Michael Willy, Data Engineer

MICHAELWILLY.COM | LINKEDIN.COM/IN/MICHAELWILLY/ | GITHUB.COM/CHRISWILLY

- Develop analytical models with Python & SQL to deliver global scientific research data and metadata inference
- Automate models on high-dimensional data into databases consumed by notebook and dashboard interfaces
- Holds over five years of experience developing data projects across diverse internal organizations and virtual compute platforms

Skills

Python (Polars SQL, Pandas, NumPy, Scikit-Learn, SQLAlchemy, Plotly), Git, Power BI
SQL (Oracle, MS SQL, SQLite), Bash & PowerShell, Cron, Airflow, Docker, Amazon S3

Experience

(1) **Data Engineer Applications**, `Research Chemical Company` 2023
- Operation and data automation of corporate R&D Electronic Lab Notebook (ELN)
- SQL accuracy validation for migrations, partitioning updates, and data sanitization for developers
- Build automated Python Polars SQL-context csv report generation on Linux & Microsoft platforms
- SQL & Python LDAP and regex reports for user patterns presented in Power BI visual dashboards

(2) **Data Science & Engineer**, `Research Chemical Company` 2020
- Developed Python & SQL pipelines to extract lab documents transformed into database model
- Built Fast Fourier Transform (FFT) analysis & plotting pipeline command line and notebook interface
- Built visualization with Dash Plotly, Jupyter Notebooks, and Microsoft Power BI to create a process visualization valued by another vendor as a $300k quote

(3) **Pilot Project Engineer**, `Research Chemical Company` 2018
- Integrated and automated time series and part wise database reporting using SQL & Python on virtual plant environment
- Designed machine user interfaces, instrumentation, and the data model for a new lab factory
- Managed contract equipment installation and start up for nine pilot machines valued at $800k capital

(4) **Research Engineer**, `Research Chemical Company` 2014
- Scaled-up organic light emitting diode (OLED) purification and programmed robotic thin-film device fab

(5) **Application Engineer**, `Semiconductor Manufacturing Company` 2012
- Scaled-up manufacturing and develop metrology applications for chemical vapor deposition diamond

Work Projects

(1) Fast Fourier Transform Numpy & SciPy signal processing automation with Plotly dashboard, JSON API, and runtime SQLite database

(2) Developed a semi-supervised machine learning graph model to estimate profiles from process inputs trained on semiconductor wafer results

(3) Data pipeline from distributed labs using a regular expression file tree search. User view into database was made with PHP site calling a parameterized SQL report Excel workbook using a command line Python API. SQL queries for activity on Power BI dashboard.

Education

- MS Applied & Computational Mathematics, University of Washington, Seattle, WA (2023)
- BS Chemical Engineering, Math Minor, Drexel University, Philadelphia, PA (2011)

MICHAELWILLY.COM | LINKEDIN.COM/IN/MICHAELWILLY/ | GITHUB.COM/CHRISWILLY

Date: November 17, 2023.