

**RYO KAMOI**

ryokamoi@psu.edu

<https://ryokamoi.github.io/>**Research Interests**

---

Detecting and mitigating errors in LLM responses (inference-time scaling, self-correction), synthetic datasets, textual entailment, fact verification, text summarization, and vision-language models.

**Education**

---

**Pennsylvania State University** - Ph.D. Student in Computer Science Aug 2023 - (May 2028)

Advisor: Dr. Rui Zhang

**University of Texas at Austin** - Master of Science in Computer Science Aug 2020 - Dec 2022

Advisor: Dr. Greg Durrett, Mentor: Tanya Goyal

**Keio University**, Japan - Bachelor of Engineering in Statistics Apr 2016 - Mar 2020

Advisor: Dr. Kei Kobayashi, Top student in the Department of Mathematics (Keio Engineering Foundation Award)

**Research Internships**

---

**Microsoft**, Redmond, WA – Research Intern at the Office of Applied Research May - Aug 2025

**Amazon**, Cambridge, UK – Applied Scientist Intern at the Alexa Team Jul - Dec 2021

- Developed an interpretable answer quality evaluation framework for responses from chatbots.

**Selected Papers**

<https://scholar.google.com/citations?user=4OWTLKAAAAAJ>

**Ryo Kamoi**, Yusen Zhang, Sarkar Snigdha Sarathi Das, Ranran Haoran Zhang, Rui Zhang. 2024. [VisOnlyQA: Large Vision Language Models Still Struggle with Visual Perception of Geometric Information](#). *arXiv preprint arXiv:2412.00947*.

**Ryo Kamoi**, Yusen Zhang, Nan Zhang, Jiawei Han, Rui Zhang. 2024. [When Can LLMs Actually Correct Their Own Mistakes? A Critical Survey of Self-Correction of LLMs](#). *TACL*.

**Ryo Kamoi**, Sarkar Snigdha Sarathi Das, Renze Lou, Jihyun Janice Ahn, Yilun Zhao, Xiaoxin Lu, Nan Zhang, Yusen Zhang, Ranran Haoran Zhang, Sujeeth Reddy Vummanthala, Salika Dave, Shaobo Qin, Arman Cohan, Wenpeng Yin, Rui Zhang. 2024. [Evaluating LLMs at Detecting Errors in LLM Responses](#). *In COLM*.

**Ryo Kamoi**, Tanya Goyal, Juan Diego Rodriguez, and Greg Durrett. 2023. [WiCE: Real-world Entailment for Claims in Wikipedia](#). *In EMNLP (main)*.

**Ryo Kamoi**, Tanya Goyal, Juan Diego Rodriguez, and Greg Durrett. 2023. [Shortcomings of Question Answering Based Factuality Frameworks for Error Localization](#). *In EACL (main)*.

**Academic Services**

---

**Reviewer** (\*outstanding reviewer)

CVPR 2025, EMNLP 2024\*, COLM 2024, ARR (2023 - present), AISTATS 2021, BMVC 2020

**Organizer** - MASC-SLL 2025

**Volunteer** - EMNLP 2024

**NLP Colloquium JP** - Organizing a biweekly online talk event about NLP and CL

2024 -

## Educational Experience

---

### Mentorship

Avitej Iyer (M.S. student at Penn State University) Spring 2025

### Teaching Assistant, Penn State University

CMPSC 442: Artificial Intelligence Spring 2024

CMPSC 448: Machine Learning and AI Fall 2023, Fall 2024

## Media Mentions

---

[Self-Correction in Large Language Models](#). **Communications of the ACM**. Feb 26, 2025

## Honors and Awards

---

Keio University Global Fellowship - Scholarships for my graduate study at UT Austin 2020

Keio Engineering Foundation Award - Graduation with highest honors (First place in the Dept. of Mathematics) 2020

Japan Student Services Organization (JASSO) Exchange Student Scholarship 2018

## Invited Talks

---

Matuo-Iwasawa Lab at the University of Tokyo - Lecture about Self-Correction of LLMs 2024

NLP Colloquium JP - Talk about my research in detecting mistakes in LLM responses 2023

Nagoya NLP Seminar at Nagoya University - Talk about my research in fact verification and textual entailment 2023

## Skills

---

**Programming Languages:** Python

**Natural Languages:** English, Japanese

## Other Papers

---

Renze Lou, Hanzi Xu, Sijia Wang, Jiangshu Du, **Ryo Kamoi**, Xiaoxin Lu, Jian Xie, Yuxuan Sun, Yusen Zhang, Jihyun Janice Ahn, Hongchao Fang, Zhuoyang Zou, Wenchao Ma, Xi Li, Kai Zhang, Congying Xia, Lifu Huang, Wenpeng Yin. 2025. [AAAR-1.0: Assessing AI's Potential to Assist Research](#). In *ICML 2025*, *Best Paper Award at the 2nd AI4Research workshop @ AAAI 2025*.

Yusen Zhang, Wenliang Zheng, Aashrith Madasu, Peng Shi, **Ryo Kamoi**, Hao Zhou, Zhuoyang Zou, Shu Zhao, Sarkar Snigdha Sarathi Das, Vipul Gupta, Xiaoxin Lu, Nan Zhang, Ranran Haoran Zhang, Avitej Iyer, Renze Lou, Wenpeng Yin, Rui Zhang. 2025. [HRScene: How Far Are VLMs from Effective High-Resolution Image Understanding?](#) *arXiv preprint arXiv:2504.18406*.

Sarkar Snigdha Sarathi Das, **Ryo Kamoi**, Bo Pang, Yusen Zhang, Caiming Xiong, Rui Zhang. 2025. [GReaTer: Gradients over Reasoning Makes Smaller Language Models Strong Prompt Optimizers](#). In *ICLR 2025*.

Jihyun Janice Ahn, **Ryo Kamoi**, Lu Cheng, Rui Zhang, Wenpeng Yin. 2024. [Direct-Inverse Prompting: Analyzing LLMs' Discriminative Capacity in Self-Improving Generation](#). *arXiv preprint arXiv:2407.11017*.

Yilun Zhao, Yitao Long, Hongjun Liu, Linyong Nan, Lyuhao Chen, **Ryo Kamoi**, Yixin Liu, Xiangru Tang, Rui Zhang, Arman Cohan. 2023. [DocMath-Eval: Evaluating Numerical Reasoning Capabilities of LLMs in Understanding Long Documents with Tabular Data](#). In *ACL 2024*.

- Yusen Zhang, Nan Zhang, Yixin Liu, Alexander Fabbri, Junru Liu, **Ryo Kamoi**, Xiaoxin Lu, Caiming Xiong, Jieyu Zhao, Dragomir Radev, Kathleen McKeown, Rui Zhang. 2024. [Fair Abstractive Summarization of Diverse Perspectives](#). In *NAACL 2024*.
- Ryo Kamoi** and Kaname Tomite. 2021. [Efficient Unknown Object Detection with Discrepancy Networks for Semantic Segmentation](#). In *the NeurIPS 2021 Workshop on Machine Learning for Autonomous Driving*.
- Ryo Kamoi** and Kei Kobayashi. 2020. [Why is the Mahalanobis Distance Effective for Anomaly Detection?](#) *arXiv preprint arXiv:2003.00402*.
- Ryo Kamoi** and Kei Kobayashi. 2020. [Out-of-Distribution Detection with Likelihoods Assigned by Deep Generative Models Using Multimodal Prior Distributions](#). In *The AAAI's Workshop on Artificial Intelligence Safety*.
- Ryo Kamoi**, Kei Kobayashi. 2019. [Likelihood Assignment for Out-of-Distribution Inputs in Deep Generative Models is Sensitive to Prior Distribution Choice](#). *arXiv preprint arXiv:1911.06515*.