

RICHARD WOLFF

Cranford, NJ 07106

Professional Summary

Staff-level data engineering and product leader with 10+ years of experience building and optimizing enterprise data platforms, including Databricks-based pipelines, CI/CD automation, and governed healthcare data solutions. Proven track record of designing complex ETL/ELT workflows, improving distributed system performance, and delivering measurable impact such as 44% annual cloud cost savings, 31% higher team productivity, and 95% faster data ingestion. Skilled in Python, SQL, PySpark, Unity Catalog, observability, and cross-functional leadership, with hands-on experience applying data quality controls, monitoring, logging, and compliance-minded access management in regulated environments.

Professional Experience

CDL Technical Product Owner, Johnson & Johnson (08/2024 - Present)- Defined the CDL product vision, data architecture, and strategic roadmap for operational analytics and user experience, using data-driven initiatives to improve performance monitoring, workflow optimization, and organizational effectiveness.

- Managed a \$2.2M budget with a focus on resource optimization and cost control, aligning delivery with strategic goals while maintaining timely project execution.
- Saved 44% of annual cloud costs by deactivating unused workflows, implementing cost controls, and optimizing pipelines, while also improving team FinOps awareness through education.
- Increased tech team productivity by 31% in Q1 2025 by streamlining processes and refining the codebase, strengthening engineering efficiency and delivery.
- Reduced data pipeline runtime by 15% by implementing optimized scheduling algorithms, improving distributed data processing efficiency and faster data processing capabilities.

Digital Product Manager, Johnson & Johnson (06/2022 - 07/2024)- Led the end-to-end product lifecycle for 5 AI/data science, 3 business intelligence, and 1 RPA use cases, defining product vision, KPIs, prioritization, product design, adoption, reliability, and impact evaluation to align delivery with business objectives.

- Built automated data pipelines and data visualizations for the Engineering & Property Services function, delivering \$650K in cost savings and ~3,000 labor hours in efficiency gains.
- Partnered with Product Managers and the Data Product Development team to improve the Product Manager playbook by adding milestones, metrics, and outcomes that strengthened technical delivery processes and user experience.

- Established business relationships to identify business goals, optimize business processes, prioritize the product roadmap, and translate product complexity into cross-functional strategies aligned with industry trends.
- Translated user stories into development roadmaps to fulfill business requirements and support execution planning.

Data Science & Engineering Team Lead, Johnson & Johnson (03/2021 - 06/2022)- Led a team of seven Data Scientists and Engineers, aligning OKRs with business partners and leadership while building the base data architecture for a new data science organization and generating more than \$300K in savings in the first year.

- Implemented the data science development process to increase experiment throughput, strengthen trust in the methodology, and streamline development timelines through sprint planning.
- Managed IT partnerships and reduced DevOps pipeline cycle time by 33%, improving delivery speed for data science and engineering work.
- Drove change management for the Data Science & Engineering team by facilitating education seminars, challenges, and walkthroughs to support adoption of new processes and practices.

Senior Data Scientist, Capgemini (10/2019 - 03/2021)- Developed technical solutions to improve product performance, including a digital twin proof of concept and NLP models for safety issue prediction and anomaly detection.

- Created end-to-end data pipelines using PySpark and Azure, improving data ingestion performance by 95% through asynchronous methodologies.
- Worked with cross-functional teams across account management and consulting to sell solutions and mentor team members in building deployable data science models.
- Managed Python package development with a team of developers using CI/CD, Git/Bitbucket, and Jira to improve the development pipeline and ensure quality.

Data Scientist, AJ Madison (08/2018 - 10/2019)- Rolled out a recommendation system to improve product discoverability and consumer engagement with relevant products, applying data science and analytics to support product performance decisions.

- Developed and implemented data-driven strategies using statistics and network analytics to improve product performance and decision-making.
- Built and managed Power BI dashboards for sales, marketing, merchandising, and warehousing teams to monitor key metrics, surface insights, and support operational reporting.
- Developed a data management platform with data dictionaries to document data sources and transformations, improving data governance and enabling efficient use of data for product innovation.
- Mentored junior analysts on the Python data stack, strengthening technical skills and supporting team development.

Senior Insights Analyst, Tempur-Sealy International (01/2017 - 08/2018)- Presented business analysis findings to executive stakeholders at the C, EVP, and VP levels, serving as the analytics team

representative and aligning insights with leadership priorities.

- Led planning and execution of analytics initiatives across the organization, managing stakeholder expectations to deliver successful data and analytics outcomes.
- Established a data governance framework to improve data quality and ensure fit-for-purpose data usage in a regulated enterprise environment.
- Developed media mix models to optimize ad spend and improve ROI, applying data-driven analysis to support marketing investment decisions.
- Defined and implemented data-driven reporting and analytics strategies, including ETL activities, reporting process automation, and executive-level dashboard development to improve efficiency and scalability.

Category Analyst, Tempur-Sealy International (07/2014 - 01/2017)- Created both ETL and ELT pipelines to feed daily data into batch scoring scripts, applying joins, cleansing, imputations, case statement matches, and depth search algorithms across hierarchies.

- Managed append-only incremental loads and SCD Type 2 dimension tables, supporting reliable data transformation workflows for North America deliver analytics.
- Rewrote pipelines that were not following PySpark best practices, reducing runtime and improving performance in distributed data processing.
- Worked with the North America deliver analytics team as a data scientist/data engineer, building production data pipelines that supported daily batch scoring and analytics.
- Handled healthcare data within the CDL environment, emphasizing data ownership and appropriate access controls when ingesting and transforming sensitive data.

Territory Manger (07/2012 - 07/2014)- Led data engineering for healthcare analytics pipelines in a player-coach capacity, designing and implementing ETL/ELT workflows that fed daily data into batch scoring scripts and supported production data science operations.

- Built and maintained Databricks-based pipelines using PySpark and SQL notebooks, including append-only incremental loads, SCD Type 2 dimension tables, joins, cleansing, imputations, case statement matching, and hierarchy-based depth search logic.
- Rewrote non-optimized PySpark pipelines to reduce runtime and improve performance, applying debugging and optimization practices to distributed data processing workloads.
- Implemented CI/CD automation within the Databricks workspace using declarative automation bundles, with full logging captured in Delta tables to support observability and operational traceability.
- Supported healthcare data governance and compliance by using Unity Catalog for metadata management and deploying row filters, column masks, and auditable infrastructure-as-code access controls aligned to PHI handling requirements.

Outside Sales Manager (02/2010 - 07/2012)- Led North America deliver analytics data science/data engineering work, building ETL/ELT pipelines that fed daily data into batch scoring scripts and supported production data science workloads.

- Designed and implemented transformations ranging from joins, cleansing, imputations, and case statement matches to depth search algorithms across hierarchies, supporting append-only incremental loads and SCD Type 2 dimension tables.
- Rewrote pipelines that were not following PySpark best practices, improving runtime and strengthening distributed data processing performance.
- Supported pipeline development, orchestration, and production support, combining data engineering and data science responsibilities in a player-coach role.

Education

University of Massachusetts Lowell, B.S.B.A, Marketing/Management
Data Science Process Alliance, DSPA Team Lead+ Certification, 11/2022
Product School, Product Manager Certificate, 10/2022
Springboard, Data Science Bootcamp, 07/2018

Skills

Databricks, Data Engineering, ETL/ELT Pipelines, Distributed Data Processing, CI/CD, Test-Driven Development, Version Control, RESTful APIs, Data Governance, Monitoring, Logging, Data Quality Validation, Healthcare Data, HIPAA, PHI, Agile, Scrum, Technical Leadership, Architectural Standards, Python, SQL, Jira, Bitbucket, Power BI, Tableau