

# RIKIYA TAKEHI

web page

rikiya.takehi@fuji.waseda.jp

Tokyo, Japan

Github

## SUMMARY

I am a fourth-year undergraduate student supervised by Prof. Tetsuya Sakai at Waseda University, Japan. Previously, I was a guest researcher at [NIST Retrieval Team](#) in the US for one year, working with [Dr. Ian Soboroff](#) and [Dr. Ellen Voorhees](#) on the evaluation of search and texts. My research interest revolves around counterfactual evaluation, rankings, recommendations, and evaluation of search and text. I also collaborate closely with [Yuta Saito](#) of Cornell Univ., working on counterfactual evaluation and recommendations.

## EDUCATION

4/2021 – Present    **Waseda University - Major in Computer Science and Engineering**    Tokyo, Japan

## PUBLICATIONS

1. **Rikiya Takehi**, Kosuke Kawakami, Masahiro Asami, Yuta Saito. 2025. **General Framework for Off-Policy Learning with Partially-Observed Reward**. In *Proceedings of the International Conference on Learning Representations (ICLR '25)*. (acceptance rate = 32%)  
OpenReview |
2. **Rikiya Takehi**, Ellen M. Voorhees, Tetsuya Sakai, Ian Soboroff. 2025. **LLM-Assisted Relevance Assessments: When Should We Ask LLMs for Help?** In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '25)*. (acceptance rate = 21%)  
arXiv |
3. **Rikiya Takehi\***, Koichi Tanaka\*, Ren Kishimoto, Masahiro Nomura, Riku Togashi, Yuta Saito. 2025. **Optimizing Categorical Proportions in Calibrated Recommendations** (under review at **RecSys '25**)
4. **Rikiya Takehi**, Akihisa Watanabe, and Tetsuya Sakai. 2023. **Open-Domain Dialogue Quality Evaluation: Deriving Nugget-level Scores from Turn-level Scores** In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region (SIGIR AP '23)*. (acceptance rate = 36%)  
code | poster | slides | proceedings

## EXPERIENCE

- 10/2023 – 9/2024    **Guest Researcher, NIST - National Institute of Standards and Technology**    Maryland, USA
- Conducting research in the retrieval group with **Dr. Ian Soboroff & Dr. Ellen Voorhees**.
  - Worked on *evaluation of generated text via nugget-based evaluation*.
  - Conducted research on the efficient usage of LLMs along with limited manual annotations for building test collections. (full paper at **SIGIR 2025**)
- 9/2024 – Present    **Research Intern, CyberAgent AI Lab.**    Tokyo, Japan
- I proposed a framework to optimize the categorical proportions, instead of relying on a heuristic like calibrated recommendations (under review at **RecSys 2025**)
  - Proposing an algorithm on online dating applications, where instead of maximizing matches, we maximize the number of user retention, using real-world dataset of the company (on-going project)
- 9/2023 – Present    **Research Intern, Hakuodo Tech. Inc.**    remote
- I proposed a framework that leverages secondary rewards to help the optimization for **off-policy learning for partially observed rewards**. (full paper at **ICLR 2025**.)
- 7/2023 – Present    **Research Intern, Hanjukukaso Inc.**    remote
- Advisor: **Yuta Saito** (PhD student at Prof. Joachims Lab, Cornell Univ.)
  - Working on **fairness in ranking** and **off-policy evaluation & learning** mainly tackling **RecSys problems**.
- 4/2023 – Present    **Student Researcher, Sakai Lab - Waseda Univ.**    Tokyo, Japan
- Advisor: **Prof. Tetsuya Sakai**
  - Conducting research mainly on **evaluation of IR**. (one on-going project)
  - Research on **evaluation of dialogue systems**. (short paper at **SIGIR AP 2023**)

## CURRENT PROJECTS (SELECTED)

- **Mood-Aware Recommendations** (first authoring)
  - Collaboration with Prof. **Fernando Diaz** of CMU.

- This study introduces a new insight that naive ranking algorithms do not maximize click rates in the recommendation settings when considering user mood. We formulate this problem and propose a method that theoretically maximizes the click probability of a user.
- **Optimizing Matching Markets for User Retention** (co-first authoring)
  - Collaboration with **Yuta Saito** et al.
  - Recommendations on online dating platforms usually aim to maximize the number of matches. However, this often leads to the retention of some users, and this is undesired because matching applications are usually subscription services. We study a way to maximize the number of users that stays on the platform.

## TALKS

---

- **National Institute of Informatics** (Hosted by **Prof. Noriko Kando**) - *Using LLMs as Assistants for Building Test Collections (+Trends and Problems of Test Collections)*
- **University of Maryland College Park** (Hosted by **Prof. Douglas Oard**) - *Nugget-Based Evaluation and the Use of LLMs*

## AWARDS

---

- **Project Research Award**, Waseda University (Mar 2025)  
Recognized for exemplary research activity as an undergraduate. (awarded to 2 out of 190 students in Dept.)

## LEADERSHIP & PROFESSIONAL SERVICE

---

**Co-organizer of the Product Search and Recommendations Track** ([website here](#)):

- Co-organizing with **Prof.ChengXiang Zhai** (Illinois), **Prof.Michael Ekstrand** (Drexel), **Dr.Surya Kallumadi** (Lowe's), **Dr.Alessandro Magnani** (Walmart), and **Dr.Daniel Campos** (Snowflake)
- Will be on a TREC Track in 2025. This will be the first track for recommendations at TREC.

### Panelist at NTCIR 2025:

- I was invited as a panelist at **NTCIR 2025** to discuss LLMs for evaluation.
- Along with **Prof.Mark Sanderson**, **Prof.Charles Clarke**, and **Dr.Ian Soboroff**.

## SKILLS

---

- Programming languages
  - Experienced with Python (along with PyTorch, Tensorflow, Huggingface, etc)
  - Familiar with C++, HTML, CSS, Java, JavaScript
- Japanese - native, English - fluent (TOEFL iBT: 114), French - proficient (CEFR B2)