


# Guanghai Qin

Curriculum Vitae

Updated July 1, 2024

CONTACT	Department of Computer Science Johns Hopkins University 3400 N Charles Street Baltimore, MD 21218, USA	Email: <a href="mailto:gqin2@jhu.edu">gqin2@jhu.edu</a> Website: <a href="https://gqin.me">https://gqin.me</a> ORCID: 0000-0002-3009-8614 
EDUCATION	<b>Johns Hopkins University</b> M.S.E. in Computer Science Ph.D. in Computer Science Advisor: <a href="#">Prof. Benjamin Van Durme</a>	Maryland, USA Aug 2019 – May 2023 Aug 2019 – July 2024 (expected)
	<b>Peking University</b> B.S. in Physics (primary major) B.S. in Computer Science (secondary major) <i>Physics GPA: 3.52; CS GPA: 3.72 .</i>	Beijing, P. R. China Sept 2015 – Jun 2019 Sept 2015 – Jun 2019
	<b>Microsoft Research Lab (MSR)</b> Research Intern. Mentor: <a href="#">Corby Rosset</a> I researched efficient methods for long-context LLMs. I proposed a method to compress the context of LLaMA with a compression ratio of up to 20x with minimal performance tradeoffs.	May 2023 – Aug 2023
	<b>Microsoft Semantic Machines</b> Research Intern. Mentor: <a href="#">Anthony Platanios</a> I studied a new research problem for user action predictions. I built a 2TiB dataset from GitHub and implemented a model based on graph neural networks (GNNs) to predict user actions (e.g. issues).	May 2022 – Aug 2022
PROFESSIONAL EXPERIENCE	<b>Johns Hopkins University</b> Visiting Scholar. Mentor: <a href="#">Hongyuan Mei</a> and <a href="#">Jason Einser</a> . I worked on temporal event stream modeling. We proposed a particle smoothing solution to sample events from a neural Hawkes process, which may interact with a deductive temporal database such as Datalog.	Jun 2018 – Nov 2018
	<b>Microsoft Research-Asia (MSRA)</b> Research Intern. Mentor: <a href="#">Jin-Ge Yao</a> and <a href="#">Chin-Yew Lin</a> . I proposed a Semi-HMMs-based statistical model for grounding natural language to structured data, which can be used to induce templates for data-to-text generation.	Nov 2017 – Jun 2018
HONORS	Best Short Paper Awardee in NAACL	2021
	PKU Scholar in Physics (Peking University)	2019
	Outstanding Reviewer in EMNLP	2019

Award for Scientific Research (Peking University)	2018
May 4th Scholarship (Peking University)	2016 and 2018
Merit Student (Peking University)	2017
Award for Academic Diligence (Peking University)	2016
Chinese Physics Olympiad (CPhO) Silver Medalist	2014

## PROFESSIONAL ACTIVITIES

Reviewer for NeurIPS (2019, 2020, 2021, 2022, and 2023)  
 Reviewer for ICLR (2019, 2020, 2021, 2023, and 2024)  
 Reviewer for ICML (2020 and 2021)  
 Reviewer for ACL (2021)  
 Reviewer for EMNLP (2019, 2020, 2021, and 2022)  
 Reviewer for NAACL (2024)  
 Reviewer for AAAI (2021)  
 Reviewer for AKBC (2020)

## SKILLS

*Programming languages:* Python, Rust, JAVA, C/C++, Shell,  $\LaTeX$ , and SQL.  
*Machine learning tools:* PyTorch, Lightning AI, DeepSpeed, and FAISS.  
*Network/Web/Database:* Front- and back-end. MySQL database.

## TEACHING EXPERIENCE

TA for *Machine Learning* (CS 601.475) at Johns Hopkins University in 2022.

## PUBLICATIONS

13. **Guanghui Qin**, Corby Rosset, Ethan C Chau, Nikhil Rao, and Benjamin Van Durme. [Dodo: Dynamic Contextual Compression for Decoder-only LMs](#). In *Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL)*. 2024.
12. Yiyang Lin\*, Dhiman S Pal\*, Parijat Banerjee, Tatsat Banerjee, **Guanghui Qin**, Yu Deng, Jane Borleis, Pablo A Iglesias, and Peter Devreotes. [Ras suppression potentiates rear actomyosin contractility-driven cell polarization and migration](#). In *Nature Cell Biology (NCB)*. 2024.
11. **Guanghui Qin** and Benjamin Van Durme. [Nugget: Neural Agglomerative Embeddings of Text](#). In *Proceedings of the Conference on 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 *Proceedings of the 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 *Proceedings of Conference of the 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 *Proceedings of the Conference on 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 Steven White, and Benjamin Van Durme. [LOME: Large Ontology Multilingual Extraction](#). In *Proceedings of Conference of the 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-based Alignment](#). In *Transactions of the Association for Computational Linguistics (TACL)*, 9:494-509. 2021.
5. Hongyuan Mei, **Guanghui Qin**, Minjie Xu, and Jason Eisner. [Neural Datalog Through Time: Informed Temporal Modeling via Logical Specification](#). In *Proceedings of the Conference on 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 *Proceedings of Workshop on Structured Prediction for NLP*. 2020.
3. Hongyuan Mei, **Guanghui Qin**, and Jason Eisner. [Imputing Missing Events in Continuous-Time Event Streams](#). In *Proceedings of the Conference on 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 *Proceedings of the Conference on 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 *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP, demo)*. 2018.

## PREPRINTS

3. Abe Bohan Hou, Orion Weller, **Guanghui Qin**, Eugene Yang, Dawn Lawrie, Nils Holzenberger, Andrew Blair-Stanek, Benjamin Van Durme. [CLERC: A Dataset for Legal Case Retrieval and Retrieval-Augmented Analysis Generation](#). In *arXiv*. 2024.
2. Corby Rosset, Ho-Lam Chung, **Guanghui Qin**, Ethan C Chau, Zhuo Feng, Ahmed Awadallah, Jennifer Neville, and Nikhil Rao. [Researchy Questions: A Dataset of Multi-Perspective, Decompositional Questions for LLM Web Agents](#). In *arXiv*. 2024.
1. 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 *arXiv*. 2024.

Asterisk \* indicates equal contribution

## CITATIONS

Visit my Google Scholar profile to view my citations:

<https://scholar.google.com/citations?user=ke8JkLgAAAAJ>