WENJIA LI

521 W Sycamore, Denton, TX, 76201 | +1 208 874 9195 | wli@uidaho.edu

DEGREES

BACHELOR DEGREE OF SCIENCE | 2012-2016 | CHANG'AN UNIVERSITY

- Major: Geographic Information System
- Related coursework: Geographic Information System; Database; Data Structure; Algorithms; Introduction of Software Engineering; Introduction to Geographic Information Systems; Cartography; etc.

DOCTOR DEGREE (MASTER-DOCTOR COMBIND PROGRAM) | 2016-2022 | CHINA UNIVERSITY OF GEOSCIENCES

- Major: Surveying and Mapping (DOCTOR; 2018-2022)/ Software Engineering (MASTER: 2016-2018)
- Related coursework: Software Engineering; Systems design; Software Testing; Scientific and technical writing; Frontiers in Mapping Science and Technology; etc.

RESEARCH INTERESTS

Geosciences, Text Mining, Natural Language Processing, Semantic Web, Ontology, Spatiotemporal Analysis and Deep Learning

PROFESSIONAL SKILLS

PROGRAMMING LANGUAGE

- Python; R; Machine Learning & AI (Experienced in developing and deploying machine learning models using frameworks like TensorFlow or PyTorch)

SOFTWARE DEVELOPMENT

- ArcGIS; protégé; neo4j; Microsoft Office; VOSviewer; Draw; GitHub.

EXPERIENCE

POSTDOCTORAL FELLOW | UNIVERSITY OF IDAHO | 2022.12 - PRESENT

- Utilized Python to conduct a detailed comparison between the Mindat and IMA Master mineral lists, achieving a comprehensive alignment and discrepancy analysis to enhance the accuracy and completeness of mineralogical databases.
- Developed the Stratigraphic Knowledge Graph (StraKG), a comprehensive semantic framework integrating geoscientific data from diverse sources using advanced NLP techniques (e.g., Transformer Models) and graph database (Neo4j), facilitating enhanced accessibility and queryability of stratigraphic information.
- Conducted a bibliometric analysis in Mathematical Geosciences, utilizing data mining and visualization tools to identify and map key research trends, collaborations, and impact within the field, providing valuable insights into its evolutionary trajectory and scholarly network.

- Led the GeoSymbolNet project, innovatively applying data augmentation techniques and deep learning models (e.g., CNNs) to accurately decipher and classify complex geological map symbols, significantly improving the efficiency and reliability of geological data interpretation.

TEACHING ASSISTANT | UNIVERSITY OF IDAHO | SPRING 2023

- CS 479/579: Data science

TEACHING ASSISTANT | UNIVERSITY OF IDAHO | FALL 2023

- CS 489: Semantic Web and Open Data

TECHNICAL STAFF | CHINA UNIVERSITY OF GEOSCIENCES | 2018-2022

- Implemented advanced NLP techniques (e.g., BERT, hypernetwork, Few-shot learning) for entity and relationship extraction from unstructured geoscientific data, enhancing data analysis workflows by accurately identifying geological entities and attributes, thus facilitating richer data insights and knowledge discovery.
- Developed an innovative methodology using BERT and SoftMax for the nuanced modeling of lengthy, unstructured geological hazard reports, coupled with LDA for visualizing geohazard interdependencies, enhancing the accuracy of geohazard classification and supporting emergency decision-making processes.

GRANTS

- **PI**, GeoSymbolNet: Leveraging Data Augmentation to Decipher Geological Map Symbols. FARR mini research. $2,500.
- **Travel grant**: $2,000 travel grant from NSF-funded PyRATES workshop (linked.earth/FROGS) at University of Southen California.
- **Travel grant**: $1500 travel grant from the FARR (FAIR in ML, AI-readiness and Reproducibility) Research Coordination Network (farr-rcn.org).
- **Technical staff**, EarthCube Capabilities: OpenMindat - Open Access and Interoperable Mineralogy Data to Broaden Community Access and Advance Geoscience Research, funded by National Science Foundation (NSF), 2021-2024, PI: Xiaogang MA, $792,475.00.
- **Student Co-I**, Research on the construction method of 3D geological model based on multimodal data understanding and fusion, funded by National Natural Science Foundation of China (NSFC), 2019-2022, PI: Liang Wu.
- **Technical staff**, Geoscience Knowledge Map Representation Model and Collaborative Construction of Group Wisdom, funded by National Natural Science Foundation of China (NSFC), Technical staff, 2021-2023, PI: Zhong Xie.
- **Technical staff**, Research on spatio-temporal topic association retrieval method for geological big data, funded by National Natural Science Foundation of China (NSFC), 2017-2020, PI: Zhong Xie.

PUBLICATIONS

JOURNAL

- **Wenjia LI**, et al. Mapping the Evolution of Mathematical Geoscience Research with Big Literature Data and Context-Aware Text Mining. Computers & Geosciences. [*Under Review*].
- **Wenjia LI**, et al. Geological Object Recognition in Geological Maps Through Data Augmentation and Transfer Learning Techniques. [*Editing*].

- Jolyon Ralph, Pavel Martynov, Xiaogang Ma*, David Von Bargen, **Wenjia Li**, et al. Identifier service in the Mindat database: Persistent and meaningful access to massive records of minerals and other natural materials[J]. Information. [*Under Review*].
- Jiyin Zhang, C Clairmont, Xiang Que, **Wenjia Li**, Weilin Chen, Chenhao Li, X Ma. Streamlining Geoscience Data Analysis with an Llm-Driven Workflow. [*Under Review*].
- **Wenjia LI**, et al. A hybrid knowledge graph for efficient exploration of lithostratigraphic information in open text data. Applied Computing and Geosciences. 2024.
- Weilin Chen, Xiaogang Ma, Zhe Wang, **Wenjia Li**, et al. Exploring neuro-symbolic AI applications in geoscience: implications and future directions for mineral prediction. Earth Science Informatics. 2024.
- **Wenjia Li**, Liang Wu, Qinjun Qiu*, et al. Deep learning and network analysis: Classifying and visualizing geologic hazard reports[J]. Journal of Earth Science. 2022.
- **Wenjia Li**, Kai Ma, Liang Wu, et al. Chinese Word Segmentation Based on Self-Learning Model and Geological Knowledge for the Geoscience Domain. Earth and Space Science. 2021.
- Hao Liu, Qinjun Qiu, Liang Wu, **Wenjia Li**, et al. Few-shot learning for name entity recognition in geological text based on GeoBERT. Earth Science Informatics. 2022.
- Can Zhang, **Wenjia Li**, Liang Wu, et al. A multi-granularity knowledge association model of geological text based on hypernetwork. Earth Science Informatics. 2021.
- Bin Wang, Liang Wu, **Wenjia Li**, et al. A semi-automatic approach for generating geological profiles by integrating multi-source data. Ore Geology Reviews. 2021.
- **Wenjia Li**, Liang Wu, Zhong Xie, et al. Ontology-based question understanding with the constraint of Spatio-temporal geological knowledge. Earth Science Informatics. 2019.
- Qinjun Qiu, Zhong Xie, Liang Wu, **Wenjia Li**. Geoscience Keyphrase Extraction Algorithm Using Enhanced Word Embedding. Expert Systems with Applications. 2019.
- Qinjun Qiu, Xie Zhong, Liang Wu, Liufeng Tao, **Wenjia Li**. BiLSTM-CRF for geological named entity recognition from the geoscience literature. Earth Science Informatics. 2019.
- Qinjun Qiu, Zhong Xie, Liang Wu, **Wenjia Li**. DGeoSegmenter: A dictionary-based Chinese word segmenter for the geoscience domain. Computers & Geosciences. 2018.
- Kai Ma, Liang Wu, Liufeng Tao, **Wenjia Li**. Matching Descriptions to Spatial Entities Using a Siamese Hierarchical Attention Network. IEEE Access. 2018.

CONFERENCE

- **Wenjia Li**, et al. A Multi-granularity Geological Objects (Multi-GeoO) Knowledge Representation Framework. 21st Annual Conference of the International Association for Mathematical Geosciences, 2023.
- **Wenjia Li**, et al. Visualizing the Dynamics of Mathematical Geosciences: An Integrated Study of Domain-Specific Dictionaries and Journal Publications. American Geophysical Union's Fall Meeting, 2023.
- Anirudh Prabhu, Michael L Wong, Shaunna M M Morrison, Alexandra Ostroverkhova, Maureen Clark, Hao Zhong, Trygve J Prestgard, **Wenjia Li**, Jason R Williams, Sarah Shi, Jennifer Mays, Robert Hazen. From detecting agnostic biosignatures to characterizing chondrites: How network science is perfect for making scientific discoveries with geochemical data [**Invited**]. American Geophysical Union's Fall Meeting, 2023.
- Chenhao Li, Xiaogang Ma, Weilin Chen, Jiyin Zhang, **Wenjia Li**. Building a New Recommender System for Open Data by Using Trust ranking and Social Trust Network. American Geophysical Union's Fall Meeting, 2023.
- Weilin Chen, Xiaogang Ma, Jiyin Zhang, Xiang Que, **Wenjia Li**, Chenhao Li, Zhe Wang. Neuro-symbolic AI in Mineral Prediction: A Literature Review. American Geophysical Union's Fall Meeting, 2023.

- Jiyin Zhang, Jolyon Ralph, David Von Bargen Von Bargen, Xiaogang Ma, Xiang Que, **Wenjia Li**. Advanced Machine Interface to Open Mineral Data: The OpenMindat Data API and Python Package. American Geophysical Union's Fall Meeting, 2023.
- Alexandra Ostroverkhova, Maureen Clark, Anirudh Prabhu, Trygve J Prestgard, **Wenjia Li**, Hao Zhong, Jennifer Mays, Robert Hazen, Kerstin Lehnert, Shaunna M M Morrison. Unraveling the Enigma of Ungrouped Chondrites: A Data Science Approach for Exploring their Origins. American Geophysical Union's Fall Meeting, 2023.
- Alexandra Ostroverkhova, Maureen Clark, Anirudh Prabhu, Shaunna M M Morrison, **Wenjia Li**, Jennifer Mays, Kerstin Lehnert. Using Data Science Methods to Explore the Origins of Ungrouped Carbonaceous Chondrites. GSA (Geological Society of America) Connects Meeting. 2023.
- Jiyin Zhang, Xiang Que, Xiaogang Ma, **Wenjia Li**, Jingyi Huang, Weilin Chen. Data Exploration Supported by Open and Fair Data from MINDAT. GSA (Geological Society of America) Connects Meeting. 2023.
- **Wenjia Li**, et al. Semantic Retrieval Facing on Spatio-temporal Geological Big-data. American Geophysical Union's Fall Meeting, 2018.

POSTER

- **Wenjia Li**, Xiaogang Ma. Trends of research 'hotspots' reflected in the IAMG journals' author keywords. 22st Annual Conference of the International Association for Mathematical Geosciences, 2023.
- **Wenjia Li**, Xiaogang Ma. Constructing a Stratigraphic Knowledge Graph (StraKG) with Multi-source Data to Better Understand the Earth's Rock Layers. Earth Science Information Partners (ESIP).

PROFESSIONAL SERVICES

Present	Review Editor for journals: Expert Systems with Applications; Geoscience Frontiers; Ore Geology Reviews; Earth and Space Science; Applied Computing and Geosciences
2023.10 – 2024.02	OSPA COORDINATOR AGU 2023 MEETING