

Matthew Nelson

Head of Data & Analytics · AI-native platform builder

Calgary, Canada · matthewpeternelson@gmail.com · +1-403-667-3022 · <https://mnelson.ca> · Github: <https://github.com/matthewpnelson> · LinkedIn: <https://linkedin.com/in/matthewpeternelson>

Data and analytics leader, 9+ years. P.Eng, Berkeley MIDS. I architect modern data platforms (Dagster, Snowflake, Polars, DuckDB, dbt) and ship AI-native products — LLM-in-the-loop pipelines, agentic developer tooling with Claude Code, MCP servers, and structured-output systems that replace brittle rules with robust inference. Comfortable leading a team through a data transformation, architecting the Terraform, and shipping the React frontend that sits on top.

SELECTED IMPACT

Architected a production Dagster platform at StackDX across 15+ US states and 400+ transformation assets — 1.4B+ rows at the mart layer, 400M+ in downstream exposures.

Cut critical-dataset processing time via targeted Polars and DuckDB rewrites, incremental materializations, and coalesced fact/mart layers.

Founded twochannel.ai: an AI-native monorepo where LLM-in-the-loop pipelines and a custom Claude Code toolchain maintain 16k+ products across 2k+ brands with minimal manual work.

Owned Canlin's annual NI 51-101 corporate reserves for 5 years; technical revisions added \$730M NPV10 over 4 years.

Built the dbt Cloud / Snowflake platform that became Canlin's foundation for operational data, BI, and ML; automation saved 50+ hours/week across Operations.

EXPERIENCE

Data Platform Engineer · Stack Technologies Ltd.

Jun 2025 — Present

www.stackdx.com/

Lead engineer on the USA data platform rebuild — replacing a legacy Python ETL stack with a production-grade Dagster-native architecture on AWS, with Polars and DuckDB doing the heavy transformation work. Operating at scale: 1.4B+ rows at the mart layer, 400M+ rows in downstream exposures, ~47GB of parquet across landing/intermediate/marts. Operating as an AI-native developer: custom Claude Code skills, sub-agents and hooks act as force multipliers across scaffolding, data inspection, formatting and review — measurably increasing the rate at which new state pipelines and transformation assets ship.

Modern Data Platform Architecture

Architected production Dagster platform across 15+ US states and 400+ transformation assets (1.4B+ mart rows, 400M+ exposure rows), replacing a legacy Python ETL stack with a schema-contracted, testable modern data platform.

Cut processing time on critical datasets through targeted Polars/DuckDB rewrites, incremental materializations, and coalesced fact/mart layers.

Designed reusable exposure/coalesce patterns so legacy mart schemas ride on top of the new fact-table architecture without consumer churn.

AI-native Developer Workflow

Built a project-specific Claude Code toolchain — sub-agent, slash commands, auto-format hook — that measurably increases the rate of shipping new pipelines.

Profiled my own Claude Code usage to surface high-frequency patterns (65% parquet inspection) and promoted them into first-class agents/commands.

Cloud Infrastructure & DevOps

Deployed the platform's AWS footprint with Terraform IaC across staging and production, migrating Batch workloads to ECS for efficiency and cost.

Implemented CI/CD (GitHub Actions, ruff, pyright, asset schema-contract tests) and the branching/deployment automation that makes staging → prod boring.

Framework & Team Enablement

Standardized 100+ Dagster assets with reusable framework patterns and wrote the docs so the team scales beyond me.

Raised the engineering bar: schema contracts, type-safe Python, architectural docs, and review workflows befitting an enterprise-grade data platform.

Founder & Principal Engineer · Twochannel

Jan 2021 — Present

twochannel.ai

Designed and built Twochannel end-to-end: product catalog, ingestion pipelines, search, recommendations, and frontend. The catalog — 16k+ products across 2k+ brands — is maintained almost entirely through AI-native workflows: LLM-in-the-loop scrapers, structured-output enrichment, semantic dedup, and a Claude Code toolchain operating against a full monorepo. A working proof that a single engineer with modern AI tooling can run the data engineering of a team.

AI-assisted product catalog at scale

Brand-scrapers → LLM normalizer → reviewed dedup pipeline now maintains 16k+ products across 2k+ brands in Directus with minimal manual intervention.

Claude Code is the primary IDE for Twochannel: skills, sub-agents, hooks, and MCP servers do the undifferentiated heavy lifting; human focus stays on product decisions.

Dedup is LLM-proposed, human-confirmed — a deliberate trade of throughput for correctness after an early over-merge taught me the cost of silent merges.

Platform Architecture

Turborepo monorepo shipping Next.js (Vercel), Directus (Railway), Algolia search, Neo4j relationships, and Doppler-managed secrets across three hosting targets.

Restructured E2E suite to 39 lean tests passing green on prod; scoping GDPR account-deletion ahead of EU launch.

Senior Data Analytics Engineer · Paramount Resources Limited

Apr 2024 — May 2025

paramountres.com

Senior member of the Data Analytics and Integration team — integrating best practices across the full data stack, mentoring ICs, and advocating for a modern data platform. Delivered polished, robust solutions spanning custom web apps, modern data-stack PoCs, data engineering pipelines, and high-fidelity dashboards. Toolkit: Databricks, PowerBI, Spotfire, dbt Core, cube.js, DuckDB; SQL, Python, R, DAX.

Dashboards & Corporate Reporting

Multi-level Netback & Operating Cost PowerBI dashboard with embedded planning/forecasting, from corporate view to well level.

Companion Corporate Capital dashboard mirroring the Netback structure for capital-project visibility.

Redesigned and automated the weekly well production report end-to-end (Spotfire + Power Automate).

Custom Web Applications

Streamlit app replacing manual zipper-frac sequencing with an algorithm optimizing surface efficiency, frac azimuth, and crane-bank rotation.

Streamlit gas-plant development planner taking forecasts + simulation parameters and returning required development.

Data Governance & Platform Advocacy

Local modern data stack PoC (Python / dbt / cube.js / DuckDB / PowerBI in Docker) replacing legacy R + CSV processes.

Data-governance-committee advocate for cloud-first analytics; educated senior leadership on modern stack design.

Data & Advanced Analytics Lead · Canlin Energy Corporation

Feb 2022 — Apr 2024

canlinenergy.com/

Designed, built, and ran Canlin's industry-leading data stack (dbt Cloud, Snowflake, DataRobot, Tableau, Streamlit) and led the Integrated Remote Operating Centre (IROC) analytics. Owned descriptive and predictive solutions end-to-end, mentored junior professionals into data roles, and made the business case to senior leadership for sustained investment in analytics.

Corporate Data Warehouse on dbt Cloud + Snowflake

Built the dbt Cloud / Snowflake platform that became Canlin's foundation for operational data, BI, and advanced analytics.

Integrated SCADA, Production, Reserves, Financial, and Accounting systems into a single source of truth.

Snowflake Streams/Tasks/Snowpark for efficient in-warehouse processing, automation, and ML.

Data Science & ML Ops

XGBoost 30-day production forecaster for 8,000+ wells, inferring daily into corporate reporting.

2-year gas/oil forecasts across the full well set for long-term corporate reporting.

Streamlit SME-labeling app + classification model ranking well risk across the portfolio, with explanations.

Anomaly classifier for gas-well events (hydrates, liquid loading, mechanical failures).

Home-grown ML-ops layer (Streamlit + Snowflake + NocoDB) for labeling and data input, with DataRobot for productionized models.

Notable Deliverables

Daily auto-forecasting of 8,000+ wells (with downtime/maintenance event handling) powering near-term corporate reporting and marketing

Well-Production-Check dashboard with statistical call-to-actions — used by fast-loop production engineers daily

Gas-well anomaly detection in production

Real-time well status / production dashboards for the Integrated Remote Operating Centre (IROC)

Facility outage impact auto-assessment using networkx + Streamlit, linked to auto-forecasting

Automated weekly production report saving 50+ hours/week across Operations

Automated Lease Operating Statements

Corporate Data Scientist & Reserves Manager · Canlin Energy Corporation

Oct 2017 — Jan 2022

canlinenergy.com/

Drove Canlin's transition to a self-service data model by pairing data engineering with petroleum-data expertise — implementing Tableau as the company-wide reporting tool and data mart, advocating for repeatable BI, advanced analytics, and automation. Owned the annual NI 51-101 corporate reserves process, improving accuracy and unlocking substantial NPV gains through technical revisions.

Tableau data-warehouse implementation

Designed and managed the implementation of Tableau as a corporate-wide reporting tool and data warehouse

Served as the in-house expert for Tableau and all things data

Mentored every Tableau Creator/Explorer on data wrangling, cleaning, and visualization

Educated senior management on the necessity to transition to a self-service data model

Advocated for advanced analytics adoption across the business

Published 100+ data sources, 50 workbooks, and 20 Prep flows to the corporate data mart in one year.

NI 51-101 Corporate Reserves, 5 years

Managed all aspects of the annual NI 51-101 compliant corporate reserves process — the lifetime-value evaluation of every current and identified future asset

Automated the reserves data lifecycle (generate → clean → compare → visualize → report → share) for employees, external evaluators, and auditors

Technical revisions added +\$730M NPV10 over 4 years through rigorous data and engineering work.

Corporate Acquisitions & Divestitures

Core member of the A&D team during a depressed-pricing divestiture cycle

Dispositions helped bring corporate debt from \$120M to \$0

Built data flows and dashboards for rapid evaluation of acquisition opportunities

Exploitation / Development Engineer, Foothills & South · Centrica Energy Canada

Jul 2014 — Sep 2017

www.centrica.com/

Exploitation engineering across the Foothills and Southern Alberta assets. Authored strategic recommendations on gas-plant longevity that set 5–10-year corporate direction for the Foothills business.

Foothills gas-plant longevity analysis

Analyzed gathering systems for every Centrica-operated gas plant in southwest Alberta to model multi-scenario abandonment horizons

Provided strategic recommendations that set 5–10-year corporate direction for Foothills assets

Built a VBA-driven Opportunity Tracker with staging/approval workflow for development-team planning

Developed a long-term Panther River development plan grounded in completion-system evolution

Production & Exploitation Engineer · Pengrowth Energy

Aug 2010 — Jun 2014

On-site production engineering at the Lindbergh SAGD project plus exploitation engineering for Heavy Oil and Shallow Gas assets in Southern Alberta. Built multi-dimensional steam-chamber tracking from thermocouple data and designed drill programs grounded in analog-pool analysis.

Highlights across both roles

On-site production/operations engineering at Lindbergh SAGD — process flow diagrams, long-term transmitter trending, pump monitoring

Multi-dimensional steam-chamber tracking from thermocouple data across injectors, producers, and observation wells

Drill-program design for Heavy Oil and Shallow Gas assets, including multilateral-horizontal feasibility in thinner pay

Comprehensive comparison of every current SAGD project in Alberta

Engineering Co-op Student · Pengrowth Energy (Co-op)

2006 — 2010

Multiple co-op terms as part of the University of Waterloo Engineering program — 8 months at Lanmark Engineering, 4 months at the Olds Sour Gas Plant, and 12 months in the Pengrowth office as an Exploitation Engineering co-op.

SELECTED PROJECTS

twochannel.ai · twochannel.ai

2021 — Current

AI-assisted HiFi product catalog & intelligent stereo-system designer. Founder project; see the Twochannel entry under Experience for the full detail. 16k+ products across 2k+ brands maintained through LLM-augmented ingestion, semantic dedup, and a Claude-Code-first developer workflow. Stack: Next.js, Directus, Algolia, Neo4j, Turborepo.

AI-native monorepo developed primarily through Claude Code (skills, sub-agents, hooks, MCP)

LLM-in-the-loop ingestion + human-reviewed semantic dedup

Next.js frontend on Vercel, Directus CMS on Railway, Algolia search, Neo4j relationships

Doppler single-source-of-truth secret management across Vercel / Railway / GitHub Actions
39-test Playwright E2E suite, green on production

StackDX USA Data Platform · www.stackdx.com

2025 — Current

Production-grade USA oil & gas data platform at StackDX. Dagster-native orchestration and transformations, Polars/DuckDB performance layer, Terraform-managed AWS infra. Developed AI-first with a purpose-built Claude Code toolchain (parquet-inspector sub-agent, /inspect-parquet + /explore-data + /scaffold-layer commands, auto-format hook).

15+ US states, 400+ Dagster transformation assets, 1.4B+ rows at the mart layer

Targeted Polars + DuckDB rewrites and incremental materializations cutting processing time on critical datasets

Reusable exposure/coalesce-sources patterns for legacy-schema compatibility

AWS infra-as-code (RDS, ECS, S3) with Terraform

Custom Claude Code skills, agents, and hooks driving developer velocity

Instill · instillmeditation.ca

2017 — Current

A modern and refined meditation and lifestyle brand. Currently in transition; historically a full-stack React app for online course content.

Vedic Meditation Directory · learnvedicmeditation.co

2017 — 2022

Helped students find their local teacher of Vedic Meditation. Ran through multiple versions — from a full-stack React app down to the final iteration: a super-simple Next.js static site powered by a single Google Doc so teachers could self-manage listings.

SKILLS

AI Engineering & Agentic Development: Claude Code (agents, skills, hooks, slash commands, sub-agents), MCP servers & clients, Anthropic / OpenAI / Perplexity APIs, Vercel AI SDK, Structured outputs & schema-adherent LLM calls, Prompt caching & evals, RAG, Agent frameworks (autogen, crew), LLM-in-the-loop data pipelines, Editors: Claude Code, Cursor, Windsurf

Data Engineering & Modern Data Stack: dbt (Core & Cloud), Dagster, Snowflake, Snowpark, Polars, DuckDB, Tableau Prep, cube.js & dbt semantic layers, pandas, SQL, Schema contracts

Cloud, Infra & DevOps: AWS (ECS, RDS, S3), Terraform, Docker, GitHub Actions CI/CD, Vercel, Railway, Supabase, Doppler, pre-commit / ruff / pyright

Data Visualization & BI: Tableau, PowerBI, Spotfire, Plotly, D3.js, matplotlib, Seaborn, Dash, Bokeh, ggplot2

Machine Learning & NLP: scikit-learn, XGBoost, PyTorch, TensorFlow / Keras, Hugging Face, nltk, Time-series forecasting, Classification, Anomaly detection

Databases: Snowflake, PostgreSQL, DuckDB, Neo4j, networkX, MongoDB, SQL Server, Oracle, Hive

Web App Development: Next.js, React, Astro, Directus (headless CMS), Algolia, Node.js, Express, Tailwind CSS, Playwright, Streamlit, Flask, Dash, Turborepo / npm workspaces

ML Ops: Snowflake & Snowpark, DataRobot, Streamlit labeling apps, Active learning workflows

Programming Languages: Python, SQL, TypeScript, JavaScript, Bash, R

Upstream Oil & Gas: Exploitation / Development Engineering, Production Engineering, Operations, Corporate reporting, A&D evaluation

Petroleum Reserves Evaluation: COGEH, NI 51-101

EDUCATION

Master of Information and Data Science · University of California, Berkeley May 2016 — Jun 2018 · GPA 3.888

Bachelor of Applied Science, Chemical Engineering · University of Waterloo Sep 2005 — Jun 2010

CERTIFICATIONS

Professional Engineering Designation (P.Eng) — Association of Professional Engineers and Geoscientists of Alberta (APEGA) (2014)

AWARDS

Randy Duxbury Memorial Award — University of Waterloo (2010)

Ontario International Education Opportunity Scholarship — University of Waterloo (2007)