Unified Data Platforms and the Rise of Data Platform Engineering

dagster

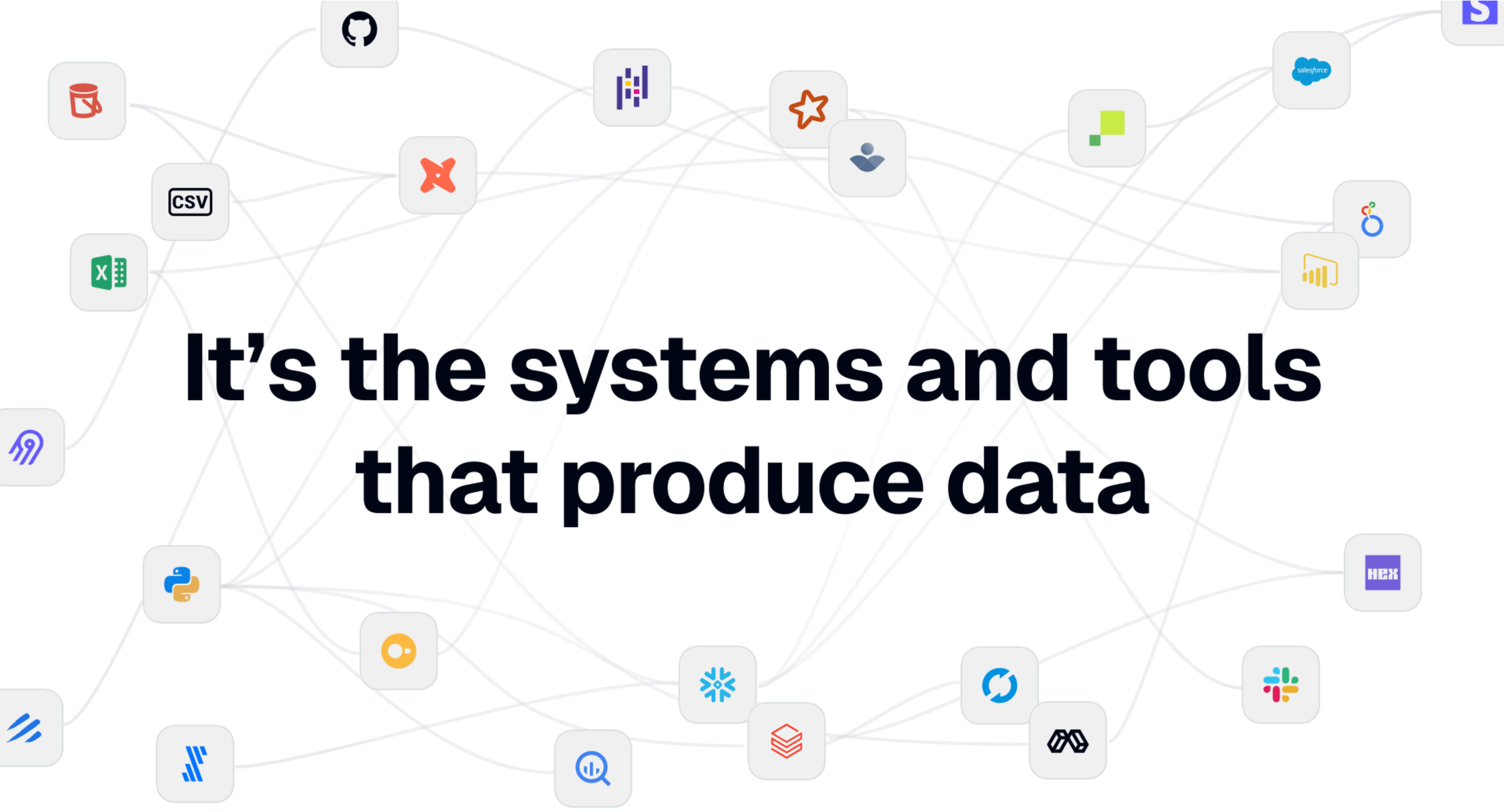
Dagster as the Unified Control Plane for your Data Platform

dagster

Day 1

Dagster as the Unified Control Plane for your **Data Platform**

Every company has a data platform



It’s the systems and tools that produce data

Your Data Platform

Context-specific

Costly: time, money, and energy

Foundational: operations, decisions, AI

dagster labs

Our mission is to empower every organization to build a productive, scalable data platform.

Data-platform orientation is spreading

Formal job titles and org structures

Dagster is the natural tool of choice

dagster

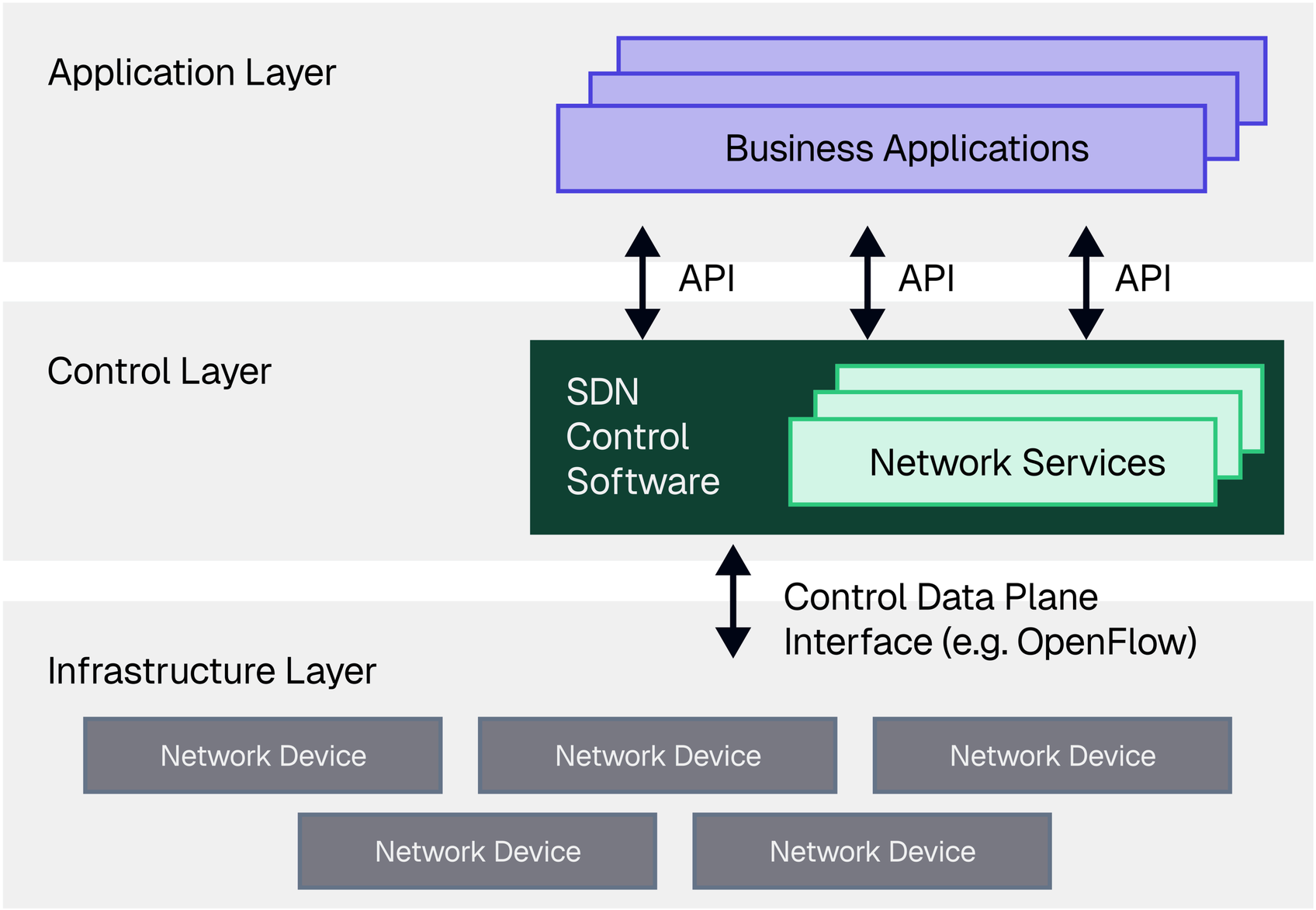
Unified Control Plane for your Data Platform

dagster

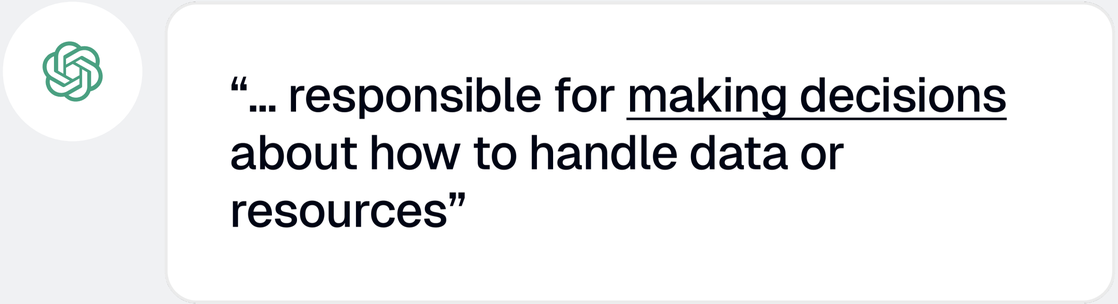
Unified **Control Plane** for your Data Platform

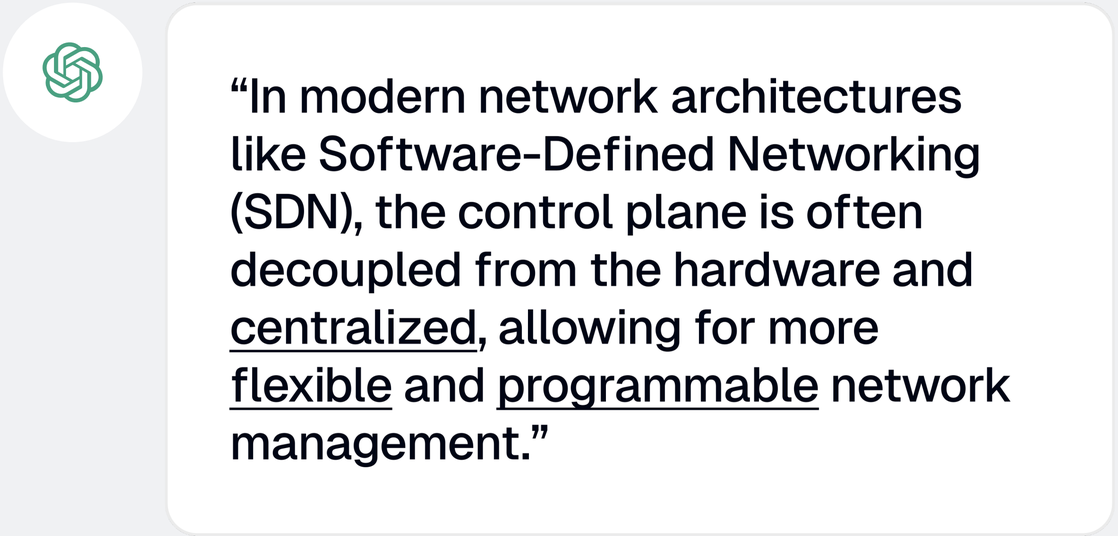
Control planes exist in many domains

What does control plane even mean?


Application Layer
Business Applications
API
API
API
Control Layer
SDN
Control
Software
Network Services
Control Data Plane
Interface (e.g. OpenFlow)
Infrastructure Layer
Network Device
Network Device
Network Device
Network Device
Network Device

What do the LLMs say?





Let’s adapt it to data platforms

“... responsible for orchestrating data or resources”

“In modern data platforms, the control plane is often decoupled from data processing and centralized, allowing for more flexible and programmable data management.”

The Control Plane

The Control Plane

An Orchestrator

Programmable

Flexible

Centralized

An Orchestrator

An Orchestrator

Decides when to run computations, in what order, and with what configuration

Asset-oriented: Defines, in software, the assets that should exist.

Better user experience and a reliable system of record

Flexible and Programmable

Control Layer

Asset Definitions

Orch e st rat i o n

Application Layer

L ine age

Catalog

Operations

Custom Tools

Control Layer

Asset Definitions

Orch e st rat i o n

Application Layer

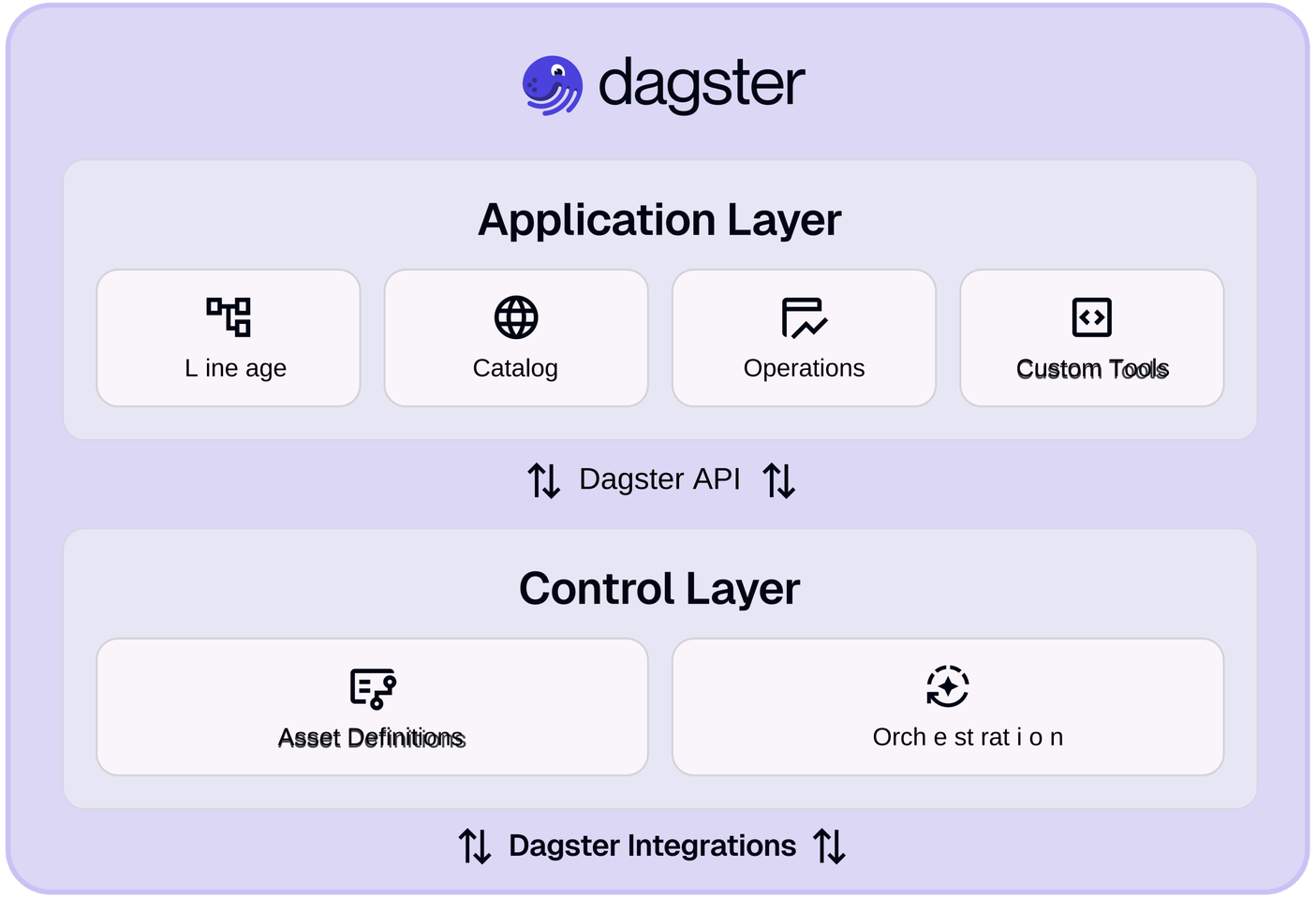
- L ine age
- Catalog
- Operations
- Custom Tools

Control Layer

- Asset Definitions
- Orch e st rat i o n

Data Processing Layer

- dbt
- SDF
- S l ing
- DLT
- Snow fl a ke
- Sp a rk
- Databricks
- Modal



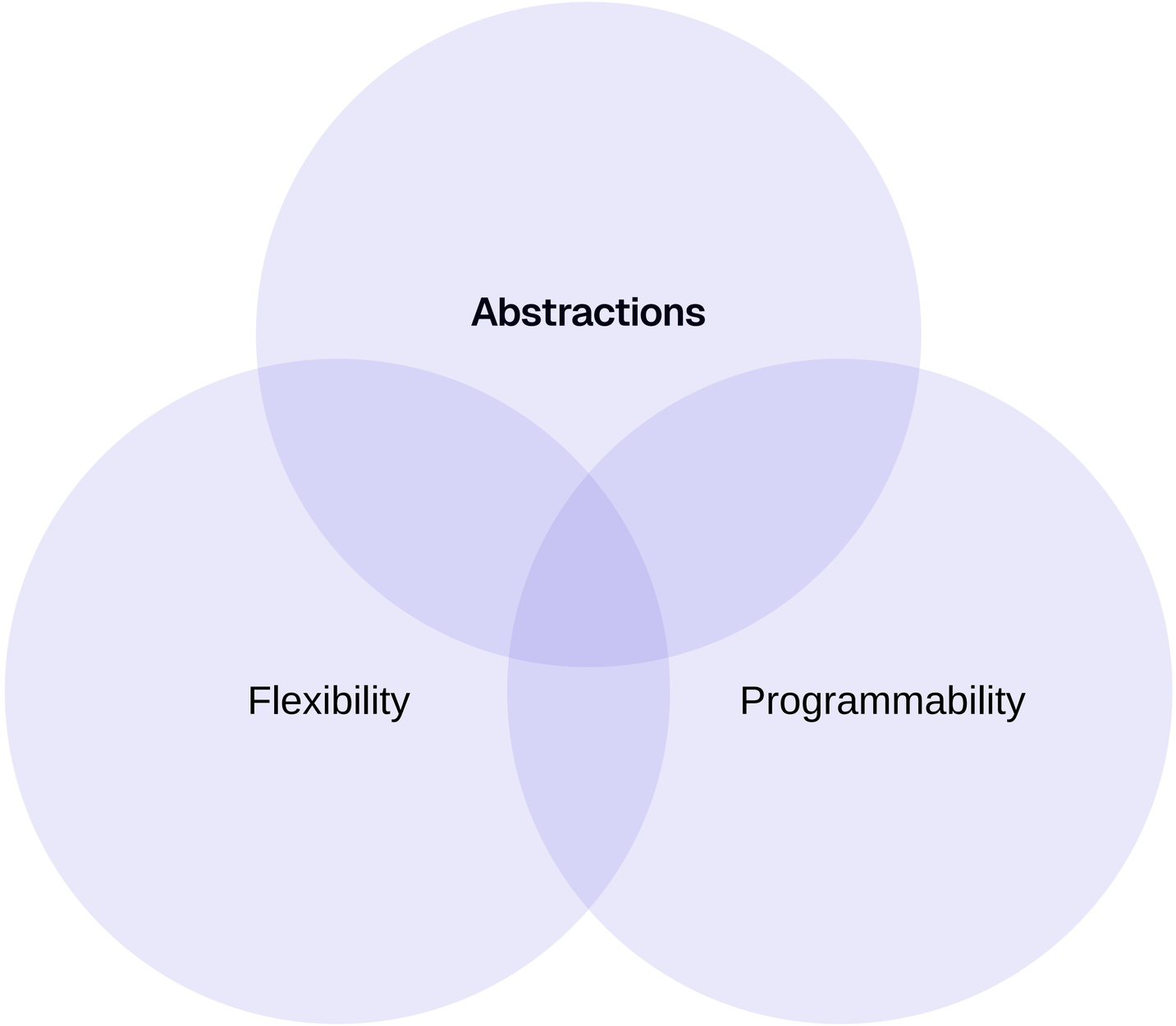
Data Processing Layer

dbt | SDF | S l ing | DLT

Snow fl a ke | Sp a rk | Databricks | Modal

Flexible and Programmable

- Any tool
- Used by any practitioner
- Deployed to any infrastructure


Abstractions
Flexibility
Programmability

Without the right abstractions

You don’t get enough richness and value

Without programmability

You cannot handle the complexity

Without flexibility

You cannot cover all use cases and contexts

dagster

Unified Control Plane for your Data Platform

dagster

Unified Control Plane for your Data Platform

Data Practitioners

Autonomy

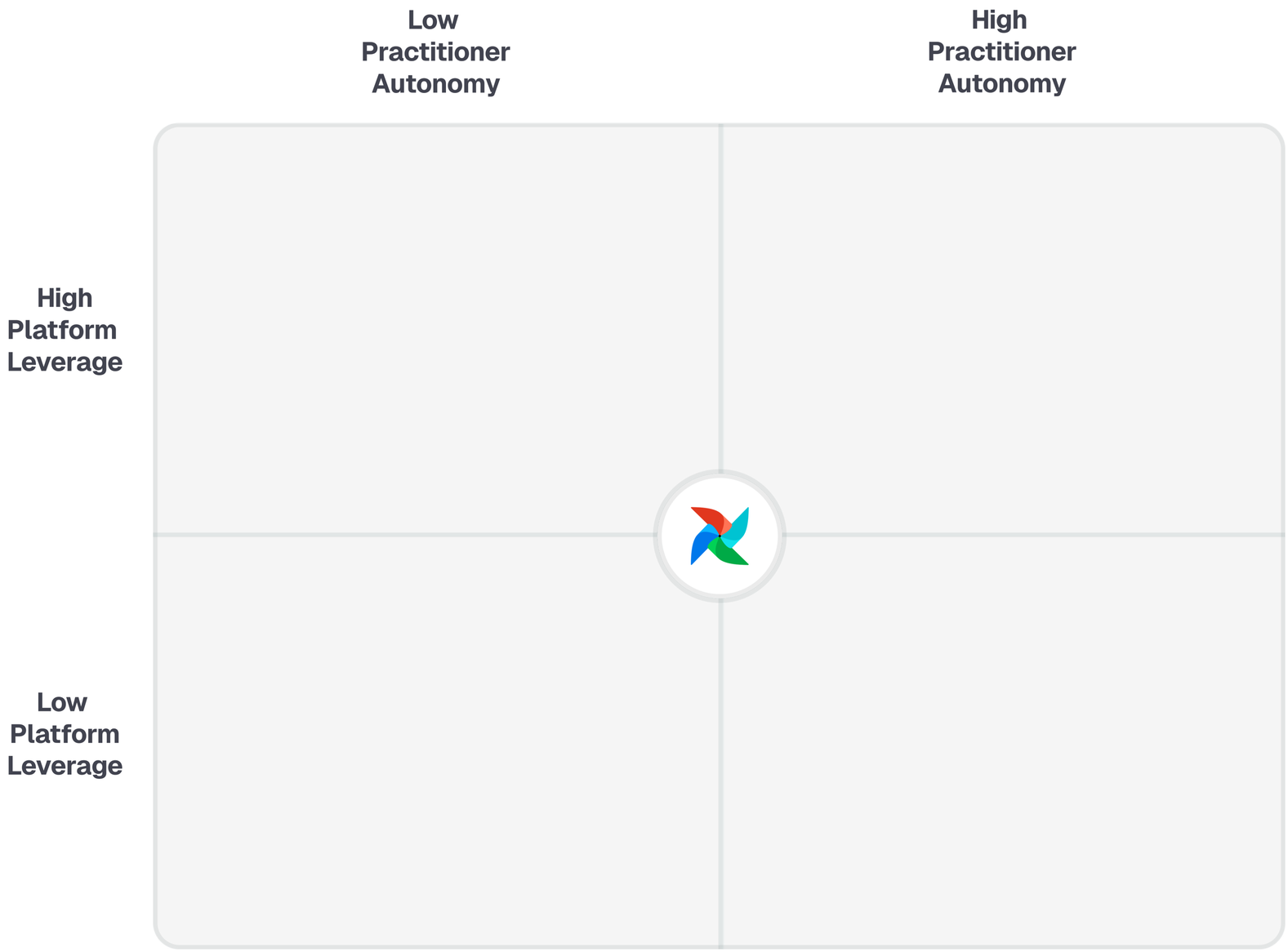
- Focus on business logic in tool of choice
- Self-serve their pipelines in a productive SDLC
- Work as close to their business stakeholders as possible

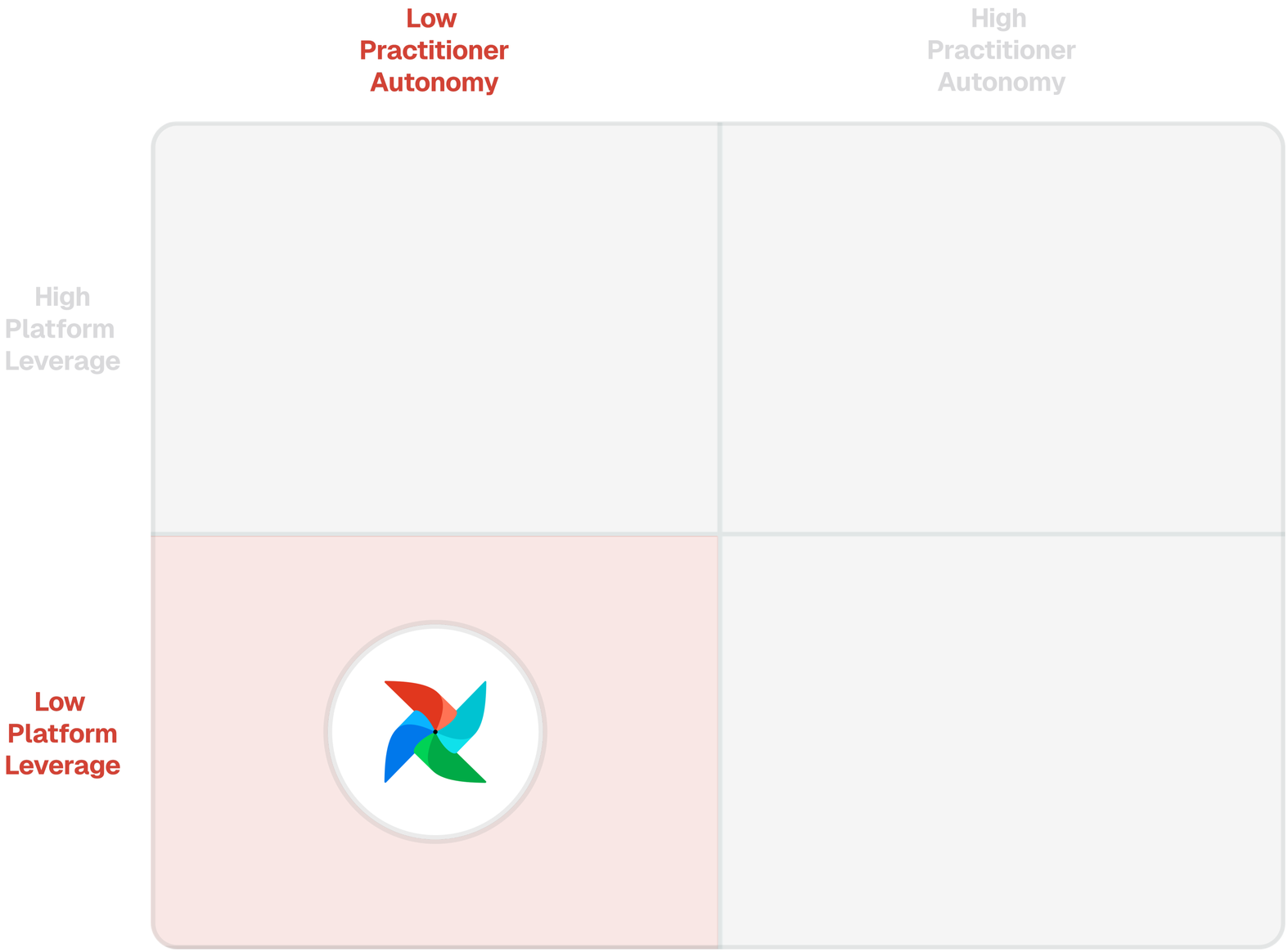
Data Platform Team

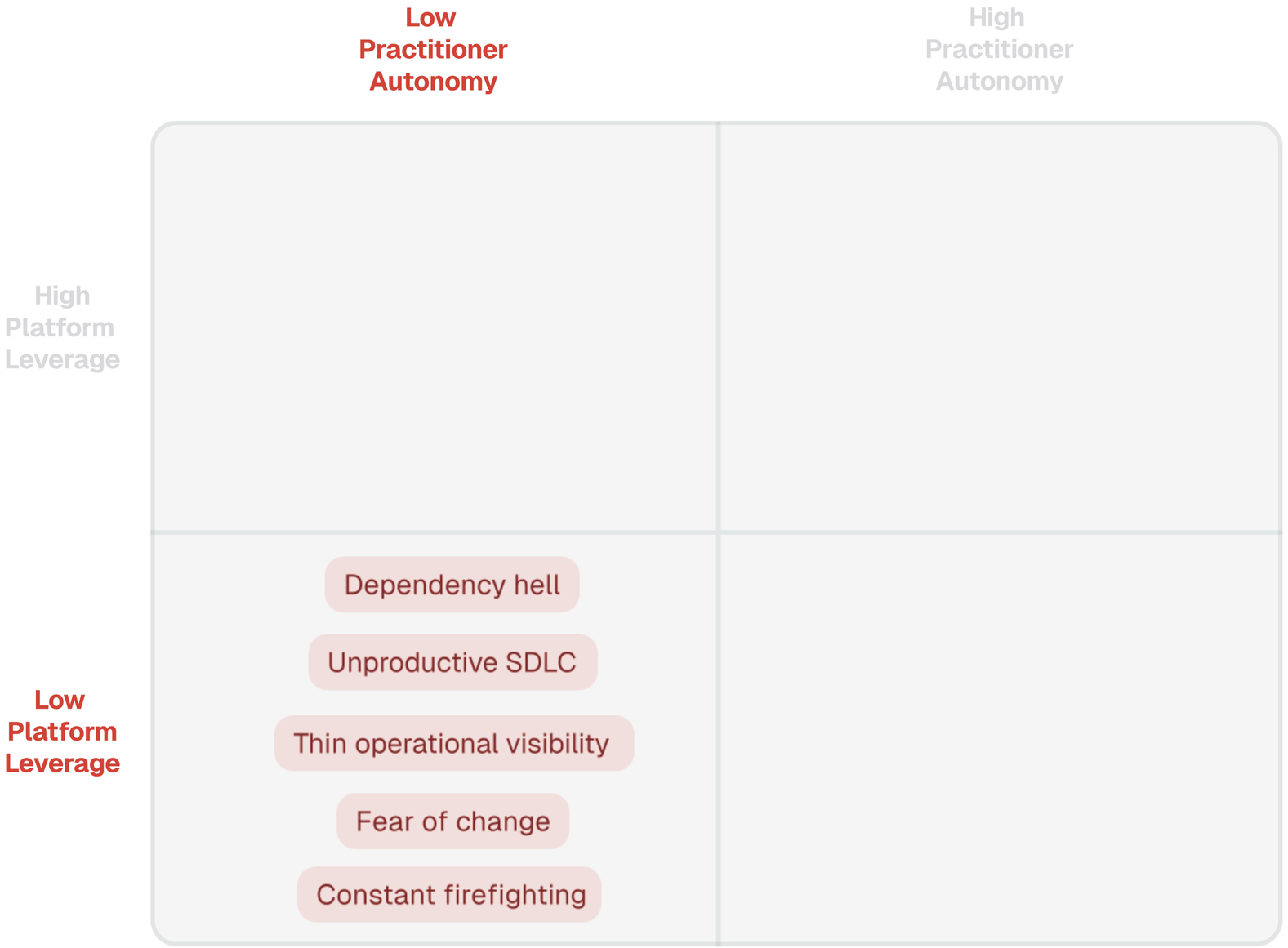
Leverage

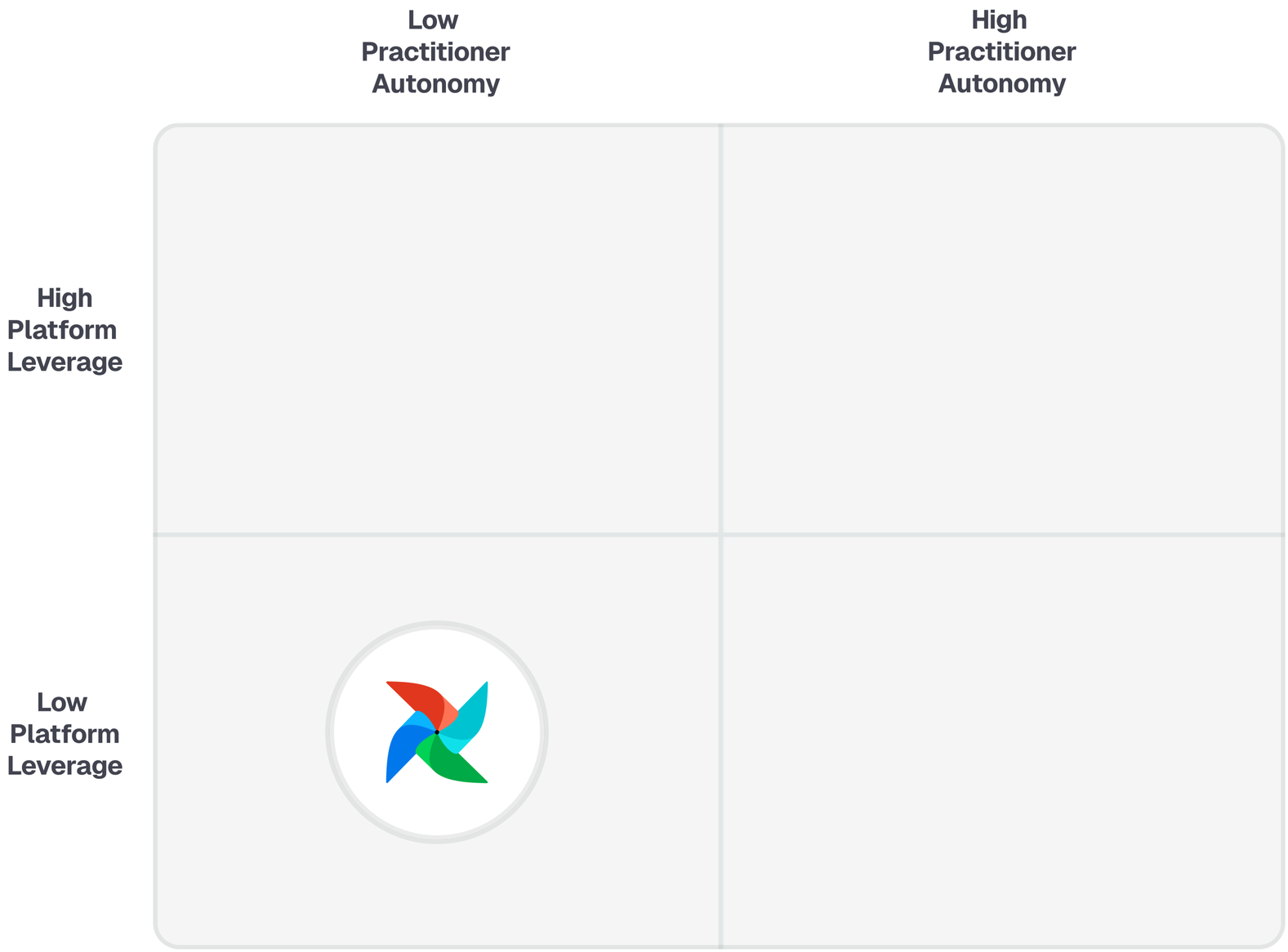
- Ensure practitioners are self-sufficient and productive
- Focus on high-impact, platform-wide improvements
- Partner with cross-cutting policy teams (finance, compliance)

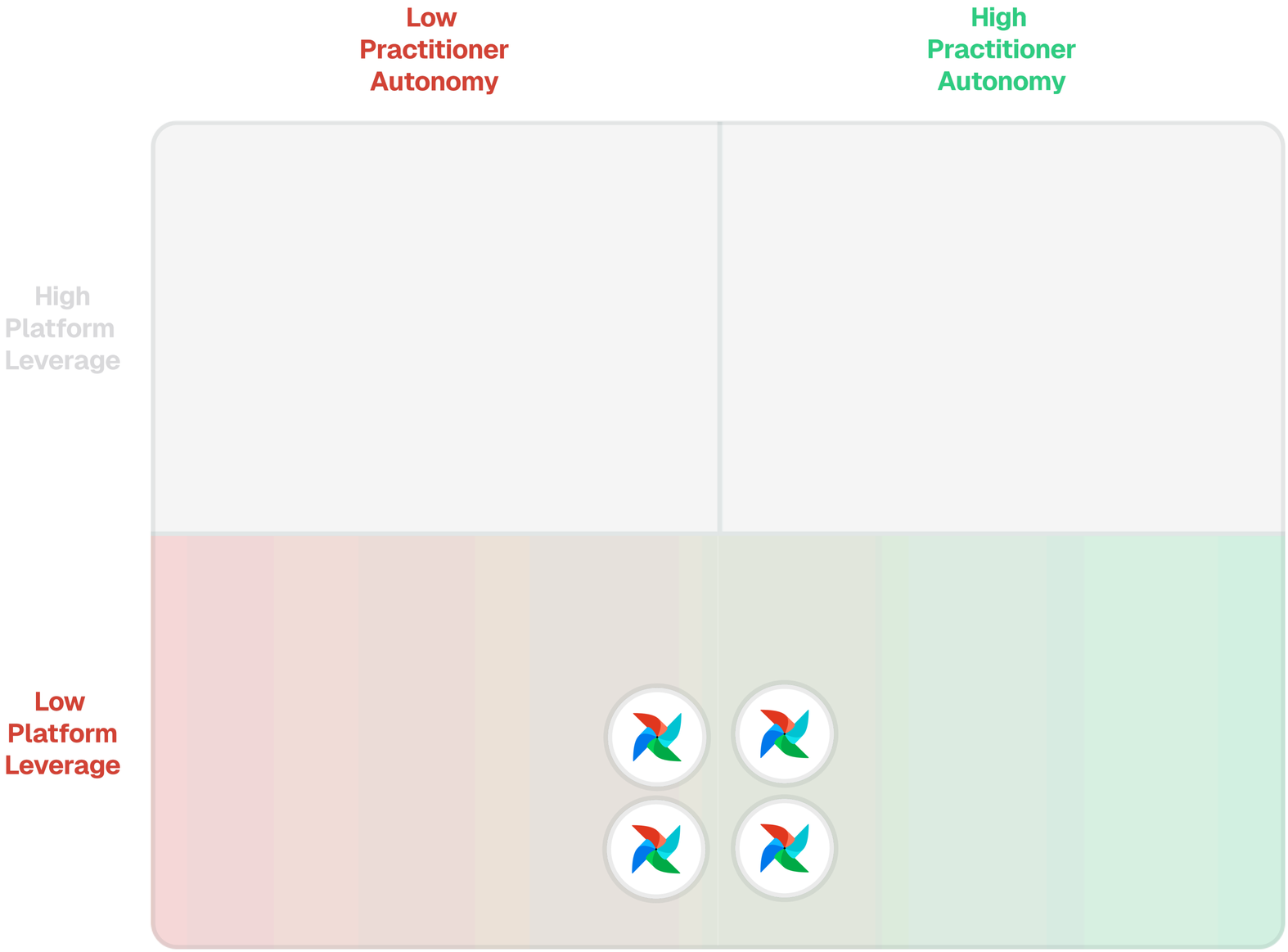
Low
Practitioner
Autonomy
High
Practitioner
Autonomy
High
Platform
Leverage
Low
Platform
Leverage


Low
Practitioner
Autonomy
High
Practitioner
Autonomy
High
Platform
Leverage
Low
Platform
Leverage


Low
Practitioner
Autonomy
High
Practitioner
Autonomy
High
Platform
Leverage
Low
Platform
Leverage


Low Practitioner Autonomy
High Practitioner Autonomy
High Platform Leverage
Low Platform Leverage
Dependency hell
Unproductive SDLC
Thin operational visibility
Fear of change
Constant firefighting


Low
Practitioner
Autonomy
High
Practitioner
Autonomy
High
Platform
Leverage
Low
Platform
Leverage

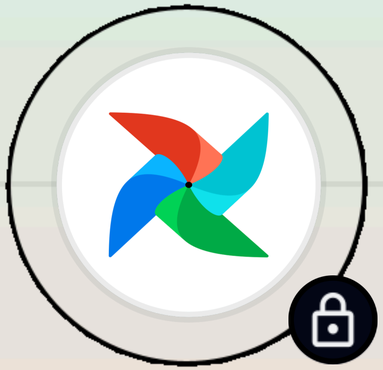

Low
Practitioner
Autonomy
High
Practitioner
Autonomy
High
Platform
Leverage
Low
Platform
Leverage

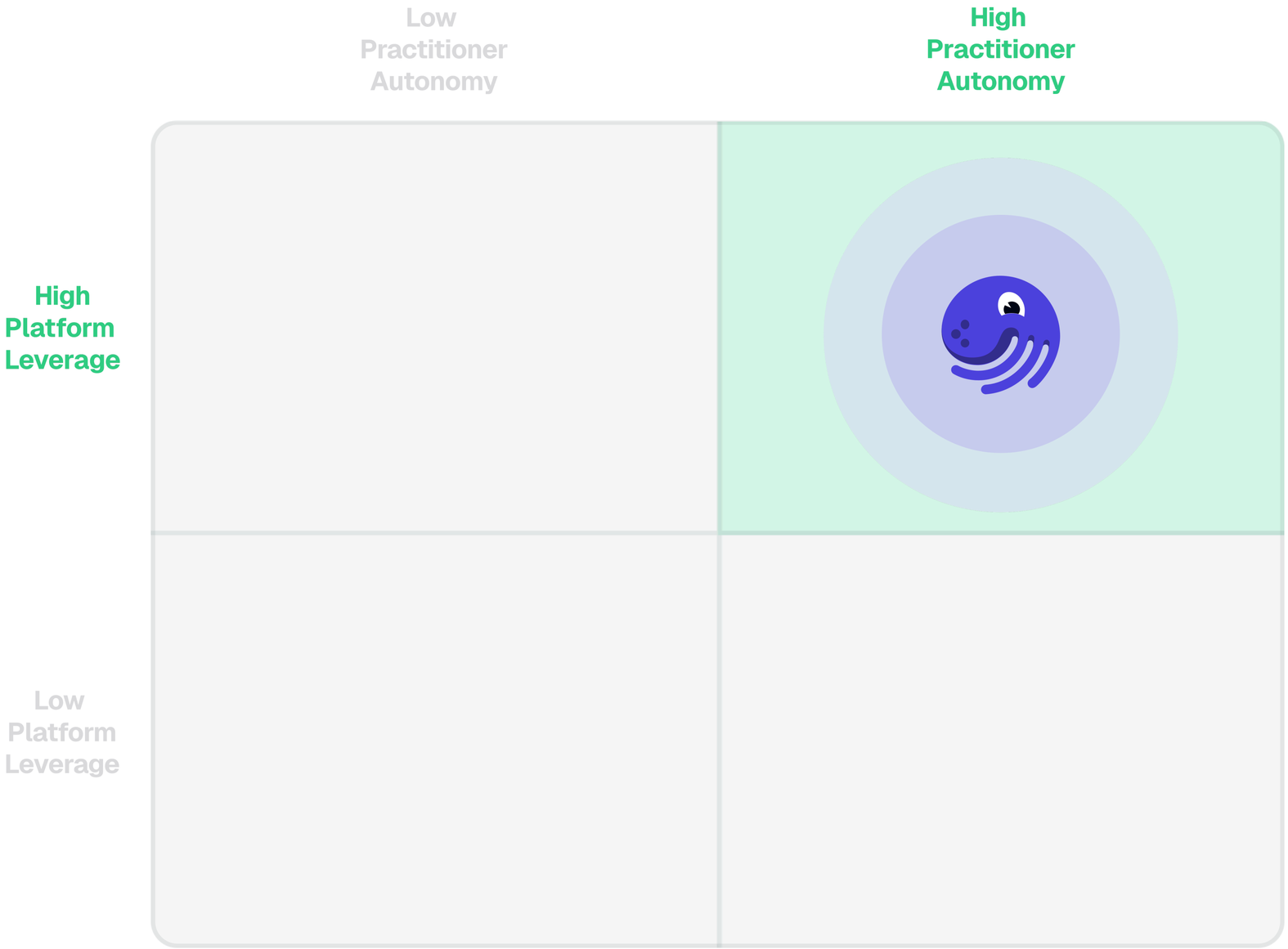
Low Practitioner Autonomy

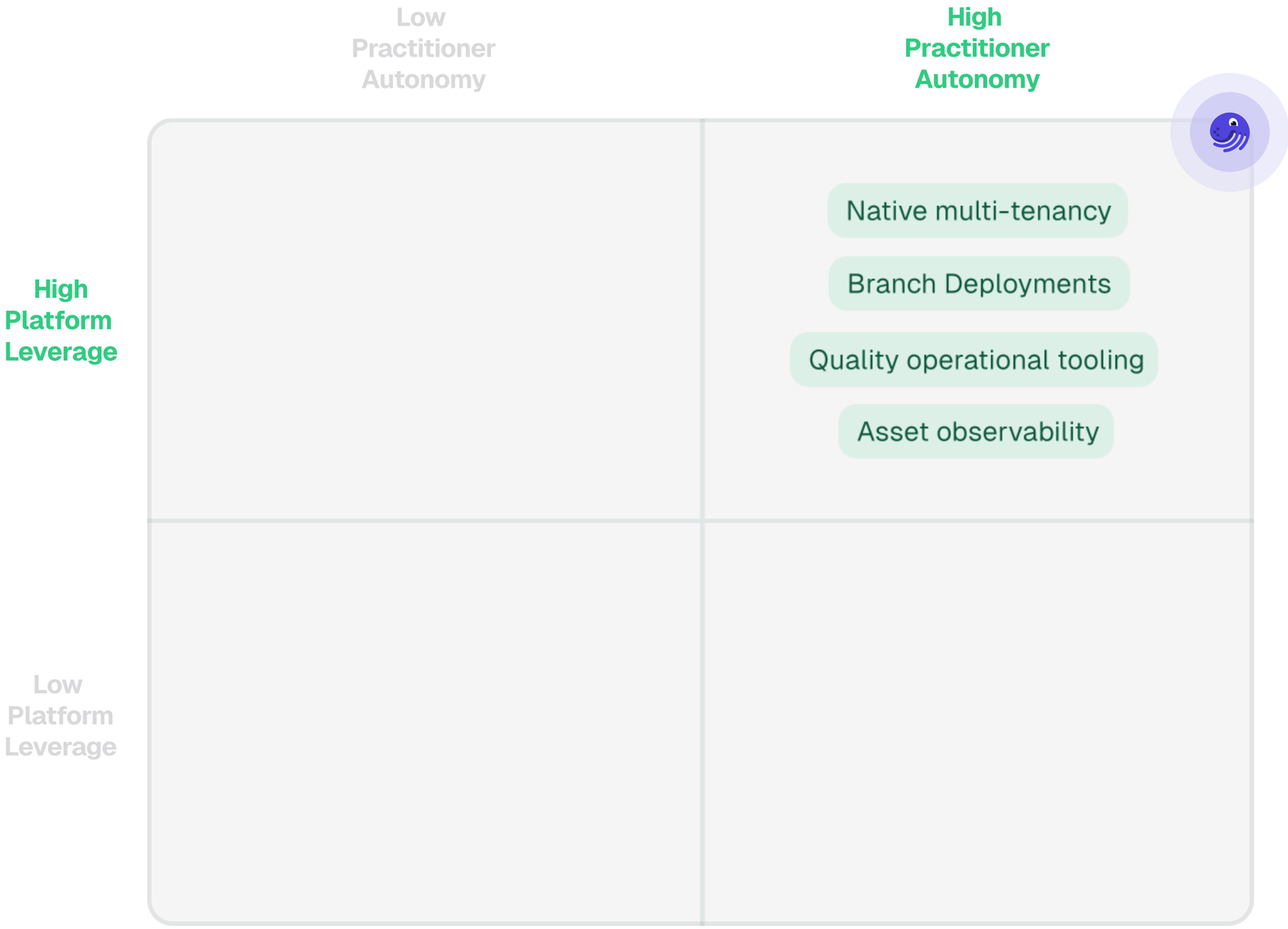
High Practitioner Autonomy

High Platform Leverage

Low Platform Leverage

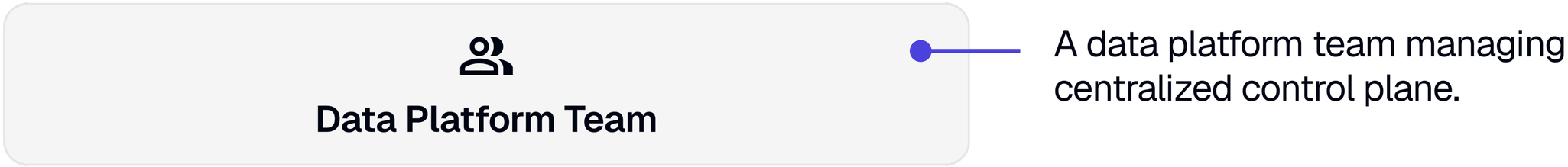



Low
Practitioner
Autonomy
High
Practitioner
Autonomy
High
Platform
Leverage
Low
Platform
Leverage

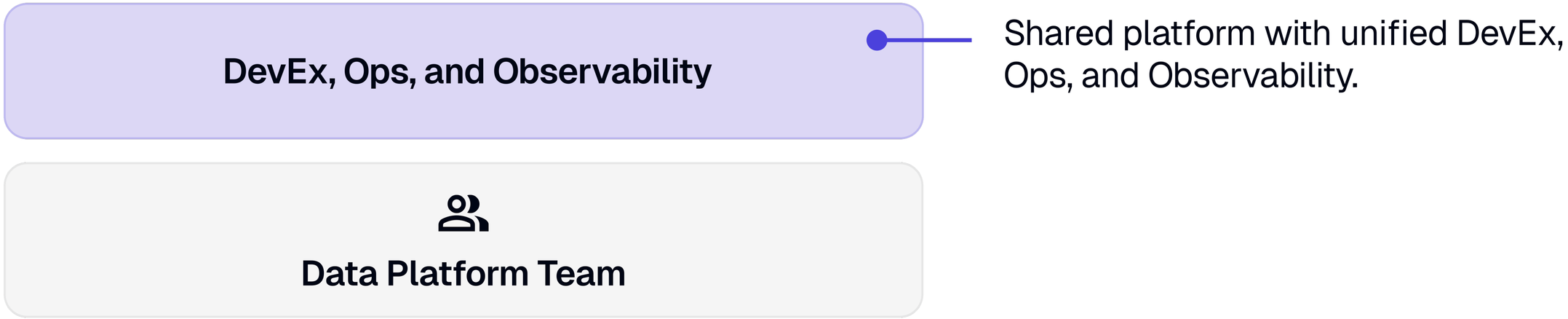

Low Practitioner Autonomy
High Practitioner Autonomy
High Platform Leverage
Low Platform Leverage
Native multi-tenancy
Branch Deployments
Quality operational tooling
Asset observability

Organizing the platform

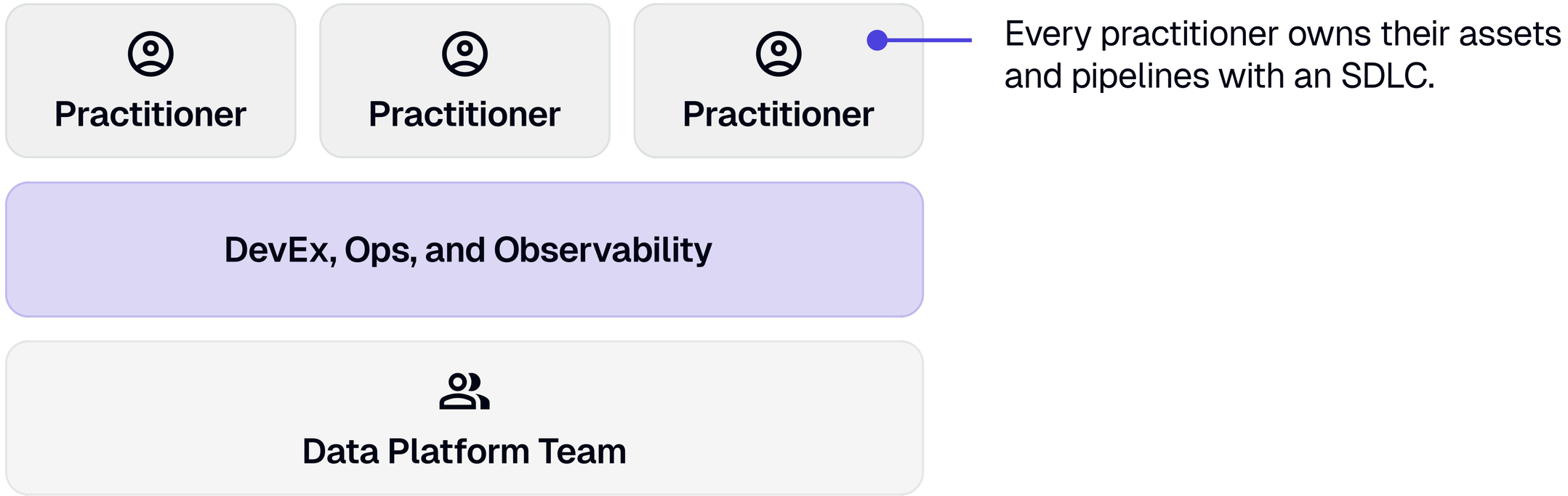
Organizing the platform



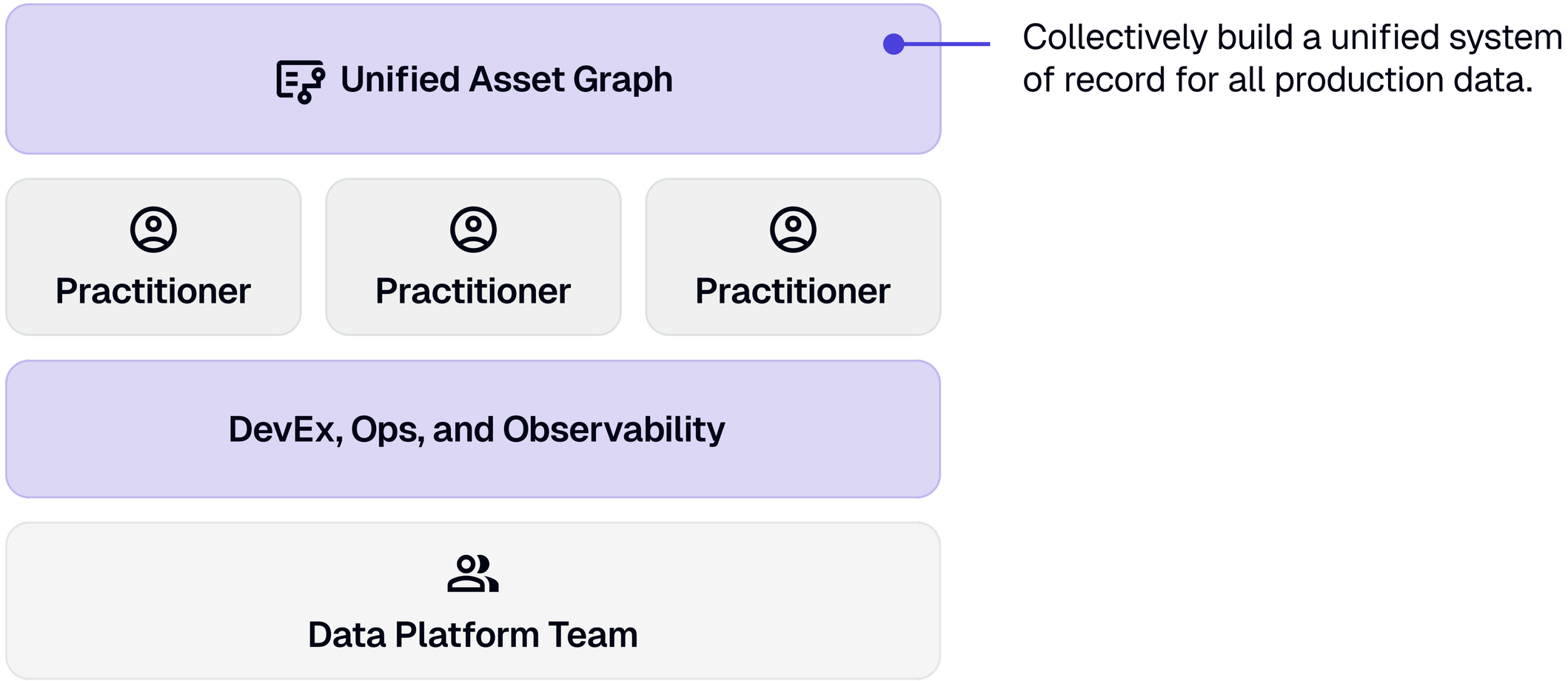
Organizing the platform

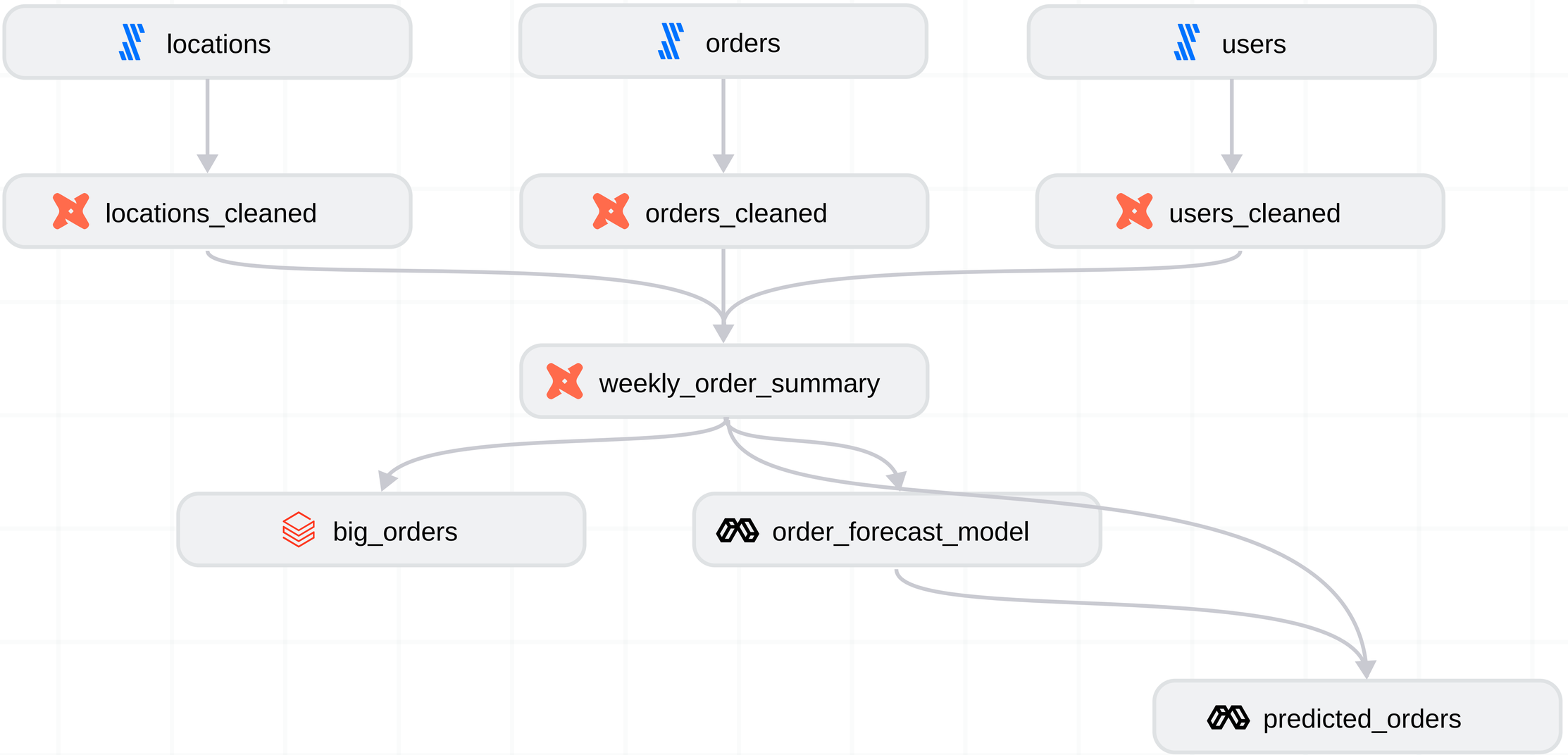


Organizing the platform



Organizing the platform




locations
orders
users
locations_cleaned
orders_cleaned
users_cleaned
weekly_order_summary
big_orders
order_forecast_model
predicted_orders

dagster

dagster

It's your Data Platform

✓ Foundational
✓ Costly
✓ Context-specific

It's the Control Plane

✓ Makes decisions
✓ Flexible & Programmable
✓ Centralized

It's a Unified System

✓ DX Workflow
✓ Operational tools
✓ Asset graph

dagster

Thank You