

# Yingcheng Sun

University of North Carolina at Greensboro  
Department of Computer Science  
Greensboro, NC 27412

Phone: (216) 302-9575  
Email: [y\\_sun4@uncg.edu](mailto:y_sun4@uncg.edu)  
Homepage: [yingchengsun.github.io/academic](http://yingchengsun.github.io/academic)

## Research Interest

Natural Language Processing, Information Retrieval, and Machine Learning, with their applications in clinical informatics that help solving important healthcare problems

## Academic Experience

**Assistant Professor, University of North Carolina at Greensboro** Aug. 2022 - Present  
Department of Computer Science

**Postdoctoral Research Scientist, Columbia University** Jan. 2020 - Jun. 2022  
Department of Biomedical Informatics

## Education

**Ph.D. in Computer Science, Case Western Reserve University** Dec. 2019

**M.E. in Computer Science, Xidian University, China** Apr. 2013

**B.E. in Software Engineering, Xidian University, China** Jul. 2010

## Grants

1. 2026 Carolina Asia Center \$1,000
2. 2025 Chancellor's Initiative for Transformative Research (CITR) award \$75,000
3. 2025 Unreal Fellowship Research Fund \$5,000
4. 2025 Draelos Scholars Research Fund \$1,000
5. 2024 DjangoCon US Opportunity Grant \$1,000
6. 2024 Generative AI Implementation Grant \$1,500
7. 2023 UNCG Research Award \$5,000
8. 2021-2022 Google Cloud Credit Grant \$10,000

## Selected Publications

### Journal Papers

\* Note that **JAMIA** and **JB**I are leading health informatics journals with impact factor 10.2 and 7.942

1. McGuirt, J.T., Steeves, E.A., Labban, J.D., Pfammatter, A.F., Allen, K., Kopper, R., **Sun, Y.** and Gustafson, A., 2024. Multi-Method Formative Evaluation of a Digital Online Grocery Shopping Assistant Among Special Supplemental Nutrition Program for Women, Infants, and Children Participants. *Journal of Nutrition Education and Behavior*, 56(8), pp.556-568.
2. McGuirt, J., Steeves, B.A., Kopper, R., **Sun, Y.**, Edamadaka, V., Allen, K. and Gustafson, A., 2023. Rural Online Shopping Assistant (ROSA): Mixed Methods Formative Evaluation With Low-Income Rural Shoppers. *Journal of Nutrition Education and Behavior*, 55(7), pp.87-88.
3. Kang, T., **Sun, Y.\***, Kim, J.H., Ta, C., Perotte, A., Schiffer, K., Wu, M., Zhao, Y., Moustafa-Fahmy, N., Peng, Y. and Weng, C., EvidenceMap: a three-level knowledge representation for medical evidence computation and comprehension. *Journal of the American Medical Informatics Association (JAMIA)*, p.ocado36. 2023.(\*co-first author)
4. **Sun, Y.**, Butler, A., Diallo, I., Kim, J.H., Ta, C., Rogers, J.R., Liu, H. and Weng, C., A Framework for Systematic Assessment of Clinical Trial Population Representativeness Using Electronic Health Records Data. *Applied Clinical Informatics*, 12(04), pp.816-825, Sep. 2021.
5. **Sun, Y.**, Butler, A., Lin, F., Liu, H., Stewart, L.A., Kim, J.H., Idnay, B.R.S., Ge, Q., Wei, X., Liu, C. and Yuan, C., The COVID-19 Trial Finder. *Journal of the American Medical Informatics Association (JAMIA)*, 28(3), pp.616-621, Mar. 2021.
6. **Sun, Y.**, Butler, A., Stewart, L.A., Liu, H., Yuan, C., Southard, C.T., Kim, J.H. and Weng, C., Building an OMOP common data model-compliant annotated corpus for COVID-19 clinical trials. *Journal of biomedical informatics (JBI)*, 118, p.103790, Jun. 2021.
7. **Sun, Y.**, Guo, F., Kaffashi, F., Jacono, F.J., DeGeorgia, M. and Loparo, K.A., INSMA: An integrated system for multimodal data acquisition and analysis in the intensive care unit. *Journal of biomedical informatics (JBI)*, 106, p.103434, Jun. 2020.
8. **Sun, Y.**, Kolacinski, R. and Loparo, K., Transitive Topic Modeling with Conversational Structure Context: Discovering Topics that are Most Popular in Online Discussions. *International Journal of Semantic Computing*, 14(02), pp.273-293, 2020.
9. Fang, Y., Idnay, B., **Sun, Y.**, Liu, H., Chen, Z., Marder, K., Xu, H., Schnall, R., and Weng, C., Combining Human and Machine Intelligence for Clinical Trial Eligibility Querying, *Journal of the American Medical Informatics Association (JAMIA)*, 2022.
10. Liu, H., Chi, Y., Butler, A., **Sun, Y.** and Weng, C., A knowledge base of clinical trial eligibility criteria. *Journal of Biomedical Informatics (JBI)*, 117, p.103771, Apr. 2021.
11. Lee, J., Liu, C., Kim, J., Chen, Z., **Sun, Y.**, Rogers, J. and Weng, C., Deep Learning for Rare Disease: A Scoping Review. *Journal of Biomedical Informatics (JBI)*, 2022 Oct 17:104227.
12. Kim, J.H., Butler, A.M., Ta, C.N., **Sun, Y.**, Maurer, M.S. and Weng, C. The potential role of EHR data in optimizing eligibility criteria definition for cardiovascular outcome trials. *International Journal of Medical Informatics*, 156, p.104587, Dec. 2021.
13. Kury, F., Butler, A., Yuan, C., Fu, L.H., **Sun, Y.**, Liu, H., Sim, I., Carini, S. and Weng, C., Chia, a large annotated corpus of clinical trial eligibility criteria. *Scientific data*, 7(1), pp.1-11, Aug. 2020.

### Conference Papers

1. Kottam, S., Annem, S., **Sun, Y.** (2025, December). Error-Aware Text-to-SQL Generation for Clinical Trial Eligibility Criteria Querying in EHR Databases. In *2025 IEEE International Conference on Big Data (Big Data)* (pp. 3271–3280). IEEE.
2. Astraikis, J., Martin, R., Es, A. V., Kottam, S., **Sun, Y.** (2025, December). Benford-Constrained Multimodal Learning for Few-Shot Fake Image Detection. In *2025 IEEE International Conference on Big Data (Big Data)* (pp. 3940–3947). IEEE.
3. Hayes, K., **Sun, Y.** (2025, December). Modality-Aware Multimodal Fusion for Sleep Event Classification. In *2025 IEEE International Conference on Big Data (Big Data)* (pp. 4136–4141). IEEE.
4. **Sun, Y.**, Tang, J., Butler, A., Liu, C., Fang, Y. and Weng, C., Interactive Similarity-Based Search of Clinical Trials. *World Congress on Health and Biomedical Informatics (MEDINFO' 21)*, October 2-4, 2021
5. Fang, Y., Kim, J.H., Idnay, B.R., Garcia, R.A., Castillo, C.E., **Sun, Y.**, Liu, H., Liu, C., Yuan, C. and Weng, C., Participatory Design of a Clinical Trial Eligibility Criteria Simplification Method. In *31st Medical Informatics Europe Conference (MIE'21)*, May 29-31, 2021.
6. Li, X., Liu, H., Kury, F., Yuan, C., Butler, A., **Sun, Y.**, Ostropolets, A., Xu, H. and Weng, C., A Comparison between Human and NLP-based Annotation of Clinical Trial Eligibility Criteria Text Using The OMOP Common Data Model. In *AMIA Annual Symposium Proceedings (AMIA'21)*, November 14 - 18, 2021.
7. **Sun, Y.**, Kolacinski, R. and Loparo, K., Eliminating search intent bias in learning to rank. In *2020 IEEE 14th International Conference on Semantic Computing (ICSC'20)*, IEEE, February 3-5, 2020
8. **Sun, Y.**, Loparo, K. and Kolacinski, R., Conversational structure aware and context sensitive topic model for online discussions. In *2020 IEEE 14th International Conference on Semantic Computing (ICSC'20)*, IEEE, , February 3-5, 2020. (Best paper)
9. **Sun, Y.** and Loparo, K., Opinion Spam Detection Based on Heterogeneous Information Network. In *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI'19)*, IEEE, Nov 4-6, 2019. (Acceptance Ratio =  $160/617 = 26\%$ )
10. **Sun, Y.**, Cai, X. and Loparo, K., Learning - based Adaptation Framework for Elastic Software Systems. In *2019 IEEE 31st International Conference on Software Engineering and Knowledge Engineering (SEKE'19)*, IEEE, Jul 10-12, 2019.
11. **Sun, Y.** and Loparo, K., Topic Shift Detection in Online Discussions using Structural Context. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC'19)*, IEEE, Jul 15-19, 2019.
12. **Sun, Y.** and Loparo, K., Information Extraction from Free Text in Clinical Trials with Knowledge-Based Distant Supervision. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC'19)*, IEEE, Jul 15-19, 2019.
13. **Sun, Y.** and Loparo, K., A Clicked-URL Feature for Transactional Query Identification. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC'19)*, IEEE, Jul 15-19.
14. **Sun, Y.** and Loparo, K., Context Aware Image Annotation in Active Learning with Batch Mode. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC'19)*, IEEE, Jul. 2019.
15. **Sun, Y.** and Loparo, K., A Common Gene Expression Signature Analysis Method for Multiple Types of Cancer. In *2019 19th Industrial Conference on Data Mining*, Springer, Jul 17-21, 2019.

16. **Sun, Y.** and Loparo, K., 2019, July. Knowledge-guided Text Structuring in Clinical Trials. In 2019 *19th Industrial Conference on Data Mining*, Springer, Jul 17-21, 2019.
17. **Sun, Y.** and Loparo, K., Context Aware Image Annotation in Active Learning. In 2019 *19th Industrial Conference on Data Mining*, Springer, Jul 17-21, 2019.

### Patents

1. Li, Q., **Sun, Y.**, Zou, Y., Xu, J. and Xi, B., 2012, Agent Based Intelligent Meta Search Engines System, CN Patent No. 102902800 A.
2. Li, Q., Zou, Y., **Sun, Y.**, Xu, J. and Xi, B., 2013. A Personalized Method of Agent Based Meta Search Engines, CN Patent No.103593413 A.

## Teaching Experience

### University of North Carolina at Greensboro

CSC 429/629 Artificial Intelligence	Spring 2024
CSC 410/610 Big Data and Machine Learning	Fall 2025, Spring 2023, Fall 2022
CSC 446/646 Natural Language Processing	Spring 2025, Spring 2024, Spring 2023
CSC 454/654 Algorithm Analysis and Design	Fall 2024, Fall 2023
CSC 490 Senior Capstone	Fall 2025, Spring 2025, Fall 2024, Fall 2023

## Invited talks and Press Coverage

1. "Augmented Intelligence for Clinical NLP: From Clinical Trial Optimization to Evidence Appraisal", Texas Tech University, Mar 14, 2025.
2. "AI in Healthcare: Transforming Medicine with Intelligent Technology", the retirement community Friends Homes, Greensboro, NC, Nov 21, 2023.
3. "Augmented Intelligence for Evidence-based Medicine: Optimizing Clinical Trial Eligibility Criteria Design", AI-Enhanced Clinical Trials Summit, Boston, Dec. 6, 2022.
4. "Improving Clinical Information Retrieval for Medical-based Evidence", University of California San Diego, Nov 13, 2021.
5. "Towards Improved Information Retrieval for Evidence-based Medicine", Health Informatics Conference (JIS Go Live 2021), Nov 15, 2021.
6. "COVID-19 trial finder provides simplified search process for COVID-related clinical trials" (*EurekaAlert*, 05,2020; *Medical Press*, 05,2020; *Columbia DBMI*, 05,2020)

## Professional Activities

**Reviewer:** Journal of Biomedical Informatics (JBI), Journal of the American Medical Informatics Association Open (JAMIA Open), Applied Clinical Informatics (ACI), International Journal of Medical Informatics (IJMI), American Medical Informatics Association Annual Symposium (AMIA'20, AMIA'21, AMIA'22), KDD'26