

PROCHETA SEN

Web page: <https://procheta.github.io/>
Github Code Repositories: <https://github.com/procheta>
DBLP: <http://dblp.uni-trier.de/pers/hc/s/Sen:Procheta>
Google Scholar: <https://scholar.google.com/citations?hl=en&user=hRKyQdoAAAAJ>
Email: procheta.sen@liverpool.ac.uk

WORK

<i>Lecturer,</i> University of Liverpool, United Kingdom.	July 2022 - Present
<i>Postdoctoral Researcher,</i> University College London, United Kingdom.	November 2021 - July 2022
<i>Research Assistant,</i> University of Glasgow, United Kingdom.	March 2021 - September 2021

EDUCATION

Topic: Proactive Information Retrieval <i>Doctor of Philosophy</i> , Dublin City University, Dublin, Ireland. Supervised by: Prof. Gareth J. F. Jones.	Nov 2016 - September 2021
<i>Master of Technology</i> , Computing Science Indian Statistical Institute, Kolkata, India	July 2016
<i>Bachelor of Technology</i> , Computing Science Institute of Engineering and Management, West-Bengal, India	Jun 2013

GRANTS

- Impact Accelerator Grant, University of Liverpool 2024.
- Impact Accelerator Grant, University of Liverpool 2023.
- EPSRC Student Summer Internship Fund 2023.
- Knowledge Transfer Partnership Fund with Carpenter Solutions Pvt. Limited, UK.

TEACHING

- Data Mining and Visualisation, University of Liverpool

PROFESSIONAL SERVICE

- Area Chair (Ethics, Bias, Fairness Track) Coling 2025
- Ethics Committee lead at the Department of Computer Science at the University of Liverpool.
- Tutorial on Explainability in Information Retrieval SIGIR 2023, FIRE 2023.
- Reviewer: AAAI, ECIR, SIGIR, CIKM, ACL, TOIS, Artificial Intelligence and Law Journal
- Organiser of NLP for Social Good Symposium 2023,2024 at the University of Liverpool
- Organiser of RCD-FIRE 2020: Retrieval From Conversational Dialogues in FIRE 2020
- Organiser of PIR-CLEF 2018: Evaluation Of Personalised Information Retrieval in CLEF 2018

PhD. Students:

- Lingfang Li (Primary Supervisor) : Explaining Large Language Models (August 2023-Present)
- Zhilu Zhang (Secondary Supervisor) : Leveraging Financial News for Portfolio Optimization (September 2022-Present)
- Siqi Sun (Secondary Supervisor) : Investigating Adversarial Robustness in Language Models: Adversarial Attacks, Certification, and Defense (Successfully Defended).
- Tong Chen (Secondary Supervisor) : Fact Checking Using LLMs

SELECTED PUBLICATIONS

- Tong Chen, Zimu Wang, Yiyi Miao, Haoran Luo, Yuanfei Sun, Wei Wang, Zhengyong Jiang, Procheta Sen, Jionglon Su: MedFact: A Large-scale Chinese Dataset for Evidence-based Medical Fact-checking of LLM Responses. EMNLP 2025
- Matteo Tucat, Anirbit Mukherjee, Procheta Sen, Mingfei Sun, Omar Rivasplata: Regularized Gradient Clipping Provably Trains Wide and Deep Neural Networks. TMLR. 2025
- Matteo Tucat, Anirbit Mukherjee, Procheta Sen, Mingfei Sun, Omar Rivasplata: Regularized Gradient Clipping Provably Trains Wide and Deep Neural Networks. TMLR. 2025
- Xi Wang, **Procheta Sen**, Ruizhe Li, Emine Yilmaz: Adaptive Retrieval-Augmented Generation for Conversational Systems, NAACL (Findings) 2025
- Lingfang Li, **Procheta Sen** Unraveling the Influence of Training Data and Internal Structures in Large Language Models for Enhanced Explainability, AAAI 2025 (Student Abstract)
- Subinay Adhikary, **Procheta Sen**, Dwaipayan Roy, Kripabandhu Ghosh, A case study for automated attribute extraction from legal documents using large language models, Artificial Intelligence and Law Journal, 2024
- Siqi Sun, **Procheta Sen**, Wenjie Ruan CROWD: Certified Robustness via Weight Distribution for Smoothed Classifiers against Backdoor Attack, EMNLP (Findings) 2024
- Bhavik Chandna, **Procheta Sen** A Counterfactual Explanation Framework for Retrieval Models. <https://arxiv.org/abs/2409.00860> (Under Review).
- Xi Wang, **Procheta Sen**, Ruizhe Li, Emine Yilmaz, Simulated Task Oriented Dialogues for Developing Versatile Conversational Agents. ECIR 2024.
- Zhilu Zhang, **Procheta Sen**, Zimu Wang, Ruoyu Sun, Zhengyong Jiang, Jionglon Su, FinBPM: A Framework for Portfolio Management-based Financial Investor Behavior Perception Model. EACL 2024
- Zhilu Zhang, **Procheta Sen**, Zimu Wang, Ruoyu Sun, Zhengyong Jiang, Jionglon Su, FinBPM: A Framework for Portfolio Management-based Financial Investor Behavior Perception Model. EACL 2024
- Xiaohang Tang, Yi Zhou, Taichi Aida, **Procheta Sen**, Danushka Bollegala: A Word Sense Distribution-based approach for Semantic Change Prediction, EMNLP (Findings) 2023
- Zhengxiang Shi, **Procheta Sen**, Aldo Lipani: Lexical Entrainment for Conversational Systems, EMNLP (Findings) 2023.
- Avishhek Anand, **Procheta Sen**, Sourav Saha, Manisha Verma, Mandar Mitra:
- **Procheta Sen**, Sourav Saha, Debasis Ganguly, Manisha Verma and Dwaipayan Roy: Measuring and Comparing the Consistency of IR Models for Query Pairs with Similar and Different Information Needs, CIKM 2022
- Ishani Mondal, **Procheta Sen** and Debasis Ganguly Multi-Objective Few-shot Learning for Fair Classification, CIKM 2021
- **Procheta Sen**, Debasis Ganguly and Gareth Jones I know what you need: Investigating Document Retrieval Effectiveness with Partial Session Contexts, ACM TOIS 2021
- **Procheta Sen**, Debasis Ganguly, Manisha Verma and Gareth Jones The Curious Case of IR Explainability: Explaining Document Scores within and across Ranking Models, SIGIR 2020
- **Procheta Sen**, Debasis Ganguly Towards Socially Responsible AI: Cognitive Bias-Aware Multi-Objective Learning, AAAI 2020
- **Procheta Sen**, Debasis Ganguly, and Gareth Jones Word-Node2Vec: Improving Word Embedding with Document-Level Non-Local Word Co-occurrences, NAACL 2019

- **Procheta Sen**, Debasis Ganguly, and Gareth Jones Procrastination is the Thief of Time: Evaluating the Effectiveness of Proactive Search Systems, SIGIR 2018
- **Procheta Sen**, Debasis Ganguly, and Gareth Jones Tempo-Lexical Context driven Word Embedding for Cross-Session Search Task Extraction, NAACL 2018
- Gabriella Pasi, Gareth J. F. Jones, Keith Curtis, Stefania Marrara, Camilla Sanvitto, Debasis Ganguly, **Procheta Sen** Overview of the CLEF 2018 Personalised Information Retrieval Lab (PIR-CLEF 2018)