Ben Drury

Senior Data Automation Specialist | Streamlining Processes Through Automation | Python • SQL

LinkedIn | Portfolio | GitHub

I am a passionate, self-driven and commercially minded Data and Automation Specialist, bridging business priorities and technical execution. I design and implement practical solutions which have delivered real impact to internal teams in terms of streamlining processes, improving operational efficiency and enabling a small team to manage and be responsible for 5M+ products across 600+ suppliers. I thrive on tackling challenges quickly, learning what's needed along the way and creating measurable improvements that benefit both the business and its customers.

Career History

Senior Data Automation Analyst - Zoro UK Ltd

October 2023 - Present

Within the Merchandising team, I lead the creation of end-to-end data and automation solutions that streamline ingestion, enrichment, and categorisation for 5M+ products across 600+ suppliers. I build scalable workflows, mentor both technical and non-technical colleagues, and manage team members to ensure reliable, high-impact delivery. By taking ownership of whatever approach is needed, whether data pipelines, automating workflows, or AI solutions, I've delivered practical solutions that improve operational efficiency, data quality, and content workflows.

- Designed and led the implementation of an LLM-powered attribute extraction system processing over 4M attributes at 96% accuracy - automating a core merchandising workflow and scaling product enrichment for the team.
- Developed a vector-based product grouping model using embeddings, DBSCAN clustering and LLMs, auto-grouping 70% of SKUs (300K+ families) with 90%+ accuracy, driving consistency and product discovery across the site and freeing up 30+ hours/week of manual work.
- Built a hybrid categorisation pipeline leveraging embedding-based vector similarity for semantic relevance and BM25 keyword matching to auto-classify products into 2K+ categories with 84% accuracy that reduced manual work by 50%.
- Partnered with merchandising and data teams to align automation with commercial KPIs, accelerating supplier ETL and stock feeds, improving catalog completeness and customer experience across a multi-million pound supplier portfolio.
- Presented how we implemented our AI and automated use cases to senior leaders across Zoro UK and the wider umbrella including Grainger plc, Zoro US, MonotaRO (Japan), sparking further conversations and sessions.
- Maintained BigQuery to Looker pipelines and built team dashboards to track the impact of solutions build, content fill rates and quality errors.
- Mentored a junior analyst into a mid-level role, supporting the delivery of Al-driven normalisation strategies and a multimodal Al model for corrosive product identification.
- Delivered internal SQL and automation training across multiple teams to foster a data-driven culture.

Data Automation Analyst - Zoro UK Ltd

November 2022 - October 2023

Responsible for developing and maintaining Python-based data pipelines and automation tools to standardise and enrich product data. Focused on improving reliability, efficiency, and scalability across product onboarding, product gathering and manipulation.

- Automated product data preparation and validation workflows, eliminating manual handling in upload processes.
- Engineered large-scale supplier data ingestion into the PIM via Python + SQL, enhancing product completeness and search relevance.
- Developed a recommendation engine (TF-IDF + cosine similarity), A/B tested via Unleash, increasing click-through rates and released to 100% of our customers.

 Supported early ML enrichment experiments that evolved into production attribute extraction and categorisation systems.

Content Quality Analyst - Zoro UK Ltd

March 2022 - November 2022

Responsible for analysing product data and developing the foundational automation and technology that now underpins the Merchandising team's product workflows. Implemented Python-based solutions to scale and streamline product onboarding and enrichment processes, improving data quality and enhancing the customer experience on-site.

- Developed Python tools to support our Merchandising team and automate product processes including a
 mass image downloader responsible for over 1 million image downloads, product title standardisation
 processor and google sheet to big query pipelines for reporting capabilities.
- Built onboarding automation tools, eliminating manual quality checks and repetitive preparation tasks, cutting average handover-to-launch time to 3 days, driving the business growth to 5M+ products as a result.
- Documented and optimised enrichment processes, improving data accuracy and throughput across product lines.
- Authored training guides to build team capability in Python-based workflows.

Education

De Montfort University, Leicester

Master of Science (MSc) - Data Analytics (Distinction)

October 2020 - February 2022

De Montfort University, Leicester

Bachelor of Science (BSc Hons) - Computer Science (First Class)

October 2017 - July 2020

Stamford College, Stamford

BTEC Extended Diploma Level 3 - Information Technology (Triple Distinction*)

September 2015 - June 2017

Strengths & Skills

Languages: Python (Pandas, NumPy, scikit-learn), SQL

Data Engineering: BigQuery, GCP (Vertex AI, Cloud Functions, Cloud Run), ETL Design, Data Modelling **Automation & APIs:** FastAPI, Flask, Selenium, BeautifulSoup, Scrapy, Power Automate, App Scripts

Machine Learning: LLMs (Gemini, OpenAI, Claude), NLP, Classification, Recommendation Systems, Embeddings,

LangChain, DSPy

Tools: GitHub, Jira, Looker Studio, Postman, Excel/Google Sheets

Soft Skills: Agile delivery, mentorship, data governance, stakeholder collaboration

Project Highlights

Product Attribute Extraction API: Deployed an LLM-powered service (FastAPI + Cloud Run) converting unstructured product data into structured BigQuery tables, scaling our attribute fill rates by over 30%. **Product Image vs Title Validation Pipeline:** Developed a scheduled Cloud Run job comparing image/text embeddings to detect mismatched visuals and improve content accuracy.

Certifications

AI & ML: TensorFlow – Working with Images; Transformers: Text Classification; Advanced NLP with Python; Generative AI (Google Cloud)

Automation & Data Tools: Text Analytics and Predictions with Python; Microsoft Power Automate Essential Training; Databrick Fundamentals (Lake House Fundamentals, Al Agents, Generative Al)

Hobbies & Interests

Enjoys home DIY projects, fitness, family life and caring for a tropical fish aquarium.