Matthew Nizol

Senior Data Professional

matthew@nizol.net
810.623.3010
linkedin.com/in/mnizol
github.com/mnizol

Experience

Lead Data Engineer *Jul 2021 – Present*
Senior Data Engineer *Jan 2021 – Jun 2021*
ICPSR at University of Michigan | Ann Arbor, MI

- Configured and deployed data engineering infrastructure and tools including AWS Aurora, JupyterLab, and Apache Airflow using Terraform, Gitlab CI, Docker, Helm, and Kubernetes
- Collaborated with technical leads to develop architectural recommendations, set naming standards, and design blueprints for multimodal file ingestion, bulk metadata export, and curation workflows
- Developed Python scripts, Jupyter notebooks, and Airflow DAGs to automate data pipelines
- Assured high quality using automated CI/CD pipelines with unit tests, coverage analysis, and linting

Director, Data Operations *Jul 2018 – Dec 2020*
Manager, Data Operations *Jun 2017 – Jul 2018*
ArborMetrix | Ann Arbor, MI

- Led Data Operations team from its inception to a team of 11 engineers responsible for data engineering, data architecture & governance, database administration, and production support
- Drove EHR integration, standards support (HL7 2.x, CDA, FHIR), and DI platform redesign initiatives
- Developed data pipelines and services using NiFi, Mirth, CloverDX, Python, SQL, PL/SQL, SAS, and Bash operating in a complex cloud (AWS/Kubernetes/Docker) and colocation (Linux VM) ecosystem
- Provided production support for batch and realtime ETL jobs ingesting tens of millions of records daily into PostgreSQL, leveraging tools such as Elasticsearch, Kibana, Grafana, and Opsgenie
- Guided design of robust, harmonized data models to protect data integrity and promote reuse

Senior Data Analyst *Aug 2015 – Jun 2017*
ArborMetrix | Ann Arbor, MI

- Lead developer on analytic pipeline written in SAS to compute price-standardized, risk-adjusted episode-of-care measures based on over 1 billion Medicare and commercial insurance claim lines
- Wrote complex code in SAS, SQL, and PL/SQL to analyze unplanned readmissions, patient safety indicators, and other quality measures based on statewide MI and CA hospital discharges
- Refactored legacy analytic code, reducing run time by over 90% while preserving correctness

NSF Graduate Research Fellow *Aug 2013 – Aug 2015*
Michigan State University | East Lansing, MI

- Extended a state-of-the-art model satisfiability and instantiation algorithm via novel transformations
- Implemented research in a Python-based tool available at github.com/mnizol/ormpy

Senior Clinical Programmer *Sep 2010 – Aug 2013*
United BioSource Corporation | Ann Arbor, MI

- Used SAS to analyze safety and efficacy of Phase I-III oncology and cardiovascular clinical trials via creation of CDISC-compliant data sets (SDTM, ADaM), tables, listings, and figures
- Programmed efficacy analyses including progression-free survival, CA-125 response, and AUC of lipids
- Championed internal process improvement by developing a custom SAS unit testing framework

Services Information Developer II *May 2002 – Aug 2010*
EDS / Hewlett-Packard | Flint / Pontiac, MI

- Developed warranty analysis pipelines in SAS on Unix to compute linear time-to-repair projections
- Lead programmer on GM's warranty, special policy, and campaign reserve adequacy forecasts, which were reviewed by internal and external auditors and senior GM financial leadership
- Designed a novel join algorithm to assign claims to suppliers based on complex mappings
- Designed a method to harmonize claims categorization across GM's global warranty systems

Education

Master of Science, Computer Science *Aug 2013 – Aug 2015*
Michigan State University
GPA: 4.0/4.0 | NSF GRFP Fellow

Bachelor of Science, Computer Science *Aug 1999 – May 2003*
University of Michigan at Dearborn
GPA: 3.92/4.0 | Honors Program Student | Graduated with High Distinction

Skills

Languages: Python | SAS | SQL | PL/SQL | Bash | C

Tools:

- *Interop/ETL*: NiFi | Mirth | CloverDX | Airflow
- *API*: Postman | Flask | curl
- *Streaming:* Kafka
- *Log:* Elasticsearch | Splunk
- *Alerts:* Opsgenie | Kibana | Grafana
- *Editors:* Vim | Emacs | VSCode | Jupyter
- *VCS*: Git | SVN | Liquibase
- *Testing & QA:* Pytest | SASUnit
- *DevOps*: Terraform | Gitlab CI | Helm

Environment: Linux | Windows | K8s / Docker

AWS: S3 | RDS | Lambda | ECR | EKS | Secrets Mgr

Data:

- *Persistence*: PostgreSQL | EDB | Oracle
- *Modeling:* ER | UML | ORM
- *Formats*: XML | JSON | YAML
- *Standards*: HL7 | UMLS | CDISC

Innovation & Leadership:

- *Algorithms:* Design | Complexity analysis
- *Team*: Recruiting | Growth | Mentoring
- *Vision*: Strategic planning

Publications

- M. Nizol, L.K. Dillon, R.E.K. Stirewalt. "Toward Tractable Instantiation of Conceptual Data Models Using Non-Semantics-Preserving Model Transformations." MiSE 2014.
- "A Simple Approach to the Automated Unit Testing of Clinical SAS Macros." PharmaSUG 2013.
- "Using a Double DOW Loop to Compute Progression Free Survival." MWSUG 2011.
- "Tables as Trees: Merging with Wildcards Using Tree Traversal and Pruning." MWSUG 2010.

Awards

Fellowships & Scholarships:

- NSF Graduate Research Fellowship
- MSU University Distinguished Fellowship
- UM Dearborn Chancellor's Scholarship
- Robert C. Byrd Scholarship

Academic & Professional Recognition:

- ArborMetrix Team^3 Award
- Upsilon Pi Epsilon Honor Society
- UM Dearborn Emerging Leaders Program
- James B. Angell Scholar