Junheng Hao

CONTACT INFORMATION	Microsoft Azure Cognitive Services Research *E-mail:* junhenghao@microsoft.com *Mobile:* +1 (424)355-5950 *Office:* Microsoft Bravern-2 Building, 11025 NE 8th St, Bellevue, WA 98004	*Homepage* *LinkedIn* *Google Scholar* *Semantic Scholar* *DBLP Profile*

CURRENT EMPLOYMENT

Researcher, Microsoft Bellevue, WA   Oct 2022 - Current
Knowledge and Language Team at Microsoft Azure Cognitive Services Research group

RESEARCH INTERESTS

Knowledge is power; knowledge graphs (KGs) are power to next-generation AI analytics. My research interests lie at the intersection of graph learning, data mining, natural language processing and machine learning with a focus on knowledge graphs, especially learning symbolic and semantic structures inside KGs and multimodal AI.

EDUCATION

University of California Los Angeles (UCLA), CA, United States   Sept 2022
Ph.D. in Computer Science
Thesis: Incorporating ontological information in knowledge graph learning and applications
Advisors: Yizhou Sun, Wei Wang
Area of Study: Knowledge graph, graph mining, natural language processing, machine learning, bioinformatics, recommender systems.
Labs: Scalable Analytics Institute (ScAi), UCLA Data Mining Lab

Tsinghua University, Beijing, China   May 2017
B. Eng. in School of Information Science and Technology
B. Sc (Econ, Minor) in School of Economics and Management

PROFESSIONAL EXPERIENCE (INDUSTRY)

Research Intern at Microsoft Research (MSR), Redmond, WA   June 2021 - Sept 2021
- **Mentors:** Chieh-Han Wu, Zhihong (Iris) Shen, Ye-Yi Wang, Jennifer Neville
- **Project:** KG-enhanced document representation learning
  - **Overview:** Enhancing document pretrained representations with infused document knowledge graph (DocKG), including Microsoft Academic Graph (Content + Graph).
  - **Deliverables:** One research technical preprint [13].

PhD Research Intern at IBM Research AI, San Jose, CA   June 2020 - Sept 2020
- **Mentor:** Chuan Lei, Berthold Reinwald, Fatma Ozcan
- **Project:** Ontology Matching by Utilizing Graph Neural Networks
  - **Overview:** Empowering hybrid graph neural networks (GNNs) for ontology matching between relational databases and standard healthcare ontologies.
  - **Deliverables:** One paper [1] published at KDD 2021 with corresponding invention filed and partially shipped in *IBM Micromedex solutions (Watson Health)*.

Applied Scientist Intern/Student Researcher at Amazon, Seattle, WA June 2019 - Dec 2019
- **Mentors:** Tong Zhao, Luna Xin Dong, Christos Faloutsos
- **Project:** Diversified Complementary Recommendation on Product Graph
  - **Overview:** Enabling diversified complementary recommendation from web-scale product graphs and hierarchical product ontology.
  - **Deliverables:** One paper [3] published at CIKM 2020 and deployed in *Amazon-wide product complementary recommendation engine*.

Research Intern at **NEC Lab America**, Princeton, NJ   June 2018 - Sept 2018
- **Mentors:** Lu-An Tang, Zhichun Li, Haifeng Chen
- **Project:** Enterprise Knowledge Graph Fusion and Adaptation

- **Overview:** Graph-based Multi-source graph knowledge transfer and network fusion on enterprise engines for malicious process detection.
- **Deliverables:** One research technical preprint [12] under review.

PUBLICATION SUMMARY

As of Sept. 2022: I have accomplished 10+ Papers published papers and submissions, among which I first-authored papers published on top-tier venues (KDD, CIKM, BCB) from interdisciplinary domains across machine learning, data mining and bioinformatics, with 260+ Citations. More up-to-date publication record can be found in *Google Scholar* and/or *Semantic Scholar*. Several research works and internship projects (such as P-Companion[3]) have been successfully deployed in market-scale product.

PUBLICATIONS, SUBMISSIONS & PREPRINTS

[1] *Multi-source Inductive Graph Knowledge Transfer*
Junheng Hao, Lu-An Tang, Yizhou Sun, Zhengzhang Chen, Haifeng Chen, Junghwan Rhee, Zhichun Li and Wei Wang.
Proceedings of Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML-PKDD). September 2022.

[2] *Metadata-Induced Contrastive Learning for Zero-Shot Multi-Label Text Classification*
Yu Zhang, Zhihong Shen, Chieh-Han Wu, Boya Xie, **Junheng Hao**, Ye-Yi Wang, Kuansan Wang and Jiawei Han.
Proceedings of The Web Conference (WWW) 2022. April 2022.

[3] *MEDTO: Medical Data to Ontology Matching using Hybrid Graph Neural Networks*
Junheng Hao, Chuan Lei, Abdul Quamar, Vasilis Efthymiou, Fatma Ozcan, Yizhou Sun, Wei Wang.
Proceedings of 27th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD, Applied Data Science Track). August 2021.

[4] *JEDI: Circular RNA Prediction based on Junction Encoders and Deep Interaction among Splice Sites*
Jyun-Yu Jiang, Chelsea J.-T. Ju, **Junheng Hao**, Muhao Chen, Wei Wang.
Proceedings of the 29th annual international conference on Intelligent Systems for Molecular Biology and the 20th annual European Conference on Computational Biology (ISMB-ECCB). September 2021.

[5] *P-Companion: Framework for Diversified Complementary Product Recommendation*
Junheng Hao, Tong Zhao, Jin Li, Luna Xin Dong, Christos Faloutsos, Yizhou Sun, Wei Wang.
Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM), Applied Research Track. October 2020.

[6] *Bio-JOIE: Joint Representation Learning of Biological Knowledge Bases*
Junheng Hao, Chelsea J.-T. Ju, Muhao Chen, Yizhou Sun, Carlo Zaniolo, Wei Wang.
Proceedings of The 11th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM BCB 2020), September 2020. **Best Student Paper Award**.

[7] *Universal Representation Learning of Knowledge Bases by Jointly Embedding Instances and Ontological Concepts*
Junheng Hao, Muhao Chen, Wenchao Yu, Yizhou Sun, Wei Wang.
Proceedings of 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD, Research Track). August 2019.

[8] *Accident Impact Analysis in Traffic Safety and Mobility Using Group Network Features*
Chen-Shuo Sun, Xin Pei, **Junheng Hao**, Zuo Zhang.
Transportation Research Part B: Methodological (TRB), 2018.

[9] *Normal/Abnormal Heart Sound Recordings Classification Using Convolutional Neural Network*
Tanachat Nilanon, Jiayu Yao, **Junheng Hao**, Yan Liu.
Proceedings of the 43rd Computing in Cardiology Conference (CinC). December 2016.

[10] *A Data Driven Approach for Evaluation of Urban Accident Impacts*
Chen-Shuo Sun, **Junheng Hao**, Xin Pei, Zuo Zhang.
Proceedings of IEEE Conference on Intelligent Transportation Systems (ITSC), December 2016.

[11] *A Mobility-Aware Deep Learning Model for Long-Term COVID-19 Pandemic Prediction and Policy Impact Analysis*
Danfeng Guo, Zijie Huang, **Junheng Hao**, Yizhou Sun, Wei Wang, Demetri Terzopoulos.
Prerpint, under review.

[12] *MurderBookKG: Automatic Case Knowledge Graph Construction*
Junheng Hao, Jingyue Shen, P. Jeffrey Brantingham, Wei Wang. Prerpint, under review.

[13] *EHIA: Empowering Homicide Analytics with MurderBook Knowledge Graphs and Domain-specific Language Models*
Junheng Hao, Jingyue Shen, Craig D. Uchida, Yizhou Sun, P. Jeffrey Brantingham, Wei Wang. Prerpint, under review.

[14] *Modularized Paradigm For ML Application Development and Collaboration*
Junheng Hao, Yicong Huang, Alexander Taylor, Wei Wang, Chen Li.
Preprint.

[15] *Neighbor Aggregative Network Embedding*
Yunsheng Bai, Ruchi Jain, **Junheng Hao**, Yang Qiao, Zixia Weng, Yizhou Sun, Wei Wang.
Prerpint.

PATENT
APPLICATIONS

[16] *OntoGNN: Hybrid Graph Neural Networks for Ontology Matching.*
Chuan Lei, **Junheng Hao**, Vasilis Efthymiou, Fatma Ozcan, Abdul Quamar. U.S. Patent Application. (Sept. 2021)

ACADEMIC
SERVICES

Conference Program Committee / Reviewer

• ACL Rolling Review	2022
• Conference on Neural Information Processing Systems (NeurIPS)	2021,2022
• SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)	2019-2021
• International Conference on Machine Learning (ICML)	2020-2022
• AAAI Conference on Artificial Intelligence (AAAI)	2021-2023
• TheWebConf / International World Wide Web Conference (WWW)	2022
• International Conference on Learning Representations (ICLR)	2020-2022
• International Joint Conferences on Artificial Intelligence (IJCAI)	2020-2022
• IEEE International Conference on Data Engineering (ICDE)	2020
• Conference on Empirical Methods in Natural Language Processing (EMNLP)	2019
• SIAM International Conference on Data Mining (SDM)	2019-2022
• Conference on Information and Knowledge Management (CIKM)	2019-2022
• European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)	2019-2022

Program Committee Board

• International Joint Conferences on Artificial Intelligence (IJCAI)	2022-2024

Journal Reviewer
- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
- IEEE Transactions on Big Data (TBD)
- ACM Transactions on Intelligent Systems and Technology (TIST)
- ACM Transactions on Knowledge Discovery from Data (TKDD)

Conference Volunteer
- International Conference on Learning Representations (ICLR)   2021
- SIKKDD Conference on Knowledge Discovery and Data Mining (KDD),   2019, 2020
- Conference on Empirical Methods in Natural Language Processing (EMNLP)   2020, 2021
- Conference on Neural Information Processing Systems (NeurIPS)   2018, 2020

INVITED TALKS
- July 2022: Google Brain.
  Invited tech talk: *Recommendations on Documents, Products and More? A Knowledge Graph Approach*

- June 2022: Microsoft, Azure Cognitive Services Research Group.
  Invited tech talk: *Knowledge Augmented Applications: NLP, Bioinformatics and Recommendation*

- Dec 2021: Coupang, Ranking, Discovery and Personalization.
  Invited tech talk: *Knowledge Graphs Meets Product Recommendation: One Deep Learning Solution*

- Oct 2021: UCLA CS Data Science Seminar
  Graphs, Transformers, and When They Meet Biology (AlphaFold2). [Video] [Slides]

- Aug 2021: Microsoft Search, Assistant and Intelligence (MSAI).
  Graph learning session: *Knowledge Graph with Ontology Learning and Applications*. [Slides]

- July 2019: Amazon, Product Graph.
  Invited talk: *Representation Learning on Knowledge Graphs: Embedding, Logic Rules and Graph Neural Networks* (with Yizhou Sun). [Slides]

HONORS AND AWARDS
- **Best Student Paper Award (ACM BCB)**   2020
- NAACL D&I Subsidy Award   2022
- ACL DEI Award   2022
- SIGIR Student Travel Grant (CIKM)   2020
- Student Travel Award (KDD)   2019, 2020
- UCLA Graduate Division Fellowship   2018-2019

TEACHING
- CS M146: Introduction to Machine Learning (Instructor: Sriram Sankararaman).
  Teaching Associate / Head TA, Winter 2021   Evaluation Score: **8.2**/9.0

- CS145: Introduction to Data Mining (Instructor: Yizhou Sun)
  Teaching Associate / Head TA, Fall 2020   Evaluation Score: **8.0**/9.0

- CS32: Introduction to Computer Science II, Data Structures (Instructor: David Smallberg & Carey Nachenberg)
  Teaching Assistant, Spring 2019 & Winter 2019

- CS145: Introduction to Data Mining (Instructor: Yizhou Sun)
  Teaching Assistant, Fall 2018

EXTRACURRICULAR ACTIVITIES
- UCLA Volunteer Income Tax Assistance (VITA) Organization: *IRS-certified volunteer*
- UCLA Bruin Mental Health Advisory Committee: *Representative and advocate volunteer*
- UCLA GUM: Graduate-Undergraduate Mentorship Program: *STEM Graduate Mentor*

- UCLA CS PhD Mentorship Program: *Seinor PhD Mentor*
- Queer In AI: *Member and workshop volunteer*

SKILLS
- **Programming:** Python (PyTorch, TensorFlow), C/C++, LaTeX, MATLAB, SQL
- **Language:** Mandarin (Native), English (Proficient), Spanish (Basic)