Omdia by informa techtarget

MAY 2026

Mission-Critical in the AI Era: How Oracle AI Database Delivers Diamond-Grade Reliability for Agentic AI Workloads

Stephen Catanzano, Principal Analyst

Abstract: AI agents are moving from experiment to production faster than most enterprise infrastructure is ready to support them. As these systems take on autonomous, multi-step workflows with direct business consequences, the definition of "mission-critical" is being rewritten. Oracle AI Database addresses this shift by delivering a converged data architecture with extreme availability, built-in cyber resilience, and global multicloud reach, engineered specifically for the agentic AI era. This Showcase examines why mission-critical reliability is now inseparable from enterprise AI strategy and how Oracle's approach positions organizations to build durable, highly scalable AI infrastructure before the workload demands it.

The architecture problem hiding behind every stalled AI production deployment

There is a pattern playing out quietly across enterprise AI programs right now. A team builds something that works beautifully in the lab: a customer-facing agent, an internal knowledge tool, an automated workflow. Then, they try to take it to production, and the wheels come off, resulting in latency spikes, governance gaps, agents accessing data they should not, and agentic AI applications crashing midstream. The fundamental data platform cannot scale to keep up with hundreds or thousands of agents, and, when something goes wrong, the organization has no clean recovery path.

80% of organizations viewed AI agents as a top or high-priority compared with other AI initiatives.[1]

This is not a model problem or a prompt engineering problem. It is a fundamental architecture problem, and it is becoming the defining obstacle between AI experimentation and AI at scale.

Most enterprise AI stacks are assembled from parts. A vector database here, a relational database and a document store there, orchestration on top, agent memory managed externally, and governance bolted on after the fact. Each component made sense as a standalone choice, but as AI moves from controlled pilot to mission-critical workload, the gaps between those components become a source of failure. When an AI agent needs to retrieve context, query structured data, look up a vector embedding, and validate a user's permissions, each of those steps crosses a network boundary into a separate system, with latency compounding at every hop. In an agentic workflow, reasoning

[1] Source: Enterprise Strategy Group (now Omdia) Research Report, *AI Agents: The Game-changing Generative AI Use Case*, August 2025.

across dozens of steps, this cascades possibly into system-level failure. What felt fast in testing becomes sluggish, unreliable, and expensive in production.

Beyond latency, fragmentation creates governance blind spots. When a vector store has no awareness of data classification policies, and an orchestration layer does not enforce access controls, there is a compliance gap that no amount of application-level patching can fully close. In addition, as AI agents become embedded in operational apps that run 24/7, such as financial and supply chain management systems, failures that cause downtime are unacceptable.

The case for a converged data architecture

A converged data architecture, such as Oracle AI Database 26ai and Autonomous AI Database, addresses this challenge at its root. Rather than connecting disparate systems through pipelines that introduce latency and consistency risk, a converged data platform brings all data types (i.e., structured, unstructured, JSON documents, and now vectors, as well as all development styles) into a single unified engine with common governance, security, access control, and availability. Nothing has to cross a network boundary to be joined together because it was never separate in the first place.

55% of organizations reported data quality and consistency as a top challenge for their data and analytics efforts.[2]

Vectorization, increasingly the backbone of AI's ability to work with unstructured content, belongs inside this converged model. When vectors are stored alongside the data they describe and use, with the same indexing, security policies, and query engine, retrieval becomes dramatically faster and more coherent. An application can query a customer record and its semantic embedding in a single operation, with a single set of permissions enforced in one place. There is no need to maintain consistency between business data stored in one system and embeddings stored in another.

The emergence of agentic AI makes this even more pressing. Most agents are stateful. They need to remember context across steps, access data dynamically, and take actions with real business consequences. Agent memory that lives outside the database introduces the same problems as any other external integration: latency, consistency risk, and governance gaps. When agent memory, vector search, structured query, and AI orchestration all live within a single converged platform, agents become faster, safer, highly available, and far easier to govern.

Mission-critical is no longer optional

What has changed for the AI era is not the definition of "mission-critical." It has always been about eliminating downtime and preventing data loss. What has changed is who is affected when those standards are not met.

[2] Source: Omdia Research Report, *Optimizing Cloud Analytics Costs in an Agentic AI Future*, October 2025.

When an AI agent is handling customer interactions, routing transactions, or driving automated business decisions, downtime is not a performance inconvenience. It is a business continuity problem. Failures in agentic workflows cascade in ways that are harder to detect and roll back than failures in traditional applications. A dropped database connection in a conventional app returns an error. A dropped connection in an agentic workflow can leave a partially executed multi-step process in an indeterminate state, with consequences that propagate across systems before anyone realizes something went wrong.

Modern workloads add further pressure. Agentic AI requires sub-millisecond data retrieval and massive concurrency. Increasingly sophisticated ransomware threats require database-integrated data protection, not perimeter defenses that attackers have learned to route around. And rapidly changing international regulations, including data residency requirements and operational resilience mandates such as DORA, require a distributed, highly scalable, and highly available architecture that can comply with data sovereignty requirements across regions.

Oracle's approach: Diamond-grade reliability built for agentic workloads

Oracle has been building toward this intersection longer than most. Its Autonomous AI Database, now eight years into production and managing 90 billion queries per hour, underpins some of the world's most demanding enterprise applications. That operational track record matters when evaluating a platform for AI workloads that cannot tolerate failure.

> **"Oracle's mission-critical AI database capabilities deliver the always-on availability, data protection, and governance that agentic workloads demand to operate safely in production."**
>
> - Stephen Catanzano, Principal Analyst, Omdia

Oracle's Maximum Availability Architecture (MAA) has long been the reference standard for mission-critical database deployments. With the latest Oracle AI Database release, Oracle has introduced diamond-grade availability, a new tier that combines Exadata, Real Application Clusters (RAC), Active Data Guard with GoldenGate replication or Globally Distributed AI Database with Raft replication, and Zero Data Loss Recovery solutions. Diamond-tier MAA is specifically engineered for ultra-critical workloads that require local high-availability recovery in seconds and regional disaster recovery in under ten seconds with zero data loss.

Application Continuity, another MAA capability, hides outages from users entirely by automatically replaying in-flight work across failures, again without requiring any changes to application code. With Oracle AI Database 26ai, this capability delivers faster query failover and lower database CPU usage, ensuring that the agentic workflows running on top of the database do not surface the infrastructure failures beneath them.

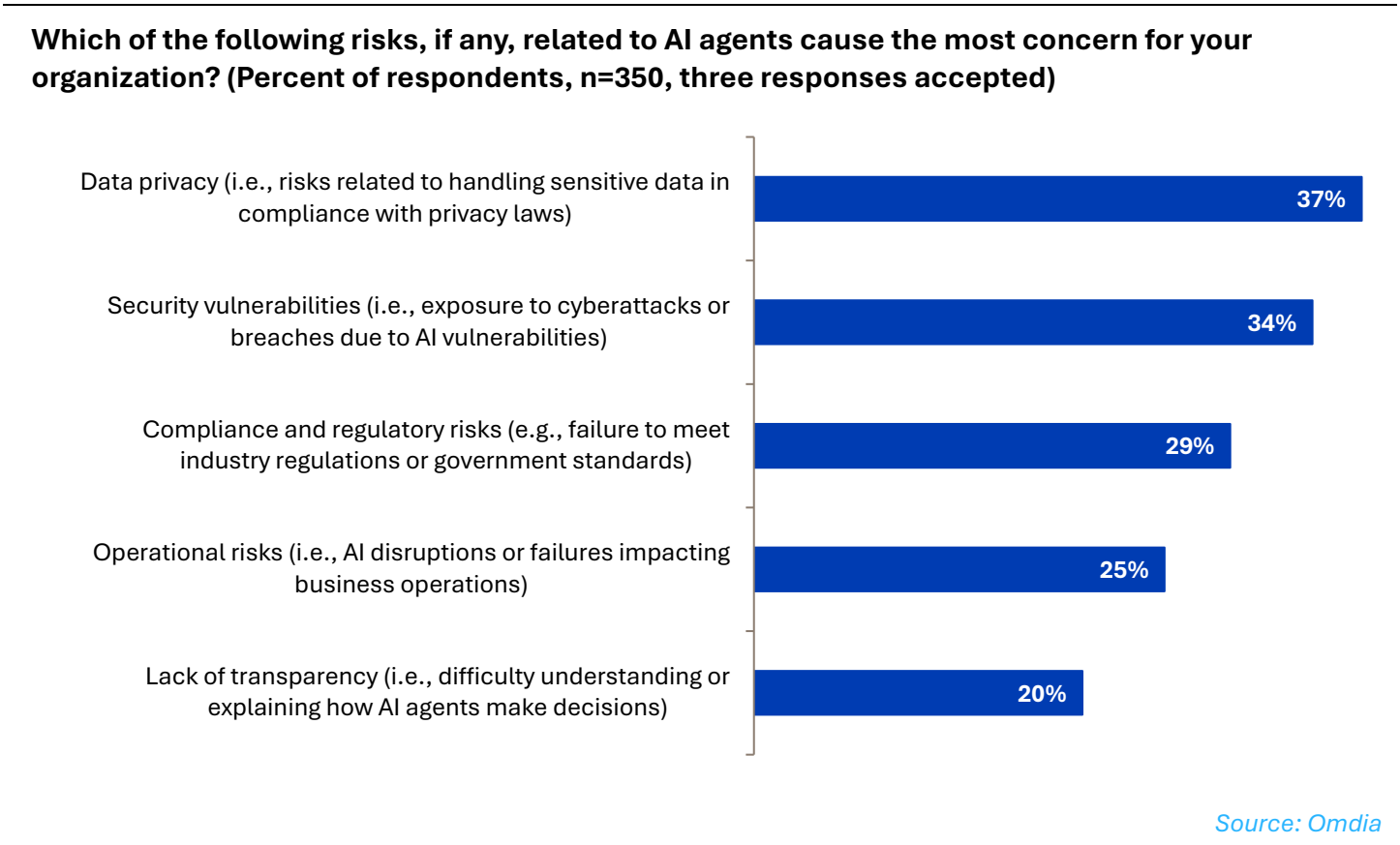
In addition to availability, agents need a high-speed infrastructure to reason and carry out actions in real time. Oracle does not disappoint. The performance numbers behind this architecture are significant. Exadata delivers I/O latency as low as 14 microseconds—up to 70 times faster than competitors for 8KB operations—and can accelerate AI vector query throughput by up to 30 times by offloading vector search to Exadata's

intelligent storage layer, according to Oracle data. Oracle True Cache extends this by serving read queries from an in-memory cache close to the application without requiring any changes to the application itself. Cached data is automatically synchronized with the primary database, delivering several millisecond response times going back to the primary database. In addition, reads of data from a local True Cache instance can complete significantly faster than being delayed by hundreds of milliseconds of network-induced latency when the primary database is in a remote location.

Security that goes beyond the perimeter

Oracle's security approach is organized around what it calls a Risk 360 model, which simultaneously addresses six types of threats: AI-powered attacks, deployment vulnerabilities, configuration drift, user privilege abuse, sensitive data exposure, and application-level risks. This matters because AI-driven attacks do not stay in one lane. Ransomware has grown more sophisticated, and AI is accelerating both the scale and the targeting precision of those attacks. Our research shows the top three concerns for organizations relating to AI agents are data privacy (37%), security vulnerabilities (34%), and compliance and regulatory risk (29%).[3]

Figure 1. Organizations' concerns surrounding AI agents



Source: Omdia

[3] Source: Enterprise Strategy Group (now Omdia) Research Report, *AI Agents: The Game-changing Generative AI Use Case*, August 2025.

Oracle's response combines tight access controls, advanced encryption designed to slow down quantum-computing-based attacks, and database-native protection and mitigation against ransomware.

In the age of agentic AI, it is critical that both agents and the users they represent cannot access data they are not supposed to. Since agentic workflows change rapidly, it is nearly impossible to implement application-level access controls in a timely manner. Organizations that want to rapidly leverage agentic AI and keep their data secure need security implemented at the data level, independent of the user, agent, or application that wants to access it. Oracle has addressed this need in the agentic AI era through what it calls Deep Data Security, which enables agents to access only the same data as the users on whose behalf they are acting, at the row, column, or cell level. This simplifies data security because access rules are implemented once, not for every application or agent.

Oracle's security approach also protects against "harvest now, decrypt later" threats posed by quantum computing. By introducing NIST-compliant quantum-resistant cryptography using TLS 1.3 with ML-KEM and default AES-256 encryption for data at rest, Oracle becomes one of the first enterprise database platforms to address this emerging threat, making it more difficult to decrypt data that is stolen today with future quantum computers.

Additionally, Oracle's efforts to improve data security include database-native protection and automated recovery of Oracle data with the Zero Data Loss Recovery Appliance and Zero Data Loss Autonomous Recovery Service. These platforms and services provide real-time protection against database changes with a near-zero recovery point objective, database-aware validation of backups, and rapid recovery to a clean-room environment after an attack, helping mitigate any attacks that do occur.

Multicloud to support deployment flexibility

Oracle AI Database is available across all leading cloud environments, including Oracle Cloud, Microsoft Azure, AWS, and Google Cloud, with 200 or more regions live or planned across the four platforms. This deployment model enables organizations to conveniently pay for services using existing cloud vendor credits, operate with sub-millisecond latency in their preferred environment, and benefit from full Exadata Cloud automation without compromising on security, reliability, or performance.

Customer adoption of the Oracle AI Database Mission-Critical portfolio

The following organizations have adopted multiple Oracle AI Database mission-critical technologies:

- **Thomson Reuters (NA, financial services):** This customer had the following to say about the portfolio: "Oracle Autonomous AI Database provides autoscaling, self-tuning, self-repairing, and self-encryption so that we have zero downtime to securely process critical financial transactions while we focus on delivering new capabilities to our customers."
- **Renesas (Japan/EMEA, industrial manufacturing):** A global leader of semiconductors, Renesas achieves 24/7 global availability and low-latency access to applications anywhere in the world with Exadata Database Service on OCI.

- **Children's Hospital of Los Angeles (NA, healthcare):** This leading children's hospital modernizes back-office PeopleSoft applications with OCI's Autonomous AI Database for higher availability and reports 98% faster response times.
- **NHS (EMEA, healthcare):** England's services provider for the health system sees better insights with Exadata Database on OCI and saves hundreds of thousands of pounds sterling per year with cloud automation and data integration for greater agility.
- **SEFE (EMEA, energy):** This customer had the following to say about Oracle's portfolio: "Operating in Europe's dynamic energy market, we can't afford downtime or compromise on performance. Oracle AI Database@Azure delivers the cloud resilience, reliability, and speed we need while allowing us to maximize the value of our Microsoft Azure investment. We saw business-process performance improvements of 20% to 45%, exceeding our ROI projections."
- **Vodafone (EMEA, telecom):** OCI Dedicated Region gives the carrier the cost, security, performance, uptime, scalability, and agility benefits of OCI across thousands of Oracle databases and related applications within its own data centers and on its own network. Additionally, Vodafone can confidently select the best cloud for the application and business data and run workloads natively on Oracle Autonomous AI Database, Oracle Exadata Database Service, or Oracle Base Database Service from Azure data centers.
- **NRI (Japan, financial services):** Nomura Research Institute (NRI) is modernizing its mission-critical back-office solutions for brokerage firms with OCI Dedicated Region to build a secure 24/7 platform. NRI said the following: "With Exadata Database Service's CPU scaling, we have managed resources efficiently and have seen a 60% performance improvement."
- **B3 (Brazil, financial services):** Brazil's financial market infrastructure leader uses Exadata Database Service and OCI Kubernetes Engine to bring the São Paulo stock exchange to the cloud.

Conclusion

The organizations building durable AI foundations today are the ones that will have a meaningful advantage when the current wave of AI experimentation settles into production reality. The organizations that do not address the architecture problem now will be refactoring under pressure at the worst possible time.

Omdia recommends that enterprise data and AI leaders evaluate Oracle AI Database as a strategic infrastructure platform for securely running mission-critical agentic AI, transaction processing, analytics, and mixed workloads, with particular focus on MAA Diamond-grade resilience, Deep Data Security for data layer protection against unauthorized agent access to private business data, and the Zero Data Loss Recovery portfolio for mitigating ransomware risks. Further, since agent memory, vector search, structured query, and AI orchestration all live within Oracle AI Database's converged architecture, agents can become faster, more secure, highly available, and far easier to govern. These days, the infrastructure conversation is just as consequential as the model conversation, and, for agentic workloads, it needs to come first. Oracle is leading the discussion.

Get in touch: www.omdia.com askananalyst@omdia.com