

# JUNWEN AN

Github [◇ anjw2020@mail.sustech.edu.cn](#) [◇ Homepage](#)

## EDUCATION

---

### Southern University of Science and Technology

B.Eng. in Computer Science and Technology.

GPA: 3.86/4.00 (Top 5%).

Sept. 2020 - Jun. 2024 (Expected)

Shenzhen, China

### National University of Singapore

School of Computing Summer Workshop, AI/ML for Financial Services track.

Grade: A.

May 2022 - Jul. 2022

Singapore

### University of Notre Dame

Visiting Student.

Courses: Compilers and Language Design, Human-Computer Interaction.

Aug. 2023 - Dec. 2023 (Expected)

IN, United States

## PUBLICATIONS

---

1. **[ICSE'24]** Ling Jiang, **Junwen An**, Huihui Huang, Qiyi Tang, Sen Nie, Shi Wu, and Yuqun Zhang. “BinaryAI: Binary Software Composition Analysis via Intelligent Binary Source Code Matching”. [\[paper\]](#)
2. **[ICSE'24 Demo]** Ningzhi Tang\*, **Junwen An\***, Meng Chen, Aakash Bansal, Yu Huang, Collin McMillan, and Toby Jia-Jun Li. “CodeGRITS: A Research Toolkit for Developer Behavior and Eye Tracking in IDE”. [\[paper\]](#) [\[website\]](#)

## RESEARCH EXPERIENCE

---

### Research on Programmer Behavior and Programming Interface

Advised by [Prof. Toby Jia-Jun Li](#)

Keywords: Empirical Study, Developer Tool, Human Factors in SE, AI4SE

Aug. 2023 - Present

SaNDwich Lab

- Developed a research toolkit for JetBrains platform IDEs that tracks and collects programmer eye traces and IDE interactions. [\[Code\]](#) [\[Doc\]](#)
- Conducting empirical studies on programmer behavior while debugging AI-generated code.
- Prototyping contextual programming interface that enhances navigation and understanding of a code repository.

### An Extensive Study on Software Composition Analysis (SCA)

Advised by [Prof. Yuqun Zhang](#)

Keywords: Software Security, Software Evolution, AI4SE

Sept. 2022 - Aug. 2023

SUSTech ARiSE Lab

- Evaluated contributing factors that lead to false positives in large-scale code feature-based SCA tools.
- Worked on binary-to-source software composition analysis tool, which utilized LLM-based code embedding and virtual address locality, in cooperation with [Tencent Keen Lab](#).
- Contributed to the design and implementation of SCA algorithms, large-scale labeled dataset collection, and extensive evaluation of SCA tools.

### Binary Code Diffing via Graph Matching

Advised by [Prof. Yuqun Zhang](#)

Keywords: Graph Matching, Binary Analysis, Reverse Engineering

Jan. 2023 - Jun. 2023

SUSTech ARiSE Lab

- Studied and evaluated [Diaphora](#), the status quo open-source binary diffing plugin for IDA disassembler.
- Categorized different diffing approaches and evaluated binary diffing tools and algorithms using our dataset.
- Proposed and implemented seed-and-extend graph matching algorithm, resulting in 12% performance increase than DeepBinDiff, the SOTA academic diffing work.

## SELECTED PROJECTS

---

- Built compiler for a C-like programming language that compromises advanced features in C standard.
- Implemented lexical and syntax analysis, semantics checking, IR, and x86 assembly code generation.
- Supported advanced features like nested arrays, code formatting, and scope checking.

- SUSTech school forum web application deployed on Vercel, tested and used by 30+ real users.
- Designed and implemented user-friendly UI with Tailwind and Chakra UI.
- Integrated AI features, e.g., toxicity detection and recommendations, based on Milvus and OpenAI API.

## TEACHING EXPERIENCE

---

- |   |             |
|---|-------------|
| • Teaching Assistant of CS202: Computer Organization, SUSTech | Spring 2023 |
| • Teaching Assistant of CS207: Digital Logic, SUSTech         | Fall 2022   |

## SELECTED AWARDS

---

- |  |            |
|--|------------|
| • Honourable Mention in Mathematical Contest in Modeling (Top 20%) | 2023       |
| • First-Class Merit Student Scholarship, SUSTech (Top 5%)          | 2022       |
| • Second prize in National Mathematical Contest in Modeling        | 2022       |
| • Second-Class Merit Student Scholarship, SUSTech (Top 15%)        | 2021, 2023 |

## SKILLS

---

<b>Programming Languages</b>	Java, JavaScript/TypeScript, C/C++, Python, HTML/CSS.
<b>Frameworks</b>	React, Next.js, Astro, Spring Boot, Pytorch, Android.
<b>Tools</b>	Git, Docker, Figma, Adobe Photoshop, Adobe Premiere.