



“Hexaware delivers a full-spectrum GenAI development platform that combines agentic architecture, observability tooling and hybrid deployment flexibility to enable scalable, secure and production-grade AI solutions across enterprise environments.”

Gowtham Kumar Sampath

Hexaware

Overview

Hexaware is headquartered in Mumbai, India. It has more than 31,500 employees across 28 countries. In FY24, the company generated \$1.4 billion in revenue, with Financial Services as its largest segment. Hexaware's GenAI delivery model integrates multiagent architecture with containerized platform elements across the RapidX, Tensai and AgentVerse frameworks. Its development approach follows an eight-stage lifecycle that includes guardrail design, context configuration, model selection, testing and orchestration. The architectures support hybrid LLM deployment, tool abstraction and reusable agent composition.

Strengths

Multimodal integration across agent stack:

Hexaware's deployments extend beyond text-centric prompts, embedding multimodal workflows through agents interacting with vision, image, document and tabular data. OCR agents using Vision LLaMA 3.2 and markdown generators allow extraction from diverse file formats, while ReasonAI agents integrate with structured databases for graph-based reasoning. Task-specific token optimization agents further enhance performance by reducing costs and maintaining fidelity across modalities.

Hybrid LLM flexibility and deployment modes:

Hexaware supports cloud-native, on-premises and hybrid LLM deployment through its modular LLM Hub. Agents interface with models from Cohere,

Claude, Azure OpenAI and LLaMA 3.2, with fallback orchestration strategies built into the deployment logic. This setup assists compliance-focused clients with local inference for fine-tuned models while supporting scalable token operations in public clouds.

Agent-oriented architecture: Hexaware's architecture centers on a multiagent execution framework that includes persona-based, reasoning-based and system-level agents operating under real-time orchestration. The AgentVerse platform leverages containerized microservices for each agent, supported by registries that manage agent lifecycle, tool access and telemetry feedback.

Caution

Hexaware should enhance its tooling for prompt versioning, rollback and performance monitoring to improve traceability and consistency across use cases.

While its orchestration logic is robust, codifying agent deployment into structured CI/CD pipelines would advance automation maturity and reduce operational friction.