Lu He

Email: lu.he@uci.edu

Website: luheholly.com

Phone: 949-572-7351

PhD Candidate
Department of Informatics
University of California, Irvine

RESEARCH INTERESTS

<u>Substantive</u>: Health informatics, human-computer interaction, social media mining, health communication, computer-supported cooperative work (CSCW)

<u>Methodological</u>: Natural language processing, applied machine learning & deep learning, social network analysis, statistical analysis

EDUCATION

2017-present	**University of California, Irvine**
	Ph.D. in Informatics
	Advisor: Kai Zheng, Ph.D.
2013-2017	**University of Minnesota, Twin Cities**
	Bachelor of Science with Distinction in Computer Science

PUBLICATION

Peer-reviewed Journals

[J7]   **He L**, He C. Help Me #DebunkThis: Unpacking Individual and Community's Collaborative Work in Information Credibility Assessment. Accepted to CSCW 2022, forthcoming

[J6]   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, forthcoming)

[J5]   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. Accepted to Computers in Human Behavior, 2021. DOI: https://doi.org/10.1016/j.chb.2021.107019

[J4]   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 Volume 5, CSCW (October 2021), 27 pages. DOI: https://dl.acm.org/doi/abs/10.1145/3479555

[J3]   **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. J Am Med Inform Assoc. 2021. (* equal contribution) PMID: 33690794

         Editor's Choice & Featured

[J2]   **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. J Am Med Inform Assoc. 2021;28(6):1125–34. PMID: 33355353

[J1]   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 Volume 1, CSCW (November 2017), 24 pages. DOI:https://doi.org/10.1145/3134708

Peer-reviewed Conference Proceedings

[C8]   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. AMIA Annu Symp Proc. 2022. (in press)

[C7]   **He L***, Song T*, Jiang Y, Yu P, Gong Y. To Improve Supportive Care for Patients Taking Oral Anticancer Agents. Proceedings of the 2021 World Congress on Health and Biomedical Informatic (MEDINFO 2021) (* equal contribution)

[C6]   **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. AMIA Annu Symp Proc. 2020;544–53. PMID: 33936428

         Student Paper Competition Finalist

[C5]   **He L**, Zheng K. How do general-purpose sentiment analyzers perform when applied to health-related online social media data? Proceedings of the 2019 World Congress on Health and Biomedical Informatics (MEDINFO 2019).

         Student Best Paper Nomination

[C4]   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 2019).

[C3]   Chi C, **He L**, Ravvaz K, Weissert J, P. Tonellato. Optimized Decision Support Rules of Precision Warfarin Treatment. 2018 Pacific Symposium on Biocomputing (PSB 2018).

[C2]   Fan Y, **He L**, Zhang R. Evaluating Automatic Methods to Extract Patients' Supplement Use From Clinical Reports. IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2017).

[C1]   Fan Y, **He L**, Pakhomov S, Melton G, Zhang R. Classifying Supplement Use Status in Clinical Notes. 2017 American Medical Informatics Association Joint Summits (AMIA 2017).

2nd Place in Student Paper Competition

Extended abstracts, workshops, and posters

[A4]   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. Poster accepted to AMIA Annu Symp Proc. 2022. To Appear.

[A3]   **He L**, Cheng Y, Zhou T, Xian Y. Investigating the Narratives of Anti-Asian Hate Speech on Twitter During the COVID-19 Pandemic. Podium Abstract, AMIA Annu Symp Proc. 2021

[A2]   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 2021)

[A1]   Yin T, **He L.** Challenges of applying sentiment analysis on health-related social media data. Southern California Natural Language Processing Symposium (SoCal 2019)

Best Poster Award

MANUSCRIPTS IN-PROGRESS

[P1]   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. *Under Review (* indicates equal contribution)*

[P2]   **He L,** Yin T, Zheng K. They May not Work! An Evaluation of Eleven Sentiment Analysis Tools on Seven Social Media Datasets. *Major revision and resubmit at Journal of Biomedical Informatics (JBI)*

EXPERIENCES (SELECTED)

2022   **Graduate Research Assistant**

-present   Advisor: Helen Ma, MD, Kai Zheng, Ph.D.

Project: To use NLP methods to extract environmental factors from clinical notes and assist predictive modeling of patient outcomes among a Veteran population.

2021   **AMIA Workgroup Volunteer**

-present   Advisor: Titus Schleyer, DMD, Ph.D.

Project: To conduct a comprehensive survey of the landscape of Fast Healthcare Interoperability Resources (FHIR) implementations.

2020   **Graduate Research Assistant**

-present   Advisors: Caryn Bradley, PT, Ph.D., William Tang, Ph.D., Kai Zheng, Ph.D., Anton Palma, Ph.D.

Project: Working closely with researchers and practitioners in physical therapy, biomedical engineering, and biostatistics, I combine their medical knowledge to develop clinical algorithms to predict days to discharge for infants in Neonatal Intensive Care Unit (NICU) based on their feeding performance.

2017   **Graduate Research Assistant**

-present   Department of Informatics, University of California, Irvine

Advisors: Kai Zheng, Ph.D., Yunan Chen, Ph.D.

Project 1: Conducted a systematic literature review on the use of computational tools for healthcare social media studies. Empirically evaluated the validity of tools. [C5, A1, J2]

Project 2: Characterized emergency department frequent attenders using cluster analysis. [C4]

Project 3: Developing computer-assisted qualitative analysis pipeline to extract useful information from free-text data. [C6]

Project 4: Analyzing public responses on social media to the COVID-19 pandemic. [J3, A3]

2018.6-9   **Bioinformatics Programmer**

Bakar Computational Health Sciences Institute (BCHSI),

University of California, San Francisco

Advisor: Gundolf Schenk, Ph.D.

Project: Developed test units for a program to identify Protected Health Information (PHI) in clinical notes. Improved the performance of the program to be able to run on 60 million notes.

GRANT ACTIVITIES

2022-2023   UCI-Internal (Ma, Zheng)

UCI ICTS/VA Long Beach Informatics Pilot Award, $50,000
Outcomes from the National Database of Lymphoid Malignancies in
US Veterans Using Data from the Veterans Health Administration, the
Largest Integrated Healthcare System.

Role: Grant writing and research assistance

2022-2023   NSF (Bradley, Tang)

STTR Phase 1, $255,990

Role: Research assistance

2018   UCI-Internal (Zheng)

Academic Senate Council on Research, Computing and Libraries
(CORCL), $3,500

Developing a domain-specific sentiment analyzer for studying health-
related social media data

Role: Grant writing and research assistance

FELLOWSHIPS, HONORS & AWARDS

2022   Graduate Dean's Dissertation Fellowship, UC Irvine

2021   Editor's Choice of Journal of American Medical Informatics
             Association (JAMIA) [J3]

2020   Student Paper Competition Finalist, American Medical Informatics
             Association Annual Symposium (AMIA 2020) [C6]

2019   Best Poster Award (1/40), Southern California Natural Language
             Processing Symposium (SoCal 2019) [W1]

2019   Student Best Paper Nomination, 2019 World Congress on Health and
             Biomedical Informatics (MedInfo '19) [C2]

2018   CRA-W Grad Cohort Workshop, Computing Research Association

2017   Graduate Dean Recruitment Fellowship, Department of Informatics,
             UC Irvine

2017   Dean's Award, Department of Informatics, UC Irvine

2017   2nd Place in Student Paper Competition, 2017 American Medical
             Informatics Association Joint Summits (AMIA 2017) [C1]

2014	Undergraduate Special Recognition, Department of German, Nordic, Slavic and Dutch, University of Minnesota
2013	Maroon Global Excellence Scholarship (four-year), University of Minnesota
2013	Dean's List (three-year), University of Minnesota

TEACHING

Instructor of Record

| 2022 | ICS 33 (Intermediate Programming) |

Teaching assistant

| 2018 | IN4MTX 151 (Project Management) |
| 2017 | CSci 1913 (Introduction to Algorithms and Data Structures) |

SKILLS

Research methods

Qualitative

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

Quantitative

Survey, applied machine learning & deep learning, natural language processing, log analysis, social network analysis, statistical analysis, linear models, generalized linear models, longitudinal data analysis, linear mixed effect models

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), R (ggplot, tidyverse, tidytext, dplyr, cluster, mltools, sna, igraph)

Databases

MySQL, PostgreSQL

Misc.

LaTex, Github, Jupyter Notebook, Amazon Web Services

Languages

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

SERVICES
Journal Editing
JAMIA Student Editorial Board Member (2022-2023)

Program Chair
Conferences

ICHI 2022 : 10th IEEE International Conference on Healthcare Informatics

Reviewer
Conferences

AMIA Annual Symposium (2018-2022), AMIA Informatics Joint Summit (2019), CHI Late Breaking Work (2020), The Pacific Asia Conference on Information Systems (2021), CHI (2022)

Journals

Journal of American Medical Informatics Association (2022-2023), 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), IEEE Access (2022)

Volunteer

AMIA Symposium Student Volunteer (2017, 2021), AMIA Year-in-Review Student Working Group Volunteer (2020), Women in Academia Reading Group Organizer (2020)

TALKS & PRESENTATIONS
Invited Talks
2019.4.8   "Characterizing Frequent Attenders of Emergency Department Using Cluster Analysis"

Medical Intelligence and Innovation Institute (MI3), Children Hospital of Orange County (CHOC)

Guest Lectures
2021.11.29   "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 (Host: Dr. Yang Gong)

2021.10.21   "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 (Host: Dr. Sahiti Myneni)

Conference Presentations

2021.11.3   Investigating the Narratives of Anti-Asian Hate Speech on Twitter During the COVID-19 Pandemic"

AMIA Annual Symposium 2021, San Diego, California

2020.11.16   "What Do Patients Care About? Mining Fine-grained Patient Concerns from Online Physician Reviews Through Computer-Assisted Multi-level Qualitative Analysis"

AMIA Annual Symposium 2020, Virtual Event (Recorded, Co-present with Changyang He)

2020.11.15   "What Do Patients Care About? Mining Fine-grained Patient Concerns from Online Physician Reviews Through Computer-Assisted Multi-level Qualitative Analysis"

AMIA Annual Symposium 2020, Student Paper Competition (Virtual Live Presentation)

2019.8.25   "Characterizing Frequent Attenders of Emergency Department Using Cluster Analysis"

MedInfo 2019, Lyon, France

2019.8.23   "How Do General-purpose Sentiment Analyzers Perform on Health-related Social Media Data?"

MedInfo 2019, Lyon, France

MEDIA COVERAGE

2021.12   KXAN News: Do face masks work? Here are 49 scientific studies that explain why they do

2021.8   Women in AMIA: Episode 20 - AMIA as a Catalyst for Collaboration

MENTORING

Tasmima Khan (Undergraduate, UCI, 2021.4-2021.11)

Tingjue Yin (Undergraduate, UCI, 2019.3-2020.6) [A1, J2]

Peilin Gan (Undergraduate, UCI, 2020.1-2020.12)

Xinchen Zhang (Undergraduate, UCI, 2020.1-2020.6)

Tianyang Zhou (Undergraduate, UCI, 2020.3-2020.6) [A3]

Ya Cheng (Undergraduate, UCI, 2020.3-2020.6) [A3]

Yongxu Xian (Undergraduate, UCI, 2020.3-2020.6) [A3]

Daniel Davies (Undergraduate, UCI, 2019.1-2019.6)

Su In Lee (Undergraduate, I-SURF Program, 2019.6-2019.12)

Yiji Bae (Undergraduate, I-SURF Program, 2019.6-2019.12)

Joohee Kwon (Undergraduate, I-SURF Program, 2019.6-2019.12)

Haotian Hu (Undergraduate, UCI, 2019.10-2019.12)

Last Update: June 2022