

Elias Velazquez

Data Engineer AWS

Buenos Aires, Argentina

Email | Website | LinkedIn | GitHub

RESUMEN PROFESIONAL

Ingeniero de Datos especializado en plataformas de datos nativas en la nube, pipelines de alto rendimiento y arquitecturas orientadas a eventos en AWS y Snowflake. Trayectoria comprobada en el escalado de sistemas de producción para manejar **más de 1.5 TB de ingesta mensual** y un **crecimiento de volumen diario de 25x YoY**, manteniendo arquitecturas resilientes y rentables. Experto en la optimización de cargas de trabajo de cómputo intensivo —reduciendo los tiempos de ejecución del **warehouse** en un **85%** mediante `write_pandas`— y en el desacople de cuellos de botella mediante mensajería asíncrona (SQS). Experimentado en desarrollo ETL/ELT con Python, aislamiento de datos **multi-tenant** y gobernanza programática de data warehouses.

HABILIDADES CENTRALES

- **Lenguajes:** Python (Avanzado: OOP, Asyncio/Threading, Type Hinting, Logging Estructurado), SQL (CTEs, Funciones Ventana, Joins), Bash.
- **AWS:** Lambda, S3, ECS/Fargate, SQS, EventBridge, DynamoDB, Step Functions, API Gateway, IAM, CloudWatch.
- **Ingeniería de Datos y Herramientas:** Snowflake, Pandas, Boto3, Docker, Git, Pipelines ETL/ELT, Arquitectura Multi-Tenant, Evolución de Esquemas.
- **Fundamentos de DataOps y DevOps:** CI/CD con GitHub Actions, Terraform (IaC básico), Automatización en Linux.

EXPERIENCIA PROFESIONAL

Data Engineer — Vordentech

Febrero 2024 – Presente

- **Pipeline 1: Motor de Procesamiento de Datos Legacy Multi-Tenant**
 - Diseñé y desplegué una plataforma de procesamiento orientada a eventos y contenerizada (ECS/Fargate, S3, EventBridge) para la ingesta y análisis de archivos financieros legacy de ancho fijo de entre 4 y 5 GB, procesando hasta 5 millones de filas por ejecución.
 - Desarrollé una pipeline en Python/Pandas basada en flujos y fragmentos (**chunks** de 100k filas) para aislar y generar conjuntos de datos multi-tenant destinados hasta a 13 clientes independientes de forma concurrente y sin sobrecarga de memoria.
 - Construí analizadores sintácticos (**parsers**) personalizados basados en posiciones para procesar especificaciones complejas de estilo COBOL, codificaciones de fechas no estándar y numéricos con signo, reduciendo el tiempo de ejecución del servicio legacy en un ~35% mediante la refactorización de código y la optimización estructural.
- **Pipeline 2: Arquitectura de Almacenamiento e Ingesta de Alto Rendimiento**

- **Diseñé un motor en AWS Lambda guiado por esquemas** que infiere la evolución de esquemas y aprovisiona automáticamente tablas en Snowflake, escalando sin fricciones para soportar un **crecimiento de volumen diario de 25x YoY** (de ~2 GB a más de 51 GB/día).
 - **Construí resiliencia en la *pipeline*** para absorber un **incremento mensual (MoM) del 65%**, eclipsando con éxito **1.5 TB de ingesta mensual** mientras se mantuvo una estricta estabilidad y eficiencia de costos.
 - **Reduje los tiempos de ejecución de ETL en un ~85%** (de 15–20 minutos a solo 2–3 minutos) al reemplazar las escrituras nativas de Python por el método `write_pandas` del conector de Snowflake.
 - **Diseñé una tubería de métricas asíncrona** donde la Pipeline 2 descarga los metadatos de ejecución en AWS SQS, desacoplando el procesamiento del almacenamiento para eliminar la saturación de escritura en el *warehouse* (Snowflake) y garantizar inserciones ordenadas bajo tráfico pesado.
 - **Implementé las mejores prácticas de DataOps** (Docker, *logging* estructurado, alertas) y colaboré con partes interesadas (*stakeholders*) globales para alinear la arquitectura con las necesidades de producción.
-

EDUCACIÓN

- Tecnicatura en Ciencia de Datos e Inteligencia Artificial — Universidad Nacional de Chilecito (Mar 2026 – Actualidad)
 - Cursando: estadística, fundamentos de machine learning, análisis de datos y bases matemáticas para ciencia de datos.

CERTIFICACIONES

- **AWS Certified Solutions Architect Associate** (otorgada en Sep 2026 - Expira en Sep 2029)

IDIOMAS

- Inglés: C1
- Español: Nativo