Jianyu Wang

Full-stack Software Engineer & AI Practioner **Phone**: 732-692-0731

Email: swordfishwjy@gmail.com Personal Site & LinkedIn & Google Scholar

SUMMARY

- 4+ years of industry full-stack experience with AI-native dev tools
- 6+ years of research experience on machine learning, computer system and networking
- 9+ publications in IEEE/ACM top-tier journals and conferences + 1 US patent granted

SKILLS

- **Programming Languages**: TypeScript, JavaScript, Python, Java, C#, C++, SQL
- **Development Tools**: Next.js, React, React Native, Expo, Flask, Spring Boot, Google Juice, MySQL, MongoDB, Supabase, GraphQL, Docker/Kubernetes, Redis, Cloudflare, DataDog
- **CI/CD**: AWS EC2/EKS/Serverless/S3, Google Cloud Platform, Azure DevOps
- **AI Engineering**: Langchain, LangSmith, MCP, PostgreSQL with pgvector, Pinecone

WORK EXPERIENCE

Full-Stack Software Engineer, *IXL Learning, US*, 06/2023 – Present

- *Carsondellosa.com:* Rebuilt site with Next.js and GraphQL APIs for sale of 10000+ physical and digital products.
- *ABCya.com:* Implemented A/B testing, subscription, and rate limiting to host online games for 1+ million students.
- *RosettaStone.com:* Led authentication consolidation for 3+ million individual, school, and enterprise users login through a single portal, reducing end-to-end latency by 30%.
- *Emmersion.ai:* Developed AI-assessed language certification on Microsoft Azure for 34,000 enterprise customers.
- *Vocabulary.com*: Integrate Mixpanel and consent management for product analytics on global traffic.

Cloud Engineer, *Tencent America, Palo Alto, US*, 05 – 08/2021

- Implemented a parallel H.265 video transcoding workflow with FFmpeg on Tencent Cloud.
- Created web-based visualization platform for media processing workflow and progress with Vis.js and Flask
- Validated workflow configurability and efficiency on 32 concurrent cloud VMs for parallel video transcoding.

Software Engineer, *Tencent America, Palo Alto, US*, 05 – 08/2020

- Implemented a prototype of Network-based Media Processing (NBMP) standard (ISO 23090-8), offering RESTful APIs to construct custom media processing workflows, such as transcoding, subtitling, and adaptive streaming.
- Developed JSON-based workflow schemas and config validation using Python libraries, as a guard between user requests and NBMP APIs.

RESEARCH & DEVELOPMENT PROJECTS

HomeX: AI-powered Home Energy Advisor, 2026

- Built home energy analysis and advisor that turns room photos and utility bills into personalized savings plans with multimodal models, Next.js, Vercel AI SDK, and Supabase.
- Developed an AI gateway to cut re-run latency from ~110s to ~9s with freshness caching that skips the vision processing on unchanged photos.
- Tracked LLM cost per user, run, and pipeline stage by logging each AI API call's token usage to a Supabase ledger and pricing it with SQL rate tables.

Grady Sensor: Measuring seed moisture content , 2025 (funded by Oregon State University)

- Built the web dashboard and token-authenticated device API for IoT seed-moisture sensor, for sensor-data ingestion, usage metering, and role-based access, with Next.js, TypeScript, and Postgres.
- Integrated Stripe Checkout for device and measurement-plan purchases, using signature-verified, idempotent webhooks to provision devices and plan allowances.

Smart Home Devices Privacy Inference and Defense, 2024 (funded by University of Missouri)

- Developed a novel machine learning algorithm with Random Forest and autoencoder models to identify human behaviors by sniffing IoT wireless network traffic.
- Designed packet padding, traffic injection, and traffic shaping techniques to reduce the feasibility of malicious inference close to zero.

Intelligent Resource Allocation for Heterogeneous Edge-cloud System, 2023 (funded by University of Missouri)

- Modeled the patterns of computing requests from heterogeneous IoT devices with Semi-Supervised machine learning models, such as One Class-SVM and One Class Neural Network, using Tensorflow and Keras framework.
- Implemented a rational and fair resource allocation model by formulating a novel Markov Decision Process that allocates computing resources of 100 edge servers to 500 IoT devices.

EDUCATION

University of Missouri, MO
Ph.D. & M.S. in Computer Science, September 2016 – May 2023

Rutgers University, NJ
M.S. in Electrical & Computer Engineering, September 2014 – May 2016

AWARDS

- Google DevFest Florida, First Place Award, 2026
- Best Paper Award, IEEE International Conference on Cloud Networking, 2024
- U.S. Bank Cybersecurity Scholarship for academic year 2020-2021