

TERRY KUO

+ 45 9194 6396

trrk.dev@gmail.com

Visa Status: Legal Residence in Denmark

Senior Data Engineer

Senior Data Professional with 5 years of experience and 5 Cloud Certification, bridging the gap between robust data engineering and advanced analytics. Specialized in **Databricks, Python, SQL**, and **software testing** with a proven record of designing high-quality data pipeline. Led Augmented by an M.Sc. in IT and Cognition, I possess a unique blend of technical expertise and human behavior.

Driven by a passion for solving problems at the intersection of human behavior and data I aim to combine my strong technical foundation in Python, Databricks, and SQL with my analytical mindset to deliver fit-for-purpose models, ensuring high data integrity and effective financial crime prevention.

SKILLS

- Python
- SQL
- Databricks
- ETL Pipeline
- Version Control
- CI/ CD + Azure DevOps
- Business Data Analytics
- Workflow Orchestration

CERTIFICATIONS

- Databricks Professional Data Engineer
- Azure Data Scientist (DP-100)
- Azure AI Engineer Fundamental (AI-900)
- Azure Data Fundamental (DP-900)
- AWS Practitioner (CLF-C02)

EDUCATION

MASTER OF SCIENCE, IT AND COGNITION

Copenhagen University 2019 - 2021

Bridge cognitive theories with computational modeling; specialized in Machine Learning, Natural Language Processing, and Python.

BACHELOR OF SCIENCE, PSYCHOLOGY

National Taiwan University 2009 - 2013

Strong foundation in Quantitative Analysis and experimental design for data-driven insights.

WORK EXPERIENCE

DATA ENGINEER, BUSINESS INSIGHTS

PostNord, Copenhagen 2024 - Present

- Translated complex business logic into production code while managing **CI/CD** release processes for seamless deployments.
- Modernized reporting architecture by re-engineering legacy logic into optimized SQL on **Databricks**.
- Orchestrated scalable data workflows via **Azure Data Factory** to ensure automated and reliable data delivery.

IT CONSULTANT, DATA & INSIGHTS

Capgemini, Copenhagen 2021 - 2024

Legacy Migration Project in TDC A/S

- Executed the migration of 30+ mission-critical legacy reports from T-SQL to **Databricks SQL**, establishing modern data processing standards.
- Achieved a 10x performance gain in **ETL pipelines** by refactoring **Databricks** notebook structures and optimizing Spark execution plans.
- **Led a QA team** to engineer an automated testing pipeline using **Python**, significantly reducing manual testing hours while ensuring data integrity across migrated solutions.

Analytics Project in Essity AB

- Acted as a technical liaison between Marketing and Analytics departments, resolving workflow misalignments and fostering cross-functional collaboration.
- Architected a modern, scalable **data warehouse** solution from the ground up to replace legacy infrastructure.
- Achieved a 13x performance improvement in daily data processing by re-engineering pipelines using **PySpark, Databricks, and Azure Data Lake**.
- Empowered analysts by deploying high-impact PowerBI solutions, directly enabling data-driven strategic decisions across the organization.

Fraud Prevention Project in Nets A/S

- Engineered a high-volume **Financial Crime Prevention (FCP) data product (PSD2)** currently utilized by major Danish banks, enhancing transaction monitoring capabilities.
- Orchestrated a massive-scale migration of TB-level data from 10+ disparate databases using **Hive SQL**, ensuring data integrity and readiness for advanced statistical analysis.
- Collaborated on defining modelling methodologies, specifically driving improvements in feature selection and loss function optimization to ensure models were robust and fit-for-purpose.

MACHINE LEARNING INTERN

Neurons Inc., Copenhagen

2020 - 2021

- Developed a **multimodal Computer Vision model** to predict human visual fixation, enhancing prediction accuracy by engineering an audio extraction feature pipeline.
- Built and orchestrated end-to-end data pipelines on **Google Cloud Platform (GCP)** using **Python** and **TensorFlow** to handle large-scale video datasets (~120 GB).
- Optimized distributed training performance and data throughput by identifying bottlenecks with TensorFlow Profiler, significantly reducing model training time.
- Designed efficient preprocessing workflows for unstructured video data, ensuring high-quality input for deep learning models.

RESEARCH ASSISTANT

Tsinghua University, Beijing

2018 - 2019

- Engineered **Deep Learning** models using **Python** and **PyTorch** to simulate auditory attention mechanisms, successfully bridging Cognitive Science theories with AI implementation.
- Built robust data preprocessing pipelines for unstructured datasets (**Audio and Image**), ensuring high-quality input features for neural network training.
- Authored and published a peer-reviewed paper as the First Author. Publication: *Inferring Cocktail Mechanisms of Auditory Attentional Modulation with Deep Neural Networks, Neural Computation (2022)*.

PROJECTS

UX-AGENT

Work in Progress

Architected an end-to-end, scalable UX research platform using Generative AI and CI/CD automation. The goal is to streamline nuanced user behavior research with LLM Agent system to significantly accelerates product feedback cycles.

LEARN DANISH APP

Work in Progress

Built an end-to-end Danish language learning platform that integrates many cognitive science theories into main feature design. I utilized Vibe Coding for rapid iteration and established a robust CI/CD pipeline.