Hyunji Cho

Analytics Engineer

hyunji.cho.uk@gmail.com | linkedin.com/in/hyunjic | github.com/HCHODUBLIN
Dublin, Ireland | UK Settled Status

PROFILE

Analytics Engineer with 6+ years of experience designing data models, building automated pipelines, and delivering analytical platforms in complex, multi-stakeholder environments. I specialise in transforming fragmented, messy data into well-governed, decision-ready structures using Snowflake, dbt, and Python. My background in sustainability research and EU programme delivery means I bring both engineering depth and strong stakeholder communication — translating business and regulatory requirements into scalable, tested, and documented data systems.

TECHNICAL SKILLS

Data Modelling	Medallion architecture (Bronze/Silver/Gold), dimensional modelling, entity resolution, semantic layers, data quality frameworks
Data Engineering	Snowflake, dbt (Core), Azure (Data Factory, SQL, Blob Storage), Kafka, Airflow, ETL/ELT pipeline design
Programming	SQL (expert), Python (pandas, API integration, NLP/LLM), R, Git/GitHub
AI / ML	LLM classification pipelines, prompt engineering, human-in-the-loop evaluation, agent-based architectures
BI & Reporting	Power BI, data storytelling, KPI frameworks, executive reporting
DevOps & Quality	CI/CD (GitHub Actions), automated testing (dbt tests, schema validation), Docker, documentation standards

EXPERIENCE

Lead Analytics Engineer   Mar 2024 – Present
Sharing Solutions (Trinity College Dublin spin-out) • EU Horizon Europe CULTIVATE (€10M+)

- Designed and built a data platform spanning 105 European cities from inception, defining data models, ingestion patterns, transformation logic, and quality frameworks for a complex, multi-partner EU programme.
- Architected a dbt-based medallion data warehouse (Bronze/Silver/Gold) on Snowflake with comprehensive data quality testing, entity resolution logic, and version-controlled transformation layers.
- Developed LLM-based classification pipelines processing 210,000+ records, implementing iterative quality evaluation (precision/recall/F1) and human-in-the-loop validation to improve accuracy from 32% to 74.5%.
- Designed and implemented automated ingestion pipelines using Azure Data Factory, Python, and SQL, with CI/CD via GitHub Actions ensuring reproducibility and governance compliance.
- Translated complex EU sustainability regulations (CSRD/GRI) into structured data models and automated KPI computation, producing policy-ready reporting outputs.
- Created comprehensive technical documentation including architecture decision records, entity-relationship diagrams, and data dictionaries to ensure transparency and team alignment.
- Served as Technical Lead for the data workstream, coordinating a cross-functional team while acting as trusted advisor to senior stakeholders across 20+ partner organisations.

Senior Data Analytics Engineer   Nov 2021 – Feb 2024
EcoWise Ltd • EU Horizon Europe DivAirCity (€10.8M)

- Designed analytics architectures for urban environmental sensor networks (air quality, noise, mobility), integrating heterogeneous datasets from multiple European cities under GDPR governance constraints.
- Built automated ETL pipelines with real-time data streaming (Kafka) for IoT sensor data, reducing manual processing by ~10 hours per week and enabling consistent cross-city analysis.
- Owned the full data lifecycle: ingestion, schema design, quality assurance, transformation, Power BI dashboards, and client-facing analytical reports for public- and private-sector stakeholders.
- Led analytical work packages contributing to securing €20M in Horizon Europe funding through robust, data-driven proposals and technical documentation.
- Assessed existing municipal data systems, identified architectural gaps, and delivered remediation plans aligned with local governance and IT environments.

Policy Consultant   Apr 2019 – Oct 2021
University College London • UK2070 Commission

- Developed quantitative data models for regional inequality and food insecurity using Python, machine learning, and spatial analysis, delivering evidence-based insights for national policy decisions.
- Designed automated data pipelines and reproducible analytical workflows processing large-scale socio-economic datasets across concurrent research programmes.
- Authored analytical briefs for UK parliamentary audiences, with findings directly cited in the English Devolution White Paper. Translated complex datasets into actionable recommendations for non-technical government stakeholders.

EDUCATION

MSc Computing — The Open University (Distributed Systems & Data Engineering)   Expected 2027
PhD Urban Planning — University College London (Network Analysis & Spatial Data Modelling)   2019
MSc Landscape Architecture — Seoul National University (Environmental Spatial Analysis)   2013

CERTIFICATIONS

Azure Fundamentals (AZ-900) — Microsoft, Jan 2026
Expert Data Scientist Programme — Pairview London, Mar 2022
CSRD Fundamentals — CSRD Institute, Nov 2025

SELECTED PUBLICATIONS

ACM WSDM 2026 — Interactive deployment and evaluation of a large-scale LLM-driven data classification system
ACM CIKM 2024 — Evaluation of LLM-based web retrieval and classification systems (precision, recall, F1)