

Learn to Use Databricks for the full ML Lifecycle

Rafi Kurlansik, Sr. Solutions Architect

databricks **Lakehouse Platform**

Simple | Open | Collaborative

Data Science &
Engineering

BI &
SQL Analytics

Machine Learning

Real-time Data
Applications

Data Management & Governance



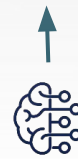
Open Data Storage



Structured



Semi-structured



Unstructured



Streaming

Reliable | Scalable | Secure



databricks **Lakehouse Platform**

Simple | Open | Collaborative

Data Science & Engineering

BI & SQL Analytics

Machine Learning

Real-time Data Applications

Our focus today

Data Management & Governance



Open Data Storage



Structured



Semi-structured



Unstructured



Streaming

Reliable | Scalable | Secure



Essential Capabilities for Full ML Lifecycle

To lower risk and maintain stability of our ML pipeline, we need to think about:

- **Robust Data Processing and Management**
- **Secure Collaboration**
- **Testing**
- **Monitoring**
- **Reproducibility**
- **Documentation**

...for code, data and models.

Business Context: Customer Retention

You are on a marketing analytics team and you have a lot of **demographic** and **historical service data** on your customers that have **churned**, which has been put into a SQL Analytics dashboard.

The data team has been asked by business stakeholders if you can go further and **predict** which customers are likely to churn. Knowing this will allow the business to take action and **retain revenue**.

Sounds simple enough. What steps do we need to take?

Operational Workflow At a Glance:



Data Engineer



Data Scientist



ML Engineer

Operational Workflow:
Data Engineering



Data Engineer

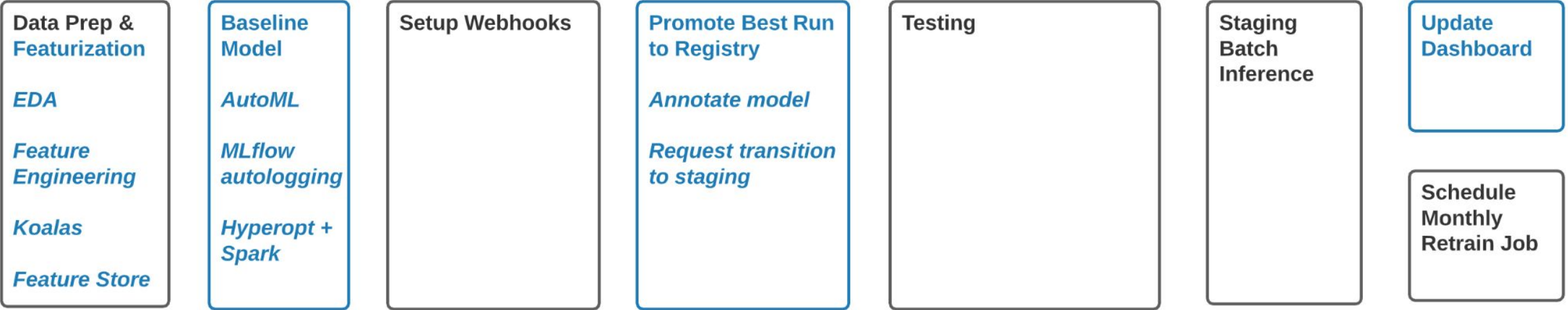


Data Scientist



ML Engineer

Operational Workflow:
Data Scientist



Data Engineer

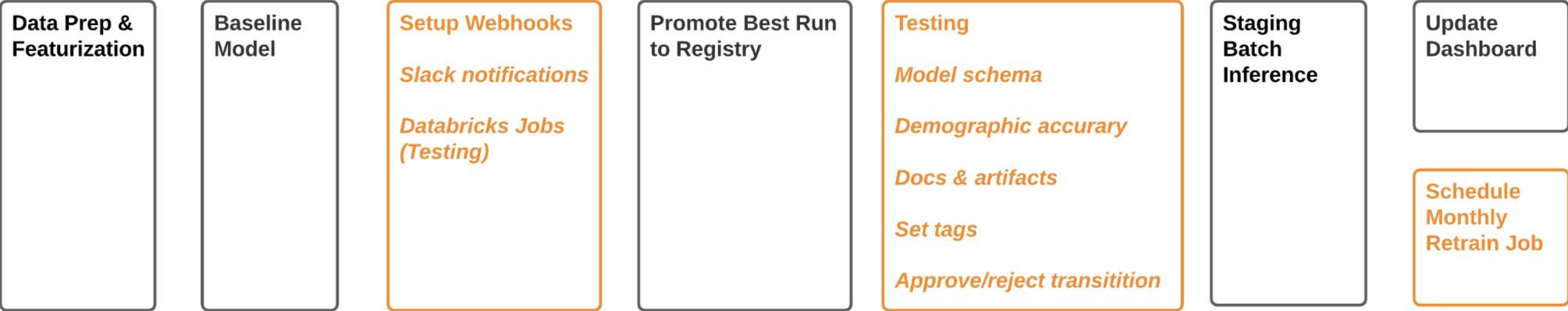


Data Scientist



ML Engineer

Operational Workflow:
ML Engineering



Data Engineer

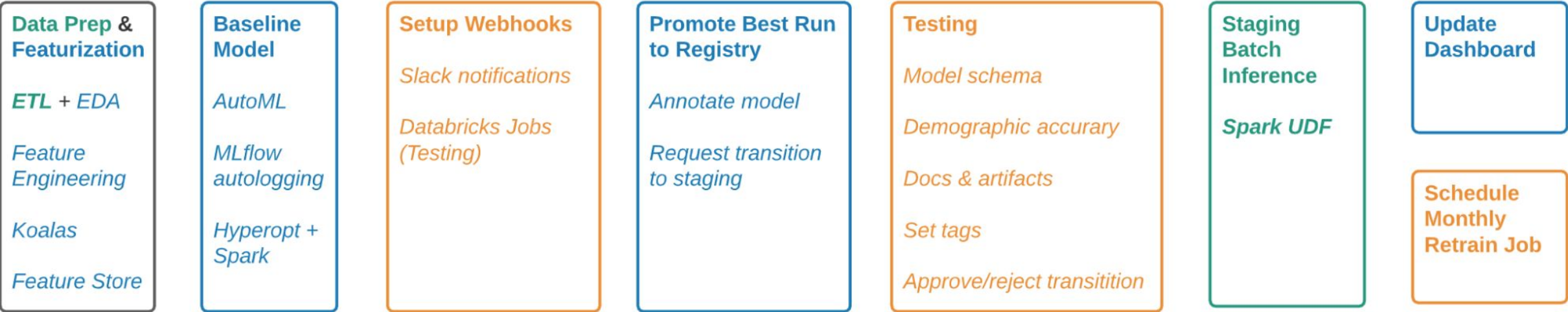


Data Scientist



ML Engineer

Operational Workflow:
Data Team



Data Engineer



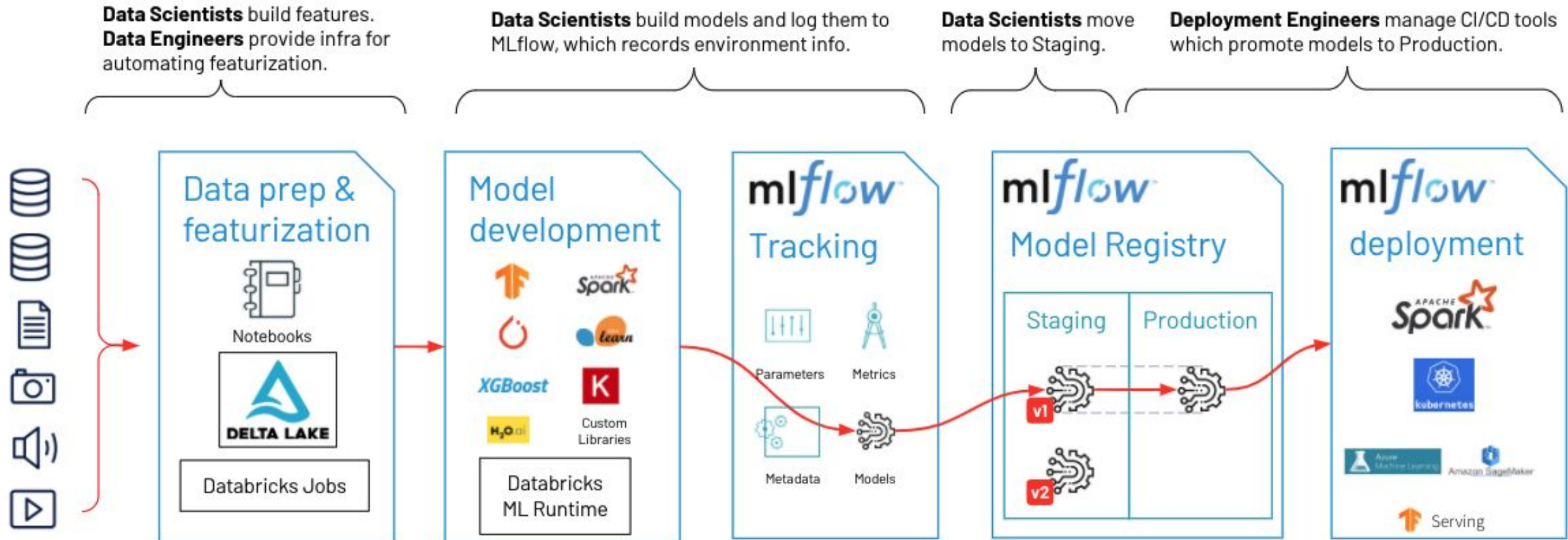
Data Scientist



ML Engineer

To the demo we go!

The Full ML Lifecycle



To learn more:

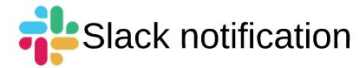
We'll be releasing a series of blogs on MLOps and ML Engineering throughout 2021:

- *The Need for Data-centric ML Platforms*
- *Selecting Technologies and Platforms for Data Science and Machine Learning*
- *Model and Data Monitoring on Databricks*
- ... and more

Check out the other MLOps talks at DAIS!

Thank you!

Automation with webhooks



Promote Best Run to Registry

