NOVARTIS

Data Engineer

Job ID
REQ-10093347

十一月 24, 2026

India

摘要

#LI-Hybrid

Location: Hyderabad, India

About The Role:
This role is an opportunity to build the trusted, scalable, and AI-ready data foundations required to deliver Corporate Affairs insights and AI solutions. The Data Engineer will design, develop, and optimize enterprise data platforms and pipelines that ingest, transform, curate, govern, and serve structured, semi-structured, and unstructured data. Working closely with AI Engineering, Product Management, Architecture, Analytics, and Corporate Affairs stakeholders, the role will make high-quality data available for reporting, semantic search, Retrieval-Augmented Generation (RAG), and other data- and AI-driven use cases.

The Data Engineer will be accountable for the end-to-end engineering and production performance of assigned data products and pipelines. Success will be measured by reliable delivery, data quality, scalability, security, governance, maintainability, and the ability to move data from source systems into trusted, reusable consumption layers.

About the Role

Key Responsibilities:

- **Design, build, test, deploy, and support scalable enterprise data pipelines** for structured, semi-structured, and unstructured data, covering ingestion, transformation, validation, orchestration, monitoring, lineage, error handling, and both batch and streaming ETL/ELT workloads.
- **Architect and implement modern lakehouse or data warehouse solutions** using Databricks and/or Snowflake integrated with AWS services, applying Medallion Architecture (Bronze, Silver, and Gold) or equivalent layered patterns for raw ingestion, validation, enrichment, curation, and business consumption.
- **Develop data engineering solutions** using Python, Pandas, DataFrames, SQL, notebooks, and distributed data-processing frameworks such as Apache Spark or PySpark, alongside AWS-based data solutions using Amazon S3, AWS Glue, Amazon Redshift, and other relevant AWS services.
- **Process, integrate, and normalize heterogeneous data formats and sources**, including relational data, JSON, XML, APIs, documents, text, logs, and other enterprise content.
- **Implement comprehensive data quality and governance controls**, covering completeness, accuracy, consistency, validity, uniqueness, timeliness, reconciliation, anomaly detection, metadata, lineage, cataloging, ownership, classification, access control, privacy, retention, and auditability.
- **Optimize data pipelines and platforms** for performance, availability, scalability, maintainability, and cloud cost, while creating reusable frameworks, components, standards, and data products that support analytics, Power BI/Tableau, AI, RAG, and semantic-search use cases.
- **Where appropriate, integrate AI services and Agentic AI capabilities**, including Amazon Bedrock, into data pipelines for classification, extraction, enrichment, summarization, entity identification, and metadata generation.
- **Where applicable, enable graph, knowledge-graph, vector database, and semantic-retrieval solutions** through entity extraction, entity resolution, relationship modeling, graph-oriented data preparation, parsing, chunking, embedding generation, metadata enrichment, indexing, and lifecycle management.
- **Ensure engineering and production excellence** through testing, CI/CD, documentation, monitoring, production support, and disciplined release management for data solutions.
- **Demonstrate independence, accountability, strong problem solving, and continuous learning**, proactively managing technical risks, dependencies, and production issues.

Essential Requirements:

- Bachelor's degree in Computer Science, Data Engineering, Information Technology, Engineering, Data Science, or a related field, or equivalent practical experience.
- 8-12 years of overall experience in data engineering, software engineering, data platforms, or related technology roles, including substantial hands-on experience delivering production data solutions.
- Strong hands-on experience with Databricks and/or Snowflake in enterprise environments.

- Proven experience implementing Medallion Architecture (Bronze, Silver, and Gold) or a comparable layered lakehouse/data-platform architecture.
- Strong experience designing and building enterprise ETL/ELT pipelines, including ingestion, transformation, orchestration, validation, monitoring, recovery, and production support.
- Strong programming skills in Python, including practical experience with Pandas, DataFrames, and notebook-based development.
- Strong SQL skills and experience with data modeling, schema design, transformation, performance tuning, and query optimization.
- Hands-on experience with AWS data services, particularly Amazon S3, AWS Glue, and Amazon Redshift.
- Proven experience processing structured, semi-structured, and unstructured data from heterogeneous enterprise sources.
- Strong practical knowledge of data governance and data quality, including metadata, lineage, cataloging, validation, reconciliation, security, and access control.
- Experience with Git, automated testing, CI/CD, data observability, monitoring, and operational support for production data pipelines.
- Ability to collaborate with AI Engineers, analysts, architects, product managers, BI developers, security teams, and business stakeholders.

Desirable Requirements:

- Hands-on experience with both Databricks and Snowflake, including workload optimization, governance, and secure data sharing.
- Relevant certifications such as Databricks Certified Data Engineer Associate or Professional, SnowPro Core or Advanced, or AWS Certified Data Engineer - Associate.
- Experience with Amazon Bedrock or other AI/LLM services embedded in data pipelines for document processing, classification, extraction, summarization, enrichment, or metadata generation.
- Knowledge of Agentic AI patterns and the use of agents or AI-assisted workflows within enterprise data-processing solutions.
- Experience with vector databases, embeddings, chunking, vector search, hybrid retrieval, semantic search, and data preparation for Retrieval-Augmented Generation.
- Experience with graph databases such as Neo4j, knowledge graphs, entity resolution, relationship modeling, and graph-based retrieval.
- Experience developing dashboards or analytical products using Microsoft Power BI or Tableau.
- Experience with streaming platforms, orchestration tools, data catalogs, master/reference data, and infrastructure-as-code.

Why Novartis: Helping people with disease and their families takes more than innovative science. It takes a community of smart, passionate people like you. Collaborating, supporting and inspiring each other. Combining to achieve breakthroughs that change patients ' lives. Ready to create a brighter future together? https://www.novartis.com/about/strategy/people-and-culture

Benefits and Rewards: Learn about all the ways we ' ll help you thrive personally and professionally.
Read our handbook (PDF 30 MB)

部门
Corporate Affairs

Business Unit
Communications

地点
India

站点
Hyderabad (Office)

Company / Legal Entity
IN10 (FCRS = IN010) Novartis Healthcare Private Limited

Functional Area
Data and Digital

Job Type
Full time

Employment Type
Regular

Shift Work
No

var kPlayer = KalturaPlayer55802022 || KalturaPlayer; var config = { targetId: "kalturaplayer6b05addd4057d799852854", provider: { widgetId: "10m7rm1pm", partnerId: "2076321", uiConfId: "55802022" }, playback: { autoplay: false, autopause: false, muted: false, loop: false }, sources: { options: {}, startTime: 0 }, disableUserCache: "true", plugins: {}, sources: { options:

{}, startTime: 0 }, ui: { showCCButton: false, settings: { showQualityMenu: true, showSpeedMenu: false }, components: { fullscreen: { disableDoubleClick: false } }, uiComponents: [{ presets: ['Playback', 'Live'], area: 'BottomBarRightControls', replaceComponent: 'Fullscreen', get: kPlayer.ui.components.Remove }] } }; // Check and add plugins only if they exist if (kPlayer.plugins["download"]) { config.plugins.download = { disable: true }; } if (kPlayer.plugins["transcript"]) { config.plugins["playkit-js-transcript"] = { position: "right", // Default: bottom;('left', 'right', 'top', 'bottom') to enable transcript. expandMode: "over", // Default: alongside;('alongside', 'hidden', 'over') expandOnFirstPlay: false, showTime: true, downloadDisabled: false, printDisabled: false, disable: true }; } if (kPlayer.plugins["preventSeek"]) { config.plugins.preventSeek = { preventSeekForward: false, preventSeek: false }; } config.plugins.floating = { disable: true }; if (kPlayer.plugins["navigation"]) { config.plugins.navigation = { position: "right", expandMode: "over", expandOnFirstPlay: false, visible: false }; } if (kPlayer.plugins["hotspots"]) { config.plugins['playkit-js-hotspots'] = { disable: true }; } if (kPlayer.plugins["moderation"]) { config.plugins['playkit-js-moderation'] = { disable: true }; } if (kPlayer.plugins["info"]) { config.plugins['playkit-js-info'] = { disable: true }; } if (kPlayer.plugins["share"]) { config.plugins.share = { disable: true }; } config.ui.uiComponents = []; if (kPlayer.plugins["googleAnalytics"]) { config.plugins.googleTagManager = {}; config.plugins.googleTagManager.customEventsTracking = {}; config.plugins.googleTagManager.containerId = 'GTM-57RJQ5'; config.plugins.googleTagManager.customEventsTracking.custom = []; config.plugins.googleTagManager.customEventsTracking = { preset: { coreEvents: true, UIEvents: false, playlistEvents: false, castEvents: false } }; }

// Ensure the global player registry array always exists, regardless of embed type. window.kalturaPlayerVideos = window.kalturaPlayerVideos || []; try { var thumbEmbedPromise = thumbnailEmbed({config, mediaInfo: {entryId: "1dgfvmafo"}}); // thumbnailEmbed() returns a Promise that resolves with the player instance // when the user clicks the thumbnail. Use .then() to capture the player directly. thumbEmbedPromise .then(function(player) { window.kalturaPlayerVideos.push(player); // Notify kaltura datalayer.js that a new player is ready so it can // attach custom event listeners immediately, regardless of when // the user clicked the thumbnail relative to page load. document.dispatchEvent(new CustomEvent('kalturaPlayerReady', { detail: { player: player } })); }) .catch(function(error) { console.error(error); }); } catch (e) { console.error(e.message) }

Accessibility and accommodation

Novartis is committed to working with and providing reasonable accommodation to individuals with disabilities. If, because of a medical condition or disability, you need a reasonable accommodation for any part of the recruitment process, or in order to perform the essential functions of a position, please send an e-mail to diversityandincl.india@novartis.com and let us know the nature of your request and

your contact information. Please include the job requisition number in your message.

Job ID
REQ-10093347

Data Engineer

Apply to Job



Job ID
REQ-10093347

Data Engineer

Apply to Job

Source URL:
https://www.novartis.com.cn/careers/career-search/job/details/req-10093347-data-engineer

List of links present in page

1. https://www.novartis.com/about/strategy/people-and-culture
2. https://www.novartis.com/sites/novartiscom/files/novartis-life-handbook.pdf
3. mailto:diversityandincl.india@novartis.com
4. https://novartis.wd3.myworkdayjobs.com/en-US/NovartisCareers/job/Hyderabad-Office/Data-EngineerREQ-10093347
5. https://novartis.wd3.myworkdayjobs.com/en-US/NovartisCareers/job/Hyderabad-Office/Data-EngineerREQ-10093347