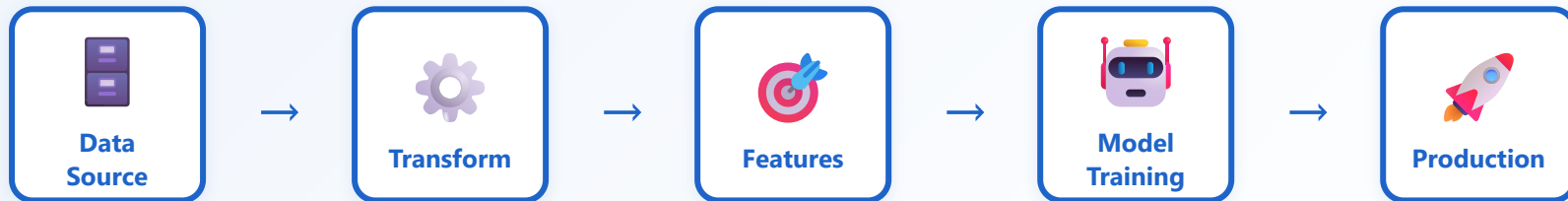


## Data Lineage

Track the complete history and flow of data through ML pipelines



### **Data Provenance**

Where did training data originate?

### **Transformation Chain**

All preprocessing and feature engineering steps

### **Model Lineage**

Which data version trained which model version?

### **Impact Analysis**

If data source changes, which models are affected?

### **Tools & Platforms**

MLflow Tracking

DataHub

Apache Atlas

DVC