Zhen Wang

: zhenwang9102@gmail.com
: https://zhenwang9102.github.io

RESEARCH INTERESTS

I am interested in empowering current AI systems with more explicit and human-understandable knowledge, aiming to make them more generalizable, interpretable, and efficient. My research lies in the nexus of natural language processing, deep learning, and data mining, where I study a "full stack" of the **knowledge-centric AI** from the ground up: acquisition, representation, transfer, and reasoning. My long-term research goal is to transfuse the strengths of human knowledge systems (e.g., intuitive physics, commonsense reasoning) to the next evolution of AI systems.

- **Knowledge Acquisition**: Structured information extraction from text and graphs, knowledge graph construction, knowledge distillation from language models
- **Knowledge Representation**: Word representation learning, graph embedding learning, graph neural networks, commonsense concept learning
- **Knowledge Transfer**: Transfer learning, multi-task learning, knowledge distillation, domain adaptation and generalization, few-shot learning
- **Knowledge Reasoning**: Multi-hop reasoning over text and graphs (KG reasoning, complex QA), neuro-symbolic reasoning, commonsense reasoning
- **Applications**: Natural language interfaces (dialogue systems, question answering), controllable text generation, clinical NLP (e.g., knowledge discovery on clinical notes), Bioinformatics

WORK EXPERIENCE

Postdoctoral Scholar — San Diego, CA, USA
Jointly Hosted by UCSD, MBZUAI and CMU — 2023-Present
Mentor: Eric Xing and Zhiting Hu

Independent Course Instructor — Columbus, OH, USA
OSU, Dept. of Computer Science & Engineering — Autumn 2022
CSE 5525: Foundations of Speech & Language Processing (Undergrad & Graduate)

Research Intern, MIT-IBM Watson AI Lab — Cambridge, MA, USA
Hosted by Rameswar Panda and Yoon Kim — Summer 2022
Topic: Multi-task prompt tuning for efficient adaptation of large language models

Research Intern, Microsoft Research Redmond — Redmond, WA, USA
Hosted by Nebojsa Jojic — Summer 2021
Topic: Coherence boosting and prompt calibration for GPT-3

Research Intern, NEC Labs America — Princeton, NJ, USA
Hosted by Bo Zong — Summer 2020
Topic: Learning word embeddings by commonsense knowledge reasoning

Graduate Research Associate — Columbus, OH, USA
OSU, Dept. of Computer Science & Engineering — 2018–2022
Advisor: Huan Sun

Graduate Teaching Assistant — Columbus, OH, USA
OSU, Dept. of Computer Science & Engineering — Autumn 2017
CSE 2111: Modeling and Problem Solving with Spreadsheets and Databases

EDUCATION

The Ohio State University (OSU) — Columbus, OH, USA
PhD, Dept. of Computer Science & Engineering — 2022
Advisor: Huan Sun
Thesis: Knowledge-centric NLP: Acquisition, Representation, Transfer, and Reasoning

Université Nice Sophia Antipolis (UNS) — Nice, France
Visiting Student, Polytech Nice Sophia — 2015
Individual Study Advisor: Lionel Fillatre and Michel Barlaud

China University of Petroleum (UPC) — Qingdao, China
B.Eng. in Electronic Information Engineering — 2015
Graduated as 1st-Ranked Student in the Major Overall Ranking

PUBLICATIONS (* equal contribution)

- **Multitask Prompt Tuning Enables Parameter-Efficient Transfer Learning**
  Zhen Wang, Rameswar Panda, Leonid Karlinsky, Rogerio Feris, Huan Sun, Yoon Kim
  *Proceedings of **ICLR 2023*** [Paper]
- **GPT Is Becoming a Turing Machine: Here Are Some Ways to Program It**
  Ana Jojic, Zhen Wang, Nebojsa Jojic
  [arXiv]
- **ThinkSum: Probabilistic Reasoning Over Sets Using Large Language Models**
  Batu Ozturkler, Nikolay Malkin, Zhen Wang, Nebojsa Jojic
  Under Review [arXiv]
- **Frustratingly Simple Entity Tracking with Effective Use of Multi-Task Learning Models**
  Janvijay Singh, Fan Bai, Zhen Wang
  *Proceedings of **EACL 2023*** [arXiv]
- **Knowledge Transfer between Structured and Unstructured Sources for Complex Question Answering**
  Lingbo Mo*, Zhen Wang*, Jie Zhao, Huan Sun
  NAACL 2022 Structured and Unstructured Knowledge Integration (SUKI) [Paper]
- **Learning Interpretable Word Representations by Commonsense Knowledge Reasoning**
  Zhen Wang, Bo Zong, Wei Cheng, Xuchao Zhang, Yanchi Liu, Wenchao Yu, Jingchao Ni, Haifeng Chen, Huan Sun
  Under Review
- **Coherence Boosting: When Your Pretrained Language Model is Not Paying Enough Attention**
  Nikolay Malkin, Zhen Wang, Nebojsa Jojic
  *Proceedings of **ACL 2022*** [Paper]
- **Bootstrapping a User-Centered Task-Oriented Dialogue System**
  Shijie Chen, Ziru Chen, Xiang Deng, Ashley Lewis, Lingbo Mo, Samuel Stevens, Zhen Wang, Xiang Yue, Tianshu Zhang, Yu Su, Huan Sun
  1st Proceedings of Alexa Prize TaskBot (Alexa Prize 2021) [Paper]
- **Modeling Context Pair Interaction for Pairwise Tasks on Graphs**
  Zhen Wang, Bo Zong, Huan Sun
  *Proceedings of **WSDM 2021*** (**Acceptance Rate: 18.6%, Long**) [Paper]

- **Rationalizing Medical Relation Prediction from Corpus-level Statistics**
  Zhen Wang, Jennifer Lee, Simon Lin, Huan Sun
  *Proceedings of **ACL 2020*** (**Acceptance Rate: 22.7%, Long**) [Paper]
- **SurfCon: Synonym Discovery on Privacy-Aware Clinical Data**
  Zhen Wang, Xiang Yue, Soheil Moosavinasab, Yungui Huang, Simon Lin, Huan Sun
  *Proceeding of **SIGKDD 2019*** (**Research Track, Acceptance Rate 14.2%, Oral**) [Paper]
- **Graph Embedding on Biomedical Networks: Methods, Applications, and Evaluations**
  Xiang Yue, Zhen Wang, Jingong Huang, Srinivasan Parthasarathy, Soheil Moosavinasab, Yungui Huang, Simon M. Lin, Wen Zhang, Ping Zhang, Huan Sun
  ***Bioinformatics**, Volume 36, Issue 4, 15 February 2020, Pages 1241–1251* (**Impact factor: 5.610**) [Paper]
- **A Comprehensive Study of StaQC for Deep Code Summarization**
  Jayavardhan Reddy Peddamail, Ziyu Yao, Zhen Wang, Huan Sun
  *Proceedings of **SIGKDD 2018*** (**Deep Learning Day, SPOTLIGHT**) [Paper]
- **Hessian Regularized Sparse Coding for Human Action Recognition**
  Weifeng Liu, Zhen Wang, Dapeng Tao, Jun Yu
  *Proceedings of **MMM 2015**, Sydney, Australia* [Paper]

HONORS AND AWARDS

- Third-Place Honor, Inaugural Alexa Prize TaskBot Challenge 2022
- Graduate Research Award, CSE, OSU 2022
- Graduate Student Research Poster Award (Top 5), CSE, OSU 2021
- SIGIR Student Travel Grant 2021
- **Rising Star in Data Science**, CDAC, University of Chicago 2020
- SIGKDD Student Travel Award 2019
- Excellent Bachelor Degree Thesis Award in Shandong Province, China 2016
- Excellent Graduate Thesis Award of UPC, China 2015
- **China Scholarship Council (CSC) Scholarship** 2015
  - A fully supported visiting program in Polytech Nice Sophia, Nice, France
- **National Scholarship**, China 2014
- Soong Ching Ling Foundation (SCLF) Scholarship, China 2013
- National Scholarship for Encouragement, China 2012

SERVICE

- Area Chair / Senior PC
  - NLPCC 2023
- Program Committee:
  - ACL Rolling Review
    - 2021 (Oct., Nov.)
    - 2022 (Jan., April, Sep., Oct., Dec.)
    - 2023 (Feb.)
  - NAACL (2021, 2022 SUKI Workshop)
  - EMNLP (2021, 2022)
  - ACL (2021, 2023)

  - ICML 2023
  - NeurIPS 2023
  - KDD 2023
  - AAAI 2023
  - NLPCC (2020, 2021, 2022)
- External Reviewer: KDD (2019, 2020), ACL 2018, ICDM 2018

TEACHING EXPERIENCE

- **CSE 5525: Foundations of Speech and Language Processing**
  ⊢ *OSU, Department of Computer Science & Engineering*
  Instructor, Autumn 2022 (~30 undergrad and graduate students)
- **Summer School Tutorial: Natural Language Processing in Deep Learning**
  ⊢ *OSU, Foundations of Data Science and AI Community of Practice*
  Panelist, June 1 - June 3, 2022 (>180 attendees)
- **CSE 2111: Modeling and Problem Solving with Spreadsheets and Databases**
  ⊢ *OSU, Department of Computer Science & Engineering*
  Teaching Assistant, Autumn 2017
- **Panel Discussion: 2001: A Space Odyssey - Science Fiction vs Science Fact**
  ⊢ *OSU, Department of Astronomy*
  Panelist, Feb. 2021

INVITED TALKS

- "Towards Efficient and Robust Practice of Large Language Models", Autumn 2022, UCSD & MBZUAI & CMU
- "Towards Efficient and Robust Practice of Large Language Models", Autumn 2022, University of Washington
- "Rationalizing Relation Prediction from Corpus-level Statistics", ACL 2020, July, Online
- "Modeling Context Pair Interaction for Pairwise Tasks on Graphs", WSDM 2021, March, Online
- "SurfCon: Synonym Discovery on Privacy-Aware Clinical Data", KDD 2019, August 6, Anchorage, Alaska, USA

REFERENCES

Available upon request