

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, SQL, JavaScript, Perl
Databases	MongoDB, MySQL, Postgres, BigQuery, BigTable, Snowflake, Redshift, EMR
Frameworks & platforms	HuggingFace, LangChain, LangGraph, LlamaIndex, Databricks
Data engineering	Kafka, Cloud Composer (Airflow), Google Dataproc (Spark), Pub/Sub, SNS/SQS, regex
Data sharing	FastAPI, Flask, Google Looker (Data Studio), Tableau, ggplot, matplotlib, Prometheus
DevOps	Git, Docker, Kubernetes, MLflow, CI/CD, Infrastructure-as-code (Terraform, CloudFormation)

Professional Experience

Solutions Architect, Applied AI	Research Triangle Institute, RTP, NC	Jul 2023 - present
<ul style="list-style-type: none">Built a retrieval-augmented generation (RAG) framework to enhance knowledge retrieval and response accuracyOptimized LLM-driven solutions using Chain of Thought (CoT) prompting strategies, incorporating backpropagation and reinforcement learning to refine intermediate reasoning stepsDesigned and built hybrid search solutions integrating full-text Lucene search with vector searchDesigned and built event-driven ELT pipelines to extract and transform proprietary data to uncover novel business insightsDesigned and built FastAPI interfaces to enable seamless integration of modular components within a data platformContainerized applications and orchestrated real-time streaming workflows to ensure reliable, efficient data processingLed development of LLM-powered chatbot utilizing memory management compliant with strict security requirements.Developed proof-of-concept demos, prototypes, and custom blueprints, tailoring AI/ML solutions to customer needsProvided technical enablement and best practices for hybrid search and retrieval-augmented generation (RAG), helping teams optimize AI-powered retrieval systems.		
Solutions Architect, Enterprise Presales	MongoDB, New York, NY	Jan 2022 - Jul 2023
<ul style="list-style-type: none">Partnered with sales teams to identify and increase workloads for enterprise customers in healthcare and financial servicesDesigned scalable, event-driven solutions to meet business requirements for digital payments, consumer lending, and the management of electronic health records and biometric dataAdvised customers on best practices in data modeling, distributed computing, and event-driven architecturesProvided technical enablement, delivered tailored demonstrations, and built proof-of-concepts for technical validation		
Senior Data Scientist, Research Computing	Research Triangle Institute, RTP, NC	Jun 2015 - Jan 2022
<ul style="list-style-type: none">Led multiple cross-institutional teams of analysts and software developers as technical lead and project managerDesigned and built a scalable cloud-based framework to support event-driven bioinformatics workflowsBuilt ETL pipeline to ingest and process multimodal biomedical data, improving reproducibility and data accessibilityDesigned and built interactive dashboards to enable data-driven decision-making for critical stakeholdersInterfaced with non-technical stakeholders to identify business needs and requirements		
Data Engineer	Oncotect, Raleigh, NC	Jan 2021 - Jan 2022
<ul style="list-style-type: none">Designed and built event-driven pipelines to aggregate and prepare experimental biomedical data for		
Data Scientist	Johns Hopkins Medical Institutions, Baltimore, MD	Oct 2011 - Jun 2015
<ul style="list-style-type: none">Lead development team on Stand Up 2 Cancer Epigenetics Dream Team (https://www.su2c.org/)Built event-driven pipelines to aggregate and prepare public health and cancer genomic data for downstream analyses		

Education

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