Zifeng Wang

Co-Founder & Head of AI
Keiji AI Inc.

zifengwang.ai@gmail.com
https://zifeng.wang

EDUCATION

Ph.D. Computer Science, University of Illinois Urbana-Champaign, Urbana, USA, 2021-2025

M.S. Data Science, Tsinghua University, Beijing, China, 2018-2021

B.S. Engineering, Tongji University, Shanghai, China, 2014-2018

APPOINTMENTS

2025- Keiji AI
Co-Founder, Head of AI

2021-2025 University of Illinois Urbana-Champaign
Research Assistant, Supervisor: Jimeng Sun

2023 Amazon Web Service
Applied Scientist Intern, Mentor: Zichen Wang

2023 Medidata, Dassault Systèmes
Research Intern, Mentor: Mandis Beigi

2022 Amplitude
Research Intern, Mentor: Cao Xiao

2019-2021 Tencent
Research Intern, Mentor: Yefeng Zheng

HIGHLIGHTS

Research

I am dedicated to building artificial intelligence (AI) methods to accelerate scientific discovery in medicine, with a focus on AI-driven literature research [1] [2], data science research [3], and clinical research [4].

My research has been featured by Nature, NIH News, POLITICO, Azure Government, etc.

Startup

I cofounded Keiji AI since 2025, dedicated to building TrialMind, an AI agent platform for clinical trials [Demo Demo1, Demo2]. Our product is now used by leading pharmaceutical organizations, including Medidata, Regeneron, Takeda, Abbvie, Guardant Health, etc.

SELECTED PAPERS

The full publication list is available at Google Scholar.

Literature research

2025 **Wang, Z.**, Jin, Q., Lin, J., Gao, J., Pradeepkumar, J., Jiang, P., Danek, B., Lu, Z., Sun, J., "TrialPanorama: Database and Benchmark for Systematic Review and Design of Clinical Trials." In: *arXiv preprint arXiv:2505.16097* (2025)

2025 **Wang, Z.**, Cao, L., Jin, Q., Chan, J., Wan, N., Afzali, B., Cho, H.-J., Choi, C.-I., Emamverdi, M., Gill, M. K., "A foundation model for human-AI collaboration in medical literature mining." In: *arXiv preprint arXiv:2501.16255* (2025)

2024 **Wang, Z.**, Cao, L., Danek, B., Zhang, Y., Jin, Q., Lu, Z., Sun, J., "Accelerating Clinical Evidence Synthesis with Large Language Models." In: *arXiv preprint arXiv:2406.17755* (2024)

Data science research

2025 **Wang, Z.**, Danek, B., Sun, J., "BioDSA-1K: Benchmarking Data Science Agents for Biomedical Research." In: *arXiv preprint arXiv:2505.16100* (2025)

2024 **Wang, Z.**, Danek, B., Yang, Z., Chen, Z., Sun, J., "Can Large Language Models Replace Data Scientists in Biomedical Research?" In: *arXiv preprint* (2024)

Clinical research

2025 **Wang, Z.**, Gao, J., Danek, B., Theodorou, B., Shaik, R., Thati, S., Won, S., Sun, J., "Compliance and Factuality of Large Language Models for Clinical Research Document Generation." In: *arXiv preprint arXiv:2504.00934* (2025)

2024 Jin, Q., **Wang, Z.**, Floudas, C. S., Chen, F., Gong, C., Bracken-Clarke, D., Xue, E., Yang, Y., Sun, J., Lu, Z., "Matching Patients to Clinical Trials with Large Language Models." In: *Nature Communications* (2024)

2023 Das, T., **Wang, Z.**, Sun, J., "Twin: Personalized clinical trial digital twin generation." In: *KDD'23*. 2023, pp. 402–413

2023 **Wang, Z.**, Theodorou, B., Fu, T., Xiao, C., Sun, J., "PyTrial: Machine learning software and benchmark for clinical trial applications." In: *arXiv preprint arXiv:2306.04018* (2023)

2023 **Wang, Z.**, Xiao, C., Sun, J., "AutoTrial: Prompting Language Models for Clinical Trial Design." In: *EMNLP'23*. 2023, pp. 12461–12472

2023 **Wang, Z.**, Xiao, C., Sun, J., "SPOT: sequential predictive modeling of clinical trial outcome with meta-learning." In: *ACM-BCB'23*. 2023, pp. 1–11

2022 **Wang, Z.**, Sun, J., "Trial2Vec: Zero-Shot Clinical Trial Document Similarity Search using Self-Supervision." In: *Findings of EMNLP'22*. 2022

Healthcare & Machine learning

2023 **Wang, Z.**, Wang, Z., Srinivasan, B., Ioannidis, V. N., Rangwala, H., Anubhai, R., "BioBridge: Bridging Biomedical Foundation Models via Knowledge Graphs." In: *ICLR'24*. 2023

2023 Wen, Y., **Wang, Z.**, Sun, J., "Mindmap: Knowledge graph prompting sparks graph of thoughts in large language models." In: *ACL'24*. 2023

2023 **Wang, Z.**, Gao, C., Xiao, C., Sun, J., "MediTab: Scaling Medical Tabular Data Predictors via Data Consolidation, Enrichment, and Refinement." In: *IJCAI'24*. 2023

2022 **Wang, Z.**, Sun, J., "Transtab: Learning transferable tabular transformers across tables." In: *NeurIPS'22* 35 (2022), pp. 2902–2915

2022 **Wang, Z.**, Wu, Z., Agarwal, D., Sun, J., "MedCLIP: Contrastive Learning from Unpaired Medical Images and Text." In: *EMNLP'22*. 2022

2022 **Wang, Z.**, Sun, J., "PromptEHR: Conditional Electronic Healthcare Records Generation with Prompt Learning." In: *EMNLP'22*. 2022, pp. 2873–2885

2022 **Wang, Z.**, Sun, J., "Survtrace: Transformers for survival analysis with competing events." In: *ACM-BCB'22*. 2022, pp. 1–9

2021 **Wang, Z.**, Yang, Y., Wen, R., Chen, X., Huang, S.-L., Zheng, Y., "Lifelong learning based disease diagnosis on clinical notes." In: *PAKDD'21*. Springer. 2021, pp. 213–224

2020 **Wang, Z.**, Wen, R., Chen, X., Cao, S., Huang, S.-L., Qian, B., Zheng, Y., "Online disease diagnosis with inductive heterogeneous graph convolutional networks." In: *WWW'21*. 2021, pp. 3349–3358

INVITED TALKS

2025 Medical scientific discovery in the era of LLMs, AI for Health Webinar.

2025 LLMs for clinical trial participant recruitment, IT Roundtable@Clinical Research Forum.

2024 Bridging biomedical foundation models, invited by Xuegong Lab@Tsinghua University.

2023 Bridging biomedical foundation models, invited by AI4Science@ByteDance.

2023 Automate clinical trial design with large language models, invited by Medidata.

2023 Medical vision-language modeling, invited by University of Toronto.

2023 Medical vision-language modeling, invited by MedAI@Stanford University.

2023 Medical vision-language modeling, invited by NCBI@NIH.

2023 Unifying large language models and knowledge graphs, invited by Titan Lab@AWS.

AWARDS

2023 Yee Memorial Fellowship, UIUC

2022 NeurIPS Scholar Award, NeurIPS

2022 Yunni & Maxine Pao Memorial Fellowship, UIUC

2021 Best Student Paper Award, PAKDD

SERVICE

Academic Journal and Conference Peer Review

Program Committee: NeurIPS (2022, 2023, 2024), EMNLP (2022, 2023), IJCAI (2022, 2023, 2024), AAAI (2022, 2024, 2025), NAACL (2024), COLM (2024), ICLR (2024), ICML (2025), ACL (2023), KDD (2023), NLPCC (2024)

Reviewer: Nature Communications, IEEE Transactions on Pattern Analysis and Machine Intelligence, Journal of Artificial Intelligence Research, Bioinformatics, Journal of the American Medical Informatics Association, ACM Computing Surveys

TEACHING

2024 Summer	TA: BSE740 Artificial Intelligence in Medicine, UIUC
2022 Spring	TA: CS598 Deep Learning for Healthcare, UIUC
2020 Summer	TA: Optimization Models and Applications, Tsinghua University
2020 Spring	TA: Bayesian Learning and Data Analysis, Tsinghua University
2019 Fall	TA: Learning from Data, Tsinghua University

(Updated on June 6, 2025.)