

## Sr Data Engineer - Job Description

---

La posición admite postulaciones de los siguientes países: [Argentina](#) / [Colombia](#)

### Información sobre el área de Cloud Economics

El área de Cloud Economics es la responsable de implementar la [práctica de Cloud FinOps](#) dentro de MELI. En este sentido, su scope es el diseño e implementación de la estrategia de gobierno y optimización del consumo cloud de todo MELI. Éste es un desafío muy interesante dado el volumen y la complejidad que tiene este consumo en nuestra compañía, y generalmente tenemos que pensar en soluciones a preguntas o problemas sobre lo que no hay mucho escrito previamente.

El área está conformada por 5 equipos conformados en torno a un enfoque multidisciplinario, y participan desarrolladores, data engineers, data scientists, analistas de BI y FinOps engineers, entre otros profesionales.

En esta oportunidad estamos haciendo una búsqueda de un Sr Data Engineer, para sumarse a nuestro equipo de Data Engineering.

Este equipo se enfoca en el diseño y desarrollo de arquitecturas de datos, asociadas principalmente al consumo y billing de recursos cloud. Buscamos que estas arquitecturas permitan una explotación eficiente de los datos que contienen, identificando los puntos críticos del proceso y generando un esquema de alerta y comunicación ante determinadas eventualidades, y que atiendan las necesidades de los clientes internos para fortalecer el proceso de toma de decisiones.

Actualmente tenemos nuestros pipelines de datos alojados completamente en ambientes cloud, y son explotados a través de distintas herramientas de BI por diferentes equipos, para quienes la oportunidad e integridad de la información es sumamente relevante. Dicha información es consumida tanto para el entrenamiento de modelos predictivos, como también para alimentar procesos internos de MeLi (por ejemplo, esquemas de alocação de costos o de alerta).

Los datos que manejamos están principalmente asociados al billing generado por el consumo de recursos cloud de distintos proveedores, por lo que trabajamos con distintas fuentes, las que suelen tener diferente granularidad temporal y requieren de diferentes procesos de transformación previo a su almacenamiento (hot y cold storage).

El objetivo del equipo es disponibilizar los datos de forma eficiente y con una mirada de negocio, para fortalecer el proceso de toma de decisiones de los equipos de IT de MeLi.

### ¿Cuáles son los desafíos del rol?

Los desafíos del rol

- Proponer y desarrollar arquitecturas de datos dentro de los conceptos de ciclo de vida del dato y gobierno de datos.
- Diseñar modelos de datos e implementarlos en ambiente cloud (AWS o GCP principalmente).
- Diseñar y desarrollar procesos de ETL y pipelines de datos basados en eventos (event-based data).
- Estar en capacidad de realizar seguimiento a los diferentes pipelines de procesamiento de datos que se ejecutan a diario y, en caso de fallos, efectuar el análisis y contactar al equipo responsable para realizar el ajuste pertinente.
- Desarrollo y mantenimiento de Lake-houses.
- Desarrollo de soluciones de BI/data visualization.
- Interactuar con clientes internos para realizar el desarrollo funcional de las soluciones.

- Participar en la realización de PoCs y MVPs relacionados a proyectos del área
- Conocimiento y experiencia en la implementación de mejores prácticas en el manejo de datos.
- Contribuir en el diseño de pipelines en ambientes cloud, que involucren grandes volúmenes de datos.
- Trabajar con requerimientos de escalabilidad, observabilidad y tolerancia a fallos.
- Interactuar con stakeholders para identificar requerimientos y coordinar el armado de soluciones para dar respuesta a esas necesidades.

## ¿Cuáles son los requisitos?

- En términos de soft skills, sería deseable que los candidatos:
  - Sean personas curiosas y con una fuerte capacidad analítica.
  - Cuenten con buena capacidad de comunicación para transmitir ideas de manera efectiva.
  - Disfruten de desenvolver su trabajo como parte de equipos multidisciplinarios y de generar vínculos con otros equipos (IT, Finanzas y Negocio).
  - Tengan interés por desenvolverse en un ambiente que requiere el continuo aprendizaje y la innovación.
- Respecto de conocimientos técnicos, en términos generales es deseable que los candidatos cuenten con conocimientos en:
  - Arquitecturas de datos, ciclo de vida del dato ,gobierno del dato, linaje.
  - Modelamiento de data lake.
  - Modelado multidimensional (kimball / Inmon).
  - Articulación de herramientas cloud para proponer arquitectura de datos.
  - Conocimientos y experiencia práctica avanzada en SQL. BigQuery es muy deseable.
  - Conocimiento en manipulación, análisis de datos, desarrollo de scripts y automatización con Python.
  - Experiencia en el uso de herramientas de orquestación de flujos de ETL (Airflow, Kubeflow, mlflow, etc.).
  - Familiaridad con herramientas de Reportes dashboards y kpis.
  - Conocimiento en uso de CVS (Git).
  - Cloud Environments (AWS / GCP / Others).
  - Familiaridad con herramientas de monitoreo como Datadog, New Relic, CloudWatch etc.
  - Conocimiento es uso y extracción de datos a partir de API RESTful y arquitectura basada en microservicios.
  - Server Management (Docker / Kubernetes)
  - Asimismo, se valorará que cuenten con conocimientos en:
    - Buenas prácticas de arquitectura (DAMA)
    - DBT
    - Great Expectations
    - CI/CD
    - Conocimientos y experiencia práctica en No-SQL
    - Herramientas de Big Data (Hadoop, Spark, Hive, Kafka)
    - Experiencia práctica con streaming de datos
    - Metodologías ágiles