

YONGPU (AARON) YANG

(202) 498-9099 | yongpuy@upenn.edu | Washington, DC

EDUCATION

University of Pennsylvania, School of Engineering and Applied Science <i>Master of Science in Engineering in Data Science (MSE-DS)</i>	Expected Aug 2027
George Washington University <i>Master of Science in Data Science</i>	May 2025
Shanghai Normal University <i>Bachelor of Economics in Financial Engineering</i>	June 2019

EXPERIENCE

Accredited Labs — Dallas, TX (Remote) <i>AI Software Designer</i>	June 2025 – Present
Design and develop AI-powered software systems for document processing, quality assurance, and recruitment automation.	
<ul style="list-style-type: none">Developed AI-powered resume screening tool with prompt injection security, reducing recruiter review time by 60% while improving candidate quality across 100s of applicants per positionBuilt certificate validation system processing 170,000+ calibration records using Python, BigQuery, and Google Cloud Storage, ensuring ISO/IEC 17025 complianceCreated automated QC workflows implementing validation rules for test points and technician certifications, decreasing data entry errors by 40%Designed ETL pipelines consolidating multi-branch calibration data, enabling real-time reporting and one-click invoice generation	

George Washington University — Washington, DC <i>Graduate Research Assistant</i>	Jan 2025 – May 2025
Conducted research on multi-agent AI systems and large language models for advanced text generation and comprehension.	
<ul style="list-style-type: none">Developed multi-agent systems with AWS Bedrock and open-source LLMs (LLaMA) for advanced text generation and comprehensionBuilt scalable NLP pipelines using LangGraph, LangChain, and RAG on large-scale datasets (WikiText, COCO)Utilized Gensim for topic modeling and word embeddings, incorporating XAI (SHAP, LIME) to enhance model transparency	

PROJECTS

Cubyt Equipment Scraper	Oct 2025
<ul style="list-style-type: none">Built Python web scraper on Google Cloud VM to extract 700,000+ equipment records from Cubyt databaseAutomated data extraction of model specifications, manufacturer details, and test point configurations using Selenium and BeautifulSoup	
IndySoft Data BI Dashboard	Jul 2025
<ul style="list-style-type: none">Developed Streamlit analytics dashboard integrating IndySoft calibration database with BigQuery for real-time reportingCreated certificate data browser and extraction tools for multi-branch operations	

SKILLS

Programming & Frameworks: Python (Pandas, NumPy, Scikit-Learn, PyTorch, TensorFlow), SQL, R, C, NLTK, Gensim, FastAPI, Selenium, BeautifulSoup
AI & Machine Learning: LLMs, RAG, LangChain, LangGraph, Multi-Agent Systems, NLP, Computer Vision, Fine-Tuning, Prompt Engineering, XGBoost, Random Forest, SHAP, LIME, Transformers
Cloud & Data: AWS (EC2, S3, Bedrock), Google Cloud Platform (BigQuery, GCS, Firestore), ETL Pipelines, Data Warehousing, Hadoop, Spark, SQL Server
Tools & MLOps: Docker, Git/GitHub, Streamlit, Jupyter Notebook, Model Deployment, Linux CLI, CI/CD, REST APIs, Tableau, Power BI