

Automated Feature Discovery to Power Your Feature Store

Feature Store Summit

Oct 11, 2022



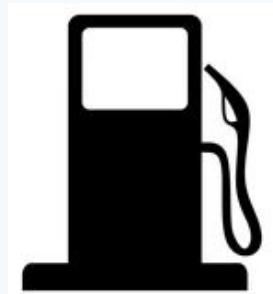
Features and Feature Store

Feature Discovery



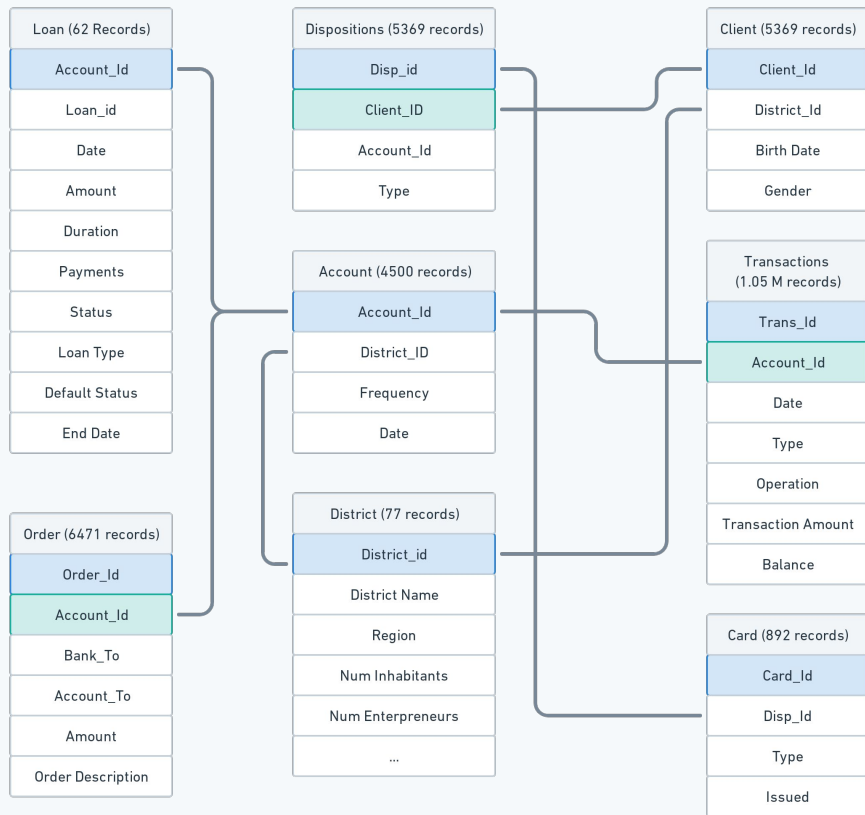
Features need to be drilled and extracted from raw data

Feature Store



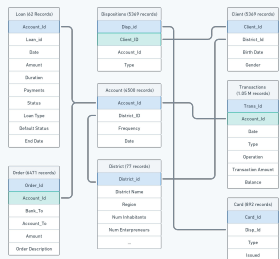
Feature is the fuel for ML

Feature Discovery / Feature Engineering



- Czech Bank loan data
- Predict client loan defaults
- 8 tables, different cardinalities
- Timestamped-transactions

Feature Discovery / Feature Engineering



Transform **multiple tables**
into a flat feature table

Data Quality and Cleansing

- Data Schema Matching
- String Value Canonicalization
- Missing Values & Outliers
- Data Skewness

Multi-table Transformation

- Normalize / Denormalize
- Auto-Join / Aggregation
- Temporal relationship
- Table combinations

Transform **multiple columns**
to higher-order features

Value Transformation

- Categorical Encoding
- Numeric Binarization
- Timestamp Featurization

Higher-Order /Combinatorial Features

- Neural Encoding (e.g. Auto-encoder)
- Wavelet Transformation

Bricks



Feature Store Delta Live Table

4 Register features into Feature Store / Delta Live Table



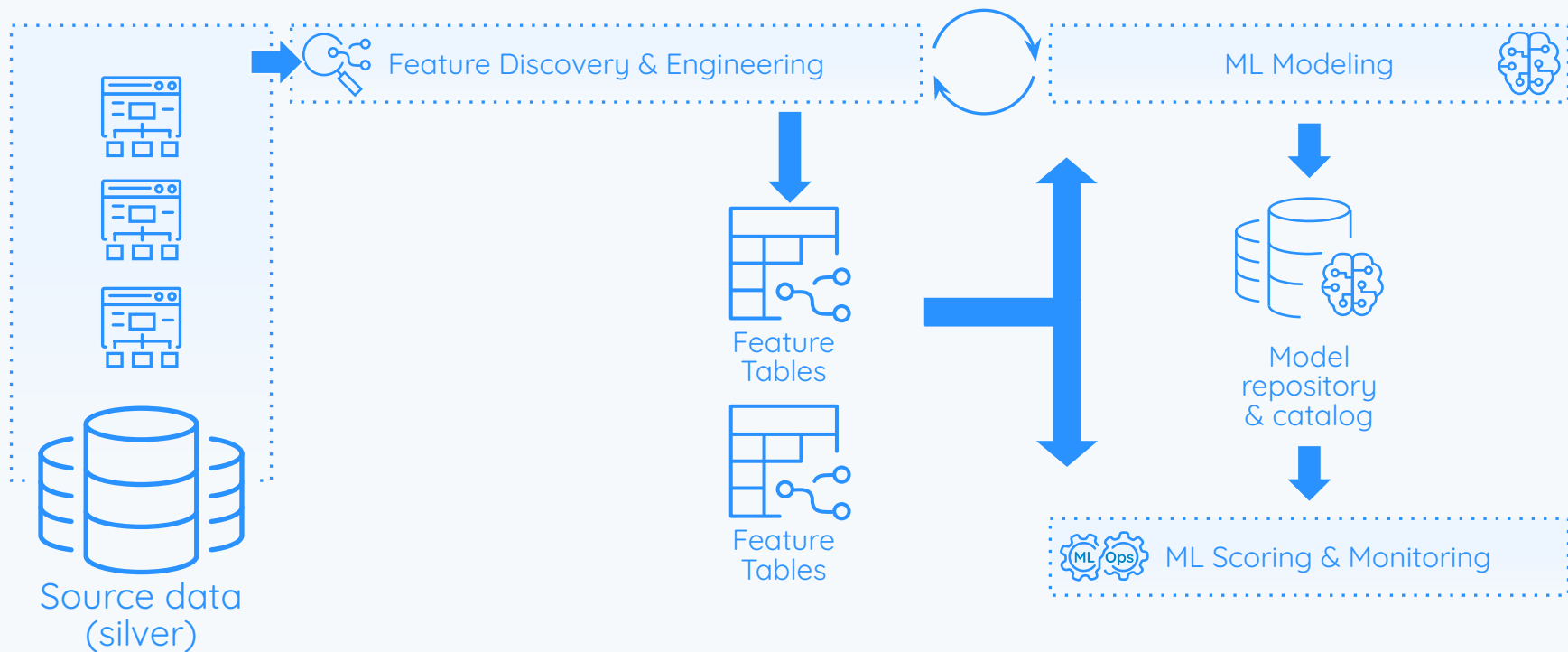
DELTA LAKE

0 Prepare Silver data and metadata

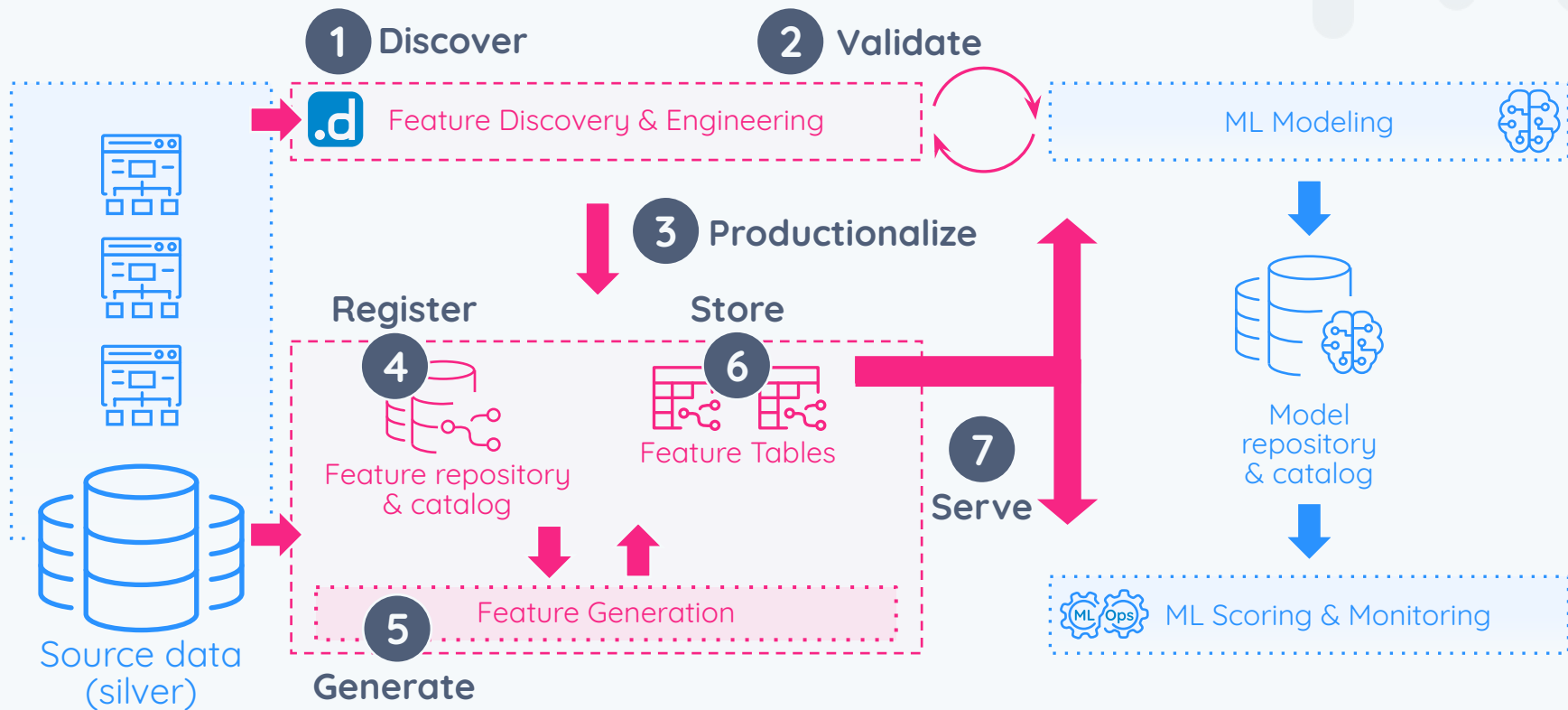


3 Run Feature Factory on Databricks

Conventional AI/ML Development



Feature Discovery with Feature Store



Award-winning feature discovery platform

2015
Automated
Feature
Engineering

2018
HQ established
in San Mateo,
CA

2019
Closed Series A
(Investment to
date \$43M)

2020
AWS Competency,
ISV Accelerate & MS
Azure partnerships

2022
Closed Series B
(Investment to
date \$74.6M)



“dotData’s exceptional feature engineering capabilities make it especially suitable for the most challenging use cases.” **Forrester**





From Tailor Made
To Mass Production



dotData
Feature Discovery



Feature Factory + Store
is A New Paradigm

Thank you!

<https://dotdata.com>
contact@dotdata.com
lulu.liu@dotdata.com