

Ian Laurel Pastene

Senior Data Engineer · Data Tech Lead

AWS Certified

AWS Glue

PySpark

Spark SQL

Banking Data Platforms

Santiago, Chile · ianlaurelpastene@gmail.com · linkedin.com/in/ian-laurel-pastene

PROFESSIONAL SUMMARY

AWS-certified Senior Data Engineer and Data Tech Lead with 6+ years building production pipelines for banking, risk, regulatory compliance, CRM and commercial banking. Hands-on with AWS Glue, PySpark, Spark SQL, S3, Athena, Apache Iceberg and layered Data Lake architectures, with a path from banking BI/ETL and Big Data toward cloud platforms on AWS. Designed a reusable framework adopted by 6 teams that cut development time from 3 days to 1, technically leads a data engineering team and drives modernization with Data Mesh and AI-assisted engineering.

FEATURED CERTIFICATION

AWS Certified Data Engineer — Associate
Amazon Web Services

ACTIVE · 2025 – 2028

TECH STACK

CORE AWS DATA ENGINEERING	AWS Glue, Glue Crawlers, S3 (lifecycle, versioning, replication), Glue Data Catalog, Athena, Athena Workgroups, Lake Formation, IAM, KMS, Lambda, Step Functions, EventBridge, SQS, CloudWatch, DynamoDB
PROCESSING & DATA PLATFORMS	PySpark, Spark SQL, Apache Spark, Apache Iceberg, Scala, SQL, HQL, ETL/ELT, Data Lake, Data Mesh, medallion, batch/intraday/near real-time, incremental loads (partition comparison), partitioning, data contracts
ORCHESTRATION, DEVOPS & GOVERNANCE	Glue Jobs, Step Functions, EventBridge, Control-M, DataStage, CloudFormation, CodePipeline, CodeBuild, CodeCommit, Git, Bitbucket, GitHub, Jenkins, SonarQube, CI/CD, containers, data quality, lineage, DoR/DoD, Jira, Scrum
CLOUD & LEGACY	GCP (BigQuery, Dataflow, Dataproc), Azure Data Lake, Hive, HDFS, SQL Server, IBM DB2, Parquet, ORC, Avro, CSV
AI-ASSISTED ENGINEERING	Amazon Q, Kiro, GenAI for code review, troubleshooting, technical documentation and delivery acceleration

PROFESSIONAL EXPERIENCE

Data Tech Lead — Banco Itaú Feb 2026 – Present
NTT DATA Europe & Latam · Santiago, Chile · Hybrid

Data initiatives for ItaúX Personas, Empresas and CRM: commercial platforms for customers, campaigns, portfolio, collections and analytics.

- Designed a reusable PySpark/Spark SQL framework adopted by 6 teams, cutting development from ~3 days to 1 and standardizing pipelines on AWS Glue, S3 and Athena.
- Delivered ~30 production processes on layered Data Lake architectures, with ~60 partitioned tables for analytics consumption and commercial publication.
- Processed variable-scale loads with validations, monitoring and quality controls, ensuring traceability and consistency for analytics consumption.
- Lead a team of 4 data engineers across three workstreams, defining technical design, standards and production-readiness criteria.
- Delivered domain-based Data Mesh capabilities with Apache Iceberg tables to production as part of the first CRM team to do so: central catalog, CI/CD and governed consumption.
- Introduced AI-assisted workflows (Amazon Q, Kiro) to accelerate code review, troubleshooting and technical documentation.

Technical Lead — Banco Itaú Jul 2025 – Jan 2026
NTT DATA Europe & Latam · Santiago, Chile · Hybrid

Technical workstream of ItaúX Empresas: corporate (PJ Minorista) client view for commercial officers.

- Optimized critical production processes, cutting execution time by up to 80% and improving processing windows and data availability.
- Implemented and deployed pipelines on the SDLF Framework with AWS Glue, PySpark and Spark SQL, automating the deployment of Glue Jobs, triggers and schedules via CloudFormation (YAML) and CodePipeline.
- Structured technical roadmaps by business impact, technical dependencies and production feasibility.
- Led the technical workstream, enabling data capabilities for corporate clients through to production.

Senior Data Engineer — Banco Itaú Mar 2024 – Jun 2025
NTT DATA Europe & Latam · Santiago, Chile · Hybrid

Data core of ItaúX Personas: 360° customer view for executive management, campaigns, credit and collections.

- Built pipelines on AWS Glue, PySpark and Spark SQL, integrating official banking sources into the 360° view of retail customers.
- Implemented an event-driven data transfer architecture with SQS (including DLQ), Lambda and AWS Glue, ensuring decoupled and fault-tolerant ingestion.
- Built batch, intraday, monthly and historical processes on the SDLF Framework with incremental loads by partition comparison, improving data quality and availability for analytics consumption.
- Implemented transformations and validations for customers, products, credit, collections and campaigns, ensuring commercial data consistency.
- Validated production releases together with architecture, governance and QA, reducing rework and inconsistencies.

EXPERIENCE (CONT.)

Site Reliability / Data Platform Engineer

Equifax · Regional LATAM

Regional operation of DataFabric flows across 6 countries (PY, PE, CL, UY, PR, DR).

Built and maintained ingestion plugins with Java, GCP Dataflow, Dataproc and BigQuery for regional DataFabric reporting and publication flows.

Operated and monitored cloud processes across 6 countries, overseeing lineage, publication, availability and data consistency.

Automated testing and deployments with Jenkins, GitHub, Maven and containers, improving release stability.

Resolved production incidents and troubleshoot regional flows, sustaining multi-country operational continuity.

Mar 2023

—

Oct 2023

Data Engineer — Equifax Dominican Republic

23people · Santiago, Chile · Remote

External consulting for Equifax RD on DataFabric flows for loading, preparation and publication of information.

Developed ingestion plugins with Java, DataFabric and GCP, integrating and validating data for reporting and analysis.

Updated schemas and ETL transformations, improving traceability and operational continuity of flows.

Automated testing and deployments with Jenkins, GitHub, Maven and containers, reducing release risk.

Aug 2022

—

Mar 2023

Data Engineer / BI & Big Data Consultant

FactorIT · Santiago, Chile

Data consulting for four banking clients —BCI, Banco de Chile, BCP and Interbank— in BI, ETL, Big Data, Data Lake, risk, compliance and CRM.

Banco de Chile (Risk & Compliance): built Big Data processes with Scala, Spark SQL, Huemul Solutions, Hive, HDFS and Control-M, supporting risk, provisions, compliance and Data Lake.

BCP (Big Data Ingestion): implemented ingestion and ETL on Cloudera/Hadoop with PySpark, HQL, Hive, HDFS, DataStage, Shell and Jenkins, certifying loads under Scrum.

Interbank (Data Platform): developed ingestion and transformation on Hadoop with PySpark and Hive, integrating sources into analytics consumption layers.

BCI (Regulatory Cell): developed regulatory ingestion with internal frameworks WOLF/BDU, Scala, Bitbucket and Jenkins into an Azure bronze/silver/gold architecture.

BCI (CRM & Commercial Management): built ETL with IBM InfoSphere DataStage, SQL Server and business rules into Salesforce for campaigns and customer management.

Delivered development, production support and QA coordination across multi-project banking cells, adapting to changing scope, priorities and clients.

Mar 2019

—

May 2022

EDUCATION

Computer Engineering · Duoc UC

2015

—

2018

CERTIFICATIONS

AWS Certified Data Engineer — Associate · Amazon Web Services

2025

—

2028

COURSES & TRAINING

AWS AI & GenAI: Amazon Q Developer & Kiro Basics · Udemy

Jun 2026

GenAI Academy — Yellow Belt L2 · NTT DATA

Aug 2025

GenAI Academy — Yellow Belt L1 · NTT DATA

Jan 2025

GenAI Academy — White Belt · NTT DATA

Dec 2024

Scrum Master Professional Certificate · CertiProf

May 2022

Path to Excellence in Project Management · Pontificia Universidad Católica de Chile

Mar 2020

LANGUAGES

Spanish — Native

English — Full Professional