

LEAD DATA ENGINEER

Lantern Pharma is seeking a talented and highly motivated **Lead Data Engineer** to build and enhance our RADR[®] AI platform that can predict drug response, therapeutic benefit and survival outcome of cancer patients. You will work directly with our RADR[®] team to take ownership of developing the engineering side of our platform, developing a backend data infrastructure to both support Lantern's internal AI scientists and platform as well as external-facing partnerships. You will have the opportunity to work with cutting edge data tools and be part of a team working on the bleeding edge of AI-driven oncology.

The ideal candidate will have an entrepreneurial mindset, be comfortable with creative problem solving, and will be excited to build engineering infrastructure from the ground-up in a flexible, fast-paced environment. You will be a key voice in the evolution of Lantern's science and platform, offering opportunities for growth and impact.

RESPONSIBILITIES:

- Build and maintain scalable data pipelines and infrastructure to support AI/ML drug discovery projects
- Develop and manage a data catalog that incorporates diverse biomedical data sources, including clinical and preclinical data
- Drive the development and delivery of ML ready data products
- Deliver datasets, insights, and ML models as part of an agile, cross-functional team
- Identify and track key performance indicators to be shared both internally and with company stakeholders
- Take a leadership role in ML operations using frameworks such as MLflow
- Collaborate with clinical and preclinical research teams to maintain a unified process for data ingestion and tracking
- Maintain a flexible, hands-on approach while working with cutting-edge tools and frameworks in data engineering and MLOps
- Assist in setting data engineering strategy.

REQUIRED QUALIFICATIONS:

- 5+ years experience working in engineering/development environments
- Understanding of modern engineering design principles (distributed systems, stateless processes, etc)
- Proficient working knowledge of relational and non-relational databases
- Proficient working knowledge of AWS
- Experience using GenAI tools for automated data meta-tagging and querying
- Experience using NLP, LLM, or ontological methods to resolve named entities and synonyms for data standardization
- Proficiency in programming languages (e.g., Python, R, Java, Scala, SQL). Python strongly preferred
- Familiar with version control systems.
- Strong leadership and communication skills
- Work and meet at Plano office several days a week

PREFERRED QUALIFICATIONS:

- Experience building and maintaining cloud-native, scalable data infrastructure (AWS or similar).
- Experience working with public biomedical databases or in the biopharmaceutical industry.
- Knowledge of biological data types and biomedical data structures
- Experience building ETL pipelines in a cloud native environment
- Familiarity with Data Lakehouse architectures

Lantern provides multiple growth opportunities and as an early team member, your work will have a direct impact on precision oncology that can transform drug development. In addition to attractive compensation, we offer employees the opportunity for competitive health, dental & vision insurance, stock options in a public company, an opportunity to take leadership on new and meaningful projects, & involvement with leading conferences & industry trade shows.

>>> Contact: jobs@lanternpharma.com