ABIN SHAKYA

abinshakyaa@gmail.com ◇ 2254411070 ◇ Website ◇ LinkedIn ◇ New York, NY

ABOUT

AI4Science researcher and full-stack software engineer specializing in machine learning, scientific software, DevOps, high-performance computing, and systems engineering. Currently a Postdoctoral Research Scientist at Columbia University, developing atomistic machine-learning models, agentic workflows, literature-curation platforms, scientific databases, and closed-loop discovery systems. Experienced in building scalable research infrastructure and applying AI, first-principles simulations, and molecular dynamics to problems in materials science, mineral physics, earth science, and polymer science.

PROFESSIONAL AND RESEARCH EXPERIENCE

Postdoctoral Research Scientist *June 2026 – Present*
Columbia University, New York, NY

- Working on fine-tuning atomistic machine learning models to predict interatomic forces and perform phonon calculations for mantle minerals, enabling studies of lattice vibrations, dynamical stability, and vibrational contributions to thermodynamic properties.
- Developed LitCurate, an open-source, configuration-driven LLM pipeline adaptable to literature-curation tasks across scientific domains, and applied it to create a searchable, provenance-tagged database of 1,334 lower-mantle EOS records from 205 publications. (https://litcurate.com)
- Design and maintain reproducible scientific workflows for large-scale simulations, data processing, and machine-learning–assisted materials/mineral informatics.
- Mentor interns and graduate students on HPC systems, scientific computing environments, and automation of computational research workflows.

Research Assistant (Ph.D.) *Aug 2021 – May 2026*
Louisiana State University, Baton Rouge, LA *Supervisor: Dr. Bijaya B. Karki*

- Developed an AI agent capable of parsing and analyzing scientific PDFs to extract polymer-related data and automatically generate a structured database of polymer properties and force fields for knowledge discovery and retrieval.
- Designed and implemented a GNN-based regression model using coarse-grained (MARTINI) topology graphs to predict structure–property and structure–interaction relationships in polymer–drug systems.
- Developed Weakly Supervised Graph Neural Networks for scalable **3D phase segmentation** in molecular dynamics simulations; accepted for Oral Presentation at KDIR 2025.
- Built Graph Neural Network-accelerated Molecular Dynamics to explore atom distribution in Earth's core.
- Developed an **active learning** framework for molecular dynamics that enhanced model generalization and made it feasible to model combinatorially complex chemical spaces.
- Developed deep learning-based potential energy models for molecular dynamics, facilitating simulations of Earth's interior, and authored the paper Insights into core-mantle differentiation from bulk Earth melt simulations. DOI: 10.1038/s41598-024-69873-8
- Built a two-level binning method to detect compositional shifts and verify core-mantle segregation—starting with spatial binning and refining clusters using hierarchical grouping and graph connectivity.

Senior Software Engineer *2019 – Sep 2021*
Sparrow Private Limited, Lalitpur, Nepal

- Built a Service Layer middleware that integrated SOAP and REST APIs from multiple vendors and exposed a consistent REST API to clients.
- Developed a mock environment for clients to mimic flight, hotel, and other booking flows without hitting production vendors.
- Optimized the flight search engine with caching and parallel request execution.
- Implemented AWS S3 uploads using presigned URLs for secure object access.
- Extended Python's suds SOAP library to handle read timeouts, connection timeouts, and connection errors.
- Implemented dynamic API throttling to balance security and usability under load.
- Built a custom Celery task scheduler to automate financial reports, transaction reconciliation, and daily/weekly/monthly email summaries to clients.
- Migrated the codebase from Django 1.8 to Django 2.2.
- Deployed and managed the staging server for client testing; oversaw technical documentation and mentored junior engineers.
- Assisted the accounting and finance team in identifying discrepancies in financial reports.

Software Engineer *Oct 2017 – 2019*
Sparrow Private Limited, Lalitpur, Nepal

- Developed client-facing and admin interfaces in AngularJS.
- Built REST APIs with Django REST Framework for CRUD operations across flight booking, hotel booking, and electricity bill payment.
- Integrated a web SDK to support online payments.
- Built a generic parser for SOAP-based XML responses and a custom API permissions module on top of Django REST Framework.
- Worked with the operations team as a database administrator to deliver required data, resolved customer-support tickets, and documented APIs in Swagger.

Intern *Feb 2017 – Sep 2017*
Leapfrog Technology, Inc., Kathmandu, Nepal

- Worked on real-time stock market prediction using neural networks and published the results in an IEEE conference paper.

EDUCATION

Louisiana State University, Baton Rouge — *Advisor: Dr. Bijaya Karki*
PhD in Computer Science : 4.0/4.0 GPA — *Aug 2021 - May 2026*

Tribhuvan University, Kathmandu, Nepal — *Advisor: Dr. Subarna Shakya*
BE in Computer Engineering: 3.6 GPA (US Scale) — *Nov 2013 - Nov 2017*

SKILLS

Programming Languages:	**Python**, R, Typescript, Java, C, C++, SML, SAS, SQL, MATLAB
Agentic Workflows:	LangGraph, LangChain, MCP, RAG, Tool Calling, Structured Outputs, Multi-Agent Orchestration, FastAPI, LangSmith, Vector DBs
Frameworks/Libraries:	**PyTorch, TensorFlow**, PyG, Django, Flask, PySpark, Pandas, NumPy, Scikit-Learn
Graph Machine Learning:	GCN, GAT, GraphSAGE, PNA, MPNNs, Graph Transformers, Equivariant GNNs, Link Prediction, Node Classification, Spatio-Temporal GNNs, PyG, Graph Sampling
Generative AI:	Pretraining and Fine-Tuning of Foundation Models, Normalizing Flows, Probabilistic Diffusion Models, VAEs, GANs, Transformers (GPT, BERT, T5, ViT)
Tools/Technologies/DB:	HPC, SLURM, Nginx, Redis, Amazon S3, SageMaker, Snakemake, MongoDB, SparkSQL, Apache Kafka, Apache Hadoop, RabbitMQ, PostgreSQL, LaTeX
Statistics:	A/B Testing, Hypothesis Testing, Bayesian Data Analysis, Time Series Analysis, MCMC, Gibbs Sampling, Monte Carlo Methods
MD / DFT Simulation:	Density Functional Theory (DFT), Ab Initio MD, VASP, LAMMPS, GROMACS, DeepMD-kit, DP-GEN, VMD

PAPERS, AWARDS AND RECOGNITIONS

- Shakya, A. (2026). LitCurate: A Configuration-Driven AI-Assisted Framework for Scientific Database Construction with an Application to Lower-Mantle Equation-of-State Data (Under Review)
- Shakya, A. (2026). Machine Learning Based Models for Simulation and Analysis of Bulk Earth Melt System. *Ph.D. Dissertation*, Louisiana State University and Agricultural & Mechanical College.
- Paudel, A. *, Shakya, A. *, Nag, S., Karki, B.B., and An, Y. (2026). A structured dataset of polymers in molecular dynamics simulations extracted by LLM-assisted text mining. *Nature Scientific Data.* (**Co-first author**).
- Shakya, A.*, Karki, B.B. (2025). Weakly Supervised Graph Neural Networks for Scalable 3D Phase Segmentation in Molecular Dynamics Simulations. *KDIR 2025.*
- Shakya, A.*, Ghosh, D.B., Jackson, C., Morra, G., Karki, B.B. (2024). Insights into core-mantle differentiation from bulk Earth melt simulations. *Scientific Reports.* DOI: 10.1038/s41598-024-69873-8.
- Karki, B.B., Jackson, C., Mallick, E., Shakya, A., Ghosh, D.B., Morra, G. A comparative study of nitrogen incorporation in silicate, metallic, and bulk Earth melts at high pressure. *Earth and Planetary Science Letters*, 671, 119653.
- Shakya, A.*, Pokharel, A., Bhattarai, A., Sitikhu, P., Shakya, S. (2018). Real-Time Stock Prediction Using Neural Network. *8th International Conference on Confluence 2018 (IEEE).* ISBN: 978-1-5386-1720-5.
- Enrolled in a Fully Funded PhD Program under NASA Emerging Worlds Research Grant, Computer Science, Louisiana State University, 2021/22.
- **National Rank: 26$^{\text{th}}$** (out of 15,000+) in Engineering Entrance Examination — Awarded Full Merit Scholarship.