Yinqi Sun — Curriculum Vitae

https://billsun.dev sunyinqi0508@gmail.com github.com/sunyinqi0508

Education

Purdue University, Department of Computer Science — **West Lafayette, IN**
Ph.D. in Computer Science, Database Research Group — *Aug 2023 -*

New York University, Courant Institute of Mathematical Sciences — **New York, NY**
M.S. in Computer Science, GPA: 3.84/4.0 — *Sept 2020 - Dec 2021*

Shandong University, Elite Class — **Shandong, China**
B.S. in Computer Science — *Sept 2013 - June 2017*

Standardized Test Scores
TOEFL iBT: 113

Publications

- Y. Su, **Yinqi Sun**, Minjia Zhang, Jianguo Wang, "**Vexless: A Serverless Vector Data Management System Using Cloud Functions**" (SIGMOD'24)
- Yunhai Wang, Y. Wang, **Yinqi Sun**, L. Zhu, K. Lu, C. Fu, M. Sedlmair, O. Deussen, and B. Chen "**Revisiting Stress Majorization as a Unified Framework for Interactive Constrained Graph Visualization**", in IEEE Transactions on Visualization and Computer Graphics (**IEEE Vis'17**), vol. 24, no. 1, pp. 489-499, Jan. 2018, doi: 10.1109/TVCG.2017.2745919.
- Yunhai Wang, Y. Wang, **Yinqi Sun***, H. Zhang*, C. Fu, M. Sedlmair, B. Chen, O. Deussen "**Structure-aware Fisheye Views for Efficient Large Graph Exploration**", in IEEE Transactions on Visualization and Computer Graphics (**IEEE Vis'18**), vol. 25, no. 1, pp. 566-575, Jan. 2019, doi: 10.1109/TVCG.2018.2864911.
- Z. Yuan, **Yinqi Sun**, Dennis Shasha, "**Forgetful Forests: Data Structures for Machine Learning on Streaming Data under Concept Drift**", in MDPI Algorithms, 2023; 16(6):278, doi: 10.3390/a16060278

Ongoing Works

Teaching

Graduate Teaching Assistant for CS348, Information Systems — *Fall 2023, Spring, Fall 2024, Spring 2025*

- Assisted in homework/exams designs and gradings, held office hours and recitation sessions.
- Developed automatic grading scripts on Gradescope for homework and lab exams.

Research

Purdue University, *Ph.D. Student* — *Aug 2023 –*
Advisor: Prof. Jianguo Wang — `https://www.cs.purdue.edu/homes/csjgwang/`

- Conducting research on vector databases, machine learning, and serverless computing.
- Developed a machine learning powered vector database tuning system that automatically suggests the best parameters for the vector index given user preferences and hardware constraints.

New York University, *Research Assistant* — *Jan 2022 – Aug 2023*
Advisor: Prof. Dennis Shasha — `https://cs.nyu.edu/~shasha`

- Designed and implemented AQuery++, an in-memory column-store database featuring compiled query execution that can be used for analytical and time-series workloads. The system supports user plugins and UDFs for extensibility. It also supports triggers for automated data ingestion and analytics.
- Involved in the design of an incremental decision tree algorithm. Built an application that runs incremental machine learning workloads on top of AQuery++.
- AQuery++ is currently being used as an education platform for the Advanced Database Systems course (CSCI-GA.2434) at NYU.

Columbia University, *Research Staff* *Aug 2018 – Jan 2020*
Advisor: Prof. Eugene Wu `http://www.cs.columbia.edu/~ewu/`

- Conducted research related to Data Provenance, database architecture and compiled query execution.
- Proposed a physical-level lineage capturing scheme that reduces the lineage capture overhead and lineage size by:
  - taking advantage of patterns that each physical operators in the query plan reorganizes the data,
  - using lazy lineage materialization

Shandong University, *Research Assistant* *Sept 2016 – Aug 2018*
Advisor: Prof. Yunhai Wang `https://www.yunhaiwang.net`

- Conducted researches around graph visualization, dimensionality reduction and vector similarity search.
- Worked on improving over Stress Majorization (MDS) algorithm to allow for custom constraints (Vis'17) and a fisheye view algorithm that focuses on minimizing structural distortion (Vis'18).
- Worked on code optimization for lab projects and parallel algorithms on GPU (CUDA).
- Studied saliency detection algorithms based on kNN (LSH and randomized kd-tree forest) that are used in Vector field visualization and built an interactive vector field visualizer.

Projects

Raytracing in GLSL *Feb - May 2021*

- Real-time interactive ray tracing with refractions in a fragment shader of WebGL. Built from ground up without external libraries, more descriptions on demo page. see:
  https://billsun.dev/graphics/hw4 Other weekly projects like this: https://billsun.dev/graphics/

AQuery++ Database *Jun - Nov 2022*

- A database management system featuring Column-Store, Compiled Query Execution, Flexible User Defined Functions, builtin timeseries functions and much more. see: **https://github.com/sunyinqi0508/AQuery2**

Interactive Vector Field Visualization Tool *2017-2018*

- A Vector Field Visualization tool written with QT (C++) that has kNN based saliency highlighting using KD tree and LSH implemented in CUDA kernels. see: **https://github.com/sunyinqi0508/cuLines**

ChocoPy Compiler *Feb - May 2021*

- A compiler in Java that compiles ChocoPy (subset of Python) programs into RISC-V assembly.
  see: **https://git.billsun.dev/bill/Chocopy**

Skills

- Programming Languages:
  Proficient in: C, C++, Python, JavaScript, Java, SQL, CUDA, GLSL.
  Familiar with: Swift, C#, Matlab, Scala, Pascal, (x86, ARM, RISCV) Assembly, Bash
- Professional Skills:
  LaTeX, Git, Docker, PyTorch, *nix, Windows, Adobe Illustrator, Photoshop
- Languages:
  Native speaker of Chinese.
  Fluent in English. (**TOEFL 113**)
  Conversational with Japanese. (JLPT N2).

Work

eBay Inc., *Software Engineering Intern* *May 2025 – Aug 2025*