McWilliams School of Biomedical Informatics at UTHealth Houston

10/12/2024

NAME: Sunyang Fu, PhD, MHI

PRESENT TITLE(S): Assistant Professor, Department of Clinical and Health Informatics, Associate Director of Team Science, Center for Translational AI Excellence and Applications in Medicine (TEAM-AI), McWilliams School of Biomedical Informatics, UTHealth Houston

WORK ADDRESS: 7000 Fannin St #600, Houston, TX 77030 McWilliams School of Biomedical Informatics, The University of Texas Health Science Center at Houston

CITIZENSHIP: China

UNDERGRADUATE EDUCATION:
The University of Iowa, Henry B. Tippie College of Business — Iowa City, IA
Bachelor of Business Administration, Management Information Systems, Finance,
August 2010 - December 2014

GRADUATE EDUCATION:
The University of Minnesota – Twin Cities — Minneapolis, MN
Doctor of Philosophy, Bioinformatics and Computational Biology
September 2018 – December 2021

The University of Michigan – Ann Arbor — Ann Arbor, MI
Master of Health Informatics, Health Informatics
September 2015 - April 2017

POSTGRADUATE TRAINING:
National Institute of Health - AIM-AHEAD Program, — Houston, TX
Leadership Fellow, Artificial Intelligence/Machine Learning,
2023 – 2024

ACADEMIC & ADMINISTRATIVE APPOINTMENTS:
Assistant Professor, Department of Clinical and Health Informatics
Associate Director of Team Science, Center for Translational AI Excellence and Applications in Medicine (TEAM-AI), McWilliams School of Biomedical Informatics, UTHealth Houston

PROFESSIONAL ORGANIZATIONS:
National Organizations:

- American Medical Informatics Association (AMIA)
  - Member-at-Large: Public Health Informatics Working Group
- Network for Investigation of Delirium: Unifying Scientists (NIDUS)
- AIM-AHEAD Consortium
- Evolve to Next-Gen ACT (ENACT) Network
  - Lead, ENACT NLP WG – Delirium Interest Group
- Observational Health Data Sciences and Informatics Network
- Institute of Electrical and Electronics Engineers
- American Association for Cancer Research

Program Committee Member

- AMIA Informatics Summit 2024, SPC
- IEEE International Conference on Bioinformatics and Biomedicine 2022, PC
- HealthNLP 2020, PC
- IEEE International Conference on Bioinformatics and Biomedicine 2020, PC

- BIBM'18 Knowledge Discovery in Translational Biomedical Informatics (KDTBI), PC
- IEEE International Conference on Bioinformatics and Biomedicine 2019, PC
- Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, PC
- Proceedings of the BioCreative/OHNLP Challenge, PC

HONORS AND AWARDS:

2024	Winner, The National A2 Pilot Awards competition, the Artificial Intelligence and Technology Collaboratories (AITC) for Aging Research program.
2024	NIDUS II LOI Award, Network for Investigation of Delirium: Unifying Scientists
2024	SPC Award, American Medical Informatics Association
2024	2nd Place of Chemotherapy Treatment Timelines Extraction Challenge, NAACL - Clinical NLP 2024
2023	Leadership Fellow, NIH AIM-AHEAD
2023	Research Spotlight, Mayo Clinic Discovery's Edge
2023	Monthly Trainee Spotlight, American Society for Clinical Pharmacology and Therapeutics
2023	Editor Featured Research Highlight, Translational Bytes, Clinical and Translational Science
2022	2nd Place of Student Paper Competition, American Medical Informatics Association (AMIA) 2022 Informatics Summit
2022	Best Student Paper Finalist, American Medical Informatics Association (AMIA) 2022 Informatics Summit
2022	Community Forum Speaker, National COVID Cohort Collaborative (N3C)
2022	Faculty-nominated outstanding student speaker, Bioinformatics and Computational Biology Research Symposium, University of Minnesota
2016	Chair, Health Informatics, Public Health Student Assembly, University of Michigan
2015	University of Michigan NRIMHI Scholarship, University of Michigan
2014	Earl A. Wimmer Scholarship, University of Iowa
2014	Hubert E. Storer Engineering Student Entrepreneurial Start-up Award Finalist
2013-2014	Co-president, UIBC Student Organization, University of Iowa
2012-2014	University of Iowa Dean's List, University of Iowa
2010-2014	University of Iowa National Scholarship, University of Iowa

EDITORIAL POSITIONS:

- Review editor - Frontiers in Public Health

SERVICE ON NATIONAL GRANT REVIEW PANELS, STUDY SECTIONS, COMMITTEES:

- UK Research and Innovation - Enhancing biomedical and health-related data and digital platform resources (https://www.ukri.org/opportunity/enhancing-biomedical-and-health-related-data-and-digital-platform-resources/)
- Modelling for Pandemic Preparedness for The Netherlands Organisation for Health Research and Development (ZonMw)

SERVICE ON THE UNIVERSITY OF TEXAS HEALTH SCIENCE CENTER AT HOUSTON COMMITTEES:

- MSBMI - Faculty development committee member

CURRENT TEACHING RESPONSIBILITIES

- 3 PhD students (1 MSBMI, 1 SPH, and 1 UMN)

MENTORING ACTIVITIES:

Individual	Position	Timeframe	Outcomes Current Status
Priyank Raj	PhD Student, UTHealth	Oct 2024 - Present	
Zhiyi Yue	PhD Student, UTHealth	Oct 2024 - Present	
Taylor Hawkins	PhD Student, BICB University of Minnesota	Jan 2022 - Present	2 AMIA papers
Ethan Zhang	Mayo Clinic high school mentorship program	2023 - 2024	Admission to Northwestern University, Honor Class
Kai Zhang	Summer Intern, Mayo Clinic	Jan 2023 – May 2023	Nature Medicine publication; PhD Student, Lehigh University
Bhavani Singh	Informatics Specialist, Mayo Clinic	Sep 2019 – June 2021	Sr. Data Science Analyst at Mayo Clinic
Rex Su	Summer Intern, Mayo Clinic	May 2019 – Aug 2019	AMIA Abstract; MHI at the University of Michigan
Kang Kai	Summer Intern, Mayo Clinic	May 2019 – Aug 2019	AMIA Abstract; MSI at the University of Michigan
Andrew Wen	Informatics Specialist, Mayo Clinic	2017 - 2018	Informatics Specialist, Department of Health Sciences Research

CURRENT GRANT SUPPORT:

Project title: Detection of Adverse Drug Event Using NLP Among Older Adults with Heart Failure
Role: mPI
Funding agency: The National a2 Pilot Awards, NIA
Percentage of full-time effort: 10%
Award period: $200,000 (direct)
Award amount: 01/01/2025 – 12/31/2025
Brief description: This project aims to develop a natural language processing model to detect adverse drug event among older adults with heart failure.

Project title: AI-Empowered Differential Detection of Patients with Delirium, Dementia, and DSD
Role: PI
Funding agency: NIDUS II Pilot Award, NIDUS (R33AG071744)
Percentage of full-time effort: 18%
Award period: $40,000 (direct)
Award amount: 11/01/2024 – 10/31/2025
Brief description: This project aims to develop an AI model to provide differential detection of patients with delirium and dementia.

Project title: Facilitate observational studies of alzheimer's disease and alzheimer's disease-related dementias using ontology and natural language processing
Role: Co-I
Funding agency: NIA, RF1AG 72799-01
Percentage of full-time effort: 25%
Award period: $954,000 (direct and indirect)
Award amount: 05/01/2021 – 04/30/2025
Brief description: This project aims to develop informatics solutions for EHR-based ADRD research.

PAST GRANT SUPPORT:

Project title: The AIM-AHEAD Fellowship Program in Leadership
Role: PI
Funding agency: AIM-AHEAD Consortium, 3OT2OD032581-01S1-207
Percentage of full-time effort: 25%

Award period: $50,000 (direct and indirect)
Award amount: 09/18/2023 – 06/17/2024
Brief description: The training program aims to develop participants who actively engage in mentored didactic and experiential educational activities, acquiring the leadership competencies necessary to promote and achieve the strategic imperatives of AIM-AHEAD.

PUBLICATIONS:

Abstracts

1. Brannon ES, Reynolds T, Fu S, Harden K, Okullo D, Ramani S. An eHARS Dashboard for State HIV Surveillance. Online Journal of Public Health Informatics. 2016 Mar 24;8(1).
2. Tibbo ME, Wyles CC, Maradit Kremers H, Fu S, Wang Y, Sohn S, Berry DJ, Lewallen DG. Use of Natural language processing (NLP) tools to identify and classify periprosthetic femur fractures 7th ISAR Congress, Reykjavik, Iceland. 2018.
3. Wyles CC, Tibbo ME, Maradit Kremers H, Fu S, Sohn S, Berry DJ, Lewallen DG. Natural language processing (NLP) tools enable facile and accurate capture of surgical approach from operative reports in total hip arthroplasty (THA) 7th ISAR Congress, Reykjavik, Iceland. 2018.
4. Jinseok AL*, Fu S*, VG Vydiswaran, Supervised Learning Approach to Link Prediction in FDA Adverse Event Reporting System (FAERS) Database Network. InAMIA annual symposium proceedings 2018. American Medical Informatics Association. (*equal contribution)
5. Fu S, Leung LY, Raulli AO, Kallmes DF, Kinsman KA, Nelson K, Clark MS, Luetmer PH, Kingsbury PR, Kent DM, Liu, H. Natural Language Processing for the Identification of Silent Brain Infarcts from Neuroimaging Reports. InAMIA annual symposium proceedings 2018. American Medical Informatics Association.
6. Fu S. infoFACT: An Informatics Framework for Automated Data Abstraction. InAMIA annual symposium proceedings 2018. NLP-Working Group. American Medical Informatics Association.
7. Welsh JL, Fu S, Liao J, Sugg SL, Scott-Conner CE, Weigel RJ, Erdahl LM, Lizarraga IM. Long-term patient satisfaction with cosmetic outcome and psychosocial wellbeing after breast conserving therapy is affected only by lumpectomy volume. InCANCER RESEARCH 2016 Feb 15 (Vol. 76). 615 CHESTNUT ST, 17TH FLOOR, PHILADELPHIA, PA 19106-4404 USA: AMER ASSOC CANCER RESEARCH.
8. Juhn Y, Moon S, Wi CI, Fu S, Weston J, Porcher J, Johnson EF, Seol HY, Sohn S, Ryu E, Liu H. Automated chart review for identifying pre-and peri-natal risk factors associated with childhood asthma. InA60. PEDIATRIC ALLERGY AND ASTHMA 2018 May (pp. A2032-A2032). American Thoracic Society.
9. Tibbo ME, Wyles CC, Maradit-Kremers H, Fu S, Wang Y, Sohn S, Berry DJ, Lewallen DG. USE OF NATURAL LANGUAGE PROCESSING TOOLS TO IDENTIFY AND CLASSIFY PERIPROSTHETIC FEMORAL FRACTURES FROM ELECTRONIC HEALTH RECORDS. InOrthopaedic Proceedings 2018 Oct (Vol. 100, No. SUPP_13, pp. 55-55). The British Editorial Society of Bone & Joint Surgery.
10. Zhao Y, Fu S, Larson NB, Decker PA, Chamberlain AM, Roger VL, Liu H, Bielinski SJ. Validation of Phenotyping Algorithms for Stroke from Electronic Health Records Using Natural Language Processing EPI | Lifestyle 2019 Scientific Sessions.
11. Fu S, Leung LY, Peterson KJ, Lehman VT, Rydberg CH, Madan N, Lanfranchi M, Kallmes DF, Luetmer PH, Kent DM, Liu H. A Case Study of Examining the Information Loss Among Three Clinical Data Sources for Ensuring Transparent and Safe AI Models. Mayo Clinic Artificial Intelligence Symposium, Rochester, MN. May 23, 2019.
12. Zhao Y, Fu S, Bielinski SJ, Decker P, Chamberlain AM, Roger VL, Liu H, Larson NB. Abstract P259: Using Natural Language Processing and Machine Learning to Identify Incident Stroke From Electronic Health Records. Circulation. 2020 Mar 3;141(Suppl_1):AP259-.
13. Leung LY, Fu S, Nelson J, Kallmes DF, Luetmer PH, Liu H, Kent DM. Examining the Information Loss Between Neuroimages and Neuroimaging Reports for Detection of Silent Brain Infarcts and White Matter Disease Using Artificial Intelligence Technologies. Stroke. 2020 Feb;51(Suppl_1):A135-

14. Wang L, Fu S, Wen A, Ruan X, He H, Liu S, Moon S, Mai M, Riaz I, Wang N, Yang P. Evaluation of mCODE Coverage in EHR: a Scoping Review of Cancer Natural Language Processing. In2022 IEEE 10th International Conference on Healthcare Informatics (ICHI) 2022 Jun 11 (pp. 517-518). IEEE.
15. He H, Fu S, Wang L, Wen A, Liu S, Liu H. MedTator: A Serverless Web-based Tool for Corpus Annotation. In2022 IEEE 10th International Conference on Healthcare Informatics (ICHI) 2022 Jun 11 (pp. 530-531). IEEE.
16. Oian R, Fu S, Liu H. Evaluation of Document-Level Identification of Pulmonary Nodules in Radiology Reports Using FLAIR Natural Language Processing Framework. In2022 IEEE 10th International Conference on Healthcare Informatics (ICHI) 2022 Jun 11 (pp. 515-516). IEEE.
17. Bhandarkar AR, Onyedimma C, Jarrah R, Fu S, Liu H, Bydon M. 170 An Integrated Voice Recognition and Natural Language Processing Platform to Automatically Extract Thoracolumbar Injury Classification Score (TLICS) Features from Radiology Reports. Neurosurgery. 2022 Apr 1;68(Supplement_1):50-1.
18. Fu S, Wang L, Odedina FT, Liu H. Abstract A020: Assessment of EHR data quality variabilities among different racial groups in the cohort of de novo stage IV breast cancer. InCancer Epidemiology, Biomarkers & Prevention 2023 Jan 1 (Vol. 32, No. 1_Supplement, pp. A020-A020). The American Association for Cancer Research.

Refereed Original Articles in Journals

1. Spanheimer PM, Armstrong JG, Fu S, Liao J, Regenbogen SE, Byrn JC. Robotic proctectomy for rectal cancer: analysis of 71 patients from a single institution. The International Journal of Medical Robotics and Computer Assisted Surgery. 2017 Dec 1;13(4).
2. Wang Y, Afzal N, Fu S, Wang L, Shen F, Rastegar-Mojarad M, Liu H. MedSTS: a resource for clinical semantic textual similarity. Language Resources and Evaluation. 2018 Aug 28:1-6.
3. Tibbo ME, Wyles CC, Fu S, Sohn S, Lewallen DG, Berry DJ, Maradit-Kremers H. Use of Natural Language Processing Tools to Identify and Classify Periprosthetic Femur Fractures. The Journal of Arthroplasty. 2019 Jul 24.
4. Fu S, Leung LY, Wang Y, Raulli AO, Kallmes DF, Kinsman KA, Nelson KB, Clark MS, Luetmer PH, Kingsbury PR, Kent DM. Natural language processing for the identification of silent brain infarcts from neuroimaging reports. JMIR medical informatics. 2019;7(2):e12109.
5. Wyles CC, Tibbo ME, Fu S, Wang Y, Sohn S, Kremers WK, Berry DJ, Lewallen DG, Maradit-Kremers H. Use of Natural Language Processing Algorithms to Identify Common Data Elements in Operative Notes for Total Hip Arthroplasty. JBJS. 2019 Oct 15.
6. Wen A, Fu S, Moon S, El Wazir M, Rosenbaum A, Kaggal VC, Liu S, Sohn S, Liu H, Fan J. Desiderata for delivering NLP to accelerate healthcare AI advancement and a Mayo Clinic NLP-as-a-service implementation. npj Digital Medicine. 2019 Dec 17;2(1):1-7.
7. Wang L, Olson JE, Bielinski SJ, St Sauver JL, Fu S, He H, Cicek MS, Hathcock MA, Cerhan JR, Liu H. Impact of diverse data sources on computational phenotyping. Frontiers in Genetics. 2020 Jun 3;11:556.
8. Fu S, Leung LY, Raulli AO, Kallmes DF, Kinsman KA, Nelson KB, Clark MS, Luetmer PH, Kingsbury PR, Kent DM, Liu H. Assessment of the impact of EHR heterogeneity for clinical research through a case study of silent brain infarction. BMC Medical Informatics and Decision Making. 2020 Dec;20(1):1-2.
9. Fu S, Chen D, He H, Liu S, Moon S, Peterson KJ, Shen F, Wang L, Wang Y, Wen A, Zhao Y. Clinical Concept Extraction: a Methodology Review. Journal of Biomedical Informatics. 2020 Aug 6:103526.
10. Carlson LA, Jeffery MM, Fu S, He H, McCoy RG, Wang Y, Hooten WM, St Sauver J, Liu H, Fan J. Characterizing Chronic Pain Episodes in Clinical Text at Two Health Care Systems: Comprehensive Annotation and Corpus Analysis. JMIR medical informatics. 2020;8(11):e18659.
11. Fu S, Lopes GS, Pagali SR, Thorsteinsdottir B, LeBrasseur NK, Wen A, Liu H, Rocca WA, Olson JE, St Sauver J, Sohn S. Ascertainment of delirium status using natural language processing from electronic health records. The Journals of Gerontology: Series A. 2020

Oct 30.
12. Wang Y, Fu S, Shen F, Henry S, Uzuner O, Liu H. The 2019 n2c2/OHNLP Track on Clinical Semantic Textual Similarity: Overview. JMIR Medical Informatics. 2020;8(11):e23375.
13. Wen A, Wang L, He H, Liu S, Fu S, Sohn S, Kugel JA, Kaggal VC, Huang M, Wang Y, Shen F. An aberration detection-based approach for sentinel syndromic surveillance of covid-19 and other novel influenza-like illnesses. Journal of Biomedical Informatics. 2020 Dec 12;113:103660.
14. Shen F, Liu S, Fu S, Wang Y, Henry S, Uzuner O, Liu H. Family History Extraction From Synthetic Clinical Narratives Using Natural Language Processing: Overview and Evaluation of a Challenge Data Set and Solutions for the 2019 National NLP Clinical Challenges (n2c2)/Open Health Natural Language Processing (OHNLP) Competition. JMIR Medical Informatics. 2021 Jan 27;9(1):e24008.
15. Fu S, Wyles CC, Osmon DR, Carvour ML, Sagheb E, Ramazanian T, Kremers WK, Lewallen DG, Berry DJ, Sohn S, Kremers HM. Automated detection of periprosthetic joint infections and data elements using natural language processing. The Journal of arthroplasty. 2021 Feb 1;36(2):688-92.
16. Zhao Y, Fu S, Bielinski SJ, Decker PA, Chamberlain AM, Roger VL, Liu H, Larson NB. Natural Language Processing and Machine Learning for Identifying Incident Stroke From Electronic Health Records: Algorithm Development and Validation. Journal of medical Internet research. 2021 Mar 8;23(3):e22951.
17. Leung LY, Fu S, Luetmer PH, Kallmes DF, Madan N, Weinstein G, Lehman VT, Rydberg CH, Nelson J, Liu H, Kent DM. Agreement between neuroimages and reports for natural language processing-based detection of silent brain infarcts and white matter disease. BMC neurology. 2021 Dec;21(1):1-5
18. Mastorakos G, Khurana A, Huang M, Fu S, Tafti AP, Fan J, Liu H. Probing Patient Messages Enhanced by Natural Language Processing: A Top-Down Message Corpus Analysis. Health Data Science. 2021 Jun 16;2021.
19. Gao CC, Suarez NR, Toloza FJ, Zuniga AS, McCarthy SR, Boehmer KR, Yao L, Fu S, Brito JP. Patients' Perspective About the Cost of Diabetes Management: An Analysis of Online Health Communities. Mayo Clinic Proceedings: Innovations, Quality & Outcomes. 2021 Oct 1;5(5):898-906.
20. Pagali S, Fu S, Lindroth H, Sohn S, Burton MC, Lapid M. Delirium occurrence and association with outcomes in hospitalized COVID-19 patients. International psychogeriatrics. 2021 Sep 23:1-5.
21. He H, Fu S, Wang L, Liu S, Wen A, Liu H. MedTator: a serverless annotation tool for corpus development. Bioinformatics. 2022 Jan 4.
22. Zong N, Li N, Wen A, Ngo V, Yu Y, Huang M, Chowdhury S, Jiang C, Fu S, Weinshilboum R, Jiang G. BETA: a comprehensive benchmark for computational drug–target prediction. Briefings in Bioinformatics. 2022 Jun 2.
23. Zong N, Wen A, Moon S, Fu S, Wang L, Zhao Y, Yu Y, Huang M, Wang Y, Zheng G, Mielke MM. Computational drug repurposing based on electronic health records: a scoping review. NPJ digital medicine. 2022 Jun 14;5(1):1-8.
24. Wang L, Fu S, Wen A, Ruan X, He H, Liu S, Moon S, Mai M, Riaz IB, Wang N, Yang P. Assessment of Electronic Health Record for Cancer Research and Patient Care Through a Scoping Review of Cancer Natural Language Processing. JCO Clinical Cancer Informatics. 2022 Aug;6:e2200006.
25. Han P, Fu S, Kolis J, Hughes R, Hallstrom BR, Carvour M, Maradit-Kremers H, Sohn S, Vydiswaran VV. Multicenter Validation of Natural Language Processing Algorithms for the Detection of Common Data Elements in Operative Notes for Total Hip Arthroplasty: Algorithm Development and Validation. JMIR medical informatics. 2022 Aug 31;10(8):e38155.
26. Ruan X, Fu S, Storlie CB, Mathis KL, Larson DW, Liu H. Real-time risk prediction of colorectal surgery-related post-surgical complications using GRU-D model. Journal of Biomedical Informatics. 2022 Sep 24:104202.
27. Pagali SR, Kumar R, Fu S, Sohn S, Yousufuddin M. Natural Language Processing CAM Algorithm Improves Delirium Detection Compared With Conventional Methods. American Journal of Medical Quality. 2022 Oct 26:10-97.

28. Wyles CC, Fu S, Odum SL, Rowe T, Habet NA, Berry DJ, Lewallen DG, Maradit-Kremers H, Sohn S, Springer BD. External Validation of Natural Language Processing Algorithms to Extract Common Data Elements in THA Operative Notes. The Journal of Arthroplasty. 2022 Oct 22.
29. Fu S, Vassilaki M, Ibrahim OA, Petersen R, Pagali S, St Sauver J, Moon S, Wang L, Fan JW, Liu H, Sohn S. Quality Assessment of Functional Status Documentation in EHRs Across Different Healthcare Institutions. Frontiers in Digital Health.:196.
30. Fu S, Wang L, Moon S, Zong N, He H, Pejaver V, Relevo R, Walden A, Haendel M, Chute CG, Liu H. Recommended Practices and Ethical Considerations for Natural Language Processing-Assisted Observational Research: A Scoping Review. Clinical and translational science. 2022 Dec 7.
31. Wang AY, Leung LY, Puttock EJ, Luetmer PH, Kallmes DF, Nelson J, Fu S, Zheng C, Liu H, Chen W, Kent DM. Stratifying Future Stroke Risk with Incidentally Discovered White Matter Disease Severity and Covert Brain Infarct Site. Cerebrovascular Diseases. 2023;52(1):117-22.
32. Kent DM, Leung LY, Zhou Y, Luetmer PH, Kallmes DF, Nelson J, Fu S, Puttock EJ, Zheng C, Liu H, Chen W. Association of Incidentally Discovered Covert Cerebrovascular Disease Identified Using Natural Language Processing and Future Dementia. Journal of the American Heart Association. 2023 Jan 3;12(1):e027672.
33. Moon S, Liu S, Kshatriya BS, Fu S, Moser ED, Bielinski SJ, Fan J, Liu H. Assessing document section heterogeneity across multiple electronic health record systems for computational phenotyping: A case study of heart-failure phenotyping algorithm. Plos one. 2023 Mar 31;18(3):e0283800.
34. Dang Y, Li F, Hu X, Keloth VK, Zhang M, Fu S, Amith MF, Fan JW, Du J, Yu E, Liu H. Systematic design and data-driven evaluation of social determinants of health ontology (SDoHO). Journal of the American Medical Informatics Association. 2023 Jun 10:ocad096.
35. Fu S, Wen A, Liu H. Clinical Natural Language Processing in Secondary Use of EHR for Research. InClinical Research Informatics 2023 Jun 15 (pp. 433-451). Cham: Springer International Publishing.
36. Wang L, He H, Wen A, Moon S, Fu S, Peterson KJ, Ai X, Liu S, Kavuluru R, Liu H. Acquisition of a Lexicon for Family History Information: Bidirectional Encoder Representations From Transformers–Assisted Sublanguage Analysis. JMIR Medical Informatics. 2023 Jun 27;11:e48072.
37. Wen A, He H, Fu S, Liu S, Miller K, Wang L, Roberts KE, Bedrick SD, Hersh WR, Liu H. The IMPACT framework and implementation for accessible in silico clinical phenotyping in the digital era. npj Digital Medicine. 2023 Jul 21;6(1):132.
38. Liu S, Wen A, Wang L, He H, Fu S, Miller R, Williams A, Harris D, Kavuluru R, Liu M, Abu-el-Rub N. An open natural language processing (NLP) framework for EHR-based clinical research: a case demonstration using the National COVID Cohort Collaborative (N3C). Journal of the American Medical Informatics Association. 2023 Aug 14:ocad134.
39. Vassilaki M, Fu S, Christenson LR, Garg M, Petersen RC, St Sauver J, Sohn S. Characterizing Performance Gaps of a Code-Based Dementia Algorithm in a Population-Based Cohort of Cognitive Aging. Journal of Alzheimer's Disease. 2023 Aug 24(Preprint):1-0.
40. Bhandarkar AR, Onyedimma C, Jarrah RM, Ibrahim S, Fu S, Liu H, Bydon M. An Integrated Voice Recognition and Natural Language Processing Platform to Automatically Extract Thoracolumbar Injury Classification Score Features From Radiology Reports. World Neurosurgery. 2023 Dec 15.
41. Fu S, Jia H, Vassilaki M, Keloth VK, Dang Y, Zhou Y, Garg M, Petersen RC, St Sauver J, Moon S, Wang L. FedFSA: Hybrid and federated framework for functional status ascertainment across institutions. Journal of Biomedical Informatics. 2024 Apr 1;152:104623.
42. Fu S, Wang L, He H, Wen A, Zong N, Kumari A, Liu F, Zhou S, Zhang R, Li C, Wang Y. A taxonomy for advancing systematic error analysis in multi-site electronic health record-based clinical concept extraction. Journal of the American Medical Informatics Association. 2024 May 14:ocae101.
43. Faust L, Wilson P, Asai S, Fu S, Liu H, Ruan X, Storlie C. Considerations for Quality Control Monitoring of Machine Learning Models in Clinical Practice. JMIR Medical

Informatics. 2024 Jun 28;12(1):e50437.

44. Chowdhury S, Chen Y, Li P, Rajaganapathy S, Wen A, Ma X, Dai Q, Yu Y, Fu S, Jiang X, He Z. Stratifying heart failure patients with graph neural network and transformer using Electronic Health Records to optimize drug response prediction. Journal of the American Medical Informatics Association. 2024 Aug;31(8):1671-81.
45. Ren Y, Wu Y, Fan JW, Khurana A, Fu S, Wu D, Liu H, Huang M. Automatic uncovering of patient primary concerns in portal messages using a fusion framework of pretrained language models. Journal of the American Medical Informatics Association. 2024 Aug;31(8):1714-24.
46. Zhang K, Zhou R, Adhikarla E, Yan Z, Liu Y, Yu J, Liu Z, Chen X, Davison BD, Ren H, Huang J, Chen C, Zhou Y, Fu S, Liu W, Liu T, Li X, Chen Y, He L, Zou J, Li Q, Liu H, Sun L. A generalist vision–language foundation model for diverse biomedical tasks. Nature Medicine. 2024 Aug 7:1-3.
47. Wen A, Wang L, He H, Fu S, Liu S, Hanauer DA, Harris DR, Kavuluru R, Zhang R, Natarajan K, Pavinkurve NP. A Case Demonstration of the Open Health Natural Language Processing Toolkit From the National COVID-19 Cohort Collaborative and the Researching COVID to Enhance Recovery Programs for a Natural Language Processing System for COVID-19 or Postacute Sequelae of SARS CoV-2 Infection: Algorithm Development and Validation. JMIR Medical Informatics. 2024 Sep 9;12:e49997.
48. Tam TY, Sivarajkumar S, Kapoor S, Stolyar AV, Polanska K, McCarthy KR, Osterhoudt H, Wu X, Visweswaran S, Fu S, Mathur P. A framework for human evaluation of large language models in healthcare derived from literature review. NPJ Digital Medicine. 2024 Sep 28;7(1):258.

Full length, Peer-reviewed, Conference articles

1. Lash M, Fu S, Wang S, Zhao K. Early Prediction of Movie Success—What, Who, and When. InInternational Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction 2015 Mar 31 (pp. 345-349). Springer, Cham.
2. Rastegar-Mojarad M, Liu S, Wang Y, Afzal N, Wang L, Shen F, Fu S, Liu H. BioCreative/OHNLP Challenge 2018. InProceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics 2018 Aug 15 (pp. 575-575). ACM
3. Tafti AP, Fu S, Khurana A, Mastorakos GM, Poole KG, Traub SJ, Yiannias JA, Liu H. Artificial intelligence to organize patient portal messages: a journey from an ensemble deep learning text classification to rule-based named entity recognition. In2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2019 Nov 18 (pp. 1380-1387). IEEE.
4. Fu S, Carlson LA, Peterson KJ, Wang N, Zhou X, Peng S, Jiang J, Wang Y, Sauver JS, Liu H. Natural Language Processing for the Evaluation of Methodological Standards and Best Practices of EHR-based Clinical Research. AMIA Summits on Translational Science Proceedings. 2020;2020:171.
5. Zhao Y, Yu H, Fu S, et al. Data-driven Sublanguage Analysis for Cancer Genomics Knowledge Modeling: Applications in Mining Oncological Genetics Information from Patient's Genetic Reports. AMIA Summits on Translational Science Proceedings. 2020; 2020: 221
6. Chen D, Jiang J, Fu S, Demuth G, Liu S, Schaeferle GM, Wilson PM, Habermann E, Larson DW, Storlie C, Liu H. Early Detection of Post-Surgical Complications using Time-series Electronic Health Records. InAMIA Annual Symposium Proceedings 2021 (Vol. 2021, p. 152). American Medical Informatics Association.
7. Ibrahim OA, Fu S, Vassilaki M, Petersen RC, Mielke MM, St Sauver J, Sohn S. Early Alert of Elderly Cognitive Impairment using Temporal Streaming Clustering. In2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2021 Dec 9 (pp. 905-912). IEEE.
8. Fu S, Wen A, Schaeferle GM, Wilson PM, Demuth G, Liu S, Ruan X, Liu S, Storlie C, Liu H. Assessment of Data Quality Variability across Two EHR Systems through a Case Study of Post-Surgical Complications. InAMIA Annual Symposium Proceedings 2021. American Medical Informatics Association
9. Fu S, Ibrahim OA, Wang Y, Vassilaki M, Petersen RC, Mielke MM, St Sauver J, Sohn S.

Prediction of Incident Dementia Using Patient Temporal Health Status. Studies in health technology and informatics. 2022 Jun 6;290:757-61.

10. Fu S, Wen A, Pagali S, Zong N, St Sauver J, Sohn S, Fan J, Liu H. The Implication of Latent Information Quality to the Reproducibility of Secondary Use of Electronic Health Records. Stud Health Technol Inform. 2022 Jun 6;290:173-177. doi: 10.3233/SHTI220055. PMID: 35672994.
11. Fu S, Wen A, Schaeferle GM, Wilson PM, Demuth G, Ruan X, Liu S, Storlie C, Liu H. Assessment of Data Quality Variability across Two EHR Systems through a Case Study of Post-Surgical Complications. InAMIA Annual Symposium Proceedings 2022 (Vol. 2022, p. 196). American Medical Informatics Association.
12. Miller K, Fu S, Abah K, Escarria AM, Peterson K, Liu H, Hart L, Tan N, Huang M. Aspect-based Sentiment Analysis of Radiology Patient Experience Surveys: A Cohort Study. In2022 IEEE 10th International Conference on Healthcare Informatics (ICHI) 2022 Jun 11 (pp. 90-96). IEEE.
13. Ibrahim OA, Fu S, Vassilaki M, Mielke MM, St Sauver J, Petersen RC, Sohn S. Detection of Dementia Signals from Longitudinal Clinical Visits Using One-Class Classification. In2022 IEEE 10th International Conference on Healthcare Informatics (ICHI) 2022 Jun 11 (pp. 211-216). IEEE.
14. Liu Z, Garg M, Fu S, Sarkar S, Vassilaki M, Petersen RC, Sauver JS, Sohn S. Harnessing Transfer Learning for Dementia Prediction: Leveraging Sex-Different Mild Cognitive Impairment Prognosis. In2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2023 Dec 5 (pp. 2097-2100). IEEE.
15. Ibrahim OA, Garg M, Fu S, Vassilaki M, Petersen RC, St Sauver J, Sohn S. Dementia Prediction in Older Adults Using Sex-Specific Health Trajectory Clustering. Studies in Health Technology and Informatics. 2024 Jan 1;310:850-4.
16. Wang L, Lu Q, Li R, Fu S, Liu H. Wonder at Chemotimelines 2024: MedTimeline: An End-to-End NLP System for Timeline Extraction from Clinical Narratives. InProceedings of the 6th Clinical Natural Language Processing Workshop 2024 Jun (pp. 483-487).

Books chapters

- Fu, S., Wen, A., Liu, H. (2023). Clinical Natural Language Processing in Secondary Use of EHR for Research. In: Richesson, R.L., Andrews, J.E., Fultz Hollis, K. (eds) Clinical Research Informatics. Health Informatics. Springer, Cham. https://doi.org/10.1007/978-3-031-27173-1_21
- Social Computing, Behavioral-Cultural Modeling, and Prediction, 8th International Conference, SBP 2015, Washington, DC, USA, March 31-April 3, 2015. Proceedings, Early Prediction of Movie Success— What, Who, and When Michael Lash, Sunyang Fu**,** Shiyao Wang, Kang Zhao Pages 345-349

Other Professional Communications

International/National

- *Standardizing NLP-related contextual metadata for clinical and translational research*. OHDSI NLP WG at OHDSI Global Symposium 2024
- *Anticancer treatment trajectory and patterns in patients receiving chemotherapy for cancer*. OHDSI NLP WG at OHDSI Global Symposium 2024
- *Recommended practices and ethical considerations for natural language processing-assisted observational research: A scoping review*. AMIA NLP WG Webinar 2023
- *An information extraction framework for cohort identification using electronic health records*. AMIA Summits on Translational Science Proceedings. 2013;2013:149 (online).
- *A Case Study of Examining the Information Loss Among Three Clinical Data Sources for Ensuring Transparent and Safe AI Models*, Mayo Clinic AI Symposium – Rochester, MN 05/2019
- *Natural Language Processing for the Identification of Silent Brain Infarcts from Neuroimaging Reports*, San Francisco, CA, 10/2018

- *infoFACT: An Informatics Framework for Automated Data Abstraction*, San Francisco, CA, 10/2018
- Supervised Learning Approach to Link Prediction in FDA Adverse Event Reporting System (FAERS) Database Network, San Francisco, CA, 10/2018

Reginal

- Machine learning journal club AI project Show Case – "Assess the Heterogeneity of EHR-based Clinical Research, Rochester, MN, 12/2018
- Data Delivery Tools Coordination Group Update – "The Use of Informatics Tools for Clinical Research", Rochester, MN, 10/2018
- Machine learning journal club – "infoFACT: An informatics Framework for Data Abstraction from Clinical Text towards High-throughput and Reproducible EHR- based Research" Rochester, MN, 09/2018
- ADVANCE Seminar – "Patient Subtyping via Time-Aware LSTM Networks" Rochester, MN, 09/2017
- Department of Hematology, Health Outcome Group – "Determining the risk factors for monoclonal gammopathy of undetermined significance and predicting its progression to lymphoplasmacytic malignancies" Rochester, MN 08/2019