

# Shoaib Jameel

## CONTACT INFORMATION

Room: 5A.529,  
School of Computer Science and Electronic Engineering,  
University of Essex,  
Colchester,  
Essex  
Postcode: CO4 3SQ,  
United Kingdom

*Mobile:* +44 7440532550  
*E-mail:* shoaib.jameel@essex.ac.uk  
*Fax:* +44 (0)1634 888900  
*Website:* <https://bashthebuilder.github.io/>  
*Google Scholar:* <https://scholar.google.co.uk/citations?user=gyy0YOIAAAAJ&hl=en>

## RESEARCH INTERESTS

Probabilistic Graphical Models, Approximate Posterior Inference, Bayesian Nonparametric Statistics, Probabilistic Topic Modeling, Web Search and Information Retrieval, Natural Language Processing, Social Networks and Urban Human Computing, Learning-to-Rank, Word Embeddings, Deep Neural Networks

## WORK EXPERIENCE

**University of Essex, United Kingdom** *(January 2020 - Present)*  
School of Computer Science and Electronic Engineering  
Lecturer in Computer Science and Artificial Intelligence (Teaching and Research)

**University of Kent, United Kingdom** *(December 2017 - December 2019)*  
School of Computing, Medway  
Lecturer in Computing (Teaching and Research)  
Served as a Deputy Admissions Officer (2018–2019) (Medway)

**Cardiff University, United Kingdom** *(July 2015 - November 2017)*  
School of Computer Science and Informatics  
Research Associate (postdoctoral researcher) with Professor Steven Schockaert

**The Chinese University of Hong Kong (CUHK)** *(November 2014 - July 2015)*  
Department of Systems Engineering and Engineering Management (SEEM)  
Postdoctoral Fellow with Prof. Wai Lam at CUHK and research collaboration with Dr. Xing Xie at Microsoft Research Asia (MSRA)  
Postdoctoral work was partly funded by Microsoft Research Asia (MSRA)

**The Chinese University of Hong Kong (CUHK)** *(August 2014 - October 2014)*  
Department of Systems Engineering and Engineering Management (SEEM)  
Research Assistant with Prof. Wai Lam and Prof. MENG, Mei Ling (Chairman and Professor, SEEM)

## EDUCATION

**The Chinese University of Hong Kong (CUHK)** *(August 2009 - July 2014)*  
Department of Systems Engineering and Engineering Management (SEEM)  
PhD. in Information Systems  
**Thesis Title:** Latent Probabilistic Topic Discovery for Text Documents Incorporating Segment Structure and Word Order

- Advisor: Prof. Wai Lam
- Areas of research: Web search and Information Retrieval, Topic Modeling
- Committee: Prof. MENG, Mei Ling (Former Chairman and Professor, SEEM), Professor CHENG, Chun Hung (Former Associate Professor, SEEM), Dr. Michael C.L. Chau (Associate Professor, School of Business, Faculty of Business and Economics, The University of Hong Kong)

**Sikkim Manipal Institute of Technology (SMIT)** (*September 2004-July 2008*)  
Bachelor of Technology (B. Tech.) in Computer Science and Engineering (CSE)  
Country: India

- Thesis Focus: Load Balancing and Scheduling in Cluster Computing
- Advisors: Prof. Mrinal Kanti Ghose (Head, CSE, SMIT), Fredi B. Zarolia (Information Technology Services, Tata Steel Limited, India)
- Co-Advisor: Dr. Marimuthu Muruganant (*Cantab.*), Research and Development Division, Tata Steel (currently Director at Steel Academics, Research and Consultancy Center (SARCC), Sydney, Australia)
- Other research topics studied: Natural Language Processing and Web search, Artificial Immune System
- Ranked among the top 10% in the entire batch
- Selected based on “merit” through the “All India Engineering Entrance Examination (AIEEE)” ranking
- My undergraduate thesis was awarded an **S** (*Superb!*) grade

**Kerala Samajam Model School (KSMS)** (*March 1989 - March 2003*)  
Country: India

- Indian Certificate of Secondary Education (ICSE - 2001)
  - Percentage - 88.88%
  - Ranked Eighth in the entire school (*out of approximately 250 students*)
- Indian School Certificate (ISC - 2003)
  - Percentage - 82.00%
  - Ranked Fourth in the Science stream (*out of approximately 120 students*).  
(*Missed out on the third spot by a score of 0.5*)

**CONFERENCE As Lecturer in Computer Science with different university affiliations**

**PUBLICATIONS**

1. Hamad Zogan, Imran Razzak, **Shoaib Jameel** and Guandong Xu. 2021. *DepressionNet: Learning Multi-modalities with User Post Summarization for Depression Detection on Social Media*. The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2021). Accepted. Full paper.
2. Kun Yan, Zied Bouraoui, Ping Wang, **Shoaib Jameel**, Steven Schockaert. 2021. *Aligning Visual Prototypes with BERT Embeddings for Few-Shot Learning*. ACM International Conference on Multimedia Retrieval 2021 (ICMR 2021). Accepted. Full paper.
3. António Correia, Daniel Schneider, **Shoaib Jameel**, Hugo Paredes, and Benjamim Fonseca. 2021. *Experimental Investigation of the Factors Influencing Researchers’ Adoption of Crowdsourcing and Machine Learning: An Initial Survey*. 20th International Conference on Intelligent Systems Design and Applications (ISDA 2020). Online conference. Full paper.
4. António Correia, Diogo Guimaraes, Dennis Paulino, **Shoaib Jameel**, Daniel Schneider, Benjamim Fonseca and Hugo Paredes. 2021. *AuthCrowd: Author Name Disambiguation and Entity Matching using Crowdsourcing*. 24th IEEE International Conference on Computer Supported Cooperative Work in Design (IEEE CSCWD 2021). Online conference. Full paper.
5. Kun Yan, Zied Bouraoui, Ping Wang, **Shoaib Jameel**, Steven Schockaert. 2021. *Few-Shot Image Classification With Multi-Facet Prototypes*. International Conference on Acoustics, Speech, & Signal Processing (ICASSP). Online conference.

Full paper.

6. António Correia, **Shoaib Jameel**, Daniel Schneider, Hugo Paredes, and Benjamim Fonseca. 2020. *A Workflow-Based Methodological Framework for Hybrid Human-AI Enabled Scientometrics*. IEEE BigData. 2021. Online conference. Full paper.
7. Zihao Fu, Lidong Bing, Wai Lam, **Shoaib Jameel**. *Dynamic Topic Tracker for KB-to-Text Generation*. 2020. International Conference on Computational Linguistics (COLING). Online conference. Full paper.
8. António Correia, Benjamim Fonseca, Hugo Paredes, Daniel Schneider, **Shoaib Jameel**. *On the Socio-Technical Aspects of Crowdsourcing as a Mechanism for Academic Study: Conceptual Framework and Scoping Review*. 2020. The Hawaii International Conference on System Sciences (HICSS), Hawaii, USA. (**Nominated for the “Best Paper Award”**)
9. António Correia, Dennis Paulino, Hugo Paredes, Benjamim Fonseca, **Shoaib Jameel**, Daniel Schneider and Jano de Souza. *Scientometric Research Assessment of IEEE CSCWD Conference Proceedings: An Exploratory Analysis from 2001 to 2019*. Accepted. Full paper. Dalian, China. *Conference postponed due to SARS-COV-2 outbreak*.
10. António Correia, Hugo Paredes, Benjamim Fonseca, **Shoaib Jameel**. *SciCrowd: A crowd-AI system for supporting crowd work in scientific research*. 2019. In: Science and Technology Summit in Portugal. Poster.
11. António Correia, Benjamim Fonseca, Hugo Paredes, Daniel Schneider, **Shoaib Jameel**. *Development of a Crowd-Powered System Architecture for Knowledge Discovery in Scientific Domains*. 2019. 2019 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1372–1377, Bari, Italy.
12. António Correia, Hugo Paredes, Daniel Schneider, **Shoaib Jameel**, Benjamim Fonseca. *Towards Hybrid Crowd-AI Centered Systems: Developing an Integrated Framework from an Empirical Perspective*. 2019. 2019 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 4013–4018, Bari, Italy.
13. **Shoaib Jameel**, and Steven Schockaert. 2019. *Word and Document Embedding with vMF-Mixture Priors on Context Word Vectors*. In the 57th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 3319–3328, Florence, Italy.
14. Jose Camacho-Collados, Luis Espinosa-Anke, **Shoaib Jameel**, Steven Schockaert. *A Latent Variable Model for Learning Distributional Relation Vectors*. 2019. In the 28th International Joint Conference on Artificial Intelligence (IJCAI), 4911–4917, Macao, China,
15. António Correia, **Shoaib Jameel**, Daniel Schneider, Benjamim Fonseca and Hugo Paredes. 2019. *The Effect of Scientific Collaboration on CSCW Research: A Scientometric Study*. In the 23rd IEEE International Conference on Computer Supported Cooperative Work in Design (IEEE CSCWD), pp. 149–161, Porto, Portugal.
16. **Shoaib Jameel**, Zihao Fu, Bei Shi, Wai Lam, and Steven Schockaert. 2019. *Word Embedding as Maximum A Posteriori Estimation*. In the 33rd AAAI Conference on Artificial Intelligence (AAAI), pp. 6562–6569, Honolulu, Hawaii, USA. [PDF]
17. Zied Bouraoui, **Shoaib Jameel**, and Steven Schockaert. 2018. *Relation Induction in Word Embeddings Revisited*. In the 27th International Conference on Computational Linguistics (COLING), pp. 1627–1637. Santa Fe, New Mexico.

[PDF] (One of the “Area Chair Favourites”)

18. **Shoaib Jameel**, Zied Bouraoui and Steven Schockaert. 2018. *Unsupervised Learning of Distributional Relation Vectors*. In the 56th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 23–33. Melbourne, Australia. [PDF]

#### During doctoral and postdoctoral work (August 2009 - 2017)

1. **Shoaib Jameel**, and Steven Schockaert. 2017. *Modeling Context Words as Regions: An Ordinal Regression Approach to Word Embedding*. In the SIGNLL Conference on Computational Natural Language Learning (CoNLL), pp. 123–133, Vancouver, Canada. [PDF]
2. Bei Shi, Wai Lam, **Shoaib Jameel**, Steven Schockaert and Kwun Ping Lai. 2017. *Jointly Learning Word Embeddings and Latent Topics*. In the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), pp. 375–384, Tokyo, Japan. [PDF]
3. **Shoaib Jameel**, Zied Bouraoui, and Steven Schockaert. 2017. *MEMBER: Max-Margin Based Embeddings for Entity Retrieval*. In the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), pp. 783–792, Tokyo, Japan. [PDF]
4. Zied Bouraoui, **Shoaib Jameel**, and Steven Schockaert. 2017. *Inductive Reasoning about Ontologies Using Conceptual Spaces*. In Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI), pp. 4364–4370, San Francisco, California, USA. [PDF]
5. **Shoaib Jameel**, Steven Schockaert. 2016. *D-GloVe: A Feasible Least Squares Model for Estimating Word Embedding Densities*. In Proceedings of the 26th International Conference on Computational Linguistics (COLING), pp. 1849–1860, Osaka, Japan. [PDF]
6. **Shoaib Jameel**, Yi Liao, Wai Lam, Steven Schockaert, Xing Xie. 2016. *Exploring Urban Lifestyles Using a Nonparametric Temporal Graphical Model*. In Proceedings of the ACM SIGIR International Conference on the Theory of Information Retrieval (ICTIR), pp. 251–260, Newark, Delaware, USA. [PDF]
7. Yi Liao, Wai Lam, **Shoaib Jameel**, Steven Schockaert, Xing Xie. 2016. *Who Wants to Join Me? Companion Recommendation in Location Based Social Networks*. In Proceedings of the ACM SIGIR International Conference on the Theory of Information Retrieval (ICTIR), pp. 271–280, Newark, Delaware, USA. [PDF]
8. **Shoaib Jameel**, and Steven Schockaert. 2016. *Entity Embeddings with Conceptual Subspaces as a Basis for Plausible Reasoning*. In Proceedings of the European Conference on Artificial Intelligence (ECAI), pp. 1353–1361, The Hague, Netherlands. [PDF]
9. Steven Schockaert, and **Shoaib Jameel**. 2016. *Plausible Reasoning based on Qualitative Entity Embeddings*. In Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI). (Invited Paper), pp. 4078–4081, New York, USA. [PDF]
10. Pengfei Liu, **Shoaib Jameel**, King Keung Wu, and Helen Meng. 2016. *Learning Track Representation and Trends for Conference Analytics*. In Proceedings of the Hawaii International Conference on System Sciences (HICSS-49), pp. 1671–1680, Hawaii, USA. [PDF]
11. Yi Liao, **Shoaib Jameel**, Wai Lam, and Xing Xie. 2015. *Abstract Venue Concept Detection from Location-Based Social Networks*. In Proceedings of the Asia

- Information Retrieval Societies Conference (AIRS), pp. 147–157, Brisbane, Australia. [PDF]
12. **Shoaib Jameel**, Wai Lam, Steven Schockaert, and Lidong Bing. 2015. *A Unified Posterior Regularized Topic Model with Maximum Margin for Learning-to-Rank*. In Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM), pp. 103–112, Melbourne, Australia. [PDF]
  13. Pengfei Liu, **Shoaib Jameel**, Wai Lam, Bin Ma, and Helen Meng. 2015. *Topic Modeling for Conference Analytics*. In Proceedings of the 16th Annual Conference of the International Speech Communication Association (INTERSPEECH). [PDF]
  14. **Shoaib Jameel**, Wai Lam, and Lidong Bing. 2015. *Nonparametric Topic Modeling using Chinese Restaurant Franchise with Buddy Customers*. In Proceedings of the 37th European Conference on Information Retrieval (ECIR), pp. 648–659, Vienna, Austria. [PDF]
  15. Lidong Bing, Wai Lam, **Shoaib Jameel**, and Chunliang Lu. 2014. *Website Community Mining from Query Logs with Two-phase Clustering*. In Proceedings of the 15th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing), pp. 201–212, Kathmandu, Nepal. [PDF]
  16. **Shoaib Jameel**, and Wai Lam. 2013. *A Nonparametric N-Gram Topic Model with Interpretable Latent Topics*. In Proceedings of the Ninth Asia Information Retrieval Societies Conference (AIRS), pp. 74–85, Singapore. [PDF]
  17. **Shoaib Jameel**, and Wai Lam. 2013. *An Unsupervised Topic Segmentation Model Incorporating Word Order*. In Proceedings of the 36th Special Interest Group on Information Retrieval (ACM-SIGIR), pp. 203–212, Dublin, Ireland. [PDF]
  18. **Shoaib Jameel**, and Wai Lam. 2013. *An N-gram Topic Model for Time-Stamped Documents*. In Proceedings of the 35th European Conference on Information Retrieval (ECIR), pp. 292–304, Moscow, Russia. [PDF]
  19. **Shoaib Jameel**, Xiaojun Qian, and Wai Lam. 2012. *N-gram Fragment Sequence Based Unsupervised Domain-Specific Document Readability*. In Proceedings of the 24th International Conference on Computational Linguistics (COLING), pp. 1309–1326, Mumbai, India. [PDF]
  20. **Shoaib Jameel**, Wai Lam, Xiaojun Qian, and Ching-man Au Yeung. 2012. *An unsupervised technical difficulty ranking model based on conceptual terrain in the latent space*. In Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL), pp. 351–352, Washington D.C., USA. [PDF]
  21. **Shoaib Jameel**, and Xiaojun Qian. 2012. *An Unsupervised Technical Readability Ranking Model by Building a Conceptual Terrain in LSI*. In Proceedings of the International Conference on Semantics, Knowledge and Grids (SKG), pp. 39–46, Beijing, China. [PDF]
  22. **Shoaib Jameel**, Wai Lam, and Xiaojun Qian. 2012. *Ranking Text Documents based on Conceptual Difficulty using Term Embedding and Sequential Discourse Cohesion*. In Proceedings of the 2012 IEEE/WIC/ACM International Conference on Web Intelligence (WI), pp. 145–152, Macau, China. [PDF]
  23. **Shoaib Jameel**, Wai Lam, Ching-man Au Yeung, and Sheaujiun Chyan. 2011. *An Unsupervised Ranking Method based on a Technical Difficulty Terrain*. In Proceedings of the 20th ACM international conference on Information and knowledge management (CIKM), pp. 1989–1992, Glasgow, Scotland. (*Winner: Best Paper Award at the Beijing-Hong Kong International Doctoral Forum-2011*)

#### During undergraduate study (2004-2008)

1. **M. Shoaib Jameel**, Anubhav, Nilesh Singh, Nitin Kumar Singh, Tejbanta Singh Chingtham, and M. K. Ghose. 2009. *An Intelligent Automatic Text Summarizer*. In Proceedings of the First International Conference on Intelligent Human Computer Interaction, pp. 223–230, Hyderabad, India.
2. **M. Shoaib Jameel**, Amar Akshat, and Tejbanta Singh Chingtham. 2008. *Enhancements in Query Evaluation and Page Summarization of The Thinking Algorithm*. In Proceedings of the International Symposium on Information Technology (ITSim), pp. 1–8, Kuala Lumpur, Malaysia.

#### JOURNAL PUBLICATIONS

##### As Lecturer

1. Jianlong Zhou, Hamad Zogan, Shuiqiao Yang, **Shoaib Jameel**, Guandong Xu, Fang Chen. 2021. *Detecting Community Depression Dynamics Due to COVID-19 Pandemic in Australia*. Accepted. IEEE Transactions on Computational Social Systems.
2. António Correia, Daniel Schneider, **Shoaib Jameel**, Hugo Paredes, Benjamim Fonseca. 2019. *‘Caught Between Worlds’: Leveraging Crowd Science and AI for Supporting Scientific Work Practices*. Accepted. Computer Supported Cooperative Work: The Journal of Collaborative Computing and Work Practices

#### During postdoctoral work (2014-Present)

1. **Shoaib Jameel**, Wai Lam, and Lidong Bing. 2015. *Supervised Topic Models with Word Order Structure for Document Classification and Retrieval Learning*. In Information Retrieval Journal (IRJ). 18 (4): 283-330.
2. Lidong Bing, Shan Jiang, Wai Lam, Yan Zhang, and **Shoaib Jameel**. 2015. *Adaptive Concept Resolution for Document Representation and Its Applications in Text Mining*. In Knowledge-Based Systems 74 (2015): 1-13.
3. Lidong Bing, Wai Lam, Tak-Lam Wong, and **Shoaib Jameel**. 2015. *Web Query Reformulation via Joint Modeling of Latent Topic Dependency and Term Context*. In ACM Transactions on Information Systems (TOIS). Volume 33, Issue 2.

#### During undergraduate study (2004-2008)

1. **M. Shoaib Jameel**, and Tejbanta Singh Chingtham. 2009. *Compounded Uniqueness Level: Geo-Location Indexing Using Address Parser*. In International Journal of Computer Theory and Engineering. pp. 27–34
2. **M. Shoaib Jameel**, Marimuthu Muruganant, and Tejbanta Singh Chingtham. 2009. *Deploying CPU Load Balancing in the Linux Cluster Using Non-Repetitive CPU Selection*. In International Journal of Computer Theory and Electrical Engineering, pp. 228–234.

#### BOOK CHAPTERS

##### As Lecturer in Computing at the University of Kent

1. António Correia, **Shoaib Jameel**, Hugo Paredes, Benjamim Fonseca, Daniel Schneider 2019. *Hybrid Machine-Crowd Interaction for Handling Complexity in Science: A Scaffolding Design Framework*. Accepted. Springer Human-Computer Interaction Series

#### UNDERGRAD RESEARCH GRANT

**Undergraduate Seed Research Grant:** My undergraduate university awarded me an INR (Indian Rupees) 10,000 seed grant (as PI) to support my future research. The grant could be used to buy research equipment or for conference travel. The award was prestigious and very competitive in the sense that I was the only student in my batch of 120 computer science students to be awarded this seed funding based on my

undergraduate research accomplishments in the final year of my undergraduate study. Quality of the research proposal followed by a consistent academic performance were also some of the criteria used to select the candidate.

## PhD GRANTS

- Scholarship name: BT and CSEE, University of Essex Doctoral Scholarship  
Duration: Three years  
Role: Main Supervisor

## INNOVATE UK RESEARCH GRANTS

- Proposal title: The utilisation of a novel combination of natural language processing (NLP) and image analysis to accelerate and automate the collection of Open Source Intelligence (OSINT), providing end users with advanced insight into individuals without invading individual privacy.  
Funding name: Knowledge Transfer Partnership (KTP)  
Duration: Three years  
Role: Lead Academic or Principal Investigator  
Amount: £350k
- Proposal title: Develop AI methods to optimise interactions with customers.  
Funding name: Knowledge Transfer Partnership (KTP)  
Duration: Two years  
Role: Support Academic (40%: highest percentage among all members)  
Amount: £123k
- Proposal title: The development of a new CPD tracker using AI and embedded machine learning to track and enhance performance of all staff. To gather data from the CPD tracker that will provide insight into effective strategies offering senior leaders a clear path for development driving industry-wide improvement.  
Funding name: Knowledge Transfer Partnership (KTP)  
Duration: Two years  
Role: Support Academic  
Amount: £230k  
Note: This project was unfortunately withdrawn by the company midway as a result of financial losses incurred by it due to the COVID19 pandemic.
- Proposal title: The design and implementation of an integrated smart platform and bespoke client-side application for process integration (site plan material sharing, management and version control), enhancing effective office-site communications in the property development industry.  
Funding name: Knowledge Transfer Partnership (KTP)  
Duration: Two years  
Role: Support Academic or Co-Investigator (with 40% role, same as the Principal Investigator)  
Amount: £250k

## INNOVATION GRANTS

- Proposal title: **NG-CrAI: A Next-Gen Crowd-AI Framework for Scientific Discovery**  
Funding name: **INESC TEC's Internal Seed Projects Grant**  
Amount approved: £35,938  
Duration: 2 years  
Role: UK collaborator  
Year: 2020–2021 Current international cooperation under the topics of the project with University of Essex (UK) and the Federal University of Rio de Janeiro (Brazil) will be enhanced by the participation of Prof. Shoaib Jameel and Prof. Jano Moreira de Sousa as external consultants. A short stay in Portugal (1 week) per year is required for each consultant.

- Proposal title: **Artificially Intelligent RoleCatcher, FinTex**  
Funding name: **Essex Innovation Voucher**  
Amount approved: £10,000  
Duration: Six months  
Role: Principal Investigator  
Year: 2020–2021

## INDUSTRY FUNDING

- Proposal title: **Mining Time-Aware Urban Living Styles via Latent Semantic Concept Analysis**  
Funding name: **Microsoft Research Asia Urban Informatics Research Fund, 2014**  
Amount approved: USD 20,000  
Duration: 1 year  
The content of this proposal was exclusively designed and written by me. I introduced a new dimension to the existing MSRA project called LifeSpec, which is available internally at MSRA. In particular, I added a temporal component when generating the urban human lifestyle patterns. To this end, I proposed an entirely new temporal model that improved upon the original urban human computing model proposed at MSRA. The proposed model improved upon the time and space complexities, quality of the urban patterns generated, handling the sparsity in the data and improving upon the quantitative tasks. I designed a non-parametric temporal graphical model and formulated the equations by myself. The two main challenges were technical formulations and the evaluation plans. I wrote the draft of the proposal with full technical details and the intuitions behind the model and how it could be useful to MSRA. Besides, I designed and prepared the slides to present the proposal to an audience from Microsoft Research, who came to judge all competing proposals. The committee applauded the idea and the model very much at the end of the presentation. The proposal cited my works which were published in the leading conferences to convince the committee that we were competent enough to accomplish the required research task within a year.

## RESEARCH COUNCIL GRANTS

- Proposal title: **Social-Aware News Trend Discovery via Joint Detection of Latent Information Structure from News and Social Media Text Content**  
Funding name: **HKRGC General Research Fund, 2014**  
Amount approved: HKD 692,894  
Completion date: 31-12-2017  
The RGC funding is among the few large research aids given by the Hong Kong government to some of the best research proposals. In this proposal, I handled and crafted the complete draft of the model's technicalities. The central theme of the proposal was to learn hierarchical human living and mobility patterns using data obtained from the online social networks and apply the trained model to conduct event modeling and group recommendation. To this end, I designed the graphical model, the generative process and the mathematical formulations. I wrote the draft which included the technical details, related work, the evaluation plan and the intuitions behind the model to convince the reviewers that the task and the proposed models are noteworthy and important enough to be funded.
- Proposal title: **Multi-Modal Image Style Transfer: Automatically Geo-Localising Reading Material Digital Artwork for Increased Reader Engagement**  
Funding name: Global Challenges Research Fund (GCRF)  
Amount approved: £30,0000



Start date: January 2021  
End date: July 2021

## PhD THESIS EXAMINATIONS

- External examiner - committee also included Professor Udo Kruschwitz, University of Essex (*currently Chair in Information Science at the University of Regensburg*)
  1. Student: Andrius Mudinas, Birkbeck, University of London
  2. Year: 2019
  3. Supervisors: Dr. Dell Zhang and Professor Mark Levene
- External Thesis Examiner - Sikkim Manipal University, India
  1. Student: Papri Ghosh
  2. Year: 2019
  3. Supervisors: Dr. Tejbanta Chingtham and Professor MK Ghose

## INVITED TALKS

- Keynote speaker at the 2021 BT AI Global Festival (Online event)
- Lectures in the Summer School at Peking University, Beijing (July – 2019). I was recently invited by the Department of Intelligent Computing and Sensing Laboratory at Peking University to deliver lectures in their summer school in Beijing, China. Nominations were first sent by the School of Computing, University of Kent to Peking University which included my name as the only junior academic staff followed by names of the senior academic members, i.e., senior lecturers to professors. Academics from other institutions in the UK and the EU had also sent their nominations. In the end, I was the **only academic** selected to deliver the lectures from the UK/EU region
- Digital Humanities Workshop, CHASE Group (October – 2018)
- BenevolentAI (October – 2018)
- Centre for Language and Linguistics, University of Kent (March – 2018). I delivered the talk along with Professor Alessandro Vinciarelli from the University of Glasgow
- Department of Computer Science and Information Systems, Birkbeck, University of London (March – 2018)
- School of Mathematics, Cardiff University (November – 2017)
- Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, SEEM Seminar Series (December – 2017)
- Workshop talk: “Probabilistic Topic Models” at the School of Computer Science and Informatics, Cardiff University (June – 2017)

## REVIEWER SERVICE

### Conferences

- Programme Committee Member:
  - ACS/IEEE International Conference on Computer Systems and Applications (AICCSA-2014)
  - Hawaii International Conference on System Sciences (HICSS-2015)
  - Russian Summer School in Information Retrieval (RussIR-2016)
  - World Wide Web Conference (WWW-2017, WWW-2018)
  - Knowledge Discovery and Data Mining (SIGKDD-2021)

- The North American Chapter of the Association for Computational Linguistics (NAACL-2021)
- Special Interest Group on Information Retrieval (SIGIR-2017, SIGIR-2018, SIGIR-2019, SIGIR-2020, SIGIR-2021)
- International Joint Conference on Artificial Intelligence (IJCAI-2017, IJCAI-2018, IJCAI-2019, IJCAI-2020)
- ACM International Conference on Information and Knowledge Management (CIKM-2017, CIKM-2018, CIKM-2020)
- Annual UK Workshop on Computational Intelligence (UKCI-2017)
- International Conference on Computational Linguistics (COLING-2018, COLING-2020)
- AAAI Conference on Artificial Intelligence (AAAI-2019, AAAI-2020, AAAI-2021)
- European Conference on Information Retrieval (ECIR-2018, ECIR-2019, ECIR-2020)
- Conference on Empirical Methods in Natural Language Processing (EMNLP-IJCNLP 2019, EMNLP-2020)
- Asia-Pacific chapter of the Association for Computational Linguistics (AACL-IJCNLP 2020)
- Web Search and Data Mining (WSDM 2021)
- European Chapter of the Association for Computational Linguistics (EACL-2021)
- Annual Meeting of the Association for Computational Linguistics (ACL-2019, ACL-2020)
- External (Sub) Reviewer
  - Conference on Artificial Intelligence (AAAI-2017 and AAAI-2016)
  - International Joint Conference on Artificial Intelligence (IJCAI-2016 and IJCAI-2015)
  - Conference on Information and Knowledge Management (CIKM-2016)
  - World Wide Web Conference (WWW-2015 and WWW-2014)
  - Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-2015)
  - Special Interest Group on Information Retrieval (SIGIR-2015, SIGIR-2014, SIGIR-2013 and SIGIR-2012)
  - Association for Computational Linguistics (ACL-2014)
  - Knowledge Discovery and Data Mining (KDD-2014 and KDD-2013)
  - IEEE International Conference on Data Mining series (ICDM-2013)

## Journals

- WIREs Data Mining and Knowledge Discovery - *Awarded a certificate highlighting my reviewing contributions*
- Entropy
- Fuzzy Sets and Systems
- Transactions on Information Systems (TOIS)

- ISPRS International Journal of Geo-Information
- Information Retrieval Journal
- The Computer Journal
- Computer Speech & Language - *Outstanding Reviewer recognition*
- Journal of Experimental & Theoretical Artificial Intelligence
- Information Processing and Management (IP and M)
- Knowledge-Based Systems (KBS) - Elsevier
- Neural Networks - Elsevier
- Transactions on Asian Language Information Processing (TALIP)
- Transactions on Knowledge and Data Engineering (TKDE)
- Information (*Member of the Reviewer Board*)
- Computational Linguistics (*Member of the Reviewer Board*)
- IEEE Transaction on Computational Social Systems

## JOURNAL EDITORIAL

**Guest Editor:** IEEE Transaction on Computational Social Systems

Topic: Special Issue on Computational Social Systems for COVID-19 Emergency Management and Beyonds

Year: 2020

## AWARDS AND ACHIEVEMENTS

### The Chinese University of Hong Kong

- Finalist in the ACM-Hong Kong Best Research Awards-2013 held at Hong Kong University of Science and Technology, Hong Kong. *Only two best research candidates from the entire Faculty of Engineering of CUHK were selected to represent the university in the final.* The event was organized by ACM (Hong Kong Chapter) and Microsoft Research Asia
- SIGIR-2013 Student Travel Award
- Student Volunteer at SIGIR-2013 (received registration fee waiver)
- ECIR-2013 Student Accommodation Grant offered by Yandex, Moscow
- Best paper award at Beijing-Hong Kong International Doctoral Forum-2011
- Department of Systems Engineering and Engineering Management Postgraduate Scholarship - 2009

### Undergraduate and before

- Letter of Appreciation for excellent performance at Tata Steel R&D (India) - 2009
- Nominated for the prestigious Overseas Research Students Awards Scheme (ORSAS), University of London (*declined*) - 2009
- Featured in the Indian national news for my ThinkSearch algorithm - 2008. I was interviewed one-on-one by a journalist from a leading Indian national newspaper, "Hindustan Times". The interview was published on the front page of my city's newspaper
- Scored a GPA of 10/10 in the final semester during undergraduate study - 2008
- Best paper award at TechFest, Sikkim Manipal Institute of Technology - 2006
- Third prize in SUPW/Science Workshop, Kerala Samajam Model School - 2002
- Second prize in Rotary Club of Jamshedpur Computer Quiz - 2001

- Consolation prize in SUPW/Science Workshop, Kerala Samajam Model School - 2001
- Awarded at the Annual Academic Prize Night for scoring 90% and above in the respective courses in ICSE (Indian Certificate of Secondary Education - Year 10) examination - 2001
- Commendation certificates for an excellent academic record in each terminal examination - 1996-2000
- Ranked among top three in the entire school - 1996-2000
- Awarded at the Annual Academic Prize Night for an excellent academic track record in the entire year consistently from 1996 till 2000
- Third prize in Inter-School Academic Quiz, Kerala Samajam Model School - 2000
- First prize in Science Quiz, Kerala Samajam Model School - 1999
- Featured in the national news for securing the top rank in school - 1996. The newspaper carried a lengthy story about me when I stood first in the entire batch for the first time. What was so inspiring about the story was that I was just a below average student a year back
- Best student in Discipline in the entire year in Taekwondo - 1995
- Rewarded numerous times in Taekwondo championships, belt examination and camps

## TEACHING EXPERIENCE

### University of Essex, Colchester

1. CE807 - Text Analytics (Module Lead)
2. CE306/CE706 - Information Retrieval (Second Module Lead)

### University of Kent, Medway

1. CO334 - People and Computing (Module Convener)
2. CO547 - Agile Software Development and Security (Module Convener)
3. CO639 - Electronic Commerce
4. CO644 - Semantic Web

### Students' feedback about my teaching at the University of Kent:

Immediately sent by a student after my first lecture ever in Kent (Module CO639) to his tutor who is also a lecturer in the School of Computing at Medway: *"Afternoon mrs, I would just like to put in a small word of gratitude & say today's lecture felt very engaging, not saying that yours or XXX (other lecturer's name made confidential) aren't but, considering it was his (didn't get his name) first time lecturing here, he engaged well I believe - he ensured we understood what was briefed in the powerpoints and just made things easier. Perhaps it was today's topic, that made the teaching easier, I don't know but I just wanted you to pass my regards to him, I just felt more engaged than usual today... thank you"*.

This is what the tutor of the student had to say: *"Congratulations Shoaib! This feedback is from a student who never ever contacts me usually (and hes my tutee!) and has not really been engaging with several courses. We need to observe your teaching rather than the other way around, it seems :)"*

Another student from a different module that I taught next semester (Module CO334). It shows that I have maintained a good teaching track record:

*“Hi, Shoaib, I am contacting you regarding my review of your teaching style. I really enjoy attending your lecture and prefer this approach to teaching, however, I feel sometimes it can go off topic. I enjoy sharing my opinion and having the class getting more involved gives the lecture a better atmosphere. Furthermore I think using videos makes it easier to write notes and is more informative than having 80+ slides to go through. I would love to see examples incorporated into this video to showcase what you’re saying rather than having to see you draw. Thanks.”*

Another student from the same module as above:

*“Dear Shoaib, Just wanted to say that as part of your People and Computing lecture on a Friday morning, I really do enjoy your teaching. You seem more charismatic than almost all of the other lecturers and I think the way you teach is engaging too. The only feedback i can give is not about you but is infact about the module being a little basic. I understand that you cannot change the stuff that has to be taught so maybe try and engage the class by creating interactive activities. I’m not too sure what other people in the class would think about it but I know it would help me become more awake as it is a Friday morning after all! Hope that can be of some help,”*

This is what one mature student had to say about my teaching in CO639:

*“Hi Shoaib, I wanted to let you know that I really enjoyed the assignment on cryptocurrencies. I really like your style of teaching. The idea of a discussion first gets you thinking and the slides afterwards cement it in your mind. The only negative I would say is you ask too much if your teaching style is okay. I totally understand why you do this but it can be a little distracting. I think it would be better if you just rolled with it and if there was an issue I’m sure the students would let you know. I like you as a person and a lecturer and wish you all the best in the future.”*

The comment below shows that I have consistently maintained my innovative teaching methods in my second year as Lecturer. A final year student from the module Semantic Web that I am teaching in 2019.

*“Hi Shoaib, Just wanted to say I loved the lecture today! And your methods have helped us a lot. Thank you! ”*

### **The Chinese University of Hong Kong**

#### **Introduction to Engineering - I (ENGG1100)**

- Supervised students in designing and constructing a robotic car for automatic navigation
- Taught students about the basics of 3D printing and using SolidWorks

#### **Information Systems Management (SEEM 3490)**

- Delivered lectures on Information Security and Cryptography
- Conducted revision classes covering the entire syllabus before the final examination (*the most challenging part!*)
- Supervised student projects

#### **Data Structures (CSC2100E/F)**

- Delivered lectures on Computational Complexity, Pseudo-code writing, Searching

and Sorting

- Formulated interesting assignment questions based on real-life problems
- Conducted revision classes covering the entire syllabus before the final examination (*again the most challenging part!*)
- Supervised student programming projects

### **Sikkim Manipal Institute of Technology**

Data Structures - Functions and Pointers in C

- Organized student workshops covering all aspects of “Functions” and “Pointers” in the C programming language

### **GRADUATED RESEARCH SUPERVISION**

- **PhD students:** Liao Yi (CUHK) and Shelan Jeawak (Cardiff University) - I can be best described as a “technical advisor” to them. My responsibility is to help them in mathematical modeling, scientific programming and experiment design with various analyses. With Yi, I have published papers in conferences such as AIRS and ICTIR.
- **Masters (MPhil):** Sam Lai (CUHK) - I have supervised Sam for a year and have taught him about data analysis and crawling for the online social networks. I have also trained him in probabilistic topic modeling programming in Java. He has successfully defended his MPhil thesis recently, and subsequently he will take the role of a research assistant in my PhD thesis supervisor’s group. We have jointly published a paper in a flagship information retrieval conference recently (SIGIR).

### **CURRENT RESEARCH SUPERVISION**

- **PhD student:** António Correia, Ph.D. candidate in Computer Science, University of Trás-os-Montes e Alto Douro, Vila Real, Portugal Researcher at INESC TEC, Porto, Portugal. (*Research Co-Supervisor*). *Awarded: PARSUK Xperience scholarship*
- **PhD student:** Hamad Zogan, PhD candidate in School of Computer Science & Advanced Analytics Institute, Faculty of Engineering & IT, University of Technology Sydney. Australia. (*Research Co-Supervisor*)
- **PhD student:** Mahsa Abazari Kia, School of Computer Science and Electronic Engineering, The University of Essex. (*Research Supervisor*)
- **Research Associate and PhD student:** Mozghan Talebpour, KTP Associate. (*Research Supervisor*)
- **PhD student:** Reyhaneh Hashempour, School of Computer Science and Electronic Engineering, The University of Essex. (*Research Co-Supervisor*)

### **COLLEGIALLY**

- **CO320: Introduction to Object-Oriented Programming** Took over Java tutorial classes for the undergraduate students; covered on behalf of a staff who abruptly resigned leaving no one to cover the required tutorials. (University of Kent in 2019)
- **CO334: People and Computing** Took over the lectures for the staff who was off work due to paternity reasons. (University of Kent in 2018)
- **CO600: Final-year Undergraduate Major Projects** Supervised a group because the main supervisor went on a sabbatical leave. (University of Kent in 2018)

- **Admissions Officer** Took over the main undergraduate responsibility because the admissions officer was offered the Head of School role. (University of Kent in 2018)
- **New Lecturer interviews** Took the responsibility on a very short notice of taking care of the applicants who visited the campus for the new Lecturer interviews. (University of Kent in 2018)
- **Examination script marking** Marked final-year written undergraduate examination answer scripts on behalf of an academic staff who suddenly took leave. (University of Kent in 2018 and 2019)
- **Data Science group logo design** Took the responsibility for designing the group's logo. (University of Kent in 2018)

**PROFESSIONAL EXPERIENCE** **Tata Steel Limited, Jamshedpur, India**

Intern at the Research and Development Division (R&D) *January 2008 - June 2009*

- Designed, developed and deployed the CPU load balancing architecture for the Linux cluster
- Designed, developed and deployed a multi-threaded local Intranet crawler
- Designed and developed the “document-streaming” algorithm
- Estimation of yield-point from hot-strip mill data
- Drafted a detailed technical report on text classifiers

**Tata Steel Limited, Jamshedpur, India**

Intern at the Automation Division

*July 2007 - August 2007*

- Used Data Mining to come up with the relationship between steel casting parameters and the billet caster

**GRADUATE COURSES UNDERTAKEN**

- Techniques for Data Mining (CSC 5180)
- Text Mining Models & Applications (SEG 5680)
- Advanced Database & Information Systems (SEG 5010)
- Knowledge Systems (SEEM 5470 )
- Logistics & Transport Planning (SEEM 5600)

**PROFESSIONAL TRAINING**

- Postgraduate Certificate for Higher Education programme (PGCHE) - **Status: Awarded (*Fellow of the HEA*)**. Two compulsory and two optional modules completed thereby fulfilling the requirements for the PGCHE. *Out of four modules, DISTINCTION in two.*
- Communication Skills Training
- Teaching Assistant Training

**SERVICES AND PARTICIPATION** **University of Kent**

- Mentored three students at ACL-2019
- Data Science group web page and blog maintainer
- Undergraduate admissions work which includes participation in Open Days (usually Saturdays), Applicant Days, Undergraduate interviews, Ad-hoc undergraduate interviews, admission meetings, undergraduate campus tours, UCAS online applications processing, processing undergraduate student applicant day invitations - *This is a major duty and the most important administration work in the*

*school. I continued to maintain a strong teaching and research track records despite taking up this time consuming role*

- Participated in organizing the new faculty recruitment
- Session Chair at AAAI-2019 (Robotics Track)
- Panel member in student plagiarism hearing
- Helping the university with student campus recruitment
- Panel member in Medway disciplinary committee

#### **Conference Organization Experience**

- “Board of members for Conference Programme editorial”, Thermec’2013, Las Vegas, Dec. 2 to 6, 2013
- Organizing Chair, Beijing-Hong Kong International Doctoral Forum (2011)
- Technical Paper Reviewer, Beijing-Hong Kong International Doctoral Forum (2010-2011)

#### **Other Services**

- University Open Day planning, organization and execution (CUHK)
- Student Volunteer at the 2014 Big Data Workshop jointly organized by KAIST-Microsoft Research Collaboration Center (KMCC), Korea; Department of Computer Science, Korea Advanced Institute of Science and Technology; the Key Lab of High Confidence Software Technologies (CUHK Sub-Lab), Ministry of Education, China; and the Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong (SEEM-CUHK) from April 22-23, 2014
- Constituted a Special Interest Group (SIG) in information retrieval at SMIT
- Regularly assisted the computer laboratory in-charge with hardware and software maintenance at SMIT
- Active member of the students’ technical body, ACCESS at SMIT
- Performed several social services such as teaching computers to backward people in India
- Lifetime member of the Numismatics Club of Jamshedpur, India
- An active participant in debates, spelling contests, etc

#### **PROGRAMMING SKILLS** Programming and Scripting Languages

- C, C++, MATLAB, R, Python
- SED, AWK, PERL, Linux Shell Scripting

#### Productivity Applications

- L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub>, MS Office, Libreoffice, Vim

#### Operating Systems

- Linux and its variants, Microsoft Windows, Sun Solaris