Panu Alaluusua

Data Engineer | Transitioning to AI Data Engineering | M.Sc. (Finance) Oulu, Finland |
panu.alaluusua@gmail.com | linkedin.com/in/panu-alaluusua | panualaluusua.fi

Profile

Data Engineer (5 years) with strong expertise in finance-sector data platforms, transitioning to **AI Data Engineering**. I combine technical expertise (Azure, Databricks) with deep understanding of the finance industry (M.Sc. Finance), enabling me to build mission-critical data products that meet business and regulatory requirements.

Databricks Certified Generative AI Engineer (2026), specializing in RAG applications, Vector Search technologies, and LLM-based solutions. I build reliable data products – from risk management pipelines to liquidity and balance sheet management – combining data platform architecture (Data Mesh) with AI-augmented development methodologies.

My approach is based on close partnership with business stakeholders, ensuring data quality and traceability, and developing platforms toward domain-oriented architectures.

Core Technical Expertise

Area	Technologies & Concepts
AI & GenAI (Emerging)	**RAG Applications**, **Vector Search**, **MLflow**, **Model Serving**, Unity Catalog, LLM Chains, **Agentic Workflows**
Data Platform Architecture	**Data Mesh**, **Azure Databricks**, **Apache Spark**, Delta Lake, ETL/ELT, PySpark, Python (Advanced), SQL
Finance Domain	**Pension Insurance**, **Risk Management Data**, Investment Operations, Middle Office, Actuarial Calculations, Regulation (Solvency II, IFRS)
Cloud & DevOps	**DataOps**, **Azure Data Factory**, **Terraform (IaC)**, Azure DevOps, CI/CD, Git
Quality & Governance	**Legacy Migrations**, Parallel Run Testing, Automated Reconciliation, Data Governance, Audit Trail

Work Experience

Solita | Data Engineer

March 2021 – Present | Oulu, Finland

Serving as Data Engineer in mission-critical data projects for a major Finnish pension insurance company, responsible for technical implementation and team guidance. Building and modernizing data platforms that

serve investment operations, risk management, actuarial calculations, and regulatory reporting.

Strategic Impact & Data Platform Development:

- **Data Platform Architecture Modernization (Data Mesh):**

    - Led the transition from legacy monolith to **domain-oriented** data platform.
    - Enabled **self-service capabilities** for finance and Middle Office teams, reducing IT dependencies.
    - Designed architecture to segregate **sensitive financial data** into dedicated domains (data governance).

- **Legacy Migrations & Quality Assurance:**

    - Led multiple critical system modernizations (e.g., Risk data, Actuarial liability data).
    - Implemented **parallel run testing** with automated reconciliation → proved **100% data consistency** and gained business confidence for production deployment.
    - Ensured **data integrity and traceability** (audit trail), critical for regulatory reporting.

Key Projects:

- **Actuarial Liability & Risk Data Pipelines:**

    - Built scalable data pipelines that calculate and aggregate data for **actuarial liability and solvency calculations**.
    - *Technology:* Databricks, PySpark, Azure Data Lake, Delta Lake.
    - *Impact:* Automated regulatory reporting (Solvency II, IFRS), ensured data accuracy and auditability.

- **Financial Data Aggregation System:**

    - Built critical aggregation engine for **multi-currency cash balances** and financial instruments.
    - Implemented **fault-tolerant mechanisms** (fallback) ensuring valid data delivery to downstream systems despite source data delays.
    - *Result:* Secured continuity of daily financial operations.

- **Risk Data Pipeline for Portfolio Management:**

    - Developed data pipelines delivering **risk metrics and scenario calculations** from risk management systems to Middle Office and portfolio managers.
    - Optimized system architecture to consistently meet critical **10:00 AM SLA deadline** for portfolio management decision-making.
    - *Feedback loop:* Fast iteration cycles based on direct Middle Office feedback.

- **Investment Data Integration Platform:**

    - Built real-time **market data API** and portfolio management system integrations.
    - Enabled up-to-date data flow for investment decision-making.

- **EUDR Regulatory Integration (EU Deforestation Regulation):**

    - Developed modular Python application for EU Deforestation Regulation compliance management.

- ○ *Architecture:* Designed **secure, segregated** solution on Azure Web Apps platform (Docker, Terraform), ensuring strict separation for auditability.
- ○ *Automation:* Replaced manual process with automated `GET → MODIFY → POST` integration.

AI-Augmented Data Development Methodologies:

- Actively leverage **LLM-assisted coding** (Cursor, Claude Code) specifically in **data development** – building data pipelines, transformations, and testing.
- Apply **Spec-Driven Development** and **Context Modeling** approaches, enabling fast iteration cycles and high code quality in data product development.
- This approach has proven particularly effective in complex data development projects where requirements evolve rapidly.

University of Oulu | Doctoral Student & Research Assistant

June 2019 – December 2020 | Oulu, Finland

Data science and advanced analytics researcher focused on higher education metrics.

- **Machine Learning & Analytics:** Trained XGBoost predictive models and performed feature engineering for learning analytics.
- **Data Pipelines:** Managed data cleaning, exploratory data analysis (EDA), and visualization workflows using R and Python.
- **Impact:** Produced customized ad-hoc analyses beyond standard BI tools to support decision-making.

Education

Master of Science (Finance) | *University of Oulu (2013 – 2020)*

- **Major:** Finance | **Minor:** Statistics/Econometrics
- **Focus Areas:** Financial econometrics, risk management, quantitative analysis, financial mathematics.
- **Relevance:** Deep understanding of financial instruments, risk models, and time series analysis. Fluent in "actuarial language" and effective communication with both finance professionals and data scientists. Statistical foundation enables understanding data quality requirements (distribution validation, change detection).

Certifications

- **Databricks Certified Generative AI Engineer Associate** (February 2026)
  - ○ *Expertise:* Generative AI Solutions, RAG Applications, Vector Search, Model Serving, MLflow, Unity Catalog, LLM Chains.
- **Databricks Certified Data Engineer Associate** (October 2022, expired 2024)

Projects & Portfolio

- **Engineering Portfolio:** Detailed case examples of my work, GenAI projects, and architectures. → panualaluusua.fi

Languages

- **Finnish:** Native
- **English:** Fluent (professional working proficiency, documentation, certifications)

Interests & Hobbies

- **Generative AI & RAG:** Building RAG applications and exploring LLM-based agents in personal projects.
- **Investing:** Active equity investor and market follower – understanding financial markets from a practical perspective.
- **Cycling:** E-cycling competitor and enthusiast (featured in passion projects on portfolio).