

Hossein A. (Saeed) Rahmani

CONTACT INFORMATION
✉ hossein.rahmani.22@ucl.ac.uk
🐙 github.com/rahmanidashti
in [linkedin.com/in/rahmanidashti](https://www.linkedin.com/in/rahmanidashti)
🏠 rahmanidashti.github.io/
G Google Scholar

UCL Centre for Artificial Intelligence
Dept. of Computer Science
90 High Holborn
University College London
United Kingdom, WC1V 6BH

🔗 **Internship:** *Actively seeking internship or student research opportunities throughout the year. Moreover, my position at the Alan Turing Institute does not have any conflict with the internship positions.*

RESEARCH INTERESTS

<input type="checkbox"/> Multi-Agent Evaluation and Self-Improvement	<input type="checkbox"/> Synthetic Data Generation
<input type="checkbox"/> LLM-as-a-Judge	<input type="checkbox"/> LLM Evaluation and Benchmark
<input type="checkbox"/> LLM Task Planning	<input type="checkbox"/> LLM Personalization
<input type="checkbox"/> Responsible RecSys/Search	<input type="checkbox"/> Search and Retrieval

EDUCATION



University College London <i>Ph.D. in Computer Science</i> <ul style="list-style-type: none">• Affiliation: UCL Web Intelligent Group• Advisors and Mentors: Prof. Emine Yilmaz and Nick Craswell	London, United Kingdom Dec. 2022 – Nov. 2026 (<i>Expected</i>)
University of Zanjan <i>M.Sc. in Computer Engineering (Distinction)</i> <ul style="list-style-type: none">• Thesis: Context-aware Point-of-Interest Recommender System• Advisors and Mentors: Prof. Mohsen Afsharchi and Dr. Mitra Baratchi	Zanjan, Iran Sep. 2016 – Feb. 2020

INTERNSHIP EXPERIENCE

The Alan Turing Institute PhD Enrichment Student <ul style="list-style-type: none">• Team: Understanding the Capabilities of Foundation Models• Working on Reasoning and Self-Improvement	London, United Kingdom Jan. 2025 – Jun. 2025
Thomson Reuters Labs Applied Scientist Intern - NLP <ul style="list-style-type: none">• Supervised by Prof. Dell Zhang and Dr. Grace Lee• Working on LLMs for Legal Tasks Evaluation	London, United Kingdom Sep. 2023 – July. 2024
Amazon Alexa AI Alexa Prize TaskBot Challenge <ul style="list-style-type: none">• \$250,000 Research Grant	London, United Kingdom Apr. 2021 – Dec. 2022
Università della Svizzera italiana (USI) Research Intern, Information Retrieval Group <ul style="list-style-type: none">• Supervised by Prof. Fabio Crestani• Working on Context-aware Recommendation Systems	Lugano, Switzerland Sep. 2019 – Dec. 2019








MEDIA COVERAGE

Microsoft Research Blog [e] Highlighted <i>JudgeBlender</i> as a cost-effective technique for automatic judgment.	2025
Thomson Reuters Innovation [e] Featured my work on <i>LLM-as-a-Judge</i> as part of my internship at TR Labs as an innovation.	2025
Thomson Reuters Lab Medium [e] Published my articles on <i>Can LLMs Replace Human Judgement in Evaluation?</i>	2025






- UKRI CDT In Foundational AI**  2024
Two Foundational AI CDT students awarded places on the Turing Enrichment Scheme
- Microsoft Research Blog**  2024
Highlighted *SynDL* as a resource for synthetic evaluation in information retrieval.

SELECTED
PUBLICATIONS
Full list, see


Preprints & Under Submission

22. **Are Large Language Models Task-based Self-Correction?** *Working Progress*
Hossein A. Rahmani, Satyapriya Krishna, Marwa Abdulhai, Emine Yilmaz
21. **Benchmarking LLMs for Task Understanding and Planning** *Working Progress*
Hossein A. Rahmani, Fanghua Ye, Bin Wu, Marzieh Fadaee, Emine Yilmaz
20. **Towards Self-Improvement through Multi-agent Debate Feedback** *Under Review*
Hossein A. Rahmani, Emine Yilmaz, Nick Craswell, Bhaskar Mitra
19. **Do LLMs Favor Towards Synthetic Test Collection for Evaluation?** *Under Review*
Hossein A. Rahmani, Varsha Ramineni, Emine Yilmaz, Nick Craswell, Bhaskar Mitra
18. **Judging the Judges: A Collection of LLM-Generated Relevance Judgements** *arXiv 2025*
Hossein A. Rahmani, Clemencia Siro, Mohammad Aliannejadi, Nick Craswell, Charles L. A. Clarke, Guglielmo Faggioli, Bhaskar Mitra, Paul Thomas, Emine Yilmaz
  
17. **Large Language Models for Legal Tasks Evaluation** *Under Review*
Hossein A. Rahmani, Risako Owan, Grace E. Lee, Sally Gao, Frank Schilder, Dell Zhang, Emine Yilmaz
16. **LLMJudge: LLMs for Relevance Judgments** *arXiv 2024*
Hossein A. Rahmani, Emine Yilmaz, Nick Craswell, Bhaskar Mitra, Paul Thomas, Charles L. A. Clarke, Mohammad Aliannejadi, Clemencia Siro, Guglielmo Faggioli
 
15. **Understanding the Role of User Profile in the Personalization of LLMs** *arXiv 2024*
Bin Wu, Zhengyan Shi, Hossein A. Rahmani, Varsha Ramineni, Emine Yilmaz
 

(Selected) Peer-reviewed Conference

14. **JudgeBlender: Ensembling Judgments for Automatic Relevance Assessment** *TheWebConf, 2025*
Hossein A. Rahmani, Emine Yilmaz, Nick Craswell, Bhaskar Mitra

13. **SynDL: A Large-Scale Synthetic Test Collection for Passage Retrieval** *TheWebConf, 2025*
Hossein A. Rahmani, Xi Wang, Emine Yilmaz, Nick Craswell, Bhaskar Mitra, Paul Thomas
  
12. **Beyond Internal Data: Constructing Complete Datasets for Fairness Testing** *NeurIPS, 2024*
Varsha Ramineni, Hossein A. Rahmani, David Berber, Emine Yilmaz
Algorithmic Fairness through the Lens of Metrics and Evaluation at NeurIPS 2024
11. **LLM4Eval: Large Language Model for Evaluation in IR** *SIGIR, 2024*
Hossein A. Rahmani, Clemencia Siro, Mohammad Aliannejadi, Nick Craswell, Charles L. A. Clarke, Guglielmo Faggioli, Bhaskar Mitra, Paul Thomas, Emine Yilmaz


10. **Synthetic Test Collections for Retrieval Evaluation** *SIGIR, 2024*
Hossein A. Rahmani, Nick Craswell, Emine Yilmaz, Bhaskar Mitra, Daniel Campos
 
9. **Natural Language User Profiles for Transparent and Scrutable Recommendations** *ACL, 2024*
Jerome Ramos, Hossein A. Rahmani, Xi Wang, Xiao Fu, Aldo Lipani
 
8. **Clarifying the Path to User Satisfaction: An Investigation into Clarification Usefulness** *EACL, 2024*
Hossein A. Rahmani, Xi Wang, Mohammad Aliannejadi, Mohammadmehdi Naghiaei, Emine Yilmaz
7. **Improving Conversational Recommendation Systems via Bias Analysis and Language-Model-Enhanced Data Augmentation** *EMNLP, 2023*
Xi Wang, Hossein A. Rahmani, Jiqun Liu, Emine Yilmaz
6. **Do LLMs Selectively Encode the Goal of an Agent's Reach?** *TOM@ICML, 2023*
Laura Ruis, Arduin Findeis, Herbie Bradley, Hossein A. Rahmani, Kyoung Whan Choe, Edward Grefenstette, Tim Rocktaschel
5. **A Survey on Asking Clarification Questions Datasets in Conversational Systems** *ACL, 2023*
Hossein A. Rahmani, Xi Wang, Yue Feng, Qiang Zhang, Emine Yilmaz, Aldo Lipani
4. **When and What to Ask Through World States and Text Instructions** *IGLU@NeurIPS, 2023*
Zhengxiang Shi, Jerome Ramos, To Eun Kim, Xi Wang, Hossein A. Rahmani, Aldo Lipani
3. **Towards Confidence-aware Calibrated Recommendation** *CIKM, 2022*
Mohammadmehdi Naghiaei, Hossein A. Rahmani, Mohammad Aliannejadi, Nasim Sonboli
 
2. **Exploring the Impact of Temporal Bias in Point-of-Interest Recommendation** *RecSys, 2022*
Hossein A. Rahmani, Mohammadmehdi Naghiaei, Ali Tourani, Yashar Deldjoo
 
1. **Experiments on Generalizability of User-Oriented Fairness in Recommender Systems** *SIGIR, 2022*
Hossein A. Rahmani, Mohammadmehdi Naghiaei, Mahdi Dehghan, Mohammad Aliannejadi
 
0. **CPFair: Personalized Consumer and Producer Fairness Re-ranking for Recommender Systems** *SIGIR, 2022*
Mohammadmehdi Naghiaei, Hossein A. Rahmani, Yashar Deldjoo
 

 AWARDS, GRANTS, AND FUNDING	Alan Turing Enrichment Schema Placement Award, The Alan Turing Institute, UK	2024
	Ranked 1st in TREC iKAT 2023, LLMs for Conversational Search	2023
	NeurIPS 2022 IGLU Challenge Research Prize	2022
	Amazon Alexa Prize Competition Research Grant	2022
	UCL PhD Studentship, UKRI CDT in Foundational AI, UCL	2022 – 2026
	School of AI Scholarship, Pi School, Italy	2021
	Student Registration Grant for attending CSCW	2021
	Student Registration Grant for attending CIKM	2021
	Third-rank Graduate Amongst MSc Students, University of Zanjan, Iran	2020

	Student Registration and Travel Grant for attending MediaEval Benchmark	2019
	Student Registration Grant for attending CLEF	2019
	SERI Mobility Grant, Zurich University of Applied Sciences (ZHAW), Switzerland	2019
	Data Science Winter School Grant, University of Bonn, Germany	2018
	Third-rank Graduate Amongst BSc Students, University of Zanjan, Iran	2016
	Ranked 3rd in Technical University Olympiad, Technical and Vocational University, Iran	2013
	Honorary Admission to Bachelor Program, Technical University of Shiraz, Iran	2013
WORK EXPERIENCE	Virtual Manufacturing Inc. Applied Scientist	California, US Oct. 2021 – Nov. 2022
	TVConal Machine Learning Engineer	Singapore, Singapore Jul. 2021 – Oct. 2021
	Pi School Research Scholar	Room, Italy Apr. 2021 – Jun. 2021
	Sanay Co. Data Scientist	Zanjan, Iran Apr. 2020 – Jan. 2021
TEACHING EXPERIENCE	University College London , Teaching Assistant <i>Statistical Natural Language Processing</i> (Prof. Pontus Stenetorp) <i>Artificial Intelligence for Sustainable Development</i> (Prof. Maria Perez-Ortiz) <i>Foundation of Artificial Intelligence</i> (Dr. Sahan Bulathwela) <i>Information Retrieval and Data Mining</i> (Prof. Emine Yilmaz)	London, United Kingdom
INVITED TALKS	Fairness and Bias in Recommender Systems QIAU, Iran	2023
	Generability of User Fairness in Recommender Systems Mila x Vector Institute, Canada	2023
	Personalized Context-Aware POI Recommendation Universita della Svizzera Italiana, Switzerland	2019
SERVICES	Organizational Activities	
	• Lead-organizing LLM4Eval Workshop at WSDM 2025	Feb. 2025 – Mar. 2025
	• Organizing LLMJudge Challenge at SIGIR 2024	Feb. 2024 – Aug. 2024
	• Co-organizing LLM4Eval Workshop at SIGIR 2024	Feb. 2024 – Jul. 2024
	• Co-organizing TREC Deep Learning Track 2023	Jul. 2023 – Nov. 2023
	• WITalk: WI Talk Series on NLP and IR, UCL AI Center	Oct. 2023 – Feb. 2024
	Program Committee and/or Reviewer	
	• ML Venues: ICLR (2025), NeurIPS (2024), AAAI (2023, 2024)	
	• NLP Venues: ACL ARR (Feb. 2023 – Present), ACL (2023, 2024), EMNLP (2022 – 2024)	
	• IR Venues: ECML/PKDD (2021), WSDM (2021, 2025), SIGIR (2022 - 2025), ECIR (2022 – 2024)	
SKILLS	• Programming Languages: Python, Java, C# • Tools: PyTorch, HuggingFace, NumPy, SciPy, Pandas, Seaborn, NLTK, Matplotlib, Flask, RESTful API • Database Systems: MS SQL Server, MySQL, SQLite, MongoDB • Soft Skills: Time Management, Team working, Problem-solving, Fast-Learning • Languages: Persian (Native), Tabari (Fluent), English (Proficient), Arabic (Familiar)	