KAMI WAZA

intel®

Intel® + Kamiwaza: Secure, efficient AI where your data lives

kamiwaza.ai 2025

Intel® + Kamiwaza: Secure, efficient AI where your data lives

Organizations with sensitive data face a fundamental challenge: they need the insights that artificial intelligence (AI) can provide, but traditional AI solutions often require moving data to centralized processing environments, creating security and compliance risks. Meanwhile, the computational demands of modern AI strain infrastructure resources and budgets.

Kamiwaza, powered by Intel® Gaudi® 3 accelerators, offers a breakthrough solution. This technology brings AI processing directly to the data, eliminating unnecessary transfers while delivering exceptional performance and efficiency. By processing information where it already lives — whether on-premises, in the cloud, or at the edge — organizations can maintain complete control over sensitive data while still leveraging the full power of modern AI.

The partnership between Kamiwaza and Intel® combines cutting-edge AI orchestration with industry-leading hardware acceleration, resulting in a solution that:

- Keeps sensitive data secure by processing it within existing security boundaries

- Transforms legacy data formats into modern, queryable assets

- Reduces processing time from days to minutes for complex analytical tasks

- Decreases energy consumption by up to 40% compared to competing solutions

- Enables non-technical users to access complex data through natural language

- Scales seamlessly from single workstations to enterprise-wide deployments

This solution brief explores how Kamiwaza and Intel® Gaudi® 3 deliver these capabilities and how organizations across government, healthcare, finance, and other regulated industries can implement this technology to turn their most sensitive data into actionable intelligence without compromise.

The challenge: AI for sensitive data

The security dilemma

As AI becomes essential to organizational strategy, many enterprises face a critical dilemma:

- **Security requirements** — Sensitive data often can't leave secure environments due to contractual obligations, intellectual property protection, or regulatory guidelines, like:

  - Health Insurance Portability and Accountability Act (HIPAA)

  - General Data Protection Regulation (GDPR)

  - Cybersecurity Maturity Model Certification (CMMC)

- **AI processing demands** — Traditional AI approaches require centralized processing and data movement, creating security vulnerabilities, compliance risks, and operational inefficiencies.

- **Legacy data formats** — Many organizations have decades of valuable data locked in legacy formats that are difficult to access, analyze, or integrate with modern AI systems.

- **Technical expertise gaps** — Data science expertise is scarce, making it difficult for organizations to extract value from complex datasets without specialized skills.

Infrastructure limitations

Even when organizations commit to AI adoption, implementation challenges persist:

- **Performance constraints** — Traditional hardware struggles to process the massive datasets required for effective AI, leading to analysis delays that impact decision-making.

- **Energy consumption** — AI workloads are notoriously power-hungry, straining data center capacity and conflicting with sustainability goals.

- **Cost barriers** — The expense of specialized AI hardware and the expertise to implement it creates significant barriers to entry for many organizations.

- **Fragmented data environments** — Enterprise data typically exists across multiple locations and formats, making comprehensive analysis difficult without risky centralization.

Organizations need a solution that addresses these challenges comprehensively — one that brings AI capabilities to data rather than forcing data to move to centralized AI systems, all while delivering performance without excessive costs or complexity.

The solution: Kamiwaza + Intel® Gaudi® 3

Overview

Kamiwaza, powered by Intel® Gaudi® 3 accelerators, creates a unified solution for secure, efficient AI processing across distributed data environments. This partnership combines:

- **Kamiwaza's AI orchestration technology** — Intelligently processes data where it resides, converts legacy formats, and provides natural language access to complex datasets.

- **Intel® Gaudi® 3 accelerators** — Deliver exceptional performance and efficiency specifically optimized for AI workloads, with industry-leading networking capabilities for distributed processing.

The result is a comprehensive solution that enables organizations to analyze sensitive data securely, efficiently, and cost-effectively without compromising on performance or capabilities.

Key capabilities

Process your data where it lives
- Analyze sensitive information without moving it outside secure environments

- Maintain compliance with industry regulations and security policies

- Deploy consistently across on-premises, cloud, and edge environments

- Connect previously siloed data without centralizing it

Transform legacy data

- Convert outdated formats into modern, queryable assets
- Unlock insights from decades of historical information
- Automatically clean and normalize data across sources
- Create unified views of previously fragmented information

Accelerate time-to-insight

- Process billions of data points in minutes rather than days
- Enable natural language queries for non-technical users
- Automatically correlate data across multiple sources
- Deliver analysis in interactive timeframes

Optimize resource use

- Reduce energy consumption by up to 40% compared to competing solutions
- Lower total cost of ownership with better price-performance ratio
- Scale efficiently as needs grow without infrastructure overhauls
- Minimize infrastructure footprint through improved efficiency

Technical architecture.

Kamiwaza architecture

Kamiwaza's technology stack consists of three primary components: the distributed data engine, inference mesh, and a natural language interface.

Distributed data engine

- Intelligently locates and accesses data across environments
- Converts legacy formats to modern, standardized structures
- Implements automatic data cleaning and normalization
- Ensures data remains within security boundaries during processing

Inference mesh

- Orchestrates AI workloads across distributed environments
- Optimizes resource allocation based on workload requirements
- Provides unified access to distributed processing resources
- Manages fault tolerance and resilience for AI tasks

Natural language interface

- Translates business questions into optimized queries

- Enables non-technical users to access complex data

- Integrates with existing business intelligence tools

- Presents results in accessible, actionable formats

Intel® Gaudi® specifications

Intel® Gaudi® 3 accelerators represent a significant advancement in AI processing capabilities.

Processing capabilities

- 64 Tensor Processor Cores (TPCs) and 8 Matrix Multiplication Engines (MMEs)

- 1.8 PFlops of FP8 and BF16 compute

- 4x AI compute improvement for BF16 over previous generation

- Optimized for dense matrix operations common in AI workloads

Memory & bandwidth

- 128GB HBM2e with 3.7 TB/s bandwidth (1.5x increase over Gaudi® 2)

- 96MB on-chip SRAM for frequently accessed data

- Advanced memory management for complex AI models

Networking & scaling

- 24x 200GbE industry-standard RoCE Ethernet ports

- 2x networking bandwidth compared to the previous generation

- Native support for RoCE v2 protocol

- Scales from 1 to hundreds of nodes without proprietary interconnects

Software integration

- Seamless integration with popular AI frameworks (PyTorch, TensorFlow)

- Support for Hugging Face and other model repositories

- Optimized libraries for common AI operations

- Open architecture for flexible deployment

Integration architecture

Kamiwaza and Intel® Gaudi® 3 integrate through a reference architecture that ensures optimal performance and security.

Deployment patterns

- Direct integration with existing data infrastructure

- Non-invasive implementation requiring minimal changes to data sources

- Flexible deployment across on-premises, cloud, and edge environments

- Support for hybrid architectures with data spanning multiple domains

Security integration

- Preservation of existing security boundaries and controls

- Integration with enterprise identity and access management

- Encryption throughout the processing pipeline

- Comprehensive audit logging for compliance requirements

Optimization layer

- Automatic workload tuning for Intel® Gaudi® 3 capabilities

- Dynamic resource allocation based on workload characteristics

- Predictive scaling to anticipate processing requirements

- Continuous performance monitoring and adaptation

Performance & security benchmarks

Performance benchmarks

Data processing speed

- **Legacy format conversion** — 500GB/hour (33,000 files processed in under 4 hours)

- **Query response time** — 90% of queries completed in under 10 seconds

- **Complex analysis (billion+ data points)** — 10-minute completion versus days with traditional methods

- **Batch processing improvement** — 85% reduction in processing time compared to CPU-only systems

Scalability

- Linear performance scaling to 64 Gaudi® 3 accelerators

- Near-linear scaling to 256 accelerators for larger workloads

- Minimal performance degradation with heterogeneous data sources

- Consistent performance across on-premises and cloud deployments

Energy efficiency

- 40% lower power consumption compared to leading GPU alternatives

- 65% improvement in performance-per-watt over CPU-only environments

- 85% reduction in cooling requirements due to optimized thermal design

- Average 52% decrease in total energy consumption for equivalent workloads

Security capabilities

Data protection

- Zero data movement outside established security boundaries

- End-to-end encryption for all processing activities

- Granular access controls tied to enterprise identity systems

- Complete audit trail of all data access and processing

Compliance support

- HIPAA compliance for healthcare deployments

- FISMA/FedRAMP capabilities for government environments

- GDPR-ready processing that respects data sovereignty

- SOC 2 compliance for commercial implementations

Resilience

- Automatic failover for hardware failures

- Self-healing capabilities for software components

- Graceful degradation under extreme loads

- 99.99% availability in production environments

Customer success story: Department of Homeland Security.

Challenge

The Department of Homeland Security (DHS) needed to transform its approach to analyzing weather data for emergency preparedness and response. Their challenges included:

- Processing 1.3 billion weather data points spanning 90 years

- Converting outdated GEMPAK format files into modern, queryable formats

- Enabling non-technical emergency response teams to access insights

- Meeting strict government requirements for security and energy efficiency

- Reducing analysis time from days to minutes to enable rapid response

Solution

DHS implemented Kamiwaza, powered by Intel® Gaudi® 3 accelerators, to transform their weather analysis capabilities:

- Deployed Kamiwaza to orchestrate AI processing across distributed data sources

- Used Intel® Gaudi® 3 accelerators for high-performance, energy-efficient processing

- Implemented natural language querying for emergency responders

- Created automated processes to correlate weather anomalies with historical events

- Developed a system that automatically retrieves news and reports related to similar weather patterns

Results

The implementation delivered transformative results for DHS:

- **Reduced analysis time** — Processing that previously took days now completes in minutes.

- **Improved accessibility** — Non-technical team members can now query complex weather data.

- **Enhanced prediction** — The system now automatically correlates current conditions with historical patterns.

- **Energy efficiency** — DHS now meets stringent government requirements for sustainable computing.

- **Expanded capabilities** — DHS now provides insights to organizations from transportation companies to theme park operators.

> **"It was extremely important that this project run on hardware that was efficient for both energy and water consumption, which is why Dell servers with Intel® Gaudi® accelerators were an ideal fit. This transformation has revolutionized our emergency preparedness."**
> **— Sunny Wescott, Chief Meteorologist, DHS**

Implementation scenarios

Government intelligence analysis

Scenario

Government agencies need to analyze secure information across multiple security domains while maintaining strict compliance with security regulations.

Implementation

- Deploy Kamiwaza within private networks at multiple security levels
- Intel® Gaudi® 3 accelerators installed in secure data centers
- Processing orchestrated to maintain data within appropriate security boundaries
- Automated correlation of information across domains where permitted
- Natural language access for intelligence analysts without specialized data science skills

Outcome

- 85% reduction in analysis time for complex intelligence datasets
- Elimination of risky data transfers between security domains
- Enhanced analytical capabilities without increasing security risks
- Improved accessibility of insights for decision-makers

Healthcare research & patient care

Scenario

Healthcare organizations need to leverage AI across patient records, research data, and operational systems while maintaining HIPAA compliance.

Implementation

- Deploy Kamiwaza across clinical and research environments

- Intel® Gaudi® 3 accelerators installed in existing data centers

- Processing configured to keep patient data within secure boundaries

- Integration with existing clinical systems and research databases

- Natural language interface for clinicians and researchers

Outcome

- Identification of patterns across decades of patient records in minutes

- HIPAA-compliant processing that never exposes PHI

- 78% faster research data analysis without specialized expertise

- Improved diagnostic accuracy through historical pattern analysis

Financial services & fraud detection

Scenario

Financial institutions need real-time fraud detection while processing sensitive transaction data that can't leave secure environments.

Implementation

- Kamiwaza deployed across transaction processing environments

- Intel® Gaudi® 3 accelerators integrated with existing fraud detection systems

- Processing optimized for real-time analysis of transaction patterns

- Integration with regulatory reporting and compliance systems

- Automated alerts based on historical and emerging fraud patterns

Outcome

- 94% reduction in false positive fraud alerts

- 65% improvement in fraud detection rates

- Regulatory compliance with data locality requirements

- Processing of 10,000+ transactions per second with millisecond latency

Critical infrastructure protection

Scenario

Utilities and industrial facilities need to analyze operational technology data for cybersecurity and efficiency without exposing control systems.

Implementation

- Kamiwaza deployed at the edge, near operational technology systems
- Intel® Gaudi® 3 accelerators installed in hardened industrial environments
- Air-gapped processing that never exposes control systems to external networks
- Integration with existing SCADA and industrial control systems
- Automated anomaly detection and correlation with historical patterns

Outcome

- Real-time detection of operational anomalies before they impact service
- 72% reduction in false alarms for security events
- Preservation of air-gapped security while enabling advanced analytics
- Predictive maintenance capabilities, reducing outages by 45%

Deployment options

On-premises deployment

Architecture

- Kamiwaza software deployed on an existing enterprise infrastructure
- Intel® Gaudi® 3 accelerators integrated into data center environments
- Direct connection to on-premises data sources
- Integration with existing security infrastructure
- Optional connection to cloud resources for hybrid scenario

Ideal use cases

- Organizations with significant existing data center investments
- Environments with strict data sovereignty requirements
- Situations requiring maximum control over infrastructure

- Applications with predictable, consistent workloads

Implementation considerations

- Minimum two Gaudi® 3 accelerators recommended for production environments
- Standard rack space and cooling requirements
- Integration with existing orchestration systems (Kubernetes, Red Hat® OpenShift®)
- On-site implementation support available

Cloud deployment

Architecture

- Kamiwaza deployed on cloud infrastructure
- Intel® Gaudi® 3 accelerators accessed through cloud provider offerings
- Direct integration with cloud-native data services
- Secure connection to on-premises data sources where needed
- Elastic scaling based on workload demands

Ideal use cases

- Organizations embracing cloud-first strategies
- Applications with variable or unpredictable workloads
- Development and testing environments
- Scenarios requiring global distribution of processing

Implementation considerations

- Available on major cloud providers (specifics vary by provider)
- Configurable security to meet compliance requirements
- Consumption-based pricing options available
- Remote implementation support provided

Edge deployment

Architecture

- Kamiwaza deployed on edge servers or appliances
- Intel® Gaudi® 3 accelerators in compact, ruggedized configurations
- Direct connection to local data sources and sensors

- Optional connection to central infrastructure for aggregation

- Optimized for intermittent connectivity scenarios

Ideal use cases
- Remote operations with limited connectivity

- Industrial environments requiring local processing

- Scenarios with high-volume, low-latency data sources

- Applications requiring processing before transmission

Implementation considerations
- Ruggedized hardware options for challenging environments

- Simplified deployment requiring minimal local expertise

- Autonomous operation with central management

- Field-deployable implementation services available

Hybrid deployment

Architecture
- Kamiwaza orchestrating across on-premises, cloud, and edge environments

- Intel® Gaudi® 3 accelerators distributed according to workload requirements

- Unified management across all deployment locations

- Intelligent workload routing based on data location and security requirements

- Consistent security model spanning all environments

Ideal use cases
- Organizations with data distributed across multiple environments

- Complex regulatory environments with varying data requirements

- Global operations with regional data sovereignty concerns

- Gradual migration strategies from on-premises to cloud

Implementation considerations
- Unified management console for all environments

- Consistent security and compliance across deployments

- Flexible scaling across deployment models

- Comprehensive implementation planning services available

Getting started

Assessment process

The journey to implementing Kamiwaza with Intel® Gaudi® 3 begins with a comprehensive assessment.

Data environment assessment

- Inventory of data sources, formats, and volumes

- Evaluation of security requirements and boundaries

- Assessment of existing infrastructure and integration points

- Identification of priority use cases and success metrics

Technical readiness evaluation

- Infrastructure readiness for Gaudi® 3 integration

- Network capacity and connectivity evaluation

- Security and compliance gap analysis

- Skills assessment and training needs identification

Business value mapping

- Quantification of current process inefficiencies

- Identification of high-impact use cases

- ROI projection based on benchmarked performance

- Implementation roadmap development

Implementation process

A typical implementation follows these key phases.

Implementation phase	Timeline	Activities
Phase 1: Proof of concept	4-6 weeks	• Installation of Dell PowerEdge XE9680 with Intel® Gaudi® 3 and Kamiwaza in a controlled environment • Integration with representative data sources • Validation of performance and security capabilities • Demonstration of key use cases

Implementation phase	Timeline	Activities
Phase 2: Pilot implementation	6-12 weeks	• Deployment in a production environment with a limited scope • Integration with prioritized data sources • User training and feedback collection • Performance optimization and tuning
Phase 3: Production expansion	3-6 months	• Scaled deployment across target environments • Integration with a full range of data sources • Advanced feature enablement • Establishment of operational processes and governance
Phase 4: Continuous optimization	ongoing	• Regular performance analysis and tuning • Expansion to additional use cases • Integration of new data sources • Adoption of new capabilities as released

Support & services.

Kamiwaza and Intel® provide comprehensive support throughout the implementation journey.

Service category	Offerings
Technical support	• 24/7 support for production environments through Dell ProSupport • Direct access to implementation specialists • Regular maintenance and updates • Performance monitoring and optimization
Professional services	• Implementation planning and execution from Dell Services • Custom integration development • Performance optimization • Security hardening and compliance validation

Service category	Offerings
Technical support	• 24/7 support for production environments through Dell ProSupport • Direct access to implementation specialists • Regular maintenance and updates • Performance monitoring and optimization
Professional services	• Implementation planning and execution from Dell Services • Custom integration development • Performance optimization • Security hardening and compliance validation
Training and enablement	• Administrator training for deployment and management • Developer training for integration and customization • End-user training for natural language interface • Executive briefings on capabilities and roadmap

About the companies

Intel®

Intel® (Nasdaq: INTC) is an industry leader, creating world-changing technology that enables global progress and enriches lives. Inspired by Moore's Law, we continuously work to advance the design and manufacturing of semiconductors to help address our customers' greatest challenges. By embedding intelligence in the cloud, network, edge and every kind of computing device, we unleash the potential of data to transform business and society for the better.

Kamiwaza

Kamiwaza revolutionizes how organizations harness AI while keeping their data secure. Our technology brings AI processing directly to the data, rather than the other way around, seamlessly orchestrating AI workloads across on-site systems, cloud environments, and network edges. By processing AI exactly where data lives — whether behind strict firewalls or regulatory boundaries — Kamiwaza empowers organizations to deploy cutting-edge AI with unprecedented security, speed, and control.

A proud graduate of the Intel® Liftoff program, Kamiwaza continues to partner with Intel® and Dell to deliver transformative AI solutions for organizations with the most demanding security and performance requirements.

Intel®, the Intel® logo, and Gaudi® are trademarks of Intel Corporation or its subsidiaries.