FEATURE STORE SUMMIT

12-13 OCTOBER | 08:30 AM - 4:00 PM PT

ORGANIZED BY HOPSWORKS

databricks

© 2021 Databricks Inc. - All rights reserved

Databricks Feature Store

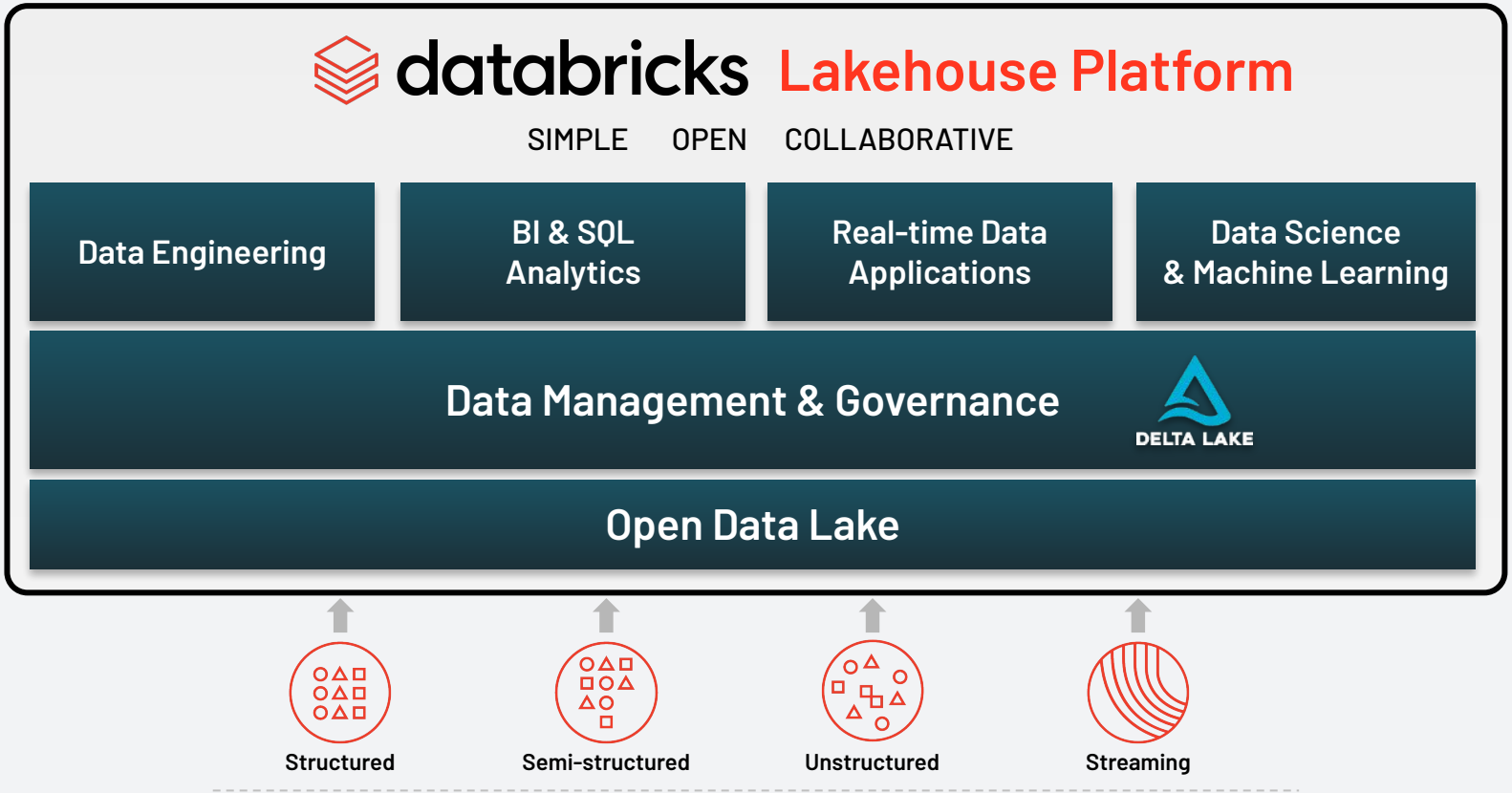
co-designed with a Data and ML-ops platform

Mani Parkhe
ML Tech Lead
Databricks

Avesh Singh
Sr. Software Engineer
Databricks

databricks

© 2021 Databricks Inc. - All rights reserved

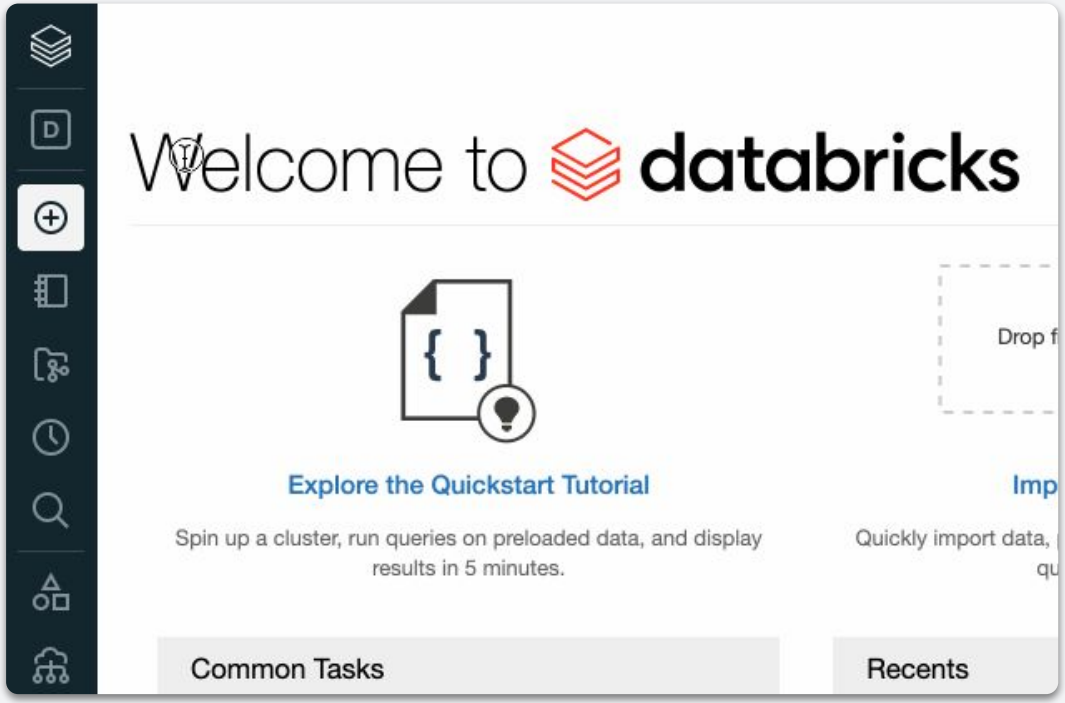


databricks

© 2021 Databricks Inc. - All rights reserved

Feature Stores for ML

Databricks Machine Learning



All the power of Databricks, plus:

- ML ready environments (e.g. for TensorFlow)
- Experiment Tracking + Reproducibility
- Model Interpretability/Understanding
- Model Registry (Model Lifecycle Management)
- Model Serving
- AutoML
- Feature Store + Feature Serving

databricks

© 2021 Databricks Inc. - All rights reserved

Feature Stores for ML

The data problem in AI

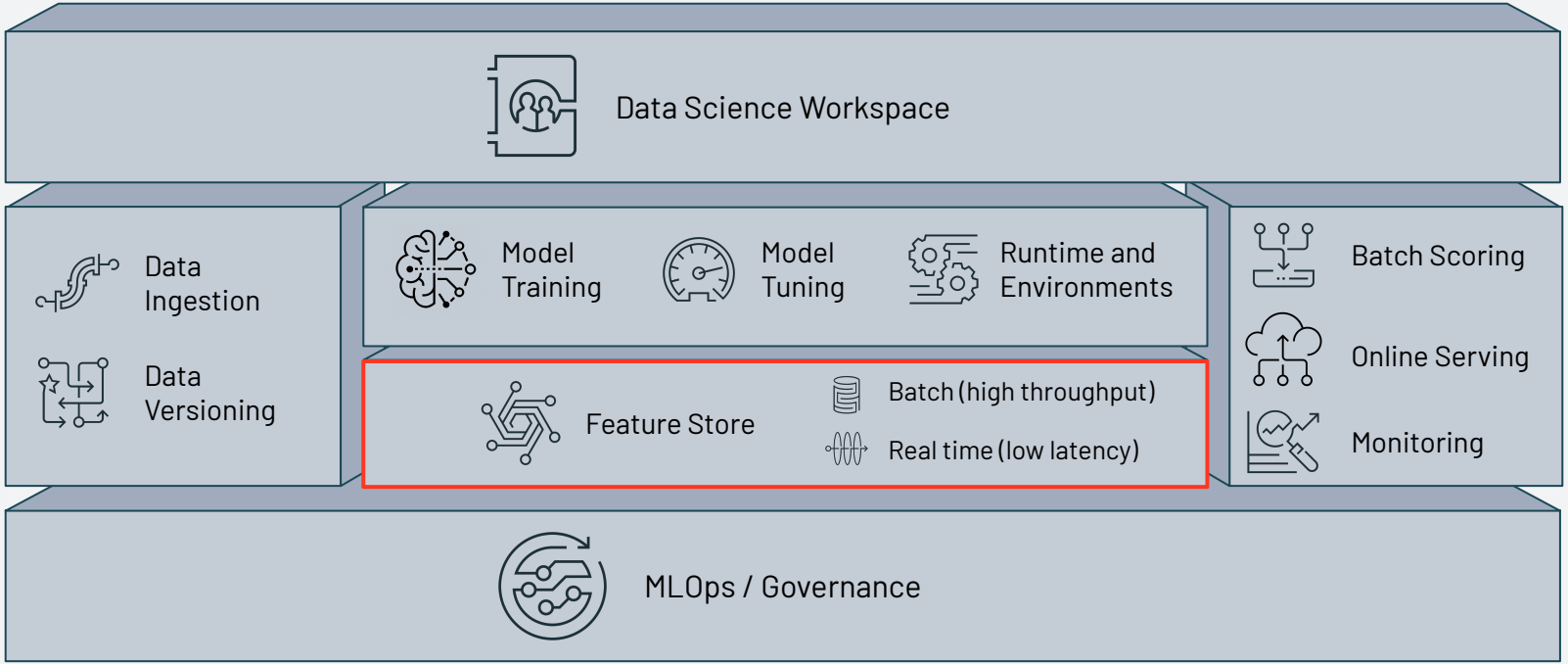
1. **Complexity** of data models, data types, custom computation logic
2. **Online/offline skew** affecting model performance
3. **Reusability** and **Discoverability** of existing features
4. **Proprietary data formats**, **tooling**, and **vendor lock in**

databricks

© 2021 Databricks Inc. - All rights reserved

Feature Stores for ML

Databricks Feature Store



databricks

© 2021 Databricks Inc. - All rights reserved

Feature Stores for ML

Databricks Feature Store

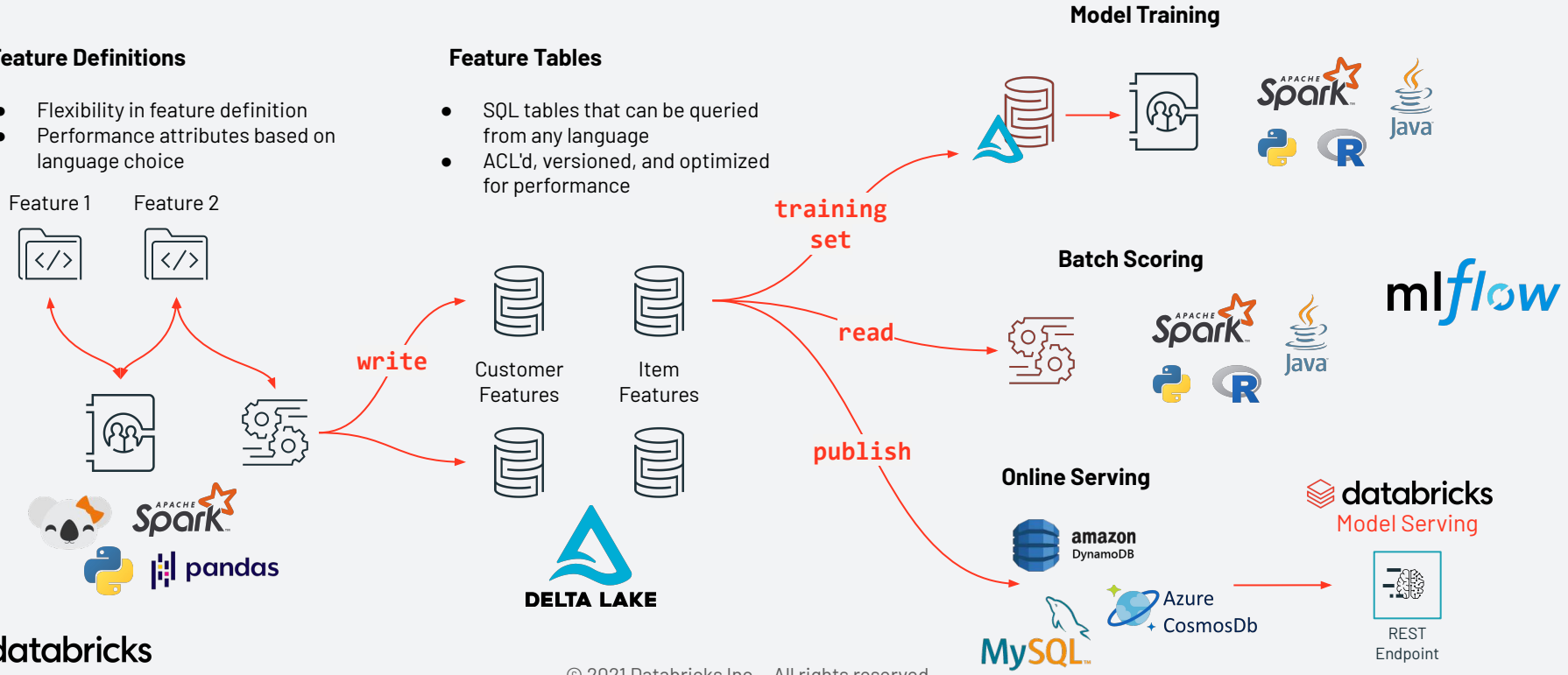
1. Open data and model formats

Feature Definitions

- Flexibility in feature definition
- Performance attributes based on language choice

Feature Tables

- SQL tables that can be queried from any language
- ACL'd, versioned, and optimized for performance



databricks

© 2021 Databricks Inc. - All rights reserved

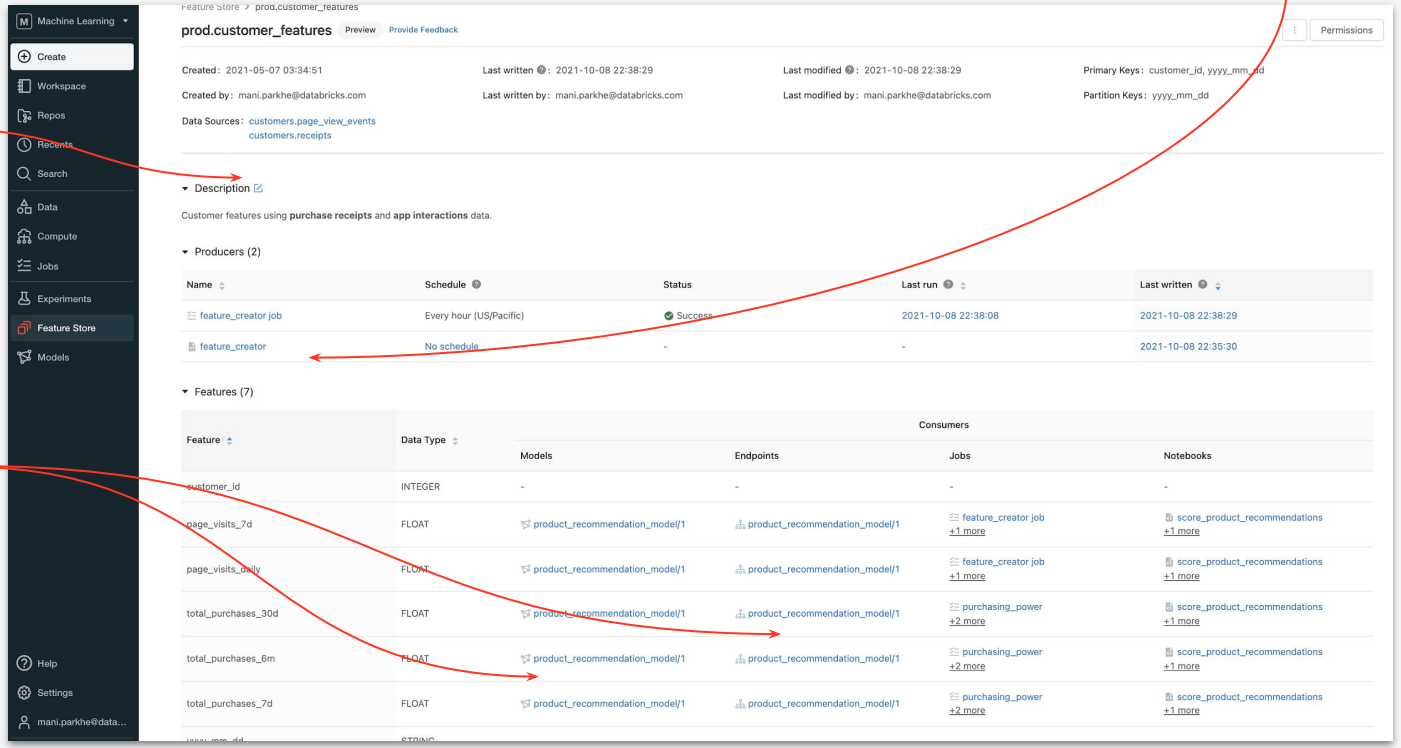
Databricks Feature Store

2. Backward and Forward Lineage Tracking

Scheduled jobs and notebooks that compute the features

Table and file sources used for feature computation

Consumers like model, endpoints, jobs, .. etc that use these features.

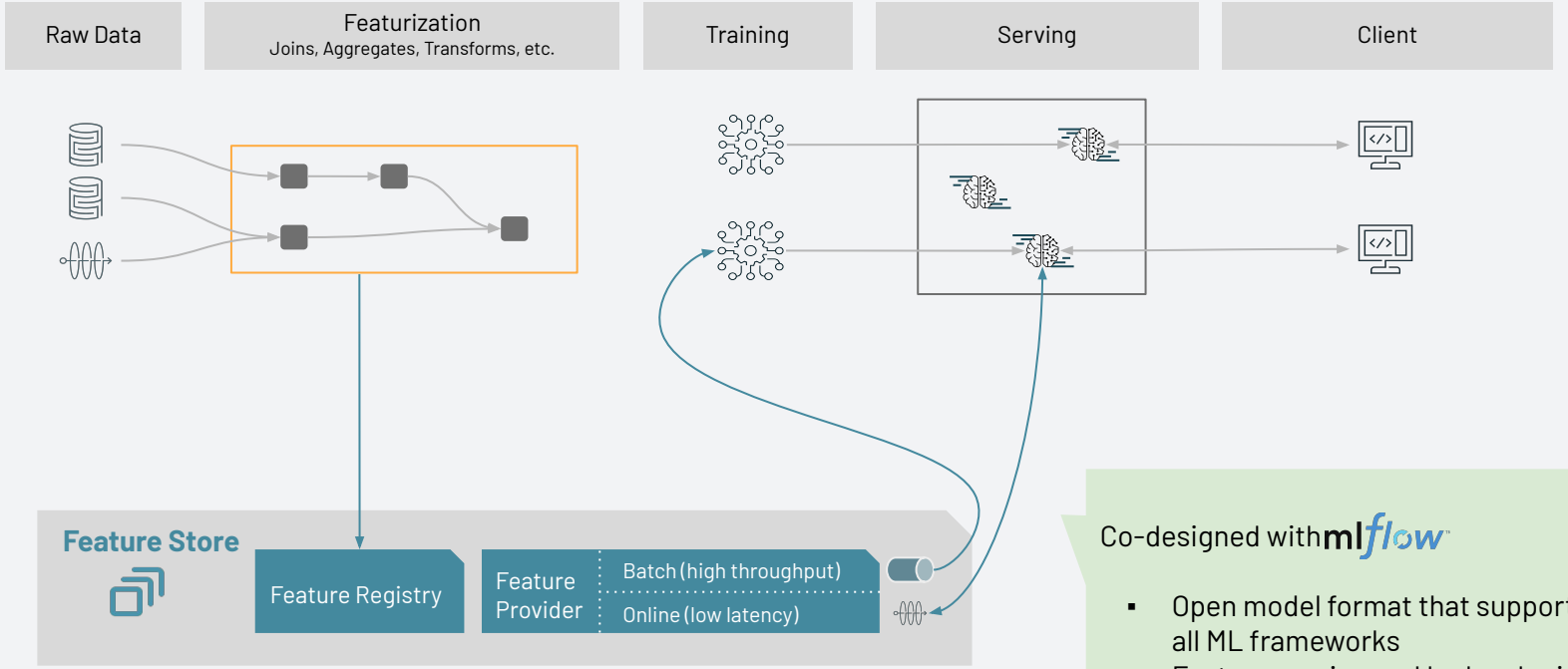


© 2021 Databricks Inc. - All rights reserved

Feature Stores for ML

Databricks Feature Store

3. Seamless integration with MLflow and in-built feature lookup



Co-designed with mlflow

- Open model format that supports all ML frameworks
- Feature version and lookup logic hermetically logged with Model

databricks

© 2021 Databricks Inc. - All rights reserved

Thank you!

Do you have any questions?

Mani Parkhe
mani@databricks.com

databricks

© 2021 Databricks Inc. - All rights reserved