Shenglai Zeng

Personal Website | Google Scholar | zengshe1@msu.edu | (+1)517-9747616

RESEARCH INTEREST

I am a third-year PhD student at Michigan State University, advised by Professor Jiliang Tang. My current research interests are mainly about RAG/Agentic AI, large language models(LLMs), information retrieval (IR), and Trustworthy AI. I also worked on federated learning for years. I have won the Best Paper Award of IEEE Transactions on Cloud Computing, 2023, nominated for the best paper candidates of ACL-2026. I also have industry experience on Meta AI, Amazon Science(Rufus), and Baidu(Search).

EDUCATION

Michigan State University — **East Lansing, U.S**
DSE Lab/PhD students in Computer Science and Engineering — *Sept 2023-Present*

Advisor: Jiliang Tang
Lab: Data Science and Engineering Lab
Research Direction: Trustworthy AI, Large language models(LLMs)

University of Electronic Science and Technology of China — **Chengdu, China**
Yingcai Honor School/B.Sc in Computer Science and Engineering — *Sept 2019-Present*

CGPA: 3.98/4.00
Weighted Average: 93.97/100(1st among 100 students)
Honors: The Most Outstanding Students Award of UESTC

PREPRINTS

- **Shenglai Zeng**, Qirui Wang, Kai Guo, Xinnan Dai, Xianxuan Long, Hui Liu
  Magnifying What Matters: Attention-Guided Adaptive Rendering for Visual Text Comprehension
  Submitted to NiPS 2026
- Kai Guo, Xinnan Dai, **Shenglai Zeng**, Harry Shomer, Haoyu Han, Yu Wang, Jiliang Tang
  Beyond Static Retrieval: Opportunities and Pitfalls of Iterative Retrieval in GraphRAG
  Submitted to ACL-ARR
- Xinnan Dai, Kai Guo, Chung-Hsiang Lo, **Shenglai Zeng**, Jiayuan Ding, Dongsheng Luo, Subhabrata Mukherjee, Jiliang Tang
  GraphGhosts: Tracing Reasoning Structures Behind Large Language Models
  Submitted to ACL ARR
- Yuping Lin, Pengfei He, Haoran Zhao, **Shenglai Zeng**, Zhen Xiang, Hui Liu, Jiliang Tang
  Structural Alignment Faking: Hiding Malicious Capabilities in Mixture-of-Experts LLMs
  Pre-print
- Jie Ren, Han Xu, Pengfei He, Yingqian Cui, **Shenglai Zeng**, Jiankun Zhang, Hongzhi Wen, Jiayuan Ding, Hui Liu, Yi Chang, Jiliang Tang
  Copyright Protection in Generative AI: A Technical Perspective
  Pre-print

PUBLICATIONS

IMPACT: 1000+ Citations

- **Shenglai Zeng***, Jiankun Zhang*, Kai Guo, Xinnan Dai, Hui Liu, Jiliang Tang, Yi Chang
  Fix Before Search: Benchmarking Agentic Query Visual Pre-processing in Multimodal Retrieval-augmented Generation
  ICML-2026
- **Shenglai Zeng**, Tianqi Zheng, Chuan Tian, Dante Everaert, Yau-Shian Wang, Yupin Huang, Michael J. Morais, Rohit Patki, Jinjin Tian, Xinnan Dai, Kai Guo, Monica Xiao Cheng, Hui Liu
  Attn-GS: Attention-Guided Context Compression for Efficient Personalized LLMs
  ACL-2026(Oral), Best Paper Candidate
- Xinnan Dai, Kai Yang, Cheng Luo, **Shenglai Zeng**, Kai Guo, Jiliang Tang
  When Do Hallucinations Arise? A Graph Perspective on the Evolution of Path Reuse and Path Compression
  ICML-2026
- Kai Guo, Xinnan Dai, Zhibo Zhang, Nuohan Lin, **Shenglai Zeng**, Jie Ren, Haoyu Han, Jiliang Tang
  Why Retrieval-Augmented Generation Fails: A Graph Perspective
  KDD-2026
- Jie Ren, Zhenwei Dai, Xianfeng Tang, Yue Xing, **Shenglai Zeng**, Hui Liu, Jingying Zeng, Qiankun Peng, Samarth Varshney, Suhang Wang, Qi He, Charu C. Aggarwal, Hui Liu
  Keeping an Eye on LLM Unlearning: The Hidden Risk and Remedy
  NIPS-2025
- **Shenglai Zeng**, Jiankun Zhang, Pengfei He, Jie Ren, Tianqi Zheng, Hanqing Lu, Han Xu, Hui Liu, Yue Xing, Jiliang Tang
  Mitigating the privacy issues in retrieval-augmented generation (rag) via pure synthetic data
  EMNLP-2025-main
- Jiankun Zhang*, **Shenglai Zeng***, Jie Ren, Tianqi Zheng, Hui Liu, Xianfeng Tang, Yi Chang
  Beyond Text: Unveiling Privacy Vulnerabilities in Multi-modal Retrieval-Augmented Generation
  EMNLP-2025-main
- Kai Guo, Harry Shomer, **Shenglai Zeng**, Haoyu Han, Yu Wang, Jiliang Tang
  Empowering GraphRAG with Knowledge Filtering and Integration
  EMNLP-2025-main
- **Shenglai Zeng**, Pengfei He, Kai Guo, Tianqi Zheng, Hanqing Lu, Yue Xing, Hui Liu
  Towards Context-Robust LLMs: A Gated Representation Fine-tuning Approach
  ACL-2025-main
- Bo Wang, Weiyi He, Pengfei He, **Shenglai Zeng**, Zhen Xiang, Yue Xing, Jiliang Tang
  Unveiling privacy risks in llm agent memory
  ACL-2025-main
- Bingheng Li, Zhikai Chen, Haoyu Han, **Shenglai Zeng**, Jingzhe Liu, Jiliang Tang
  Unveiling Mode Connectivity in Graph Neural Networks
  KDD-2025
- **Shenglai Zeng**, Jiankun Zhang, Bingheng Li, Yuping Lin, Tianqi Zheng, Dante Everaert, Hanqing Lu, Hui Liu, Yue Xing, Monica Xiao Cheng, Jiliang Tang
  Towards Knowledge Checking in Retrieval-augmented Generation: A Representation Perspective
  NAACL-2025-main(Oral)
- Jie Ren, Kangrui Chen, Yingqian Cui, **Shenglai Zeng**, Hui Liu, Yue Xing, Jiliang Tang, Lingjuan Lyu
  Six-cd: Benchmarking concept removals for benign text-to-image diffusion models
  CVPR-2025
- Pengfei He, Yue Xing, Han Xu, Jie Ren, Yingqian Cui, **Shenglai Zeng**, Jiliang Tang, Makoto Yamada, Mohammad Sabokrou
  Stealthy Backdoor Attack via Confidence-driven Sampling

TMLR

- Jie Ren, Yaxin Li, **Shenglai Zeng**, Han Xu, Lingjuan Lyu, Yue Xing, Jiliang Tang
  Unveiling and mitigating memorization in text-to-image diffusion models through cross attention
  ECCV-2024
- Han Xu,Jie Ren,Pengfei He, Yingqian Cui, **Shenglai Zeng**, Hui Liu, Jiliang Tang, Amy Liu
  On the Generalization of Training-based ChatGPT Detection Methods
  EMNLP-2024-Findings
- **Shenglai Zeng***, Yaxin Li*, Jie Ren, Yiding Liu, Han Xu, Pengfei He, Yue Xing, Shuaiqiang Wang, Jiliang Tang, Dawei Yin
  Exploring Memorization in Fine-tuned Language Models
  ACL-2024
- **Shenglai Zeng***, Jiankun Zhang*, Pengfei He, Yue Xing, Yiding Liu, Han Xu, Jie Ren, Shuaiqiang Wang, Dawei Yin, Yi Chang, Jiliang Tang
  The Good and The Bad: Exploring Privacy Issues in Retrieval-Augmented Generation (RAG)
  ACL-2024-findings
- Pengfei He, Han Xu, Jie Ren, Yingqian Cui, **Shenglai Zeng**, Hui Liu, Charu Aggarwal, Jiliang Tang
  Sharpness-aware Data Poisoning Attack
  ICLR-2024 Spotlight
- Juanhui Li, Harry Shomer, Haitao Mao, **Shenglai Zeng**, Yao Ma, Neil Shah, Jiliang Tang, Dawei Yin
  Evaluating graph neural networks for link prediction: Current pitfalls and new benchmarking
  NIPS-2023 Benchmark
- **Shenglai Zeng**, Zonghang Li, Hongfang Yu, Zhihao Zhang, Long Luo,Bo Li, Dusit Niyato
  HFedMS: Heterogeneous Federated Learning with Memorable Data Semantics in Industrial Metaverse
  IEEE Transactions on Cloud Computing, 2023 Best Paper
- **Shenglai Zeng**, Zonghang Li, Hongfang Yu, Yihong He, Zenglin Xu, Dusit Niyato, Han Yu
  Heterogeneous Federated Learning via Grouped Sequential-to-Parallel Training
  International Conference on Database Systems for Advanced Applications (DASFAA-2022)
- Jiaqi Wang*, **Shenglai Zeng***, Zewei Long, Yaqing Wang, Houping Xiao, Fenglong Ma
  Knowledge-Enhanced Semi-Supervised Federated Learning for Aggregating Heterogeneous Lightweight Clients in IoT
  SDM-2023

Patent

- **Shenglai Zeng**, Zonghang Li, Yihong He, Xun Zhang, Hongfang Yu, Gang Sun
  "A Hierarchical User Training Management Architecture and Training Strategy for Non-i.i.d Data" Chinese patent

RESEARCH EXPERIENCE

DSE Lab, Michigan State University **Lansing, U.S**
Research Assistant/Research on Trustworthy AI and LLM-safety *Sept 2023 - Present*

-Advisor: Professor Jiliang Tang

- Identify and mitigate the real privacy issues of LLMs. (Trustworthy AI)
- Integrate IR with LLMs/MLLMs (RAG/MRAG)
- Deeper understanding of underlying mechnism behind LLMs.
- Leaverage LLMs to enhance/empower challanging applications and tasks. (Agentic AI)

Meta Reality Labs(Incoming) **Seattle, USA**

Research Scientist Intern/Research on RAG/Agentic AI *Summer 2026*

-Mentor: Yang Xiao, Luna Xin Dong

- Incoming research about Multi-modal Agentic AI.

Search Science Team, Amazon **CA, USA**

Applied Scientist Intern/Research on Retrieval-augmented Generation and personalization *Summer 2024, 2025*

-Mentor: Tianqi Zheng, Dante Everaert

-Manager: Hanqing Lu, Chuan Tian, Monica Xiao Cheng

- Representation finetuning for robust RAG [ACL-2025].
- Utilize LLMs' internal behavior to conduct knowledge checking in RAG [NAACL-2025(oral)].
- Privacy of multi-modality RAG [EMNLP-2025]
- Synthetic data for RAG [EMNLP-2025]

Search Science Team, Baidu.Inc **Beijing, China**

Research Intern/Research on the memorization of LLM *May 2023 - May 2024*

-Mentor: Dr. Yiding Liu and Dr. Dawei Yin

- Investigating the memorization behavior and privacy implications of fine-tuned LLMs[ACL-2024].
- Currently worked on privacy risks of Retrieval LMs and AI-agents[ACL-2024].

Intelligent Networking and Applications Research Center, UESTC **Chengdu, China**

Research Assistant/Research on the Optimization of Federated Learning *Sept 2020 - Jun 2023*

-Mentor: Professor Hongfang Yu and Dr.Zonghang Li

- Proposed a novel idea of Sequential-to-Parallel training in FL.
  One conference Paper accepted by DASFAA 2022(First author)
- Investigated the application of FL in Industrial Metaverse.
  One journal paper accepted by IEEE TCC(First author).

The Pennsylvania State University **Pennsylvania, USA**

Online Intern/Research on Semi-supervised Federated Learning *Jun 2021 - Jun 2022*

-Mentor: Professor Fenglong Ma

- Implemented a semi-supervised federated learning system combined with novel personalized punning and structure-aware collaborative distillation techniques.
  Paper accepted by to SDM 2022(Co-First Author)
- Currently focusing on FL with different model structures.

University of British Columbia **Vancouver, Canada**

Summer Intern/Federated Data Evaluation with Unlearning *Jun 2022 - Sept 2022*

-Mentor: Professor Xiaoxiao Li

- Try to use the concept of cooperative game to evaluate the importance of data of participating users in federated learning.
- Try to accelerate the evaluation process and straggler problem using federated unlearning.

The University of Chicago **Chicago, USA**

Online Intern/Research on the IOT & Sensing Security *Mar 2020 - Mar 2021*

-Mentor: Shinan Liu(PhD candidate)

- Tried to use audio data collected by microphone to reconstruct user's state of motion during recording time.
- Proposed a mathematical-physical model to explain the correlation between different sensors' responses to motion.

AWARDS AND ACHIEVEMENTS

- **Best Paper Candidates, ACL-2026**
- **Best Paper Award, IEEE Transactions on Cloud Computing (JCR Q1)**
- **Spotlight Paper Award, ICLR**
- **The Most Outstanding Students Award of UESTC** (Highest honor in UESTC, Only10 students are awarded)
- **National Scholarship** in the session of 2019-2020.(Highest honor of undergraduate student)
- WAC Scholarship in the session of 2020-2021.(Only 10 undergraduate students in UESTC are awarded each year)
- 1st Outstanding Academic Scholarship in 2020,2021,and 2022.(Top 5 % students)

Services

- **Program Committee or Reviewer:** IEEE TMC, IEEE Trans on Service Computing, IEEE TPDS, IEEE TKDE, IoT-J, IEEE Trans on Information Forensics Security, MICCAI Workshop on Distributed, Collaborative and Federated Learning (DeCaF-2022), The ACM Transactions on Information Systems (TOIS),IEEE Transactions on Vehicular Technology(TVT),IEEE TKDD, ACL-ARR,NIPS
- Reviews **30+** papers.