

Stephen Y. Hwang

(224) 730-9044
stephenyounghwang@gmail.com
linkedin.com/in/hwangstephen

Technical certifications

AWS Certified Solutions Architect (Associate) MongoDB Database Administrator MongoDB Developer (Associate)

Skills

Programming languages	Python, R/Shiny, JavaScript, SQL, Perl
Databases	MongoDB, MySQL, Postgres, BigQuery, BigTable, Snowflake, Redshift, EMR
Data engineering	Kafka, Cloud Composer (Airflow), Google Dataproc (Spark), Pub/Sub, SNS/SQS
APIs, Visualization & Monitoring	FastAPI, Google Looker (Data Studio), Tableau, ggplot, matplotlib, Prometheus
DevOps, MLOps	Git, Docker, Kubernetes, MLflow, Infrastructure-as-code (Terraform, CloudFormation)

Professional Experience

Solutions Engineer (Applied AI/ML & Search) Research Triangle Institute, RTP, NC July 2023 - present

- Architected and deployed AI/ML enterprise applications for semantic classification and data harmonization
- Leveraged Chain-of-Thought to optimize retrieval-augmented generation (RAG) pipelines for domain-specific applications
- Designed and implemented complex ETL/ELT pipelines to extract business value from proprietary data
- Designed and built scalable FastAPI interfaces to connect modular components in data platform
- Built and orchestrated containerized applications using FastAPI within ETL workflows to streamline model training, evaluation, and inference
- Developed hybrid search solutions combining full-text Lucene search with vector search to augment information retrieval
- Designed and developed LLM-powered chatbots requiring session and memory management while adhering to data privacy and security requirements
- Provided technical enablement and documentation of platform architecture

Enterprise Presales Solutions Engineer MongoDB, New York, NY Jan 2022 - Jul 2023

- Collaborated with sales teams to support and onboard enterprise customers in healthcare and financial services
- Designed and deployed scalable data layer solutions to fulfill business and compliance requirements for digital payments, consumer lending, and management of electronic health records and biometric data
- Advised customers on best practices for data modeling, distributed computing, and event-driven architectures
- Led technical discovery sessions, delivered tailored demonstrations, and built proof-of-concepts for customer validation

Senior Bioinformatics Data Scientist Research Triangle Institute, RTP, NC Jun 2015 - Jan 2022

- Designed and implemented a scalable, event-driven cloud framework to automate bioinformatics workflows
- Led cross-functional teams of analysts and developers as a technical lead and program manager
- Developed and optimized ETL pipelines to ingest, process, and integrate multimodal biological and clinical data, enabling reproducibility and improving data accessibility and downstream analytics.
- Built interactive dashboards to enable data-driven decision-making for critical stakeholders
- Translated complex technical concepts into actionable insights for non-technical stakeholders, ensuring alignment between business requirements and bioinformatics solutions.

Lead DevOps Engineer Oncotect, Raleigh, NC Jan 2021 - Jan 2022

- Designed and established pipelines to ingest, process and store experimental biomedical data

Bioinformatics Data Scientist Johns Hopkins Medical Institutions, Baltimore, MD Oct 2011 - Jun 2015

- Utilized Python to develop visualizations and to implement machine learning methods for exploratory analysis
- Developed and maintained data pipelines to promote data interoperability across disparate public health and lab systems
- Developed pipeline to aggregate public health and genomic data, improving the legacy workflow by 7000%

Education

Johns Hopkins University	M.S. Bioinformatics, Aug 2013	Baltimore, MD
Loyola University	B.S. Bioinformatics, May 2011	Chicago, IL