

Shuning Zhang



My research focuses on **human-AI interaction**, especially **safety, privacy and security aspects**. I address these challenges by distilling human insights and develop human-in-the-loop algorithms. I publish 11 first (co-first) authored CCF-A full papers on CHI, CSCW, Ubicomp conferences, and receive awards spanning top-tier conferences like IJCAI, CCS and Ubicomp. I interned at Bytedance, Tencent, Zhipu, and have connections in universities with Ant Group, Huawei, etc.

Education

Tsinghua University

PhD. Student

- Scholarship: National Scholarship

Sep. 2023 – Present

Institute for Network Sciences and Cyberspace

Tsinghua University

Bachelor

- Scholarship: National Scholarship (3/210), HAODUSHU Scholarship (1/210), ZHONGSHIMO Scholarship (3/210)

Sep. 2019 – June. 2023

Computer Science and Technology

Internship Experience

Tencent, CDG

Designed multi-agent retrieval and generation frameworks for Text2SQL and Text2DSL, improving accuracy by **14%**.

Jun. 2025 - Aug. 2025

Bytedance, Overseas education

Improved math solving models Intern-7B by **5%** through process-oriented reward and rejection fine-tuning.

Mar. 2024 - Jun. 2024

Visiting Experience

Remote research assistant (Due to COVID19)

Proposed adaptive neural networks on physiological signals for emotion classification

Jun. 2022 - Aug. 2022

University of Washington

Remote research assistant (Due to COVID19)

Designed REACTGENIE, a typescript framework for developers to facilitate multimodal apps' building (CHI24)

Jun. 2022 - Aug. 2022

Stanford University

Remote research assistant (Due to COVID19)

Implemented template based NER model and benchmarked the model's transferability (ACL23)

Jun. 2021 - Aug. 2021

UCLA

Selected Publications

CoVer: Conflict-aware Claim Verification

(In Submission ACL ARR 26) **Shuning Zhang**, Dai Shi, Bohao Chu, Hui Wang, Yuwei Chuai, Yifan Wang, Jingruo Chen, Simin Li, Xin Yi, Hewu Li

MultiCom: Multi-Agent Simulation-Based Community Note Evaluation

(In Submission AAAI 27) Changxi Wen*, **Shuning Zhang***, Bohao Chu, Yuwei Chuai, Hui Wang, Xiaolong Li, Dai Shi, Xin Yi

Towards Benchmarking and Assessing Visual Naturalness of Physical World Adversarial Attacks

(CVPR23) Simin Li, **Shuning Zhang**, Gujun Chen, Dong Wang, Pu Feng, Jiakai Wang, Aishan Liu, Xin Yi, Xianglong Liu

PrivWeb: Unobtrusive and Content-aware Privacy Protection For Web Agents

(CHI26) **Shuning Zhang***, Yutong Jiang*, Rongjun Ma, Yuting Yang, Mingyao Xu, Zhixin Huang, Xin Yi, Hewu Li

Exploring Collaboration Patterns and Strategies in Human-AI Co-creation Through the Lens of Agency: A Scoping Review of Top-tier HCI Literature

(CSCW25) **Shuning Zhang**, Hui Wang, Xin Yi

Actual Achieved Gain and Optimal Perceived Gain: Modeling Human Take-over Decisions Towards Automated Vehicles' Suggestions

(CHI25) **Shuning Zhang**, Xin Yi, Shixuan Li, Chuye Hong, Gujun Chen, Jiarui Liu, Xueyang Wang, Yongquan 'Owen' Hu, Yuntao Wang, Hewu Li

PrivCAPTCHA: Interactive CAPTCHA to Facilitate Effective Comprehension of APP Privacy Policy

(CHI25) **Shuning Zhang**, Xin Yi, Shixuan Li, Haobin Xing, Hewu Li

The Manipulative Power of Voice Characteristics: Investigating Deceptive Patterns in Mandarin Chinese Female Synthetic Speech

(IMWUT25) **Shuning Zhang***, Han Chen*, Yabo Wang, Yiqun Xu, Jiaqi Bai, Yuanyuan Wu, Shixuan Li, Xin Yi, Hewu Li

Towards Aligning Personalized Conversational Recommendation Agents with Users' Privacy Preferences

(🏆 CCS25 HAIPS workshop distinguished paper) **Shuning Zhang**, Ying Ma, Jingruo Chen, Simin Li, Xin Yi, Hewu Li

SituGuard: LLM-based Fine-grained Smart Glass Privacy Control in Home Environments

(🏆 Ubicomp25 best student challenge award) **Shuning Zhang**, Ying Ma, Qucheng Zang, Yongquan Hu

Synlab: A Spatio-Temporal Aggregation Framework for Collaborative Visual Misinformation Detection in Videos

(🏆 IJCAI25 DDL Workshop best paper candidate) **Shuning Zhang**, Linzhi Wang, Ruokai Zhao, Yiqun Xu, Yuanyuan Wu, Yabo Wang, Simin Li, Xin Yi, Hewu Li

Explainable AIGC Misinformation Video Identification Via Vision-based Language Models With Human-initiated Insights

(IJCAI25 DDL workshop) **Shuning Zhang**, Yiqun Xu, Ruokai Zhao, Yuanyuan Wu, Yabo Wang, Linzhi Wang, Simin Li, Xin Yi, Hewu Li

Evaluating the Privacy Valuation of Personal Data on Smartphones

(IMWUT24) Lihua Fan*, **Shuning Zhang***, Yan Kong, Xin Yi, Yang Wang, Xuhai 'Orson' Xu, Chun Yu, Hewu Li, Yuanchun Shi

Squeez'In: Private Authentication on Smartphones based on Squeezing Gestures

(CHI23) Xin Yi, **Shuning Zhang**, Ziqi Pan, Louisa Shi, Fengyan Han, Yan Kong, Hewu Li, Yuanchun Shi

Technical Skills, Language Skills, and Interests

Programming Languages: Python, Javascript, Java, C Script

Language: Chinese (Native), English (GRE 332+5, Toefl 109)

Interests: Marathon, Chinese Folk Dance