Bin Liu, Ph.D.

binliu14qd@gmail.com — 310-497-5766 — GitHub — Website

Career Focus

Generative AI · Large Language Models (LLM) · Vision-Language Models (VLM) · Multimodal AI · Embodied Intelligence · AI Agents · Physics AI · AI for Science

Education

North Carolina State University, Raleigh, NC
Ph.D. in Physics, Concentration: *AI in Physics, Generative Modeling* *2022*
M.S. in Phsysics *2018*
M.S. in Electrical Engineering, Concentration: *Computational Intelligence* *2018*
B.S. in Mechanical Engineering *2015*
B.S. in Physics *2015*

Technical Skills

- **Languages & Web**: Python, C++, C#, Java, TypeScript, JavaScript, SQL, HTML/CSS, MATLAB.
- **Frameworks & Libraries**: PyTorch, TensorFlow, JAX, HuggingFace, Diffusers, Transformers, TRL, FastAPI, React, Vite, CUDA, Scikit-Learn.
- **Agentic AI & Graph**: LangChain, LangGraph, LlamaIndex, AgentScope, Neo4j, GraphRAG.
- **Cloud & MLOps**: AWS (SageMaker, EC2), Azure (AI Studio, DevOps), Docker, Kubernetes, CI/CD, Git.
- **Generative AI Tools**: ComfyUI, Gradio, GPT-SoVITS, Ollama, n8n, Dify.

Professional Experience

BKO AI, Houston, TX
Strategic Leadership Role:
Senior AI Manager *Jan 2026 – Present*
AI Technology Manager *May 2025 – Jan 2026*

- **Executive Strategy:** Define and execute the strategic AI roadmap, aligning GenAI, multi-modal, and Knowledge Graph initiatives with business goals to drive innovation and operational efficiency.
- **Product Leadership:** Direct the lifecycle of all Agentic AI products on the enterprise platform, including Deep-search, Wide-search, Unified Vision, GUI automation, and Self-executing task agents.
- **Team Management:** Manage a cross-functional team of AI researchers, software engineers, and graph specialists; improved project delivery timelines by **30%** through Agile implementation and roadmap alignment.
- **MLOps Governance:** Established standardized MLOps practices including CI/CD pipelines, model monitoring, and secure Azure-based deployment, ensuring reproducibility and governance.

Technical Core Role:
Principal AI Research Engineer *Oct 2024 – Present*

- **Agentic Swarm Architecture:** Architected an AI agentic swarm system integrated within the Knowledge Graph Explorer, enabling autonomous cross-collaboration between agents for complex query resolution.
- **GraphRAG Implementation:** Spearheaded the design of fine-tuned **multi-LLM-agents** using **GraphRAG** and **Neo4j** on Azure. Streamlined CAD blueprint retrieval and budget estimation, improving planning accuracy by **55%** and reducing retrieval latency by **40%**.
- **Strategic Knowledge Graphs:** Designed Context Graphs for Tag-mapping and Failure Mode and Effects Analysis (FMEA), significantly enhancing agent decision-making capabilities and trading signal tracking.
- **Multi-Modal Pipeline:** Led the development of a pipeline integrating thermal sensors, IoT devices, and text reports for predictive maintenance. Achieved **92% accuracy**, reducing equipment downtime by **40%** and preventing **$15M** in annual losses.

Shell International Exploration and Production, Houston, TX
AI Research Scientist *Oct 2022 – Oct 2024*

- **Legal AI Automation:** Revolutionized contract generation by fine-tuning **LLaMA-3.2** with **LoRA**, **RAG**, and **LlamaParse**. Reduced processing time from 2 weeks to **10 minutes**, saving **$8M annually**. Implemented RLHF and an internal reflection mechanism to ensure legal accuracy.
- **3D Vision at Scale:** Developed state-of-the-art multiclass segmentation algorithms using transformer-based LVMs for 10TB of 4K-resolution CT-scan data. Achieved **0.841 mIoU** on severe class imbalances, generating **$20M** in annual revenue growth.
- **Generative Super-Resolution:** Applied diffusion models for rock image super-resolution (up to 32K), accelerating processing from 3 days to **15 minutes** (99% efficiency boost).

North Carolina State University, Raleigh, NC
Ph.D. Research Assistant *August 2019 - October 2022*

- **JWST Data Analysis:** Built deep learning pipelines to analyze multidimensional spectral data from the James Webb Space Telescope (EIGER Collaboration).
- **Bayesian Inference:** Reduced measurement errors by **80%** in intergalactic medium studies using generative deep learning, achieving reconstruction error rates as low as **0.025%**.
- **GUI Development:** Created deep learning pipelines and GUI software to analyze multidimensional spectral data from James Webb Space Telescope, a $10B project exploring the deep Universe using PyTorch.

Selected Technical Projects

- **E-Commerce Generative AI Suite** (2022 – Present): Integrated **StableDiffusionXL**, **Flux.1**, **ControlNet**, and **SegmentAnything2** to generate controllable 8K product images for 50+ clients.
- **Virtual Influencer Pipeline** (2022 – Present): Engineered a system using **GPT-SoVITS** and **Kling AI** to generate multilingual virtual influencers, reducing video customization time to 6 hours.
- **Digital Twin with Gaussian Splatting** (2023 – 2024): Reconstructed sneaker digital twins using **NeRF** and **3D Gaussian Splatting** with 97% geometric accuracy.
- **Autonomous Navigation System** (2020 – 2022): Developed a navigation system using NVIDIA Jetson, LIDAR, and ZED cameras with **vSLAM**, achieving 0.05s latency.

Awards & Publications

- **Top 10 Breakthroughs of the Year**, EIGER Collaboration for James Webb Space Telescope, 2023. [LINK]
- **B. Liu**, R. Bordoloi, *A Deep Learning Approach to Quasar Continuum Prediction*, Monthly Notices of the Royal Astronomical Society, Volume 502, Issue 3, April 2021, Pages 3510–3532. [PAPER]
- **B. Liu**, *Ph.D. Thesis: Study of Intergalactic Medium Using Spectroscopy and Photometry with Deep Learning*, 2022. [PAPER]
- R. Bordoloi, R. A. Simcoe, J. Matthee, D. Kashino, R. Mackenzie, S. J. Lilly, A.-C. Eilers, **B. Liu**, D. DePalma, M. Yue, and R. P. Naidu, *EIGER IV. The Cool 104 K Circumgalactic Environment of High-redshift Galaxies Reveals Remarkably Efficient Intergalactic Medium Enrichment*, The American Astronomical Society, Vol 963, Number 1, Pages 28, 2024. [PAPER]
- B. Greig, S.E.I. Bosman, F. Davies, D. Durovcikova, H. Fathivavsari, **B. Liu**, R. A. Meyer, Z. Sun, V. D'Odorico, S. Gallerani, A. Mesinger, and Y.-S. Ting, *Blind Quasar Reconstruction Challenge: Exploring Methods to Reconstruct the Lyman-Alpha Emission Line of Quasars*, Monthly Notices of the Royal Astronomical Society, Volume 533, Issue 3, September 2024, Pages 3312–3343. [PAPER]
- D. Kashino, S. J. Lilly, J. Matthee, R. Mackenzie, A.-C. Eilers, R. Bordoloi, R. A. Simcoe, R. P. Naidu, M. Yue, **B. Liu**, *EIGER VII. The evolving relationship between galaxies and the intergalactic medium in the final stages of reionization*, The Astrophysical Journal, Volume 997, Issue 2, id.280, 26 pp [PAPER]
- R. Bordoloi, **B. Liu**, *rbcodes v0.2: JWST/NIRCam Grism Spectroscopic Analysis API*, Zenodo, October 2022. [LINK]

Invited Talks

- **Grounding AI Agents with Root Cause Analysis beyond MCP Tools**, Future AI. *November 2025, Houston, TX*
- **Failure Mode and Effects Analysis Graph as MCP Server for AI Agents**, Industrial AI Nexus.*November 2025, Houston, TX*
- **Industrial AI Agents and Agent-as-a-Service**, Houston Data and AI. *September 2025, Houston, TX*
- **Empowering AI Agents with MCP and A2A Protocols**, Houston Data and AI. *June 2025, Houston, TX*
- **AI Agents and Future**, Houston Data and AI. *May 2025, Houston, TX*
- **STORM AI and UI Agent systems**, Chevron Corporation - Athena Search Seminar. *December 2024, Houston, TX*
- **Digital Rock Smart Segmentation 2.0**, Shell International Exploration and Production - Knowledge Sharing Session. *October 2024, Houston, TX*
- **8K-Resolution Segmenting Every Grain**, Shell International Exploration and Production - Digital Rock Modeling Session. *April 2024, Houston, TX*
- **Generative Diffusion Models in Action**, Shell International Exploration and Production - Generative-AI Reading Group. *June 2023, Houston, TX*
- **Explore the Epoch of Reionization Using Deep Learning**, The Statistical Challenges in Modern Astronomy VII. *June 2021, Virtual*
- **A Deep Learning Approach to Quasar Continuum Prediction**, The 236th American Astronomical Society Meeting. *June 2020, Virtual*