

Shubham Chatterjee

✉ shubham.chatterjee94@gmail.com • 🌐 <https://homepages.inf.ed.ac.uk/schatte4/index.html>

🐙 <https://github.com/shubham526> • **in** www.linkedin.com/in/shubham-chatterjee

EDUCATION

University of New Hampshire, Durham, USA **Completed: September 2022**
Doctor of Philosophy (PhD) in Computer Science **GPA: 3.83/4**

University of New Hampshire, Durham, USA **Completed: December 2020**
Master of Science (MS) in Computer Science **GPA: 3.83/4**

University of Calcutta, Kolkata, India **Completed: May 2017**
Master of Science (MSc) in Computer Science **GPA: 8.54/10**

University of Calcutta, Kolkata, India **Completed: May 2015**
Bachelor of Science (BSc) in Computer Science **GPA: 8.43/10**

WORK EXPERIENCE

Postdoctoral Research Associate, GRILL Lab. Mentor: Dr. Jeff Dalton **October 2023 – present**
University of Edinburgh, Scotland

- Research in neural IR and interactive conversational IR.
- Leading, developing and maintaining collaborations with academic/industry partners.
- Organization, supervision, mentoring, and training of PhD students.

Postdoctoral Research Associate, GRILL Lab. Mentor: Dr. Jeff Dalton **December 2022 – September 2023**
University of Glasgow, Scotland

- Research in neural IR and interactive conversational IR.
- Leading, developing and maintaining collaborations with academic/industry partners.
- Organization, supervision, mentoring, and training of PhD students.

Postdoctoral Research Associate, TREMA Lab. Mentor: Dr. Laura Dietz **September 2022 – November 2022**
University of New Hampshire, Durham, USA

- Researched and improved entity-oriented understanding of long-form text.
- Researched and improved learning query-specific entity embeddings for IR tasks.
- Took a leading role in the organization, supervision, mentoring, and training of PhD students.

Teaching Assistant, Department of Computer Science **August 2017 – August 2020**
University of New Hampshire, Durham, USA

- Taught labs and recitations for undergraduate and graduate computer science courses.
- Collaborated with faculty to create new coursework for courses.
- Supported students via regular in-person and remote (Zoom) office hours.

RESEARCH INTERESTS

- Neural Entity-Oriented IR
- Information Extraction for Text Understanding
- Entity Ranking
- Knowledge Graphs for IR
- Representation Learning for IR
- Conversational IR and Question Answering
- LLMs for IR

PROJECTS

Document Ranking with Entity-based Query Understanding

Postdoc Project [Ongoing]

December 2022 – Present

- Proposed a new method of learning query-specific document representations that incorporate knowledge about what is relevant for understanding the query.
- Showed that the approach can outperform several hard neural and non-neural baselines on three large-scale document ranking datasets by 40–60%.

Personal Knowledge Graphs for Interactive Conversational Search

Postdoc Project [Ongoing]

December 2022 – Present

- Leading a collaborative project with the Radboud University, Netherlands.
- Collecting dataset for personalized recommendation (food/music/etc.), to be useful for an interactive conversational assistant.
- Involved in supervision and mentoring of PhD students and staff associated with the project at Radboud University and University of Glasgow.

Generative Relevance Modelling with LLM-Generated Reports

Postdoc Project [Ongoing]

February 2022 – Present

- Using an LLM to generate a diverse set of subtopics and related content for a given query.
- Studying the application of LLMs for reasoning about relevance in IR.

Relevance Feedback Using Large Language Models

Postdoc Project [Ongoing]

February 2022 – Present

- Studied the application of generative models for pseudo-relevance feedback in IR.
- Showed that LLMs can generate “relevant” documents that can improve document retrieval performance by 30–40% on two large-scale document ranking datasets. Paper published at SIGIR 2023.

Learning Query-Specific Entity-Oriented Latent Spaces for IR

Postdoc Project

May 2022 – November 2022

- Proposed a new method of learning query-specific latent entity spaces that are useful for IR.
- Showed that the approach can outperform several hard neural and non-neural baselines on two large-scale entity ranking datasets by 40–60%. Paper under review.

Predicting Entities as Guides for Long-Form Text Similarity

PhD Project

March 2022 – May 2022

- Proposed a new method for entity-oriented text understanding. Specifically, given a query/question/etc, it is possible to predict entities that act as “guides” to the most relevant document. Studied this in the context of the entity aspect linking task.
- Showed that the approach can outperform several hard neural and non-neural baselines on a large-scale dataset by 40%. Paper published at CIKM 2022.

Query-Specific BERT Entity Embeddings

PhD Project

May 2021 – October 2021

- Showed that using query-specific entity embeddings in an entity ranking system can outperform several neural and non-neural entity ranking systems by 13–42% on two large-scale entity ranking test collections.
- Showed that query-specific entity embeddings can promote relevant entities to the top of the ranking. Paper published at SIGIR 2022.

Entity Retrieval Using Fine-Grained Entity Aspects

PhD Project

January 2021 – May 2021

- Proposed a new entity ranking method that derives entity relevance indicators by understanding the meaning of the entity in the context of the query (entity aspects).
- Showed that the use of fine-grained entity aspects can improve performance by 50% over the current state-of-the-art on a large-scale entity ranking test collection. Paper published at SIGIR 2021

Entity Support Passage Retrieval

PhD Project

January 2019 – December 2019

- Designed a joint probabilistic model which captures the joint relevance of (1) entity to query, (2) passage to query

and (3) passage to entity.

- o Achieved a 21% improvement in terms of Mean Average Precision as compared to state-of-the-art baselines for the task. Paper published at ICTIR 2019.

RECENT PUBLICATIONS

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

- o Pooja Oza, Shubham Chatterjee, and Laura Dietz. Neural Entity Context Models. In *Proceedings of the the 12th International Joint Conference on Knowledge Graphs, IJCKG '23*. Association for Computing Machinery, 2023
- o Iain Mackie, Shubham Chatterjee, and Jeff Dalton. Generative Relevance Feedback with Large Language Models. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '23*. Association for Computing Machinery, 2023
- o Iain Mackie, Shubham Chatterjee, and Jeffrey Dalton. Generative and Pseudo-Relevant Feedback for Sparse, Dense and Learned Sparse Retrieval. In *Proceedings of the Workshop on Large Language Models' Interpretation and Trustworthiness (LLMIT), CIKM '23*. CEUR, 2023
- o Iain Mackie, Ivan Sekulic, Shubham Chatterjee, Jeffrey Dalton, and Fabio Crestani. GRM: Generative Relevance Modeling Using Relevance-Aware Sample Estimation for Document Retrieval. arXiv, <https://arxiv.org/abs/2306.09938>, 2023
- o Shubham Chatterjee and Laura Dietz. Predicting Guiding Entities for Entity Aspect Linking. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management, CIKM '22*. Association for Computing Machinery, 2022
- o Shubham Chatterjee and Laura Dietz. BERT-ER: Query-Specific BERT Entity Representations for Entity Ranking. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '22*. Association for Computing Machinery, 2022
- o Laura Dietz, Shubham Chatterjee, Connor Lennox, Sumanta Kashyapi, Pooja Oza, and Ben Gamari. Wikimarks: Harvesting Relevance Benchmarks from Wikipedia. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '22*. Association for Computing Machinery, 2022
- o Shubham Chatterjee. An Entity-Oriented Approach for Answering Topical Information Needs. In *Proceedings of the 44th European Conference on Information Retrieval, ECIR '22*. Springer, 2022
- o Shubham Chatterjee and Laura Dietz. Entity Retrieval Using Fine-Grained Entity Aspects. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '21*. Association for Computing Machinery, 2021
- o Shubham Chatterjee and Laura Dietz. Why Does This Entity Matter? Support Passage Retrieval for Entity Retrieval. In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval, ICTIR '19*. Association for Computing Machinery, 2019

INVITED TALKS AND GUEST LECTURES

- o BBC Data Dates Series, UK Invited Talk
- o Curai Health, USA Invited Talk
- o Radboud University, The Netherlands Invited Talk
- o Glasgow IR Seminar Series, UK Invited Talk
- o University of New Hampshire, Durham, USA Guest Lecture

TEACHING

I was a teaching assistant in the Department of Computer Science at the University of New Hampshire, Durham, USA where I assisted with the following courses:

- o CS 415 | Introduction to Computer Science I Undergraduate
- o CS 416 | Introduction to Computer Science II. Undergraduate
- o CS 414 | From Problems to Algorithms to Programs. Undergraduate
- o CS 417 | From Programs to Computer Science. Undergraduate
- o CS 515 | Data Structures and Introduction to Algorithms. Undergraduate
- o CS 619 | Introduction to Object-oriented Design and Development. Undergraduate
- o CS 671 | Programming Language Concepts and Features. Undergraduate
- o CS 853 | Information Retrieval. Undergraduate/Graduate

MENTORING

- o Neelabha Banerjee | Christ University, India
- o Iain Mackie | University of Glasgow, UK
- o Paul Owoicho | University of Glasgow, UK

MSc Data Science
PhD Computer Science
PhD Computer Science

REVIEWING SERVICES

- o European Conference on Information Retrieval (ECIR).
- o ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR).
- o ACM International Conference on Information and Knowledge Management (CIKM).
- o ACL Rolling Review.
- o ACM Transactions on Information Systems (TOIS).
- o Springer Artificial Intelligence Review (AIR).

ORGANIZING

- o Organizing the new TREC Interactive Knowledge Assistance track.
- o Organized a tutorial on “Neuro-Symbolic Approaches for IR” at ECIR 2023 and SIGIR 2023.
- o Area chair for the Information Retrieval track at the 12th International Joint Conference on Knowledge Graphs (IJCKG 2023).

AWARDS

- o Outstanding reviewer award. ECIR 2022.
- o Dissertation Year Fellowship, University of New Hampshire, Durham. 2021.
- o Student travel grants to attend conferences (ICTIR 2019, ECIR 2022, SIGIR 2022).
- o INSPIRE Fellowship, Government of India. 2012.

REFERENCES

1. **Laura Dietz.** *Associate Professor.*
Department of Computer Science,
University of New Hampshire, Durham, USA.
@ dietz@cs.unh.edu
2. **Jeff Dalton.** *Associate Professor.*
School of Informatics,
University of Edinburgh, Scotland, UK.
@ jeff.dalton@ed.ac.uk
3. **Arjen de Vries.** *Professor.*
Department of Computer Science,
Radboud University, The Netherlands.
@ a.devries@cs.ru.nl
4. **Debabrata Datta.** *Assistant Professor.*
Department of Computer Science,
St. Xavier's College (Autonomous), Kolkata, India.
@ debabrata.datta@sxccal.edu
5. **Anal Acharya** *Assistant Professor*
Department of Computer Science,
St. Xavier's College (Autonomous), Kolkata, India.
@ anal.acharya@sxccal.edu