

# Haihua Chen, Ph.D.

Clinical Assistant Professor, Data Science

Department of Information Science, University of North Texas  
3940 North Elm, Suite E292 Denton, Texas 76203-5017

Email: [haihua.chen@unt.edu](mailto:haihua.chen@unt.edu)  
[Google Scholar](#)

## Research Interests and Expertise

---

- Data science: applied data science, applied machine learning
- Data quality: data quality evaluation and improvement, data quality in machine learning
- Natural language processing: information retrieval, recommendation system, legal NLP
- Data mining: academic data mining, innovation measurement, social media analysis
- Informatics: health informatics, legal informatics

## Education

---

|             |   |
|-------------|---|
| 2017 – 2021 | <b>University of North Texas</b><br>Ph.D. in Information Science concentrated on Data Science (Advisors: Dr. Jiangping Chen & Dr. Junhua Ding)<br>Dissertation Title: Data Quality Evaluation and Improvement for Machine Learning<br>Great Graduates |
| 2014 – 2017 | <b>Wuhan University</b><br>M.S. in Information Science (Advisor: Dr. Wei Lu)<br>Fellowship for Graduate Students  |
| 2010 – 2014 | <b>Central China Normal University</b><br>B.S. in Information Science & B.S. in English Literature<br>Honor Graduates   |

## Professional Experience

---

|                 |   |
|-----------------|---|
| 01/2022-present | <b>Clinical Assistant Professor</b> , University of North Texas                                   |
| 01/2019-01/2022 | <b>Research Assistant &amp; Teaching Fellow</b> , University of North Texas                       |
| 05/2018-01/2022 | <b>Lab Manager</b> , Intelligent Information Access Lab, University of North Texas                |
| 10/2020-08/2021 | <b>Vice President</b> , Data Science Organization, University of North Texas                      |
| 01/2018-01/2020 | <b>Officer</b> , International Graduate Students Research Organization, University of North Texas |
| 09/2017-01/2019 | <b>Research Assistant &amp; Teaching Assistant</b> , University of North Texas                    |
| 09/2014-06/2017 | <b>Graduate Research Assistant</b> , Wuhan University   |

## Department Service

---

|                 |  |
|-----------------|--|
| 09/2022-present | <b>Faculty Search Committee</b> , Data Science Program and Health Informatics Program, University of North Texas |
| 09/2019-01/2022 | <b>Faculty-Staff-Student Council</b> , Department of Information Science, University of North Texas              |

## Publications

---

(\*) indicates the corresponding author

### Articles In-progress

- [1] TBD Feng, Y., Qiu, M., Xuan, L., Tao, D., **Chen, H.**, & Ding, J. Exploring Fairness and Stereotype of Skin-toned Emoji Representations in Pre-trained Embeddings.
- [2] TBD Nguyen, H., Maganti, R., **Chen, H.**, Hossain, T., & Ding, J. Identifying High-quality Informative Comments for Software Review Summarization. IEEE AITest 2023.
- [3] TBD **Chen, H.**, Chen, J., & Ding, J. Data Quality in Machine Learning and Big Data: A Comprehensive Survey.
- [4] TBD Aledhari, M., **Chen, H.**, Hossain, T., & Ding, J. Auto Grader for Unstructured Open-Ended Questions Using Sentence Bidirectional Encoder Representation from Transformers.
- [5] TBD Madsen, M., Nguyen, H., Ding, J., & **Chen, H.\***. Evaluating Text Summarization: Dimensions, Metrics, Approaches, and Challenges. IEEE AITest 2023.

### Articles Under Review

- [1] *J. King Saud Univ.* Wang, Z., Peng, S., Huang, J., Chen, J., & **Chen, H.\***. ICDIA-NCM: A Framework for Interdisciplinary Concept Drift Identification and Analysis based on the Normal Cloud Model. *Journal of King Saud University - Computer and Information Sciences*. Under review. IF = 9.006
- [2] *ESWA* Wang, Z., Zhang, H., Chen, J., & **Chen, H.\***. Measuring the novelty of scientific literature through contribution sentence analysis using deep learning and cloud model. *Expert Systems with Applications*. Minor revision. IF = 8.655
- [3] *Knowledge-Based Systems* Wang, Z., Peng, S., Chen, J., Zhang, X., & **Chen, H.\***. ICAD-MI: Interdisciplinary concept association discovery from the perspective of metaphor interpretation. *Knowledge-Based Systems*. Under review. IF = 8.139
- [4] *SCM* Wang, Z., Chen, J., Chen, J., & **Chen, H.\***. Identifying interdisciplinary topics and their evolution using deep learning. *Scientometrics*. Under review. IF = 3.238
- [5] *JDE* **Chen, H.**, Cleveland, A., Katragadda, N., Maddi, S., Chung, H., & Phan, T. (2020). Dental Informatics over time: A 20-years longitudinal bibliometrics analysis. *Journal of Dental Education*. Under review. IF = 2.313
- [6] *IEEE Intelligent Systems* Feng, Y., Qiu, M., Zhou, W., **Chen, H.**, & Ding, J. Deep Learning Based Intelligent System for Ancient Character Recognition. *IEEE Intelligent Systems*. Under review. IF = 6.744.
- [7] *ACM Computing Surveys* Huang, J., **Chen, H.**, Yu, F., & Lu, W. From Detection to Application: Recent Advances in Understanding Scientific Tables and Figures. *Review paper*. Under review. IF = 14.32.
- [8] *DHQ* Han, Y., Maganti, R., Kim, J., Chen, J., & **Chen, H.\***. "The Long Journey Home": A Case Study of Learning Japanese American Internment History Using Natural Language Processing. *Digital Scholarship in the Humanities*. Under review.
- [9] *IEEE COMPSAC 2023* Feng, Y., Vanam, S., Cherukupally, M., Zheng, W., Qiu, M., & **Chen, H.** Investigating Code Generation Performance of ChatGPT with Crowdsourcing Social Data. Under review.

### Refereed Journal Articles

- [J16] *IDD* Nguyen, H., **Chen, H.**, Chen, J., Kargozari, K., & Ding, J. Construction and Evaluation of a Domain-Specific Knowledge Graph for Knowledge Discovery. *Information Discovery & Delivery*. Accepted. IF = 1.942
- [J15] *IEEE Access* Tran, N., **Chen, H.**, Bhuyan, J., & Ding, J. (2022). Data Curation and Quality Evaluation for Machine Learning-Based Cyber Intrusion Detection. *IEEE Access*. IF = 3.367
- [J14] *JBI* Zhang, L., Lu, W., **Chen, H.**, Huang, Y., & Cheng, Q. (2022). A comparative evaluation of biomedical similar article recommendation. *Journal of Biomedical Informatics*, 104106. IF = 8.00

- [J13] *IEEE TR* **Chen, H.**, Piepeta, L. F., & Ding, J. (2022). Construction and Evaluation of a High-Quality Corpus for Legal Intelligence Using Semiautomated Approaches. *IEEE Transactions on Reliability*, 71(2), 657-673. IF = 4.424
- [J12] *IP&M* **Chen, H.**, Wu, L., Chen, J., Lu, W., & Ding, J. (2022). A comparative study of automated legal text classification using random forests and deep learning. *Information Processing & Management*, 59(2), 102798. IF = 7.466
- [J11] *SCM* **Chen, H.\***, Nguyen, H., & Alghamdi, A. (2022). Constructing a high-quality dataset for automated creation of summaries of fundamental contributions of research articles. *Scientometrics*, 1-15. IF = 3.238
- [J10] *SCM* Wang, Z., Wang, K., Liu, J., Huang, J., & **Chen, H.\*** (2022). Measuring the innovation of method knowledge elements in scientific literature. *Scientometrics*, 127(5), 2803-2827. IF = 3.238
- [J9] *SCM* Zhang, Y., Zhao, R., Wang, Y., **Chen, H.**, Mahmood, A., Zaib, M., ... & Sheng, Q. Z. (2022). Towards employing native information in citation function classification. *Scientometrics*, 1-21. IF = 3.238
- [J8] *IEEE TR* **Chen, H.**, Chen, J., & Ding, J. (2021). Data evaluation and enhancement for quality improvement of machine learning. *IEEE Transactions on Reliability*, 70(2), 831-847. IF = 4.424
- [J7] *IJPE* Tran, N., **Chen, H.**, Jiang, J., Bhuyan, J., & Ding, J. (2021). Effect of Class Imbalance on the Performance of Machine Learning-based Network Intrusion Detection. *International Journal of Performability Engineering*, 17(9). IF = 0.437
- [J6] *JDIS* **Chen, H.\*** (2021). A New Citation Recommendation Strategy Based on Term Functions in Related Studies Section. *Journal of Data and Information Science*, 6(3), 75-98. IF = 1.771
- [J5] *JMLA* **Chen, H.**, Chen, J., & Nguyen, H. (2021). Demystifying COVID-19 publications: institutions, journals, concepts, and topics. *Journal of the Medical Library Association: JMLA*, 109(3), 395. IF = 3.18
- [J3] *TEL* **Chen, H.**, Yang, Y., Lu, W., & Chen, J. (2020). Exploring multiple diversification strategies for academic citation contexts recommendation. *The Electronic Library*, 38(4), 821-842. IF = 1.453
- [J3] *TEL* Tang, M., Chen, J., **Chen, H.**, Xu, Z., Wang, Y., Xie, M., & Lin, J. (2020). An ontology-improved vector space model for semantic retrieval. *The Electronic Library*, 38(5/6), 919-942. IF = 1.453
- [J2] *Information Sciences* Lu, W., Luo, M., Zhang, Z., Zhang, G., Ding, H., **Chen, H.**, & Chen, J. (2019). Result diversification in image retrieval based on semantic distance. *Information Sciences*, 502, 59-75. IF = 8.233
- [J1] *IDD* Zhang, Z., **Chen, H.\***, & Xiao, B. (2019). Understanding eWOM of Chinese Governments information service: a perceived value-based perspective. *Information Discovery & Delivery*, 47(4), 251-258. IF = 1.942

## Refereed Conference Papers

- [C10] *AEJMC Midwinter Conference 2023* Zhang, X., **Chen, H.\***. Assessing the Change of Country Reputation during the COVID-19 Pandemic Period on Twitter: A Case Study of China (Long paper). *The Association for Education in Journalism and Mass Communication Midwinter Conference 2023*. Accepted.
- [C9] *iConference'23* Nguyen, H., Oladapo, L., Ali, I., **Chen, H.**, & Chen, J. Fighting Misinformation: Where Are We and Where to Go? (Long paper). *iConference 2023*. Accepted.
- [C8] *ISSI'21* **Chen, H.\***, & Kanuboddu, B. N. (2021). A fine-grained annotation scheme for research contribution in academic literature. In *Proceedings of the 18th International Conference on Scientometrics and Informetrics* (pp. 241-248).
- [C7] *ASIS&T'21* Chong, M., & **Chen, H.\*** (2021). Racist Framing through Stigmatized Naming: A Topical and Geo - locational Analysis of# Chinavirus and# Chinesevirus on Twitter. In *Proceedings of the association for information science and technology*, 58(1), 70-79.
- [C6] *IEEE QRS'20* **Chen, H.**, Chen, J., & Ding, J. (2020). Data Evaluation and Enhancement for Quality Improvement of Machine Learning. *2020 IEEE 20th International Conference on Software Quality, Reliability and Security (QRS)*, 13-13. (**Best Paper Award Nominate**)

- [C5] CCKS'19 **Chen, H.**, Cao, G., Chen, J., & Ding, J. (2019, August). A practical framework for evaluating the quality of knowledge graph. In *China conference on knowledge graph and semantic computing* (pp. 111-122). Springer, Singapore.
- [C4] IEEE BIBE'18 Ding, J., Jin, W., & **Chen, H.** (2018, October). Regression-Based Documents Reranking for Precision Medicine. In *2018 IEEE 18th International Conference on Bioinformatics and Biengineering (BIBE)* (pp. 283-286). IEEE.
- [C3] ICSH'18 **Chen, H.**, Ding, J., Chen, J., & Cao, G. (2018, July). Designing a novel framework for precision medicine information retrieval. In *International Conference on Smart Health* (pp. 167-178). Springer, Cham.
- [C2] ICSH'18 Chen, J., Chen, M., Qu, J., **Chen, H.**, & Ding, J. (2018, July). Smart and connected health projects: Characteristics and research challenges. In *International Conference on Smart Health* (pp. 154-164). Springer, Cham.
- [C1] ICKM'17 Zhang, Q., Lu, W., Yang, Y., **Chen, H.**, & Chen, J. (2017). Automatic identification of research articles containing data usage statements. In *Knowledge Discovery and Data Design Innovation: Proceedings of the International Conference on Knowledge Management (ICKM 2017)* (pp. 67-87).

## Workshops, Posters, Blogs

- [W13] JCDL 2023 Wang, Z., **Chen, H.\***, Zhang, C., Lu, W., Wu, J. JCDL2023 Workshop: Innovation Measurement for Scientific Communication (IMSC) in the Era of Big Data. *Workshop proposal*. Accepted.
- [W12] JCDL 2023 Zhang, C., Zhang, Y., Mayr, P., Lu, W., Suominen, A., **Chen, H.**, Ding, Y. JCDL2023 Workshop: Joint Workshop of the 4th Extraction and Evaluation of Knowledge Entities from Scientific Documents (EEKE2023) and the 3rd AI + Informetrics (AII2023). *Workshop proposal*. Accepted.
- [W11] *Information Matters* **Chen, H.** (2022, September 23). Why does data quality matter? How to evaluate and improve data quality for machine learning systems. *Information Matters*, 2(9).
- [W10] EEKE@JCDL'22 Wang, Z., Chen, J., Chen, J., & **Chen, H.\*** (2022). Interdisciplinary Topics Extraction and Evolution Analysis. In *Proceedings of 3rd Workshop on Extraction and Evaluation of Knowledge Entities from Scientific Documents (EEKE2022)*, June 20-24, 2022, Cologne, Germany and Online.
- [W9] ASIS&T'22 Han, Y., Maganti, R., Chen, J., & **Chen, H.\*** (2022). Uncover Marginalized Narratives of Japanese American Incarceration: An Annotation Scheme for Natural Language Processing and Data Analytics. *Proceedings of the Association for Information Science and Technology*.
- [W8] ASIS&T'21 You, J., Shaik, N., & **Chen, H.\*** (2021). Data Mining on COVID - 19 Vaccines: Side Effects. *Proceedings of the Association for Information Science and Technology*, 58(1), 869-871.
- [W7] TREC'21 Nguyen, H., **Chen, H.**, Prasad, B., Zhao, H., Ding, J., Chen, J., & Cleveland, A. (2021). UNTIIA Lab at TREC 2021 - Clinical Trial. In *TREC*.
- [W6] IEEE Big Data'20 Tang, M., Su, C., **Chen, H.**, Qu, J., & Ding, J. (2020, December). SALKG: A Semantic Annotation System for Building a High-quality Legal Knowledge Graph. In *2020 IEEE International Conference on Big Data (Big Data)* (pp. 2153-2159). IEEE.
- [W5] ASIS&T'20 Tang, M., Chen, J., & **Chen, H.** (2020). Smart bookshelf for library book management. *Proceedings of the Association for Information Science and Technology*, 57(1), e396.
- [W4] IEEE QRS'20 Luo, M., Yu, F., & **Chen, H.\*** (2020, December). Result Diversification in Clinical Case Reports Retrieval based on Main Finding. In *2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C)* (pp. 213-220). IEEE.
- [W3] IEEE QRS'20 Tang, M., Chen, J., & **Chen, H.** (2020, December). SemOIR: an ontology-based semantic information retrieval system. In *2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C)* (pp. 204-208). IEEE.
- [W2] JCDL'18 Yang, Y., **Chen, H.\***, Lu, W., & Ayala, B. R. (2018, May). Diversifying citation contexts in academic literature for knowledge recommendation. In *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries* (pp. 397-398).
- [W1] TREC'17 Viswavarapu, L. K., Chen, J., Cleveland, A. D., & **Chen, H.** (2017). UNT Precision Medicine Information Retrieval at TREC 2017. In *TREC*.

## Research Grants

- Co-PI (PI: Dr. Mohammed Aledhari). Auto Grader for Unstructured Open-Ended Questions using NLP and Multimodal ML Transformer-based Methods. In preparation.
- Co-PI (PI: Dr. Yunhe Feng). NSF Expanding AI Innovation through Capacity Building and Partnerships (ExpandAI), National Science Foundation. In preparation.
- Co-PI (PI: Dr. Sharad Sharma). NSF Build and Broaden (B2) *Enhancing Social, Behavioral and Economic Science Research and Capacity at Minority-Serving Institutions*, National Science Foundation. Under review.
- Co-PI (PI: Dr. Sharad Sharma). NSF I/UCRC Planning Grant “*Center for Emerging System Dependability Assurance (CESDA)*”, National Science Foundation. Under review.
- Co-PI (PI: Dr. Xiaoqun Zhang). The UNT VPRI Intramural Research Seed Grant (RSG) “*Assessing Country Reputation during the COVID-19 Pandemic Period on Twitter: A Case Study of China*”. Amount: \$10,000. Under review.
- 2023-2025. Faculty Mentor (PI: Dr. Ting Xiao). NSF REU Site “*Beyond Language: Training to Create and Share Vector Embeddings across Applications*”, National Science Foundation. Amount: \$398,000. Under review.
- 2022-2025. Co-PI (PI: Dr. Junhua Ding). NSF HSI Implementation and Evaluation Project “*Developing a High-Quality Academic Environment for Broadening Participation of Hispanic Students in Computing*”, National Science Foundation. Award Amount: \$499,608.
- 2022. PI. UNT COI Seed Funding “*Towards a large-scale and high-quality corpus for legal argument mining*”, College of Information, University of North Texas. Award Amount: \$9,975.
- 2021. PI. Google Cloud Research Credits for project “*Building a knowledge-driven Legal IR, Automatic Evidence and Argument Generation System for Decision Making*”, Google. Award Amount: \$5,000.
- 2021. PI. Google Cloud Education Grant - google cloud platform credits for course “*INFO 5731: Computational Methods for Information Systems*”. Google. Award Amount: \$1,550.
- 2020. PI. Google Cloud COVID-19 Research Credits Program - google cloud platform credits for project “*Demystifying COVID-19 publications*”. Award Amount: \$13,800.
- 2020. PI. UNT Graduate Student Research Award. Award Amount: \$800.
- 2020. PI. Google Cloud Education Grant - google cloud platform credits for course “*INFO 5731: Computational Methods for Information Systems*”. Award Amount: \$2,800.
- 2019. PI. Google Cloud Education Grant - google cloud platform credits for course “*INFO 4206: Information Retrieval*”. Award Amount: \$5,000.
- 2019-2022. Co-PI (PI: Dr. Sisi Gui). NSSFC project “*Query Intent for Academic Search*”, National Social Science Foundation of China. Award Amount: \$28,000.
- 2018-2020. Co-PI (PI: Dr. Qikai Cheng). NSFC project “*Diversifying Citation Recommendation based on Semantic Analysis of Scholar Text*”, National Natural Science Foundation of China. Award Amount: \$28,000.
- 2017-2020. Co-PI (PI: Dr. Wei Lu). NSFC project “*Multi-Semantic Information Fusion based Citation Recommendation*”, National Natural Science Foundation of China (NSFC). Award Amount: \$72,000.

## Travel Fundings

---

- 2022. UNT Department of Information Science Travel Grant (ASIS&T 2022). Award Amount: \$1,000.
- 2022. UNT Department of Information Science Travel Grant (IEEE CISOSE 2022). Award Amount: \$1,200.
- 2021. UNT Travel Grant (ASIS&T 2021). Award Amount: \$500.
- 2021. UNT Department of Information Science Travel Grant (ASIS&T 2021). Award Amount: \$1,000.
- 2021. UNT Department of Information Science Travel Grant (TREC 2021). Award Amount: \$35.
- 2021. The International Society for Scientometrics and Informetrics (ISSI) Student Travel Award. Award Amount: €115.
- 2020. UNT Department of Information Science Travel Grant (IEEE QRS 2020). Award Amount: \$1,500.
- 2019. UNT Department of Information Science Travel Grant (WWW 2019). Award Amount: \$600.
- 2018. UNT Travel Grant (TREC 2018). Award Amount: \$1,200.
- 2018. UNT Department of Information Science Travel Grant (ALA 2018). Award Amount: \$800.
- 2018. UNT Department of Information Science Travel Grant (IEEE/ACM JCDL 2018). Award Amount: \$1,000.
- 2017. UNT Department of Information Science Travel Grant (ICKM 2017). Award Amount: \$300.

## Honors and Awards

---

- 2022. Great Grads Spring 2022, College of Information, University of North Texas.
- 2021 - 2022. UNT COVID-19 Student Success Award (Spring 2022, Summer 2021, Fall 2021), University of North Texas. Award Amount: \$1,500.
- 2021. UNT Toulouse Graduate School Summer Award, University of North Texas. Award Amount: \$1,605.
- 2021. Linda Schamber Writing Award for paper *"Topic sentiments toward the COVID-19 vaccine on Twitter"*, College of Information, University of North Texas. Award Amount: \$250.
- 2021. The Melba S. Harvill Endowed Scholarship, Department of Information Science, University of North Texas. Award Amount: \$1,000.
- 2020. The Dewey E. Carroll Graduate Fellowship Award, Department of Information Science, University of North Texas. Award Amount: \$4,000.
- 2020. The Donald B. and Ana D. Cleveland Medical Informatics Endowed scholarship, Department of Information Science, University of North Texas. Award Amount: \$1,000.
- 2019. The Mark E. Rorvig Endowed Graduate Fellowship Award, Department of Information Science, University of North Texas. Award Amount: \$4,000.
- 2019. Outstanding Reviewer for The Electronic Library, Emerald.
- 2019. The Donald B. and Ana D. Cleveland Medical Informatics Endowed scholarship, Department of Information Science, University of North Texas. Award Amount: \$1,000.
- 2018. The Howard Griesdorf Interdisciplinary Ph.D. Award, Department of Information Science, University of North Texas. Award Amount: \$1,000.
- 2017. Honor Graduates, Excellent Student Leader Award, Wuhan University.

## Teaching

---

### **Instructor**, University of North Texas

- INFO 5810: Data Analysis and Knowledge Discovery (Summer 2022, Fall 2022, Spring 2023, Summer 2023)

### **Instructor**, University of North Texas

- INFO 5731: Computational Methods for Information Systems (Spring 2020 - Spring 2023)

### **Instructor**, University of North Texas

- INFO 5502: Analytic Tools, Techniques and Methods (Spring 2022, Summer 2022)

### **Instructor**, University of North Texas

- INFO 4206: Information Retrieval (Spring 2019)

### **Course Designer**, University of North Texas

- INFO 5506: Applications of Artificial Intelligence in Health (Spring 2024)

### **Course Designer**, University of North Texas

- CSCE 4290/5290: Introduction to Natural Language Processing (CIB) (Fall 2022)

### **Co-Instructor**, University of North Texas

- INFO 5717: Networked Data Modeling and Processing (Fall 2019)

### **Guest Lecturer**, University of North Texas

- LSCM 3960: Logistics and Supply Chain Management – Topic: The Applications of Machine Learning and Natural Language Processing for Supply Chain Management (Spring 2022)
- INFO 5709: Data Visualization and Communication – Topic: Deep Learning with TensorFlow and Keras (Summer 2020)
- INFO 3020: Introduction to Computation with Python – Topic: Web Scraping with Python (Summer 2020)
- INFO 5707: Data Modeling for Information Professionals – Topic: Construction and Applications of Knowledge Graphs (Fall 2019)
- INFO 5810: Data Analysis and Knowledge Discovery – Topic: Web Scraping with Python (Fall 2019)

- INFO 5501: Fundamentals of Data Analytics - Topic: Introduction to Python Tools and Packages for Data Analysis and Scientific Computing (Spring 2018)

### Teaching Assistant, University of North Texas

- INFO 5717: Dynamic WWW Structures (Fall 2018)

## Invited Talk and Presentations

---

- Talk Title: Informatics applications in multiple fields: A bibliometrics analysis, Department of Information Science, University of North Texas, Nov 16, 2022
- Talk Title: Measuring the novelty of scientific literature through contribution sentences analysis, Information Processing & Management Conference 2022, Oct 22, 2022
- Talk Title: Fine-grained semantic analysis for academic data mining and evaluation, National Science Library, University of Chinese Academic of Science, May 31, 2022
- Talk Title: The applications of machine learning and natural language processing in healthcare and medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Mar 3, 2022
- Talk Title: Building high-performance machine learning systems for domain-specific applications, Department of Information Science, University of North Texas, Dec 21, 2021
- Talk Title: Construction and evaluation of high-quality corpus for legal intelligence using semi-automated approaches, Department of Information Science, University of North Texas, Nov 17, 2021
- Talk Title: Building high-performance machine learning systems for domain-specific applications, Sun Yat-Sen University (Online), Oct 24, 2021
- Talk Title: Information retrieval in 40 years: a bibliometric analysis of SIGIR proceedings, Department of Information Science, University of North Texas, Sep 1, 2021
- Talk Title: Information quality improvement and evaluation for building high performance machine learning systems, Department of Information Science, University of North Texas, Oct 7, 2020
- Talk Title: Information quality improvement and evaluation for building high performance machine learning systems, Wuhan University, Jul 9, 2020
- Talk Title: Knowledge graph construction, application and evaluation, Beihua University (Online), May 19, 2020
- Talk Title: Construction and applications of knowledge graphs, Department of Information Science, University of North Texas, Nov 6, 2019
- Talk Title: Research experience sharing in 2019 new Ph.D. student orientation, Department of Information Science, University of North Texas, Aug 24, 2019
- Talk Title: Diversifying the citation contexts in academic literature for knowledge recommendation, Central China Normal University, Jan 10, 2019

## Professional Service

---

- Co-editor of The Electrical Library (2022-present)
- Editorial board member:
  - Heliyon Information Science (2022-present)
  - American Journal of Information Science and Technology (2020-present)
- Guest editor:
  - The Electrical Library special issue (Innovation Measurement for Scientific Communication in the Era of Big Data), 2023 - 2024
  - Information Discovery & Delivery special issue (Information and Data Quality for Intelligent Systems), ongoing.
  - Frontiers in Big Data special issue (Data Quality for Big Data and Machine Learning), ongoing.
- Co-chair of Workshops: DQIS workshop@ IEEE QRS 2021, 2022
- Journal reviewer:
  - The Electrical Library (2017-present)
  - IEEE Access (2018-present)
  - Information Discovery and Delivery (2018-present)

- Scientometrics (2021-present)
- Computer Standards & Interfaces (2021-present)
- JAMIA Open (2021-present)
- Journal of Computational Social Science (2022-present)
- Aslib Journal of Information Management (2022-present)
- Applied Intelligence (2022-present)
- ACM Transactions on Asian and Low-Resource Language Information Processing (2022-present)
- Artificial Intelligence Review (2022-present)
- Information Processing and Management (2022-present)
- Journal of Informetrics (2022-present)
- Information Resources Management Journal (2022 - present)
- Frontiers in Public Health (2022-present)
- Neural Processing Letters (2022-present)
- Machine Learning (2022-present)
- Sensors (2022-present)
- Computational Intelligence and Neuroscience (2022-present).
- PeerJ Computer Science (2022-present).
- Mathematics (2022-present).
- Computer Animation and Virtual Worlds (2023-present).
- Complex & Intelligent Systems (2023-present).
- PC member and reviewer for Conferences and Workshops:
  - IEEE COMPSAC (2020, 2021, 2022, 2023)
  - IEEE ICSC (2020, 2021, 2022, 2023)
  - IEEE Big Data (2019, 2020, 2021, 2022, 2023)
  - IEEE QRS (2020, 2021)
  - IEEE ETCEA (2022)
  - iConference (2021)
  - JCDL (2021, 2022)
  - ASIS&T (2021)
  - ALLDATA 2022, 2023
- Workshops and Round Table Chair of the 18th International ISKO Conference (ISKO International Conference 2024).
- Website Chair of the 5th IEEE International Conference on Artificial Intelligence Testing (IEEE AITest 2023).
- Publicity Chair of the 18th ACM/IEEE Joint Conference on Digital Libraries (JCDL 2018).

## Professional Associations

---

- 2017-Present. Association for Computing Machinery (ACM)
- 2020-Present. Institute of Electrical and Electronics Engineers (IEEE)
- 2020-Present. ACM Special Interest Group on Information Retrieval (SIGIR)
- 2020-Present. Texas Computer Education Association (TCEA)
- 2020-Present. Association For Information Science and Technology (ASIS&T)
- 2021-Present. Association for the Advancement of Artificial Intelligence (AAAI)