

Akari Asai Curriculum Vitae

Carnegie Mellon University
Gates Hillman Complex, 4902 Forbes Ave
Pittsburgh, PA 15213

akariasai.github.io/
aasai@andrew.cmu.edu

Employment

07/2026 – **Carnegie Mellon University, School of Computer Science**
Assistant Professor
Language Technology Institute
Machine Learning Department (Courtesy)

07/2025 – **Allen Institute for AI**
present Research Scientist (OLMo)

Education

09/2019 – **University of Washington**
06/2025 Ph.D. in Computer Science
Paul G. Allen School of Computer Science and Engineering
Supervisor: Hannaneh Hajishirzi
Thesis: Beyond Scaling: Frontiers of Retrieval-Augmented Language Models
Committee: Yulia Tsvetkov, Luke Zettlemoyer, Pang Wei Koh, Lucy Lu Wang

03/2019 **University of Tokyo**
B.E. in Information and Communication Engineering (Electrical and Computer Engineering)
Department of Engineering
Supervisor: Kiyoharu Aizawa
Major GPA: 3.90/4, Cumulative GPA: 3.83
1st in Class (1/133), Dean's Award, Best Undergraduate Thesis Award

08/2015 – **University of California, Berkeley**
05/2016 American and International Study Program (AISP), Department of Political Science

Past Employment

09/2019 – **University of Washington**
06/2025 Research Assistant with Hannaneh Hajishirzi

09/2023 – **FAIR, Meta**
06/2025 Visiting Student Researcher with Scott Yih

07/2022 – **FAIR, Meta**
02/2023 Research Intern with Scott Yih, Sebastian Riedel

06/2021 – **Allen Institute for AI**
12/2021 Research Intern with Hannaneh Hajishirzi, Matt Gardner

04/2019 – **Salesforce Research**
08/2020 Research Intern with Kazuma Hashimoto, Caiming Xiong, Richard Socher

09/2018 – **Microsoft Research Asia**
01/2019 Research Intern with Nan Duan, Duyu Tang

03/2017 – **Megagon Labs**
05/2017 Research Intern with Alon Y. Halevy

06/2016 – **Google Japan**
10/2016 Engineering Intern at Google Search

Selected Awards, Grant and Fellowship

2025 AAAI/ACM SIGAI Dissertation Award Winner

2025 MIT Technology Review Innovators Under 35

2025 Forbes 30 Under 30 Asia in Science

2024 NeurIPS 2024 Outstanding Reviewer

2024 MIT Technology Review Innovators Under 35 Japan

2024 Amazon Research Grant

2024 Outstanding Paper Award at Annual Meeting of Japanese Association for NLP

2023 Best Paper Honorable Mention at NeurIPS Instruction-tuning Workshop

2023 Best Video Award (Most Viewed) ACL 2023

2023 Stability AI HPC Cluster Research Grant

2023 Microsoft Accelerate Foundation Models Research

2022 EECS Rising Stars

2022 The IBM Ph.D. Fellowship

2019 The Dean's Award, Best Undergraduate Thesis Award at University of Tokyo

2019 The Young Researcher Award at Annual Meeting of Japanese Association for NLP

2019 The Nakajima Foundation Doctoral Studying Abroad Scholarship (2019-2022)

2017 Women Techmakers Scholarship (Google Anita Borg Memorial Scholarship)

Peer-Reviewed Conference Publications

Google Scholar: https://scholar.google.com/citations?user=gqB4u_wAAAAJ&hl=en
Semantic Scholar: <https://www.semanticscholar.org/author/Akari-Asai/35584853>
* denotes equal contributions.

- [C32] **Akari Asai**, Jacqueline He*, Rulin Shao*, Weijia Shi, Amanpreet Singh, Joseph Chang, Kyle Lo, Luca Soldaini, Sergey Feldman, Mike D’arcy, Matt Latzke, David Wadden, Minyang Tian, Pan Ji, Shengyan Liu, Hao Tong, Bohao Wu, Yanyu Xiong, Luke Zettlemoyer, Graham Neubig, Dan Weld, Doug Downey, Wen-tau Yih, Pang Wei Koh, Hannaneh Hajishirzi. OpenScholar: Synthesizing Scientific Literature with Retrieval-augmented LMs. *Nature*, 2026. [pdf]
Featured in VentureBeat, Nature News, Science News, and UW News
Our public demo, hosted by Aiz, reached 11,000 users in the first week; used by more than 50k practitioners.
- [C31] Prahaladh Chandrahasan, Jiahe Jin, Zhihan Zhang, Tevin Wang, Andy Tang, Lucy Mo, Morteza Ziyadi, Leonardo F.R. Ribeiro, Zimeng Qiu, Markus Dreyer, **Akari Asai**, Chenyan Xiong, Deep Research Comparator: A Platform For Fine-grained Human Annotations of Deep Research Agents. *THE ACM Web Conference (WWW; Demonstrations Track)*. 2026. [pdf]
- [C30] Tianyu Cao, Neel Bhandari, Akhila Yerukola, **Akari Asai**, Maarten Sap, Out of Style: RAG’s Fragility to Linguistic Variation. *Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. 2026. [pdf]
- [C29] Zora Zhiruo Wang*, **Akari Asai***, Xinyan Velocity Yu, Frank F. Xu, Yiqing Xie, Graham Neubig, Daniel Fried. CodeRAG-Bench: Can Retrieval Augment Code Generation?. *Findings of NAACL*. 2025. [pdf]
- [C28] Xiang Yue*, Yueqi Song*, **Akari Asai**, Seungone Kim, Jean de Dieu Nyandwi, Simran Khanuja, Anjali Kantharuban, Lintang Sutawika, Sathyanarayanan Ramamoorthy, Graham Neubig. Pangea: A Fully Open Multilingual Multimodal LLM for 39 Languages. *In Proceedings of International Conference on Learning Representation (ICLR)*. 2025. [pdf]
- [C27] Go Kamoda, **Akari Asai**, Ana Brassard, Keisuke Sakaguchi. Quantifying the Influence of Evaluation Aspects on Long-Form Response Assessment. *In Proceedings of International Conference on Computational Linguistics (COLING)*. 2025.
Outstanding Paper Award at Annual Meeting of the Japanese Association for NLP
- [C26] Rulin Shao, Jacqueline He, **Akari Asai**, Weijia Shi, Tim Dettmers, Sewon Min, Luke Zettlemoyer, Pang Wei Koh. Scaling Retrieval-Based Language Models with a Trillion-Token Datas-tore. *In Proceedings of Conference on Neural Information Processing Systems (NeurIPS)*. 2024. [pdf]
- [C25] Tong Chen, **Akari Asai***, Niloofar Miresghallah*, Sewon Min, James Grimmermann, Yejin Choi, Hannaneh Hajishirzi, Luke Zettlemoyer, Pang Wei Koh. CopyBench: Measuring Literal and Non-Literal Reproduction of Copyright-Protected Text in Language Model Generation. *In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2024. [pdf]
Contributed Talk at NeurIPS 2024 Workshop on Regulatable ML
- [C24] Abhika Mishra, **Akari Asai**, Vidhisha Balachandran, Yizhong Wang, Graham Neubig, Yulia Tsvetkov, Hannaneh Hajishirzi. Fine-grained Hallucination Detection and Editing for Language Models. *In Proceedings of Conference on Language Modeling (COLM)*. 2024. [pdf]
- [C23] **Akari Asai**, Sneha Kudugunta, Xinyan Velocity Yu, Terra Blevins, Hila Gonen, Machel Reid, Yulia Tsvetkov, Sebastian Ruder, Hannaneh Hajishirzi. BUFFET: Benchmarking Large Language Models for Few-shot Cross-lingual Transfer. *In Proceedings of Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2024. [pdf]
Oral presentation

- [C22] **Akari Asai**, Zeqiu Wu, Yizhong Wang, Avirup Sil, Hannaneh Hajishirzi. Self-RAG: Learning to Retrieve, Generate, and Critique through Self-Reflection. *In Proceedings of International Conference on Learning Representations*. 2024. [pdf]
Oral presentation (Top 1% among 7k submissions)
Most Influential ICLR 2024 Papers No. 13 (Paper Digest)
Best Paper Honorable Mention at NeurIPS Instruction Workshop
Officially integrated into LangChain and LlamaIndex, 2k Github stars
Featured in Forbes
- [C21] Jungo Kasai, Keisuke Sakaguchi, Yoichi Takahashi, Ronan Le Bras, **Akari Asai**, Xinyan Yu, Dragomir Radev, Noah A. Smith, Yejin Choi, Kentaro Inui. RealTime QA: What's the Answer Right Now?. *In Proceedings of NeurIPS (Datasets and Benchmarks)*. 2023. [pdf]
- [C20] Joongwon Kim, **Akari Asai**, Gabriel Ilharco, Hannaneh Hajishirzi. TaskWeb: Selecting Better Source Tasks for Multi-task NLP. *In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2023. [pdf]
- [C19] Sheng-Chieh Lin, **Akari Asai**, Minghan Li, Barlas Oguz, Jimmy Lin, Yashar Mehdad, Wen-tau Yih, Xilun Chen. How to Train Your DRAGON: Diverse Augmentation Towards Generalizable Dense Retrieval. *In Proceedings of Findings of EMNLP*. 2023. [pdf]
- [C18] Odunayo Ogundepo, Tajuddeen R. Gwadabe, Clara E. Rivera, Jonathan H. Clark, Sebastian Ruder, David Ifeoluwa Adelani, Bonaventure F. P. Dossou, ..., **Akari Asai** ... (52 authors). AfriQA: Cross-lingual Open-Retrieval Question Answering for African Languages. *In Proceedings of Findings of EMNLP*. 2023. [pdf]
- [C17] Alex Mallen*, **Akari Asai***, Victor Zhong, Rajarshi Das, Hannaneh Hajishirzi, Daniel Khashabi. When Not to Trust Language Models: Investigating Effectiveness and Limitations of Parametric and Non-Parametric Memories. *In Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL)*. 2023. [pdf]
Oral presentation
ACL 2023 Best Video Award (Most Viewed)
Most Influential ACL 2023 Paper No 5 (Paper Digest)
- [C16] **Akari Asai**, Timo Schick, Patrick Lewis, Xilun Chen, Gautier Izacard, Sebastian Riedel, Hannaneh Hajishirzi, Wen-tau Yih. Task-aware Retrieval with Instructions. *In Proceedings of Findings of ACL*. 2023. [pdf]
- [C15] Xiaoyu Shen, **Akari Asai**, Bill Byrne, Adrià de Gispert. xPQA: Cross-Lingual Product Question Answering across 12 Languages. *In Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL; Industry track)*. 2023. [pdf]
- [C14] **Akari Asai**, Mohammadreza Salehi, Matthew E. Peters, and Hannaneh Hajishirzi. Attentional Mixtures of Soft Prompt Tuning for Parameter-efficient Multi-task Knowledge Sharing. *In Proceedings of Conference on Empirical Methods in Natural Language Processing*. 2022. [pdf]
- [C13] Xinyan Yu*, **Akari Asai***, Trina Chatterjee, Junjie Hu, Eunsol Choi. Beyond Counting Datasets: Investigating Multilingual Dataset Construction and Necessary Resources. *Findings of EMNLP*. 2022. [pdf]
- [C12] **Akari Asai**, Matt Gardner, and Hannaneh Hajishirzi. Evidentiality-guided Generation for Knowledge-Intensive NLP Tasks. *In Proceedings of Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2022. [pdf]
Oral presentation

- [C11] **Akari Asai**, Xinyan Yu, Jungo Kasai, and Hannaneh Hajishirzi. One Question Answering Model for Many Languages with Cross-lingual Dense Passage Retrieval. *In Proceedings of Conference on Neural Information Processing Systems (NeurIPS)*. 2021. [pdf]
- [C10] Takumi Kawashima, Qina Yu, **Akari Asai**, Daiki Ikami, Kiyoharu Aizawa. The aleatoric uncertainty estimation using a separate formulation with virtual residuals. *In Proceedings of International Conference on Pattern Recognition (ICPR)*. 2021. [pdf]
- [C9] **Akari Asai**, Eunsol Choi. Challenges in Information Seeking QA: Unanswerable Questions and Paragraph Retrieval. *In Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL)*. 2021. [pdf]
- [C8] Ikuya Yamada, **Akari Asai**, Hannaneh Hajishirzi. Efficient Passage Retrieval with Hashing for Open-domain Question Answering *In Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL)*. 2021. [pdf]
- [C7] **Akari Asai**, Jungo Kasai, Jonathan H. Clark, Kenton Lee, Eunsol Choi, Hannaneh Hajishirzi. XOR QA: Cross-lingual Open-Retrieval Question Answering. *In Proceedings of Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*. 2021. [pdf]
Oral presentation
- [C6] Alon Talmor, Ori Yoran, Amnon Catav, Dan Lahav, Yizhong Wang, **Akari Asai**, Gabriel Ilharco, Hannaneh Hajishirzi, Jonathan Berant. MultiModalQA: complex question answering over text, tables and images. *In Proceedings of International Conference on Learning Representations (ICLR)*. 2021. [pdf]
- [C5] Ikuya Yamada, **Akari Asai**, Hiroyuki Shindo, Hideaki Takeda, Yuji Matsumoto. LUKE: Deep Contextualized Entity Representations with Entity-aware Self-attention. *In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2020. [pdf]
Most Influential EMNLP 2020 Paper No.8 (Paper Digest)
1.4 million downloads at Hugging face
Officially integrated into Huggingface Transformers Model Class
700 Github stars
- [C4] Ikuya Yamada, **Akari Asai**, Jin Sakuma, Hiroyuki Shindo, Hideaki Takeda, Yoshiyasu Takefuji, Yuji Matsumoto. Wikipedia2Vec: An Optimized Tool for Learning Embeddings of Words and Entities from Wikipedia. *In Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP, demo)*. 2020. [pdf]
400k PyPI install
1k Github stars
- [C3] **Akari Asai**, Hannaneh Hajishirzi. Logic-guided Data Augmentation and Regularization for Consistent Question Answering. *In Proceedings of The Annual Conference of the Association for Computational Linguistics (ACL; Short)*. 2020. [pdf]
- [C2] **Akari Asai**, Kazuma Hashimoto, Hannaneh Hajishirzi, Richard Socher, Caiming Xiong. Learning to Retrieve Reasoning Paths over Wikipedia Graphs for Question Answering. *International Conference on Learning Representations (ICLR)*, 2020. [pdf]
Featured in VentureBeat
Featured in Towards Datascience (Top Trends of Graph Machine Learning in 2020)
Integrated into Salesforce COVID Search (CO-Search)

- [C1] **Akari Asai**, Sara Evensen, Behzad Golshan, Alon Halevy, Vivian Li, Andrei Lopatenko, Daniela Stepanov, Yoshihiko Suhara, Wang-Chiew Tan, Yinzhan Xu. HappyDB: A Corpus of 100,000 Crowdsourced Happy Moments *In Proceedings of International Conference on Language Resources and Evaluation (LREC)*, 2018. [pdf]
Featured in MIT Technology Review

Preprint

- [P3] Rulin Shao*, **Akari Asai***, Shannon Zejiang Shen*, Hamish Ivison*, Varsha Kishore, Jingming Zhuo, Xinran Zhao, Molly Park, Samuel G. Finlayson, David Sontag, Tyler Murray, Sewon Min, Pradeep Dasigi, Luca Soldaini, Faeze Brahman, Wen-tau Yih, Tongshuang Wu, Luke Zettlemoyer, Yoon Kim, Hannaneh Hajishirzi, Pang Wei Koh, DR Tulu: Reinforcement Learning with Evolving Rubrics for Deep Research. 2025. [pdf]
- [P2] Tong Chen, **Akari Asai**, Luke Zettlemoyer, Hannaneh Hajishirzi, Faeze Brahman, Train for Truth, Keep the Skills: Binary Retrieval-Augmented Reward Mitigates Hallucinations. Under review & NeurIPS 2025 Socially Responsible and Trustworthy Foundation Models, 2025. [pdf]
- [P1] **Akari Asai**, Zexuan Zhong, Danqi Chen, Pang Wei Koh, Luke Zettlemoyer, Hannaneh Hajishirzi, Wen-tau Yih. Reliable, Adaptable, and Attributable Language Models with Retrieval. 2024. [pdf]

Tutorials

- [T1] **Akari Asai**, Sewon Min, Zexuan Zhong, Danqi Chen. Retrieval-based Language Models and Applications. *In Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL; Tutorial)*. 2023.

Workshop Publications

- [W3] **Akari Asai**, Shayne Longpre, Jungo Kasai, Chia-Hsuan Lee, Rui Zhang, Junjie Hu, Ikuya Yamada, Jonathan H. Clark, Eunsol Choi. MIA 2022 Shared Task: Evaluating Cross-lingual Open-Retrieval Question Answering for 16 Diverse Languages. *In Proceedings of Workshop on Multilingual Information Access (MIA)*. 2022. [pdf]
- [W2] **Akari Asai**, Daiki Ikami, and Kiyoharu Aizawa. Multi-task Learning Based on Separable Formulation of Depth Estimation and Uncertainty. *In Proceedings of Conference on Computer Vision and Pattern Recognition (CVPR), Uncertainty and Robustness in Deep Visual Learning*. 2019.
- [W1] **Akari Asai**, Vivian Li, Daniela Stepanov, and Wang-Chiew Tan. A Data-Driven Approach to Understanding Happiness. *In Proceedings of ACL Widening NLP Workshop (WiNLP)*. 2018.

Teaching

Head TA at UW CSE Courses

Winter 2024 **CSE599J Data-centric ML (Graduate)**

Paul G. Allen School of Computer Science and Engineering, University of Washington
Instructor: Pang Wei Koh

Fall 2023 **CSE473 Artificial Intelligence (Undergraduate)**
Paul G. Allen School of Computer Science and Engineering, University of Washington
Instructor: Hannaneh Hajishirzi

Invited Lecturer at Courses

Spring 2026 **CS11-766 Large Language Model Applications**
School of Computer Science, Carnegie Mellon University
Instructors: Fernando Diaz, Daphne Ippolito

Spring 2025 **CS11-711 Advanced Natural Language Processing**
School of Computer Science, Carnegie Mellon University
Instructor: Sean Welleck

Fall 2024 **CS 769 Advanced Natural Language Processing**
Department of Computer Science, University of Wisconsin-Madison
Instructor: Junjie Hu

Fall 2024 **CS 11-667 Large Language Models: Methods and Applications**
School of Computer Science, Carnegie Mellon University
Instructor: Daphne Ippolito and Chenyan Xiong

Spring 2024 **CPSC 477/577 Natural Language Processing**
Department of Computer Science, Yale University
Instructor: Arman Cohan

Spring 2024 **CS 11-711 Advanced NLP**
School of Computer Science, Carnegie Mellon University
Instructor: Graham Neubig

Spring 2024 **CSE 561A: Large Language Models**
Department of Computer Science, Washington University in St. Louis
Instructor: Jiaxin Huang

Winter 2024 **CSE 447 Natural Language Processing**
Paul G. Allen School of Computer Science and Engineering, University of Washington
Instructor: Yejin Choi

Fall 2023 **CSE 473 Introduction to Artificial Intelligence**
Paul G. Allen School of Computer Science and Engineering, University of Washington
Instructor: Hannaneh Hajishirzi

Winter 2023 **CSE 447 Natural Language Processing**
Paul G. Allen School of Computer Science and Engineering, University of Washington
Instructor: Sofia Serrano

Fall 2022 **CSE 373 Data Structures and Algorithm**
Paul G. Allen School of Computer Science and Engineering, University of Washington
Instructor: Kevin Lin

Conference Tutorial Instructor

July 2023 **Tutorial on Retrieval-based LMs and Applications**
The 61st Annual Meeting of the Association for Computational Linguistics (ACL)
Co-instructors: Sewon Min, Zexuan Zhong, Danqi Chen

Mentoring

04/2022 – **Public Office hour**
09/2024 Virtual office hours open to the public (two 30-minute 1-on-1 sessions every week)
More than 90 students joined to discuss research and Ph.D. applications

PhD Students

02/2024 – **Tong Chen**
present UW CSE Ph.D. student (advised by Hannaneh Hajishirzi and Luke Zettlemoyer)
EMNLP 24
Contributed Talk at NeurIPS 2024 Regulatable ML

02/2024 – **Rulin Shao**
present PhD student (advised by Luke Zettlemoyer, Pang Wei Koh)
NeurIPS 24

09/2023 – **Joongwon (Daniel) Kim**
07/2024 PhD student (advised by Hannaneh Hajishirzi)
EMNLP 23

Predoctoral Students

02/2024 – **Kevin Farhat**
12/2024 UW CSE Undergraduate Student → Research Engineer at Aiz

03/2023 – **Abhika Mishra**
10/2024 UW CSE Undergraduate Student → Meta
COLM 24

01/2023 – **Go Kamoda**
01/2025 Tohoku University (Japan) Master Student → Ph.D. Student at SOKENDAI
COLING 25
Outstanding Paper Award at Annual Meeting of Japanese Association for NLP

04/2022 – **Alex Mallen**
06/2023 UW CSE Undergraduate Student → Research Engineer at Eleuther AI
ACL 23
Best Video Award and Oral Presentation; Outstanding Senior Student Award

04/2021 – **Velocity Yu**
06/2022 UW CSE Undergraduate Student → Ph.D. Student at University Southern California
NeurIPS 21; Findings of EMNLP 23; NAACL 2024

Service

Academic Services within the University

2021 – 2023 **Ph.D. Admission Student Area Chair (NLP)**
Managed Ph.D. Admission student reviewing process for the entire UW NLP group

2021 – 2023 **Ph.D. Admission Committee Member**
Reviewed Ph.D. application materials for applicants in NLP

- 2022 – 2023 **Diversity and Inclusion Committee Student Representative**
Represented student perspectives at monthly DEI meetings of the department
 - 2022 – 2023 **Faculty Hiring Student DEI Reviewer**
Played a key role in integrating DEI considerations into the UW CSE faculty hiring process
 - 2022 **Ph.D. Pre-Application Mentorship Service (PAMS) Mentor**
Mentored students from underrepresented groups in preparing their Ph.D. applications
 - 2023 – 2025 **UW NLP Seminar Organizer**
Organized and invited speakers for the annual UWNLP seminar series
2023: Chenhao Tan, Tri Dao, He He, Tatsu Hashimoto, Jacob Andreas, Chelsea Finn
2024: Yoon Kim, Peter Henderson, John Hewitt, Diyi Yang, Daphne Ippolito, Been Kim
 - 2023 – 2025 **H2lab Undergraduate Interns Hiring Chair**
Managed undergraduate interns recruitment process at Hannaneh Hajishirzi group.
- Research Community
- 12/2025 **NeurIPS Competition Organizer**
Massive Multi-Modal User-Centric Retrieval-Augmented Generation (MMU-RAG)
 - 10/2025 **Workshop Organizer**
The 1st Workshop on LLM for Scientific Discovery (COLM)
 - 04/2025 **Workshop Organizer**
The 4th Workshop on Knowledge-Augmentation for LMs and NLP Methods (NAACL)
 - 07/2022 **Workshop Lead Organizer**
The first Workshop on Multilingual Information Access (NAACL)
 - 07/2022 **Shared Task Lead Organizer**
NAACL 2022 MIA Shared Task on Multilingual Open-Retrieval Question Answering
 - 07/2023 **Tutorial Instructor**
ACL 2023 Tutorial on Retrieval-based Language Models and Applications
 - 07/2022 **Organizing Committee**
NAACL 2022 Student Volunteer Coordinator
 - 2023 – present **Area Chair**
ACL 2025, 2026
AAACL 2025
EMNLP 2023, 2024
EACL 2023, 2025

2020 – present **Reviewer**
ACL Rolling Review (ARR) 2021-present
ACL 2020-2023, ACL 2023 System Demonstrations
EMNLP 2019-2023
COLING 2022
NeurIPS 2021-2025, NeurIPS Dataset and Benchmark 2024
ICLR 2021-2024
AKBC 2020
AAAI 2020
IJCAI 2021
COLM 2024
Workshops: EMNLP 2024 WiNLP, EMNLP 2024 Big Picture, ICLR 2024 ME-FoMo

Selected Media Coverage

- 02/2026 Nature News. *In a study, AI model OpenScholar synthesizes scientific research and cites sources as accurately as human experts* [link]
- 02/2026 Science News. *Open-source AI program can answer science questions better than humans* [link]
- 02/2026 UW News. *In a study, AI model OpenScholar synthesizes scientific research and cites sources as accurately as human experts* [link]
- 09/2025 Nature News. *Can researchers stop AI making up citations?* [link]
- 09/2025 CMU LTI News. *Asai Named to MIT Technology Review's '35 Innovators Under 35'* [link]
- 01/2025 Allen School News. *Working to solve global problems: Allen School Ph.D. student Akari Asai named one of MIT Technology Review's Innovators Under 35 Japan* [link]
- 11/2024 VentureBeat. *OpenScholar: The open-source A.I. that's outperforming GPT-4o in scientific research.* [link]
- 11/2024 MIT Technology Review. *Innovators Under 35: A trailblazer in retrieval-augmented generation research is taking on the challenge of preventing 'hallucinations' and elevating LLM reliability.* [link]
- 07/2024 Forbes. *How RAG-Powered AI Applications Have A Positive Impact On Businesses* [link]
- 02/2024 LangChain Official Blog. *Self-Reflective RAG with LangGraph* [link]
- 11/2023 Towards Data Science. *How Self-RAG Could Revolutionize Industrial LLMs* [link]
- 10/2022 Allen School News. *Lost in translation no more: IBM Fellowship winner Akari Asai asks — and answers — big questions in NLP to expand information access to all* [link]
- 02/2020 Towards Data Science. *Top Trends of Graph Machine Learning in 2020* [link]
- 02/2020 VentureBeat. *Salesforce's AI navigates Wikipedia to find answers to complex questions* [link]
- 02/2018 MIT Technology Review. *100,000 happy moments – What makes people happy? A huge database is making it possible to discern the answer at last.* [link]

Invited Talks

02/2026 Penn NLP Computational Linguistics Lunch Seminar

02/2026 Microsoft Research (Hosted by Naoto Usuyama)

01/2026 Oracle (Hosted by Avirup Sil)

12/2025 NeurIPS 2025 Competition: NeurIPS 2025 Competition: MMU-RAGent: Massive Multi-Modal User-Centric Retrieval Augmented Generation Benchmark

12/2025 NeurIPS 2025 Competition: NeurIPS 2025 Competition: MMU-RAGent: Massive Multi-Modal User-Centric Retrieval Augmented Generation Benchmark

09/2025 CMU AI for Science Workshop 2025

05/2025 The 10th Workshop on Representation Learning for NLP (Repl4NLP) at NAACL

05/2025 Foundation Models for Science Workshop (hosted by Shirley Ho)

04/2025 Flatiron Institute (hosted by Shirley Ho)

01/2025 University of Pittsburgh, CS Colloquium Series (hosted by Xiang Lorraine Li)

11/2024 MIT Technology Review Japan Innovators under 35 Summit

11/2024 Carnegie Mellon University, WInE Lab (hosted by Sherry Tongshuang Wu)

05/2024 Microsoft Research India (hosted by Srivathsan Koundinyan)

05/2024 National Institutes of Health (NIH) (hosted by Zhiyong Lu, Ashley Shin)

03/2024 Meta, Smart Glass AI Reading Group (hosted by Kai Sun)

02/2024 University College London, Web Intelligence Group (hosted by Emine Yilmaz)

12/2023 Microsoft Research, Health Futures team (hosted by Naoto Usuyama)

07/2021 Allen School Open House, NLP Research Showcase

10/2023 The University of Edinburgh (hosted by Papsquale Minervanie)

10/2023 University of Massachusetts Amherst Machine Learning and Friends Lunch (hosted by Dhruv Agarwal)

08/2023 Preferred Networks (hosted by Daisuke Okanohara)

08/2023 Okinawa Institute of Science and Technology (hosted by Makoto Yamada)

08/2023 University of Tokyo, Multi Media Processing Group (hosted by Kiyoharu Aizawa)

08/2023 The University of Queensland Data Science Seminar Series (hosted by Guido Zuccon)

07/2023 The First International Workshop on Retrieval-enhanced Machine Learning at SIGIR, Keynote Talk (hosted by Hamed Zamani)

07/2023 Japanese NLP colloquium (hosted by Naoki Otani)

04/2023 MILA, NLP Reading Group (hosted by Siva Reddy)

03/2023 The AI TALK (hosted by Jingkang Yang)
12/2022 The AI King, Keynote Talk (hosted by Jun Suzuki)
10/2022 UT Austin NLP Seminar (hosted by Eunsol Choi)
08/2022 Amazon Alexa (hosted by Qiang Ning)
10/2021 University of Amsterdam, Search Engines Amsterdam (hosted by Maarten de Rijke)
07/2021 Apple, Web Answers (hosted by Shayne Longpre)
04/2021 Japanese NLP colloquium (hosted by Naoki Otani)
01/2021 Apple, Web Answers (hosted by Shayne Longpre)
01/2021 Google Research (hosted by Markus N. Rabe, Christian Szegedy)