

# Data Architect / Data Engineer

## About Liveline

Liveline enables dramatic improvements in manufacturing performance through a unique application of artificial intelligence to provide real-time process control and predictive assistants for plant personnel. Our focus is on automating complex processes, not simply providing dashboards for managers and operators.

Our team combines experts in AI with world-class process engineers who can focus on the “last mile” with customers: Extracting data from the process and implementing controls on the shop floor. We speak the language of AI but also industrial controllers.

Our hardware and software offerings are scalable and cost-effective whether customers have one production line or hundreds, delivering an ROI that’s attractive to small and medium-sized enterprises.

We are passionate about democratizing the power of analytics and advanced automation for manufacturers of almost any size. Through our approach, producers can de-mystify complex processes and free up valuable technicians to focus on more advanced tasks instead of constantly monitoring and adjusting equipment parameters.

## Role Description

The Data Architect/Engineer is responsible for developing, deploying, and maintaining an end-to-end technology stack to manage real-time data flows. They enable advanced machine learning manufacturing control and maintain the heartbeat between the customer and Liveline. The ideal candidate is capable to take the big picture and executing an ideal data structure for real-time machine learning applications. This role will build the data pipeline from the ground up, creating an architecture that will scale as Liveline grows.

## Primary Responsibilities

- Design and implement enterprise data infrastructure to support data pipelining, modeling and visualization by providing the right data, at the right time, at an optimal cost
- Works within the Liveline team to develop requirements and ensure workflow needs are satisfied
- Ensure seamless infrastructure integration and security
- Ensure compatibility across infrastructure architecture and with customers’ environments
- Some programming background like Python, Numpy, Pandas, Sklearn, SciPy, Statsmodels, R, Java, etc
- Data Cleaning, Data Integration and Transformation, Data Reduction, Feature Engineering.
- Knowledge on Outlier Analysis, Evolution Analysis.
- Build the meaningful data visualization.

### Experience:

- Bachelor's Degree in engineering or computer science
- Five plus years of experience in enterprise data architecture
- Strong experience with conceptual, logical, and physical data modeling using databases
- Experience with data pipelines, database storage and basic ETL.
- Statistical Data Engineering and Statsmodels.
- Working experience with Windows, Linux/GNU

### Hard Skills:

- Experience with relational database, including writing complex SQL queries in PostgreSQL and MSSQL
- Proficient with cloud platforms: AWS and Azure
- Experience with time series and/or no-sql databases, including Influxdb, MongoDB
- Experience with statistical data analyses and smart dashboards
- Proven experience with data management concepts and activities including an understanding of industry best practices for Data Sourcing, Key Data Elements, Data Quality Management, Metadata, and Enterprise Data Management Policy, Process, and Procedures
- Knowledge in Time Series Data Engineering, Resampling and Frequency conversions.
- Some programming background like Python, R, Data Frame, Java, JS, etc
- Experience building robust data ingestion pipelines; Redshift/Glue, Snowflake, Hadoop/Spark/Kafka, Schema for Multidimensional Databases
- Experience with ETL, Datawarehouse, Data marts, and data integrity
- Experience with converting data sources between database formats and flat files/JSON.
- Data Loading, Storage, and File Format, Data Wrangling.

### Soft Skills:

- Demonstrated ability to work in teams to solve business problems
- Strong attention to detail and ability to document findings and convey information.
- Excellent interpersonal, analytical, computer & software skills
- Ability to explain moderately complex information to others in concise manner
- Good written and presentation skills
- Willingness and ability to travel, as necessary

## Interview questions:

Give your opinion on Hadoop vs Snowflake vs Cloud (redshift)?

How have you handled differences of opinion on the future architecture? How have you convinced or how would you convince upper level management in IT that your solution is superior?

What data better to be on cloud vs on prem?

Give us views on a corporate databases and data lakes? Do you prefer a rigid structure or open object storage? What are the use cases for these?

How do you connect disparate data warehouses?

What SQL and nosql databases have you used? Give an example of converting between databases, ie converting to or from relational.

Give examples of large architectures that you have designed?

Have you coordinated design and implementation with outside vendors or team members?

What's the last time you apologize to anyone? What's the reason?

## Whiteboard

How do you move and store time series, batch and flatfiles. Interaction between plants to a regional datacenter to a global location. On prem and cloud hybrid  
Draw the connections and interactions based on our input

Keys to success

Time series and rational databases.

Connections

Asks us about needs

Discusses experiences they have with databases

Credentials and user management

Interactions with corporate systems