Suma P. Bhat

Email: sumabhat@princeton.edu

Education

Ph.D. Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, 2010.

M.A. South and Southeast Asian Studies, University of California, Berkeley, 2000.

M.E. Electrical Engineering, Indian Institute of Science, India, 1996.

B.S. Statistics, Mangalore University, India, 1992.

Academic Positions

Research Scholar, Department of Computer Science, Princeton University, August 2025–present.

Assistant Professor, Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, January 2020–July 2025.

Research Assistant Professor, Department of Electrical and Computer Engineering, University of Illinois, Urbana-Champaign, May 2015–December 2019.

Faculty Affiliate, Department of Computer Science, University of Illinois, Urbana-Champaign, November 2020-August 2025.

Faculty Affiliate, Department of Educational Psychology, University of Illinois, Urbana-Champaign, January 2020-August 2025.

Faculty Affiliate, National Center for Supercomputing Applications, University of Illinois, Urbana-Champaign, September 2016-August 2025.

Post-doctoral Fellow, Beckman Institute, University of Illinois, 2010.

Fellowships, & Awards

Ronald W. Pratt Faculty Outstanding Teaching Award, University of Illinois, Urbana-Champaign, 2021.

Beckman Foundation Post-doctoral Fellowship, 2012-2015.

Research Interests

Natural language processing, Human-AI interaction, Learning analytics, Medical language processing.

Professional Service

Area Chair, ACL Rolling Review January 2026 cycle

Meta-reviewer, International Conference on Audio, Speech and Signal Processing, 2025

Panelist, Workshop Rethinking Children's ASR for Education, University of Buffalo, 2025

Area Chair, International Conference on Natural Language Generation, (2023, 2024)

Associate Editor, IEEE Transactions on Audio, Speech and Language Processing (2022-present).

Co-organizer of the "Conversational AI" track at the 2022 US National Academy of Engineering's Annual Frontiers of Engineering (FoE) Symposium.

Panelist for NSF Smart and Connected Health Language and Behavior Panel of Reviewers, 2021

Standing Reviewer for the Transactions of the Association for Computational Linguistics, (2019-present).

Journal Reviewer: Journal of the American Medical Informatics Association, Journal of Educational Data Mining, Speech Communication.

Davidson Institute Fellows judge to evaluate projects and identify nations brightest students for a college scholarship award (2024)

Program Committee Member:

American Medical Informatics Association (AMIA; 2019)

Association for the Advancement of Artificial Intelligence (AAAI; 2018, 2019)

Association for Computational Linguistics (ACL; 2015)

Educational Data Mining (EDM;2019)

Empirical Methods for Natural Language Processing (EMNLP; 2015-2018)

International Joint Conference on Natural Language Processing (IJCNLP; 2016)

North American Association for Computational Linguistics (NAACL; 2019)

Workshop for Innovative Use of NLP for Building Educational Applications (BEA; 2015-2019)

Conference on Language Modeling (2024)

Students

Ph.D. Advisees

Hongyu Gong (graduated in May 2020)

Tarek Sakakini (graduated in June 2020)

Wanzheng Zhu (graduated in May 2022)

Ziheng Zeng (graduated in May 2023)

Jianing Zhou (graduated in May 2025)

M.S. Advisees

Jianing Zhou (CS, 2021)

Vyom Thakkar (ECE, 2022)

Kellen Cheng (Princeton University ECE, 2025)

Ph.D. Co-advised

Jiaqi Mu (2019, advisor: Pramod Viswanath)

Genevieve Henricks (2018, advisor: Michelle Perry)

Undergraduate Advisees

James Han (Princeton University, COS, BA 2026)

Feruza Mukhammadeva (Princeton University, COS, BA 2026)

Laiba Ali (Princeton University, COS, BA 2026)

Aasha Jain (Princeton University, COS, BA 2026)

Tanzila Morshed (Princeton University, COS, BA 2023)

Eve Rosenthal (Princeton University, COS, BA 2023)

Undergraduate Mentees at UIUC

Sarah Liu (ECE class of 2019, now Microsoft Inc.)

Khushi Arora (ECE class of 2020)

Akriti Jain (ECE class of 2020)

Vyom Thakkar (ECE class of 2020)

Kshitij Gupta (CS class of 2021)

Jay Huang (ECE class of 2021)

Daniel Polyakov (ECE class of 2021)

Pooja Bhagchandani (ECE class of 2022)

Neha Prabhu (CS class of 2022)

Sreenidhi Vijayraghavan (CS class of 2023)

Srihari Nanniyur (CS major at Washington University, St. Louis, Class of 2025)

High School Mentees

Aditya Mohan (Montavista High school, CA -¿ UC Berkeley EECS class of 2022)

Aayush Bhat (Cupertino High school, CA -¿ University of Illinois Urbana-Champaign, CS class of 2024)

Siddarth Vaidyanathan (Milpitas High School, CA -¿ UC Berkeley, CS+Statistics class of 2024)

Abhinav Ganesh (Obra D. Tompkins High School, TX; UT Austin CS class of 2025)

Divya Krishnan (Ridge High School, NJ; UPenn, Mechanical Engineering class of 2025)

Sridevi Pulugurtha (TERRA Environmental Research Institute, FL; Columbia University SEAS class of 2025)

Teaching

Princeton University

Junior Research Methods (COS 324 JRW), junior-level course, Fall 2022.

Introduction to Machine Learning (COS 324), junior-level course, Fall 2022.

Project-Based Seminar: Language, Computation, and Creativity (IW 10), Spring 2023.

Project-Based Seminar: Language and Computation, Spring 2026.

University of Illinois, Urbana-Champaign

Introduction to Computing (ECE 120), a required course for incoming undergraduate students in the ECE department: Fall 2015, Fall 2016, Fall 2017, Fall 2020. Fall 2021, Fall 2023, Spring 2024, Spring 2025.

Data Science and Engineering (ECE 365), junior/senior-level technical elective, Fall 2019, Spring 2020, Spring 2021.

Mathematical Models of Language (ECE 594), graduate-level course, Spring 2021, Spring 2022.

External Service and Outreach

Judge and Plenary Speaker at hackPHS, a student organized hackathon at Princeton High School, November 2025. (attendees: 100 students from NJ/PA)

Conducted an online professional development workshop in AI and Data Science for middle- and high school teachers under the auspices of CSTA (August 2025; 40 teachers)

Conducted an online professional development workshop in AI and Data Science for middle- and high school teachers (June 2025; 20 teachers)

Invited speaker to speak about my journey in AI at Vahati, an organization helping South Asian immigrant women navigate cultural transitions in the USA (March 2025).

Invited speaker at a professional development workshop on AI and Data science for middle- and high school teachers organized by DIYA Research (October 2024; 15 teachers)

Invited speaker at the New Jersey Science Convention for middle- and high school science teachers. (October 2024)

Panelist speaking about women engineers in academia at High School Engineering Colloquium (HSEC) for female high school students exploring STEM and engineering organized by the Society of Women Engineers at Princeton, NJ (Fall 2022, Fall & Spring 2023, Spring 2024; about 200 students in all)

Co-taught a data science unit at Princeton High School, NJ, in Spring 2024 (58 students—Sophomores, Juniors and Seniors)

Co-organized the North-American Computational Linguistics Open Competition for high school students at Princeton (2023, 2024) (60 students)

Guest speaker at the Linguistics and Computer science club of Hillsborough High School, NJ in Spring 2023 (25 students—Sophomores, Juniors and Seniors)

Led two 5-day workshops on data science for high school students from Chicago public schools under the auspices of the Chicago Pre-College Science and Engineering Program (ChiS&E) in Fall 2021 and Spring 2022.

Speaker at Data Science summer programs organized by DIYA Research Inc. (2020-present)

High school research mentor (2020-present, 6 weeks in the summer; 20 students in all)

Social entrepreneurship via creative opportunities for high school students and professional development for teachers in AI and Data Science (2020-present)

Judge in the Illinois Hackathon (2018, 2019)

Grants

Current Research

Co-PI Using LLMs to decode mental states from language, 1/2026-6/2027, $60,000. Funds allocated to me: $10,000. (Intramural)

Co-PI Environment and Intent Specific Adaptive PHY Communication via Deep Learning, Office of Naval Research, 9/2024-8/2027, $1,000,000, Lead PI: Pramod Viswanath. Funds allocated to me: $350,000.

Technical Co-lead NSF-IES AI Institute for Inclusive and Intelligent Technolo-

gies for Education (INVITE), 06/2023-05/2028, $19,998,746. Funds allocated to me: $451,000

Lead PI: Building Idiomaticity into Natural Language Processing, National Science Foundation (EAGER), Aug 2022-Dec 2026, $150,000. Funds allocated to me: $150,000.

Completed Research

Co-PI Teaching tools for thought with large language models, 1/2025-6/2026, $58,000. Funds allocated to me: $10,000. (Intramural)

Consultant: Patient Notes After Simulation Encounters as a Teaching Tool for Differential Diagnostic Reasoning, 01/2023-12/2023, $97,640. Lead PI: Manajyoti Yadav, Univ. of Illinois College of Medicine, Peoria. Funds allocated to me: $7,200.

Co-PI: Development of a Chatbot for Delivering Long-Term Motivational Interviewing for Improving Exercise Adherence in Hemodialysis Patients, The Jump Applied Research for Community Health through Engineering and Simulation program, Jan 2022- May 2023, $75,000. Lead PI: Jessie Chin, Univ. of Illinois, Urbana-Champaign.

Co-PI: Automated Assessment of Written Chart Notes: Generating Reliable, Timely, and Useful Feedback, Edward J. Stemmler, MD Medical Education Research Fund of the National Board of Medical Examiners (NBME), 2021-2023, $149,497.61. Lead PI: Dr. William Bond, Jump Simulation, Peoria.

Consultant: Application of Automated Grading of Written Patient Notes after Standardized Patient Encounters and its Influence on Feedback Quality, 01/2022-12/2022, $73,000. Lead PI: Saurabh Bansal, Univ. of Illinois College of Medicine, Peoria. Funds allocated to me: $7,200.

Lead PI: Using Study Partners to Broaden Participation, Institute for Inclusion, Diversity, Equity, and Access, Grainger College of Engineering (UIUC), 2020-2022, $13,000. Funds allocated to me: $13,000.

Co-PI: Building a Motivational-Interviewing Conversational Agent (MintBot) for Promoting COVID-19 Vaccination among People with Multiple Sclerosis, The Jump Applied Research for Community Health through Engineering and Simulation program, 2021-2022, $74,992. Lead PI: Jessie Chin. Funds allocated to me: $61,900

Co-PI: COVID-19 Risk Mitigation: Interactive Automatic Counselor, University of Illinois at Urbana-Champaign, 2020-2022, $5,000. PI: Dan Morrow.

Lead PI: Improving Feedback and Efficiency: Automated Grading of Post Simulation Written Chart Notes, The Jump Applied Research for Community Health through Engineering and Simulation program, 2020-2021, $75,000. Funds allocated to me: $61,900

Co-PI: Underrepresented Student Learning in Online Introductory STEM College Courses, Institute of Education Sciences, 2018-2021, $1,399,194. Lead PI: Michelle Perry. Funds allocated to me: $258,000

Co-PI: Using Conversational Agents to Support Older Adult Learning for Health, Technology Innovation in Educational Research and Design Program, 2020-2021, $13,602. Lead PI: Dan Morrow.

Lead PI: Collaborative Research; BystanderBots: Automated Bystander Intervention for Cyberbullying Mitigation, National Science Foundation (EAGER), 2017-2021, $300,000. Funds allocated to me: $100,000.

Co-PI: IBM-ILLINOIS Center for Cognitive Computing Systems Research (C3SR), 2016-2021.

Co-PI: Supporting Self-regulated Learning in Online Education via Automatically Personalized Interventions, Technology Innovation in Educational Research and Design Program, 2020-2021, $14,997. PI: Nigel Bosch.

Lead PI: A Nuanced Model for Recognizing Levels of Conflict in Decision Making Using Natural Language Processing, Social and Behavioral Sciences Research Initiative (SBSRI) Small Grants Program, 2017-2020, $13,200. Funds allocated to me: $13,200

Co-PI: Interactive Technology Support for Patient Medication Self-Management (Round 2 funding), Jump Applied Research for Community Health through Engineering and Simulation (ARCHES), UIC/OSF Hospital, 2018-2019, $52,224. PI: Daniel Morrow. Funds allocated to me: $33,000

Collaborator: The Role of Gesture in Mathematics Learning: From Research to Practice. National Science Foundation, Science of Learning Collaborative Network, 2016-2019, $747,903. PI: Susan Goldin-Meadow.

Co-PI: Understanding Learning Behavior Patterns in MOOCs to Support Early Interventions, Illinois Learning Sciences Design Initiative, 2016-2017, $15,000. Lead PI: Michelle Perry.

Co-PI: Interactive Technology Support for Patient Medication Self-management. Jump Applied Research for Community Health through Engineering and Simulation (ARCHES), UIC/OSF Hospital, 2016-2017, $113,958. Lead PI: Daniel Morrow. Funds allocated to me: $33,000

Invited Talks

Children's ASR in Action at National AI Institutes and Centers, Panelist, Inaugural International Workshop on Rethinking Children's Automatic Speech Recognition for Education, University of Buffalo (September 2025)

Towards Enhanced Linguistic Reasoning in Language Models, CLAY Seminar at Yale University (April 2025)

Towards Robust Linguistic Reasoning in Language Models, MBZUAI, Abu Dhabi (2025)

Enhancing Linguistic Reasoning in Language Models, Indian Institute of Technology, Bombay, India (2025)

Enhancing Language Models: Overcoming Challenges in Comprehending Idiomatic Expressions, NLP Seminar at the Hitotsubashi University, Japan (2025)

Enhancing Language Models: Overcoming Challenges in Comprehending Idiomatic Expressions, Tokyo Metropolitan University NLP Seminar, Japan (2025)

Giving Language models a Leg up with Idiomaticity, Montclair State University Brown Bag Series on Language, Computation, and Information (2025)

Giving Language models a Leg up with Idiomaticity, Columbia Natural Language Processing Seminar, Columbia University (2024)

Understanding Learners Through Learning Analytics, Code.org, Online (2023)

Idiomaticity in Pre-Trained Language Models, Princeton Language and Intelligence Seminar, Princeton University (2023)

Shining a Light on Dark Jargon and Countering Hate Speech, Autumn 2021 Research Colloquium Series, University of Washington, Seattle (2021)

Knowing the Unseen: Estimating Vocabulary Size over Unseen Samples, NLP Seminar, Central Institute of Indian languages, India (2010)

Special Achievements

Best Approach Award for demonstration in the New Innovation category of the SimVentors Showcase at the International Meeting for Simulation in Healthcare, January 2023, Orlando, FL. (#1 of 52 demonstrations)

Publications

Peer-Reviewed Journal Articles

1. N. L. Schroeder, S. Zhang, J. R. Gladstone, and S. Bhat. (2026). Trust, Privacy, and Context Dependence of Large Language Models in Educational Contexts: Technical Challenges and Opportunities. Journal of Educational Computing Research.
2. S. K. Bansal, M. Yadav, J. Zhou, R. A. EbertAllen, R. M. Klute, W. F. Bond, and S. Bhat. (2025). Medical Students' Perception of Automated Note Feedback After Simulated Encounters. The Clinical Teacher, 22(6).
3. M. Perry, R. F. Azevedo, G. Henricks, R. W. Crues, and S. Bhat. (2024). Learning From Online Instructional Videos Considering Video Presentation Modes, Technological Comfort, and Students' Characteristics. International Journal of Human-Computer Interaction, 1-18.
4. W. F. Bond, J. Zhou, S. Bhat, Y. S. Park, R. A. Ebert-Allen, A. L. Ruger, and R. Yudkowsky. (2023). Automated patient note grading: Examining scoring reliability and feasibility. Academic Medicine, 98(11S), S90-S97.
5. Z. Zeng, and S. Bhat. (2022). Getting BART to Ride the Idiomatic Train: Learning to Represent Idiomatic Expressions. Transactions of the Association for Computational Linguistics, 11.
6. G.M. Henricks, M. Perry, and S. Bhat. (2021). Gender and Gendered Discourse in Two Online Science College Courses. *Computer-Based Learning in Context*, 3 (1), 1–16.
7. Z. Zeng, and S. Bhat. (2021). Idiomatic expression identification using semantic compatibility. Transactions of the Association for Computational Linguistics, 9, 1546–1562.

8. S. Yoon and S. Bhat. (2018). A Comparison of Grammatical Proficiency Measures in the Automated Assessment of Spontaneous Speech. *Speech Communication*, Volume 99, May, 221–230.

9. R. W. Crues, G. Henricks, M. Perry, S. Bhat, C. Anderson, N. Shaik and L. Angrave. (2018). How does Gender, Learning Goals, and Forum Participation Predict Persistence in a Computer Science MOOC? *ACM Transactions on Computing Education*, 18(4).

10. K. Zechner, S. Yoon, S. Bhat, and C. Leong. (2017). Comparative Evaluation of Automated Scoring of Syntactic Competence of Non-native Speakers. *Computers in Human Behavior*, Volume 76, November, 672–682.

11. S. Bhat, and S. Yoon. (2015). Automatic Assessment of Syntactic Complexity of Spontaneous Speech Scoring. *Speech Communication* (67), 42–57.

12. R. Girju, B. Beamer, A. Rozovskaya, A. Fister, and S. Bhat. (2010). A Knowledge-rich Approach to Identifying Semantic Relations Between Nominals. *Information Processing & Management* 46(5), 589–610.

Published Conference Papers

13. J. Zhou, Z. Zeng, H. Gong, and S. Bhat. (2026). Dialect ASR based on Multi-View Pseudo-Parallel Augmentation and Noise-Robust Contrastive Learning. To Appear in NeurIPS 2026.

14. S. Almheiri, B. Elbouardi, S. Z. Pranida, I. Nikishina, A. Rao, P. Krishnamurthy, M. C. Airlangga, R. A. Genadi, N. P. G. Bao, A. H. Yari, H. O. Toyin, N. Mukhituly, M. Attia, B. Hassan, A. F. Hidayatullah, T. Kuribayashi, H. Li, S. Bhat, and F. Koto. (2026). Multilingual Idioms in Sentences and Conversations Across High-, Medium-, and Low-Resource Languages. Proceedings of the Association for Computational Linguistics (ACL) 2026.

15. N. Gangwar, S. Bhat, and N. Kani. (2026). Integrating Arithmetic Learning Improves Mathematical Reasoning in Smaller Models. Proceedings of the 2026 Language Resources and Evaluation Conference.

16. S. Zhang, P. S. Meshram, P. Ganapathy Prasad, M. Israel, and S. Bhat. (2025). An LLM-Based Framework for Simulating, Classifying, and Correcting Students' Programming Knowledge with the SOLO Taxonomy. In Proceedings of the 56th ACM Technical Symposium on Computer Science Education V. 2, pp. 1681-1682. 2025.

17. N. Gangwar, S. Bhat, and N. Kani. (2024). Intermediate Fine-Tuning Improves Mathematical Reasoning in Smaller Models. In Proceedings of the 4th Workshop on Mathematical Reasoning and AI at NeurIPS'24.

18. B. Bhavya, C. Palaguachi, Y. Zhou, S. Bhat, and C. Zhai. (2024). Long-Form Analogy Evaluation Challenge. (2024). In Proceedings of the 17th International Natural Language Generation Conference: Generation Challenges, pp. 1-16, 2024.

19. J. Zhou, Z. Zeng, H. Gong, and S. Bhat. (2024). Enhancing Language Models with Idiomatic Reasoning, In Proceedings of the First Conference on Language Modeling (COLM) 2024. (29% acceptance rate)

20. J. Zhou and S. Bhat. (2024) Non-compositional Expression Generation and its Continual Learning. In Findings of ACL 2024. (22% acceptance rate)

21. J. Zhou, Z. Zeng, H. Gong, and S. Bhat. (2024). CLASP: Cross-modal Alignment Using Pre-trained Unimodal Models. In Findings of ACL 2024. (22% acceptance rate)

22. K. Cheng and S. Bhat. (2024). No Context Needed: Contextual Quandary In Idiomatic Reasoning With Pre-Trained Language Models. In *Proceedings of the 17th (2024) Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL).* (26% acceptance rate)

23. G. M. Henricks, M. Perry, and S. Bhat. (2024). The Relation Among Gender, Language, and Posting Type in Online Chemistry Course Discussion Forums. In Proceedings of the 14th International Learning Analytics and Knowledge Conference (LAK24). Kyoto, Japan. New York, NY: ACM. pp. 189-199.

24. Z. Zeng, K. Cheng, S. Nanniyur, J. Zhou, and S. Bhat. (2023). IEKG: A Commonsense Knowledge Graph for Idiomatic Expressions. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, pages 14243–14264, Singapore. Association for Computational Linguistics. (23% acceptance rate)

25. Z. Zeng and S. Bhat. (2023). Unified Representation for Non-compositional and Compositional Expressions. In Findings of the Association for Computational Linguistics: EMNLP 2023, pages 11696–11710, Singapore. Association for Computational Linguistics. (23% acceptance rate)

26. J. Zhou, Z. Zeng, H. Gong, and S. Bhat. (2023). Non-compositional Expression Generation Based on Curriculum Learning and Continual Learning.

In Findings of the Association for Computational Linguistics: EMNLP 2023, pages 4320–4335, Singapore. Association for Computational Linguistics. (23% acceptance rate)

27. J. Zhou, Z. Zeng, and S. Bhat. (2023). CLCL: Non-compositional Expression Detection with Contrastive Learning and Curriculum Learning. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 730–743, Toronto, Canada. Association for Computational Linguistics. (23.5% acceptance rate)

28. S. A. Hebbar, V. V. Nadkarni, A. V. Makkuva, S. Bhat, S. Oh, and P. Viswanath. (2023). CRISP: Curriculum based Sequential neural decoders for Polar code family. In International Conference on Machine Learning, pp. 12823-12845. PMLR, 2023. (27.9% acceptance rate)

29. E. Cox, M. Hasegawa-Johnson, S. Bhat, M. Umashankar, C. Lane, and D. Morrow. (2023). The Importance of Diverse User Goals When Designing an Automated COVID Risk Counselor. In the Proceedings of the Health Care Symposium (2023).

30. J. Zhou, V. Thakkar, R. Yudkowsky, S. Bhat, and W. F. Bond (2022). "Automatic Patient Note Assessment without Strong Supervision." In Proceedings of the 13th International Workshop on Health Text Mining and Information Analysis (LOUHI), pp. 116-126. 2022.

31. P. Hur, H. Lee, S. Bhat, and N. Bosch (2022). Using Machine Learning Explainability Methods to Personalize Interventions for Students. In Proceedings of the 15th International Conference on Educational Data Mining (EDM 2022). International Educational Data Mining Society.

32. W. Zhu and S. Bhat (2022). "Slow Service" does not imply "Great Food": Enhancing Content Preservation in Unsupervised Text Style Transfer. In *15th International Natural Language Generation Conference*.

33. J. Zhou, Z. Zeng, H. Gong, and S. Bhat (2022) Idiomatic Expression Paraphrasing without Strong Supervision. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36, No. 10, pp. 11774-11782. (15% acceptance rate)

34. J. Zhou, H. Gong, and S. Bhat, PIE: Parallel Idiomatic Expression Corpus for Idiomatic Sentence Generation and Paraphrasing. In *Proceedings of the 17th Workshop on Multiword Expressions (MWE 2021)*, pp. 33-48.

35. N. Prabhu, M. Perry, R. F. L. Azevedo, L. Angrave, and S. Bhat (2021). Study Partners Matter: Impacts on Inclusion and Outcomes. In *2021 ASEE Annual Conference and Exposition, Conference Proceedings.*

36. W. Zhu, and S Bhat (2021). Euphemistic Phrase Detection by Masked Language Model. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pp. 163-168. 2021. (11.8% acceptance rate)

37. J. Zhou, and S. Bhat (2021). Paraphrase generation: A survey of the state of the art. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 5075-5086. 2021. (25.6% acceptance rate)

38. W. Zhu, H. Gong, R. Bansal, Z. Weinberg, N. Christin, G. Fanti, and S. Bhat (2021). Self-Supervised Euphemism Detection and Identification for Content Moderation. In *2021 IEEE Symposium on Security and Privacy (SP)*, pp. 229-246. (12.1% acceptance rate)

39. E. Cox, S. Bhat, M. Hasegawa-Johnson, L. H. Chad, I. Chen, and D. Morrow (2021). "Developing An Automated COVID Risk Counselor: Situation Type and Perceived Risk." In TMS Proceedings 2021.

40. W. Zhu and S. Bhat (2021). Generate, Prune, Select: A Pipeline for Counterspeech Generation against Online Hate Speech. In Findings of ACL 2021. (14.5% acceptance rate)

41. D. Williams-Dobosz, R. Azevedo, A. Jeng, V. Thakkar, S. Bhat, N. Bosch, and M. Perry (2021). A Social Network Analysis of Online Engagement for College Students Traditionally Underrepresented in STEM. In *Proceedings of the 11th International Learning Analytics and Knowledge (LAK) Conference 2021.*

42. J. Zhou, and S. Bhat (2021). Modeling Consistency of Engagement Patterns in Online Courses. *Proceedings of the 11th International Learning Analytics and Knowledge (LAK) Conference 2021.*

43. H. Gong, A. Valido, K. Ingram, G. Fanti, S. Bhat, and D. Espelage (2021). Abusive Language Detection in Heterogeneous Contexts: Dataset Collection and the Role of Supervised Attention. In Proceedings of the AAAI Conference on Artificial Intelligence 35(17), 14804–14812. (21.4% acceptance rate)

44. H. Gong, L. Song, and S. Bhat (2020). Rich Syntactic and Semantic Information Helps Unsupervised Text Style Transfer. In *Proceedings of the 13th International Conference on Natural Language Generation*, pp. 113–119.

45. T. Sakakini, J. Y. Lee, A. Duri, R. F.L. Azevedo, V. Sadauskas, K. Gu, S. Bhat, D. Morrow, J. Graumlich, S. Walayat, M. Hasegawa-Johnson, T. Huang, A. Willemsen-Dunlap, and D. Halpin (2020). Context-Aware Automatic Text Simplification of Health Materials in Low-Resource Domains. In *Proceedings of the 11th International Workshop on Health Text Mining and Information Analysis,* pp. 115–126.

46. H. Gong, S. Bhat, and P. Viswanath (2020). Enriching Word Embeddings with Temporal and Spatial Information. In *Proceedings of the 24th Conference on Computational Natural Language Learning,* pp. 1-11. (23% acceptance rate)

47. W. Zhu, and S. Bhat (2020), GRUEN for Evaluating Linguistic Quality of Generated Text. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pp. 94–108. (13.5% acceptance rate)

48. W. Zhu, H. Gong, J. Shen, C. Zhang, J. Shang, S. Bhat, and J. Han (2020). FUSE: Multi-Faceted Set Expansion by Coherent Clustering of Skip-grams. In Proceedings of ECML-PKDD 2020. (25% acceptance rate)

49. D. Morrow, R. F. L. Azevedo, L. Sari, K. Gu, T. Sakakini, M. Hasegawa-Johnson, S. Bhat, J. Graumlich, T. Huang, A. Hariharan, Y. Shao, and E. Cox. (2020). Closing the Loop in Computer Agent/Patient Communication. In *Proceedings of the 2020 Human Factors and Ergonomics Society Annual Meeting, Chicago, IL.*

50. F. Brahman, N. Varghese, S. Bhat and S. Chaturvedi. (2020). Effective Forum Curation via Multi-task Learning. In *Proceedings of the Educational Data Mining Conference 2020.* (22% acceptance rate)

51. V. Jay, G. Henricks, C. Anderson, L. Angrave, N. Bosch, N. Shaik, D. Williams-Dobosz, S. Bhat, and M. Perry. (2020). Online Discussion Forum Help-Seeking Behaviors of Students Underrepresented in STEM. In *Proceedings of the 2020 International Conference of the Learning Sciences (ISLS).*

52. G. Henricks, S. Bhat, and M. Perry, Gender and Gendered Discourse in Online STEM College Courses.(2020). In *Proceedings of the 2020 International Conference of the Learning Sciences.*

53. H. Gong, K. Gupta, A. Jain, and S. Bhat. (2020). Illinimet: Illinois system for metaphor detection with contextual and linguistic information. In *Proceedings of the Second Workshop on Figurative Language Processing.*

54. O. Anjum, H. Gong, S. Bhat, J. Xiong, and W. Hwu. (2019). PaRE: A Paper-Reviewer Matching Approach Using a Common Topic Space. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing.* (Long paper, 24% acceptance rate)

55. T. Sakakini, H. Gong, J. Y. Lee, R. Schloss, J. Xiong, and S. Bhat. (2019). Equipping Educational Applications with Domain Knowledge. In *Proceedings of the Workshop on Innovative Use of NLP for Building Educational Applications.* (Long paper)

56. R. F. L. Azevedo, D. Morrow, K. Gu, T. Huang, M. Hasegawa-Johnson, P. Soni, S. Tang, T. Sakakini, S. Bhat, A. Willemsen-Dunlap, and J. Graumlich. (2019). The Influence of Computer Agent Characteristics on User Preferences in Health Contexts. In *Proceedings of the 2019 Human Factors and Ergonomics Society Health Care Symposium.*

57. H Gong, S. Bhat, L. Wu, J. Xiong and W. Hwu. (2019). Reinforcement Learning Based Text Style Transfer without Parallel Training Corpus, In *Proceedings of the 17th (2019) Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL).* (Long paper, 26% acceptance rate)

58. H. Gong, Y. Li, S. Bhat, and P. Viswanath. (2019). Context-Sensitive Malicious Spelling Error Correction. In *Proceedings of the Web Conference (WWW), 2019.* (Short paper, 18% acceptance rate)

59. Z. Zeng, S. Chaturvedi, S. Bhat, and D. Roth. (2019). DiAd: Domain Adaptation for Learning at Scale, In *Proceedings of the 9th International Learning Analytics and Knowledge (LAK) Conference 2019.* (Long paper, 32% acceptance rate)

60. H. Gong, J. Mu, S. Bhat, and P. Viswanath. (2018). Preposition Sense Disambiguation and Representation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP).* (Long paper, 26% acceptance rate)

61. R. F. L. Azevedo, T. Sakakini, J. Y. Lee, V. Sadauskas, K. Gu, Y. Zhang, D. Morrow, S. Bhat, M. Hasegawa-Johnson, T. S. Huang, A. Willemsen-Dunlap, D. J. Halpin, and J. Graumlich. (2018). Using Conversational Agents to Explain Medication Instructions to Older Adults. *Proceedings of the AMIA Annual Symposium.* San Francisco, CA: American Medical Informatics Association.

62. H. Gong, T. Sakakini, S. Bhat, and J. Xiong. (2018). Document Similarity for Texts of Varying Lengths via Hidden Topics. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics.* (Long paper, 25% acceptance rate)

63. R. W. Crues, N. Bosch, M. Perry, L. Angrave, N. Shaik, and S. Bhat. (2018). Refocusing the Lens on Engagement in MOOCs. In R. Luckin, K. R. Koedinger, & S. Klemmer (Eds.), *Proceedings of the 5th (2018) ACM Conference on Learning@Scale.* New York, NY: ACM. (Long paper, 22% acceptance rate)

64. R. W. Crues, N. Bosch, C. J. Anderson, M. Perry, S. Bhat, and N. Shaik. (2018). Who They Are and What They Want: Understanding the Reasons for MOOC Enrollment. In K. E. Boyer & M. V. Yudelson (Eds.), *Proceedings of the 11th International Conference on Educational Data Mining (EDM 2018).* International Educational Data Mining Society. (Long paper, 16% acceptance rate)

65. H. Gong, S. Bhat, and P. Viswanath. (2018) Embedding Syntax and Semantics of Prepositions via Tensor Decomposition. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL).* (Long paper, 32% acceptance rate)

66. N. Bosch, R. W. Crues, G. Henricks, M. Perry, L. Angrave, N. Shaik, S. Bhat, and C. J. Anderson. (2018). Modeling Key Differences in Underrepresented Students' Interactions with an Online STEM Course. In *Proceedings of the TechMindSociety'18 conference.* New York, NY: ACM. (27% acceptance rate)

67. R. F. L., Azevedo, K. Gu, Y. Zhang, V. Sadauskas, T. Sakakini, D. Morrow, M. Hasegawa-Johnson, T. S. Huang, S. Bhat, A. Willemsen-Dunlap, D. J., Halpin, J. Graumlich, and W. Schuh. (2017). Using Computer Agents to Explain Clinical Test Results. *AMIA Annual Symposium Proceedings.* Washington, DC: American Medical Informatics Association.

68. T. Sakakini, R. F. L. Azevedo, V. Sadauskas, K. Gu, Y. Zhang, S. Bhat, D. Morrow, M. Hasegawa-Johnson, T. S. Huang, A. Willemsen-Dunlap, D. J. Halpin, and J. Graumlich. (2017). Dr. Babel Fish: A Machine Translator to Simplify Providers' Language. *In AMIA Annual Symposium Proceedings.* Washington, DC: American Medical Informatics Association.

69. Z. Zeng, S. Chaturvedi and S. Bhat. (2017). Learner Affect Through the Looking Glass: Characterization and Detection of Confusion in Online Courses,

Proceedings of the Educational Data Mining Conference 2017. (Short paper, 42% acceptance rate)

70. J. Mu, S. Bhat, and P. Viswanath. (2017). Representing Sentences as Low-Rank Subspaces. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 629-634. (Short paper, 19% acceptance rate)

71. T. Sakakini, S. Bhat, and P. Viswanath. (2017). MORSE: Semantic-ally Drive-n MORpheme SEgment-er, *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 552-561. (Long paper, 25% acceptance rate)

72. J. Mu, S. Bhat, and P. Viswanath. (2017). Geometry of Polysemy. In *Proceedings of the International Conference on Learning Representation*, ICLR. (Long paper, 39% acceptance rate)

73. H. Gong, S. Bhat, and P. Viswanath. (2017). Geometry of Compositionality. In *Proceedings of Thirty-First AAAI Conference on Artificial Intelligence (AAAI).* (Long paper, 25% acceptance rate)

74. S. Bhat, S. Yoon, and D. Napolitano. (2015). Automatic Detection of Grammatical Structures from Non-Native Speech. *Proceedings of the Workshop on Speech and Language Technology in Education*, Interspeech.

75. S. Bhat, P. Chinprutthiwong and M. Perry. (2015). Seeing the Instructor in Two Video Styles: Preferences and Patterns. *Proceedings of the 8th International Conference on Educational Data Mining (EDM)*, 305–312. (Long paper, 16% acceptance rate)

76. B. Amnueypornsakul and S. Bhat. (2014). Machine-guided Solution to Mathematical Word Problems. *Proceedings of the Pacific Asia Conference on Language, Information and Computing (PACLIC).*

77. B. Amnueypornsakul, S. Bhat, P. Chinprutthiwong. (2014). Predicting Attrition Along the Way: The UIUC Model. *Proceedings of the Workshop on Analysis of Large Scale Social Interaction in MOOCs (EMNLP).*

78. S. Bhat, H. Xue and S. Yoon. (2014). Shallow Analysis Based Assessment of Syntactic Complexity for Automated Speech Scoring. *Proceedings of the Annual meeting of the Association of Computational Linguistics (ACL).* (Long paper, 26% acceptance rate)

79. S. Bhat, Statistical Stemming for Kannada. (2013). *Proceedings of the 4th Workshop on South and Southeast Asian Natural Language Processing*, International Joint Conference on Natural Language Processing.

80. S. Bhat, G. Herman. (2013). Student Perceptions of Differences in Visual Communication Mode for an Online Course in Engineering. *Frontiers in Education.*

81. S. Bhat. (2012). Morpheme Segmentation for Kannada Standing on the Shoulder of Giants. *Proceedings of the 3rd Workshop on South and Southeast Asian Natural Language Processing*, COLING.

82. S. Yoon, S. Bhat. (2012). Assessment of ESL Learners' Syntactic Competence Based on Similarity Measures. *Proceedings of the Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL).*

83. S. Yoon, S. Bhat and K. Zechner. (2012). Vocabulary Profile as a Measure of Vocabulary Sophistication. *Proceedings of 7th Workshop on Innovative Use of NLP for Building Educational Applications.*

84. S. Bhat, M. Hasegawa-Johnson and R. Sproat. (2010). Automatic Fluency Assessment by Signal-Level Measurement of Spontaneous Speech. *INTERSPEECH Satellite Workshop on Second Language Studies: Acquisition, Learning, Education and Technology.*

85. S. Bhat, R. Sproat, M. Hasegawa-Johnson and F. Davidson. (2010). Automatic Fluency Assessment Using Thin-Slices of Spontaneous Speech. *Language Technology Research Colloquium.*

86. S. Bhat, R. Sproat. (2009). Knowing the Unseen: Estimating Vocabulary Size over Unseen Samples. *Proceedings of the 47th Annual Meeting of the Association of Computational Linguistics (ACL)*, page 109–117.

87. S. Bhat, K. Church. (2008). Variable Selection for Ad Prediction. *Proceedings of the 2nd International Workshop on Data Mining and Audience intelligence For Advertising.*

88. B. Beamer, S. Bhat, B. Chee, A. Fister, A. Rozovskaya and R. Girju. (2007). UIUC: A Knowledge-rich Approach to Identifying Semantic Relations between Nominals. *Proceedings of the Semantic Evaluation Workshop (SemEval 2007).*

89. S. Bhat and R. Sproat. (2007). Consistent Estimation of the Number of Unseen Elements. *Proceedings of the Fourth Midwest Computational Linguistics Colloquium.*

Conference Presentations

90. N. L. Schroeder, J. R. Gladstone, S. Bhat, and S. Zhang. (2026). Work-In-Progress: Educational Benchmarks for Large Language Models in Computer Science Education. Proceedings of the 16th IEEE Integrated STEM Education Conference.

91. S. Zhang, T. V. Earle-Randell, P. Ganapathy Prasad, Z. Liu, Y. Shi, S. Bhat, M. Israel, and A. F. Botelho. (2026). Examining Students' Code Comprehension with LLMs in Block-and Text-Based Programming. In Proceedings of the 57th ACM Technical Symposium on Computer Science Education V. 2 (pp. 1601-1602).

92. S. Zhang, P. Ganapathy Prasad, T. V. Earle-Randell, Y. Shi, S. Bhat, M. Israel. (2026, February). Investigating High School Students' Code Comprehension and Strategy Use Across Block-Based and Text-Based Programming. In Proceedings of the 57th ACM Technical Symposium on Computer Science Education V. 2 (pp. 1603-1604).

93. S.Bhat, B. Ganesh, M. Krishnan, A. Manian, and P. Parthasarathy. (2024). Bridging Data Literacy and Social Good: A Framework for Transformative Education in AI for High School Students. Presentation at the AAAI 2024 Symposium on Increasing Diversity in AI Education and Research, Stanford, CA.

94. Bhavya, Y. Zhou, S. Sehgal, S. Bhat, and C. Zhai. (2024). Analego: Let's build analogies together! Demo presentation at the AAAI 2024 Workshop on AI for Education (AI4Ed). 2024.

95. W. Bond, S. Bhat, J. Zhou, V. Thakkar, R. Ebert-Allen, P. Bhagchandani, R. Ruger, Y. S. Park, and R. Yudkowsky. Speedy Feedback: Automated Grading of Post Simulation Written Chart Notes. Demonstration in the New Innovation category of the SimVentors Showcase at the International Meeting for Simulation in Healthcare 2023, Orlando, FL. [Best Approach Award #1 of 52]

96. S. Bhat, W. F. Bond, V. Thakkar, R. Ebert-Allen, P. Bhagchandani, and T. Sakakini (2022). Automated Assessment of Written Patient Notes for Reli-

able Medical Learner Feedback. Presentation at the International Meeting for Simulation in Healthcare. Los Angeles, CA.

97. W. F. Bond, S. Bhat, V. Thakkar, T. Sakakini, P. Bhagchandani, and R. Ebert-Allen (2021). Improving Feedback and Efficiency: Automated Grading of Post Simulation Written Chart Notes. Oral Presentation. University of Illinois Urbana-Champaign, Grainger College of Engineering, 8th Health Care Engineering Systems Symposium.

98. T. Sakakini, R. F.L. Azevedo, J. Y. Lee, A. Duri, V. Sadauskas, K. Gu, D. Morrow, S. Bhat, J. Graumlich, M. Hasegawa-Johnson, T. S. Huang, A. Willemsen-Dunlap, and D. J. Halpin. (2019). An Online Study on the Evaluation of Dr. Babel Fish: an Unsupervised Machine Translation System for Simplifying Providers Language. *Proceedings of the Tech Mind Society '19 conference.*

99. V. Jay, G. M. Henricks, C. J. Anderson, L. Angrave, N. Bosch, D. Williams, N. Shaik, S. Bhat, and M. Perry. (2019). Patterns and Outcomes of Online Help-seeking Behavior for Students Underrepresented in STEM. Presented at the annual meeting of the American Educational Research Association, San Francisco.

100. H. Gong, A. Valido, K. Ingram, S. Bhat, G. Fanti and D. Espelage. (2019). The Subtleties of Profanity in Abusive Language: Emergent Words and Meanings. Paper presented at the Fourth Workshop on Computational Methods in Online Misbehavior (Cybersafety 2019).

101. G. M. Henricks, M. Perry, and S. Bhat. (2019). Gendered Discourse in an Online STEM Course. Paper presented at the annual meeting of the American Educational Research Association, Toronto.

102. A. Valido, K. Ingram, D. L. Espelage, S. Bhat, G. Fanti. (2018). Understanding Prosocial Bystander Behavior in Cyberbullying. Poster presentation at the American Psychological Association Annual Meeting, San Francisco, CA.

103. G. M. Henricks, R. W. Crues, S. Bhat, and M. Perry. (2018). Predicting Learning by Using Students Perceptions of and Experiences with Statistics Online Course Videos. Poster presented at the annual meeting of the American Educational Research Association, New York, NY.

104. G. Henricks, R. W. Crues, S. Bhat, and M. Perry (2018). How Instructional Videos Affect Perceptions and Outcomes. Poster presented at the annual meeting of the American Educational Research Association.

105. R. W. Crues, G. M. Henricks, M. Perry, S. Bhat, C. J. Anderson, N. Shaik, and L. Angrave, (2018). How do Gender, Learning Goals, and Forum Participation Predict Persistence in a Computer Science MOOC? Paper presented at the 49th ACM Technical Symposium on Computer Science Education, Baltimore, MD.

106. R. W. Crues, G. M. Henricks, M. Perry, S. Bhat, C. J. Anderson, N. Shaik, and L. Angrave (2017). Using Closed- and Free-form Survey Responses in Parametric Models. Poster presented at the 10th International Conference on Educational Data Mining, Wuhan, China.

107. G. M. Henricks, S. Bhat, and M. Perry. (2016). Toward Impacting Learning Experiences in MOOCs. Poster presented at American Psychological Society Annual Convention, Chicago.

Patents

1. Computer-Implemented Systems and Methods for Scoring of Spoken Responses Based on Part of Speech Patterns (S. Yoon and S. Bhat), July 18, 2013, US Patent 20130185057 A1.

2. Text style transfer using reinforcement learning L. Wu, J. Xiong, H. Gong, S. Bhat, and W. Hwu, US Patent 11,314,950, 2022/4/26.

3. Abstract Meaning Representation Parsing with Graph Translation, L. Wu, J. Xiong, H. Gong, S. Bhat, and W. Hwu US Patent 11704486, 2023/7/18