
Advanced Generative AI Development on AWS

Duration: 3 Days Course Code: GK910033 Delivery Method: Company Event

Overview:

Master the implementation of production-ready generative AI solutions on AWS.

This advanced 3-day instructor-led training builds expertise across the entire generative AI stack - from foundation models to enterprise integration patterns. In addition, you will learn about advanced data processing techniques, vector database implementation and retrieval augmentation, sophisticated prompt engineering and governance, agentic AI systems and tool integration, AI safety and security measures, performance optimization and cost management strategies, comprehensive monitoring and observability solutions, testing and validation frameworks.

The course structure follows AWS's proven model for generative AI adoption, progressing from experimentation to production-ready implementations.

- **Course level:** Advanced
- **Duration:** 3 day

Updated August 2026

Company Events

These events can be delivered exclusively for your company at our locations or yours, specifically for your delegates and your needs. The Company Events can be tailored or standard course deliveries.

Target Audience:

This course is intended for those with:

- Software developers
 - Technical Professionals
-

Objectives:

■ **In this course, you will learn to do the following:**

- Develop production-ready generative AI solutions using AWS services that meet enterprise requirements for security, scalability, and reliability
 - Evaluate and select appropriate foundation models for specific business use cases, including benchmarking performance and implementing dynamic model selection architectures
 - Design and implement resilient foundation model systems with circuit breakers, cross-region deployment, and graceful degradation strategies
 - Build comprehensive data processing pipelines for multi-modal inputs, including validation workflows and optimization techniques
 - Implement sophisticated vector database solutions using Amazon Bedrock Knowledge Bases, OpenSearch, and hybrid approaches for effective retrieval augmentation
 - Create and manage advanced prompt engineering frameworks, including chain-of-thought reasoning and enterprise-wide prompt governance systems
 - Develop autonomous AI agents using Amazon Bedrock Agents, implementing complex reasoning patterns and tool integration capabilities
 - Implement comprehensive AI safety and security controls, including content filtering, privacy preservation, and adversarial testing mechanisms
 - Optimize performance and manage costs through token efficiency strategies, batching implementations, and intelligent caching systems
 - Design and implement comprehensive monitoring and observability solutions for foundation model applications
 - Create systematic testing and validation frameworks for continuous quality assurance of AI applications
 - Integrate generative AI solutions within enterprise environments using secure, compliant, and scalable architectural pattern
-

Prerequisites:

We recommend that attendees of this course have:

- AWS Technical Essentials
- Generative AI Essentials on AWS
- 2 or more years of experience building production grade applications on AWS or with open source technologies, general AI/ML or data engineering experience
- 1 year of hands-on experience implementing generative AI solutions

Testing and Certification

None

Follow-on-Courses:

- GK910029 - Machine Learning Engineering on AWS
- GK910036 - Building Agentic AI with Amazon Bedrock AgentCore

Content:

Day 1

Module 1: Foundation Model Selection and Configuration

- Enterprise foundation model evaluation framework
- Dynamic model selection architecture patterns
- Resilient foundation model system designs
- Cost optimization and economic modeling

Module 2: Advanced Data Processing for Foundation Models

- Comprehensive data validation and quality assurance
- Multi-modal data processing pipelines
- Input optimization and performance enhancement

Module 3: Vector Databases and Retrieval Augmentation

- Enterprise vector database architecture
- Advanced document processing and chunking strategies
- Sophisticated retrieval system implementation
- Hands-on Lab: Develop Retrieval Augmented Generation (RAG) Applications with Amazon Bedrock Knowledge Bases

Day 2

Module 4: Prompt Engineering and Governance

- Advanced prompt engineering frameworks
- Complex prompt orchestration systems
- Enterprise prompt governance and management
- Hands-on Lab: Develop conversation pattern with Amazon Bedrock APIs

Module 5: Implementing Agentic AI Frameworks with Amazon Bedrock AgentCore

- Agentic AI Frameworks
- Amazon Bedrock AgentCore

Module 6: AI Safety and Security

- Comprehensive content safety implementation
- Privacy-preserving AI architecture
- AI governance and compliance frameworks

Day 3

Module 7: Performance Optimization and Cost Management

- Token efficiency and cost optimization
- High-performance system architecture
- Intelligent caching systems implementation
- Hands-on Lab: Building Secure and Responsible Gen AI with Guardrails for Amazon Bedrock

Module 8: Monitoring and Observability for Generative AI

- Foundation model monitoring systems
- Business impact and value management
- AI-specific troubleshooting and diagnostics

Module 9: Testing, Validation, and Continuous Improvement

- Comprehensive AI evaluation frameworks
- Quality assurance and continuous improvement
- RAG system evaluation and optimization

Module 10: Enterprise Integration Patterns

- Enterprise connectivity and integration architecture
- Secure access and identity management
- Cross-environment and hybrid deployments

Module 11: Course wrap-up

- Next steps and additional resources
- Course summary

Additional Information:

Activities:

This course includes presentations, hands-on labs, demonstrations, and group exercises.

Further Information:

For More information, or to book your course, please call us on 00 971 4 446 4987

training@globalknowledge.ae

www.globalknowledge.com/en-ae/

Global Knowledge, Dubai Knowledge Village, Block 2A, First Floor, Office F68, Dubai, UAE