



# Sean Jung

Data Scientist / Engineer | Active TS/SCI | AWS-certified | U.S Navy Veteran

[linkedin.com/in/jungshinho](https://www.linkedin.com/in/jungshinho) | [\(201\) 528-3499](tel:(201)528-3499) | [jung2shinho@gmail.com](mailto:jung2shinho@gmail.com) | [jung2shinho](https://www.jung2shinho.com) | [seanjung.com](https://www.seanjung.com)

## TECHNICAL SKILLS

- Python | Typescript | Kubernetes | Docker | FastAPI | PyTorch | Scikit-Learn | LangGraph | MCP | LLMs | Milvus
- AWS | Google Cloud | TensorFlow | SQL | Redshift | DynamoDB | Django | React | API | gRPC | GraphRAG | Neo4j
- Linux | Bash | Sagemaker | Bedrock | Spark | Kafka | Clustering | Machine Learning | Data Visualization | Git

## EXPERIENCE

### DATA SCIENTIST – AI & DATA MODERNIZATION ..... Feb 2025 – Present

*Deloitte LLP | U.S Air Force - Informational Systems Security Office (ISSO) | Tampa, FL, USA*

- Developed and deployed an enterprise application on JWICS (via AWS EKS) to detect illegal supply chain activities on behalf of the U.S Government, supporting over +2M data analysts in counter-drug operations.
- Developed a Natural Language Processing (NLP) pipeline for entity recognition and relationship resolution to construct a 42M+ entity knowledge graph on Neo4j, derived from +12M classified reports.
- Utilized a HDBSCAN clustering algorithm to optimize query retrieval on a Milvus vector database, increasing speed by 31% while maintaining accuracy, availability and reliability for over +2M users.

### SOFTWARE ENGINEER – AI & DIGITAL TRANSFORMATION ..... Aug 2024 – Nov 2024

*Google LLC | Google Public Sector | Seattle, WA, USA*

- Deployed Google's GenAI for Enterprise (GENIE) platform using Kubernetes clusters for a public educational client, enabling the use of customized LLMs via LangGraph with retrieval-augmented generation (RAG).
- Developed a data pipeline for extracting knowledge graphs from unstructured data sources via Vertex AI, creating a comprehensive personalized learning system using Neo4j and Google Cloud Storage instances.
- Utilized Terraform modules to configure cloud infrastructure based on the project's CI/CD standards.

### DATA ENGINEER – SPACE WEATHER & HELIO-ANALYTICS ..... Jan 2024 – Jun 2024

*NASA Jet Propulsion Laboratory | Caltech | Pasadena, CA, USA*

- Implemented an ETL data pipeline to transform 87GB+ of geomagnetic data from NASA's OMNIWeb and SuperMAG databases, to train ML models for predicting space weather patterns and its effect on U.S power infrastructure. Resulted in an LSTM model with a >90% F1-score. (*2<sup>nd</sup> Place - UW ENGINE Competition 2024*)
- Delivered a React.js website for visualizing project research and analytics with an interactive Plotly dashboard.

### FACILITIES ENGINEERING & ACQUISITION DIVISION DIRECTOR ..... Aug 2022 – Jun 2023

*Naval Facilities Engineering Command | Naval Support Facility Deveselu | Deveselu, Olt, Romania*

- Refactored the internal C-based control software within a critical HVAC unit for the Navy's \$1.2B Aegis Ashore Missile Defense System to enhance reliability, increasing performance by 132% via convex optimization.
- Developed a \$7.3M base services contract with Vectrus Inc. (via task order) for the daily operations and maintenance of utilities and equipment to support base security, fire protection, and anti-terrorism efforts.

## PROJECTS

### SEMANTIC SEGMENTATION | INT. SOCIETY FOR PHOTOGRAMMETRY AND REMOTE SENSING

- Designed and implemented a FastCNN, ResUNet, and transformer-based neural network that can delineate public structures through sole visual segmentation using digital aerial images, reaching a +83% F1-score.

### HOME PRICE PREDICTOR | ZILLOW

- Deployed L1 & L2 regression models to predict housing prices, based Zillow datasets (e.g location, sq ft., lot size), and created data visualization using Plotly, Folium, & Bootstrap under Django framework with AWS EC2.

### LOAN VIABILITY | LENDING CLUB

- Developed a decision tree classifier using scikit-learn to determine bank approvals for 1412 Lending Club loan applications, based an individual's loan amount, credit score, delinquency history, and 7 other features.

### FOOD PRODUCT REVIEWS | AMAZON

- Created a logistics regression model using scikit-learn and seaborn to quantify and categorize the emotional sentiment of 1251 Amazon product reviews for evaluating the success of a given Amazon product.

## EDUCATION

University of Washington	(Seattle, WA)	M.S in Electrical Engineering (Data Science) – GPA: 3.9
United States Naval Academy	(Annapolis, MD)	B.S in Naval Architecture & Marine Engineering – Top 25%

## CERTIFICATES & LICENSES

AWS Solutions Architect (SAA-C03) Certification	Active: 02/09/2025
Professional Engineer (PE) – State of Washington and Oregon	ID: 23014628 / ID: 99743PE
Project Management Professional (PMP) – Agile and Scrum Methodology	ID: 3012182