Guanghui Qin

Curriculum Vitae — *Updated July 4, 2026*

Contact

Microsoft Building 99
14820 NE 36th Street
Redmond, Washington, 98052 USA

Email: qin.guanghui@microsoft.com
Website: https://gqin.me
ORCID: 0000-0002-3009-8614

Education

Johns Hopkins University — Maryland, USA
Ph.D. in *Computer Science* — Aug 2019 – July 2024
Advisor: Benjamin Van Durme

Peking University — Beijing, China
B.S. in *Physics* (GPA 3.52) — Sept 2015 – June 2019
B.S. in *Computer Science* (GPA: 3.72) — Sept 2015 – June 2019

Experience

Microsoft — Washington, USA
Senior Researcher, with Hoifung Poon. — Sept 2024 – Present
Research Intern, with Corby Rosset. — May 2023 – Aug 2023
Research Intern, with Anthony Platanios. — May 2022 – Aug 2022

Johns Hopkins University — Maryland, USA
Research Assistant, with Benjamin Van Durme. — Aug 2019 – July 2024
Visiting Scholar, with Jason Eisner. — June 2018 – Nov 2018

Microsoft Research-Asia — Beijing, China
Research Intern, with Chin-Yew Lin. — Nov 2017 – June 2018

Honors

Gold Reviewer	ICML, 2026
Top Reviewer	NeurIPS, 2025
Best Short Paper Awardee	NAACL, 2021
PKU Scholar in Physics	Peking University, 2019
Outstanding Reviewer	EMNLP, 2019
Silver Medalist	Chinese Physics Olympiad , 2014

Mentoring

Jiachen Tu as an intern student at Microsoft	2025
Vivek Chari, undergrad student at Johns Hopkins University	2023
Yukun Feng, master student at Johns Hopkins University	2022

Skills

Programming languages: Python, Shell, SQL, LaTeX, Rust, and Go.
Machine Learning: PyTorch; Distributed Training with PyTorch FSDP and Ray; Training Large Vision Language Models (Supervised Fine-tuning and Reinforcement Learning with verl)
Network/Web/Database: Django; PostgreSQL; NGINX and HAProxy.

Teaching

Teaching assistant for *Machine Learning* (CS 601.475), JHU 2022

Activities

Reviewer for NeurIPS	2019, 2020, 2021, 2022, 2023 & **2025**
Reviewer for ICLR	2019, 2020, 2021, 2023, 2024, 2025 & 2026
Reviewer for ICML	2020, 2021 & **2026**
Reviewer for CVPR	2026
Reviewer for ACL	2021 & 2025
Reviewer for EMNLP	**2019**, 2020, 2021, 2022 & 2026
Reviewer for NAACL	2024
Reviewer for AAAI	2021
Reviewer for AKBC	2020

(I received reviewer awards for the marked years)

Publications

22. Yu Gu[*†], Jingjing Fu[*], Xiaodong Liu, Jeya Maria Jose Valanarasu, Noel C.F. Codella, Reuben Tan, Qianchu Liu, Ying Jin, Sheng Zhang, Jinyu Wang, Rui Wang, Lei Song, **Guanghui Qin**, Naoto Usuyama, Cliff Wong, Hao Cheng, HoHin Lee, Praneeth Sanapathi, Sarah Hilado, Tristan Naumann, Javier Alvarez-Valle, Jiang Bian, Mu Wei, Khalil Malik, Lidong Zhou, Jianfeng Gao, Eric Horvitz, Matthew P. Lungren, Doug Burger, Eric Topol[†], Hoifung Poon[†], and Paul Vozila. Evaluating the robustness and readiness of large frontier models in health AI applications. In *Nature Medicine*. 2026.

21. Jiachen Tu[*], **Guanghui Qin**[*], Theodore Zhao, Jeya Maria Jose Valanarasu, Sheng Zhang, Tristan Naumann, Fan Lam, Sheng Wang, and Hoifung Poon. Masked-Diffusion Autoencoders for 3D Medical Vision Representation Learning. In *Conference on Computer Vision and Pattern Recognition (CVPR)*. 2026.

20. Timothy Ossowski[*], Sheng Zhang[*], Qianchu Liu, **Guanghui Qin**, Reuben Tan, Tristan Naumann, Junjie Hu, and Hoifung Poon. OctoMed: Data Recipes for State-of-the-Art Multimodal Medical Reasoning. In *Conference on Computer Vision and Pattern Recognition (CVPR)*. 2026.

19. Yiyan Lin, Eleana Parajón, Qinling Yuan, Siyu Ye, **Guanghui Qin**, Yu Deng, Jane Borleis, Ariel Koyfman, Pablo A. Iglesias, Konstantinos Konstantopoulos, Douglas N. Robinson, and Peter N Devreotes[†]. Ras-mediated dynamic and biphasic regulation of cell migration. In *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*. 2025.

18. Abe Bohan Hou, Orion Weller, **Guanghui Qin**, Eugene Yang, Dawn Lawrie, Nils Holzenberger, Andrew Blair-Stanek, and Benjamin Van Durme. CLERC: A Dataset for Legal Case Retrieval and Retrieval-Augmented Analysis Generation. In *Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL, findings)*. 2025.

17. Corby Rosset, Ho-Lam Chung, **Guanghui Qin**, Ethan Chau, Zhuo Feng, Ahmed Awadallah[†], Jennifer Neville[†], and Nikhil Rao[†]. Researchy Questions: A Dataset of Multi-Perspective, Decompositional Questions for Deep Research. In *ACM SIGIR*. 2025.

16. Wenting Tan, Yunmo Chen, Tongnfei Chen, **Guanghui Qin**, Haoran Xu, Heidi C. Zhang, Benjamin Van Durme, and Phillip Koehn. Streaming Sequence Transduction through Dynamic Compression. In *International Conference on Spoken Language Translation (IWSLT)*. 2025.

15. **Guanghui Qin**. Ph.D. Dissertation: Towards Efficient Long-Context Natural Language Processing. In *Johns Hopkins University Library*. 2024.

14. Yiyan Lin*, Dhiman S Pal*†, Parijat Banerjee, Tatsat Banerjee, **Guanghui Qin**, Yu Deng, Jane Borleis, Pablo A. Iglesias, and Peter N. Devreotes†. Ras suppression potentiates rear actomyosin contractility-driven cell polarization and migration. In *Nature Cell Biology*. 2024.

13. **Guanghui Qin**, Corby Rosset, Ethan Chau, Nikhil Rao, and Benjamin Van Durme. Dodo: Dynamic Contextual Compression for Decoder-only LMs. In *Annual Meeting of the Association for Computational Linguistics (ACL, oral)*. 2024.

12. Abe Hou, Zhengping Jiang, **Guanghui Qin**, Orion Weller, Andrew Blair-Stanek, and Benjamin Van Durme. L-FRESco: Factual Recall Evaluation Score for Legal Analysis Generation. In *Workshop on Generative AI and Law*. 2024.

11. **Guanghui Qin** and Benjamin Van Durme. Nugget: Neural Agglomerative Embeddings of Text. In *International Conference on Machine Learning (ICML)*. 2023.

10. **Guanghui Qin**, Yukun Feng, and Benjamin Van Durme. The NLP Task Effectiveness of Long-Range Transformers. In *European Chapter of the Association for Computational Linguistics (EACL, oral)*. 2023.

9. **Guanghui Qin** and Jason Eisner. Learning How to Ask: Querying LMs with Mixtures of Soft Prompts. In *North American Chapter of the Association for Computational Linguistics (NAACL, short, oral)*. 2021. Best Short Paper Award

8. Mahsa Yarmohammadi, Shijie Wu, Marc Marone, Haoran Xu, Seth Ebner, **Guanghui Qin**, Yunmo Chen, Jialiang Guo, Craig Harman, Kenton Murray, Aaron S. White, Mark Dredze, and Benjamin Van Durme. Everything Is All It Takes: A Multipronged Strategy for Zero-Shot Cross-Lingual Information Extraction. In *Empirical Methods in Natural Language Processing (EMNLP, oral)*. 2021.

7. Patrick Xia*, **Guanghui Qin***, Siddharth Vashishtha, Yunmo Chen, Tongfei Chen, Chandler May, Craig Harman, Kyle Rawlins, Aaron S. White, and Benjamin Van Durme. LOME: Large Ontology Multilingual Extraction. In *European Chapter of the Association for Computational Linguistics (EACL, demo)*. 2021.

6. Ryan Culkin, J. Edward Hu, Elias Stengel-Eskin, **Guanghui Qin**, and Benjamin Van Durme. Iterative Paraphrastic Augmentation with Discriminative Span Alignment. In *Transactions of the Association for Computational Linguistics (TACL)*. 2021.

5. Hongyuan Mei, **Guanghui Qin**, Minjie Xu, and Jason Eisner. Neural Datalog Through Time: Informed Temporal Modeling via Logical Specification. In *International Conference on Machine Learning (ICML, oral)*. 2020.

4. Singh Abhinav, Patrick Xia, **Guanghui Qin**, Mahsa Yarmohammadi, and Benjamin Van Durme. CopyNext: Explicit Span Copying and Alignment in Sequence to Sequence Models. In *Workshop on Structured Prediction for NLP*. 2020.

3. Hongyuan Mei, **Guanghui Qin**, and Jason Eisner. Imputing Missing Events in Continuous-Time Event Streams. In *International Conference on Machine Learning (ICML, oral)*. 2019.

2. **Guanghui Qin**, Jin-Ge Yao, Xuening Wang, Jinpeng Wang, and Chin-Yew Lin. Learning Latent Semantic Annotations for Grounding Natural Language to Structured Data. In *Empirical Methods in Natural Language Processing (EMNLP, oral)*. 2018.

1. Longxu Dou, **Guanghui Qin**, Jinpeng Wang, Jin-Ge Yao, and Chin-Yew Lin. Data2Text Studio: Automated Text Generation from Structured Data. In *Empirical Methods in Natural Language Processing (EMNLP, demo)*. 2018.

Preprints

8. Qianchu Liu*, Sheng Zhang*, **Guanghui Qin***, Yu Gu, Ying Jin, Sam Preston, Yanbo Xu, Sid Kiblawi, Wen-wai Yim, Tim Ossowski, Tristan Naumann, Mu Wei†, and Hoifung Poon†. Scaling medical imaging report generation with multimodal reinforcement learning. In *arXiv*. 2026.

7. Qianchu Liu*, Sheng Zhang*, **Guanghui Qin***, Jeya Maria Jose Valanarasu, Maximilian Rokuss, Mingyu Lu, Timothy Ossowski, Juan Manuel Zambrano Chaves, Cliff Wong, Peniel Argaw, Yashna Hasija, Mu Wei, Wen-wai Yim, Qin Liu, Zilin Jing, Jason Entenmann, Naoto Usuyama, Tristan Naumann, and Hoifung Poon. HealthAgentBench: A Unified Benchmark Suite of Realistic Agentic Healthcare Environments for Challenging Frontier AI Agents. In *arXiv*. 2026.

6. Yiyan Lin, Siyu Ye, Saki Takayanagi, **Guanghui Qin**, Yu Long, Anjelika Gasilina, Takanari Inoue, Xufeng Wu, John Hammer, Miho Iijima, Michael L. Piacentino, and Peter N. Devreotes†. Excitable signaling and membrane curvature shape the continuum of cellular protrusions. In *preparation*. 2026.

5. James Y. Huang*, Sheng Zhang*, Qianchu Liu, **Guanghui Qin**, Tinghui Zhu, Tristan Naumann, Muhao Chen, and Hoifung Poon. Be My Eyes: Extending Large Language Models to New Modalities Through Multi-Agent Collaboration. In *arXiv*. 2025.

4. Huichao Deng, Jinbo Cheng, Richard D. Fetter, **Guanghui Qin**, Jianxiu Zhang, Xing Liang, Caitlin Taylor, Mingjie Zhang, Xiandeng Wu, and Kang Shen†. ERM proteins regulate the shape and number of Endoplasmic Reticulum-Plasma Membrane Junctions in neurons. In *bioRxiv*. 2025.

3. Sheng Zhang*, Qianchu Liu*, **Guanghui Qin***, Timothy Ossowski, Aiden Gu, Ying Jin Sid Kiblawi, Sam Preston, Mu Wei, Paul Vozila, Tristan Naumann, and Hoifung Poon. X-Reasoner: Towards Generalizable Reasoning Across Modalities and Domains. In *arXiv*. 2025.

2. Sheng Zhang*, Qianchu Liu*, **Guanghui Qin***, Tristan Naumann, and Hoifung Poon. Med-RLVR: Emerging Medical Reasoning from a 3B base model via reinforcement Learning. In *arXiv*. 2025.

1. Vivek Chari, **Guanghui Qin**, and Benjamin Van Durme. KV-Distill: Nearly Lossless Learnable Context Compression for LLMs. In *arXiv*. 2025.

*Asterisk * indicates authors with equal contribution. Dagger † indicates correspondence authors.*