

John J. Guerrero

john.j.guerrero.26@dartmouth.edu | [linkedin.com/in/john-guerrero](https://www.linkedin.com/in/john-guerrero) | github.com/JohnG112358 | orcid.org

EDUCATION

Dartmouth College

Class of 2026

Major: Computer Science, Minor: Chinese | GPA: 4.00 | SAT: 1600

Hanover, NH

- Francis L. Town Scientific Prize, Computer Science: Awarded to "one meritorious and deserving student in each department of scientific study" at the end of sophomore year. Chosen as the sole computer science nominee by the department and confirmed by a college wide committee.
- Phi Beta Kappa Sophomore Prize: Awarded to students with the highest grade-point average after the completion of their sophomore year.
- Rufus Choates Scholar (2024): Awarded to the top 5% of students for academic excellence.
- Rufus Choates Scholar (2023): Awarded to the top 5% of students for academic excellence.
- Sharon K. Harvey Memorial Scholarship: Awarded to two Baltimore residents for exemplary academic performance, leadership, and community service.
- ASPIRE Scholarship: One of two students program-wide recognized for outstanding contributions to research in the Johns Hopkins University Applied Physics Laboratory ASPIRE internship program.
- US Presidential Scholars Program Semifinalist: One of 620 high school seniors nationally recognized for academic excellence, community service, and leadership.
- Achieved a perfect score on the AP Computer Science A exam, one of 369 students globally.
- Citations for Meritorious Performance: Security Engineering, Linear Algebra, Object-Oriented Programming, Advanced Modern Chinese, Expository Writing
- Relevant Coursework: Computer Networks, Computational Linguistics, Algorithms, Software Development, Linear Algebra, Discrete Math, Multivariable Calculus, Probability, Intelligence & National Security, Computer Architecture

EXPERIENCE

James O. Freedman Presidential Scholar - IoT Security & Privacy

September 2024 – Present

Dartmouth College Department of Computer Science - Advised by Timothy Pierson

Hanover, NH

- Awarded Presidential Scholars research assistantship, a distinction reserved for top Dartmouth students.
- Designing novel network protocols and methodologies to negotiate smart home device privacy settings between visitors and hosts while assuring security and trustworthiness.

Data Developer - Machine Learning Track

September 2024 – Present

Digital Applied Learning & Innovation (DALI) Lab

Hanover, NH

- Developing and deploying image classification and segmentation models for the National Park Service to analyze tide pool algae and barnacle samples, reducing tedious manual annotation.

Machine Learning & Artificial Intelligence Intern

June 2024 – Present

Johns Hopkins University Applied Physics Laboratory - Assured AI Section

Laurel, MD

- Developing frameworks to enhance AI-generated predictions for downstream decision makers by assessing the uncertainty of model output; research resulted in a group-wide presentation and a workshop paper accepted to the 2024 *NeurIPS Workshop on Bayesian Decision-making and Uncertainty*.
- Designed and implemented a scalable data pipeline to process, store, and optimize access to 400,000+ images and 2+ million bounding boxes across 4 datasets.
- Trained 3 image classification models, achieving 98% accuracy; implemented 2 uncertainty quantification methods and model optimization techniques, improving downstream decision making by 20%.
- Invited by PI to continue research during the academic year in recognition of exceptional performance and contributions.

Applied Mathematics Research Assistant

May 2024 – Present

Dartmouth College Department of Mathematics - Advised by Ethan Levien

Hanover, NH

- Invited by professor due to exemplary (top 5%) performance in Honors Probability and outstanding final research project.
- Formulating novel stochastic models to represent bacterial gene transfer, enhancing understanding of antibiotic resistance.

Machine Learning Research Assistant

March 2024 – Present

Dartmouth College Department of Computer Science - Advised by Soroush Vosoughi

Hanover, NH

- Designed, constructed and validated a LLM-based methodology to generate culturally specific data for evaluating biases in LLMs; employed this framework to create a first-of-its-kind Spanish language dataset with 252 instances of 84 stereotypes.
- Crafted 9 novel prompt injection attacks across 4 models to generate and translate multilingual stereotypical text.
- Performed TF-IDF analysis on LLM-generated Spanish text; identified 15+ biases expressed by the LLM that were included in our dataset, while only 11% of these biases were covered by existing datasets.
- Extending dataset construction framework to Spain, Peru, Colombia, Mexico and Guatemala.

Undergraduate Research Fellow - Cybersecurity

June 2023 – December 2023

National Institute of Standards and Technology - Software Quality Group

Gaithersburg, MD

- Developed novel formal models for 200+ software weaknesses and 10+ software vulnerabilities to advance AI-based cybersecurity research; Collaborated with 3 faculty researchers to integrate these models into their projects.
- Identified 11 ambiguous areas in the current standard for classifying software vulnerabilities resulting in misclassifications and proposed potential improvements.
- Research resulted in a group-wide presentation, a lab-wide poster presentation, and publication in *IT Professional (IEEE)*.
- Invited by PI to continue research during the academic year in recognition of exceptional performance and contributions.

Summer Intern - Natural Language Processing

June 2022 – August 2022

National Institutes of Health - Text Mining Group

Bethesda, MD

- Developed, fine-tuned, and validated BERT-based named entity recognition systems to identify vaccine and strain names across 300,000+ papers within the LitCovid literature hub, enhancing 10 search filters.
- Achieved an F1 score of 0.97, outperforming previous models by 0.06 and leading to deployment within LitCovid.
- Research resulted in a section-wide talk, an NIH-wide poster presentation, and publication in *Nucleic Acids Research*.

ASPIRE Intern - Machine Learning & Natural Language Processing

June 2020 – June 2022

Johns Hopkins University Applied Physics Laboratory - Analytic Capabilities Group

Laurel, MD

- Engineered deep-learning methods to identify underlying moral sentiment in social media posts, achieving an F1 score of 0.92 and enhancing sponsor analytical capabilities; Project resulted in presentation to group leadership and 20+ lab members.
- Evaluated 2 generalized relation extraction models on their capacity to extract business-specific relationships from Reddit data, identifying 5 areas for retraining; Project resulted in lab-wide poster presentation.
- Employed custom network reconnaissance tools and Kali Linux cybersecurity tools (e.g., Metasploit) to simulate the first three phases of a penetration test; Project resulted in lab-wide presentation to 100+ people.

TECHNICAL SKILLS

Programming Languages: Python, C, Java, Bash, SQL

Developer Tools: Git, GitHub, GitFlow, Make, Docker, Valgrind, pytest, Multipass, Unix/Linux development, Agile Dev

Data Science Technologies: PyTorch, scikit-learn, SciPy, NumPy, spaCy, pandas, Transformers, SHAP, Gensim, VADER, NetworkX, Matplotlib, Seaborn, NLTK, XGBoost, Hugging Face, Gradio, Apache Airflow, PostgreSQL, OpenMetadata, MinIO

Cybersecurity Tools: Wireshark, Scapy, Netcat, Nmap, Traceroute, Metasploit, Hashcat, Kali Linux

Languages: Chinese (Fair - AMCAS scale, self-evaluated), Spanish (Good - AMCAS scale, self-evaluated)

PUBLICATIONS & PRESENTATIONS

1. Gregory Canal, Vladimir Leung, **John J. Guerrerio**, Philip Sage, & I-Jeng Wang. Decision-Driven Calibration for Cost-Sensitive Uncertainty Quantification. Accepted to *NeurIPS 2024 Workshop on Bayesian Decision-making and Uncertainty*.
2. Bojanova I, & **Guerrerio JJ** (2023). Labeling Software Security Vulnerabilities. *IT Professional (IEEE)*, 25(5), 64-70.
3. Chen Q, Allot A, Leaman R, Wei C, Aghaarabi E, **Guerrerio JJ**, Xu L, & Lu Z. (2023). LitCovid in 2022: An Information Resource for the COVID-19 Literature. *Nucleic Acids Research*, 51(D1), D1512–D1518.

LEADERSHIP

Executive Member and Visit Lead

September 2023 – Present

Dartmouth Coder Dojo

Hanover, NH

- Developing curriculum materials, managing club finances, and coordinating club events engaging 20+ attendees.
- Led weekly visits to a local under-resourced high school to teach Python to 15+ students.
- Authored a successful proposal for a \$1,000 grant to expand our robotics lessons.

Analyst

September 2023 – Present

Dartmouth Investment and Philanthropy Program (DIPP)

Hanover, NH

- Developed and presented equity pitches each term for an \$850,000+ portfolio.
- Employed a DCF valuation model, comprehensive risk assessment, and value investing framework for equity evaluation.

Volunteer Tutor

June 2023 – June 2024

Norwalk Housing Authority Student Success Tutoring Program

Norwalk, CT

- Tutored elementary school students 1-on-1 in math twice a week, delivering tailored lessons and practice.
- Improved student math test scores by 33%.

Mentor

March 2023 – June 2024

Dartmouth Center for Social Impact SIBs Program

Hanover, NH

- Mentored a local elementary school student identified as needing a consistent adult mentor, providing personalized support and guidance.

PROJECTS

Mini Intrusion Detection System | *Python, Scapy*

April 2024 – May 2024

- Developed a framework to detect 10 common computer network-based attacks such as ARP poisoning, SYN floods, fixed signatures, and the ICMP "ping of death."
- Employed statistical methods to identify abnormal traffic patterns.

Generative Computer Vision | *Python, PyTorch, torchvision, HuggingFace, Gradio*

March 2024 – April 2024

- Developed a generative adversarial network to generate realistic images of the MNIST fashion dataset with limited compute.
- Implemented an algorithm to combine the artistic style of one image with the content of another to create a third image.
- Deployed the style transfer model using Hugging Face Spaces and Gradio.

Inbox Guardian | *Python, PyTorch, Hugging Face, scikit-learn, Pandas, Gmail API*

February 2024 – March 2024

- Developed a Python framework to filter spam and irrelevant emails from a user's inbox.
- Scraped a user's emails using the Gmail API; Developed and fine-tuned 6 classification methods (e.g., BERT) to learn the user's personal preferences.
- Achieved a classification F1 score of 0.92 on my own personal emails.

Tiny Search Engine | *C, Bash, Make, Valgrind*

October 2023 – November 2023

- Developed a search engine with three components: Crawler, Indexer, and Querier.
- Crawler: Crawls websites, saves their content, and recursively follows links it encounters.
- Indexer: Parses crawler output and compiles search terms from each scraped webpage.
- Querier: Executes user search queries with support for operator precedence.