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. Candidate, School of Information — *08/2018 – Present*
  - Committee: Dr. Matt Lease (co-advisor), Dr. Junyi Jessy Li (co-advisor, Dept. of Linguistics), Dr. Min Kyung Lee, Dr. Ken Fleischmann
  - Dissertation: Retrieving and Evaluating Example-based Explanations in Natural Language Processing
  - GPA: **3.99**/4.00 (Coursework)
  - **Graduate Courses**: Natural Language Processing, Research in Computational Linguistics, Information Retrieval, Machine Learning, Human-AI Interaction, 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

Selected Publications [Google Scholar]

Manuscript in Preparation/ Under review

1. **Das, Anubrata**, Manoj Kumar, Ninareh Mehrabi, Morteza Ziyadi, Anil Ramakrishna, Kai-Wei Chang, Aram Galstyan, Anna Rumshisky, and Rahul Gupta. Localizing and editing autoregressive language models for reducing toxicity. 2024. (Unver review at Conference on Language Modeling)

Refereed Conference/Journal Articles

1. Houjiang Liu*, **Das, Anubrata***, Alexander Boltz*, Didi Zhou, Daisy Pinaroc, Matthew Lease, and Min Kyung Lee. Human-centered NLP for fact-checking: Co-designing with fact-checkers using matchmaking for AI. *arXiv preprint arXiv:2308.07213*, 2023. (Accepted for CSCW 2024)(* denotes equal contribution)
2. **Das, Anubrata**, Houjiang Liu, Venelin Kovatchev, and Matthew Lease. The state of human-centered nlp technology for fact-checking. *Information Processing & Management, Special Issue on Machine and Human Factors in Misinformation Management*, 2023. (Impact Factor: 6.222)
3. Li Shi, Nilavra Bhattacharya, **Das, Anubrata**, and Jacek Gwizdka. True or false? cognitive load when reading covid-19 news headlines: an eye-tracking study. In *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*, pages 107–116, 2023
4. Li Shi, Nilavra Bhattacharya, **Das, Anubrata**, Matt Lease, and Jacek Gwizdka. The effects of interactive ai design on user behavior: An eye-tracking study of fact-checking covid-19 claims. In *ACM SIGIR Conference on Human Information Interaction and Retrieval*, pages 315–320, 2022
5. **Das, Anubrata**, Houjiang Liu, Venelin Kovatchev, and Matthew Lease. The need for human-centered design in fact-checking research. In *Information Processing & Management Conference*, 2022
6. Michael D Ekstrand, **Das, Anubrata**, Robin Burke, and Fernando Diaz. Fairness in information access systems. *Foundations and Trends® in Information Retrieval*, 16(1-2):1–177, 2022 (**SJR Score 0.73**)
7. **Das, Anubrata***, Gupta, Chitrank*, Venelin Kovatchev, Matthew Lease, and Junyi Jessy Li. ProtoTEx: Explaining model decisions with prototype tensors. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics* (*Volume 1: Long Papers*), pages 2986–2997, Dublin, Ireland, May 2022. Association for Computational Linguistics. (* denotes equal contribution)
8. Soumyajit Gupta, Gurpreet Singh, **Das, Anubrata**, and Matthew Lease. 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*, ICTIR '21, page 43–52, New York, NY, USA, 2021. Association for Computing Machinery

9. **Das, Anubrata**, Brandon Dang, and Matthew Lease. Fast, accurate, and healthier: Interactive blurring helps moderators reduce exposure to harmful content. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 8, pages 33–42, 2020
10. **Das, Anubrata**, Samreen Anjum, and Danna Gurari. Dataset bias: A case study for visual question answering. *Proceedings of the Association for Information Science and Technology*, 56(1):58–67, 2019. (Diversity and Inclusion **student best paper Award** by the School of Information, UT Austin)

Lightly Refereed Publications

1. Michael D Ekstrand, **Das, Anubrata**, Robin Burke, and Fernando Diaz. Fairness in recommender systems. In *Recommender systems handbook*, pages 679–707. Springer, 2022
2. **Das, Anubrata**, Kunjan Mehta, and Matthew Lease. Cobweb: A research prototype for exploring user bias in political fact-checking. *ACM SIGIR Workshop on Fairness, Accountability, Confidentiality, Transparency, and Safety in Information Retrieval (FACTS-IR) (8 Pages)*, 2019
3. Venelin Kovatchev, Trina Chatterjee, Venkata S Govindarajan, Jifan Chen, Eunsol Choi, Gabriella Chronis, **Das, Anubrata**, Katrin Erk, Matthew Lease, Junyi Jessy Li, et al. longhorns at dadc 2022: How many linguists does it take to fool a question answering model? a systematic approach to adversarial attacks. In *Proceedings of the First Workshop on Dynamic Adversarial Data Collection*, pages 41–52, 2022
4. Alexandra Olteanu, Jean Garcia-Gathright, Maarten de Rijke, Michael D Ekstrand, Adam Roegiest, ... **Das, Anubrata**, et al. Facts-ir: fairness, accountability, confidentiality, transparency, and safety in information retrieval. In *ACM SIGIR Forum*, volume 53, pages 20–43. ACM New York, NY, USA, 2021. (**Workshop Report**)

Preprints

1. Prakhar Singh, **Das, Anubrata**, Junyi Jessy Li, and Matthew Lease. The case for claim difficulty assessment in automatic fact checking. *arXiv preprint arXiv:2109.09689*, 2021
2. **Das, Anubrata** and Matthew Lease. A conceptual framework for evaluating fairness in search. *arXiv preprint arXiv:1907.09328*, 2019

Undergraduate Work

1. **Das, Anubrata**, Neeratyoy Mallik, Somprakash Bandyopadhyay, Sipra Das Bit, and Jayanta Basak. 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)*, pages 1–6. IEEE, 2016
2. **Das, Anubrata**, Moumita Roy, Soumi Dutta, Saptarshi Ghosh, and Asit Kumar Das. Predicting trends in the twitter social network: a machine learning approach. In *International Conference on Swarm, Evolutionary, and Memetic Computing*, pages 570–581. Springer, 2014

Research Experience

- **University of Texas at Austin** — **Austin, Texas**
  Research Assistant, Artificial Intelligence and Human Centered Computing Lab — *08/2018 – Present*
  - Advisor: Dr. Matthew Lease, Dr. Junyi Jessy Li
  - Research Areas: *see above*
- **Cisco Research, Responsible AI** — **Remote**
  Research Intern — *09/2023 – 12/2023*
  - Project: Example Selection for In-Context Learning
- **Amazon Alexa Responsible AI** — **New York City, NY**
  Applied Scientist Intern — *06/2023 – 09/2023*
  - Project: Model editing for detoxifying natural language generation
- **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*

Funding

- Evaluating Example-based Explainable Models in Large Language Models. **Amazon AWS Cloud Credit for Research**. Funding period: 11/30/2022 - 11/30/2023. **26,000 USD** (AWS Service Credits only).
- **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 - USD 7000.

Presentations

Invited Talks

Developing Language Technologies to Complement Human Capabilities

- Microsoft Research FATE Group, New York City, 02/16/2024
- McCombs School of Business, University of Texas at Austin, 02/12/2024

ProtoTEx: Explaining Model Decisions with Prototype Tensors

- Research Colloquium, UT Austin, iSchool, 09/20/2022
- iSchools European Doctoral Seminar Series, 09/16/2022
- Amazon Science Clarify Team, 05/17/2022
- NEC Laboratories Europe, 06/09/2022

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

Conference Presentations

Das, A. ProtoTEx: Explaining Model Decisions with Prototype Tensors. ACL. May 2022. Dublin, Ireland.

Das, A. You are what you tweet: Profiling users by past tweets to improve hate speech detection. iConference. March 2022. Virtual Conference.

Das, A. Exfacto: An explainable fact-checking tool. Knight Research Network Tool Demonstration Day, 2021. Virtual Conference.

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

Das, A. Dataset bias: A case study for visual question answering. ASIS&T 2019. Melbourne, Australia.

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.

Other 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.

Awards and Honors

- Annual Diversity & Inclusion Best Student paper award 05/2019
  - Das et al., ASIS&T 2019 (see the pulication section above)
  - 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

- Co-Supervising student research with Dr. Matt Lease *01/2022 - 06/2022*
  - Undergraduate thesis on Active Learning with Natural Language Rationales
  - Featured in UT Austin, College of Natural Sciences News
- 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 - 08/2021*
  - A group of ten students
  - Working on fact-checking using NLP and Human-computation methods

---

Service

- Program Committees and Reviewing
  - BlackboxNLP Workshop 2022
  - ACL Rolling Review 2022, 2023, 2024
  - AAAI AIES 2022
  - CHI 2021, 2022
  - CSCW 2021, 2022, 2023
  - The Web Conference 2021
  - Annual Meeting of the Association for Information Science and Technology: 2019, 2020
  - Journal: Information Processing and Management
- Conference Volunteer
  - ACL 2022
  - CSCW 2019
- University Committees
  - Assistant Professor Hiring Committee 2020-2021
  - Doctoral Studies Committee, School of Information, 2019-2020

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
