

# Zebin Yao

M.S. Student, Nankai-Baidu Joint Laboratory Group, Nankai University  
[zebin\\_yao@126.com](mailto:zebin_yao@126.com) — [Personal Page](#) — [GitHub](#)

## EDUCATION

---

<b>M.S. in Computer Science and Technology</b> Nankai University (NKU) <b>Supervisor:</b> Prof. Gang Wang	<b>Sep 2023 — Present</b>
<b>B.S. in Computer Science and Technology</b> Nankai University (NKU)	<b>Sep 2019 — Jun 2023</b>

## PUBLICATIONS AND MANUSCRIPTS

---

<b>ALGAS: A Low-latency GPU-Accelerated Approximate Nearest Neighbor Search System.</b> IEEE International Parallel & Distributed Processing Symposium ( <b>IPDPS</b> ), 2025.	(Third Author)
<b>Demystifying and Enhancing the Efficiency of Large Language Model Based Search Agents.</b> Annual Conference on Neural Information Processing Systems ( <b>NeurIPS</b> ), 2025.	arXiv (Submitted, Co-First Author)
<b>Dynamic Detect and Fix Hardness for Efficient Approximate Nearest Neighbor Search.</b> Proceedings of the ACM SIGMOD International Conference on Management of Data ( <b>SIGMOD</b> ), 2026.	(Under Review, Third Author)

## RESEARCH AND PROJECT EXPERIENCE

---

<b>Alibaba Group</b> <i>Research Intern</i>	<b>Nov 2024 — Jan 2025</b> Beijing, China
<ul style="list-style-type: none"><li>Addressed the long-tail recall issue in multimodal image retrieval by optimizing the ANN graph index structure, significantly improving retrieval accuracy.</li><li>Implemented an intra-query parallel algorithm leveraging multi-core CPU architecture to accelerate single-query performance under large-scale data conditions.</li></ul>	
<b>Infinity</b> <i>Open Source Contributor</i>	GitHub
<ul style="list-style-type: none"><li>Infinity is a database designed for LLM applications, offering hybrid search across dense vectors, sparse vectors, tensors, and full-text data.</li><li>Designed and implemented a disk-based vector index architecture that eliminates split-table complexity and enables efficient end-to-end indexing and querying; successfully merged into the core repository.</li></ul>	

## TEACHER ASSISTANT

---

• <b>COSC0025</b> Parallel Programming	Spring 2025, NKU
• <b>COSC0017</b> Compiler Principles	Autumn 2022, NKU

## SELECTED AWARDS

---

• Gongneng Scholarship, Nankai University	Sep 2024
• Gongneng Scholarship, Nankai University	Sep 2023
• Huawei Smart Pedestal Scholarship (sponsored by Huawei)	Oct 2022
• Innovation and Academic Excellence Scholarship, Nankai University	Oct 2021

## TECHNICAL STRENGTHS

---

- **Skills:** Vector Database, Information Retrieval, Retrieval Augmented Generation
- **Programming Languages:** C/C++, Python, Java, Golang, SQL, Shell, Markdown
- **Tools:** Git, Bash,  $\LaTeX$