

Chris Blue – Data Engineer

hellochrisblue@gmail.com | [GitHub](#)

Results-driven Data Engineer with expertise in crafting robust ETL pipelines, managing large-scale data migrations, and leveraging AWS to drive business insights and support critical analytics and AI systems.

PROFESSIONAL EXPERIENCE

Data Engineer

Los Angeles, CA

BuzzFeed Inc.

Jul 2023 – Present

- **Lead the migration of over 100TB of critical media content across BuzzFeed, Meta, and Adobe ecosystems**, using AWS S3 batch operations, AWS SDK, and Python scripting, achieving seamless transfers across all datasets.
- **Automated 10+ ETL workflows, ensuring 100% data reliability and consistency**, by leading the development of Python-based solutions with robust validation checks.
- **Reduced data migration preparation time by 50%**, designing and deploying an internal Streamlit web tool, and enabling internal IT teams to perform bulk S3 operations independently.
- **Increased backup capacity 5X to 100TB weekly**, by architecting scalable AWS-based archiving and backup solutions, maintaining these critical systems and streamlining monthly archiving of media assets.
- **Cut storage costs by 50%** through optimizing resource allocation across cloud and on-prem storage pools, enabling valuable reinvestment in BuzzFeed content initiatives.

Operations Manager

Los Angeles, CA

7till8 Wetsuits

May 2021 – May 2022

- **Surfaced insights to directly inform product development and operations strategy** through routine reporting of customer measurement and tailoring data analysis.
- **Managed a 15,000+ record dataset of customer measurements**, ensuring data integrity and accuracy to improve downstream tailoring workflows and pattern generation efficiency.
- **Optimized customer data management workflows** by streamlining input and storage processes, enhancing data consistency and reducing manual errors.

Associate Data Engineer

Los Angeles, CA

Osk Studio

Oct 2019 – May 2021

- **Improved data accessibility by 30%** for lead artists and developers by designing and implementing custom programming logic to support ML pipelines.
- **Built ETL processes to collect, clean, and transform multi-source datasets for ML model training**, ensuring high-quality, analysis-ready data for critical initiatives.
- **Used Python (Pandas, NumPy) to perform advanced data cleaning and manipulation**, resolving inconsistencies and ensuring data quality for ML model inputs.

TECHNICAL SKILLS

Data Engineering & ETL: Pipeline Design & Optimization, Data Modeling (Relational & Dimensional), OLAP Concepts, Data Quality & Validation, Data Integrity, Data Governance, Performance Tuning, Cost Optimization

Programming and Query Languages: Python, SQL, Bash/Shell Scripting, REST API Integration, SparkSQL

Cloud and Big Data Technologies: AWS (S3, Glue, EMR, Redshift, Lambda, Athena), Spark

Tools & DevOps: Git/GitHub, CI/CD Concepts, Streamlit, Grafana, Agile/Scrum Framework

Chris Blue – Data Engineer

hellochrisblue@gmail.com | [GitHub](#)

Results-driven Data Engineer with expertise in crafting robust ETL pipelines, managing large-scale data migrations, and leveraging AWS to drive business insights and support critical analytics and AI systems.

EDUCATION

Certified Scrum Product Owner®

Dec 2024

Scrum Alliance, Inc., Colorado, USA

Bachelor of Fine Arts

May 2015

New York University, New York, USA

*Recipient of the **Richard Hirsch Memorial Award** for outstanding scholastic attainment in the arts.*