

Associate Director, AI Foundations Engineering (3 Openings)

Job ID

REQ-10086336

Sep 30, 2026

LOC_US

About the Role

Key Responsibilities:

- Lead the architecture, design, and engineering delivery of the organization's agentic AI platform and related solutions.
- Drive the transformation of existing products, processes, and workflows into scalable agentic AI-enabled capabilities.
- Oversee end-to-end AI engineering operations, including development, deployment, monitoring, reliability, and continuous improvement.
- Partner with cross-functional stakeholders to define technical roadmaps, platform standards, and solution priorities aligned to business goals.
- Establish engineering best practices for AI system performance, scalability, security, governance, and maintainability.
- Guide the evaluation, integration, and optimization of AI models, orchestration frameworks, and supporting platform components.
- Build, mentor, and lead high-performing engineering teams while fostering innovation, accountability, and technical excellence.
- Identify opportunities to accelerate value delivery through reusable AI services, automation, and operational efficiencies.

Essential Requirements:

- Bachelor's degree in computer science, engineering, or a related field, or equivalent practical experience. An advanced degree is a plus but not required.
- 5+ years of professional software engineering experience, including at least 1–2 years building applications on large language models, with at least one agentic system deployed beyond a demo.
- Hands-on experience building AI agents on AWS. This includes Amazon Bedrock (models, Agents, Knowledge Bases, Guardrails) and Bedrock AgentCore or comparable approaches for agent runtime, memory, tool access, and identity.
- Proficiency with at least one agent framework, such as Strands Agents, LangGraph, LlamaIndex, CrewAI, or a custom orchestration layer. You should be able to explain when a simple workflow beats a multi-agent design.
- Strong Python skills plus experience with AWS services commonly used in agent architectures: Lambda, Step Functions, ECS/EKS, API Gateway, DynamoDB, S3, and IAM.
- Experience with retrieval-augmented generation (RAG), including chunking strategies, embeddings, vector stores (OpenSearch, pgvector, or similar), and improving retrieval quality.
- Experience designing tool use and function calling so agents can connect to APIs, databases, and enterprise systems. Familiarity with Model Context Protocol (MCP) is preferred.
- A practical approach to evaluation and observability: test sets for agent behavior, LLM-as-judge or human review loops, tracing (CloudWatch, OpenTelemetry, LangSmith, or similar), and monitoring cost and latency.
- Ability to take an ambiguous business problem to a working prototype quickly, then harden it for real users with security, guardrails, error handling, and CI/CD using CDK or Terraform.
- Clear communication of technical tradeoffs to product and business partners, and comfort working in a small, fast-moving team.

Novartis Compensation Summary:

The salary for this position is expected to range between \$145,600 and \$270,400 per year.

The final salary offered is determined based on factors like, but not limited to, relevant skills and experience, and upon joining Novartis will be reviewed periodically. Novartis may change the published salary range based on company and market factors.

Your compensation will include a performance-based cash incentive and, depending on the level of the role, eligibility to be considered for annual equity awards.

US-based eligible employees will receive a comprehensive benefits package that includes health, life and disability benefits, a 401(k) with company contribution and match, and a variety of other benefits. In addition, employees are eligible for a generous time off package including vacation, personal days, holidays and other leaves.

Role Requirements

Division

DIV_IU

Business Unit

General Management

Location

LOC_US

Site

Remote Position (USA)

Company / Legal Entity

U014 (FCRS = US014) Novartis Pharmaceuticals Corporation

Functional Area

FCT_MM

Job Type

Full time

Employment Type

Regular

Shift Work

No

[Apply to Job](#)

Job ID

REQ-10086336

Associate Director, AI Foundations Engineering (3 Openings)

[Apply to Job](#)

Source URL: <https://prod1.jobapi.novartis.com.cn/req-10086336-associate-director-ai-foundations-engineering-3-openings>

List of links present in page

1. <https://prod1.jobapi.novartis.com.cn/req-10086336-associate-director-ai-foundations-engineering-3-openings>

2. https://novartis.wd3.myworkdayjobs.com/en-US/Novartis_Careers/job/Remote-Position-USA/Associate-Director--AI-Foundations-Engineering--3-Openings-_REQ-10086336-1
3. https://novartis.wd3.myworkdayjobs.com/en-US/Novartis_Careers/job/Remote-Position-USA/Associate-Director--AI-Foundations-Engineering--3-Openings-_REQ-10086336-1