Hao Liu

Columbia University
Department of Biomedical Informatics
New York, NY 10032

Phone: (669) 224-0090
Email: hl2659@columbia.edu
Homepage: dr-haoliu.github.io

Research Interest

Data Mining, Natural Language Processing, Machine Learning, and Ontology Engineering, with their applications in clinical informatics

Academic Experience

Postdoctoral Research Scientist, Columbia University   Jun. 2020 - Present
Department of Biomedical Informatics
Mentor: Prof. Chunhua Weng

Education

Ph.D. in Computer Science, New Jersey Institute of Technology, USA   Jun. 2020
Advisor: Prof. Yehoshua Perl
Co-advisor: Prof. James Geller

M.S. in Electrical Engineering, Columbia University, USA   Dec. 2012

B.S. in Electrical and Computer Engineering, New York Institute of Technology, USA   May. 2011

Teaching Interests

Fundamental or intermediate courses in Computer Science and Data Science

Advanced courses in Natural Language Processing and Deep Learning

Teaching Experience

Teacher, Columbia University

G4099 Biomedical Informatics seminar   Spring 2021

Adjunct Professor, New York Institute of Technology

DTSC 115 - Introduction to Data Science (online course development and lecture)   Spring 2020
CSCI 415 - Introduction to Data Mining   Spring 2020
CSCI 636 - Big Data Analytic   Fall 2019

Teaching Assistant, New Jersey Institute of Technology

CS 115 - Introduction to C++	Spring 2018
CS 106 - Introduction to Python	Fall 2017, 2015
CS 610 - Data Structure & Algorithm	Spring 2017
CS 435 - Advance Data Structure & Algorithm Design	Spring 2017, Fall 2016
CS 431 - Introduction to Database System	Spring 2016

Publications

Journal Papers

  * *Note that **JAMIA** and **JBI** are leading health informatics journals with impact factors 7.942 and 8.000*

  [#]*: equal contribution first author*

1. **H. Liu**, S. Carini, Z. Chen, S. Phillips Hey, I. Sim, and C. Weng, "Ontology-based categorization of clinical studies by their conditions," *Journal of biomedical informatics (JBI)*, vol. 135, p. 104235, Nov. 2022.

2. **H. Liu**, Y. Chi, A. Butler, Y. Sun, and C. Weng, "A knowledge base of clinical trial eligibility criteria," *Journal of biomedical informatics (JBI)*, vol. 117, p. 103771, May 2021.

3. **H. Liu**, Y. Perl, and J. Geller, "Concept placement using BERT trained by transforming and summarizing biomedical ontology structure," *Journal of biomedical informatics (JBI)*, vol. 112, p. 103607, Dec. 2020.

4. L. Franks[#], **H. Liu**[#], M. S. V. Elkind, M. P. Reilly, C. Weng, and S. M. Lee, "Misalignment between COVID-19 hotspots and clinical trial sites," *Journal of the American Medical Informatics Association (JAMIA)*, vol. 28, no. 11, pp. 2461–2466, Oct. 2021.

5. Y. Fang, B. Idnay, Y. Sun, **H. Liu**, Z. Chen, K. Marder, H. Xu, R. Schnall, and C. Weng, "Combining human and machine intelligence for clinical trial eligibility querying," *Journal of the American Medical Informatics Association (JAMIA)*, vol. 29, no. 7, pp. 1161–1171, Jun. 2022.

6. Y. Sun, A. Butler, I. Diallo, J. H. Kim, C. Ta, J. R. Rogers, **H. Liu**, and C. Weng, "A Framework for Systematic Assessment of Clinical Trial Population Representativeness Using Electronic Health Records Data," *Applied Clinical Informatics*, vol. 12, no. 4, pp. 816–825, Aug. 2021.

7. L. Zheng, Y. Perl, Y. He, C. Ochs, J. Geller, **H. Liu**, and V. K. Keloth, "Visual comprehension and orientation into the COVID-19 CIDO ontology," *Journal of biomedical informatics (JBI)*, vol. 120, p. 103861, Aug. 2021.

8. Y. Sun, A. Butler, L. A. Stewart, **H. Liu**, C. Yuan, C. T. Southard, J. H. Kim, and C. Weng, "Building an OMOP common data model-compliant annotated corpus for COVID-19 clinical trials," *Journal of biomedical informatics (JBI)*, vol. 118, p. 103790, Jun. 2021.

9. Y. Sun, A. Butler, F. Lin, **H. Liu**, L. A. Stewart, J. H. Kim, B. R. S. Idnay, Q. Ge, X. Wei, C. Liu, C. Yuan, and C. Weng, "The COVID-19 Trial Finder," *Journal of the American Medical Informatics Association (JAMIA)*, vol. 28, no. 3, pp. 616–621, Mar. 2021.

10. J. H. Kim, C. N. Ta, C. Liu, C. Sung, A. M. Butler, L. A. Stewart, L. Ena, J. R. Rogers, J. Lee, A. Ostropolets, P. B. Ryan, **H. Liu**, S. M. Lee, M. S. V. Elkind, and C. Weng, "Towards clinical data-driven eligibility criteria optimization for interventional COVID-19 clinical trials," *Journal of the American Medical Informatics Association (JAMIA)*, vol. 28, no. 1, pp. 14–22, Jan. 2021.

11. L. Zheng, Y. Chen, H. Min, P. L. Hildebrand, **H. Liu**, M. Halper, J. Geller, S. de Coronado, and Y. Perl, "Missing lateral relationships in top-level concepts of an ontology," *BMC Medical Informatics and Decision Making*, vol. 20, no. Suppl 10, p. 305, Dec. 2020.

12. F. Kury, A. Butler, C. Yuan, L.-H. Fu, Y. Sun, **H. Liu**, I. Sim, S. Carini, and C. Weng, "Chia, a large annotated corpus of clinical trial eligibility criteria," *Scientific Data*, vol. 7, no. 1, p. 281, Aug. 2020.

13. G. Elhanan, C. Ochs, J. L. V. Mejino Jr, **H. Liu**, C. J. Mungall, and Y. Perl, "From SNOMED CT to Uberon: Transferability of evaluation methodology between similarly structured ontologies," *Artificial intelligence in medicine*, vol. 79, pp. 9–14, Jun. 2017.

Conference Papers

1. **H. Liu**, N. Moustafa-Fahmy, C. Ta, and C. Weng, "Can Race-sensitive Biomedical Embeddings Improve Healthcare Predictive Models?" *AMIA 2023 Informatics Summit*(accepted), March 2023.

2. A. Turfah, **H. Liu**, L. A. Stewart, T. Kang, and C. Weng, "Extending PICO with Observation Normalization for Evidence Computing," *Studies in Health Technology and Informatics*, vol. 290, pp. 268–272, Jun. 2022.

3. Z. Chen, **H. Liu**, S. Liao, M. Bernard, T. Kang, L. A. Stewart, and C. Weng, "Representation and Normalization of Complex Interventions for Evidence Computing," *Studies in Health Technology and Informatics*, vol. 290, pp. 592–596, Jun. 2022.

4. C. Liu, **H. Liu**, C. Ta, J. Roger, A. Butler, J. Lee, J. Kim, N. Shang, and C. Weng, "Evaluation of Criteria2Query: Towards Augmented Intelligence for Cohort Identification," *Studies in Health Technology and Informatics*, vol. 290, pp. 297–300, Jun. 2022.

5. F. Lin, **H. Liu**, P. Moon, and C. Weng, "A sample size extractor for RCT reports," *Studies in Health Technology and Informatics*, vol. 290, pp. 617–621, Jun. 2022.

6. Z. Chen, Y. Fang, **H. Liu**, and C. Weng, "Data-Driven Modeling of Randomized Controlled Trial Outcomes," *Studies in Health Technology and Informatics*, vol. 294, pp. 392–396, May 2022.

7. X. Li[#], **H. Liu**[#], F. Kury, C. Yuan, A. Butler, Y. Sun, A. Ostropolets, H. Xu, and C. Weng, "A Comparison between Human and NLP-based Annotation of Clinical Trial Eligibility Criteria Text Using The OMOP Common Data Model," *AMIA Summits on Translational Science Proceedings*, vol. 2021, pp. 394–403, May 2021.

8. Y. Fang, J. H. Kim, B. R. Idnay, R. Aragon Garcia, C. E. Castillo, Y. Sun, **H. Liu**, C. Liu, C. Yuan, and C. Weng, "Participatory Design of a Clinical Trial Eligibility Criteria Simplification Method," *Studies in Health Technology and Informatics*, vol. 281, pp. 984–988, May 2021.

9. Z. Chen, **H. Liu**, A. Butler, A. Ostropolets, and C. Weng, "Potential Role of Clinical Trial Eligibility Criteria in Electronic Phenotyping," *Studies in Health Technology and Informatics*, vol. 281, pp. 148–152, May 2021.

10. **H. Liu**, Y. Perl, and J. Geller, "Transfer Learning from BERT to Support Insertion of New Concepts into SNOMED CT," *AMIA Annual Symposium Proceedings*, vol. 2019, pp. 1129–1138, 2019.

11. L. Zheng, **H. Liu**, Y. Perl, and J. Geller, "Training a Convolutional Neural Network with Terminology Summarization Data Improves SNOMED CT Enrichment," *AMIA Annual Symposium Proceedings*, vol. 2019, pp. 972–981, 2019.

12. **H. Liu**, J. Geller, M. Halper, and Y. Perl, "Using Convolutional Neural Networks to Support Insertion of New Concepts into SNOMED CT," *AMIA Annual Symposium Proceedings*, vol. 2018, pp. 750–759, Dec. 2018.

13. L. Zheng, **H. Liu**, Y. Perl, J. Geller, C. Ochs, and J. T. Case, "Overlapping Complex Concepts Have More Commission Errors, Especially in Intensive Terminology Auditing," *AMIA Annual Symposium Proceedings*, vol. 2018, pp. 1157–1166, Dec. 2018.

14. **H. Liu**, P. L. Hildebrand, and Y. Perl, "Enrichment of SNOMED CT ophthalmology component to support EHR coding," *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018.

15. **H. Liu**, L. Zheng, Y. Perl, J. Geller, and G. Elhanan, "Can a convolutional neural network support auditing of nci thesaurus neoplasm concepts?" *International Conference on Biological Ontology* , 2018.

16. **H. Liu**, L. Chen, L. Zheng, Y. Perl, and J. Geller, "A Quality Assurance Methodology for ChEBI Ontology Focusing on Uncommonly Modeled Concepts," *International Conference on Biological Ontology* , 2018.

17. L. Zheng, C. Ochs, J. Geller, **H. Liu**, and Y. Perl, "Multi-layer Big Knowledge visualization scheme for comprehending neoplasm ontology content," *2017 IEEE International Conference on Big Knowledge (ICBK)*, 2017.

18. **H. Liu**, L. Zheng, Y. Perl, Y. Chen, and G. Elhanan, "Correcting Ontology Errors Simplifies Visual Complexity," *Studies in Health Technology and Informatics*, vol. 245, p. 1330, 2017.

19. **H. Liu**, Z. Dong, and H. Gu, "Microblogging as a social sensing tool," *Proceedings of the 11th IEEE International Conference on Networking, Sensing and Control*, 2014.

Papers in Submission

1. **H. Liu**, A. Soroush, J.G. Nestor, E. Park, B. Idnay, Y. Fang, J. Pan, S. Liao, M. Bernard, Y. Peng, and C. Weng, "CliVER: Toward Scientific Claim Verification on Clinical Trial Publications," *Journal of biomedical informatics (JBI)*, 2022. (in submission to special issue Semantics-enabled Biomedical Literature Analytics)

2. Y. Fang, **H. Liu**, B. Idnay, C. Ta, K. Marder, and C. Weng, "A Data-Driven Approach to Optimizing Clinical Study Eligibility Criteria," *Journal of biomedical informatics (JBI)*, 2022. (under review)

Honors

Honorable Mention for *LitCoin Natural Language Processing (NLP) Challenge*, National Center for Advancing Translational Sciences and National Institutes of Health (NIH) (2022)

Ph.D. awarded with full GPA (4.0), NJIT, Newark, NJ (2020)

B.S. awarded with President's list, NYIT, New York, NY (2011)

Presentations

Conference and workshop presentations

AMIA Joint Summits on Translational Science, "A Comparison between Human and NLP-based Annotation of Clinical Trial Eligibility Criteria Text Using The OMOP Common Data Model", Mar. 2021.

AMIA Annual Symposium, "Transfer Learning from BERT to Support Insertion of New Concepts into SNOMED CT", Nov. 2019.

AMIA Annual Symposium, "Using Convolutional Neural Networks to Support Insertion of New Concepts into SNOMED CT", Nov. 2018.

International Conference on Biological Ontology, "Can a convolutional neural network support auditing of nci thesaurus neoplasm concepts?", Aug. 2018.

Invited talks and university presentations

1. Seminar talk in the College of Engineering & Computing Sciences at **New York Institute of Technology**, "NLP-based knowledge sharing and evidence retrieval of clinical studies", March 29, 2022.
2. Guest lecture for HINF 5016: Natural Language Processing in Health at **Weill Cornell Medicine**, "NLP-based quality assurance of ontologies and knowledge sharing of clinical trials", March 9, 2022.

Grants and Contracts

As Principal Investigator

Google Cloud Credit Grant: A knowledge base for Clinical Trials: $5,000, 11/15/2021-11/14/2022

Involved in Proposal Writing

NIH R01: "RESCUE: Rare Disease Detection and Escalation Support via a Learning Health System", 2021 (Impact Score: 28)

Professional Activities

Scientific Program Committee

The 3rd International Workshop on Quality Assurance and Enrichment of Biological and Biomedical Ontologies and Terminologies
IEEE International Conference on Bioinformatics and Biomedicine (BIBM'2021)

Reviewer for journals

Journal of Biomedical Informatics (JBI)
Journal of the American Medical Informatics Association
Journal of the American Medical Informatics Association Open (JAMIA Open)
Applied Clinical Informatics (ACI)
BMC Supplements

BMC Medical Informatics and Decision Making
Clinical and Experimental Medicine
Scientific Reports

Reviewer for conferences

American Medical Informatics Association Annual Symposium (AMIA'19, '20, '21, '22)
American Medical Informatics Association Informatics Summit (AMIA'21, '22, '23)
The IEEE International Conference on Bioinformatics and Biomedicine (BIBM'21)

Last updated: January 2, 2023