

# Zheng Yu

[dataisland@outlook.com](mailto:dataisland@outlook.com) | +1 312-520-5936

---

## EDUCATION

### Northwestern University

*Ph.D. in Computer Science (Advisor: Prof. Xinyu Xing)*

### Shanghai Jiao Tong University

*Bachelor of Science in Computer Science | ACM Honors Program*

Evanston, IL, USA

*Sep 2022 – March 2026*

Shanghai, China

*Sep 2018 – Jun 2022*

---

## CORE COMPETENCIES

### High-Impact Open-Source Leadership

*5,000+ Combined GitHub Stars | Industry & Government Recognition*

- **Project Leadership:** Created [GPT-Fuzzer \(500+ stars\)](#), an LLM robustness testing framework integrated into Microsoft Azure's PyRIT security toolkit. Developed [PatchAgent \(100+ stars\)](#), an AI-powered automated program repair system that achieved DARPA AIXCC finalist status among 50+ international teams.
- **Core Maintainer Contributions:** Served as core developer for [Qiling Framework \(5,000+ stars\)](#), implementing hardware emulation modules. Contributed security patches and performance improvements to critical infrastructure projects including Linux Kernel, PHP interpreter, Assimp, GPAC, and Yasm assembler.
- **Competitive Programming Excellence:** [Silver medalist at China National Olympiad in Informatics \(NOI 2017\)](#), ranking top 100 among 10,000+ participants nationwide. Leveraged algorithmic expertise as Teaching Assistant for competitive programming and data structures courses at Shanghai Jiao Tong University.
- **Full-Stack Engineering:** Architected and deployed end-to-end systems across the technology stack. Expert proficiency in front-end technologies (JavaScript, React, CSS), back-end development (C/C++, Python, Java, Go, Rust), and cloud-native infrastructure (Docker, Kubernetes, CI/CD pipelines) for production-grade applications.

### Pioneering Agentic System Design for Security Purpose

*IEEE S&P & USENIX Security Publications | DARPA AIXCC Finalist*

- **Patch Backporting Agent:** First-authored [PortGPT](#), an LLM-based agent system for automated backporting with hierarchical workflow design to mitigate hallucination and context limitations. Published at IEEE S&P 2026 and backported 9 Linux kernel patches from mainline to stable versions, garnering coverage by tech media.
- **Automated Vulnerability Repair:** First-authored [PatchAgent](#) with innovative middleware layer between LLM and execution environment, enabling enhanced tool utilization. Accepted at USENIX Security 2025 and repaired 10 real-world CVEs across open-source projects including Pcapplusplus, libredwg, Assimp, libssh2, and HDF5.
- **End-to-End DARPA AIXCC System:** Core team member of 42-b3yond-6ug at DARPA AIXCC, architecting [BugBuster](#)—an LLM-powered system for automated vulnerability discovery and repair. Advanced to finalist position among 50+ international teams.

### Deep Expertise in Low-Level Security & Systems

*50+ CVE Discoveries | Compiler-Based Defenses | Highly-Cited AI Security Research*

- **Compiler & Kernel Security:** First-authored [ShadowBound](#) and co-authored [CAMP](#) for userspace and kernel hardening, both published at USENIX Security. Implemented custom LLVM optimization passes, showing mastery of compiler and low-level memory management across privilege boundaries.
- **Vulnerability & Exploit Study:** Discovered [50+ CVEs](#) across userspace applications and the Linux kernel through systematic security analysis. [Top 5 finalist at DEFCON CTF Finals \(2022-2023\)](#) with team StrawHat, showcasing deep understanding of binary exploitation and defense mechanisms.
- **AI Security Research Impact:** Co-authored [GPT-Fuzzer](#) (published as LLM-Fuzzer at USENIX Security 2024), which has become a foundational work in LLM jailbreak assessment. The framework was integrated into Microsoft Azure's PyRIT security toolkit.

---

## PUBLICATIONS

*\* denotes equal contribution*

### Patch Validation in Automated Vulnerability Repair

*Zheng Yu, Wenxuan Shi, Xinqian Sun, Zheyun Feng, Meng Xu, Xinyu Xing*  
*arXiv 2603.06858*

### pPatch: Automated Vulnerability Unpatching

*Tianyi Jing, Pengyu Ding, Meng Xu, Yinhao Hu, Zheng Yu, Dongliang Mu*  
*ACM International Conference on the Foundations of Software Engineering (FSE) 2026*

## PortGPT: Towards Automated Backporting Using Large Language Models

Zheng Yu\*, Zhaoyang Li\*, Jingyi Song, Meng Xu, Yuxuan Luo, Dongliang Mu

IEEE Symposium on Security and Privacy (S&P) 2026

## PatchAgent: A Practical Program Repair Agent Mimicking Human Expertise

Zheng Yu, Ziyi Guo, Yuhang Wu, Jiahao Yu, Meng Xu, Dongliang Mu, Yan Chen, Xinyu Xing

USENIX Security Symposium 2025

## ShadowBound: Efficient Heap Memory Protection Through Advanced Metadata Management and Customized Compiler Optimization

Zheng Yu, Ganxiang Yang, Xinyu Xing

USENIX Security Symposium 2024

## LLM-Fuzzer: Scaling Assessment of Large Language Model Jailbreaks

Jiahao Yu, Xingwei Lin, Zheng Yu, Xinyu Xing

USENIX Security Symposium 2024

## CAMP: Compiler and Allocator-based Heap Memory Protection

Zhenpeng Lin, Zheng Yu, Ziyi Guo, Simone Campanoni, Peter Dinda, Xinyu Xing

USENIX Security Symposium 2024

## FIRST: Exploiting the Multi-Dimensional Attributes of Functions for Power-Aware Serverless Computing

Lu Zhang, Chao Li, Xinkai Wang, Weiqi Feng, Zheng Yu, Quan Chen, Jingwen Leng, Minyi Guo, Pu Yang, Shang Yue

IEEE International Parallel and Distributed Processing Symposium (IPDPS) 2023

## PROFESSIONAL EXPERIENCE

---

### DepthFirst, AI. Inc

Security Research Engineer

San Francisco, CA, US

March 2026 – Present

- **Web & Application Security:** Developing LLM-based agent systems for automated vulnerability discovery in web applications, targeting both frontend and backend components written in high-level languages. Agents autonomously identify and exploit classes such as XSS, SQLi, SSRF, IDOR, authentication bypass, and business-logic flaws.
- **Low-Level System Security:** Building automated vulnerability discovery systems for programs written in low-level languages (C/C++), including memory corruption bugs (buffer overflows, use-after-free, format strings) and logic vulnerabilities in system software.
- **Automated Penetration Testing:** Designing end-to-end agentic pipelines that conduct full penetration testing engagements autonomously—covering reconnaissance, attack surface enumeration, exploitation, and impact verification across heterogeneous environments.

### Certified Kernel Tech LLC

Security Researcher

New York, NY, US

Jan 2026 – March 2026

- Developed LLM-based agents for automated smart contract vulnerability discovery and analysis
- Conducted smart contract vulnerability research across multiple platforms including Solidity, Move, and Solana

### University of Waterloo

Visiting Graduate Researcher – CrySP Lab (Advisor: Prof. Meng Xu)

Waterloo, ON, Canada

Jun 2025 – Sep 2025

- Conducted patch validation research to verify the correctness of vulnerability repairs across open-source projects.
- Designed and evaluated techniques for detecting incomplete or incorrect patches.

### JD.COM, Inc.

Security Engineer – Application Security Team

Beijing, China

Jun 2021 – May 2022

- Conducted vulnerability research and security assessments for large-scale e-commerce infrastructure.
- Developed automated security testing tools and contributed to internal security training programs.

## TEACHING EXPERIENCE

---

### Operating Systems (COMP\_SCI 343)

Graduate Teaching Assistant, Northwestern University

Sep 2025 – Dec 2025

Instructor: Dr. Branden Ghena

### Introduction to Computer Security (COMP\_SCI 350)

Graduate Teaching Assistant, Northwestern University

Sep 2024 – Dec 2024

Instructor: Prof. Xinyu Xing

### Programming Design and Data Structures (CS151)

Undergraduate Teaching Assistant, Shanghai Jiao Tong University

Jun 2019 – Sep 2019

Instructor: Prof. Huiyu Weng

## SELECTED PROJECTS

---

- PortGPT** | *LLM-based Automated Patch Backporting* [\[GitHub\]](#)
- Engineered an LLM-powered system for automatically backporting security patches across software versions, addressing a critical challenge in software maintenance that traditionally requires significant manual expert effort.
  - Implemented novel prompt engineering techniques and validation mechanisms to ensure backport correctness.
- PatchAgent** | *AI-Powered Program Repair System* [\[GitHub\]](#)
- Built an intelligent program repair agent that mimics human debugging workflows using large language models, achieving state-of-the-art performance on standard benchmarks and recognition as a DARPA AIxCC finalist.
  - Designed multi-stage repair pipeline incorporating static analysis, dynamic testing, and iterative refinement.
- ShadowBound** | *Compiler-Based Memory Safety* [\[GitHub\]](#)
- Developed novel heap memory protection mechanism using advanced metadata management and LLVM-based compiler optimizations, achieving significant performance improvements over existing solutions.
  - Implemented custom LLVM passes for fine-grained memory instrumentation with minimal runtime overhead.
- GPT-Fuzzer** | *LLM Security Testing Framework* [\[GitHub\]](#)
- Created scalable fuzzing infrastructure for evaluating LLM safety properties and jailbreak resistance, now integrated into Microsoft Azure's PyRIT security toolkit used by enterprise customers worldwide.
  - Pioneered template-based generation techniques for systematic exploration of prompt injection vulnerabilities.
- Qiling Framework** | *Advanced Binary Emulation* [\[GitHub\]](#)
- Core contributor to multi-platform binary emulation framework with 5,000+ GitHub stars. Implemented comprehensive MCU emulation module supporting STM32, GigaDevice, and SiFive microcontrollers across ARM Cortex-M and RISC-V architectures.
  - Extended firmware analysis capabilities and integrated AFL fuzzing support for embedded systems testing.
- Pymx Compiler** | *Educational Compiler Infrastructure* [\[GitHub\]](#)
- Designed and implemented a complete compiler for Java-like language targeting RISC-V architecture, featuring lexical analysis, parsing, semantic analysis, optimization passes, and code generation to RV32IM assembly.
- RISC-V CPU Implementation** | *Hardware Design* [\[GitHub\]](#)
- Architected RISC-V processor in Verilog HDL implementing Tomasulo algorithm for out-of-order execution, with instruction cache, load buffer, and branch prediction for enhanced performance on FPGA platforms.

## HONORS & AWARDS

---

- CSAW Applied Research Competition Finalist / Runner-up in Technical Impact Award** 2025  
*NYU Tandon School of Engineering – PatchAgent Project*
- USENIX Security Student Grant Recipient** 2024, 2025  
*USENIX Association*
- DARPA AIxCC Advanced Finals** 2024  
*Top Finalist – Team 42-beyond-6ug*
- ACM CCS Student Grant Recipient** 2024  
*ACM SIGSAC Conference on Computer and Communications Security*
- DEFCON CTF Finals – 5th Place** 2023  
*Team StrawHat (7th Place in 2022)*
- Outstanding Graduate Award** 2022  
*Shanghai Jiao Tong University*
- Zhiyuan Honor Scholarship** 2018–2021  
*Shanghai Jiao Tong University (Top 2% annually)*
- China National Olympiad in Informatics – Silver Medal** 2017  
*China Computer Federation (CCF) – Top 100 Nationally*

## CONFERENCE PRESENTATIONS

---

- Reversing MCU with Firmware Emulation** Dec 2022  
*BlackHat Europe* London, UK

## ACADEMIC SERVICE

---

### **Program Committee Member**

#### *AI/ML/SE Conferences*

- LLM4Code Workshop, co-located with ICSE – 2026
- International Conference on Learning Representations (ICLR) – 2025, 2026
- Conference on Neural Information Processing Systems (NeurIPS) – 2024, 2025
- International Conference on Machine Learning (ICML) – 2025
- Artificial Intelligence and Statistics (AISTATS) – 2025
- AAAI Undergraduate Consortium (AAAI-UC) – 2025
- ICECI – 2024

### **Artifact Evaluation Committee Member**

#### *Security & Systems Conferences*

- ACM Conference on Computer and Communications Security (CCS) – 2023, 2024, 2025
- USENIX Security Symposium – 2024, 2025
- Network and Distributed System Security Symposium (NDSS) – 2025
- USENIX Annual Technical Conference (ATC) – 2024
- Operating Systems Design and Implementation (OSDI) – 2024
- International Symposium on Software Testing and Analysis (ISSTA) – 2024

### **Peer Reviewer**

#### *Academic Journals*

- IEEE Transactions on Dependable and Secure Computing (TDSC)
- PeerJ Computer Science
- IEEE Transactions on Information Forensics and Security (TIFS)

## TECHNICAL SKILLS

---

**Programming Languages:** C/C++ (Expert), Python (Expert), Rust, Java, JavaScript/TypeScript, Go, OCaml, Verilog HDL

**Security & Binary Analysis:** IDA Pro, Ghidra, Angr, Unicorn Engine, Qiling Framework, Binary Ninja, GDB

**AI/ML Frameworks:** PyTorch, LangChain, Transformers, OpenAI API, Anthropic API

**Systems & Compilers:** LLVM/Clang, GCC, Linux Kernel Development, QEMU, Unicorn

**Development Tools:** Git, Docker, Kubernetes, CI/CD (GitHub Actions, Jenkins), Vivado, VSCode, Emacs

**Databases & Infrastructure:** MySQL, PostgreSQL, Redis, Apache Spark, Elasticsearch

**Languages:** English (Fluent), Chinese (Native)