JIE HUANG

✉ jeffrey.huangjie@gmail.com | ⌂ https://jeffhj.github.io/

EDUCATION

University of Illinois at Urbana-Champaign   Urbana, USA
Ph.D. in Computer Science   Aug. 2020 - May 2024

- Advisor: Prof. Kevin C.C. Chang

Sun Yat-sen (Zhongshan) University   Guangzhou, China
B.Eng. in Computer Science & Technology   Aug. 2016 - June 2020

EMPLOYMENT

xAI   Palo Alto, USA
Member of Technical Staff   July 2024 -

- Invent and initialize many core recipes & data, across posttraining, midtraining, reasoning, and multimodal, covering domains including but not limited to general instruction following, general reasoning/thinking, knowledge update, expert domain, video understanding/captioning, reward modeling and RLHF.
- Grok-2/2.5: Co-contributor to posttraining/midtraining [blog 1] [blog 2]. Lead posttraining of the open-source Grok 2.5 model 🍍 [xai-org/grok-2].
- Grok-3: Lead the development of the first-ever model to break 1400 score and achieve #1 across all categories in LMSYS Chatbot Arena 🍫 [blog] [X post].
- Grok-3.5/4: Co-lead reasoning posttraining 🫐 [blog].
- Video Gen 🎬: Lead video agent, captioning, RL.

Google DeepMind   Kirkland/Mountain View, USA
Research Intern   May 2023 - Oct. 2023

- Contributor to the early efforts on Gemini factuality. Build early prototypes for search agent.
- Reasoning and factuality research: Self-Correction, SAFE.

NVIDIA   *Remote*, USA
Research Intern   Jan. 2023 - May 2023

- Develop RAVEN: a SOTA retrieval-augmented encoder-decoder model with in-context learning capability.

PUBLICATIONS

Google Scholar: `https://scholar.google.com/citations?user=GIoPkMoAAAAJ&hl=en`

(*equal contribution, †mentored student)

31. Xinyu Lian*, Yinfang Chen*, Runxiang Cheng, **Jie Huang**, Parth Thakkar, Minjia Zhang, Tianyin Xu. Large Language Models as Configuration Validators. The 47th IEEE/ACM International Conference on Software Engineering (**ICSE**). 2025. [code]

30. **Jie Huang**, Wei Ping, Peng Xu, Mohammad Shoeybi, Kevin C.C. Chang, Bryan Catanzaro. RAVEN: In-Context Learning with Retrieval-Augmented Encoder-Decoder Language Models. The 1st Conference on Language Modeling (**COLM**). 2024. [code]

29. **Jie Huang***, Xinyun Chen*, Swaroop Mishra, Huaixiu Steven Zheng, Adams Wei Yu, Xinying Song, Denny Zhou. Large Language Models Cannot Self-Correct Reasoning Yet. The Twelfth International Conference on Learning Representations (**ICLR**). 2024. [media coverage 1] [2] [3] [4] [Top of Hacker News]

28. Ziyi Chen†, Xiaocong Yang, Jiacheng Lin, Chenkai Sun, Kevin C.C. Chang, **Jie Huang**. Cascade Speculative Drafting for Even Faster LLM Inference. The 38th Conference on Neural Information Processing Systems (**NeurIPS**). 2024. [code]

27. Jerry Wei*, Chengrun Yang*, Xinying Song*, Yifeng Lu*, Nathan Hu, **Jie Huang**, Dustin Tran, Daiyi Peng, Ruibo Liu, Da Huang, Cosmo Du, Quoc V. Le. Long-form Factuality in Large Language Models. The 38th Conference on Neural Information Processing Systems (**NeurIPS**). 2024. [code]

26. Yifan Song, Da Yin, Xiang Yue, **Jie Huang**, Sujian Li, Bill Yuchen Lin. Trial and Error: Exploration-Based Trajectory Optimization of LLM Agents. The 62nd Annual Meeting of the Association for Computational Linguistics (**ACL**). 2024. [code]

25. **Jie Huang**, Kevin C.C. Chang. Citation: A Key to Building Responsible and Accountable Large Language Models. Findings of the 2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL Findings**). 2024.

24. Hanyin Shao*†, **Jie Huang***, Shen Zheng†, Kevin C.C. Chang. Quantifying Association Capabilities of Large Language Models and Its Implications on Privacy Leakage. Findings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (**EACL Findings**). 2024. [code]

23. Nishant Balepur†, **Jie Huang**, Kevin C.C. Chang. Expository Text Generation: Imitate, Retrieve, Paraphrase. The 2023 Conference on Empirical Methods in Natural Language Processing (**EMNLP**). 2023. [code]

22. Nishant Balepur†, **Jie Huang**, Kevin C.C. Chang. Text Fact Transfer. The 2023 Conference on Empirical Methods in Natural Language Processing (**EMNLP**). 2023. [code]

21. **Jie Huang**, Kevin C.C. Chang. VER: Unifying Verbalizing Entities and Relations. Findings of the 2023 Conference on Empirical Methods in Natural Language Processing (**EMNLP Findings**). 2023. [code]

20. Haoran Li*, Dadi Guo*, Wei Fan, Mingshi Xu, **Jie Huang**, Fanpu Meng, Yangqiu Song. Multi-step Jailbreaking Privacy Attacks on ChatGPT. Findings of the 2023 Conference on Empirical Methods in Natural Language Processing (**EMNLP Findings**). 2023. [code]

19. Yuxiang Liu, **Jie Huang**, Kevin C.C. Chang. Ask To The Point: Open-Domain Entity-Centric Question Generation. Findings of the 2023 Conference on Empirical Methods in Natural Language Processing (**EMNLP Findings**). 2023. [code]

18. Kerui Zhu†, **Jie Huang**, Kevin C.C. Chang. Descriptive Knowledge Graph in Biomedical Domain. The 2023 Conference on Empirical Methods in Natural Language Processing (**EMNLP Demo**). 2023. [demo]

17. **Jie Huang**, Kevin C.C. Chang. Towards Reasoning in Large Language Models: A Survey. Findings of the 61st Annual Meeting of the Association for Computational Linguistics (**ACL Findings**). 2023. [paperlist]

16. Chenzhengyi Liu*†, **Jie Huang***, Kerui Zhu†, Kevin C.C. Chang. DimonGen: Diversified Generative Commonsense Reasoning for Explaining Concept Relationships. The 61st Annual Meeting of the Association for Computational Linguistics (**ACL**). 2023. [code]

15. **Jie Huang**, Kevin C.C. Chang, Jinjun Xiong, Wen-mei Hwu. Can Language Models Be Specific? How?. Findings of the 61st Annual Meeting of the Association for Computational Linguistics (**ACL Findings**). 2023. [code]

14. **Jie Huang**, Yifan Gao, Zheng Li, Jingfeng Yang, Yangqiu Song, Chao Zhang, Zining Zhu, Haoming Jiang, Kevin C.C. Chang, Bing Yin. CCGen: Explainable Complementary Concept Generation in E-Commerce. 2023.

13. Shen Zheng*†, **Jie Huang***, Kevin C.C. Chang. Why Does ChatGPT Fall Short in Providing Truthful Answers?. I Can't Believe It's Not Better Workshop (**ICBINB@NeurIPS**). 2023.

12. Zining Zhu, Haoming Jiang, Jingfeng Yang, Sreyashi Nag, Chao Zhang, **Jie Huang**, Yifan Gao, Frank Rudzicz, Bing Yin. Situated Natural Language Explanations. Natural Language Reasoning and Structured Explanation Workshop (**NLRSE@ACL**). 2023.

11. **Jie Huang***, Kerui Zhu*†, Kevin C.C. Chang, Jinjun Xiong, Wen-mei Hwu. DEER: Descriptive Knowledge Graph for Explaining Entity Relationships. The 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP**). 2022. [code]

10. **Jie Huang**, Hanyin Shao†, Kevin C.C. Chang, Jinjun Xiong, Wen-mei Hwu. Understanding Jargon: Combining Extraction and Generation for Definition Modeling. The 2022 Conference on

Empirical Methods in Natural Language Processing (**EMNLP**). 2022. [code]

9. **Jie Huang***, Hanyin Shao*†, Kevin C.C. Chang. Are Large Pre-Trained Language Models Leaking Your Personal Information?. Findings of the 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP Findings**). 2022. [code] [media coverage 1] [2] [interview from MIT Tech Review]

8. Pritom Saha Akash, **Jie Huang**, Kevin C.C. Chang. Coordinated Topic Modeling. The 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP**). 2022. [code]

7. Fanghua Ye, Xi Wang, **Jie Huang**, Shenghui Li, Samuel Stern, Emine Yilmaz. MetaASSIST: Robust Dialogue State Tracking with Meta Learning. The 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP**). 2022. [code]

6. **Jie Huang**, Kevin C.C. Chang, Jinjun Xiong, Wen-mei Hwu. Open Relation Modeling: Learning to Define Relations between Entities. Findings of the 60th Annual Meeting of the Association for Computational Linguistics (**ACL Findings**). 2022. [code]

5. Pritom Saha Akash, **Jie Huang**, Kevin C.C. Chang, Yunyao Li, Lucian Popa, ChengXiang Zhai. Domain Representative Keywords Selection: A Probabilistic Approach. Findings of the 60th Annual Meeting of the Association for Computational Linguistics (**ACL Findings**). 2022. [code]

4. **Jie Huang**, Kevin C.C. Chang, Jinjun Xiong, Wen-mei Hwu. Measuring Fine-Grained Domain Relevance of Terms: A Hierarchical Core-Fringe Approach. The 59th Annual Meeting of the Association for Computational Linguistics (**ACL**). 2021. [code]

3. **Jie Huang***, Zilong Wang*, Kevin C.C. Chang, Wen-mei Hwu, Jinjun Xiong. Exploring Semantic Capacity of Terms. The 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP**). 2020. [code]

2. **Jie Huang**, Chuan Chen, Fanghua Ye, Weibo Hu, Zibin Zheng. Nonuniform Hyper-Network Embedding with Dual Mechanism. ACM Transactions on Information Systems (**TOIS**). 2020. [code]

1. **Jie Huang**, Xin Liu, Yangqiu Song. Hyper-Path-Based Representation Learning for Hyper-Networks. The 28th ACM International Conference on Information and Knowledge Management (**CIKM**). 2019. [code]

TALKS & INTERVIEWS

Large Language Models Cannot Self-Correct Reasoning Yet
Cohere for AI. Jan. 2024
AI TIME. Jan. 2024

Navigating Cognitive and Ethical Boundaries of Large Language Models
UIUC NLP Talk Series. Nov. 2023

RAVEN: In-Context Learning with Retrieval Augmented Encoder-Decoder Language Models
IBM Research. Sept 2023

Towards Responsible and Accountable Large Language Models:
Navigating Privacy Risks and Ethical Challenges
University of Toronto. July 2023

Knowledge, Reasoning & Ethics in (Large) Language Models
UIUC ML Reading Group. April 2023

Towards Reasoning in Large Language Models
NVIDIA NLP Reading Group. March 2023
Tencent AI Lab. Feb. 2023

Are Large Pre-Trained Language Models Leaking Your Personal Information?
Interview from MIT Technology Review. July 2022

PROFESSIONAL SERVICES

Area Chair / Action Editor

- ACL 2024; NAACL 2024; EMNLP 2024; ACL Rolling Review (2023-)

Program Committee Member / Reviewer

- ICLR 2025; NeurIPS 2024; ACL (2022-2023); EMNLP (2022-2023); NAACL 2022; AAAI (2023-2024); EACL (2023-2024); LoG 2022; ACL Rolling Review (2021-2023)
- EMNLP 2023 Industry Track
- AAAI 2024 Workshop on Privacy-Preserving Artificial Intelligence
- NeurIPS 2023 Workshop on Socially Responsible Language Modelling Research
- ACL 2023 Workshop on Natural Language Reasoning and Structured Explanations
- ICML 2023 Workshop on Generative AI and Law
- ICML 2022 Pre-training Workshop
- IEEE/ACM Transactions on Audio, Speech, and Language Processing
- ACM Computing Surveys

Community Services

- CS Graduate Ambassador, UIUC CS Visit Day for Prospective Graduate Students, 2023
- Chairman & Vice Chairman, Baidu Club, Sun Yat-sen University (2017-2019)

TEACHING

Co-Instructor of CS 598: Understanding and Using Large Language Models, Fall 2023, UIUC

MENTORED STUDENTS

Nishant Balepur (UIUC BS → UMD PhD), 2022-2023, Factuality in Text Generation

Chenzhengyi Liu (UIUC MSCS), 2022-2023, Commonsense Reasoning

Hanyin Shao (UIUC ECE MS), 2021-2023, Privacy in Language Models

Shen Zheng (UIUC MSCS), 2022-2024, LLM Evaluation, LLM Agents

Kerui Zhu (UIUC MSCS → UIUC PhD), 2021-2024, Knowledge Graph, Retrieval-Augmented LLMs

SELECTED HONORS & AWARDS

Yunni & Maxine Pao Memorial Fellowship, The Grainger College of Engineering	2023-2024
Outstanding Graduation Thesis, Sun Yat-sen University	2020
Outstanding Graduate in National Program for Top-Notch Undergraduate in Basic Sciences	2020
National Scholarship (Top 1%)	2017,2019
CCF Elite Collegiate Award, China Computer Federation	2018
Third Prize, 16th National Conference on Software and Applications Prototype System Competition (Free-Style Category)	2017
First Prize, Chinese Chemistry Olympiad (CChO, Provincial Competition Area, rank 8)	2015