

Qingyun Wang

📧 eaglew.github.io ✉️ qwang16@wm.edu 🌐 EagleW

CURRENT EMPLOYMENT

Assistant Professor, Department of Data Science, **William & Mary**

August 2025 -

EDUCATION

University of Illinois at Urbana-Champaign (UIUC), Champaign, Illinois, USA

Ph.D. in Computer Science

August 2019 - May 2025

Doctoral Advisor: Heng Ji

Thesis Committee: Jiawei Han, Dilek Hakkani-Tur, Graham Neubig, Han Zhao, Tom Hope

Thesis Title: **AI4Scientist: Accelerating and Democratizing Scientific Research Lifecycle**

Rensselaer Polytechnic Institute (RPI), Troy, New York, USA

B.S. with double degrees in Computer Science, Mathematics

August 2015 - May 2019

Honors: **summa cum laude**, **Dean's Honor List** Fall 2015 through Spring 2019

Advisor: Heng Ji

PUBLICATIONS

Google Scholar: <https://scholar.google.com/citations?user=HQcZOHMAAAAJ&hl=en>

Semantic Scholar: <https://www.semanticscholar.org/author/Qingyun-Wang/1786863>

** denotes equal contributions*

Peer-reviewed Conference and Journal Publications

- [C18] Junsheng Huang, Zhitao He, Sandeep Polisetty, Qingyun Wang, May Fung. **MAC-Tuning: LLM Multi-Compositional Problem Reasoning with Enhanced Knowledge Boundary Awareness**. *EMNLP 2025*
- [C17] Charles Yu, Qingyun Wang, Yuting Hu, Jinjun Xiong, Heng Ji. **LORE: Continual Logit Rewriting Fosters Faithful Generation**. *EMNLP 2025 Findings*
- [C16] Minghao Liu, Zhitao He, Zhiyuan Fan, Qingyun Wang, May Fung. **MedEBench: Revisiting Text-instructed Image Editing on Medical Domain**. *EMNLP 2025 Findings*
- [C15] Jiayi He, Hehai Lin, Qingyun Wang, Yi Fung, Heng Ji. **Self-Correction is More than Refinement: A Learning Framework for Visual and Language Reasoning Tasks**. in *ACL 2025 Findings*
- [C14] Yuting Hu, Dancheng Liu, Qingyun Wang, Charles Yu, Heng Ji, Jinjun Xiong. **Automating Knowledge Discovery in Scientific Literature via LLMs: A Dual-Agent Approach with Progressive Ontology Prompting**. in *IJCAI 2025*
- [C13] Yumeng Wang, Zhiyuan Fan, Qingyun Wang, Yi Fung, Heng Ji. **CALM: Unleashing the Cross-Lingual Self-Aligning Ability of Language Models**. *NAACL 2025 Findings*
- [C12] Sha Li, Revanth Gangi Reddy, Khanh Duy Nguyen, Qingyun Wang, Yi Fung, Chi Han, Jiawei Han, Kartik Natarajan, Clare R. Voss, Heng Ji. **Schema-Guided Culture-Aware Complex Event Simulation with Multi-Agent Role-Play**. *EMNLP 2024 Demo*
- [C11] Qingyun Wang, Doug Downey, Heng Ji, Tom Hope. **SciMON: Scientific Inspiration Machines Optimized for Novelty**. *ACL 2024*

- [C10] Kexuan Xin, **Qingyun Wang**, Junyu Chen, Pengfei Yu, Huimin Zhao, and Heng Ji. **Gene-Metabolite Association Prediction with Interactive Knowledge Transfer Enhanced Graph for Metabolite Production**. *BIBM 2024*
- [C9] Hongyi Liu, **Qingyun Wang**, Payam Karisani, Heng Ji. **Name Tagging Under Domain Shift via Metric Learning for Life Sciences**. *NAACL 2024*.
- [C8] **Qingyun Wang**, Zixuan Zhang, Hongxiang Li, Xuan Liu, Jiawei Han, Huimin Zhao, Heng Ji. **Chem-FINESE: Validating Fine-Grained Few-shot Entity Extraction through Text Reconstruction**. *EACL 2024 Findings*
- [C7] **Qingyun Wang**, Manling Li, Hou Pong Chan, Lifu Huang, Julia Hockenmaier, Girish Chowdhary, Heng Ji. **Multimedia Generative Script Learning for Task Planning**. *ACL 2023 Findings*, [Findings Spotlight Presentation](#)
- [C6] Wenhao Yu, Chenguang Zhu, Zaitang Li, Zhiting Hu, **Qingyun Wang**, Heng Ji, Meng Jiang. **A Survey of Knowledge-Enhanced Text Generation**. *ACM Computing Surveys 2022*
- [C5] **Qingyun Wang**, Manling Li, Xuan Wang, Nikolaus Parulian, Guangxing Han, Jiawei Ma, Jingxuan Tu, Ying Lin, Haoran Zhang, Weili Liu, Aabhas Chauhan, Yingjun Guan, Bangzheng Li, Ruisong Li, Xiangchen Song, Yi Fung, Heng Ji, Jiawei Han, Shih-Fu Chang, James Pustejovsky, David Liem, Ahmed Elsayed, Martha Palmer, Jasmine Rah, Clare Voss, Cynthia Schneider, Boyan Onyshkevych. **COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation**. *NAACL 2021 System Demonstrations*, [Best Demo Award](#) 🏆
- [C4] **Qingyun Wang**, Qi Zeng, Lifu Huang, Kevin Knight, Heng Ji, Nazneen Fatema Rajani. **ReviewRobot: Explainable Paper Review Generation based on Knowledge Synthesis**. *INLG 2020*
- [C3] **Qingyun Wang**, Lifu Huang, Zhiying Jiang, Kevin Knight, Heng Ji, Mohit Bansal, Yi Luan. **PaperRobot: Incremental Draft Generation of Scientific Ideas**. *ACL 2019*
- [C2] **Qingyun Wang**, Xiaoman Pan, Lifu Huang, Boliang Zhang, Zhiying Jiang, Heng Ji, Kevin Knight. **Describing a Knowledge Base**. *INLG 2018*
- [C1] **Qingyun Wang***, Zhihao Zhou*, Lifu Huang, Spencer Whitehead, Boliang Zhang, Heng Ji, Kevin Knight. **Paper Abstract Writing through Editing Mechanism**. *ACL 2018*

Preprints and Papers Under Review

- [P5] Ruiling Xu, Yifan Zhang, **Qingyun Wang**, Carl Edwards, Heng Ji. **oMeBench: Towards Robust Benchmarking of LLMs in Organic Mechanism Elucidation and Reasoning** in *submission of ARR 2025*
- [P4] Yuji Zhang, **Qingyun Wang**, Cheng Qian, Jiateng Liu, Chenkai Sun, Denghui Zhang, Tarek Abdelzaher, Chengxiang Zhai, Preslav Nakov, Heng Ji. **Atomic Reasoning for Scientific Table Claim Verification** in *submission of ARR 2025*
- [P3] Bohao Wu, **Qingyun Wang**, Yue Guo. **Explain Less, Understand More: Jargon Detection via Personalized Parameter-Efficient Fine-tuning** in *ARR 2025*
- [P2] Ruochen Li, Teerth Patel, **Qingyun Wang**, Xinya Du. **MLR-copilot: Autonomous machine learning research based on large language models agents** in *arXiv 2024*
- [P1] Cheng Li, Yi Fung, **Qingyun Wang**, Chi Han, Manling Li, Jindong Wang, Heng Ji. **MentalArena: Self-play Training of Language Models for Mental Health Diagnosis and Treatment**. in *submission of arXiv 2025*

Tutorials

- [T4] **Qingyun Wang**, Yao Ge, Qingyu Chen, Zhiyong Lu. **Towards a Human-AI Collaborative Medical Research Lifecycle**. Tutorial at **AMIA 2025**
- [T3] **Qingyun Wang**, Carl Edwards, Heng Ji, Tom Hope. **Towards a Human-Computer Collaborative Scientific Paper Lifecycle: A Pilot Study and Hands-On Tutorial**. Tutorial at **LREC-COLING 2024**
- [T2] Carl Edwards, **Qingyun Wang**, Heng Ji. **Language + Molecules**. Tutorial at **EACL 2024**
- [T1] Wenhao Yu, Meng Jiang, Zhiting Hu, **Qingyun Wang**, Heng Ji, Nazneen Rajani. **Knowledge-Enriched Natural Language Generation**. Tutorial at **EMNLP 2021**

Extended Abstracts, Workshop Papers, and Technical Reports

- [A7] Carl Edwards, **Qingyun Wang**, Lawrence Zhao, Heng Ji. **L+M-24: Building a Dataset for Language+Molecules @ ACL 2024**. *Language + Molecules Workshop@ACL 2024*
- [A6] **Qingyun Wang**, Semih Yavuz, Xi Victoria Lin, Heng Ji, Nazneen Fatema Rajani. **Stage-wise Fine-tuning for Graph-to-Text Generation**. *ACL-IJCNLP 2021 SRW*, [Ranked 1st in Paper with Code at the time of publication](#)
- [A5] Manling Li, Ying Lin, Tuan Manh Lai, Xiaoman Pan, Haoyang Wen, Lifu Huang, Zhenhailong Wang, Pengfei Yu, Di Lu, **Qingyun Wang**, Haoran Zhang, Qi Zeng, Chi Han, Zixuan Zhang, Yujia Qin, Xiaodan Hu, Nikolaus Parulian, Daniel Campos, Heng Ji, et al. **GAIA at SM-KBP 2020 - A Dockerized Multi-media Multi-lingual Knowledge Extraction, Clustering, Temporal Tracking and Hypothesis Generation System**. *NIST TAC-KBP 2020*
- [A4] Manling Li, Ying Lin, Ananya Subburathinam, Spencer Whitehead, Xiaoman Pan, Di Lu, **Qingyun Wang**, Tongtao Zhang, Lifu Huang, Heng Ji, et al. **GAIA at SM-KBP 2019-A Multi-media Multi-lingual Knowledge Extraction and Hypothesis Generation System**. *NIST TAC-KBP 2019*
- [A3] Tongtao Zhang, Ananya Subburathinam, Ge Shi, Lifu Huang, Di Lu, Xiaoman Pan, Manling Li, Boliang Zhang, **Qingyun Wang**, Spencer Whitehead, Heng Ji et al. **GAIA - A Multi-media Multi-lingual Knowledge Extraction and Hypothesis Generation System**. *NIST TAC-KBP 2018*
- [A2] Ronald Ahmed Cardenas Acosta, Thamme Gowda, Ulf Hermjakob, Nima Pourdamghani, Michael Pust, Jibiao Shen, Tian Xie, Jonathan May, Nikolaos Malandrakis, Victor Martinez, Pavlos Papadopoulos, Anil Ramakrishna, Karan Singla, Ruchir Travadi, Colin Vaz, Shrikanth Narayanan, Kenton Murray, Toan Nguyen, David Chiang, Boliang Zhang, Xiaoman Pan, Ying Lin, Di Lu, Lifu Huang, Tongtao Zhang, Kevin Blissett, Ni Zhang, Spencer Whitehead, Ananya Subburathinam, Diya Li, **Qingyun Wang**, Zhiying Jiang, Heng Ji, et al. **ELISA System Description for LoReHLT 2018**. *LoReHLT 2018*
- [A1] Jieming Ji, **Qingyun Wang**, Zev Battad, Jiashun Gou, Jingfei Zhou, Rahul Divekar, Craig Carlson, Mei Si. **A Two-Layer Dialogue Framework For Authoring Social Bots**. *1st Proceedings of Alexa Prize 2017*, [Amazon Alexa Prize Finalist](#)

PROFESSIONAL SERVICES

Workshop Organizer

- [W4] Guohao Li, Yuan He, Yi R. Fung, **Qingyun Wang**, Fangru Lin, Xingyue Huang, Alisia Lupidi, Yusheng Su, Ziyu Ye, Da Yin, Ziyi Yang. **Scaling Environments for Agents (SEA)** at **NeurIPS 2025**
- [W3] Denghui Zhang, Zhaozhuo Xu, Jimin Huang, **Qingyun Wang**, Jing Gao. **VISTA: Visionary Innovation in Standards and Technology of GenAI** at **ICDM 2025**

[W2] **Qingyun Wang**, Wenpeng Yin, Lifu Huang, Yi R. Fung, Xinya Du, Carl Edwards, Tom Hope. **2nd AI4Research: Towards a Knowledge-grounded Scientific Research Lifecycle. AAAI 2025**

[W1] Carl Edwards, **Qingyun Wang**, Heng Ji, Tom Hope, Manling Li, Lawrence Zhao. **Language + Molecules**. 1st Workshop on Language + Molecules (L+M 2024) at **ACL 2024**

Session Chair

- ACL (2025-)

Area Chair

- ACL Rolling Review (2025-), EACL Demo (2024-), EMNLP Demo (2024-)

Program Committee Member

- NAACL/ACL/EMNLP Tutorials (2025-), ACL Rolling Review (2021-), ACL (2019-), EMNLP (2019-), NAACL (2019-), EMNLP Industry (2025-), AACL (2020-), INLG (2019-), COLING (2020-), EACL (2023-), LREC (2022-), COLM (2024-), ACL Demo (2021-), EACL Demo (2023-), EMNLP Demo (2023-), COLING Demo (2025-), FLLM (2024-), CCL (2020-), NLPCC (2021-)
- ICML (2020-), NeurIPS (2021-), ICLR (2021-), AISTATS (2025-), AAAI (2020-), IJCAI (2021-), ECAI (2025-)

Journal Reviewer

- **Nature Communications Chemistry** (2024-), Journal of the American Medical Informatics Association (2025-), Bioinformatics (2024-), Transactions on Audio, Speech and Language Processing (2025-), Engineering Applications of Artificial Intelligence (2023-), Neural Processing Letters 2022, ACM Transactions on Asian and Low-Resource Language Information Processing (2020-), Neurocomputing 2019

Workshop Program Committee Member

- RepL4NLP (2024-), WiNLP (2024-), AI4Research@IJCAI 2024, KnowledgeLM @ACL 2024, KnowledgeNLP (2023-)

Community Services

- Create student mentorship initiative for **AI4Research Workshop at AAAI 2025** 2024 - 2025
- Seminar Coordinator for the campus-wide NLP Talk Series 2024
- Cloud server maintainer of Blender Lab 2023 - 2024
- **ACL 2022 D&I Special Initiative Globalization via localization** 2022
- UIUC English Corner Treasurer 2021
- Organizer of Blender Lab Reading Group 2020

SELECT INVITED TALKS

[S9] **AI4Scientist: Accelerating and Democratizing Scientific Research Lifecycle**
The Cosmic Horizons: AI-Powered Insights into the Universe, NSF-Simons AI Institute for Cosmic Origins (CosmicAI), May 2025

[S8] **LLMs Copyright Risks: Copyright and Plagiarism in AI4Science**
NAACL Tutorial: LLMs and Copyright Risks: Benchmarks and Mitigation Approaches, May 2025

[S7] **AI4Scientist: Accelerating and Democratizing Scientific Research Lifecycle**
The Australian National University, Mar 2025

- [S6] **AI4Scientist: Accelerating and Democratizing Scientific Research Lifecycle**
Yale University, Feb 2025
- [S5] **AI4Scientist: Accelerating and Democratizing Scientific Research Lifecycle**
Texas A&M University, Jan 2025
- [S4] **AI4Scientist: Accelerating and Democratizing Scientific Research Lifecycle**
National Library of Medicine, National Center for Biotechnology Information, Jan 2025
- [S3] **SciMON: Scientific Inspiration Machines Optimized for Novelty**
Elsevier, Aug 2024
- [S2] **AIScientist: Toward Automated Literature Understanding and Scientific Discovery**
Pennsylvania State University, Apr 2024
- [S1] **SciMON: Scientific Inspiration Machines Optimized for Novelty**
Oak Ridge National Laboratory, Mar 2024

TEACHING

DATA 302 Databases

Fall 2025

OPEN SOURCE PROJECTS

- Implementation of **Covid-KG** framework received more than 2,000+ downloads.
- Implementation of **PaperRobot**. 478+ stars in GitHub.
- Implementation of **Writing-editing-Network**. 73+ stars in GitHub.
- Implementation of **Scientific Inspiration Machines Optimized for Novelty**. 87+ stars in GitHub.
- Implementation of **Describing a Knowledge Base**. 52+ stars in GitHub.
- Implementation of **ReviewRobot**. 30+ stars in GitHub.
- Implementation of **Stage-wise Fine-tuning**. 26+ stars in GitHub.
- Implementation of **Multimedia-Generative-Script-Learning**. 8+ stars in GitHub.
- Implementation of **Validating Fine-Grained Few-shot Entity Extraction through Text Reconstruction** . 7+ stars in GitHub.
- Reading list **KENLG-Reading** . 521+ stars in GitHub.

MEDIA COVERAGE

- **#AAAI2025 workshops round-up 1: Artificial intelligence for music, and towards a knowledge-grounded scientific research lifecycle**
AIhub
- **Conferences & Publications - Scientific Inspiration Machines Optimized for Novelty**
Molecule Maker Lab Newsletter - Issue 35
- **AI Breakdown - SciMON: Scientific Inspiration Machines Optimized for Novelty**
AI Breakdown
- **AI News (5th June 2023)**
Samuel Albanie YouTube Blog
- **Illinois CS Papers Earn Top Honors at Recent Conferences**
CS @ Illinois News

- **Senior Student Studying Abroad Develops ‘PaperRobot’ to Read Papers and Provide Research Direction**
CaiXin Blog
- **PaperRobot: Automatic Research Assistant That Can Generate Drafts of Scientific Ideas**
Neurohive
- **Rensselaer Team Is Finalist in \$2.5 Million Amazon Alexa Prize**
RPI News

AWARDS, GRANTS AND PATENTS

Awards

- [3] **NeurIPS** Top Reviewer in 2023.
- [2] **NAACL 2021 Best Demo Paper Award.**
- [1] Entered the **finalist** in the first Amazon Alexa Prize (2016).

Grants

- [4] Google Cloud Research Credits \$5000 2025
- [3] Co-PI of **ACCESS Accelerate Project: CIS240136 AI-Assisted Automated Laboratory with Multimodal Scientific Large Language Models** with 1,500,000 ACCESS Credits 2024
- [2] PI of **ACCESS Accelerate Project: CIS230280 Fine-Grained Named Entity Recognition with Self-Validation** with 200,000 ACCESS Credits 2023
- [1] The first Amazon Alexa Prize Finalist \$100,000 2016

Patents

- [PT3] **Qingyun Wang**, Nazneen Rajani, Semih Yavuz, Xi Lin, **Structured Graph-To-Text Generation with Two Step Fine-Tuning.** *US11727210B2*, Issued Feb 2022
- [PT2] **Qingyun Wang**, **Remote control moving table.** *CN202774939U*, Issued Mar 2013
- [PT1] **Qingyun Wang**, **Soap making device using household waste oil.** *CN201817462U*, Issued Nov 2011