 MW

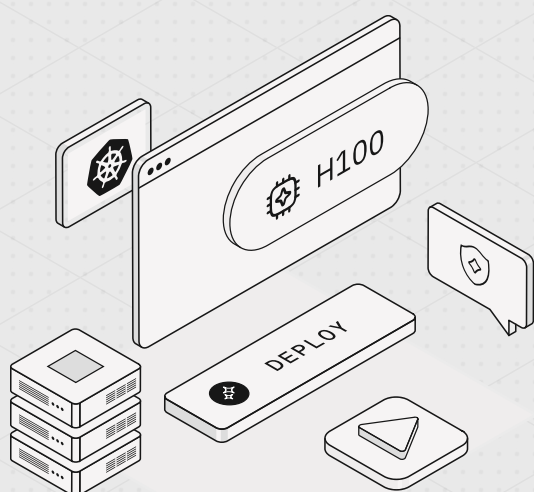
LEARN MORE



aion.xyz

aion Forge: The Most Powerful, Optimized AI Engine

FORGE



Start Instant
Deployment

With aion



Deployment Flexibility

Bare-metal, Kubernetes, & Slurm

Support for out-of-the-box or custom workflows and research

Secure Multi-tenancy

Built-in RBAC and identity management



Optimization

49%+ MFU

Performance that maximizes hardware ROI

Unified optimization

Software + hardware optimization across AI apps and infrastructure for comprehensive performance



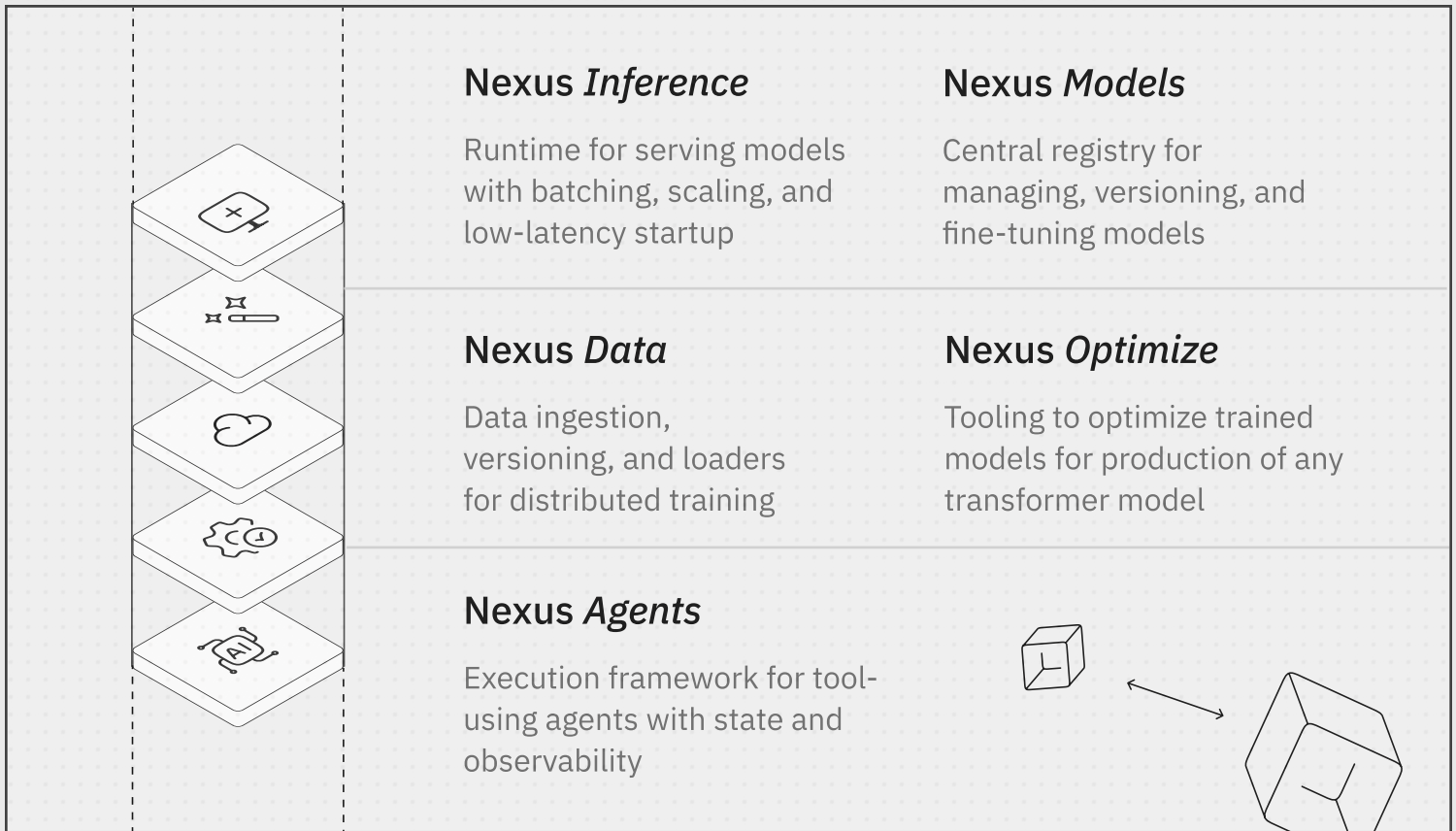
Execution

Developer-friendly






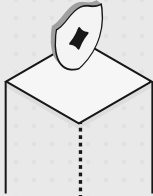

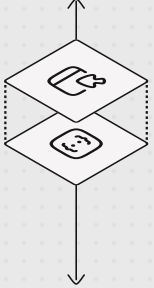

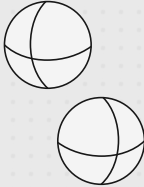


Deploy in minutes, not weeks, with extensive documentation, APIs & SDKs

Transparent pricing

Pay per usage with zero upfront commitments



Nexus Core (Always-On Control Plane)

 <p>Platform and IP Licensing Customer retains access to aion's AI models, optimization techniques, and infrastructure</p>		 <p>Observability & Granular Billing Every resource is tracked and billed per unit on a single dashboard</p>	
 <p>Enterprise Security & Compliance RBAC, SSO/SAML integration, audit logging, policy enforcement</p>		 <p>Immediate Deployment Zero-to-prod AI runtime without GPU orchestration or cluster ops</p>	
 <p>Hybrid Cloud Orchestration Unified compute layer across training, fine-tuning, and inference, using aion's private cloud or BYOC (public/on-prem)</p>		 <p>Constant Optimization Automatic scaling, security, and reliability by default</p>	

aion Factory: Owned Infrastructure for Full Control

FACTORY

No more queueing for compute.

aion provides direct access to 800 MW of owned machines.



Guaranteed Performance

Dedicated capacity with zero resource contention



Architecture Flexibility

Deployment across GPU architectures (GB200, GB300, future chips)



Custom Configurations

Specialized hardware setups for specific research workloads



Cost Advantage

Transparent pricing without intermediary markup







Sovereign & Secure

Regional data residency, ISO 27001-compliant, and zero-trust architecture for regulated industries

LEARN MORE



aion.xyz

	Traditional Software Vendors	aion: Your Full-Stack AI Partner
 Partnership	Vendors sell you tools, then abandon you.	On-demand AI engineers find the biggest wins and build custom solutions. Our platform runs it all transparently.
 Compute Access	They force you into long waitlists for GPUs, then lock you in with limited capacity.	We provide on-demand GPUs for the exact clusters you need: bare-metal, Kubernetes, or serverless.
 Pricing	They force you into lock-in contracts.	You pay-as-you-go in a usage-based system.
 Timeline	POCs take quarters to build.	We build long-term, optimized solutions in weeks, not months.

