



YONG CAO

🏠 University of Tübingen, Germany 📞 +49 152-5244-2031

✉️ yongcao2018@gmail.com 🔗 yongcaoplus.github.io 📅 Update 29th May 2026

Research Area and Interests

NLP & LLMs · Foundation-model Post-training · AI for Science · Multimodal Understanding · LLMs for Social Science (Cultural Alignment, Survey Simulation), with focus on interdisciplinary, domain-driven applications.

Education

University of Tübingen

Postdoctoral Researcher, Supervisor: Andreas Geiger.

Sep 2024 – Now

Tübingen, Germany

University of Copenhagen

Visiting Ph.D. student, Supervisor: Daniel Hershcovich.

Nov 2022 – Apr 2024

Copenhagen, Denmark

Huazhong University of Science and Technology

Ph.D. student of Computer Science, Supervisor: Min Chen.

Sep 2018 – Jun 2024

Wuhan, China

Chinese University of Hong Kong, Shenzhen

Visiting Ph.D student, Supervisor: Kai Hwang.

Jul 2020 – Nov 2020

Shenzhen, China

Sichuan University

Bachelor of Telecommunication Engineering, Rank: 1/60 (1%).

Sep 2014 – Jun 2018

Sichuan, China

Experience

Alibaba-DAMO

Intern of NLP Algorithm Researcher

Jun 2022 – Oct 2022

Hangzhou, China

Xiaomi

Intern of NLP Algorithm Researcher

Nov 2021 – Jun 2022

Wuhan, China

Deepwisdom

Intern of NLP Algorithm Engineering

Jan 2021 – Jun 2021

Shenzhen, China

Published Paper

Citations: 1118. h-index: 16. For a full list, see <https://scholar.google.com/citations?user=v8WGMQAAAAAJ>.

1. Chuqiao Li, Xianghui Xie, **Yong Cao**, Andreas Geiger, Gerard Pons-Moll. “FrankenMotion: Part-level Human Motion Generation and Composition”. *CVPR*, 2026. **Supervision**
2. Steffen Eger, **Yong Cao**, Jennifer D’Souza, Andreas Geiger, Christian Greisinger, Stephanie Gross, Yufang Hou, Brigitte Krenn, Anne Lauscher, Yizhi Li, Chenghua Lin, Nafise Sadat Moosavi, Wei Zhao, Tristan Miller. “Transforming Science with Large Language Models: A Survey on AI-assisted Scientific Discovery, Experimentation, Content Generation, and Evaluation”. *ACM Computing Surveys*, 2026.
3. Bolei Ma, **Yong Cao**, Indira Sen, Anna-Carolina Haensch, Frauke Kreuter, Barbara Plank, Daniel Hershcovich. “Too Open for Opinion? Embracing Open-Endedness in Large Language Models for Social Simulation”. *EACL main*, 2026. **Oral**
4. **Yong Cao**, Haijiang Liu, Arnav Arora, Isabelle Augenstein, Paul Röttger, Daniel Hershcovich. “Specializing Large Language Models to Simulate Survey Response Distributions for Global Populations”. *NAACL Main*, 2025. **Oral** **First**

5. Markus Flicke, Glenn Angrarbeit, Madhav Iyengar, Vitalii Protsenko, Illia Shakun, Jovan Cicvaric, Bora Kargi, Haoyu He, Lukas Schuler, Lewin Scholz, Kavyanjali Agnihotri, **Yong Cao**, Andreas Geiger. “Scholar Inbox: Personalized Paper Recommendations for Scientists”. *ACL Demo*, 2025. **Supervision**
6. Zhigen Li, Jianxiang Peng, Yanmeng Wang, **Yong Cao**, Tianhao Shen, Minghui Zhang, Linxi Su, Shang Wu, Yihang Wu, YuQian Wang, Ye Wang, Wei Hu, Jianfeng Li, Shaojun Wang, Jing Xiao, Deyi Xiong. “ChatSOP: An SOP-Guided MCTS Planning Framework for Controllable LLM Dialogue Agents”. *ACL main*, 2025.
7. Li Zhou, Taelin Karidi, Wanlong Liu, Nicolas Garneau, **Yong Cao**, Wenyu Chen, Haizhou Li, Daniel Hershcovich. “Does mapo tofu contain coffee? probing llms for food-related cultural knowledge”. *NAACL main*, 2025.
8. Haijiang Liu, Qiyuan Li, Chao Gao, **Yong Cao**, Xiangyu Xu, Xun Wu, Daniel Hershcovich, Jinguang Gu. “Beyond Demographics: Enhancing Cultural Value Survey Simulation with Multi-Stage Personality-Driven Cognitive Reasoning”. *EMNLP Main*, 2025.
9. Haijiang Liu, **Yong Cao**, Xun Wu, Chen Qiu, Jinguang Gu, Maofu Liu, Daniel Hershcovich. “Towards Realistic Evaluation of Cultural Value Alignment in Large Language Models: Diversity Enhancement for Survey Response Simulation”. *Information Processing & Management*, 2025.
10. **Yong Cao**, Ruixue Ding, Boli Chen, Xianzhi Li, Min Chen, Daniel Hershcovich, Pengjun Xie, Fei Huang. “Geo-Encoder: A Chunk-Argument Bi-Encoder Framework for Chinese Geographic Re-Ranking”. *EACL Main*, 2024. **Oral** **First**
11. **Yong Cao**, Yova Kementchedjhieva, Ruixiang Cui, Antonia Karamolegkou, Li Zhou, Megan Dare, Lucia Donatelli, Daniel Hershcovich. “Cultural Adaptation of Recipes”. *TACL*, 2024. **Oral** **First**
12. **Yong Cao**, Min Chen, Daniel Hershcovich. “Bridging Cultural Nuances in Dialogue Agents through Cultural Value Surveys”. *Findings of EACL*, 2024. **First**
13. Zhengdao Li, **Yong Cao**, Kefan Shuai, Yiming Miao, Kai Hwang. “Rethinking the Effectiveness of Graph Classification Datasets in Benchmarks for Assessing GNNs”. *IJCAI*, 2024. **Oral**
14. Li Zhou, Wenyu Chen, **Yong Cao**, et al.. “MLPs Compass: What is Learned when MLPs are Combined with PLMs?”. *ICASSP*, 2024.
15. Antonia Karamolegkou, Phillip Rust, **Yong Cao**, Ruixiang Cui, Anders Søgaard, Daniel Hershcovich. “Vision-Language Models under Cultural and Inclusive Considerations”. *HuCLLM Workshop @ ACL*, 2024.
16. **Yong Cao**, Xianzhi Li, Huiwen Liu, Wen Dai, Shuai Chen, Bin Wang, Min Chen, Daniel Hershcovich. “Pay More Attention to Relation Exploration for Knowledge Base Question Answering”. *Findings of ACL*, 2023. **First**
17. **Yong Cao**, Li Zhou, Seolhwa Lee, Laura Cabello, Min Chen, Daniel Hershcovich. “Assessing Cross-Cultural Alignment between ChatGPT and Human Societies: An Empirical Study”. *Workshop on Cross-Cultural Considerations in NLP (C3NLP)*, *EACL*, 2023. **Oral** **First**
18. Li Zhou, Laura Cabello, **Yong Cao**, Daniel Hershcovich. “Cross-cultural transfer learning for Chinese offensive language detection”. *Workshop on Cross-Cultural Considerations in NLP (C3NLP)*, *EACL*, 2023. **Oral**
19. Tianshu Hao, Jianfeng Zhan, Kai Hwang, **Yong Cao**. “Edge AiBench: Scenario-Based AI Benchmarking for Cloud/Edge/Device Computing”. *IEEE Transactions on Computers*, 2023.
20. **Yong Cao**, Wei Li, Xianzhi Li, Min Chen, Guangyong Chen, Long Hu, Zhengdao Li, Hwang Kai. “Explore More Guidance: A Task-aware Instruction Network for Sign Language Translation Enhanced with Data Augmentation”. *Findings of NAACL*, 2022. **First**.
21. Tarik Alfif, Bander Alzahrani, **Yong Cao**, Reem Alotaibi, Ahmed Barnawi, Min Chen. “Generative Adversarial Network Based Abnormal Behavior Detection in Massive Crowd Videos: A Hajj Case Study”. *Journal of Ambient Intelligence and Humanized Computing*, 2021.

22. M. Chen, Y. Jiang, **Yong Cao**, A. Y. Zomaya. “CreativeBioMan: Brain and Body Wearable Computing based Creative Gaming System”. *IEEE Systems, Man, and Cybernetics Magazine*, 2020.
23. **Yong Cao**, R. Wang, M. Chen, A. Barnawi. “AI Agent in Software-defined Network: Agent-based Network Service Prediction and Wireless Resource Scheduling Optimization”. *IEEE Internet of Things Journal*, 2019. **First**
24. Rui Wang, **Yong Cao**, Adeeb Noor, Thamer A. Alamoudi, Redhwan Nour. “Agent-enabled Task Offloading in UAV-aided Mobile Edge Computing”. *Computer Communications*, 2019.
25. M. Chen, **Yong Cao**, R. Wang, Y. Li, D. Wu, Z. Liu. “DeepFocus: Deep Encoding Brainwaves and Emotions with Multi-scenario Behavior Analytics for Human Attention Enhancement”. *IEEE Network*, 2019.
26. Jun Yang, Jiayi Lu, Yiming Miao, Lu Wang, Yiting Zhao, **Yong Cao**. “The Effective Recycling of Crashed Drone Based on Machine Intelligence”. *IWCMC*, 2018.

Preprints

1. **Yong Cao**, Markus Flicke, Haoyu He, Katrin Renz, Andreas Geiger. “Modeling Multi-Scale Scientific Impact via Heterogeneous Networks and LLMs”. *arXiv*, 2026. **First**
2. **Yong Cao**, Chuqiao Li, Xianghui Xie, Gerard Pons-Moll, Andreas Geiger. “NextMotionQA: Benchmarking and Judging Human Motion Understanding with Vision-Language Models”. *arXiv*, 2026. **First**
3. Jonathan von Rad, **Yong Cao**, Andreas Geiger. “UniComp: A Unified Evaluation of Large Language Model Compression via Pruning, Quantization and Distillation”. *arXiv*, 2026. **Supervision**
4. Haoyu He, Katrin Renz, **Yong Cao**, Andreas Geiger. “MDPO: Overcoming the Training-Inference Divide of Masked Diffusion Language Models”. *arXiv*, 2025. **Supervision**
5. **Yong Cao**, Wenyan Li, Jiaang Li, Yifei Yuan, Daniel Hershcovich. “Exploring Visual Culture Awareness in GPT-4V: A Comprehensive Probing”. *arXiv*, 2024. **First**

Academic Activities

Area Chair: ACL 2025, EMNLP 2025, ACL ARR.

Co-Organizer: Cross-Cultural Considerations in NLP workshop @ ACL 2024, @ NAACL 2025, @ ACL 2026.

Program Committee: AAAI, EACL, ACL ARR, CoLM, NeurIPS, ICLR, CVPR.

Media: “Assessing Cross-Cultural Alignment between ChatGPT and Human Societies: An Empirical Study” was picked up by Politiken, Børsen, Ekstra Bladet, P1 Morgen, TV 2 (Denmark) and Science et Avenir (France).

Talk

2026.03, “AI-assisted Scientific Discovery, Experimentation, Content Generation, and Evaluation”, EACL 2026, Tutorial, Morocco.

2025.06, “Cultural Reliability of LLMs: Can AI Truly Reflect Global Perspectives?”, WDMD Workshop @ DSN, Naples, Italy.

2024.09, “Cultural Considerations in NLP”, Xiaomi, online.

2024.03, “Cultural Considerations in Large Language Models”, Max Planck Institute for Human Development, Berlin Germany.

2024.02, “Cultural Considerations in Dialogue Systems”, University of Marburg, Germany.

2023.10, “Cultural Adaptation of Large Language Models”, University of Copenhagen, Denmark.

Teaching

2026.09 Course: Deep Learning, University of Tübingen (Upcoming).

2025.10 Seminar: LLM Research Assistants, University of Tübingen.

2025.10 Bachelor Team Project, University of Tübingen.

2025.07 AI for Science, tutorial, IT:U Summer School, Linz, Austria.

2019.01 Big Data and Internet of Things, Co-lecturer with Prof. Min Chen, Huazhong University of Science and Technology.

Projects

Human Motion Understanding | *University of Tübingen* **Sep 2025 - Now**

- Systematically probe and benchmark how well Vision-Language Models (VLMs) understand 3D human motion, designing fine-grained evaluation protocols that diagnose their capabilities and failure modes across motion semantics, temporal dynamics, and spatial reasoning.
- Propose to leverage VLMs for synthesizing out-of-distribution motion data as training-time augmentation, improving the robustness and generalization of downstream motion understanding and generation models.

Scholar Inbox | *University of Tübingen* **Sep 2024 - Now**

- Design LLM/VLM-based methods to parse and understand multimodal scientific documents, powering downstream academic services including personalized paper recommendation, user attribute analysis, and trending-topic prediction.
- Develop downstream academic tasks on top of these representations, such as scientific impact prediction, scientific question answering, and idea generation.

Cultural Adaption in LLMs | *University of Copenhagen* **Nov 2022 - Apr 2024**

- Investigate cultural differences in LLMs, quantifying the cultural alignment between LLMs and human societies based on human society surveys.
- Construct cultural adaptation benchmark datasets, and propose solutions for cultural adaptation in specific domains (e.g. recipe) and general domains (e.g. dialogue).

Modeling of Abnormal Behavior of Large-Scale Crowd | *Collaboration with KAUST* **Dec 2019 - Dec 2022**

- Collect large-scale crowd abnormal behavior benchmark dataset on hajj scenario.
- Design abnormal behavior classification algorithm based on the optical flow algorithm and GAN model.

Product of Sequence Tagging Based on AutoML | *Deepwisdom* **Mar 2021 - May 2021**

- Survey state-of-the-art AutoML and sequence-tagging approaches to inform system design.
- Develop sequence tagging product based on AutoML, and automatically construct the task pipeline in four stages: Data EDA, offline training, testing and online prediction. The algorithm is implemented based on PyTorch and Keras framework and is deployed through Docker.
- Expand the basic operators of the model to 30+, and evaluate the product performance based on 10+ benchmark.
- Seven basic NLP SaaS services are developed based on this product.

Resume Parsing System for Multi-Source Unstructured Data | *Deepwisdom* **May 2021 - Jun 2021**

- Build PDF parsing and word parsing operator based on the company's resume data, extract information based on entity extraction algorithm and rule matching, and establish a resume parsing prototype system.
- Optimize specific rules for specific field extraction (eg. educational background and work experience), and merge the algorithm into the company's platform.

Early Warning System and Intervention Strategy for Depression | *National key R&D plan* **Nov 2018 - Jun 2021**

- Develop collection and storage scheme of multi-modal dataset from real depression patients and normal volunteers.
- The depression diagnosis model was established based on EEG signal, near-infrared signal, video and audio data, and the depression diagnosis result was realized based on prediction fusion.
- Develop intelligent follow-up system, and construct the development of psychological counseling and intervention robot.

Honors and Awards

- 2025 DDSA Visiting Grant, Denmark
- 2023 DAAD AInet Fellow on Human-centered AI, Germany.
- 2023 Outstanding PhD Scholarship of HUST.
- 2021 International Youth Talent Fund by Zhejiang Lab, Hangzhou, China.
- 2020 Zhixing Scholarship of HUST.
- 2019 Outstanding Student of HUST.
- 2018 Outstanding Graduates of SCU.
- 2017,2016 National Scholarship.
- 2017 Outstanding student cadre of SCU.
- 2016 Excellent Paper Award in the National Mathematical Modeling Challenge.
- 2015 Outstanding Students of SCU.