Sunny Rai, Ph.D.

☐ +1 267 249 9835 • ✉ sunnyrai@seas.upenn.com

🌐 https://raisunny.com/

Research Experience

2022-Present	University of Pennsylvania, U.S.A Postdoctoral Researcher Advisor: Dr. Sharath C. Guntuku & Dr. Lyle H. Ungar ○ Lead development of multilingual, multimodal digital media content analysis (movies, TV shows, Youtube videos) for extracting human preferences (such as social norms) for value alignment of LLMs ○ Demonstrated cultural dichotomy in how and why shame is expressed in Hollywood and Bollywood, and its utility for extracting social norms ○ Raised 50K USD to build value-aligned LLMs capable of adapting its conversational style to match users' preferences for vaccine counseling and promoting self-referral ○ Designed and implemented a multi-phase Randomized Control Trial with 930 vaccine-hesitant parents to test persuasive abilities of off-the-shelf LLMs (improved vaccination intent by 10%) ○ Built explainable AI/ML models to examine if and how language of people with depression varies; showed racial bias in AI models; research featured in 35+ media articles including Forbes and Reuters. ○ Co-PI for R-01 grant for >811K USD (under review at NIH) for AI-driven technology for depression risk assessment and promoting help-seeking behaviors.
2024-Present	The World Bank, U.S.A Consultant, DIME AI ○ Leading the design and implementation of an LLM-driven approach to extract gender norms from multimodal multilingual digital media content.
2020-2022	École Centrale School of Engineering, Mahindra University (MU), India Assistant Professor (*tenure-track*) ○ Developed and delivered the inaugural course for Natural Language Processing ○ Taught Introduction to Programming, Data Structures and Algorithm Design & Analysis to batch of size > 300 ○ Mentored 20+ students for capstone covering areas such as Dataset curation and benchmarking, LLM evaluation, Controlled text generation, Human preferences extraction, LLM fine-tuning

Education

2014-2020	Ph.D. in Computer Engineering, University of Delhi, India Thesis: *Computational Metaphor Processing in English Language*
2012-2014	M.Tech in Information Security, GGS Indraprastha University, India
2008-2012	B.Tech in Information Technology, GGS Indraprastha University, India

Grants

Total amount = USD 71K

1. Leonard Davis Institute of Health Economics, "M2SurveyBot A MultiModal Chatbot to Understand Performance Gap in Language-based Depression Risk Assessment for Black Populations" (2024),
   Role: Co-PI, Amount: USD 19.5K

2. The Penn Medical Communication Research Institute, "A Culturally Aligned LLM-powered Conversational Agent to Promote HPV Vaccination Awareness and Self-Referral" (2024)
   Role: Co-PI, Amount: USD 50K

3. NIH R-01: "Culturally Informed Large Language Model-based Depression Assessment to Address Racial Disparities in Black Adult Mental Health". Under review (Oct'24 cycle)
   Role: Co-PI, Amount: USD 811K

4. Travel Grant for presenting research at NAACL 2016, Netaji Subhas Institute of Technology, University of Delhi.
   Amount: USD 1500

Selected Publications

Please see Google Scholar Profile here for the complete list. * indicates equal contribution.

Conferences..

1. **Rai, S.**, Zaveri, K., Havaldar, S., Nema, S., Ungar, L., & Guntuku, S. C. (2025, April). Social Norms in Cinema: A Cross-Cultural Analysis of Shame, Pride and Prejudice. In Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers) (pp. 11396-11415).

2. Jamalova, R., **Rai, S.**, Thorbecke, L. & Guntuku, S.C. (2025). High-Risk Communities: Mapping the Drivers of COVID-19 Misinformation Across U.S. Counties. WebSci'2025

3. Havaldar, S., Giorgi, S., **Rai, S.**, Talhelm, T., Guntuku, S. C., & Ungar, L. (2024, June). Building Knowledge-Guided Lexica to Model Cultural Variation. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers) (pp. 211-226).

4. Cho, Y. M., **Rai, S.**, Ungar, L., Sedoc, J., & Guntuku, S. (2023, December). An Integrative Survey on Mental Health Conversational Agents to Bridge Computer Science and Medical Perspectives. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (pp. 11346-11369).

5. Joseph, R., Liu, T., Ng, A. B., See, S., & **Rai, S**. (2023, July). NewsMet: A 'do it all' dataset of contemporary metaphors in news headlines. In Findings of the Association for Computational Linguistics: ACL 2023 (pp. 10090-10104).

6. Khaliq, M., Joseph, R., & **Rai, S**. (2021, December). #covid is war and #vaccine is weapon? COVID-19 metaphors in India. In Proceedings of the 18th International Conference on Natural Language Processing (ICON) (pp. 431-438).

7. **Rai, S.**, & Chakraverty, S. (2017). Metaphor detection using fuzzy rough sets. In Rough Sets:

International Joint Conference of Rough Sets, IJCRS 2017, Olsztyn, Poland, July 3–7, 2017, Proceedings, Part I (pp. 271-279). Springer International Publishing.

Journal ....

1. **Rai, S.**, Stade, E.C., Zhou, G., Sehgal, N. K. R., Schriger, S., Ungar, L. & Guntuku, S.C. (2025). Cross Cultural Considerations for Designing AI Mental Health Applications. R & R The Lancet Digital Health

2. Pang, D., Guntuku, S. C., Sherman, G., Liu, T., **Rai, S.**, ... & Ungar, L. H. (2025). Cultural universalities and differences in the social media Language of gender and age between China and the US. Humanities & Social Sciences Communications.

3. **Rai, S.**, Stade, E. C., Giorgi, S., Francisco, A., Ungar, L. H., Curtis, B., & Guntuku, S. C. (2024). Key language markers of depression on social media depend on race. Proceedings of the National Academy of Sciences, 121(14), e2319837121.
   Impact factor: 9.4
   🏆 **Featured Paper of the Month** – September 2024, National Institute of Drug Abuse, NIH.
   🏆 **Early career lightning talk, IC2S2 2023**

4. **Rai, S.**, Stade, E. C., Giorgi, S., Francisco, A., Ungar, L. H., Curtis, B., & Guntuku, S. C. (2024). Reply to De Freitas: Social media language explains less of the variance in depression of Black individuals than of White individuals. Proceedings of the National Academy of Sciences, 121(29), e2411458121.
   Impact factor: 9.4

5. **Rai, S.**, Stade, E. C., Giorgi, S., Francisco, A., Ganesan, A. V., Ungar, L. H., ... & Guntuku, S. C. (2024). Reply to Wang: Clarifying model performance and language markers of depression across races. Proceedings of the National Academy of Sciences, 121(31), e2410449121.
   Impact factor: 9.4

6. Southwick, L., Sharma, M., **Rai, S.**, ..., Guntuku, S.C., & Merchant, R.M. (2024). Integrating Patient-Generated Digital Data Into Mental Health Therapy: Mixed Methods Analysis of User Experience. JMIR Mental Health.
   Impact factor: 4.8

7. **Rai, S.**, Kornides, M., Morgan, J., Kumar, A., Cappella, J., & Guntuku, S. C. (2024). Detecting and monitoring concerns against HPV vaccination on social media using large language models. Scientific Reports, 14(1), 14362.
   Impact factor: 3.8

8. **Rai, S.**, & Chakraverty, S. (2020). A survey on computational metaphor processing. ACM Computing Surveys (CSUR), 53(2), 1-37.
   Impact factor: 23.8

9. **Rai, S.**, Chakraverty, S., Tayal, D. K., Sharma, D., & Garg, A. (2019). Understanding metaphors using emotions. New Generation Computing, 37, 5-27.
   Impact factor: 2

10. **Rai, S.**, Chakraverty, S., Tayal, D. K., & Kukreti, Y. (2018). A study on impact of context on metaphor detection. The Computer Journal, 61(11), 1667-1682.
    Impact factor: 1.5

11. Jain, A.♣, Tayal, D. K.♣, & **Rai, S.** (2015). Shrinking digital gap through automatic generation of WordNet for Indian languages. AI & SOCIETY, 30, 215-222.

Impact factor: 2.9
♣ - Thesis Advisors for M.Tech

Workshops

1. Havaldar, S., Cho, Y, **Rai, S.**, & Ungar, L. (2025). Culturally-Aware Conversations: A Framework & Benchmark for LLMs. NeurIPS 2025 Workshop on Evaluating the Evolving LLM Lifecycle: Benchmarks, Emergent Abilities, and Scaling

2. Gupta, A., Joseph, R. & **Rai, S.** HATS: A Hindi Analogy Test Set for Evaluating Reasoning in LLMs. The Second Workshop on Analogical Abstraction in Cognition, Perception, and Language (Analogy-Angle II). ACL 2025

3. **Rai, S.***, Shelat, K.*, Jain, D. R., Cho, Y. M., Redkar, M., Sawant, S., Ungar., Lyle H. & Guntuku, S. C. (2024). Cross-Cultural Differences in Mental Health Expressions on Social Media. In 3rd Workshop on Cross-Cultural Considerations in NLP (C3NLP), NAACL 2025. arXiv preprint arXiv:2402.11477. ♛ **Outstanding Paper Award**

4. Havaldar, S., **Rai, S.**, Singhal, B., , Liu, L., Guntuku, S. C., & Ungar, L. (2023, July). Multilingual Language Models are not Multicultural: A Case Study in Emotion. In Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis (pp. 202-214).
   ♛ **Best Paper Award**

5. Bogireddy, N. R., Suresh, S., & **Rai, S**. (2023). I'm out of breath from laughing! I think? A dataset of COVID-19 Humor and its toxic variants. SocialNLP, In Companion Proceedings of the ACM Web Conference 2023 (pp. 1004-1013).
   ♛ **Dataset for CLEF 2024 Humor Processing Tasks**

6. Aggarwal, A., **Rai, S.**, Giorgi, S., Havaldar, S., Sherman, G., Mittal, J., & Guntuku, S. C. (2023). A Cross-Modal Study of Pain Across Communities in the United States. In Companion Proceedings of the ACM Web Conference 2023 (pp. 1050-1058).

7. **Rai, S.**, Joseph, R., Thakur, P. S., & Khaliq, M. A. (2022). Identifying Human Needs through Social Media: A study on Indian cities during COVID-19. In Proceedings of the Tenth International Workshop on Natural Language Processing for Social Media (pp. 65-74).

8. **Rai, S.**, Chakraverty, S., & Tayal, D. K. (2016). Supervised metaphor detection using conditional random fields. In Proceedings of the Fourth Workshop on Metaphor in NLP (pp. 18-27).

Preprints/ Work in Progress

1. Sehgal, N., **Rai, S.**, Tonneau, M., et al. Conversations with AI Chatbots Increase Short-Term Vaccine Intentions But Do Not Outperform Standard Public Health Messaging.

[T]alks/**[P]**osters/**[A]**bstracts..

T1 **Rai, S.**, Zaveri, K. J., Havaldar, S., Nema, S., Ungar, L., & Guntuku, S. C. (2024). A Cross-Cultural Analysis of Social Norms in Bollywood and Hollywood Movies. Oral Talk at IC2S2'24, Philadelphia.

P1 **Rai, S.***, Shelat, K.*, Jain, D. R., Cho, Y. M., Redkar, M., Sawant, S., & Guntuku, S. C. (2024). Studying Differential Mental Health Expressions in India. Poster at IC2S2'24, Philadelphia.

P2 Jamalova, R., **Rai, S.**, & Guntuku, S. C. (2024). Characterizing the Geographical Landscape of COVID-19 Misinformation from X. Poster at IC2S2'24, Philadelphia

A1 Baheti, B., **Rai, S.**, Innani, S., Mehdiratta, G., Guntuku, S. C., Nasrallah, M. P., & Bakas, S. (2024). "Detecting Histologic & Clinical Glioblastoma Patterns of Prognostic Relevance", Oral presentation at the 20th edition of the European Congress on Digital Pathology.

A2 Baheti, B., **Rai, S.**, Innani, S., Mehdiratta, G., Guntuku, S. C., Nasrallah, M. P., & Bakas, S. (2023). EPCO-15. DETECTING HISTOLOGIC & CLINICAL GLIOBLASTOMA PATTERNS OF PROGNOSTIC RELEVANCE. Neuro-Oncology, 25(Supplement_5), v126-v126.

T2 **Rai, S.**, Francisco, A., Giorgi, S., Curtis, B., Ungar, L., and Guntuku, S.C. (2023). Studying Racial Heterogeneity in Language Markers of Depression.
🏆 **Early Career Plenary Talk** at IC2S2'23, Denmark. Acceptance Rate: 1.74%

P3 Aggarwal, A., **Rai, S.**, Giorgi, S., Havaldar, S., Sherman, G., Mittal, J., and Guntuku, S.C. (2023). A Cross-Modal Study of Pain Across Communities in the United States. Workshop on Trustworthy AI, Jan 5-6, 2023 - Microsoft Research, Bengaluru, India.

P4 Joseph, R., **Rai, S.**, Thakur, P.S., & Khaliq, M.A. (2022). Pandemic, People and Pattern of needs: Emerging trends in Indian society during COVID-19. IC2S2 2022.

P5 Gollapudi, U. M , Arulkumaran, R., Perla, M., Jonnalagadda, S.H., Das, A., and **Rai, S.**. "*Understanding User Vulnerability Towards Radicalization On Twitter*". IC2S2 2019.

Invited Talks

1. "Using AI to decode Gender Norms in Global Entertainment Media." **The World Bank**, June 12, 2025

2. "Key Language Markers of Depression on Social Media Depend on Race", **Society for Personality and Social Psychology (SPSP) Annual Convention 2025** (pre-conference), Denver.

3. "Why and How Shame is expressed across Cultures?: A Case Study of Hollywood and Bollywood Movies", **National Institute of Drug Abuse, NIH**, Aug 21, 2024.

4. "Shame, Social Norms and Cultural Alignment of LLMs", **The 1st Workshop on Reclaiming the Narrative: Digital Recovery, AI, and Mitigating Harm in Social Media**, ICWSM, Buffalo, NY, June 3, 2024

5. Invited Research Talk on "Using LLMs for analyzing human behaviors: A case study of social norms ", **Center for Social Norms and Behavioral Dynamics, University of Pennsylvania**. Apr. 19, 2024.

6. Cross-Cultural Study of Social Norms in Hollywood and Bollywood, **Computational Linguistics Lunch'24 (CLunch), University of Pennsylvania**.

7. Panelist for Penn Data Scientist Meetup, **Penn Data-Driven Discovery Initiative**. May 29, 2024

8. Racial Heterogeneity in Language Markers of Depression , **Computational Linguistics Lunch'23 (CLunch), University of Pennsylvania**.

9. Multimodal explainable AI for prognostic stratification of glioblastoma patients, **Penn ASSET/IBI Symposium on Trustworthy AI for Health Care**. Oct 6, 2023

10. Language Models/ NLP Chatbots, **IBI-GCB Retreat at the University of Pennsylvania**. Sept, 2023.

11. Multilingual Zero-Shot Learning, **ATAL-FDP, a Government of India Initiative**. Nov'21

Awards & Scholarships

1. Outstanding Paper Award at C3NLP, NAACL'2025
2. 1st prize at Rachleff Scholars Program, University of Pennsylvania for "Characterizing the Geographical Landscape of COVID-19 Misinformation from Twitter."
3. Best Paper Award at WASSA, ACL'2023.
4. Abstract selected for Early Career Research Plenary Talk at IC2S2, 2023, Denmark. (16 Plenaries were selected out of 918 submissions)
5. 1st at Yale Undergraduate Research Symposium for Project "Mental Health Expressions in India"
6. Awarded Grace Hopper Celebration'22 Faculty Scholarship
7. UGC-NET (2018)- Ranked in Top 6% in National level Exam for Assistant Professorship
8. Graduate Scholarship (2012, 2014) by Government of India
9. Director's Award for Academic Excellence, 2011

Service

Chair
1. Session chair for "Computational Social Science and Cultural Analytics" track, NAACL 2024.
2. Publication chair, Session chair for "Computational Sequence Modeling", ICETCI'[21,22]

Reviewer
1. Conferences: COLM'24, ACL'[23,24], ICWSM'[23-24], EACL'23, AACL'[21-22], EMNLP'[21-24]
2. Workshops: NAACL-SRW'22, ACL-SRW'24
3. Other venues: IC2S2'24
4. Journals
   (a) AI/NLP: Computational Linguistics, ACM Transactions on Asian and Low Resource Language Information Processing (TALLIP), Natural Language Engineering, IEEE Transactions on Affective Computing, Intelligent Data Analysis
   (b) Health Sciences: Frontiers in Psychology, Journal of Language and Social Psychology, Psychiatry Research, ACM Transactions on Computing for Healthcare, Communications Psychology, Scientific Reports, Human Behavior and Emerging Technologies, Hindawi, Drug and Alcohol Dependence Reports, etc.
   (c) Others: IEEE Access, Neurocomputing, Computer Applications in Engineering Education (Wiley), Neurocomputing, Expert Systems, Computer Application in Engineering Education (Wiley)

Organizer
1. Building Culturally Competent Mental Health Applications with Large Language Models, ICWSM'25.
2. Summer Institute for Computational Social Science (SICSS), University of Pennsylvania, 2023

Science outreach & selected media coverage (48+)

- Cincinnati schools are using an app to identify suicidal kids. Not everyone is convinced, The Enquirer
- Could A Bot Be Your New Therapist? How AI Has Transformed Mental Healthcare, Forbes
- AI fails to detect depression signs in social media posts by Black Americans, study finds, Reuters
- Language models may miss signs of depression in Black people's Facebook posts, Science News
- "I" Usage Predicts Depression Severity in White but Not Black Individuals, HCP Live
- AI can detect depression in white people based on their social media posts – but not in Black people, Penn study finds, Philly Voice
- AI fails to detect depression signs in social media posts by Black Americans: study, The Economic Times
- AI Falters In Identifying Depression Cues From Facebook Posts Of Black Americans: Report, Benzinga
- New study reveals AI's hidden bias in screening depression in Black people, WION
- Language of Depression Among Black Americans is Not the Same as That for White Americans, Revealing a Limitation to AI, The Crusader Newspaper Group
- What do Hollywood and Bollywood regard as Shameful?, Dr. Alice Evans, The Great Gender Divergence

Teaching Experience

Natural Language Processing	Lead Instructor, Spring'21, 22 No. of students - 20-25 undergraduate students Normalized Course feedback[1]: Course material: 0.9; Delivery: 0.9
Introduction to Programming	Co-instructor, Spring'21 No. of students: ≈ 350
Algorithm Design and Analysis	Co-instructor, Spring'21 No. of students: ≈ 350
Discrete Mathematics	Lead instructor, Spring'15,16,17 No. of students: ≈ 70
Data Structures	Lead instructor, Fall'15,16,17 No. of students: ≈ 70

Mentorship (37+)

1. Chanya (Grace) Thanglerdsumpan, UPenn
2. Graise Zhou, UPenn
3. Khanh Vy Nguyen, UPenn -> Delta Airlines
4. Khushi Shelat, UPenn -> Rupa Health
5. Khushang Jilesh Zaveri, MU -> John Hopkins University
6. Ajinkya Bhalerao, MU -> Centrale Méditerranée
7. Sai Anish Sreeramagiri, MU -> Northeastern University
8. Devireddy Maneesh Reddy, MU -> University of South Florida
9. Lagishetty Rithwik, MU -> New Jersey Institute of Technology
10. Rohan Joseph, MU -> Telestra
11. Draksharam Sundara Raghuveer -> Stony Brook University
12. Sai Anirudh Yellanki -> UC Santa Cruz
13. Smriti Suresh, MU -> Boston University
14. Bogiredddy Neha Reddy, MU -> George Mason University
15. Sakshi Reddy Methukpally, MU -> University of Virginia
16. Vinuthna Ghattamaneni, MU -> University of Houston
17. Manoj Kumar Gunturu, MU -> University of Central Florida
18. Sree Deepya Bondalapati, MU -> Machine Learning Engineer, Jocata
19. Prakruti Singh Thakur, MU -> Arizona State University
20. Mohammad Abdul Khaliq, MU -> University of Groningen
21. Heamesh Nettam Choudhary, MU -> Rutgers University
22. Pavan Mallina -> ASU -> Software Engineer, Microsoft
23. Kamakshi Moparthi, MU -> Dartmouth
24. Manikumar Perla, MU -> Amazon
25. Venkata Sai Hitesh Jonalagadda, MU -> Columbia Business School
26. Uma Maheshwari Gollapudi, MU -> Indiana University
27. Apar Garg, Netaji Subhas Institute of Technology -> Google
28. Ayush Garg, Netaji Subhas Institute of Technology -> Amazon
29. Yash Kukreti, Netaji Subhas Institute of Technology -> Amazon