SARAH M. PREUM

Email: spreum@dartmouth.edu
Website: https://persist-lab.github.io/portfolio/ || Google Scholar

SUMMARY

I am a researcher in **Natural Language Processing, Human–AI Interaction, and Computational Health**, advancing socio-technical AI that reduces cognitive burden and promotes equity in communication-rich, high-stakes domains. I have authored **55 peer-reviewed papers** (31 as the lead PI/first author) in premier Computer Science venues such as ACL, EMNLP, AAAI, CHI, IMWUT, ICWSM, and ACM Computing Surveys (impact factor: 23.8). Recognized as a **Google Research Scholar** (2025) and **EECS Rising Star** (2020) with multiple **Best Paper nominations**, I also serve as an **Associate Editor for ACM Transactions on Computing for Healthcare**. I have organized international workshops on **reliability of large language models (LLMs) and AI in healthcare**. I mentor five PhD students (two graduating by 2026) and have guided undergraduates toward placements at Google, ETH Zurich, and Harvard. My lab develops domain-informed NLP methods, systems, and metrics, and leverages large-scale multimodal behavioral and clinical data, and interdisciplinary collaboration to build and evaluate adaptive, uncertainty-aware systems for **patient–provider messaging, online health communities,** and **digital health research**. My work emphasizes **reproducibility**, resulting in **14 public datasets** and open-source resources. My research, supported by **NIH, Google, National Center for Advancing Translational Sciences, NSF, and institutional grants** totaling over $17 M (>$1.5 M to my lab), engages diverse stakeholders across the U.S. and Global South to **advance equitable communication and well-being through socio-technical AI systems**.

EMPLOYMENT

July 2021-Present	**Assistant Professor** Department of Computer Science, Dartmouth College **Technical Associate Director,** Dartmouth Center for Precision Health and AI **Adjunct Assistant Professor,** Department of Biomedical Data Science, Geisel School of Medicine, Dartmouth College **Faculty Affiliate**, Center for Technology and Behavioral Health at Dartmouth
July 2025-Present	**Visiting Research Scholar** HCI Institute, School of Computer Science, Carnegie Mellon University
July 2020-June 2021	**Post-Doctoral Fellow** HCI Institute, School of Computer Science, Carnegie Mellon University

EDUCATION

August 2020

Ph.D. in Computer Science
University of Virginia, Charlottesville, VA
Advisor: Professor John Stankovic
Thesis: *Information Extraction and Fusion for Improving Health Safety*
Awards: Commonwealth fellowship, Big Data fellowship

M.Sc. in Computer Science, CGPA: 3.97/4.00 May 2015
University of Virginia, Charlottesville, VA
Thesis: *Personalized model for temporal human behavior prediction*
Awards: Adobe Research scholarship

B.Sc. in Computer Science & Engineering, CGPA: 3.95/4.00 Feb 2013
[*Double Major Equivalent in Computer Science & Computer Engineering*]
Bangladesh University of Engineering & Technology, Bangladesh
Summa Cum Laude (**2nd** in a class of **141** students)
Awards: Best undergraduate thesis, Google student grant, Dean's list and University merit scholarship in each term

GRANTS & FUNDING

1. **PI, Google Research Scholar Award**, Title: *LLMs for patient-provider communication in underserved communities.* Total funding amount: $60K, Period: 2025-2026.
2. **PI, NIH R21**, Title: *Characterizing information needs and peer engagement regarding medication for opioid use disorder on social media* (DA059665). Total funding amount: $451K, Period: 2024-2026.
3. **Co-I**, **NIH National Center for Advancing Translational Sciences (NCATS) Program**, *Clinical and Translational Science Institute grant* (1UM1TR004772). Total funding amount: $12.47M, Period: 2025-2031.
4. **Co-PI**, **Dartmouth Evergreen Project**, Title: *Evergreen AI: A Digital Ecosystem for Student Wellness at Dartmouth*. Total funding amount to-date: $3.52M, Period: 2025-2029.
5. **PI**, **Hitchcock Foundation Grant**, Title: *Predicting respiratory symptom severity in patient portal messages using advanced natural language processing*. Total funding amount: $50K, Period: 2024-2025.
6. PI, Burke Award, Dartmouth College, Total funding amount: $30K, Period: 2022-2027.
7. PI, OpenAI research access program award, for conducting human-centric NLP research on health-related communication using OpenAI models, Total funding amount: $10K. Period: 2024-2025.

8. PI, CTBH Core Pilot. Title: *Characterize Self-Reported Treatment Information Needs Regarding Maintenance Medications for Opioid Use Disorder on Social Medi*a, Total funding amount: $19K, Period: 2022-2023.
9. PI, Google Cloud Platform research credits, for conducting research, $5K. Period: 2022-2023.
10. Co-PI, NSF SaTC Collaborative Supplement: Security and Privacy in the Lifecycle of IoT for Consumer Environments. Total funding amount: $198K, Period: 2024-2025.
11. Co-I, NIDA CTBH Pilot. Title: *A Digital Cannabis Regulatory Science Paradigm for Testing the Cognitive and Behavioral Impacts of Online Cannabis Retailer Practices*, Total funding amount: $19K, Period: 2023-2024.
12. Co-PI, University of Toronto Data Science Institute Catalyst Grant. Title: *Toward Reducing Harmful Content on social media with Pluralistic Justifications through Community-powered AI.* Total funding amount: $100K, Period: 2022-2023.

SELECTED HONORS & AWARDS

- **Google Research Scholar** (early-career faculty award), 2025
- **Best Paper Nomination,** Association for Computational Linguistics (ACL), 2025
- **Junior Faculty Fellowship**, Dartmouth College, 2025
- Finalist, The Dartmouth Innovation Accelerator for Digital Health, 2024.
- **Best Paper Nomination**, IEEE International Conference on Smart Computing, 2022
- **EECS Rising Stars, 2020**
- NSF Smart and Connected Health student award, 2017
- **Commonwealth Fellowship**, Department of Computer Science, UVA, 2013-2014
- Adobe Research Scholarship as a Grace Hopper Scholar, 2014
- **Big Data Fellowship** by Data Science Institute, UVA, 2014
- Computing Research Association women grad cohort award in 2014, 2016
- **Best Undergraduate Thesis Award**, Department of Computer Science & Engineering, Bangladesh University of Engineering & Technology, 2013
- **Merit Scholarship,** Bangladesh University of Engineering & Technology, 2009-2012
- **Dean's List Award,** Bangladesh University of Engineering & Technology, 2009-2012
- **1st place**, National Software Project Contest, Bangladesh, 2011
- **Conference Travel Grants:**
  - **AAAI** student travel grant to present paper at AAAI 2020
  - **NSF** student travel grant to present paper at PerCom 2017
  - **SIGIR** conference grant to present paper at CIKM 2015
  - **Google student grant** to present paper at ICDE 2013

PEER-REVIEWED CONFERENCE & JOURNAL PUBLICATIONS†

Total peer-reviewed publications: 55 (Journal: 13, Conference: 31), Link to Google Scholar

† In **computer science**, premier conferences are the primary venues for high-impact research, often with competitive acceptance rates comparable to or exceeding those of journals.

1. CS ranking listed top conference, Lead PI, **Nominated for best paper** 
   Joseph Gatto, Omar Sharif, Parker Seegmiller, Sarah M. Preum. *Large Language Models for Document-Level Event-Argument Data Augmentation for Challenging Role Types*. The 63rd Annual Meeting of the Association for Computational Linguistics (**ACL**) 2025.
2. CS ranking listed top conference, Lead PI
   Joseph Gatto, Parker Seegmiller, Timothy Burdick, Inas S. Khayal, Sarah DeLozier, Sarah M. Preum. *Follow-up Question Generation For Enhanced Patient-Provider Conversations*. The 63rd Annual Meeting of the Association for Computational Linguistics (**ACL**) 2025.
3. CS ranking listed top conference, Lead PI
   Omar Sharif, Joseph Gatto, Madhusudan Basak, Sarah M. Preum. *REGen: A Reliable Evaluation Framework for Generative Event Argument Extraction*. Proceedings of the International Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2025.
4. CS ranking listed top conference, Lead PI
   Omar Sharif, Joseph Gatto, Madhusudan Basak, Sarah M. Preum. *Explicit, Implicit, and Scattered: Revisiting Event Extraction to Capture Complex Arguments.* Proceedings of the International Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2024.
5. Journal with impact factor 7.4, Lead PI
   Madhusudan Basak, Omar Sharif, Sarah E. Lord, Jacob T. Borodovsky, Lisa A. Marsch, Sandra A. Springer, Edward Nunes, Charlie D. Brackett, Luke J. Archibald, Sarah M. Preum. *Information Needs for Opioid Use Disorder Treatment Using Buprenorphine Product: Qualitative Analysis of Suboxone-Focused Reddit Data*. Journal of Medical Internet Research (JMIR), 2025.
6. Journal with impact factor 7.4
   Bushra Hossain, Sarah M. Preum, Md Fazle Rabbi, Rifat Ara, Mohammed Eunus Ali. *Subreddit to Symptomatology: A Lexicon-based Approach to Extract Symptoms of Complex Conditions from Online Discourse*. Journal of Medical Internet Research (JMIR), 2025.
7. CS ranking listed top conference, Lead PI
   Eftekhar Hossain, Omar Sharif, Mohammed Moshiul Hoque, Sarah M. Preum. Deciphering Hate: Identifying Hateful Memes and Their Targets. The 62nd Annual Meeting of the Association for Computational Linguistics (**ACL),** 2024
8. CS ranking listed top conference, Lead PI
   Omar Sharif, Madhusudan Basak, Tanzia Parvin, Ava Scharfstein, Alphonso Bradham, Jacob Borodovsky, Sarah Lord, Sarah M. Preum. *Characterizing Information Seeking Events in Health-Related Social Discourse*. The 38th Annual AAAI Conference on Artificial Intelligence (**AAAI**), 2024.
9. CS ranking listed top conference
   Subigya Nepal, Arvind Pillai, William Campbell, Talie Massachi, Michael V. Heinz, Ashmita Kunwar, Eunsol Soul Choi et al. *MindScape Study: Integrating LLM and Behavioral Sensing for Personalized AI-Driven Journaling Experiences.* Proc. ACM Interact. Mob. Wearable Ubiquitous Technology (**IMWUT**). 2025
10. Joseph Gatto, Parker Seegmiller, Timothy Burdick, Sarah M. Preum. *In-Context Learning for Preserving Patient Privacy: A Framework for Synthesizing Realistic Patient Portal Messages*. Machine Learning for Healthcare (ML4H) findings 2024.

11. Core A Conference, Lead PI
    Will Romano, Omar Sharif, Madhusudan Basak, Joseph Gatto, and Sarah M. Preum. *Theme-driven Keyphrase Extraction to Analyze Social Media Discourse*. Proceedings of the AAAI Conference on Web and Social Media (**ICWSM**). 2024
12. CS ranking listed top conference
    Nur Yildirim, Susanna Zlotnikov, Andrew J. King, Lu Tang, Billie S. Davis, Jeremy M. Kahn, Sarah M. Preum, Leigh A. Bukowski, John Zimmerman. *Sketching AI Concepts with Capabilities and Examples: AI Innovation in the Intensive Care Unit.* Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (**CHI**), 2024.
13. Seegmiller, Parker, Joseph Gatto, Omar Sharif, Madhusudan Basak, and Sarah Masud Preum. *Do LLMs Find Human Answers To Fact-Driven Questions Perplexing? A Case Study on Reddit.* (2024). 1st Workshop on Reliable Evaluation of Large Language Models for Factual Information (REALInfo-2024), co-located with AAAI ICWSM'24.
14. Eftekhar Hossain, Omar Sharif, Mohammed Moshiul Hoque, Sarah M. Preum. *Align before Attend: Aligning Visual and Textual Features for Multimodal Hateful Content Detection*. European Association for Computational Linguistics Student Research Workshop, 2024.
15. Matthew J. Davis, Gokul Srinivasan, Rachael Chacko, Sophie Chen, Anish Suvarna, Louis J. Vaickus, Veronica C. Torres et al. *A deep learning algorithm to detect cutaneous squamous cell carcinoma on frozen sections in Mohs micrographic surgery: A retrospective assessment*. Experimental Dermatology 33, no. 1 (2024).
16. CS ranking listed top conference
    Arvind Pillai, Subigya Kumar Nepal, Weichen Wang, Matthew Nemesure, Michael Heinz, George Price, Damien Lekkas, Amanda C. Collins, Tess Griffin, Benjamin Buck, Sarah M. Preum, Trevor Cohen, Nicholas C. Jacobson, Dror Ben-Zeev, and Andrew Campbell. 2024. *Investigating Generalizability of Speech-based Suicidal Ideation Detection Using Mobile Phones*. Proc. ACM Interact. Mob. Wearable Ubiquitous Technology (**IMWUT**), 2024.
17. CS ranking listed top conference, Lead PI
    Parker Seegmiller, Sarah M. Preum. *Statistical Depth for Ranking and Characterizing Transformer-Based Text Embeddings.* Proceedings of the International Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2023.
18. CS ranking listed top conference, Lead PI
    Joseph Gatto, Omar Sharif, Sarah M. Preum. *Chain-of-Thought Embeddings for Stance Detection on Social Media.* Proceedings of the International Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2023.
19. Joseph Gatto, Omar Sharif, Parker Seegmiller, Philip Bohlman, Sarah M. Preum. *Text Encoders Lack Knowledge: Leveraging Generative LLMs for Domain-Specific Semantic Textual Similarity*. Generation, Evaluation & Metrics (GEM) Workshop collocated with Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2023.
20. Core A Conference, Lead PI
    Joseph Gatto, Madhusudan Basak, Sarah M. Preum. *Scope of Pre-trained Language Models for Detecting Conflicting Health Information*. International AAAI Conference on Web and Social Media (**ICWSM**), 2023.
21. Core A Conference, Lead PI

Joseph Gatto, Parker Seegmiller, Garrett Johnston, Madhusudan Basak, Sarah M. Preum. *HealthE: Recognizing Health Advice & Entities in Online Health Communities*. International AAAI Conference on Web and Social Media (**ICWSM**), 2023.

22. Parker Seegmiller, Joseph Gatto, Madhusudan Basak, Diane Cook, Hassan Ghasemzadeh, John Stankovic and Sarah M. Preum. The *Scope of In-Context Learning for the Extraction of Medical Temporal Constraints.* HealthNLP 2023 IEEE 11th International Conference on Healthcare Informatics.
23. Journal with impact factor 7.4, Lead PI
    Joseph Gatto, Parker Seegmiller, Garrett Johnston, Sarah M. Preum. *Identifying Severe Medical Queries from An Online Medical Hotline: A Transfer Learning Based Solution*. Journal of Medical Internet Research (**JMIR**), 2022.
24. M. Arif Rahman, Leon Jia, Eimara Mirza, Sarah M. Preum, Homa Alemzadeh, Ronald D. Williams, and John A. Stankovic. *EMSReACT: A Real-Time Interactive Cognitive Assistant for Cardiac Arrest Training in Emergency Medical Services*. In 2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT), pp. 120-128. IEEE, 2023.
25. J. Greenburg, Y. Lu, S. Lu, U. Kamau, R. Hamilton, J. Pettus, Sarah M. Preum, L. Vaickus, and J. Levy. (2023). *Development of an interactive web dashboard to facilitate the reexamination of pathology reports for instances of underbilling of CPT codes.* Journal of Pathology Informatics, 202.
26. Andrew J. King, Lu Tang, Billie S. Davis, Sarah M. Preum, Leigh A. Bukowski, John Zimmerman, and Jeremy M. Kahn. "*Machine learning-based prediction of low-value care for hospitalized patients*." Journal of Intelligence-based medicine 8 (2023): 100115.
27. M Arif Rahman, Sarah M. Preum, Homa Alemzadeh, Ronald D. Williams, John Stankovic. *EMS-BERT: A Pre-Trained Language Representation Model for the Emergency Medical Services (EMS) Domain*. IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE) 2023.
28. Arya Kadakia, Sarah Masud Preum, Andrew R. Bohm, Karen L. Fortuna. *Investigating the Fidelity of Digital Peer Support: A Preliminary Approach Using Natural Language Processing to Scale High-Fidelity Digital Peer Support*. Workshop on Best Practices for Scaling-Up Digital Innovations in Healthcare – Scale-IT-up 2023, 16th International Joint Conference on Biomedical Engineering Systems and Technologies.
29. **Nominated for best paper**
    Zirong Chen, Isaac Li, Haoxiang Zhang, Sarah Preum, John A. Stankovic, and Meiyi Ma. *CitySpec: An Intelligent Assistant System for Requirement Specification in Smart Cities*. International Conference on Smart Computing, 2022.
30. Journal with impact factor 23.8, First author
    Sarah M. Preum, Sirajum Munir, Meiyi Ma, Mohammad Yasar, David J. Stone, Ronald D. Williams, Homa Alemzadeh, and John Stankovic. *A Review of Cognitive Assistants for Healthcare: Trends, Prospects, and Future Directions*. **ACM Computing Surveys, 2021**
31. CS ranking listed top conference, Lead PI
    Nur Yildirim,1 John Zimmerman, Sarah M. Preum. *Technical Feasibility, Financial Viability, and Clinician Acceptance: On the many challenges to AI in Clinical Practice*. Human partnership with Medical Artificial Intelligence, AAAI Fall 2021 Symposium.

32. John Stankovic, Meiyi Ma, Sarah M. Preum, and Homa Alemzadeh. *Challenges and Directions for Ambient Intelligence: A Cyber Physical Systems Perspective*. The Third IEEE International Conference on Cognitive Machine Intelligence (IEEE CogMI), 2021.
33. Journal with impact factor 7.4
    Brooke M. Bell, Ridwan Alam, Abu Mondol, Meiyi Ma, Ifat Emi, Sarah M. Preum, Kayla Haye, John A. Stankovic, John Lach, Donna Spruijt-Metz. *Validity and Feasibility of the Monitoring and Modeling Family Eating Dynamics (M2FED) System to Automatically Detect In-Field Family Eating Behavior*. Journal of Medical Internet Research (JMIR), 2021
34. CS ranking listed top conference, First author
    Sarah M. Preum, Homa Alemzadeh, and John Stankovic. *EMSContExt: EMS Protocol-driven Concept Extraction for Cognitive Assistance in Emergency Response*. The Thirty-Second Annual Conference on Innovative Applications of Artificial Intelligence (AAAI /IAAI), 2020
35. CS ranking listed top conference
    M Arif Rahman, Sarah M. Preum, Homa Alemzadeh, Ronald D. Williams, and John Stankovic. *GRACE: Generating Summary Reports Automatically for Cognitive Assistance in Emergency Response.* The Thirty-Second Annual Conference on Innovative Applications of Artificial Intelligence (AAAI /IAAI), 2020
36. Sarah M. Preum, Sile Shu, Mustafa Hotaki, Ron Williams, John Stankovic, and Homa Alemzadeh. *CognitiveEMS: A Cognitive Assistant System for Emergency Medical Services*. ACM Special Interest Group on Embedded Systems Review, 2019
37. CS ranking listed top conference
    Sile Shu, Sarah M. Preum, Haydon M. Pitchford, Ronald D. Williams, John Stankovic, and Homa Alemzadeh. *A Behavior Tree Cognitive Assistant System for Emergency Medical Services.* IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019.
38. Meiyi Ma, Sarah M. Preum, Mohsin Ahmed, William Turneberg, Abdeltawab Hendawi, John Stankovic, *Data Sets, Modeling and Decision Making in Smart Cities: A Survey*, Transactions on Cyber Physical Systems (TCPS), 2019
39. Sarah M. Preum, Rizwan Parvez, Kai-Wei Chang, John Stankovic, *A Corpus of Drug Usage Guidelines Annotated with Type of Advice*, International Conference on Language Resources and Evaluation (LREC), 2018
40. Core A* Conference, First author
    Sarah M. Preum, Md A. Mondol, Meiyi Ma, Hongning Wang, John Stankovic, *PreCluDe: Conflict Detection in Textual Health Advice*, IEEE International Conference on Pervasive Computing and Communications (**PerCom**), 2017
41. Sarah M. Preum, Md A. Mondol, Meiyi Ma, Hongning Wang, John Stankovic, *PreCluDe2: Personalized Conflict Detection in Heterogeneous Health Applications*, Pervasive and Mobile Computing Journal, Volume 42, 226-247, 2017
42. Meiyi Ma, Sarah M. Preum, John Stankovic, *CityGuard: A Watchdog for Safety-Aware Conflict Detection in Smart Cities*, ACM/IEEE International Conference on Internet-of-things Design and Implementation (IoTDI), 2017
43. Meiyi Ma, Sarah M. Preum, William Tarneberg, Mohsin Ahmed, John Stankovic, *Detection of Runtime Conflicts among Services in Smart Cities*, IEEE Smart Computing, 2016
44. Core A Conference, First author
    Sarah M. Preum, John Stankovic, Yanjun Qi, *MAPer: A Multi-Scale Adaptive Personalized Model for Temporal Human Behavior Prediction*, ACM International Conference on Information and Knowledge Management (**CIKM**), 2015

45. Enamul Hoque, Robert Dickerson, Sarah M. Preum, Mark Hanson, Adam Barth, John Stankovic, *Holmes: A Comprehensive Anomaly Detection System for Daily In-home Activities*, International Conference on Distributed Computing in Sensor Systems (DCOSS), 2015
46. Farhana Choudhury, Mohammed Ali, Sarah M. Preum, Suman Nath, Ishat Rabban, *Scalable Visibility Color Map Construction in Spatial Databases,* Information Systems Journal, Volume 42, 89-106, 2014
47. Sarah M. Preum, Monjura Afrin, Syed I. Ahmed, *VizResearch: Linking the Knowledge of People and the People with Knowledge*. Procedia Computer Science, Volume 9, 1416-1425, 2012
48. CS ranking listed top conference, First author
    Sarah M. Preum, Farhana Choudhury, Mohammed Ali, Sarana Nutanong, *Maximum Visibility Queries in Spatial Databases*, International Conference on Data Engineering **(ICDE)**, 2013

PEER-REVIEWED ABSTRACTS

49. Brooke M. Bell, Donna Spruijt-Metz, Ridwan Alam, Abu Mondol, Meiyi Ma, Ifat Emi, Sarah M. Preum, John A. Stankovic, John Lach, Kayla Haye. *Utilizing wearable sensors and ecological momentary assessment to automatically detect in-home contextualized eating events*. 43rd Annual Meeting of the Society of Behavioral Medicine, 2022.
50. M. Arif Rahman, Sarah M. Preum, John A. Stankovic, Leon Jia, Eimara Mirza, Ronald Williams, and Homa Alemzadeh. "IMACS-an interactive cognitive assistant module for cardiac arrest cases in emergency medical service: demo abstract." In Proceedings of the 18th Conference on Embedded Networked Sensor Systems, pp. 621-622. 2020.
51. Elizabeth Shelton, Annie Chapman, Sarah M. Preum, Sile Shu, Arif Rahman, John Stankovic, Homa Alemzadeh, Poster abstract: *An Assessment of a Cognitive Assistant System Under Diverse Noise Profiles in the Context of Emergency Medical Services*, SRC TECHCON, 2019
52. Sarah M. Preum, Sile Shu, Jonathan Ting, Vincent Lin, Ron Williams, John Stankovic, Homa Alemzadeh, Poster abstract: *Towards a cognitive assistant system for emergency response*, ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS), 2018
53. Sarah M. Preum, Md A. Mondol, Meiyi Ma, Hongning Wang, John Stankovic, Demo Abstract: *Conflict Detection in Online Textual Health Advice,* International Conference on Information Processing in Sensor Networks (IPSN), 2017
54. Meiyi Ma, Sarah M. Preum, John Stankovic, Demo Abstract: *Simulating Conflict Detection in Heterogeneous Services of a Smart City,* IEEE/ACM Second International Conference on Internet-of-Things Design and Implementation (IoTDI), 2017
55. Md A. Mondol, Ifat Emi, Sarah M. Preum, John Stankovic. Poster abstract: *User authentication using wrist mounted inertial sensors*. ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), 2017.

WORK IN PROGRESS (As Lead PI)

56. Omar Sharif, Charlie D. Brackett, Luke J. Archibald, Sarah M. Preum. *Socially Constructed Treatment Plans: Analyzing Online Peer Interactions to Understand How Patients Navigate Complex Medical Conditions*. Under review.

57. Madhusudan Basak, Dipto Das, Ishtiaque Ahmed, Sarah M. Preum. *Contextualizing Information Needs in Online Health Communities: The Case of Recovery from Opioid Use Disorder through the Lens of Epistemic Injustice*. Under review.
58. Parker Seegmiller, Joseph Gatto, Sarah M. Preum. *LENS: Measuring the Effects of Natural Shifts in User Prompt Distribution on Large Language Model Performance*. Under review.
59. Joseph Gatto, Madhusudan Basak, Yash Srivastava, Phillip Bohlman, Sarah M. Preum. *Scope of Large Language Models for Mining Emerging Opinions in Online Health Discourse*. Arxiv
60. Joseph Gatto, Sarah M. Preum. *Not Enough Labeled Data? Just Add Semantics: A Data-Efficient Method for Inferring Online Health Texts.* arXiv preprint arXiv:2309.09877 (2023).

TEACHING EXPERIENCE

Graduate Level Course: [Rated between 1-5, 1 being the highest rating]

CS 274: Introduction to Machine Learning and Statistical Data Analysis

- Fall 2023: 57 students. Rating Median: 1, Mean 1.2
- Spring 2022: 6 students. Rating Median: 2, Mean 2.3

CS 189/89.35: Human-centric Large Language Models

- Spring 2025: 12 students. Rating Median: 1, Mean 1

CS 189.28: Computational Healthcare

- Spring 2025: 9 students. Rating Median: 1, Mean 1.3
- Spring 2024: 12 students. Rating Median: 1, Mean 1.9
- Winter 2023: 12 students. Rating Median: 1, Mean 1
- Fall 2021: 23 students. Rating Median: 1, Mean 1.5

Undergraduate Level Course: [Rated between 1-5: 1 being the highest rating]

CS 72/ Ling 48: Accelerated Computational Linguistics

- Winter 2025: 27 students. Rating Median: 1, Mean 2.1

CS 74: Introduction to Machine Learning and Statistical Data Analysis

- Fall 2023: 86 students. Rating Median: 1, Mean 1.7
- Spring 2022: 30 students. Rating Median: 2, Mean 2.3

Some feedback from students on my teaching at Dartmouth:

- *Professor Preum did a fantastic job of explaining core concepts clearly and communicating passion and depth in her field of study. I loved her lectures, especially the ones pertaining to her research and the landscape of AI in general.*
- *Sarah Preum is very friendly and kind. She would always be ready to explain things even for the tenth time if needed. I enjoyed talking to her outside class as well.*

- *Incredible lecturer, patient with students, great answers to questions…She explained central concepts clearly. Every time someone asks a question, she will first confirm it as a good question, which will encourage students to continue to ask.*
- *Professor Preum reviewed the material from the previous lecture at the beginning of class which was very helpful in clarifying any confusion. There were also examples of real-world applications which were interesting. Professor Preum also engaged with students to clarify if there were any questions.*
- *Great communication of concepts, good incorporation of class participation. The Kaggle competitions were wonderful! The homework felt like a good demonstration of the material.*
- *I loved the focus on relevant research and reading papers. Until this class I didn't realize that I could even understand research papers in CS, much less contribute to them and comment on them.*
- *The class had discussions sessions apart from the regular lecture hours where the professor takes time to clarify the student doubts related to any topic that was taught so far.*
- *I loved the Kaggle competitions. They were playful and creative, and yet required rigorous application of feature engineering and understanding of the datasets. In terms of real-world experience, they helped me apply theoretical knowledge from the class to a real-world problem.*
- *I feel ready to learn more about machine learning in future classes and on my own, and I am excited about the prospect of applying these new-found techniques towards all different kinds of projects!*
- *It reinforced a lot of my machine learning experience and encouraged me to continue with my computer science major.*
- *The lectures were very effective in providing the conceptual understanding of machine learning concepts. The project and homework provided the opportunity to reinforce the ideas through hands-on applications.*
- *Professor Preum provided exceptional guidance and support throughout this term in CS 74. Her passion for the subject matter and dedication to teaching have made a lasting impact on my academic experience. Taking CS 74 was truly an enriching experience, and I am grateful that I got to take the course under her.*
- *Professor Preum was wonderful -- engaging, friendly, easy to talk to, and someone I admire.*
- *I want to thank Professor Preum for such an awesome term and for making Machine Learning a class that I would literally look forward to!*
- *Dr. Preum was just on her game! She engaged with the class and knew our names! She is warm and very helpful in answering our questions.*

SELECTED MENTORING EXPERIENCE

PhD Students

- Joseph Gatto, Computer Science, Dartmouth College

- Matthew Seegmiller, Computer Science, Dartmouth College
- Omar Sharif, Computer Science, Dartmouth College
- Roshnik Rahat, Computer Science, Dartmouth College
- Bushra Hossain, Computer Science and Engineering, Bangladesh University of Engineering and Technology

Masters Students

- Reet Kothari, Computer Science, Dartmouth College
- Nikil Kumar, Computer Science, Dartmouth College
- Phil Bohlman, Computer Science, Dartmouth College
- Madhusudan Basak, Computer Science, Dartmouth College
- Yash Srivastava, Computer Science, Dartmouth College
- Yunrui Lu, Quantitative Biomedical Science, Dartmouth College

Undergraduate Students

- Vasavi Garimella, Dartmouth College, Senior Thesis, *now MD Candidate at University of Washington Medical School*
- Alphonso Bradham, Dartmouth College, *now at Google*
- Ava Scharfstein, Dartmouth College, Women in Science program, *now at Johns Hopkins*
- Garrett Johnston, Dartmouth College, Senior Thesis with Honors, *now at Google*
- Gabriel Asher, Dartmouth College, Senior Thesis with Honors, *now Researcher at Matterworks*
- Natalie Ashgriz, University of Toronto, *now at ETH Zurich*
- Zhanel Nugmanova, Dartmouth College, Undergraduate Research Scholar, *now Co-founder at Valinor Discovery*
- Love Tsai, Dartmouth College, Ada Lovelace Fellow, *now MD Candidate at Harvard Medical School*
- Rabia Saikhu, Dartmouth College, Women in Science program
- Rohan Ray, Dartmouth College
- Yasmeen Hussein, Dartmouth College, Women in Science program

PhD Thesis Committee

- Adam M. Gronewold, Thayer School of Engineering, Dartmouth College, Thesis Title: *Advances Towards a Robotic Management Vehicle Suited to Nurse Row Crops to More Efficient Outcomes*
- Yubo Luo, Department of Computer Science, University of North Carolina at Chapel Hill, Thesis tile: *Scaling up Task Execution on Resource-Constrained Systems*
- Xiaobo Guo, Department of Computer Science, Dartmouth College, Thesis Title: *Leveraging Large Language Models For Enhancing Well-Being in the Digital Age*
- Subigya Nepal, Department of Computer Science, Dartmouth College, Thesis Title: *Toward the Integration of Behavioral Sensing and Artificial Intelligence*

SELECTED INVITED TALKS

September 2025	*Generating Editable Responses for Providers to Enhance Patient Portal Communication*, Invited talk at Teledia lab, Language Technology Institute (LTI), **Carnegie Mellon University**
October 2024	*In-Context Learning for Human-AI Collaboration: Methods and Measures,* Network Science Institute Colloquium, **Northeastern University**
September 2024	*AI at Dartmouth*, panel for the Board of Trustees, Dartmouth College
February 2024	*Characterizing Information Seeking Events in Health-Related Social Discourse*. The 38th Annual AAAI Conference on Artificial Intelligence (AAAI), 2024.
February 2022	*Intelligent Assistants for Improved Health: Teaching Machines to Understand Data,* Biomedical Data Science Grand Rounds, **Geisel School of Medicine, Dartmouth College**
February 2022	*Intelligent Assistants for Improved Health: Teaching Machines to Understand Data,* Masters in Digital Arts Seminar, Department of Computer Science, Dartmouth College
April 2021	Panelist, REsearch Exposure in Socially Relevant Computing program, Syracuse University
March 2021	*Navigating academic research in Computer Science: through the lens of an outlier, Preparation for Research through Immersion, Skills, and Mentorship* (PRISM) seminar, Department of Computer Science, **University of Toronto**
February 2021	WiCS Dinner (Women in Computer Science), Department of Computer Science, Dartmouth College
September 2020	*Future Directions for Clinical Decision Support*, Clinical Decision Support Workshop hosted by the **Office of the National Coordinator for Health Information Technology**
July 2020	*Towards a cognitive assistant for emergency response*, Public Safety Broadband Stakeholder Meeting by **National Institute of Standards and Technology**
April 2020	*Information Extraction & Fusion for Improving Personal & Public Health Safety,* Senseable City Lab, **Massachusetts Institute of Technology** (MIT), Cambridge, MA
March 2020	*Information Extraction & Fusion for Improving Personal & Public Health Safety,* Human-centered Artificial Intelligence Center, **Stanford University**, Stanford, CA
March 2020	*Information Extraction & Fusion for Improving Personal & Public Health Safety,* Department of Computer Science, **Dartmouth College**, Hanover, NH
March 2020	*Information Extraction & Fusion for Improving Personal & Public Health Safety,* School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ
July 2019	*Towards a cognitive assistant for emergency response*, Public Safety Broadband Stakeholder Meeting by **National Institute of Standards and Technology**
March 2018	*Data-driven, resource-constrained, resilient decision support systems*, Bosch Research and Technology Center, Pittsburgh
February 2018	*Increasing health safety using timeseries analysis and natural language* processing, Department of Psychology, University of Virginia

March 2017 *Conflict Detection from online information using NLP*, Student mentoring session of **NSF Smart and Connected Health** (SCH) PI meeting

SELECTED MEDIA COVERAGE

- Dartmouth's Evergreen AI Puts Student Wellness To The Test, October 2025, Forbes
- AI Helps Streamline Patient Portal Messages, June 2025, Dartmouth News
- Trustees Hail Progress on Housing and Sustainability, September 2024, Dartmouth News
- AI Tools Reveal Knowledge Gaps in Addiction Treatment, February 2024, Dartmouth News
- Symposium Focuses on Digital Tech for Mental Health, September 2023, Dartmouth News

PROFESSIONAL SERVICES

Internal at Dartmouth
- Curriculum Committee, 2024-present
- Colloquium co-chair, 2023-present
- MSCS Program Director, 2023-2024
- Diversity Equity Inclusion committee, 2022-2023
- Faculty search committee, 2022-2023
- PhD Admission committee, 2021-2022

Associate Editor ACM Transactions on Computing for Healthcare (ACM Health), 2022-present

Organizing Chair/Co-chair
- Workshop on Big Data & ML in Healthcare: Emerging Challenges, co-located with IEEE conference on Big Data, 2025
- Workshop on Reliable Evaluation of LLMs for Factual Information (REAL-Info), co-located with AAAI ICWSM, 2024
- ICCPS 2023 Poster and Demo tracks, 2023

Panelist NSF CISE Panels, 2022-present

Technical Program Committee

- AAAI Conference on Artificial Intelligence (AAAI), 2025, 2023
- Senior PC member, AAAI ICWSM 2023-present
- IoTDI, 2023
- ICCPS, 2023
- PerIoT, Co-located with PerCom, 2021
- CIKM, 2022

Reviewer

- ACL Rolling Review, 2023-present
- AAAI, 2021-present
- The New England Journal of Medicine (NEJM) AI, 2025
- IEEE International Engineering in Medicine and Biology Conference, 2022
- Journal of American Medical Informatics Association (JAMIA), 2021
- Language Resources and Evaluation Conference (LREC), 2020
- IEEE Transactions on Knowledge and Data Engineering (TKDE), 2019
- American Medical Informatics Association (AMIA), 2019
- ACM Health Journal, 2021, 2020, 2019