Suchana Datta

+353 894136199 | suchana.datta@ucd.ie | https://suchanadatta.github.io | https://github.com/suchanadatta | suchana-datta-94ba942b/ | https://dblp.org/pid/270/6658.html

Myself

I am a Postdoctoral Research Fellow at Insight Centre for Data Analytics in University College Dublin, Ireland, where I am working with the Victeur team as part of the UCD Centre for Cultural Analytics. While my current research project, Victeur (an ERC-funded project), focuses on developing computational methods to explore the role of European Migrants in Victorian and Neo-Victorian Culture, my research interest broadly includes adhoc information retrieval (IR), query performance predictions (QPP), cultural analytics and explainability. Before joining as a Postdoc, I graduated with a PhD in Computer Science from University College Dublin, Ireland in 2024. My PhD supervisor was Assoc. Prof. Derek Greene from School of Computer Science, UCD. I was affiliated with the Insight Centre for Data Analytics during my PhD. While my earlier focus was on adhoc information retrieval, my PhD research aimed to exploring causality-driven IR, where I pursued novel search interfaces and algorithms towards addressing causality-driven search, i.e. where a search user wants to find out documents answering the 'Why' question, e.g. why was Osama bin Laden assassinated? etc. I am now looking forward to a suitable job role that would drive me pursuing my academic career further.

Education

University College Dublin *Dublin, Ireland*

PhD in Computer Science *May 2020 - April 2024*

- PhD topic : Causality-driven ad-hoc Information Retrieval.
- Research Interest : Information Retrieval, Natural Language Processing, Machine Learning, Cultural Analytics, Explainable AI.

Maulana Abul Kalam Azad University of Technology *Kolkata, India*

M.Tech in Computer Science & Engineering *August 2014 - September 2016*

- Thesis Topic : Malicious host recognition via Cloud Forensics.

Maulana Abul Kalam Azad University of Technology *Kolkata, India*

B.Tech in Computer Science & Engineering *August 2009 - September 2013*

- Thesis Topic : Analyzing social media networks.

Work Experience

University College Dublin *Dublin, Ireland*

Postdoctoral Fellow *February 2019 - present*

- Research project : VICTEUR (an ERC-funded project), focuses on developing computational methods to explore the role of European Migrants in Victorian and Neo-Victorian Culture.

University College Dublin *Dublin, Ireland*

Occasional Lecturer *September 2024 - present*

- Subjects taught : Data Science in Python - Undergraduate and postgraduate level.

University College Dublin *Dublin, Ireland*

Teaching Assistant / Demonstrator *September 2020 - present*

- Subjects taught : Introduction to Java, Data Structure with Java, Data Science in Python, Network Analysis.

Xcelerator *Kolkata, India*

Assistant Professor *August 2019 - April 2020*

- Subjects taught : C, JAVA, Database Management Systems, Data Structures, Operating Systems, Software Engineering.

Techno India University *Kolkata, India*

Assistant Professor *April 2018 - July 2019*

- Subjects taught : C, JAVA, DBMS, Data Structures, Operating Systems, Networking.

Prasanta Chandra Mahalanabish College *Kolkata, India*

Guest Lecturer *January 2017 - July 2018*

- Subjects taught : C, Data Structures, Algorithms.

Indian Statistical Institute *Kolkata, India*

Project-linked Person *November 2016 - April 2018*

- Research project : Unsupervised learning algorithms for deriving insight from text data and building an intelligent query suggestion system.

Maulana Abul Kalam Azad University of Technology *Kolkata, India*

Teaching Assistant *August 2014 - August 2016*

- Subjects taught : C, C++, JAVA, Data Structures, Algorithms, Theory of Computations.

Cognizant Technology Solutions *Kolkata, India*

Software Developer *July 2013 - July 2014*

- Project : Insurance database management, Gas & poer in UK, Electronics equipments, Insurance in Middle East

Publications

2025

- **Suchana Datta**, Guglielmo Faggioli, Nicola Ferro, Debasis Ganguly, Cristina Ioana Muntean, Raffaele Perego, Nicola Tonellotto: Projection-Displacement based Query Performance Prediction for Embedded Space of Dense Retrievers. TOIS 2025 [Accepted (to appear soon)].
- **Suchana Datta**, Dwaipayan Roy, Derek Greene, Gerardine Meaney: Tales and Truths: Exploring the Linguistic Journey of 19th Century Literature and Non-Fiction. ECIR 2025.
- Sourav Saha, **Suchana Datta**, Dwaipayan Roy, Mandar Mitra, Derek Greene: Combining Query Performance Predictors:A Reproducibility Study. ECIR 2025.

2024

- **Suchana Datta**, Dwaipayan Roy, Derek Greene, Gerardine Meaney: Unveiling Temporal Trends in 19th Century Literature: An Information Retrieval Approach. JCDL 2024.
- Priyangshu Datta, **Suchana Datta**, Dwaipayan Roy: RAGing Against the Literature: LLM-Powered Dataset Mention Extraction. JCDL 2024.
- **Suchana Datta**, Debasis Ganguly, Sean MacAvaney, Derek Greene: A Deep Learning Approach for Selective Relevance Feedback. ECIR 2024: 189-204

2023

- **Suchana Datta**, Debasis Ganguly, Mandar Mitra, Derek Greene : A Relative Information Gain-based Query Performance Prediction Framework with Generated Query Variants. ACM Trans. Inf. Syst. 41(2): 38:1-38:31 (2023).
- Ashutosh Singh, Debasis Ganguly, **Suchana Datta**, Craig MacDonald : Unsupervised Query Performance Prediction for Neural Models with Pairwise Rank Preferences. SIGIR 2023: 2486-2490.
- **Suchana Datta**, Debasis Ganguly, Derek Greene, Mandar Mitra : On the Feasibility and Robustness of Pointwise Evaluation of Query Performance Prediction. QPP++@ECIR 2023: 1-6.
- **Suchana Datta**, Debasis Ganguly, Josiane Mothe, Md Zia Ullah : Combining Word Embedding Interactions and LETOR Feature Evidences for Supervised QPP. QPP++@ECIR 2023: 7-12.

2022

- **Suchana Datta**, Sean MacAvaney, Debasis Ganguly, Derek Greene : A 'Pointwise-Query, Listwise-Document' based Query Performance Prediction Approach. SIGIR 2022: 2148-2153.
- **Suchana Datta**, Debasis Ganguly, Derek Greene, Mandar Mitra : Deep-QPP: A Pairwise Interaction-based Deep Learning Model for Supervised Query Performance Prediction. WSDM 2022: 201-209.
- Debasis Ganguly, **Suchana Datta**, Mandar Mitra, Derek Greene : An Analysis of Variations in the Effectiveness of Query Performance Prediction. ECIR (1) 2022: 215-229.

2021

- **Suchana Datta**, Debasis Ganguly, Dwaipayan Roy, Derek Greene : Overview of the Causality-driven Adhoc Information Retrieval (CAIR) task at FIRE-2021. FIRE 2021: 25-27.

2020

- **Suchana Datta**, Debasis Ganguly, Dwaipayan Roy, Francesca Bonin, Charles Jochim, Mandar Mitra : Retrieving Potential Causes from a Query Event. SIGIR 2020: 1689-1692.
- **Suchana Datta**, Derek Greene, Debasis Ganguly, Dwaipayan Roy, Mandar Mitra : Where's the Why? In Search of Chains of Causes for Query Events. AICS 2020: 109-120.
- **Suchana Datta**, Debasis Ganguly, Dwaipayan Roy, Derek Greene, Charles Jochim, Francesca Bonin : Overview of the Causality-driven Adhoc Information Retrieval (CAIR) task at FIRE-2020. FIRE 2020: 14-17.

2016

- **Suchana Datta**, Palash Santra, Koushik Majumder, Debashis De : An Automated Malicious Host Recognition Model in Cloud Forensics. International Conference on Recent Advancement in Computer Communication and Computational Sciences (ICRACCCS-2016).
- **Suchana Datta**, Koushik Majumder, Debashis De : DCF: A Novel Dynamic Forensic Framework towards Cloud Computing Environment. IEEE International Conference on Computing, Communication and Automation (ICCCA -2016).
- **Suchana Datta**, Koushik Majumder, Debashis De : Review on Cloud Forensics: An Open Discussion on Challenges and Capabilities. International Journal of Computer Application; Vol.–145, No.–1; July 2016.

Key Projects

University College Dublin

- ***VICTEUR.*** A ERC-funded cultural analytics-based project that aims to analyze large-scale historical and contemporary data sources to address a key unanswered societal question, how does migration impact on the cultural identity of both migrant and host communities in the historical long-term (`https://projectvicteur.com/`).
- ***Causal Information Retrieval.*** Explores causality-driven information retrieval, where I pursued novel search interfaces and algorithms towards addressing causality-driven search, i.e. where a search user wants to find out documents answering the 'Why' question, e.g. why was Osama bin Laden assassinated? etc.

Indian Statistical Institute

- ***Timeline Search.*** The project was focused on timeline search for historical events, aiming to develop unsupervised learning algorithms for deriving insight from text data and building an intelligent query suggestion system.

Cognizant Technology Solutions

- ***Insurance database management.*** The main objective of the project was to follow a unified approach for all applications to interact with database level and implementation of database access layer. Changing and addressing the flaws of SQL injection in existing coding architecture was another challenge.
- ***Developing Market Strategy for Gas & Power in UK.*** The client was one of the largest dealer in Power and utility Domain. Company was headquartered in United Kingdom, strategically aligned into two major business segments – Power and Gas. The project was for developing new applications to improve the marketing strategies and greatly improve the way of work.
- ***Electronics Equipment Manufacturing.*** The project includes maintenance of CRM application instances for four different organizations named IG (Industrial Group), AUTO (Automation, FCAL (Fluke Calibration) and FBC (Fluke Bio Medical).
- ***Leading Insurance Companies of Middle East.*** It is an insurance project involving MSCRM customizations with an on-premise version of CRM. The project is for developing an MSCRM 2013 application which can support both new business as well as renewal applications for availing the insurance policy from the broker/agent or through the customer directly.

Achievements

2024 **Best Paper Award,** Recipient of the Vannevar Bush Best Paper Award at JCDL'24.
2024 **Grant,** Recipient of a travel grant from European Research Council to present a paper at JCDL'24.
2024 **Grant,** Recipient of a travel grant from European Research Council to present a paper at ECIR'24.
2023 **Grant,** Recipient of ACM SIGIR student grant to present a paper at SIGIR'23.
2023 **Grant,** Recipient of SFI student grant to present a paper at ECIR'23.
2022 **Grant,** Recipient of ACM SIGIR student grant to present a paper at WSDM'22.
2022 **Grant,** Recipient of ACM SIGIR student grant to present a paper at SIGIR'22.
2023 **Grant,** Recipient of SFI student grant to present a paper at ECIR'22.
2020 **Grant,** Recipient of ACM SIGIR student grant to present a paper at SIGIR'20.
2020 **Doctoral Scholarship,** Science Foundation Ireland (SFI) funded PhD scholar.
2020 **Grant,** Recipient of TEQIO-II student grant to present a paper at ICCCA'20.
2014 **Post-graduation Scholarship,** Recipient of a 2year scholarship from Ministry of Human Resource Development (MHRD), India.
2014 **Post-graduation Scholarship,** Recipient of a 2year scholarship from Technical Education Quality Improvement Program - phase II (TEQIP-II)
2021 - **Program Committee,** EMNLP'21, CIKM'22, SIGIR'23, CIKM'23, FIRE'16-'20, ECIR'24, SIGIR'25, ECIR'25.

Skills

Programming C, C++, JAVA, J2EE, Python, R, Java Script, PHP, Lucene, .NET framework, ML platform.
Operating Systems Linux, MacOS, Windows.

Languages

Bengali Read, Write, Speak, Peer review
English Read, Write, Speak, Peer review
Hindi Read, Speak

References available upon request.