AI Data Engineer

Description
Mission:
At FirstIgnite, we're driving innovation by connecting cutting-edge research with industry partners ready to propel discoveries forward. Our mission is to build the most efficient software for deciphering complex scientific data and fostering valuable partnerships.

Company Status:
FirstIgnite serves a wide range of institutions and industries globally, including leading academic, research, and corporate entities. We are at the forefront of revolutionizing how research is commercialized, seamlessly integrating scientific research with industry opportunities, and enabling advances that shape the future.

Role Overview:
The AI Data Engineer will play a pivotal role in architecting and maintaining data systems that manage a diverse array of data sources. This role includes applying AI models to enhance data insights and build connections that drive our platform forward.

Responsibility Statement

- Develop and optimize scalable data pipelines for integrating data from patents, grants, clinical trials, publications, labs, and firmographic sources.
- Design and manage a Neo4j graph database to organize and connect complex data relationships, supporting the use of AI-driven techniques.
- Explore and implement additional data storage solutions to complement graph approaches, if necessary.
- Deploy machine learning solutions, ranging from standard ML models to graph-based AI models.
- Ensure a high level of data integrity, accuracy, security, and governance throughout the data lifecycle.
- Work cross-functionally to innovate and deploy AI-driven insights that support FirstIgnite's mission of industry-research engagement.
- Implement tools for data exploration and visualization to translate data insights into actionable business decisions.

Qualifications

- Proven experience with Neo4j or related graph databases and designing graph structures.
- Expertise in programming languages, particularly Python, and related data analysis and machine learning libraries (e.g., Pandas, Scikit-learn, TensorFlow, PyTorch).
- Proficiency with machine learning platforms such as Google Cloud AI Platform (Vertex AI), AWS SageMaker, or Azure Machine Learning.
- Experience with RESTful and GraphQL APIs for data integration and strategy evaluation.
- Familiarity with ETL processes for efficient data transformation and integration.
- Analytical skills for conducting cost-benefit analyses of data ingestion methods, such as bulk storage versus real-time API use.
- Understanding of Retrieval-Augmented Generation (RAG) techniques to

Hiring Manager
Cody Pawlowski

Employment Type
Full-time

Job Location
Remote work from: BR / Remote (BR)

Experience
3+ years

Base Salary
$ 45K - $ 60K

Date posted
May 15, 2025

Valid through
31.07.2025

enhance AI applications.
- Strong version control skills using Git and experience with CI/CD pipelines.
- Problem-solving skills and the ability to work independently and collaboratively with cross-functional teams.

Skills
Python, Neo4j, Natural Language Processing, Data Warehousing, ETL, Data Modeling

Why FirstIgnite?
At FirstIgnite, we're not just building software – we're connecting the dots between research and industry to catalyze innovation. Join us in a role that challenges norms, embraces forward-thinking, and puts you at the forefront of technological advancement. If you're passionate about using data to drive meaningful progress, we want to hear from you.

Role: Engineering
Eng Type: Data science

Remote: Yes
US Visas: US citizenship/visa not required
Equity: 0.10% – 0.24%