

# Gerald K. White, M.S.

<https://geraldkwhite.com> | (732) 476-8915 | geraldkwhite10@gmail.com

---

## SUMMARY

Data Scientist & ML Researcher with 3+ years leading predictive-modeling & AI projects—especially government-funded research for the U.S. Navy. Expert in experimental design, end-to-end ML pipelines (Azure, PyTorch, Hugging Face), and model evaluation via RL and conformal methods. Seeking roles in predictive analytics, ML science, or applied research.

---

## TECHNICAL SKILLS

**Languages:** Python, R, SQL, Bash

**Frameworks/Tools:** LangChain, Flask, PyTorch, Hugging Face, OpenAI Gym, Flask, FAISS

**Algorithms:** Reinforcement Learning (PPO, RLHF), Conformal Prediction, ARIMA, Bayesian Feature Selection

**MLOps/Cloud:** Git, TabPy, Docker, CI/CD

---

## PROFESSIONAL EXPERIENCE

**Research Scientist – New Jersey Institute of Technology** New Jersey, NJ **09/2023 – Present**

- Designed and deployed a Python/Postgres-backed predictive-maintenance pipeline for U.S. Navy ships, supplying outputs to a Tableau–Python cloud environment.
- Boosted predictive-maintenance accuracy by 40%, enabling proactive logistics planning.
- Architected a Flask-based RAG retrieval system to surface contextual guidance via an evaluator LLM, helping sailors complete technical forms quickly and accurately.
- Implemented robust data-governance and security protocols to safeguard sensitive Navy datasets.
- Led a three-member research team—coordinating code reviews and delivering regular presentations to Admirals and Program Directors.

**Machine Learning Engineer (Part-Time) – Materium Technologies** New Jersey, NJ **11/2024 – 04/2025**

- Developed a Flask-based Python application to tackle inverse design problem in polymer nanocomposite structures, predicting candidate recipes from optical property thresholds.
- Scaled ML infrastructure and mentored one software engineer on production deployment best practices.

**Research Assistant – Toyota Technical Institute of Chicago** Chicago, IL **02/2023 – 08/2023**

- Spearheaded NLP research to advance reinforcement-learning algorithms for large language models, leveraging a GPU cluster for large-scale experimentation and Weights & Biases to manage and track hundreds of experimental runs.
- Diagnosed shortcomings in the Proximal Policy Optimization (PPO) algorithm and designed novel variants that improved model alignment by 5–15% on key benchmarks including METEOR, BLEU, and ROUGE.
- Engineered end-to-end experimentation pipeline which preprocessed multi-gigabyte text corpora, training LLaMA and GPT models and automating evaluation against standard benchmarks.
- Synthesized results and strategic recommendations in weekly presentations to faculty and project stakeholders, driving roadmap decisions and resourcing for next-phase RL research.

**Data Scientist Intern – Lubrizol** Chicago, IL **06/2022 – 10/2022**

- Built a stacked Random Forest model in Python/R to estimate material properties, boosting accuracy by 60%.
  - Enhanced chemical time-series predictions via Bayesian feature selection & ARIMA, increasing reliability by 44%.
- 

## EDUCATION

**M.S. in Statistics** University of Chicago Sep 2021 – Jun 2023

**B.S. in Chemical Engineering** (Minor: Mathematics) Villanova University Sep 2015 – May 2019

---

## PUBLICATION

“A Computational Framework for Estimating Days of Maintenance Delay of Naval Ships,” EDBT 2025 (first author).

<https://www.openproceedings.org/2025/conf/edbt/paper-296.pdf>

---

## CERTIFICATIONS

- CompTIA SYO-701: Security+ for IT and Cybersecurity
- Microsoft AZ-900: Azure Fundamentals
- Microsoft AI-900: Azure AI Fundamentals