Lucia Woollett (Rapanova)

Data & AI Engineer

Data and AI/ML engineer with an MSc in Computer Science and about 4 years of professional experience in data engineering, data analysis, AI/ML development, with a strong foundation in Python, SQL, and large language models.

@ lucia.woollett@proton.me | luc-rap.github.io | lucia-rapanova | luc-rap | 239-486-3886

SKILLS

Languages	Data Engineering	Data & ML	AI	Tools
Python, SQL (Postgres), R, Bash	PySpark, Databricks, Apache Airflow, AWS, ETL	Pandas, scikit-learn, Data Analysis, TensorFlow, EDA	LLMs, Ollama, RAG, MCP, vector search, Natural Language Processing, GenAI	Git, Jira, Linux, Confluence, Docker

EXPERIENCE

Data Engineer, Global Marketing

Dell Technologies | 2021 – 2025 | Bratislava, Slovakia

- Technical lead for an **entity resolution project** that linked company records across 18-20 third-party vendor sources into company intelligence datasets. Worked with large, messy, heterogeneous data - solved **data quality** problems such as address and geographic parsing, domain validation, and cleaning inconsistent fields
- Built an **LLM stage** for companies that were renamed, merged, or acquired and that name matching rejected automatically: retrieved web evidence for both companies and had two on-prem open-weight models (Mistral, Llama) assess each pair independently, accepting a match only when both agreed. Roughly 60-70% of candidates sent to this stage were resolved with a confident decision, replacing manual review
- The results were used by several teams: Sales (link accounts and build company hierarchies), Marketing (campaigns and contacting customers), and Demand Intelligence and Modelling (need for cleaner data)
- Participated in a hackathon on domain validation, using web search and an LLM pipeline to improve domain data quality
- Communicated technical solutions and project progress to cross-functional stakeholders and leadership
- Other projects and responsibilities: work on marketing campaigns, data normalization, region normalization, campaign types and subtypes; built audit table to track changes of entity_ids; data ingestion and data integration; identifying potential Dell customers from business online articles; experience using Senzing for entity resolution use cases
- Key technologies: Greenplum, SQL, Python, Airflow, Bash, GitLab

EDUCATION

MSc Computer Science

Slovak University of Technology | 2022 – 2024

- Notable courses: Neural Networks, Knowledge Discovery, Statistical Methods of Experiment Evaluation, Advanced Methods of Interactive System Design, Information Retrieval
- Thesis: “Prediction of Atomic Forces and Potential Energy Surfaces with Machine Learning Methods” (Key technologies: Python, Dscribe, JupyterLab, ML, NNs, TensorFlow)
- Team Project topic: Ball Detection in Sports Videos - cooperation with Sportradar (Key technologies: PyTorch, YOLO models, CNNs, Transformers, Deep Learning, Object Detection). Won second place in Team Project Cup 2023
- Weighted average score: 1.32/5 equivalent to US GPA 3.68/4

BSc Computer Science

Slovak University of Technology | 2018 – 2022

- Notable courses: Intelligent Data Analysis, Artificial Intelligence, Database Systems, Human-Computer Interaction, Data Science Programming
- Thesis: “Improving user experience on the Web” (Key technologies: Python, UX, JupyterLab, Azure Speech Services, VADER, nltk, sentiment analysis)
- Weighted average score: 1.45/5 equivalent to US GPA 3.55/4

OTHER ACTIVITIES AND PROJECTS

Volunteer & Mentor - PyCon, Django Girls, Aj Ty v IT
2020 - 2024

Contributed to technology education initiatives through mentoring, event coordination, and preparation of learning resources

Site Business Academy, Dell Technologies
2025

10-month program in project management and executive presentations; team won 2nd place among 30 participants

The Rise of the AI Engineer (DataExpert) - Databricks bootcamp
2026

Completed 1-week bootcamp focused on Databricks, Lakehouse architecture, and Data Engineering. Built and worked with RAG pipelines, including chunking, embedding, vector search, and MCP to connect LLM with tools and data