Anubrata Das

anubrata.github.io • anubrata[at]utexas.edu
1616 Guadalupe Street, STE - 5.202 • Austin, TX, USA, 78701

Education

- **University of Texas at Austin**   **Austin, Texas**
  Ph.D. Student, School of Information   *08/2018 – Present*

  - Committee: Dr. Matt Lease *(Chair)*, Dr. Min Kyung Lee, Dr. Ken Fleischmann

  - Research: Developing Human-AI partnership for Curbing Misinformation Online: An Explainable Natural Language Processing Approach

  - GPA: **3.97**/4.00 (Coursework)
  - **Graduate Courses**: Human-AI Interaction, Natural Language Processing, Research in Computational Linguistics, Information Retrieval, Machine Learning, Supervised Teaching in Information Science

- **Indian Institute of Engineering Science and Technology Shibpur**   **Kolkata, India**
  Bachelor of Engineering, Department of Computer Science and Technology   *07/2011 – 06/2015*

  - GPA: **8.43**/10.00
  - First Class with Honors

Research Experience

- **University of Texas at Austin**   **Austin, Texas**
  Research Assistant, [Information Retrieval and Crowdsourcing Lab](.)   *08/2018 – Present*

  - Advisor: Dr. Matthew Lease
  - Research Areas: *Explainable NLP, Crowdsourcing,*
    Human-AI Interaction, Content Moderation, Automated Fact-Checking

- **Max Planck Institute of Informatics**   **Saarbrücken, Germany**
  Research Intern, Databases and Information Systems Group   *06/2019 – 08/2019*

  - Advisor: Prof. Dr. Gerhard Weikum
  - Project: *Systematic discovery of bias: A case study on Airbnb Listings*

- **Indian Institute of Management Calcutta**   **Kolkata, India**
  Research Intern, Management Information Systems Group   *10/2012 – 09/2015*

  - Advisor: Dr. Somprakash Bandyopadhyay
  - Project: *Interactive crowdsourcing on social media for micro-level need assessment using for disaster management*

- **Indian Institute of Technology Kharagpur**   **West Bengal, India**
  Research Intern, Complex Networks and Research Group   *05/2013 – 07/2013*

  - Advisor: Dr. Saptarshi Ghosh
  - Project: *Prediction of Twitter trends using Machine Learning and Data Mining Techniques*

Publications [Google Scholar]

* denotes equal contribution

Peer Reviewed Conference - Long Papers

Das, A.*, Gupta, C.*, Kovatchev, V., Lease, M., & Li, J. J. (2022, May) *ProtoTexNL: Explaining Model Decisions with Prototype Layers*. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL).

Gupta, S., Singh, G., **Das, A.**, & Lease, M. (2021, July). *Pareto Solutions vs Dataset Optima: Concepts and Methods for Optimizing Competing Objectives with Constraints in Retrieval*. In Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval (pp. 43-52). (**51%** acceptance rate)

Das, A., Dang, B., & Lease, M. (2020, October). *Fast, Accurate, and Healthier: Interactive Blurring Helps Moderators Reduce Exposure to Harmful Content.* In Proceedings of the AAAI Conference on Human Computation and Crowdsourcing (Vol. 8, No. 1, pp. 33-42). (**27%** acceptance rate)

Das, A., Anjum, S., & Gurari, D. (2019). *Dataset bias: A case study for visual question answering.* Proceedings of the Association for Information Science and Technology, 56(1), 58-67.

Peer Reviewed Conference - Short Papers

Shi, L., Bhattacharya, N., **Das, A.**, Lease, M., & Gwizdka, J. *The Effects of Interactive AI Design on User Behavior: An Eye-tracking Study of Fact-checking COVID-19 Claims.* In Proceedings of the 7th ACM SIGIR Conference on Human Information, Interaction and Retrieval (CHIIR), 2022.

Journal Articles

Ekstrand, M. D., **Das, A.**, Burke, R., & Diaz, F. *Fairness and Discrimination in Information Access Systems.* Foundations and Trends in Information Retrieval (To Appear) ([arxiv preprint](#)) (2021) (100 pages). (**SJR Score 0.73**)

Book Chapters

Ekstrand, M. D., **Das, A.**, Burke, R., & Diaz, F. *Fairness in Recommendation.* Recommender Systems Handbook (To Appear) (2021) (40 pages).

Peer Reviewed Workshops

Das, A., Mehta, K., & Lease, M. (2019). *CobWeb: A Research Prototype for Exploring User Bias in Political Fact-Checking.* In ACM SIGIR Workshop on Fairness, Accountability, Confidentiality, Transparency, and Safety in Information Retrieval (FACTS-IR), 2019. (8 pages)

Preprints / Informal Publications

Singh, P., **Das, A.**, Li, J. J., & Lease, M. (2021, September). *The Case for Claim Difficulty Assessment in Automatic Fact Checking* (8 pages).

Das, A., Lee, S., Nguyen, T. A., Kharosekar, A., Krishnan, S., Krishnan, S., Tate, E., Wallace, B. C., & Lease, M. *ExFacto: An Explainable Fact-Checking Tool.* In Knight Research Network Tool Demonstration Day, 2021. Presentation abstract.

Das, A., & Lease, M. (2019). *A Conceptual Framework for Evaluating Fairness in Search.* (4 pages) arXiv:1907.09328.

Undergraduate Work

Das, A., Mallik, N., Bandyopadhyay, S., Bit, S. D., & Basak, J. (2016, March). *Interactive information crowdsourcing for disaster management using SMS and Twitter: A research prototype.* In 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops) (pp. 1-6). IEEE.

Das, A., Roy, M., Dutta, S., Ghosh, S., & Das, A. K. (2014, December). *Predicting trends in the twitter social network: A machine learning approach.* In International Conference on Swarm, Evolutionary, and Memetic Computing (pp. 570-581). Springer, Cham.

Funding

- **Graduate School Fellowship**   *08/2018 – 05/2019*
  - Awarded by the Graduate School, University of Texas at Austin
- **UT Good Systems Grand Challenge -** Graduate Student Grant Proposal. **Anubrata Das**, Chenyan Jia, Shivam Garg. Supervisor: Dr. Min Kyung Lee. *Designing algorithmic nudge to reduce inadvertent COVID-19 misinformation sharing on social media*. Awarded - $7000.

Presentations

Invited Talks

Commercial Content Moderation and Psychological Well-Being

- TxHCI - A seminar organized by HCI Researchers across Universities in Texas, 10/02/2020
- Amazon AWS Science, 10/14/2020

- Amazon Human-in-the-loop (HILL) services team, 10/23/2020
- ACM SIGCHI Mumbai Chapter, 26th Meet, 08/28/2021

Oral Presentations

Das, A. ProtoTex: Explaining Model Decisions with Prototype Layers. Research Colloquium, School of Information, University of Texas at Austin. November 2021. Lightning Talk.

Das, A. ProtoBART: Explaining Model Decisions with Prototype Layers. TACCSTER: TACC Symposium for Texas Researchers. September 2021. Lightning Talk.

Das, A. Fast, Accurate, and Healthier: Interactive Blurring Helps Moderators Reduce Exposure to Harmful Content. AAAI HCOMP 2020. Virtual Conference.

Das, A. CobWeb: A Research Prototype for Exploring User Bias in Political Fact-Checking. ACM SIGIR Workshop on Fairness, Accountability, Confidentiality, Transparency, and Safety in InformationRetrieval (FACTS-IR), 2019. Paris, France.

Awards and Honors

- Annual Diversity & Inclusion Best Student paper award   05/2019
  - Dataset Bias: Predicting and Understanding Implications for Visual Question Answering
  - Awarded by the School of Information, University of Texas at Austin

- Spot Award - Mu Sigma Inc.   2016
  - Awarded by the Innovation and Development Team
  - Interactive visualization for Stock Market as a network

- Class of 1990 Award: Excellence in Leadership   02/2014
  - Awarded by the Global Alumni Association of BESU (now IIEST)

Teaching and Mentoring

- Teaching Assistant   *Fall 2020*
  - INF385T.3 / CS395T: Human Computation and Crowdsourcing by Dr. Matt Lease
  - Three 60-minutes Tutorials on Amazon Sagemaker Ground Truth for collecting data annotations

- Co-Supervising undergraduate research group with Dr.Matt Lease   *06/2020 - Present*
  - A group of ten students
  - Working on fact-checking using NLP and Human-computation methods

---

Service

- Reviewer
  - ACL Rolling Review 2022
  - CHI 2021, 2022
  - CSCW 2021, 2022
  - The Web Conference 2021
  - Annual Meeting of the Association for Information Science and Technology: 2019, 2020
  - Journal: Information Processing and Management

- Committee
  - Assistant Professor Hiring Committee 2020-2021
  - Doctoral Studies Committee, School of Information, 2019-2020

- Student Volunteer
  - CSCW 2019

Industry Experience

- **Microsoft**   **Hyderabad, India**
  Software Engineer   *04/2018 – 07/2018*
    - Build, debug and maintain a marketing management tool for Microsoft Universal Store
- **Microsoft**   **Hyderabad, India**
  Associate Consultant   *11/2016 – 04/2018*
    - Develop solutions for enterprise search for a Fortune 500 oil and gas corporation
- **Mu Sigma**   **Bangalore, India**
  Decision Scientist   *08/2015 – 10/2016*
    - Design and build research prototypes for algorithmic trading using machine learning

Skills

Research Methodologies: Experimental Design, User Study, Crowdsourcing,
Natural Language Processing, Machine Learning, Inferential Statistics
Programming Languages: Python, R, JavaScript, SQL
Technologies: Flask, Pytorch, Scikit-Learn, NLTK, SciPy, NumPy, Git
Survey Tools: Qualtrics
Crowdsourcing: Amazon Mechnaical Turk, AWS Sagemaker Ground Truth, AWS Augmented AI
Languages: Fluent in English and Bengali, Knowledge of Hindi