BANGZHENG LI

◇ bzhli@ucdavis.edu ◇ Google Scholar

1 Shields Ave, Davis, CA 95616

EDUCATION

University of California, Davis, Ph.D. in Computer Science, 2024 – Now
University of Southern California, Ph.D. in Computer Science 2022 – 2023
University of Illinois, Urbana-Champaign, B.S. in Mathematics and Computer Science 2018 – 2020
Beihang University, B.S. in Mathematics 2016 – 2018

SUMMARY OF QUALIFICATIONS

- **Programming skills**: Python, LaTeX, C/C++, Java, HTML, CSS
- **Developer Tools**: Git, Docker, Google Cloud Platform, VS Code, Visual Studio, PyCharm
- **Frameworks**: PyTorch, Tensorflow, OpenCV, Scikit-Learn, HuggingFace

EXPERIENCE

Research Assistant, University of California, Davis,, *Sacramento, CA* 2024 - Now

- **LUKA Lab** (advisor: Muhao Chen)
- Lead a research team in developing a Large Language Model (LLM) specifically for meta-material science, designed to answer factual questions and generate scientific hypotheses for future research. Generated response quality is favored by domain experts over GPT-4's output.
- Designed a universal framework for Multi-modal Language Models (MMLMs) in visual question answering on image details, achieving comparable or superior performance with approximately 50% fewer visual tokens than alternative methods. This adaptable framework integrates seamlessly with existing MMLMs without the need for additional training.

Research Assistant, University of Southern California, *Los Angeles, CA* 2020 - 2023

- **LUKA LAB** (advisor: Muhao Chen)
- Proposed an evaluation benchmark to assess hallucination in Large Language Models (LLMs) during complex, multi-step reasoning processes. This benchmark standardizes human question-answering into structured reasoning chains, uncovering significant semantic biases in LLMs.
- Created a model training framework that reformulates classification tasks as instructional tasks in natural language, achieving a 5% improvement over the state-of-the-art in ultra-fine-grained entity typing and demonstrating strong transfer learning capabilities.

Research Assistant, University of Illinois, Urbana-Champaign, *Champaign, IL* 2019-2020

- **DMG** (advisor: Jiawei Han)
- Developed a COVID-19 literature NER system by implementing data collection, processing, model training, and back-end deployment. Processed a corpus of COVID-related scientific and news articles to perform named entity recognition (NER) and deliver a search service based on identified entities.

REWARDS

- **Provost's Fellowship, Viterbi School of Engineering, USC** 2021 - 2022

- **Dean's List, College of Liberal Arts & Sciences, UIUC** 2019 - 2020

PROFESSIONAL SERVICES

- **Member**: Association for Computational Linguistics (ACL) 2021 - Now
- **Reviewer**: AAAI, ACL, ACL Rolling Review, EMNLP, NAACL, IJCAI 2021 - Now

PUBLICATION LIST

Preprints

- **Bangzheng Li**, Fei Wang, Ben Zhou, Nan Xu, Sheng Zhang, Hoifung Poon, Muhao Chen. "Efficient Vision Language Modeling with Semantic-guided Visual Selection" in Submission to ICML, 2025

- **Bangzheng Li**, Ben Zhou, Xingyu Fu, Fei Wang, Dan Roth, Muhao Chen. "FamiCom: Further Demystifying Prompts for Language Models with Task-Agnostic Performance Estimation" in Submission to ACL, 2025

Conference & Journal Papers

- Xingyu Fu*, Yushi Hu*, **Bangzheng Li**, Yu Feng, Haoyu Wang, Xudong Lin, Dan Roth, Noah A. Smith, Wei-Chiu Ma, Ranjay Krishna. "BLINK : Multimodal Large Language Models Can See but Not Perceive" ECCV, 2024
  Spotlight presentation at Workshop cVinW@CVPR 2024
  Spotlight paper at Harmonious
  Over 36K Total downloads on HuggingFace
  HuggingFace Paper of the day

- **Bangzheng Li**, Ben Zhou, Fei Wang, Xingyu Fu, Dan Roth, Muhao Chen. "Deceiving Semantic Shortcuts on Reasoning Chains: How Far Can Models Go without Hallucination?" NAACL, 2024

- Tenghao Huang, Ehsan Qasemi, **Bangzheng Li**, He Wang, Faeze Brahman, Muhao Chen, Snigdha Chaturvedi. "Affective and Dynamic Beam Search for Story Generation" EMNLP, 2023

- **Bangzheng Li**, Wenpeng Yin, Muhao Chen. "Ultra-fine Entity Typing with Indirect Supervision from Natural Language Inference" TACL, 2022

- James Huang, **Bangzheng Li**, Jiashu Xu, Muhao Chen."Unified semantic typing with meaningful label inference" NAACL, 2022

- Nan Xu, Fei Wang, **Bangzheng Li**, Mingtao Dong, and Muhao Chen. "Does your model classify entities reasonably? diagnosing and mitigating spurious correlations in entity typing" EMNLP, 2022

- Qingyun Wang, Manling Li, Xuan Wang, Nikolaus Parulian, Guangxing Han, Jiawei Ma, Jingxuan Tu, Ying Lin, Ranran Haoran Zhang, Weili Liu, Aabhas Chauhan, Yingjun Guan, **Bangzheng Li**, Ruisong Li, Xiangchen Song, Yi Fung, Heng Ji, Jiawei Han, Shih-Fu Chang, James Pustejovsky, Jasmine Rah, David Liem, Ahmed ELsayed, Martha Palmer, Clare Voss, Cynthia Schneider, Boyan Onyshkevych. "COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation" NAACL, 2021
  Best Demo Paper

- Carl Yang, Jieyu Zhang, Haonan Wang, **Bangzheng Li**, Jiawei Han. "Neural Concept Map Generation for Effective Document Classification with Interpretable Structured Summarization" SIGIR, 2020

- Xuan Wang, Xiangchen Song, **Bangzheng Li**, Yingjun Guan, Jiawei Han. "Comprehensive named entity recognition on CORD-19 with distant or weak supervision" ISMB, 2020

- Xuan Wang, Xiangchen Song, **Bangzheng Li**, Kang Zhou, Qi Li, Jiawei Han. "Fine-Grained Named Entity Recognized Dataset of COVID-19 Literature" BIBM, 2020