# Feature Store Usage Patterns
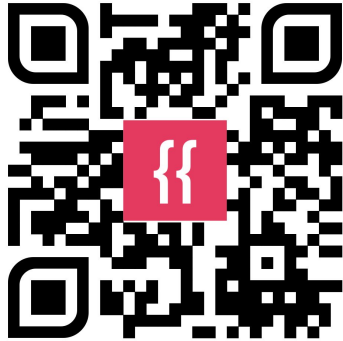
**From a Single Data Scientist to an Enterprise**

Simba Khadder, Creator, Featureform

# Who am I

**Featureform Repo**



**LinkedIn**



**Simba Khadder**
Co-founder & CEO

# Two problems to be solved

## Processing Problem

## Organization Problem

# MLOps is about organizational problems

ML Platform  ⟵  ▤ ▤ ▤ ▤ ▤  ⟶

MLOps  ⟵  👤 👤 👤 👤 👤  ⟶

# Two problems to be solved

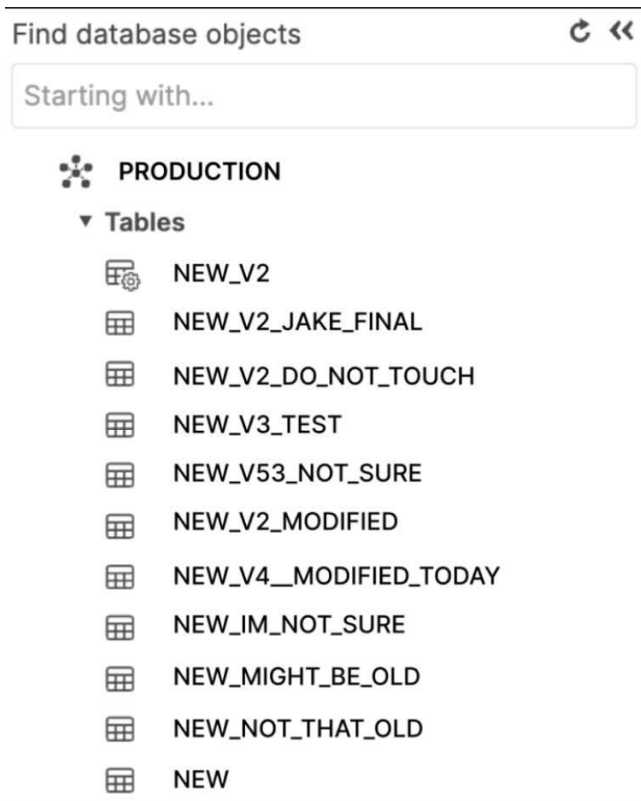## Processing Problem

## Organization Problem

# Processing Problem

- I need 10ms latency on this feature.

- I need to process and store 100TB of data

- I need streaming data to process exactly once and be up to date to the last second.

# Organization Problem

- Who owns this feature?

- What happens if X leaves the company?

- Is this feature still being used?

- What features are our ad serving model using?

- Am I allowed to use PII in this model?
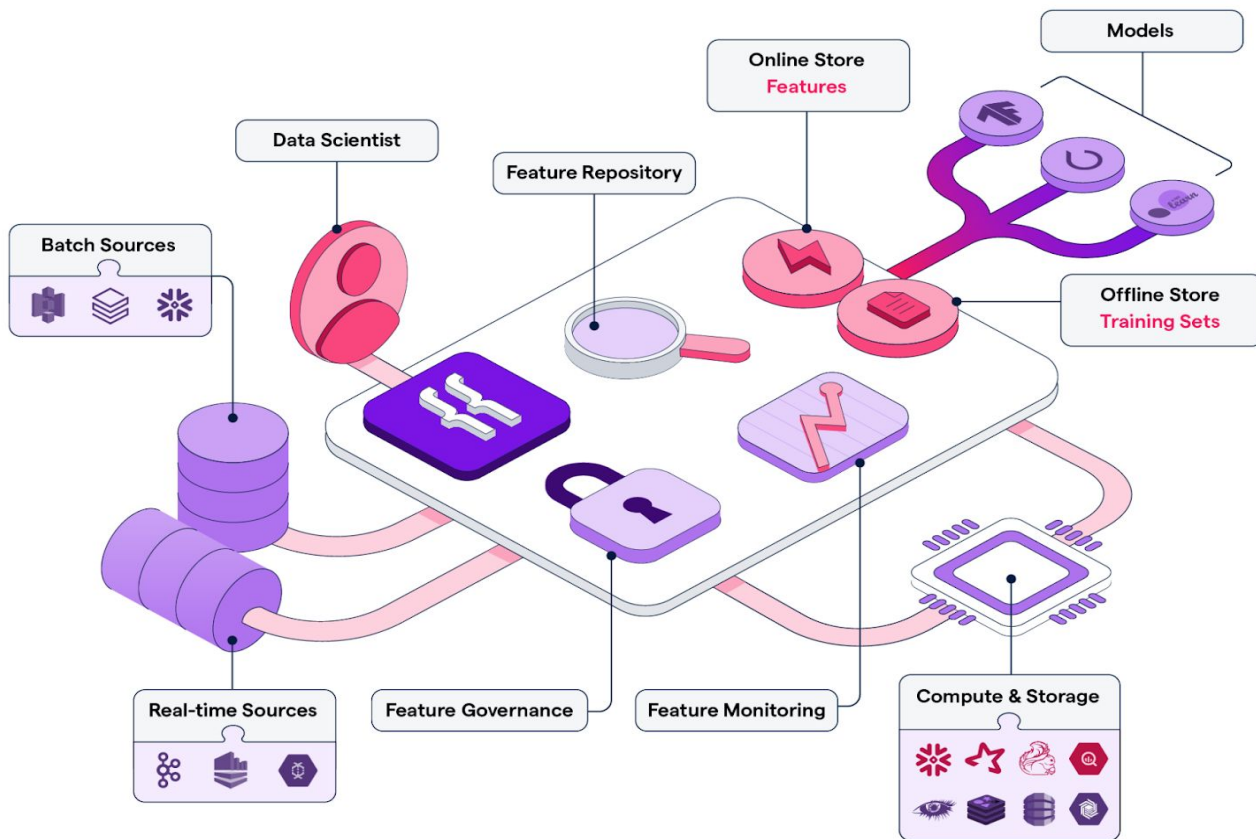
- How is this feature defined?

# Scaling up with no MLOps
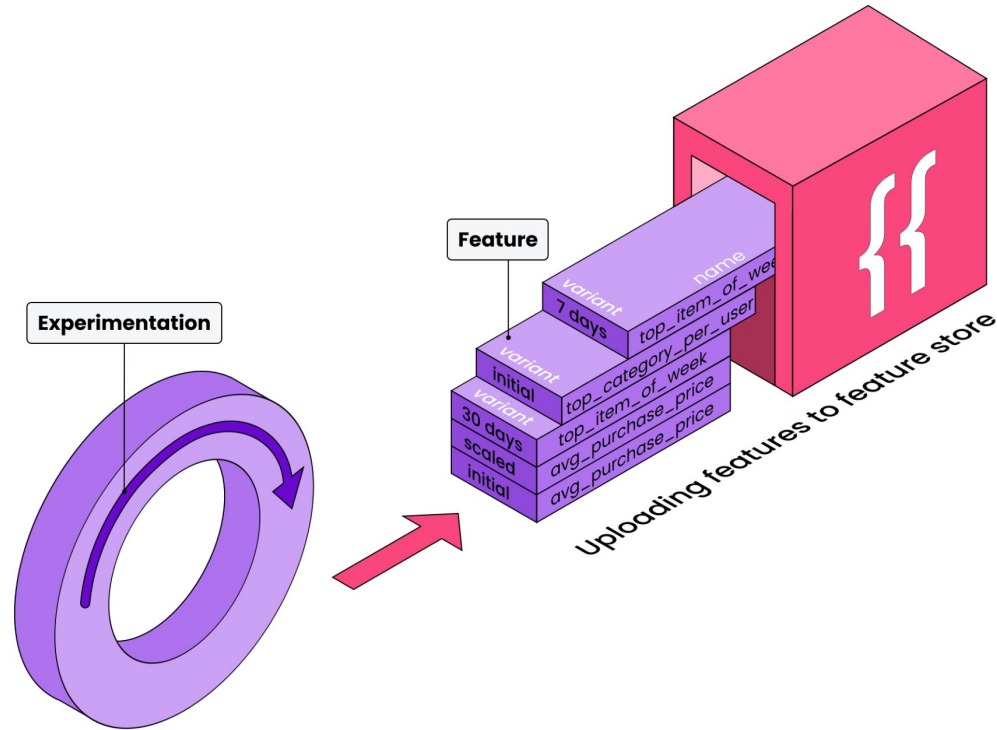
# Virtual Feature Store Brief

- Enables data scientists to **define**, **manage**, and **serve** their ML model's features

- Sits **atop your existing infrastructure** and orchestrates it to work like a **traditional feature store**
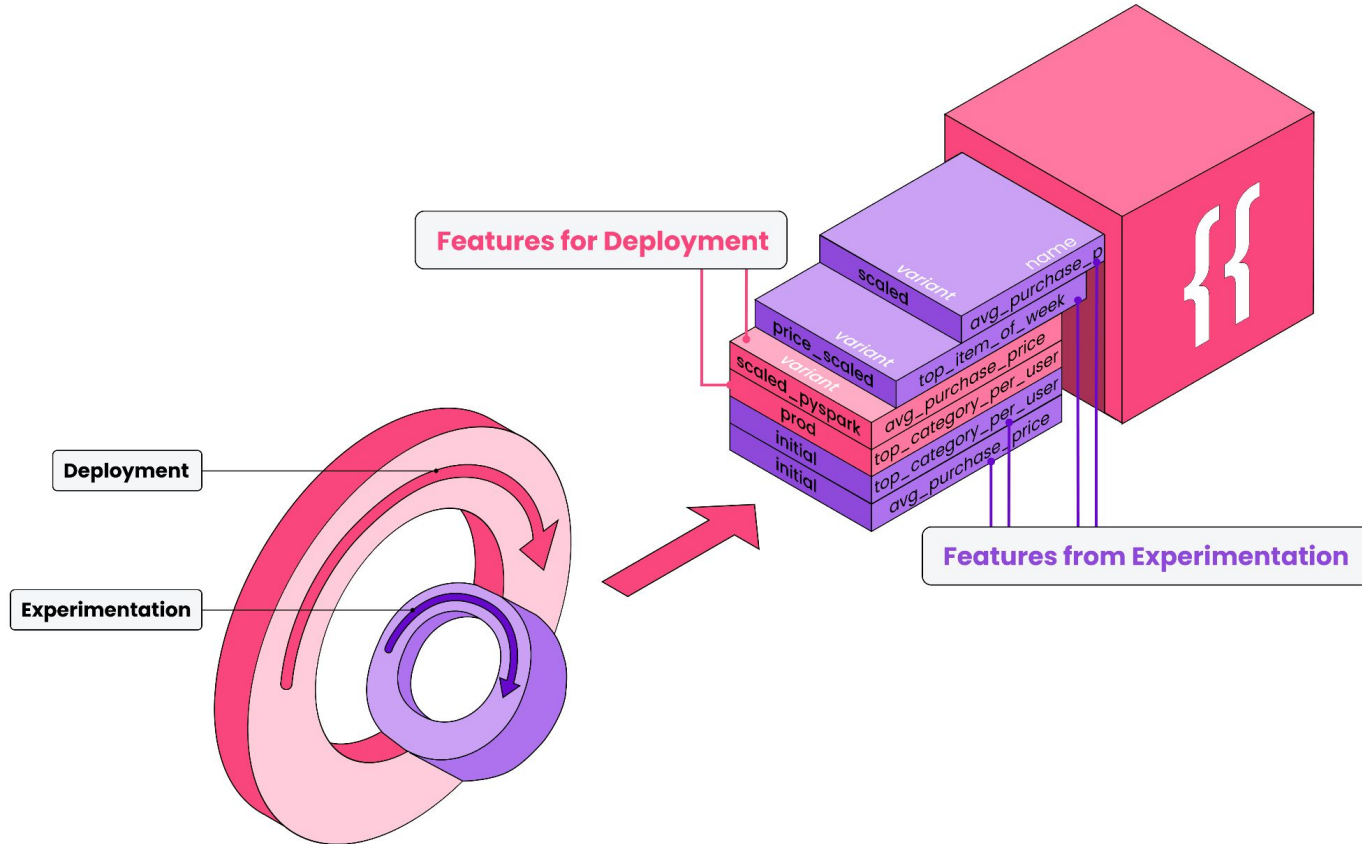
# **Virtual** Feature Store



Models

Online Store
**Features**

Data Scientist

Feature Repository

Batch Sources

Offline Store
**Training Sets**

Real-time Sources

Feature Governance

Feature Monitoring

Compute & Storage

# One Data Scientist, Locally



Experimentation

Feature

variant — name
7 days — top_item_of_week
variant
initial — top_category_per_user
variant — top_item_of_week
30 days — avg_purchase_price
scaled
initial — avg_purchase_price

Uploading features to feature store

# Purpose

- Store and organize transformations, features, and training sets in a purpose built repository

- Keep a History of Changes

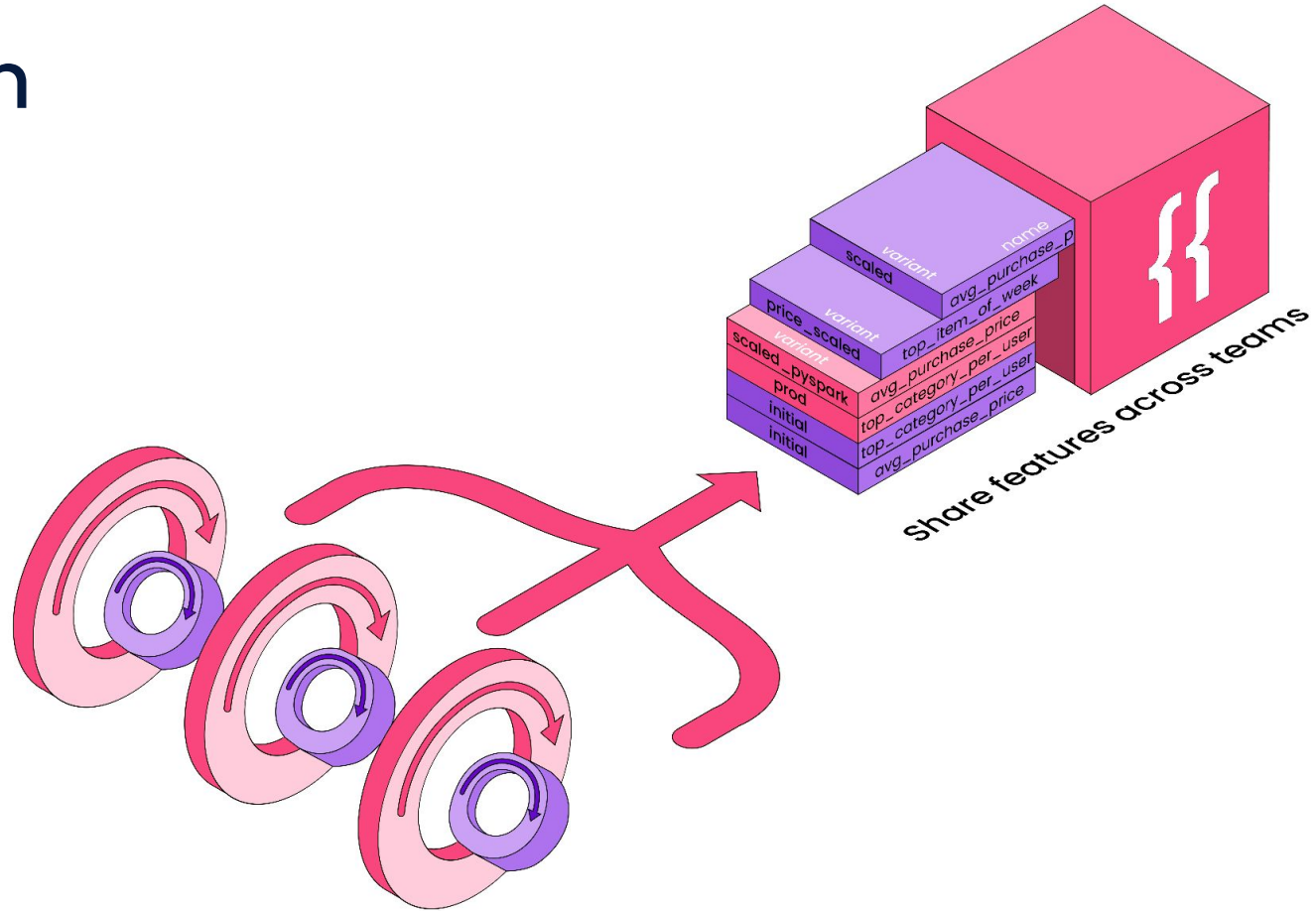- Be able to progress to a larger organization

# One Data Scientist



**Features for Deployment**

scaled
variant
name
avg_purchase_p

price_scaled
variant
top_item_of_week
avg_purchase_price

scaled_pyspark
variant
avg_category_per_user
top_category_per_user

prod
top_category_per_user

initial
avg_purchase_price

initial

**Deployment**

**Experimentation**

**Features from Experimentation**

# Purpose

- Simplify Feature Deployment

- Monitor Production Features

- Unify Experimental and Production Features
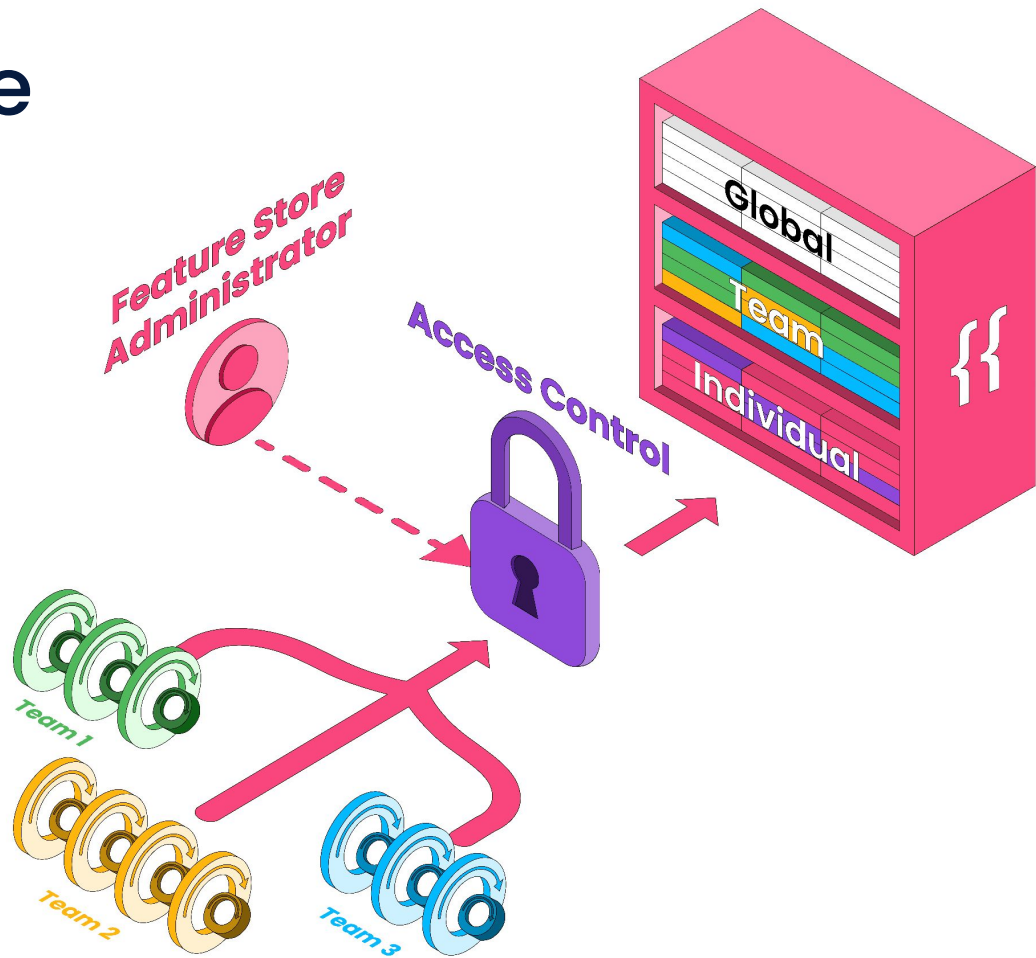
# Team



share features across teams

15

# Purpose

- Enhanced Communication & Collaboration

- Discover & Understand Each Other's Work

- Re-use & Share resources with immutable definitions
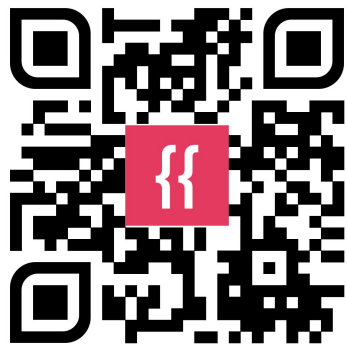
# Enterprise

# Purpose

- Hierarchical Search & Discovery

- Heterogeneous Infrastructure

- Governance, Access Control, Audit Logs

- Cross-Functional Workflows

# Thanks!

**Featureform Repo**



**LinkedIn**



**Simba Khadder**
Co-founder & CEO