6 reasons to choose Java for moving AI into production

Moving AI into production is harder than it looks. The language, framework, and ecosystem that teams choose will shape how security-focused, scalable, and cost-efficient their AI applications are for years to come. This checklist outlines 6 reasons why Java on Red Hat® OpenShift® AI is the right foundation for enterprise AI applications—from supply chain security to performance and more—and why most organizations already have the skills and infrastructure to get started.

1 Supply chain security is a real and growing risk

Python tends to dominate AI development, especially during early stages, but production environments are an entirely different challenge. PyPI—Python's primary package registry—has become an actively targeted surface in open source security, with AI-specific frameworks and software development kits (SDKs) increasingly being compromised. Choosing a language for enterprise AI production means choosing its entire ecosystem, including its security model.

The Red Hat approach to open source distribution is designed for exactly this challenge. Curated Maven repositories, service level agreement (SLA)-backed Common Vulnerabilities and Exposures (CVE) patching, and artifact signing give Java development teams a dependency stack they can rely on and a vendor accountable for securing it.

The question for architects evaluating AI platforms should no longer be whether supply chain risk is real, but whether their current ecosystem gives them any meaningful way to manage it.

2 The Java AI ecosystem is ready

In the past, the misconception that the Java AI ecosystem does not offer what organizations need to navigate the complexities of AI prevented some from adopting it into their AI application layer. Yet the Java AI ecosystem is now heavily equipped to deliver scalable AI solutions that are optimized for the enterprise and integrated into advanced orchestration tools and robust infrastructure. This allows organizations to:

- Compose and integrate large language models (LLMs) using LangChain4j and its native Quarkus extensions, and with full model neutrality, swap foundation models via configuration rather than code.
- Connect data pipelines using Apache Kafka, Flink, or Camel—which many Java enterprises already use.
- Accelerate application development using Quarkus AI extensions—a growing catalog of integrations for vector stores, embedding models, and AI services that reduce complexity and help teams focus on business value.
- Benefit from standardized support for the Model Context Protocol (MCP) and the Agent2Agent (A2A) protocol. This allows Java AI applications to dynamically discover tools, interact with external systems, and support multiagent interoperability across different platforms and frameworks.

3 Java is production-grade by default

Quarkus is a cloud-native Java framework purpose-built for container-native environments, and Red Hat build of Quarkus is available to customers. Quarkus brings production capabilities Java teams need into the development workflow. While most frameworks treat those capabilities as optional, Quarkus treats them as the starting point.

Kubernetes-native by design, Quarkus reduces memory consumption and accelerates startup—delivering efficiency gains to push against rising memory costs.[1] Native integration with OpenTelemetry and Micrometer, alongside distributed tracing, offers built-in observability. Governance features including audit trails, compliance controls, and content guardrails are standard—giving regulated industries a production-ready AI platform from Day 1.

4 Developers can use existing skills and infrastructure

The most overlooked AI advantage enterprises using Java have is what they already own. The middleware stack running many businesses today—including Red Hat JBoss Enterprise Application Platform, Red Hat Application Foundations, Apache Camel, Apache Kafka, and OpenJDK—is not a liability. It is the foundation of an enterprise AI application platform, ready to be extended.

Java developers who understand the language, along with its frameworks, and operational practices, are already within many organizations. Adding AI capabilities in Java means those developers can contribute immediately, with no retraining, no ramp-up, and no parallel team required. The expertise teams have built is their biggest advantage in AI adoption.

5 Java offers proven performance efficiency

AI has fundamentally changed the infrastructure cost equation. Increased demand for AI-optimized hardware has driven up the price of memory market-wide, and this is unlikely to be a temporary surge. In production environments, the memory footprint of AI application runtimes is no longer an engineering detail. It is a budget line. Every gigabyte recovered translates to spend avoided—captured inside the Java estate organizations already run.

Quarkus is purpose-built to use less of the resource that just got expensive. Compared to Spring Boot running the same workload, Quarkus delivers:[1]

- A 50% reduction in the memory footprint (269 MB vs. 583 MB).
- 2.7 times more transactions per second.
- 2.3 times faster startup.

6 Modernize with AI-assisted tooling, migration, and more

Java developers do not need to learn a new language—but they do still need onboarding and learning resources. The tools to get started are available today, including the Quarkus Spring Boot Migration Skill, which guides an AI agent through a complete migration and works directly inside tools such as Claude Code and Cursor. For enterprise modernization at scale, Red Hat's migration toolkit for applications provides a fully supported, integrated development environment (IDE)-native path.

Learn more about Java for AI

Discover more benefits of using Java for AI application development.

Try it today

Explore Quarkus and LangChain4j through the Quarkus community and learn how to get started.

1 Cummins, Holly. "Quarkus has great performance – and we have new evidence." Quarkus, 2 March 2026.

About Red Hat

Red Hat helps customers standardize across environments, develop cloud-native applications, and integrate, automate, secure, and manage complex environments with award-winning support, training, and consulting services.

North America	Europe, Middle East, and Africa	Asia Pacific	Latin America
1 888 REDHAT1 www.redhat.com	00800 7334 2835 europe@redhat.com	+65 6490 4200 apac@redhat.com	+54 11 4329 7300 info-latam@redhat.com

facebook.com/redhat
x.com/RedHat
linkedin.com/company/red-hat

redhat.com