

Haihua Xie

Assistant Professor | Beijing Institute of Mathematical Sciences and Applications (BIMSA)
No. 544 Hefangkou Village, Huairou District, Beijing 101408 | haihua.xie@bimsa.cn | haihuaxie@foxmail.com | +86 18710156236
Curriculum Vitae | Last updated: April 2026

RESEARCH PROFILE

Dr. Haihua Xie works at the intersection of natural language processing, large language models, knowledge engineering, and domain text big-data analysis. His recent research spans intelligent agents and social simulation, knowledge graphs and knowledge services, public-opinion dynamics, and computational methods for scholarly, media, medical, and macroeconomic text data.

RESEARCH INTERESTS

Natural language processing, large language models, agents, and social simulation.
Knowledge engineering, knowledge mining and services, and knowledge graphs.
Domain text big-data analysis, including media big data and scholarly big data.

PROFESSIONAL APPOINTMENTS

2021.10-Present	Beijing Institute of Mathematical Sciences and Applications Assistant Professor
2017.06-2021.09	Peking University & Founder Tech Group Senior Algorithm Researcher
2015.07-2017.05	Peking University & Founder Tech Group Postdoctoral Fellow

EDUCATION

2015.07-2017.05	Peking University Postdoctoral Fellow, Artificial Intelligence
2009.08-2015.05	Iowa State University Ph.D., Computer Science
2006.09-2009.07	Tsinghua University M.S., Software Engineering
2002.09-2006.07	Nanjing University of Posts and Telecommunications B.S., Electronic Information Science and Technology

PUBLICATIONS AND WRITINGS

Journal Articles

Yinzhu Cheng, **Haihua Xie**, Yaqing Wang, Miao He, and Mingming Sun. Semantic Distance Measurement Based on Multi-Kernel Gaussian Processes. *Data Analytics and Topology* 1 (2026): 133-142. DOI: 10.4310/DAT.260127000721.

Xue, Huzhi, **Haihua Xie**, and Butian Zhao. A social trust network-based hybrid fuzzy multi-criteria group decision-making approach for decarbonized and sustainable supplier selection. *Complex & Intelligent Systems* 12, no. 1 (2026): 2. (SCI, JCR Q1).

Haihua Xie, Miao He, Xiaoming Zhang, Lin Liu, and Philip S. Yu. OpenNOD: Modeling Event-Driven Opinion Dynamics in Open Social Networks. *IEEE Transactions on Big Data* (2025). DOI: 10.1109/TBDATA.2025.3650059. (SCI, JCR Q1).

Haihua Xie and Miao He. Tracking Fine-Grained Public Opinions: Two Datasets from Online Discourse on Trending Topics. *Mathematics* 13, no. 21 (2025): 3433. <https://doi.org/10.3390/math13213433>. (SCI, JCR Q1).

Huzhi Xue, **Haihua Xie**, Butian Zhao, and Jun Wang. An integrated multi-criteria decision-making approach with unknown weight information under probabilistic dual-hesitant Pythagorean fuzzy environment and its application to supplier selection. *Engineering Applications of Artificial Intelligence* 161, Part C (2025): 112202. <https://doi.org/10.1016/j.engappai.2025.112202>. (SCI, JCR Q1).

Tao Zhang, Ke Tang, and **Haihua Xie**. Construction and Macroeconomic Forecasting of the Chief Economist Confidence Index Based on Text Analysis. *Journal of Management Science in China* 2025, no. 3: 1-18.

Haihua Xie, Zhiyou Chen, Jing Cheng, Xiaoqing Lyu, and Zhi Tang. Chinese Grammar Error Detection Based on Data Enhancement and Multi-task Feature Learning. *Journal of Chinese Information Processing* 36, no. 12 (2022): 16-23.

Haihua Xie, Xuefei Chen, Yimin Du, Xiaoqing Lyu, and Zhi Tang. A Semi-supervised Method for Chinese Key Phrase Extraction Based on Statistical Features and Graph Model. *Journal of Chinese Information Processing* 36, no. 4 (2022): 57-65.

Haihua Xie, Aolin Li, Yabo Li, Zhiyou Chen, Jing Cheng, Xiaoqing Lyu, and Zhi Tang. CPLM-CSC: A Character-based Pre-trained Language Model for Chinese Spelling Checking and Correction. *Journal of Chinese Information Processing* 35, no. 5 (2021): 38-45.

Jianwei Li, Yongtao Wang, **Haihua Xie**, and Kai-Kuang Ma. Learning a Single Model With a Wide Range of Quality Factors for JPEG Image Artifacts Removal. *IEEE Transactions on Image Processing* 29 (2020): 8842-8854. (SCI, JCR Q1).

Haihua Xie, Jingwei Yang, Carl K. Chang, and Lin Liu. A statistical analysis approach to predict user's changing requirements for software service evolution. *Journal of Systems and Software* 132 (2017): 147-164. (SCI, JCR Q1).

Conference Papers and Proceedings

- Huzhi Xue and **Haihua Xie**. From Configuration to Evolution: Corporate Social Responsibility Practices and Supply Chain Effects on Environmental Innovation. 2026 IEEE World Congress on Computational Intelligence (WCCI 2026). Accepted.
- Yinzhu Cheng, Yihan Meng, **Haihua Xie**, and Mingming Sun. TopoEval: A Comprehensive Benchmark for Topological Reasoning in Foundation Models. The 64th Annual Meeting of the Association for Computational Linguistics (ACL 2026).
- Haihua Xie** and Miao He. GC-DPG: Graph-Constrained Dual-Phase Generation for Safe and Verifiable Chinese Medical Question Answering. The ACM Web Conference 2026 (WWW '26). Accepted.
- Miao He and **Haihua Xie**. Reshaping the Narrative: Linguistic Evolution and Reception in Online Storytelling Communities. CHI EA '26: Proceedings of the Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems. Article No.: 562, Pages 1-6. <https://doi.org/10.1145/3772363.37984>.
- Haihua Xie**, Yinzhu Cheng, Yaqing Wang, Miao He, and Mingming Sun. RD-MCSA: A Multi-Class Sentiment Analysis Approach Integrating In-Context Classification Rationales and Demonstrations. The 2025 Conference on Empirical Methods in Natural Language Processing (EMNLP 2025).
- Huzhi Xue, Butian Zhao, **Haihua Xie**, and Zeyu Sun. RankLLM: A multi-criteria decision-making method for LLM performance evaluation in sentiment analysis. The Twenty-fourth China National Conference on Computational Linguistics (CCL 2025).
- Haihua Xie** and Miao He. Stance Detection on Macro Topics in Social Discussions Via Multi-Factor Aggregation Analysis. 2025 International Joint Conference on Neural Networks (IJCNN 2025).
- Huzhi Xue, **Haihua Xie**, and Carl K. Chang. Assessing AI adoption for enhancing healthcare supply chain resilience: A novel hybrid interval-valued q-rung orthopair fuzzy MCDM. IEEE International Conference on Digital Health (ICDH 2024), in IEEE World Congress on SERVICES, Shenzhen, China, 2024, pp. 50-57.
- Lin Liu, **Haihua Xie**, Jingwei Yang, and Yanni Li. Building Smart Systems to Satisfice User Intention in Situ: A Look Back to the Future. IEEE International Conference on Software Services Engineering (SSE 2023), pp. 262-271.
- Haihua Xie**, Xuefei Chen, Xiaoqing Lyu, and Lin Liu. A Generalized Strategy of Chinese Grammatical Error Diagnosis Based on Task Decomposition and Transformation. China Conference on Knowledge Graph and Semantic Computing (CCKS 2023), pp. 135-147.
- Haihua Xie**, Xiaoqing Lyu, and Xuefei Chen. String Editing based Chinese Grammatical Error Diagnosis. Proceedings of the 29th International Conference on Computational Linguistics (COLING 2022), pp. 5335-5344.
- Tao Zhang, Ke Tang, and **Haihua Xie**. Construction and Macroeconomic Forecasting of a High-frequency Economist Confidence Index Based on Natural Language Processing. The Third China Macroeconomists Forum, June 2022.
- Jianwei Li, Yongtao Wang, **Haihua Xie**, and Kai-Kuang Ma. Learning Local and Global Priors for JPEG Image Artifacts Removal. IEEE Signal Processing Letters 27 (2020): 2134-2138.
- Shaoling Jing, Shibo Hong, Dongyan Zhao, **Haihua Xie**, and Zhi Tang. Combining Impression Feature Representation for Multi-turn Conversational Question Answering. CNCL 2020, pp. 57-69.
- Nuo Xu, **Haihua Xie**, and Dongyan Zhao. A Novel Joint Framework for Multiple Chinese Events Extraction. CNCL 2020, pp. 174-183.
- Haihua Xie**, Xuefei Chen, Yimin Du, Xiaoqing Lyu, and Zhi Tang. A Semi-supervised Method for Chinese Key Phrase Extraction based on Statistical Features and Graph Model. Proceedings of the Fourteenth China Conference on Knowledge Graph and Semantic Computing (CCKS 2020).
- Haihua Xie**, Zhiyou Chen, Jing Cheng, Xiaoqing Lyu, and Zhi Tang. Chinese Grammar Error Detection based on Data Enhancement and Multi-task Feature Learning. Proceedings of the Nineteenth China National Conference on Computational Linguistics (CCL 2020).
- Haihua Xie**, Zhiyou Chen, Yabo Li, Shaoling Jing, Xiaoqing Lyu, and Zhi Tang. SemBRS: A Semantic Analysis based Book Retrieval Approach. IEEE 3rd International Conference on Multimedia Information Processing and Retrieval (MIPR 2020).
- Haihua Xie**, Aolin Li, Yabo Li, Jing Cheng, Zhiyou Chen, Xiaoqing Lyu, and Zhi Tang. Automatic Chinese Spelling Correction based on Character-based Pre-trained Contextual Representations. Proceedings of the 8th CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC 2019).
- Haihua Xie**, Aolin Li, Yabo Li, Zhiyou Chen, Jing Cheng, Xiaoqing Lyu, and Zhi Tang. CPLM-CSC: Character-based Pre-trained Language Model for Chinese Spelling Checking and Correction. Proceedings of the Eighteenth China National Conference on Computational Linguistics (CCL 2019).
- Haihua Xie**, Xiaoqing Lyu, Xuefei Chen, Jinle Tong, and Zhi Tang. ES-ESens: Detection of Event Sentences based on Evaluation of the Explicitness and Significance of Information. IEEE 2nd International Conference on Multimedia Information Processing and Retrieval (MIPR 2019).
- Yitao Liu, Xiaoqing Lyu, **Haihua Xie**, Xiaotong Yan, Bei Wang, and Zhi Tang. A Method for Improving the Reliability of Causal Inference from Large-scale Data in Biomedicine. 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2018).
- Haihua Xie**, Xiaoqing Lyu, and Zhi Tang. Clustering Analysis-Based Approach to Detecting Entity Mixture in Knowledge Bases. Proceedings of the 18th ACM/IEEE Joint Conference on Digital Libraries (JCDL 2018), pp. 395-396.
- Haihua Xie**, Xiaoqing Lyu, Zhi Tang, and Xiaojun Huang. Detection of Entity Mixture in Knowledge Bases Using Hierarchical Clustering. Proceedings of the 5th CCF Conference on Natural Language Processing and Chinese Computing and 24th International Conference on Computer Processing of Oriental Languages (NLPCC-ICCPOL 2016), pp. 288-299.
- Haihua Xie**, Xiaoqing Lyu, Zhi Tang, and Mao Ye. A Methodology to Evaluate Triple Confidence and Detect Incorrect Triples in Knowledge Bases. Proceedings of the 16th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2016), pp. 251-252.
- Haihua Xie** and Carl K. Chang. Detection of New Intentions from Users Using the CRF Method for Software Service Evolution in Context-Aware Environments. Proceedings of the 39th Annual IEEE Computer Software and Applications Conference (COMPSAC 2015), pp. 385-391.
- Haihua Xie**, Carl K. Chang, Hua Ming, and Kai-Shin Lu. The Concepts and Ontology of SiSL: A Situation-Centric Specification Language. Proceedings of the 36th Annual IEEE Computer Software and Applications Conference (COMPSAC 2012), pp. 385-391.

Xiaoying Bai, Fevzi Belli, Elisa Bertino, Carl K. Chang, Atilla Elci, Cristina Cerschi Seceleanu, **Haihua Xie**, and Mohammad Zulkernine. 36th Annual IEEE Computer Software and Applications Conference. COMPSAC 2012, Izmir, Turkey, July 16-20, 2012. IEEE Computer Society, ISBN 978-1-4673-1990-4.

Xiaoying Bai, Fevzi Belli, Elisa Bertino, Carl K. Chang, Atilla Elci, Cristina Cerschi Seceleanu, **Haihua Xie**, and Mohammad Zulkernine. 36th Annual IEEE Computer Software and Applications Conference Workshops. COMPSAC 2012, Izmir, Turkey, July 16-20, 2012. IEEE Computer Society, ISBN 978-1-4673-2714-5.

Haihua Xie. Observable Context-Based User Intention Specification in Context-Aware Environments. Proceedings of the 35th Annual IEEE Computer Software and Applications Conference (COMPSAC 2011), pp. 712-715.

Tianying Li, Tong Li, **Haihua Xie**, and Lin Liu. i*-Prefer: A Tool for Preference Model-Driven Decision Making. Proceedings of the 5th International i* Workshop (iStar 2011), pp. 151-153.

Wenting Ma, Lin Liu, **Haihua Xie**, Hongyu Zhang, and Jinglei Yin. Preference Model Driven Services Selection. Proceedings of the 21st International Conference on Advanced Information Systems Engineering (CAiSE 2009), pp. 216-230.

Haihua Xie, Lin Liu, and Jingwei Yang. i*-Prefer: Optimizing Requirements Elicitation Process Based on Actor Preferences. Proceedings of the 24th ACM Symposium on Applied Computing (ACM SAC 2009), pp. 347-354.

PATENTS AND SOFTWARE COPYRIGHTS

Granted Patents

Information Value Evaluation Method. Inventors: Haihua Xie; Xuefei Chen; Jinle Tong. Publication No.: CN110866389B; granted 2021.12.17. IPC: G06F40/289; G06F40/295; G06F40/253. CPC: G06F17/278; G06F17/2775; G06F17/274.

Information Quality Evaluation Method. Inventors: Haihua Xie; Xuefei Chen; Jinle Tong. Publication No.: CN110889274B; granted 2022.02.08. IPC: G06F40/211; G06F40/295; G06F40/30. CPC: G06F17/271; G06F17/2785; G06F17/278.

Information Explicitness Evaluation Method. Inventors: Haihua Xie; Xuefei Chen; Jinle Tong. Publication No.: CN110889289B; granted 2022.05.06. IPC: G06F40/289; G06N3/04; G06F40/30. CPC: G06F17/2785; G06F17/2775; G06N3/04.

Email Information Extraction Method and Apparatus. Inventors: Haihua Xie; Xuewen Luo; Xuefei Chen; Jinle Tong. Publication No.: CN110020366B; granted 2021.06.15. IPC: G06F16/958; G06F16/903; G06F16/84. CPC: G06F17/30896; G06F17/2264.

Method for Extracting Basic Information from Scholars' Academic Profiles. Inventors: Xuefei Chen; Xuewen Luo; Haihua Xie; Jinle Tong. Publication No.: CN110019761B; granted 2021.04.30. IPC: G06F16/9535; G06F16/953; G06F16/335; G06F16/951. CPC: G06F17/30699; G06F17/30864.

Method and Apparatus for Validating Knowledge Base Triples. Inventors: Haihua Xie; Xiaojun Huang; Xiaoqing Lyu; Zhi Tang. Publication No.: CN108021595B; granted 2020.07.14. IPC: G06F16/30. CPC: G06F17/30731; G06F17/30705.

Method for Detecting Confusion Among Homonymous Instances in Knowledge Bases. Inventors: Haihua Xie; Xiaojun Huang; Xiaoqing Lyu; Zhi Tang. Publication No.: CN108009184B; granted 2021.08.27. IPC: G06F16/35; G06F16/36. CPC: G06F17/30734.

Software Copyrights

The Cooperative Research Environment (CoRE) System. Copyright owner: Iowa State University; developer: **Haihua Xie**; copyright type: GNU General Public License.

Subject Preference Requirement Modeling Software (i*-Prefer Modeller). Copyright owner: Tsinghua University; developers: Lin Liu and **Haihua Xie**; registration no.: 2009SRBJ6677.

ACADEMIC ACTIVITIES AND SERVICE

ICBS (International Congress of Basic Sciences) 2025: Session Chair; FSA Manager.

ICBS (International Congress of Basic Sciences) 2023 and 2024: Session Chair.

KDD (ACM Conference on Knowledge Discovery and Data Mining) 2026: Reviewer.

ICDH (IEEE International Conference on Digital Health) 2024: Program Chair.

ICDH (IEEE International Conference on Digital Health) 2023: Program Committee Member.

INDERSCIENCE Publishers: Program Committee Member.

IEEE SERVICES 2018 Congress: Program Committee Member.

COMPSAC 2015: MyReview Chair.

COMPSAC 2012: Conference Proceedings Chair.

COMPSAC 2011: Conference Proceedings Chair.

China Computer Federation (CCF): Professional Member.

TEACHING

2026SP	Data Mining BIMSA - Beijing University of Civil Engineering and Architecture - Beijing Information Science and Technology University, School of Digital Intelligence
2026SP	Machine Learning Practice with Python Qiuzhen College, Tsinghua University
2025AU	Python Programming for Data and Computation Qiuzhen College, Tsinghua University
2025SP	Natural Language Processing and Computational Linguistics BIMSA
2024AU	Frontiers in Large Language Models (LLMs) BIMSA
2024SP	Foundations of Large Language Models (LLMs) BIMSA

2023AU	Frontiers in Natural Language Processing BIMSA
2023SP	Fundamentals of Natural Language Processing BIMSA
2022AU	Algorithms in Natural Language Processing BIMSA
2022SP	Algorithms and Models in Natural Language Processing II BIMSA
2021AU	Algorithms and Models in Natural Language Processing I BIMSA

SUPERVISION

Ph.D. Students Supervised

Huzhi Xue - jointly supervised by BIMSA and Beihang University.
Yinzhu Cheng - jointly supervised by BIMSA and Renmin University of China.

Postdoctoral Projects Supervised

Shaoling Jing: Research on Multi-turn Conversational Question Answering Based on Reading Comprehension (2019.04-2021.03).
Nuo Xu: Research on Event Extraction and Evolution Analysis Based on Media Data (2019.04-2021.03).
Jianwei Li: Research on Image Enhancement Technology for Printed Materials Based on Deep Learning (2018.10-2020.09).

AWARDS AND HONORS

2025 Ruolin Paper Award (BIMSA, YMSC-Tsinghua, Qiuzhen-Tsinghua, Morningside Center of Mathematics-CAS).
2019 Zhongguancun High-End Leading Talent (Zhongguancun High-Level Talent Project).
2019 Outstanding Postdoctoral Fellow, 20th Anniversary of the Zhongguancun Science Park Haidian Park Enterprise Postdoctoral Research Workstation (2 recipients total).
2018 Beijing 13th Batch of Overseas High-Level Talents; Beijing Distinguished Expert (Beijing Thousand Talents Program - Young Talent Project).

RESEARCH PROJECTS

2025.01-Present	NSFC General Program: Research on Data Asset Pricing Based on Growth Options RMB 500,000; Grant No. 72473013. Role: Participated in research on the knowledge graph component.
2019.12-2023.07	National Key R&D Program of China: R&D of Knowledge Aggregation Service Technologies for Professional Content and Demonstration of Innovative Services RMB 12 million; Grant No. 2019YFB1406300. Role: Responsible for the subproject 'Automatic Resource Tag Generation and Semantic Annotation Technologies across Media and Domains'.
2018.07-2021.07	Research on Data Mining and Knowledge Extraction Technologies Based on Publishing Resource Content RMB 3 million. Role: Principal Investigator.
2015.07-2017.07	Research and Implementation of a Confidence Calculation Model for Domain Ontology Libraries RMB 200,000. Role: Principal Investigator.
2015.07-2017.07	Chinese Digital Library Xuezhi Search System RMB 5 million. Role: Principal Investigator for the Xuezhi Search System.
2016.01-2017.05	Huafu Zhi - Cloud Platform for Chinese Costume Culture RMB 5 million. Role: Responsible for resource tag generation and resource relevance calculation.
2009.01-2010.12	NSFC Project: Research on Security Requirements Engineering Technology Based on Social Modeling Role: Participant.
2008.09-2010.07	National 973 Program: Typical Business Simulation Technology and Networked Testbed for Specific Applications Role: Participant.
2007.09-2009.07	National 863 Program: Requirements Engineering Technology Based on Dynamic Optimal Decision-Making Role: Participant.