LI, Haoran

The Hong Kong University of Science and Technology
PhD Student of Computer Science
Google Scholar: [Link]

Phone: (852)91449740
Email: hlibt@connect.ust.hk
Homepage: https://hlibt.student.ust.hk/

EDUCATION

2020 - Present **The Hong Kong University of Science and Technology**
PhD in Computer Science
CGA: 3.900

2016 - 2020 **The Hong Kong University of Science and Technology**
BSc in Computer Science and Mathematics of Computer Science Track
GPA: 3.814

RESEARCH INTEREST

Privacy for NLP: privacy attacks and defenses on (Large) Language Models.

Federated Learning for NLP: raw data privacy protection during cross-silo Federated Learning.

PUBLICATIONS

Haoran Li*, Dadi Guo*, Wei Fan, Mingshi Xu, Jie Huang, Fanpu Meng, Yangqiu Song, Multi-step Jailbreaking Privacy Attacks on ChatGPT, To appear at Findings of EMNLP 2023.

Haoran Li, Mingshi Xu, Yangqiu Song, Sentence Embedding Leaks More Information than You Expect: Generative Embedding Inversion Attack to Recover the Whole Sentence, Findings of ACL 2023.

Haoran Li, Yangqiu Song, Lixin Fan, You Don't Know My Favorite Color: Preventing Dialogue Representations from Revealing Speakers' Private Personas, NAACL 2022 (Oral).

Haoran Li*, Ying Su*, Qi Hu, Jiaxin Bai, Yilun Jin, Yangqiu Song, FedAssistant: Dialog Agents with Two-side Modeling, FL-IJCAI'22 (International Workshop on Trustworthy Federated Learning in Conjunction with IJCAI 2022).

Hao Peng*, **Haoran Li***, Yangqiu Song, Vincent Zheng, Jianxin Li, Differentially Private Federated Knowledge Graphs Embedding, CIKM 2021 (oral).

Xuanchi Ren, **Haoran Li**, Zijian Huang, Qifeng Chen, Self-supervised Dance Video Synthesis Conditioned on Music, ACM MM 2020 (oral).

PREPRINTS

Haoran Li*, Dadi Guo*, Donghao Li*, Wei Fan, Qi Hu, Xin Liu, Chunkit Chan, Duanyi Yao, Yangqiu Song, P-Bench: A Multi-level Privacy Evaluation Benchmark for Language Models. Arxiv preprint, 2023.

Haoran Li*, Yulin Chen*, Jinglong Luo*, Yan Kang, Xiaojin Zhang, Qi Hu, Chunkit Chan, Yangqiu Song, Privacy in Large Language Models: Attacks, Defenses and Future Directions. Arxiv preprint, 2023.

Qi Hu, **Haoran Li**, Jiaxin Bai, Yangqiu Song. Privacy-Preserving Neural Graph Databases. Arxiv preprint, 2023.

PATENTS

Haoran Li, Yangqiu Song, Lixin Fan, Defending Chatbots from Black-Box Persona Inference Attacks, *TTC.PA*.01606, *P*2279*US*00, 2022

Haoran Li, Ying Su, Qi Hu, Jiaxin Bai, Yilun Jin, Yangqiu Song, FedAssistant: Dialog Agents with Two-side Modeling, *TTC.PA*.1560, *P*2139*US*00, 2021

INTERNSHIP

April - August, 2022 Research internship. ByteDance AI Lab, Shanghai. Work on tuning language models to solve math word problems for tutoring junior students.

TEACHING

Teaching Assistant Coordinator of CSE, HKUST (2023-2024).

Teaching Assistant of COMP4332/RMBI4310 Big Data Mining at HKUST (Spring 2021, 2022).

Teaching Assistant of COMP2011 Programming with C++ at HKUST (Fall 2022-23).

EXCHANGE

2019.01 - 2019.05 **University of Waterloo, Canada**
Computer Science exchange student with grade 87.25 for the winter semester, 2019.

AWARDS/HONORS/SCHOLARSHIPS

HKUST RedBird Academic Excellence Award for Continuing PhD Students in 2021-2023.

PGS Studentship, 2020-present.

Dean's list for School of Engineering, 2017 and 2018.

HKUST Academic Excellence Honor, 2018-2020.

University's Scholarship Scheme for Continuing Undergraduate Students, HKUST, 2017 and 2018.

HKUST Study Abroad Sponsorship, 2018.