JULIAN DEVILLE

Vista, CA 92081
julian_deville@icloud.com
https://www.linkedin.com/in/julian-deville/

PROFESSIONAL SUMMARY

Passionate data engineer in the healthcare space skilled in data wrangling, validation, and analysis. Experienced working in challenging environments, and finding innovative solutions.

EXPERIENCE

Data Integration Engineer — March 2022 - February 2023
Apixio | San Mateo, California

- Managed data operations for healthcare provider network clientele
- Supported healthcare provider clients with diverse technical skill levels in adopting and optimizing the use of the new pre-visit product
- Created a variety of Python tools, including a Jupyter-based data monitoring dashboard and numerous Pandas data wrangling and validation utilities such as file sanity checks, ICD10, HCC, and CPT code validation, date format handling, automated generation of SQL queries for metrics, and many more
- Led several automation initiatives using Airflow and Rundeck, streamlining processes
- Wrangled wide range of of data types and specs, including .csv, .tsv, .fwf, .json, Excel, Google Sheets, and most types of images to meet complex and narrow ETL requirements
- Transitioned Informatica tasks and bash scripts to AirFlow (using Python/Pandas operators), including intricate and robust data lookups, transformations, and validations such as demographics matching, patient eligibility by time and location, claims linking, and deduplication/aggregation.
- Thrived under high-pressure situations and tight deadlines, delivering detailed analyses, accurate data extractions, and managing accelerated project timelines
- Maintained software in Github mono repo
- Worked on “Hackathon” project creating chat bot in company Slack account that could fetch common project state and resource monitoring metrics

Data Integration Specialist — September 2020 - March 2022
Apixio | San Mateo, California

- Cooperated with large Payer clients to seamlessly integrate vast healthcare data into Apixio's data platform
- Supported key clients, accounting for over 60% of the company's revenue
- Wrangled .csv, Excel, and .pdf files to meet complex ETL requirements

- Sustained Informatica Cloud pipeline for validating and processing structured and unstructured data, as well as validating and routing data based on patient demographics, eligibility, and location
- Executed various ad hoc operations tasks to address client needs, including bulk export of patient records from AWS s3 glacier meeting complex criteria, swift computation of metrics via SQL and Pandas, and complex data validation
- Created novel Jupyter interfaces for self-documenting code (Markdown), and interactive dashboards for monitoring processes
- Utilized Python, Pandas, Unix tools, Jupyter, and Informatica Cloud for efficient data management and analysis, as well as Jupyter for visualization
- Worked on “Hackathon” project migrating from traditional NAS to AWS s3

Data Analyst — July 2019 - September 2020
University of Kentucky CCTS | Lexington, KY

- Partnered with physician researchers on outcomes-focused investigations
- Engaged closely with researchers to understand data needs, delivered tailored datasets, and maintained data integrity and security in accordance with HIPAA and IRB standards
- Crafted intricate SQL queries using MS SQL Server and IBM Netezza for retrieval, aggregation, and transformation of large-scale data from an on-site EDW meeting incredibly narrow and complex criteria
- Maintained SQL scripts with version control in Gitlab
- Skillfully extracted and managed data from EHR backends, while leveraging licensed databases like IBM MarketScan and Medi-Span to address diverse client requirements
- Efficiently delivered over 100 datasets within a year, making a notable impact on medical research publications that cited the provided data
- Automated considerable portion of the role with small Javascript web UI

Graduate Research Assistant — August 2017 - May 2019
Eastern Kentucky University | Richmond, KY

- Collaborated with Kihara Labs at Purdue on machine vision applications for protein binding prediction
- Implemented and fine-tuned an advanced machine vision algorithm based on academic research
- Significantly reduced runtimes from multiple weeks to 10 minutes using an innovative precomputing approach
- Implemented the algorithm in both Matlab (GNU Octave) and Python for versatility
- Contributed to the scientific community by publishing the open-source code in Source Code for Biology and Medicine (Springer)

EDUCATION

Master of Science (M.S.) - Mathematics — May 2019
Eastern Kentucky University, Richmond, KY