

Kai Spicer

517-775-0477 | kes2246@columbia.edu | kai-spicer.com | github.com/Spicerke

EDUCATION

Columbia University

Bachelor of Science in Computer Engineering

New York, NY

Aug. 2025 – May 2027

Washington & Jefferson College

Bachelor of Arts in Computer and Information Studies

Washington, PA

Aug. 2022 – May 2025

EXPERIENCE

Software Engineering Intern

L2 Labs

May 2026 – Present

New York, NY

- Owned the XML token matching subflow end to end, increasing top-10 matching accuracy by 40% through an LLM-powered retrieval pipeline with agentic refinement; accepted at ACM CIKM 2026.
- Containerized the matching service with Docker and deployed it to production for live users, building CI/CD pipelines with automated testing across local, staging, and production environments.
- Co-built and evaluated a separate agentic pipeline that extracts PDFs into a custom structured format, benchmarking each step through large-scale experiments to raise extraction accuracy from 85% to 97%.
- Shipped the full user-facing workflow for the matching subflow and delivered 35+ bug fixes across the application.

Undergraduate Research Assistant

Columbia University

Feb. 2026 – Present

New York, NY

- Engineered an automated gait labeling pipeline integrating graph neural networks (GCNs), reinforcement learning, gradient-based optimization, and human-in-the-loop learning, reducing annotation time by 20× while achieving 83% classification accuracy.
- Built the end-to-end training stack for data preprocessing, model training, evaluation, and automated label generation, enabling reproducible experimentation for mobility-aid gait classification.

SULI Research Intern

Pacific Northwest National Laboratory

May 2025 – July 2025

Richland, WA

- Engineered a lab-ready data pipeline and backend web systems in Python to standardize, manage, and visualize scientific datasets, automating reporting and improving reproducibility.
- Extracted and logged 1,000+ data points from scientific literature into a research database and automated transfers into proprietary modeling software, eliminating manual error.
- Named a Roy G. Post Foundation Undergraduate Scholar (2026) for work in nuclear waste management research.

RESEARCH & PUBLICATIONS

AidSight: Privacy-Aware Autolabeling of Pedestrian Mobility-Aid Use from Naturalistic Roadside Video

K. Spicer, Y. Fu, X. Di (First Author)

106th Annual Meeting of the Transportation Research Board (TRB)

Under Review, Aug. 2026

Aligning Real-World Document Schemas to Fine-Grained Concept Vocabularies

E. Menadjiev, I. Fisher, K. Spicer, C. Koutras, A. Bell

ACM International Conference on Information and Knowledge Management (CIKM)

Accepted, 2026

Enhancing Research Workflow with Python-Based Automation for Data Management

Poster — Waste Management Symposia, Phoenix, AZ

Mar. 2026

TECHNICAL SKILLS

ML & Data: PyTorch, TensorFlow, Scikit-learn, NumPy, pandas, Matplotlib

Methods: LLMs and agentic systems, agent benchmarking and evaluation, information retrieval, graph neural networks, reinforcement learning, human-in-the-loop learning

Languages: Python, Java, C, SQL (Postgres, NoSQL), R, HTML/CSS

Tools: Docker, CI/CD, Git, Google Cloud Platform, Flask, FastAPI, React, Node.js