

Lead Data Engineer

Waterloo, Ontario | Full-time | Compensation: TBD + Equity

Who we are

Sidian is disrupting architecture, engineering and construction (AEC). We began September 2025 with mid-sized engineering firms, raised a US\$1M pre-seed round at the end of May and are now moving into enterprise, with 3 of the top 30 global firms entering the platform in 3 months. Our mission is to lead AEC's technological revolution, bringing the same step-change already reshaping medicine, finance and law. Our purpose is simple: to make humanity's basic needs affordable and sustainable, starting with infrastructure.

Our Team

A small but mighty team of 4. James is the CEO with 4 years of experience in AEC, a varsity football player at Waterloo, and Sanford Fleming recipient for top academic students. Mankirat is the CTO, an ex-founder, and Valedictorian of engineering at Waterloo. Cardoso is a founding engineer with a Masters in AI/ML and expertise in knowledge graphs and GNNs. Finally, our beloved coops.

The role

Engineers today work with data trapped in silos 3D models, BIM files, PDF drawings, spreadsheets, Word documents, emails — that not even other engineers, let alone AI, can easily read across. At Sidian, we're building the company's brain that brings all the engineering firm's knowledge to their fingertips.

As our Lead Data Engineer, you'll design and build the pipelines that ingest these diverse, messy engineering data types and unify them into a structured, queryable knowledge graph. You'll ensure that every engineering decision is captured, connected, and retrievable — not just going forward, but reconstructed from past projects as well. The goal: make hard-won engineering knowledge a durable, leverageable asset rather than something locked in files.

You'll work directly with the founders, own core architecture, ship product, learn from customers, and shape the engineering culture. You'll turn fast-moving AI capabilities into dependable systems.

What you'll own

- Build and improve the multimodal ingestion pipelines that transform drawings, BIM, PDFs, spreadsheets, documents and project history into reliable, structured data.
- Architect and advance the company knowledge graph so engineering decisions across current and past projects are tracked, connected, and retrievable.
- Develop the retrieval and reasoning systems on top of the graph that make this knowledge usable for search, recommendations, and automation.
- Design the data models and polyglot storage that hold heterogeneous engineering data together at scale.
- Work with customers to turn ambiguous engineering workflows into simple, dependable products.

What you bring

- Strong data engineering skills in Python, with sound judgment across backend and data systems.
- Hands-on experience with knowledge graphs and graph databases (Neo4j, Cypher), including modeling complex relationships.
- Experience building production data pipelines over modern infrastructure: Apache Spark, Databricks, Parquet, and polyglot / multi-store architectures.
- Comfort ingesting and parsing messy, heterogeneous formats (3D/BIM, PDF drawings, Excel, Word, email) into structured data.
- Evidence of initiative — an independent project, open-source contribution, research build, or startup you drove from idea to working system.

Compensation and ownership

Want to build?

- Resume and GitHub portfolio
- 3-5 sentences on why you are the right person for the job and how you align with our mission.

Message: msains@sidian.ai