Will Dupree, Ph.D.
Sr. Data Scientist

Boise, Idaho ·
willdupree90@gmail.com ·
+ 1 (208) 993-0043

Professional Summary

AI R&D leader integrating machine learning, graph analytics, and multimodal systems to solve complex challenges in the defense and government domains.

Professional Experience

APTIMA, INC., WOBURN, MA (remote)
Sr. Research Engineer & Team Lead **2020-Present**

- **Lead a cross-functional team** of 5+ data scientists and AI engineers advancing Aptima's AI and R&D portfolio, integrating modern machine learning (ML) and AI solutions into defense related applications
- **Serve as Principal Investigator (PI)** on multiple government funded AI research programs (including a 3-year, $2 M effort), overseeing full project lifecycle from proposal and team direction (UX, software, and AI engineers) to final delivery.
- **Share technical results with key stakeholders**, producing sponsor briefings and technical reports with >90% customer satisfaction resulting in new and repeat funding
- **Shape Aptima's AI strategy** by collaborating across divisions and identifying, adapting, and integrating open-source innovations (e.g., Graph RAG, causal analytics, and containerized workflows) to expand organizational capabilities
- **Recognized for excellence**, receiving *Outstanding Achievement - Aptima Annual Achievement Awards (2025)* for leading an internal R&D initiative that developed a compound AI prototype for generating military training scenarios, strategically positioning the organization for modern defense training programs.

Selected Aptima Project Experience

Event Summarization Via Structured Retrieval in Geospatial Data

- **Delivered a knowledge-graph (KG) driven event summarization system** leveraging Neo4j and LLMs to enhance search, pattern matching, and context retrieval for geospatial reports
- **Extended expert-curated domain KG**, developed using structured and unstructured data relationships, using AI/ML to enhance depth and breadth of information with techniques like named entity recognition, image captioning, and hierarchical LLM summarization (e.g., page, section).
- **Reduced the time to write report content from hours to minutes**, providing automated context lookups for locations, key words, and date-based information
- **Demonstrated a proof-of-concept UI to external stakeholders**, highlighting chat-like ability and source tagging required by their sector

Compound AI for Wargaming and Scenario Artifact Creation

- **Guided the design of a compound AI system** (e.g., multi-modal RAG, agents, tools) used to support wargame scenario planning
- **Reduced manual content generation from 6-12 months down to days-weeks** by automating the creation of planning material, such as mission goals or map-like visuals
- **Ingested and structured 30 GB+ of doctrinal data**, supporting RAG workflows with multi-modal documents for aligned government-based context awareness
- **Produced a 15 K example dataset to enable future model training**, created from hypothetical-question parsing against RAG documents

Understanding Anomalous Satellite Behavior

- **Explored time forecasting of satellite data** to aid satellite analysts in detecting anomalous activity
- **Demonstrated a proof of concept on single targeted satellite**, applying Long Short Term Memory (LSTM) neural network forecasting while out performing two other model choices
- **Guided integration of machine learning tools into larger system of Dockerized software** micro-services that are focused on satellite behavior analysis

WASHINGTON STATE UNIVERSITY, PULLMAN, WA
Research and Teaching Assistant **2013-2019**

- **Led two projects** in areas of blind detection and false-alarm analysis for gravitational waves analysis
- **Formalized statistical modeling**, combining two currently used data tests to reduce the negative effect random noise has on data analytics, showing increased detection probabilities by as much as 2%

EDUCATION & CERTIFICATES

WASHINGTON STATE UNIVERSITY **Pullman, WA**
Ph.D. Physics **May 2019**

- NASA Space Grant Fellowship, 2016; Research Assistantship for Diverse Scholars Recipient, 2013

MONTANA STATE UNIVERSITY **Bozeman, MT**
B.S. Physics & Applied Mathematics **May 2013**

Northeastern University D'Amore-McKim School of Business
Certificate of Value Creation Practitioner **2023**

TECHNICAL SKILLS

Tools, Packages, & Platforms: Python, OpenAI, Instructor, Pydantic-AI, LangChain, Pandas, Numpy, Scikit-Learn, TensorFlow, PyTorch, Instructor, Git, Docker, Neo4j
Specializations: Data Analytics, Machine Learning, Deep Learning, LLM Prompt & Context Engineering, Graph RAG, Model Validation, Mathematical/Statistical Analysis, Algorithm Development, Research, Project Planning

PUBLICATIONS

Dupree, W., Volkova, S., Kao, H.-T., Markey, M., Ganberg, G., Bessey, A., Rebensky, S., Dubai, T., & Cardenas, N. (2025). *Generative AI Models, Agents and Tools for Multimodal Training Scenario Generation and Mission Planning*. Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2025.

Volkova, S., Cassani, L., Lynch, S., Kao, H.-T., Bautista, P., & **Dupree, W.** (2025). *Compound AI for Decision Advantage: Human Digital Twin Agent Teams and Populations for Military Operations*. 17th International MODSIM World Conference 2025.

Dupree, W., Rebensky, S., Smith, J., Ulery, A., & Borghetti, L. (2025). *Advancing air combat training: Innovations in virtual simulation and red agent development*. The International Training Technology Exhibition and Conference (ITEC).

Cotton, O., Morgan, J., Lucia, L., **Dupree, W.**, Coker, J., & Ewer, M. (2024). *Training ML Classification Models of Warfighter State with fNIRS.* Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2025.

Dupree, W., Penafiel, L., & Gemmer, T. (2021). *Time forecasting satellite light curve patterns using Neural Networks*. Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS).

Penafiel, L**., Dupree, W.**, & Gemmer, T. (2021). *Clustering-based uncorrelated track association*. Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS).

Dupree, W., & Bose, S. (2019). Multi-detector null-stream-based \chi^2 statistic for compact binary coalescence searches. *Classical and Quantum Gravity*, *36*(19), 195012. https://doi.org/10.1088/1361-6382/ab30cf