

Sujal Dhungana

📍 Perth, Australia ✉ sujald@dataprofessionacademy.com ☎ 0410677503 🌐 in/dhunganasujal 🌐 sujal-dhungana.netlify.app/hire-me

SUMMARY

Seasoned Data Engineer with a passion for building robust, scalable data platforms that stand the test of time. Specialising in exploring the crucial balance between developer velocity and enterprise-grade reliability. Deeply committed to "Everything as Code" and DevOps principles to build self-service tooling, standardised pipelines, and efficient orchestration systems that empower teams while embedding critical data governance, security controls, and maintainability.

SKILLS

Languages: Python, SQL, PySpark, R, C, C++, JavaScript, Bash

Platforms & Data Orchestration: AWS (EKS, ECS, Lambda, Step Functions), Kubernetes, ArgoCD, Helm, Ansible

Data Processing & Warehousing: Python, SQL, AWS Glue, DBT, ClickHouse, Apache Iceberg, AWS MSK(Managed Kafka), AWS Opensearch, DynamoDB, AWS RDS

GitOps & Infrastructure as Code: Cloudformation, CDK, Terraform, CI/CD, Azure DevOps, Gitlab, Artifactory

Monitoring & Visualisation Tools: Prometheus, Grafana, PowerBI, Quicksight

EXPERIENCE

Resident Platform Engineer (Contractor)

Telstra

October 2024 - Present, Perth, WA

- **Optimised data pipelines ingesting over 1TB/hr of log data to Clickhouse and Opensearch**, including separating IO-bound and CPU-bound operations, parallelization of operations to utilise idle CPU cores, coupled with event-based scaling of Kubernetes resources, resulted in a savings of over 50% in compute cost (valued at over \$600,000 per annum) without compromising latency and platform reliability.
- **Designed and built a suite of internal toolings and RESTful APIs** to automate the provisioning and management of K8s applications, ensuring consistency and reducing the time taken to create data pipelines from days to hours.
- **Ensured platform uptime exceeds SLA** through rigorous change management, regular rolling updates, and leading incident response, maintaining high levels of customer satisfaction.

Principal Data Engineer

Data Vanguard Pty Ltd.

April 2024 - Present, Perth, WA

- **Technical Lead:** Architected and led the development of the world's first Warehouse agnostic (Bring your own warehouse like Redshift, Athena, Snowflake, etc.), agentic Data Platform. **Empowering AI agents** with comprehensive, governed access to pipeline metadata, status, and data models to establish data provenance and enable secure, context-aware analytical reports.
- **Developed a fully automated end-to-end MLOPs pipeline** to generate item recommendations for potential bidders to one of the biggest horse auctioneers in Australia.
- **Developed a lightweight high-frequency ingestion pipeline** to seamlessly ingest IoT data into the data warehouse, enabling near real-time analytical capabilities using native AWS APIs, Lambda, SQS, DynamoDB, Glue Catalogue and Redshift.
- **Designed and deployed a Data Ingestion Platform** to standardise data ingestion and introduced Medallion layers following the modern data architecture. Reducing the time taken to onboard new data sources by 60% (as the pipeline between each layer is standardised and reused) while also reducing the maintenance effort.
- **Redesigned and developed an ingestion pipeline on AWS**, significantly enhancing reliability and maintainability while reducing operational costs by 90%, resulting in annual savings of over \$30K in cloud expenses.

Founder | Instructor

Data Profession Academy

August 2023 - Present, Perth, WA

- **Designed and delivered a 12-week data analytics curriculum** to 20+ aspiring professionals, with a majority securing their first industry role within 3 months of completion.
- **Provided mentorship to 20+ students**, both technical (Python, SQL, PowerBI) and non-technical skills (Problem Solving, Data Storytelling), empowering them to become well-rounded data analysts ready to deliver value from day one.

Senior Data Engineer

Water Corporation

May 2022 – April 2024, Perth, WA

- **Created an end-to-end MLOPs pipeline** leveraging AWS native cloud stacks (SageMaker, Code Build, Step Functions, Lambda), significantly reducing the time to production from days to minutes with minimal manual intervention. The MLOPs pipeline enabled seamless cross-account deployment of ML models with a simple click of a button while also providing appropriate governance with manual validations and approval steps in between.
- **Productionised ML model to predict Centrifuge Performance** using real-time operational data, voted one of the top papers in the OzWater conference 2024.
- **Designed and developed an end-to-end ingestion platform** that ingests data from various sources into Apache Iceberg tables in S3, incorporating data quality reporting and dimensional modelling with SCD capabilities out of the box. This innovation cut pipeline build time from 2 days to 2 hours and reduced cloud costs by 50% with a config-driven approach.
- **Developed an automated reporting platform** enabling self-serve analytics teams to build data models with SQL, reducing model build time by 50% and eliminating production delays with automated workflows.
- **Developed various large-scale data ingestion pipelines** for sources (including PI, GIS, SQL servers), and implemented business logic using PySpark in AWS Glue ETL, resulting in a data product that facilitates the management of AUD 2.3 billion worth of assets through objective risk assessment.
- **Automated data validation and quality assurance** with test scripts, reducing test times from days to hours, and ensuring the freshness and accuracy of data in AWS Redshift.

Technology Consultant

Visagio

August 2020 – May 2022, Perth, WA

- **Developed a time and attendance tool** for a tier-one mining company, providing enhanced visibility into actual workforce consumption and future forecasts. This tool empowered supervisors to optimise daily plans, resulting in over \$40 million in savings.
- **Built a strategic workforce planning tool** for the same company, enabling demographic projections over 5- and 10-year horizons. The tool accounts for current demographics and allows scenario modelling, generating realistic and actionable workforce forecasts.
- **Developed and redesigned multiple mission-critical dashboards** using Power BI and Spotfire for internal and regulatory reporting. Restructuring multiple dashboards with best practices in data modelling and transformation significantly increased responsiveness and easier to maintain.

EDUCATION

Master of Predictive Analytics

Minor in Finance and Investment • Curtin University • Bentley, Western Australia • 2020 • 8.0/10.0

Bachelor of Electronics and Communications Engineering

Tribhuvan University • Kathmandu, Nepal • 2017 • 8.0/10.0

INVOLVEMENT

Python Instructor

Nepalese Engineers in WA • July 2021 – August 2021

- Conducted a 5-week-long Python programming masterclass with more than 20 participants. The masterclass covered the basics of Python programming and specific use cases for different engineering fields, including electrical, mechanical and other engineering.

Student Volunteer

John Curtin Leadership Academy • January 2020 – December 2020

- Led a volunteer team to help the Curtin C3 project conduct various activities, including a “Design-A-Thon” conducted in May 2020.
 - Led a team to organise, advertise and coordinate an online event “WA's Sustainable Future” with speakers from Deloitte, Curtin University and CleanState and 50+ participants.
-