Lu He

Assistant Professor
Joseph J. Zilber College of Public Health
University of Wisconsin – Milwaukee

he32@uwm.edu
luhe@mcw.edu
www.luheholly.com

Academic Appointment

University of Wisconsin – Milwaukee

- Tenure-track Assistant Professor, Joseph J. Zilber College of Public Health
- Participating faculty member, Biomedical and Health Informatics, Department of Computer Science
- 2023 – Present

Medical College of Wisconsin

- Adjunct Assistant Professor, Clinical & Translational Science Institute of Southeast Wisconsin (CTSI)
- 2024 – present

Education

University of California, Irvine

- Ph.D. in Informatics, 2017–2023
- Dissertation: *Computational Analysis of Health Text*
- Committee: Kai Zheng PhD, FACMI (Chair), Yunan Chen PhD, Daniel Epstein, PhD, Helen Ma, MD

University of Minnesota, Twin Cities

- B.S. in Computer Science with Distinction, 2013–2017

Honors and Awards

- LEAD Trainee and Early Career Meeting Scholarship, American Medical Informatics Association Clinical Informatics Conference (AMIA CIC '23), 2023
- Graduate Student Consortium, Natural Language Processing Working Group, American Medical Informatics Association 2022 Annual Symposium (AMIA '22), 2022
- Second Place in Student Paper Competition, American Medical Informatics Association Annual Symposium (AMIA '22), 2022
- Graduate Dean's Dissertation Fellowship, University of California, Irvine, 2022
- Editor's Choice of Journal of American Medical Informatics Association (JAMIA), 2021
- Student Paper Competition Finalist, American Medical Informatics Association 2020 Annual Symposium (AMIA '20), 2020

- Best Poster Award, The 2019 Southern California Natural Language Processing Symposium (SoCal NLP '19), 2019
- Student Best Paper Nomination, The 2019 World Congress on Health and Biomedical Informatics (MedInfo '19), 2019
- CRA-W Grad Cohort Workshop Travel Award, Computing Research Association, 2018
- Graduate Dean Recruitment Fellowship, Department of Informatics, University of California, Irvine, 2017
- Dean's Award, Department of Informatics, University of California, Irvine, 2017
- Second Place in Student Paper Competition, American Medical Informatics Association 2017 Joint Summits, 2017
- Undergraduate Special Recognition, Department of German, Nordic, Slavic and Dutch, University of Minnesota, Twin Cities, 2014
- Maroon Global Excellence Scholarship, University of Minnesota, Twin Cities, 2013–2017
- Dean's List, University of Minnesota, Twin Cities, 2013–2016

Publications

Peer-Reviewed Journal Articles

Underline: Student or trainee I advised or co-advised.

J.15 Omranian S, **He L**, Talsma A, Scoglio A, McRoy S, Rich-Edwards J. The Role of Burnout and Health Beliefs in COVID-19 Vaccine Decisions Among Healthcare Workers: A Large Language Model-Based Text Analysis. JMIR Nursing 2025 (forthcoming).

J.14 He C, Deng Y, **He L**, Guo Q, Zhang Y, Lu Z, Li B. Engage Wider Audience or Facilitate Quality Answers? a Mixed-methods Analysis of Questioning Strategies for Research Sensemaking on a Community Q&A Site. Proceedings of the ACM on Human-Computer Interaction, CSCW. 2024 (Accepted)

J.13 He C, **He L**, Yang W, Li B. Images Connect Us Together: Navigating a COVID-19 Local Outbreak in China Through Social Media Images. Proceedings of the ACM on Human-Computer Interaction, CSCW. 2024 (Accepted)

J.12 Ye J, **He L**, Beestrum M. Implications for implementation and adoption of telehealth in low-and-middle income countries during the COVID-19 pandemic: systematic review of China's practices and experiences. npj Digit. Med. 6, 174 (2023). DOI: 10.1038/s41746-023-00908-6

J.11 Goyal J, Ng DQ, Zhang K, Chan A, Lee J, Zheng K, Hurley-Kim K, Nguyen L, **He L**, Nguyen M, McBane S, Li W, Cadiz CL. Using machine learning to develop a clinical prediction model for SSRI-associated bleeding: a feasibility study. BMC Medical Informatics and Decision Making. 2023. (Accepted)

J.10 He C, **He L**, Lu Z, Li B. "I have to use my son's QR code to run the business": unpacking senior street vendors' challenges in mobile money collection in China. Proceedings of the ACM on Human-Computer Interaction, CSCW. 2023;7. DOI: 10.1145/3579493

J.9 Griffin A, **He L**, Sunjaya A, King A, Khan Z, Douthit B, Nwadiugw M, Subbin V, Braunstein M, Nguyen V, Jaffe C, Schleyer T. Clinical, technical, and implementation characteristics of real-world health applications using FHIR. Journal of American Medical Informatics Association Open. 2022;5. DOI: 10.1093/jamiaopen/ooac077

J.8 **He L**, Yin T, Zheng K. They may not work! An evaluation of eleven sentiment analysis tools on seven social media datasets. Journal of Biomedical Informatics. 2022;132:104142. PMID: 35835437

J.7 **He L**, He C. Help me #DebunkThis: unpacking individual and community's collaborative work in information credibility assessment. Proceedings of the ACM on Human-Computer Interaction, CSCW. 2022;6. DOI: 10.1145/3555138

J.6 Su Z, **He L**, Jariwala SP, Zheng K, Chen Y. "What is your envisioned future?": towards human-AI enrichment in data work of asthma care. Proceedings of the ACM on Human-Computer Interaction, CSCW. 2022;6. DOI: 10.1145/3555157

J.5 He C, Liu H, **He L**, Lu T, Li B. More collaboration, less seriousness: investigating new strategies for promoting youth engagement in government-generated videos during the COVID-19 pandemic in China. Computers in Human Behavior. 2021;126. DOI: 10.1016/j.chb.2021.107019

J.4 He C, **He L**, Lu T, Li B. Beyond entertainment: unpacking Danmaku and comments' role of information sharing and sentiment expression in online crisis videos. Proceedings of the ACM on Human-Computer Interaction, CSCW. 2021;5. DOI: 10.1145/3479555

J.3 **He L***, He C*, Reynolds TL, Bai Q, Huang Y, Li C, Zheng K, Chen Y. Why do people oppose mask wearing? A comprehensive analysis of US tweets during the COVID-19 pandemic. Journal of American Medical Informatics Association. 2021;28(7):1564–73. PMCID: PMC7989302 (* equal contribution) (**Editor's Choice and Featured Article**)

J.2 **He L**, Yin T, Hu Z, Chen Y, Hanauer DA, Zheng K. Developing a standardized protocol for computational sentiment analysis research using health-related social media data. Journal of American Medical Informatics Association. 2021;28(6):1125–34. PMCID: PMC8200276

J.1 Ma H, Smith C.E, **He L**, Narayanan S, Giaquinto R.A, Evans R, Hanson L, Yarosh S. Write for life: persisting in online health communities through expressive writing and social support. Proceedings of the ACM on Human-Computer Interaction, CSCW. 2017;1. DOI: 10.1145/3134708

Peer-Reviewed Full-Length Conference Papers

C.10 **He L**, Omranian S, McRoy S, Zheng K. Using Large Language Models for sentiment analysis of health-related social media data: empirical evaluation and practical tips. Accepted to American Medical Informatics Association Annual Symposium 2024. Forthcoming.

C.9 He C, **He L**, Lu Z, Li B. Seeking love and companionship through streaming: unpacking livestreamer-moderated senior matchmaking in China. In: Proceedings of the 2023 ACM Conference on Human Factors in Computing Systems (CHI '23) DOI: 10.1145/3544548.3581195

C.8 Guo Y, Zhu J, Huang Y, **He L**, He C, Li C, Zheng K. Public opinions toward COVID-19 vaccine mandates: a machine learning-based analysis of U.S. tweets. American Medical Informatics Association Annual Symposium Proceedings. 2022;502–511. PMCID: PMC10148373 (**Student Paper Competition Second Place**)

C.7 **He L***, Song T*, Jiang Y, Yu P, Song L, Gong Y. To improve supportive care for patients taking oral anticancer agents. In: Proceedings of the 2021 World Congress on Health and Biomedical Informatic (MEDINFO '21) 2021;290:547–51. PMID: 35673076 (* equal contribution)

C.6 **He L**, He C, Wang Y, Hu Z, Zheng K, Chen Y. What do patients care about? Mining fine-grained patient concerns from online physician reviews through computer-assisted multi-level qualitative analysis. American Medical Informatics Association Annual Symposium Proceedings. 2020;544–53. PMCID: PMC8075539 (**Student Paper Competition Finalist**)

C.5 **He L**, Zheng K. How do general-purpose sentiment analyzers perform when applied to health-related online social media data? In: Proceedings of the 2019 World Congress on Health and Biomedical Informatics (MEDINFO '19). 2019;1208–12. PMCID: PMC8061710 (**Student Best Paper Nomination**)

C.4 Shehada ER, **He L**, Eikey EV, Jen M, Wong A, Young S, Zheng K. Characterizing frequent flyers of an emergency department using cluster analysis. In: Proceedings of the 2019 World Congress on Health and Biomedical Informatics (MEDINFO '19). 2019;158–61. PMID: 31437905

C.3 Chi C, **He L**, Ravvaz K, Weissert J, P. Tonellato. Optimized decision support rules of precision warfarin treatment. In: Proceedings of the 2018 Pacific Symposium on Biocomputing (PSB '18). 2018;23:412–23. PMID: 29218901

C.2 Fan Y, **He L**, Zhang R. Evaluating automatic methods to extract patients' supplement use from clinical reports. In: Proceedings of the 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM '16). 2016;1054–61. DOI: 10.1109/BIBM.2016.7822668

C.1 Fan Y, **He L**, Pakhomov S, Melton G, Zhang R. Classifying supplement use status in clinical notes. The 2017 American Medical Informatics Association Joint Summits (AMIA Joint Summit '17). 2017;493–501. PMCID: PMC5543386 (**Student Paper Competition Second Place**)

Extended Abstracts, Workshops, and Posters

A.10 Nambisan, P, Devatha, P, **He, L**. (2024). Chronic diseases that increase suicidal risk: Analyzing social media posts using Metamap and AI. The American Public Health Association (APHA) Annual Meeting 2024.

A.9 **He L**, Ma H. Unpacking potential biases in using Natural Language Processing (NLP) for clinical entity extraction from notes of veterans with lymphoid malignancies. Mitigating AI Risk through Ethical Data Science Workshop at the American Medical Informatics Association Annual Symposium (AMIA'23)

A.8 **He L**, Moldenhauer M, Zheng K, Ma H. Extracting and assessing environmental exposures and substance use from clinical notes for veterans with lymphoid malignancies using Natural Language Processing (NLP). 2023 Military Health System Research Symposium (MHSRS'23)

A.7 **He L**, Moldenhauer M, Zheng K, Ma H. Analyzing free-text clinical narratives for veterans with lymphoid malignancies using natural language processing (NLP). American Society of Clinical Oncology Annual Meeting (ASCO'23)

A.6 **He L**, Quang J, Sarpong K, Kuo S, O'Connel R, Spiegelman L, Rudkin S. Disparities in Telehealth Experiences during the COVID-19 Pandemic at an Academic Medical Center. Podium abstract accepted to American Medical Informatics Association Clinical Informatics Conference. 2023. (AMIA CIC'23, in press)

A.5 **He L**, Ma H, Moldenhauer M, Zheng K. Extracting clinical and non-clinical information from clinical notes for veterans with lymphoid malignancies with limited expert annotation to assist clinical research. American Medical Informatics Association Annual Symposium Proceedings. 2022. (Graduate Student Consortium)

A.4 Griffin A, **He L**, Sunjaya A, King A, Khan Z, Douthit B, Nwadiugw M, Subbin V, Braunstein M, Nguyen V, Jaffe C, Schleyer T. Assessment of real-world health applications on FHIR. American Medical Informatics Association Annual Symposium Proceedings. 2022. (in press)

A.3 **He L**, Cheng Y, Zhou T, Xian Y. Investigating the narratives of anti-Asian hate speech on Twitter during the COVID-19 pandemic. American Medical Informatics Association Annual Symposium Proceedings. 2021

A.2 Goyal J, Ng DQ, Maddhuri J, Kumar AS, Jia S, **He L**, Wisseh C, Nguyen M, Lee J, McBane S, Zheng K, Hurley-Kim K, Nguyen Lee, Chan A, Cadiz CL. Predicting adverse drug events using the All of Us cohort data: a feasibility study. In: Proceedings of the 2021 American College of Clinical Pharmacy Annual Meeting (ACCP '21). 2021.

A.1 Yin T, **He L**. Challenges of applying sentiment analysis on health-related social media data. The 2019 Southern California Natural Language Processing Symposium (SoCal NLP '19) (**Best Poster Award**)

Workshops Organized

O.1 **He L**, Liu X, Reynolds T, Subbian V, Choi Y. Research using health-related social media in the Large Language Model era: methodological advancements, novel applications, and challenges. Collaborative workshop at AMIA 2024 Annual Symposium.

Grant Activities

Active

- 2025 CTSI Pilot Award to Advance Translational Science (Do, He)
  07/01/2025–06/30/2026, $25,000
  Towards a Holistic Approach for Opioid Treatment: Harnessing Natural Language Processing to Extract Social Determinants of Health from Clinical Notes
  Role: Co-PI; PI: Phuong Do (UWM); Co-I: Stephen Brandt (Medical College of Wisconsin)

- Northwestern Mutual Data Science Institute Paving ROADS Seed Fund Program (Do)
  01/30/2025–06/30/2026, $49,985
  Towards Equity in Opioid Use Disorder Care: Natural Language Processing Analysis of Social Determinants of Health in Medication Prescription Practices
  Role: Co-I; PI: Phuong Do (UWM)

- Northwestern Mutual Data Science Institute Paving ROADS Seed Fund Program (He, Madiraju, Deshpande)
  01/30/2024–07/30/2025, $99,996
  Towards a comprehensive framework and repository for documenting and assessing biases in using Large Language Models in health
  Role: PI; Co-PI: Praveen Madiraju (Marquette University), Priya Deshpande (Marquette University)

- National Science Foundation SBIR Phase 2 (Bradley)
  03/15/2024–02/28/2026, $943,819.00
  NSF Award #2335207: *Pedi-Sync: An Integrated Biomedical Platform with Custom Algorithm for Determining Optimum Feeding Protocols for Preterm Infants*
  Role: Co-I (Site PI for UWM subaward)

- National Institute of Health/National Institute of Child Health and Human Development (Bradley)
  06/12/2024–05/31/2025, $295,897
  1R41HD113429-01A1: *Pedi-Sync: A Clinical Decision Support Tool to Improve Oral Feeding Outcomes in Preterm Infants*
  Role: Co-I (Site PI for UWM subaward)

Talks and Presentations

Invited Talks

- Annual Southeast Wisconsin Data Science (SEAWINDS) Research Symposium, Medical College of Wisconsin, April 19, 2024

- "*Computational Analysis of Health Text: Applications, Challenges, and Opportunities*", Rigor and Reproducibility Seminar Series, University of Florida Interdisciplinary T32 in Movement Disorders and Neurorestoration, October 13, 2023

- "*Computational Analysis of Health Text: Applications, Challenges, and Opportunities*", COMPSCI 870 Medical Informatics Seminar, Department of Computer Science, University of Wisconsin-Milwaukee, September 22, 2023
- "*Computational Analysis of Health Text: Applications, Challenges, and Opportunities*", Division of Computational Health Sciences, University of Minnesota, Twin Cities, Minneapolis, Minnesota, April 3, 2023
- "*Computational Analysis of Health Text: Applications, Challenges, and Opportunities*", College of Information, University of North Texas, Virtual Presentation, March 24, 2023
- "*Computational Analysis of Health Text: Applications, Challenges, and Opportunities*", Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, Ohio, March 20, 2023
- "*Computational Analysis of Health Text: Applications, Challenges, and Opportunities*", Department of Computer Science, American University, Washington DC, February 28, 2023
- "*Computational Analysis of Health Text: Applications, Challenges, and Opportunities*", Zilber School of Public Health, University of Wisconsin, Milwaukee, Milwaukee, Wisconsin, February 15, 2023
- "*Computational Analysis of Health Text: Applications, Challenges, and Opportunities*", School of Computing, DePaul University, Chicago, Illinois, February 7, 2023
- "*Characterizing Frequent Attenders of Emergency Department Using Cluster Analysis*", Medical Intelligence and Innovation Institute (MI3), Children Hospital of Orange County (CHOC), Irvine, California, April 8, 2019

Conference Presentations

- "*Disparities in Telehealth Experiences during the COVID-19 Pandemic at an Academic Medical Center*", American Medical Informatics Association Clinical Informatics Conference 2023, Chicago, Illinois, May 25, 2023
- "*Help Me #DebunkThis: Unpacking Individual and Community's Collaborative Work in Information Credibility Assessment,*" The 25th ACM Conference On Computer-Supported Cooperative Work And Social Computing, Virtual Event, November 10, 2022
- "*Extracting Clinical and Non-clinical Information From Clinical Notes for Veterans with Lymphoid Malignancies with Limited Expert Annotation to Assist Clinical Research,*" Graduate Student Consortium, American Medical Informatics Association Annual Symposium 2022, Washington DC, November 5, 2022
- "*Investigating the Narratives of Anti-Asian Hate Speech on Twitter During the COVID-19 Pandemic,*" American Medical Informatics Association Annual Symposium 2021, San Diego, California, November 3, 2021
- "*What Do Patients Care About? Mining Fine-grained Patient Concerns from Online Physician Reviews Through Computer-assisted Multi-level Qualitative Analysis,*" American Medical Informatics Association Annual Symposium 2020, Virtual Event (Recorded, Co-present with Changyang He), November 16, 2020
- "*What Do Patients Care About? Mining Fine-grained Patient Concerns from Online Physician Reviews Through Computer-Assisted Multi-level Qualitative Analysis,*" American Medical Informatics Association Annual Symposium 2020, Student Paper Competition (Virtual Live Presentation), November 15, 2020
- "*Characterizing Frequent Attenders of Emergency Department Using Cluster Analysis,*" MedInfo 2019, Lyon, France, August 25, 2019

- "*How Do General-purpose Sentiment Analyzers Perform on Health-related Social Media Data?*" MedInfo 2019, Lyon, France, August 23, 2019

Teaching Experience

Instructor of Record

University of Wisconsin-Milwaukee

- HI 744 Text Retrieval and Its Applications in Biomedicine, Zilber College of Public Health (Fall 2024, Fall 2025)
- HI 741 Essential Programming for Health Informatics, Zilber College of Public Health (Spring 2024, Spring 2025)

University of California, Irvine

- ICS 33 Intermediate Programming, Summer 2022, Department of Informatics, Donald Bren School of Information and Computer Sciences (Enrollment: 46; Evaluation: 8.6/9.0)

Teaching Assistant

- INF 151 Project Management, Fall 2018, Department of Informatics, Donald Bren School of Information and Computer Sciences, University of California, Irvine
- CSci 1913 Introduction to Data Structures and Algorithms, Fall 2016 & Spring 2017, Department of Computer Science and Engineering, College of Science and Engineering, University of Minnesota, Twin Cities

Guest Lectures

- "*Human-centered Computational Analysis of Health Texts: Applications, Challenges, and Opportunities*," School of Information, Florida State University, Tallahassee, Florida, November 15, 2022. Instructor: Dr. Zhe He
- "*Computational Analysis of Social Media Data for Health Research: Applications, Challenges, and Opportunities*," Center for Digital Health and Analytics, University of Texas Health School of Biomedical Informatics, Houston, Texas, November 29, 2021. Instructor: Dr. Yang Gong
- "*Computational Analysis of Social Media Data for Health Research: Applications, Challenges, and Opportunities*," Center for Digital Health and Analytics, University of Texas Health School of Biomedical Informatics, Houston, Texas, October 21, 2021. Instructor: Dr. Sahiti Myneni

Mentoring

- Samaneh Omranian (PhD dissertation committee member, UWM)
- Nancy Melnik (PhD dissertation committee member, UWM)
- Apoorv Prasad (PhD dissertation committee member, UWM)
- Tasmima Khan (Undergraduate, UCI, 4/2021–11/2021)
- Tingjue Yin (Undergraduate, UCI, 3/2019–6/2020) [A.1, J.2, J.8]
- Peilin Gan (Undergraduate, UCI, 1/2020–12/2020)
- Tianyang Zhou (Undergraduate, UCI, 3/2020–6/2020) [A.3]

- Xinchen Zhang (Undergraduate, UCI, 1/2020–6/2020)
- Ya Cheng (Undergraduate, UCI, 3/2020–6/2020) [A.3]
- Yongxu Xian (Undergraduate, UCI, 3/2020–6/2020) [A.3]
- Daniel Davies (Undergraduate, UCI, 1/2019–6/2019)
- Su In Lee (Undergraduate, I-SURF Program, 6/2019–12/2019)
- Yiji Bae (Undergraduate, I-SURF Program, 6/2019–12/2019)
- Joohee Kwon (Undergraduate, I-SURF Program, 6/2019–12/2019)
- Haotian Hu (Undergraduate, UCI, 10/2019–12/2019)

Professional Service

Journal Editing

- Student Editorial Board Member, Journal of the American Medical Informatics Association (JAMIA), 2022–2023

Program Committee

- The 10th IEEE International Conference on Healthcare Informatics (ICHI '22), 2022
- The 11th IEEE International Conference on Healthcare Informatics (ICHI '23), 2023

Paper Reviewing

Journals

- Journal of Medical Internet Research (2019, 2020), BMC Medical Informatics and Decision Making (2020, 2021), Journal of Healthcare Informatics Research (2020, 2021), Journal of Biomedical Informatics (2021, 2022, 2023, 2024), IEEE Access (2022), ACM Computer-Supported Cooperative Work And Social Computing (2024, 2025), Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (2024), Journal of the American Medical Informatics Association (2024), ACM Computing Surveys (2024)

Conferences

- AMIA Annual Symposium (2018, 2019, 2020, 2021, 2022), AMIA Informatics Joint Summit (2019), CHI Late Breaking Work (2020), The Pacific Asia Conference on Information Systems (2021), CHI (2022, 2023 with 1 special recognition, 2024 with 1 special recognition, 2025)

Grant Reviewing

- Ad Hoc Reviewer, Biostatistics, Epidemiology & Research Design (BERD) Methodological Innovation Pilot Grant, Clinical & Translational Science Institute of Southeast Wisconsin (CTSI), January 2024

Volunteer & Events Organization

- JAMIA Journal Club Co-Manager & Moderator, 2023
- AMIA Annual Symposium Student Volunteer, 2017, 2021
- AMIA Year-in-Review Student Working Group Volunteer, 2020
- Women in Academia Reading Group Organizer, 2020

Media Coverage

- KXAN News: Do face masks work? Here are 49 scientific studies that explain why they do, December 2021
- Women in AMIA: AMIA as a Catalyst for Collaboration, August 2021

Skills

Research Methods

Qualitative Methods

- Interview, observational study, user study, grounded theory, wizard-of-oz

Quantitative Methods

- Survey, applied machine learning and deep learning, natural language processing, log analysis, social network analysis, statistical analysis (hypothesis testing, linear models, generalized linear models, longitudinal data analysis, linear mixed effect models, generalized estimating equation, survival analysis)

Programming Languages

- Python (proficient), R (medium), Matlab (medium), Java (medium), C/C++ (medium), SQL (medium), Lisp (familiar), Clojure (familiar), Go (familiar), Ruby (familiar)

Frameworks and Packages

- Python (scikit-learn, Numpy, Pandas, NLTK, Spacy, Gensim, Keras, PyTorch, Django, Tensorflow, transformers), R (ggplot, tidyverse, tidytext, dplyr, cluster, mltools, sna, igraph)

Databases

- MySQL, PostgreSQL

Languages

- Mandarin Chinese (Native), English (Full professional proficiency), Japanese (JLPT N3), German (Limited proficiency)

Last Updated: May 2025