

# Dazhou Yu

+1-770-312-2439 | [yuqi0071@gmail.com](mailto:yuqi0071@gmail.com) | [Google Scholar](#)

 [LinkedIn](#) |  [GitHub](#)

Atlanta, GA, U.S.

## RESEARCH INTERESTS

LLM Agent, Retrieval Augmented Generation, Spatial Representation Learning and Question Answering, Data Mining, Machine Learning

## EDUCATION

- **Emory University** Aug 2024 - Aug 2026 (expected)  
*Ph.D. in Computer Science and Informatics (Advisor: Prof. Liang Zhao)* Atlanta, GA
- **Emory University** Aug 2021 - Dec 2024  
*Master in Computer Science and Informatics (Advisor: Prof. Liang Zhao)* Atlanta, GA
- **The Ohio State University** Aug 2019 - Dec 2020  
*Master in Electrical and Computer Engineering (program not completed)* Columbus, OH  
◦ GPA: 4.00/4.00
- **Tongji University** Sept 2015 - Jul 2019  
*Bachelor of Engineering in Traffic Engineering* Shanghai  
◦ GPA: 90.75/100
- **The University of Auckland** Jan 2019 - Jun 2019  
*China Scholarship Council (CSC) Outstanding Undergraduate International Exchange Program* Auckland, New Zealand  
◦ Final Year Project on Vibration Energy Harvesting (Advisor: Prof. Lihua Tang)
- **National Chiao Tung University** Jan 2017 - Jul 2017  
*Exchange student, Mechanical Engineering Department* Hsinchu  
◦ GPA: 94.3/100

## PROFESSIONAL EXPERIENCE

- **Machine Learning Engineer Intern, TikTok** Sep 2025 - Dec 2025  
San Jose, CA  
◦ Extended the team's POI embedding foundation model to multi-modal inputs, integrating POI descriptions, user reviews, popularity metrics, and images. Achieved an average 5 pp improvement in downstream classification performance in weekly online A/B tests.  
◦ Developed a RAG-based LLM agent to enable private, in-team question answering over internal POI data.
- **Applied Scientist Intern, Amazon** May 2025 - Aug 2025  
Seattle, WA  
◦ Designed and implemented M3-Agent (Multi-item, Multi-constraint, Multi-objective), an LLM agent that translates natural language shopping queries into precise, catalog-verified recommendations under multiple constraints. The system connects LLM capabilities with recommendation engines, delivering results that balance immediate needs with long-term user utility.

## SELECTED PUBLICATIONS

- [ARR (Under review)] Spatial-Agent: Agentic Geo-spatial Reasoning with Scientific Core Concepts  
Riyang Bao, Cheng Yang, **Dazhou Yu**, Zhexiang Tang, Gengchen Mai, Liang Zhao.
- [ARR (Under review)] Spatial-RAG: Spatial Retrieval Augmented Generation for Real-World Spatial Reasoning Questions  
**Dazhou Yu**, Riyang Bao, Ruiyu Ning, Jinghong Peng, Gengchen Mai, Liang Zhao.
- [Digital Discovery] ACES-GNN: Can Graph Neural Network Learn to Explain Activity Cliffs?  
Xu Chen, **Dazhou Yu**, Liang Zhao, Fang Liu. Digital Discovery, 2025, 4, 2062 - 2074, DOI: [10.1039/D5DD00012B](https://doi.org/10.1039/D5DD00012B), Impact factor: 5.6
- [ICLR] PolyhedronNet: Representation Learning for Polyhedra with Surface-attributed Graph  
**Dazhou Yu**, Genpei Zhang, Liang Zhao. The Thirteenth International Conference on Learning Representations (2025)
- [KDD] PolygonGNN: Representation Learning for Polygonal Geometries with Heterogeneous Visibility Graph  
**Dazhou Yu**, Yuntong Hu, Yun Li, Liang Zhao. KDD '24: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Pages 4012 - 4022, DOI: [10.1145/3637528.3671738](https://doi.org/10.1145/3637528.3671738)
- [KDD] Self-consistent Deep Geometric Learning for Heterogeneous Multi-source Spatial Point Data Prediction  
**Dazhou Yu**, Xiaoyun Gong, Yun Li, Meikang Qiu, Liang Zhao. KDD '24: Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Pages 4001-4011, DOI: [10.1145/3637528.3671737](https://doi.org/10.1145/3637528.3671737)
- [SIGSPATIAL] Transferable Unsupervised Outlier Detection Framework for Human Semantic Trajectories  
Zheng Zhang, Hossein Amiri, **Dazhou Yu**, Yuntong Hu, Liang Zhao, Andreas Züfle. SIGSPATIAL '24: Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems, Pages 350 - 360, DOI: [10.1145/3678717.3691324](https://doi.org/10.1145/3678717.3691324), **Best Paper Candidate**

|              |                                                                                                                                                                                                                                                                                                                                                                                        |
|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| [SDM]        | STES: A Spatiotemporal Explanation Supervision Framework<br><b>Dazhou Yu</b> , Binbin Chen, Yun Li, Suman Dhakal, Yifei Zhang, Zhenke Liu, Minxing Zhang, Jie Zhang, Liang Zhao. Proceedings of the 2024 SIAM International Conference on Data Mining (SDM), Pages 73 - 81, DOI: <a href="https://doi.org/10.1137/1.9781611978032.9">10.1137/1.9781611978032.9</a>                     |
| [CSUR]       | Going Beyond XAI: A Systematic Survey for Explanation-Guided Learning<br>Yuyang Gao, Siyi Gu, Junji Jiang, Sungsoo Ray Hong, <b>Dazhou Yu</b> , Liang Zhao. ACM Computing Surveys, Volume 56, Issue 7, Article No.: 188, Pages 1 - 39, DOI: <a href="https://doi.org/10.1145/3644073">10.1145/3644073</a> , Impact factor: 28                                                          |
| [TSAS]       | Deep Spatial Prediction via Heterogeneous Multi-source Self-supervision<br>Minxing Zhang, <b>Dazhou Yu</b> , Yun Li, Liang Zhao. ACM Transactions on Spatial Algorithms and Systems, Volume 9, Issue 3, Article No.: 20, Pages 1 - 26, DOI: <a href="https://doi.org/10.1145/3605358">10.1145/3605358</a>                                                                              |
| [ICDM]       | Deep Spatial Domain Generalization<br><b>Dazhou Yu</b> , Guangji Bai, Yun Li, Liang Zhao. 2022 IEEE International Conference on Data Mining (ICDM), DOI: <a href="https://doi.org/10.1109/ICDM54844.2022.00170">10.1109/ICDM54844.2022.00170</a>                                                                                                                                       |
| [SIGSPATIAL] | Deep Geometric Neural Network for Spatial Interpolation<br>Minxing Zhang, <b>Dazhou Yu</b> , Yun Li, Liang Zhao. SIGSPATIAL '22: Proceedings of the 30th International Conference on Advances in Geographic Information Systems, Article No.: 72, Pages 1 - 4, DOI: <a href="https://doi.org/10.1145/3557915.3561008">10.1145/3557915.3561008</a> , <b>Best Poster Runner-Up Award</b> |
| [ICDM]       | GNES: Learning to Explain Graph Neural Networks<br>Yuyang Gao, Tong Sun, Rishab Bhatt, <b>Dazhou Yu</b> , Sungsoo Hong, Liang Zhao. 2021 IEEE International Conference on Data Mining (ICDM), DOI: <a href="https://doi.org/10.1109/ICDM51629.2021.00023">10.1109/ICDM51629.2021.00023</a>                                                                                             |

## PROJECTS

- **Spatial Retrieval-Augmented Generation:** Nov 2024 - Jan 2025  

  - Designed and implemented a spatial question-answering framework integrating spatial databases (PostgreSQL, PostGIS) with retrieval mechanisms and large language models (LLMs).
  - Proposed a hybrid retrieval mechanism that combines sparse retrieval (SQL-based structured queries) with dense retrieval (LLM-powered semantic matching).
  - Proposed a multi-objective optimization framework that dynamically balances trade-offs between spatial and semantic relevance.
- **Polygon and Polyhedron Representation Learning:** Sep 2023 - Oct 2024  

  - Developed automatic script tools for converting polygons, multipolygons, and polyhedrons into comprehensive graph structures, ensuring no loss of information from the original geometric forms.
  - Collected real-world geometric shapes and transformed them into graphs for benchmarking the effectiveness of polygon and polyhedron representation learning models.
  - Designed pipelines that decompose graphs into local rigid structures and reassemble them using specialized graph neural networks. This approach achieves best performance in classification and retrieval tasks compared to contemporary shape encoding techniques.
- **Spatial (PM 2.5) Interpolation and Prediction** Sep 2021-Sep 2023  

  - Designed various fusion frameworks to align multiple data sources autonomously in a self-supervised manner, achieving the best performance in regression tasks compared to contemporary spatial prediction techniques.
  - Designed a geo-location-aware, multi-source graph neural network tailored to manage the intricacies of spatial relationships and feature heterogeneity among varied data sources, optimizing performance in complex environmental analyses.
- **Explanation-guided Deep Learning** Sep 2021-Sep 2023  

  - Designed explanation supervision frameworks that align model explanations with human-labeled counterparts, significantly enhancing model interpretability and performance.
  - Developed versatile loss functions to adapt the core technique across various domains, including solar flare forecasting using image-time series data, and molecular property prediction with graph data.

## HONORS AND AWARDS

- **Best Paper Candidate:** SIGSPATIAL, 2024
- **Student Travel Award:** KDD, 2024
- **Student Travel Award:** SDM, 2024
- **Best Poster Runner-Up Award:** SIGSPATIAL, 2022
- **Student Travel Award:** ICDM, 2022
- **First Class Scholarship:** Tongji University, 2017
- **National Scholarship:** Ministry of Education of the People's Republic of China, 2016

## SKILLS

- **Programming Languages:** Python, SQL, Java, C/C++, Matlab
- **Languages:** Chinese, English

## SERVICES

---

- **Independent reviewer:** ICML 25, KDD 24-25, ICLR 25, AISTATS 25, NeurIPS 24, Pattern Recognition, Information Fusion, IEEE Transactions on Big Data, International Journal of Applied Earth Observation and Geoinformation

## TEACHING EXPERIENCE

---

- **CS 170 Introduction to Computer Science:** TA, 2023 Spring
- **CS 526 Algorithm:** TA, 2022 Fall
- **CS 470 Data Mining:** TA, 2022 Spring