



How Airflow solves the coordination of decentralised teams at Vinted

Oscar Ligthart
Rodrigo Loredó

3.0

Vinted

Make second-hand first
choice worldwide



Millions

of registered members

20+

countries

2000+

talents

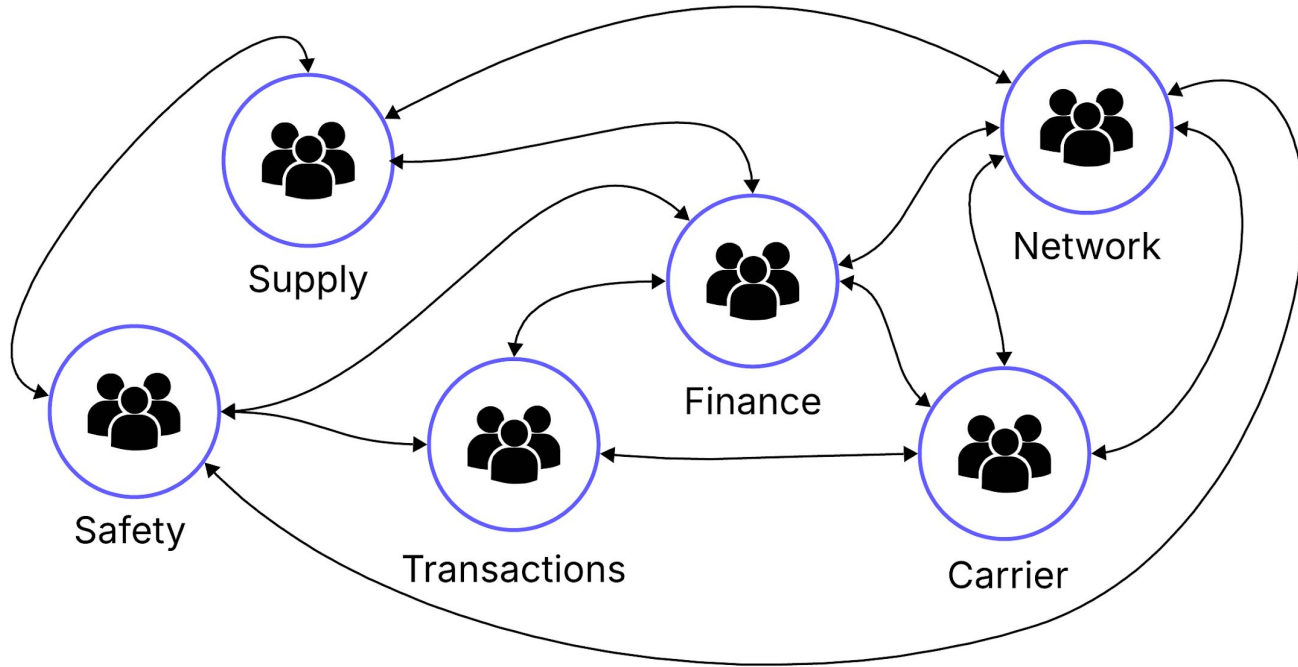
10

workplaces



2000+

talents





Supply



Finance



Network



Safety



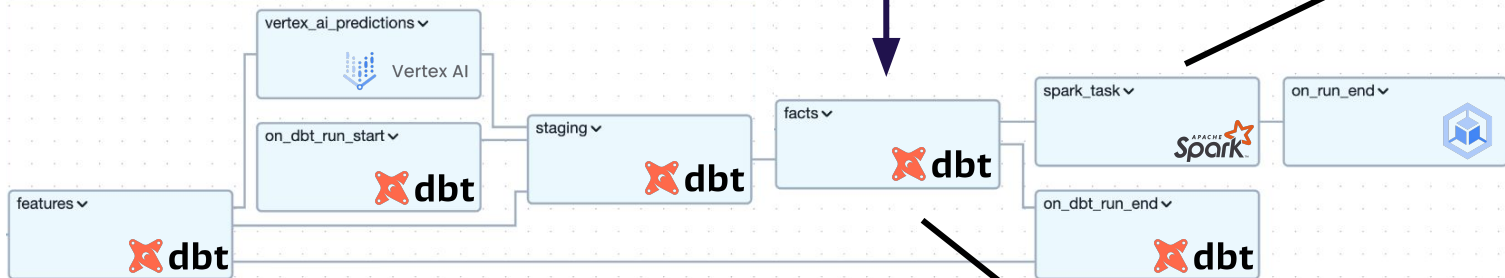
Transactions

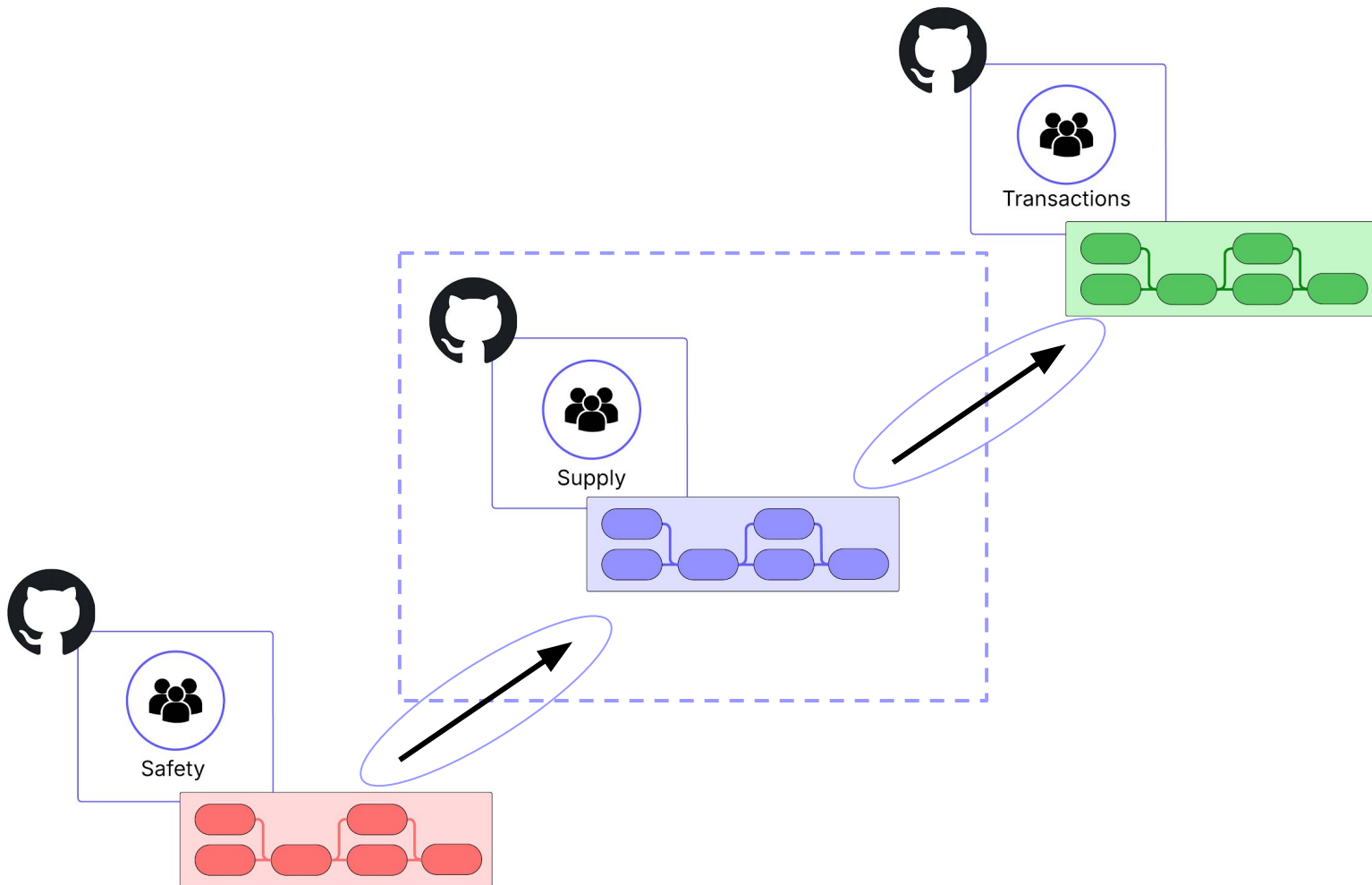


Carrier

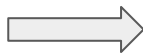
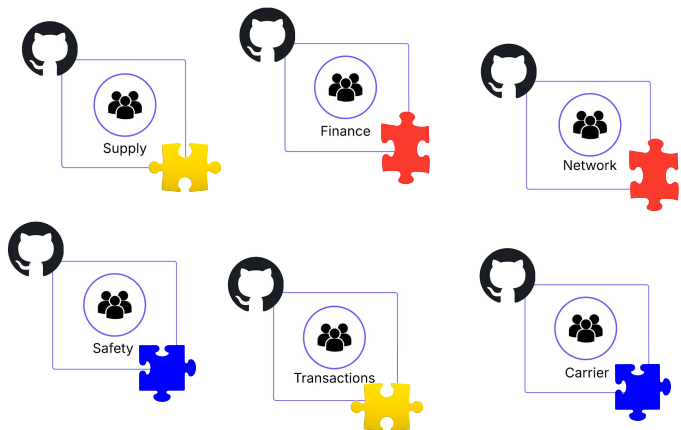


GitHub

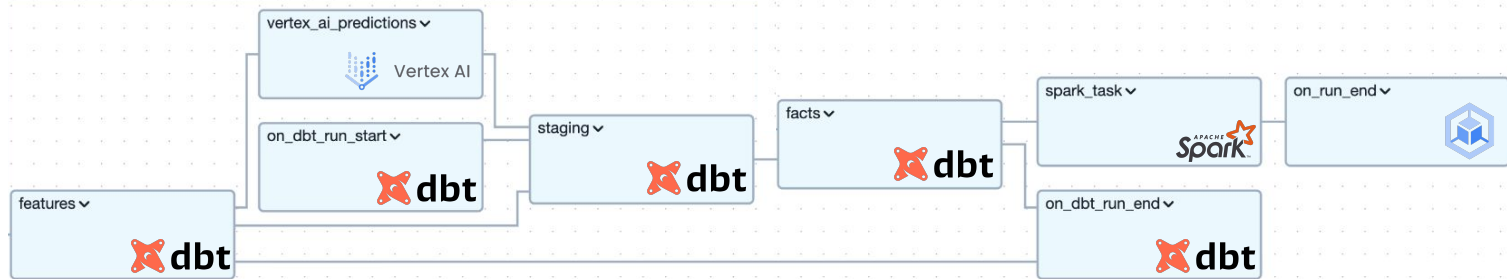


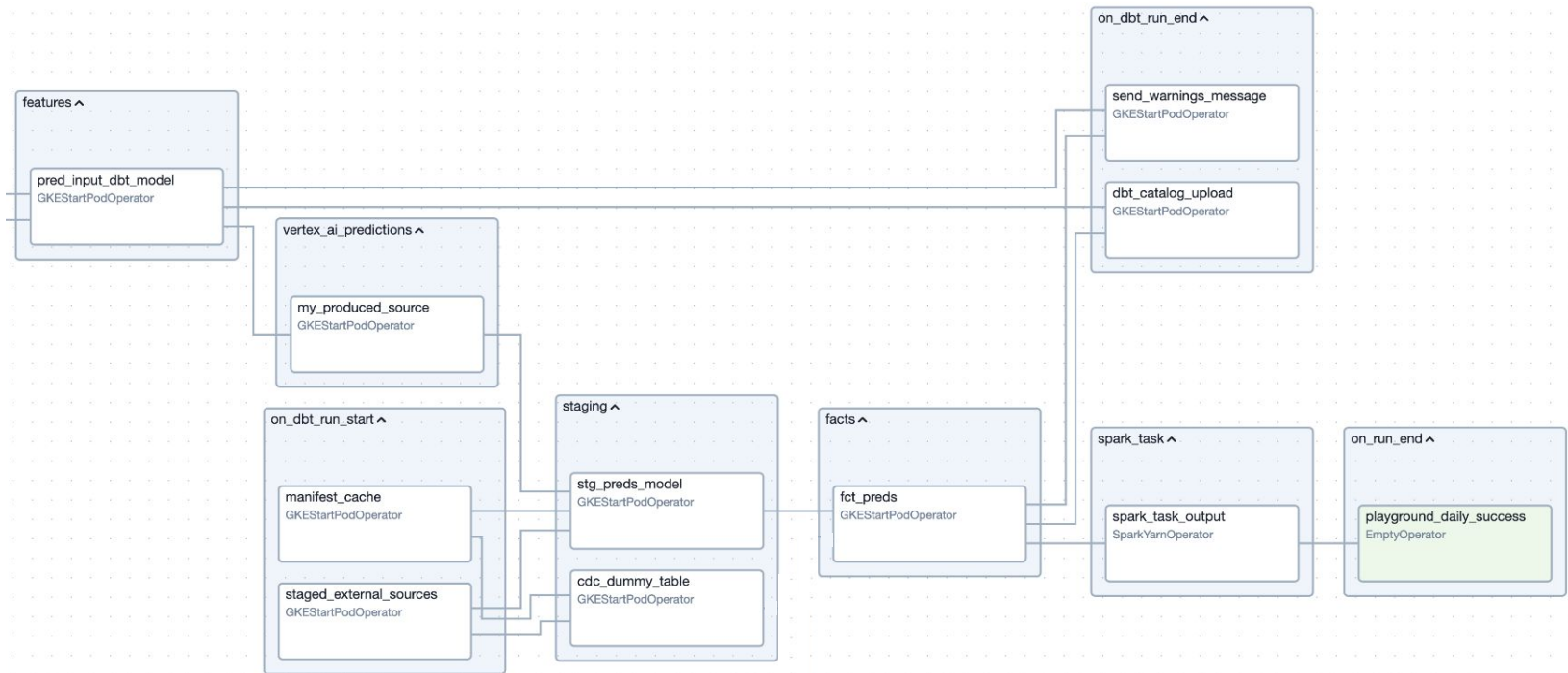






Generating DAGs



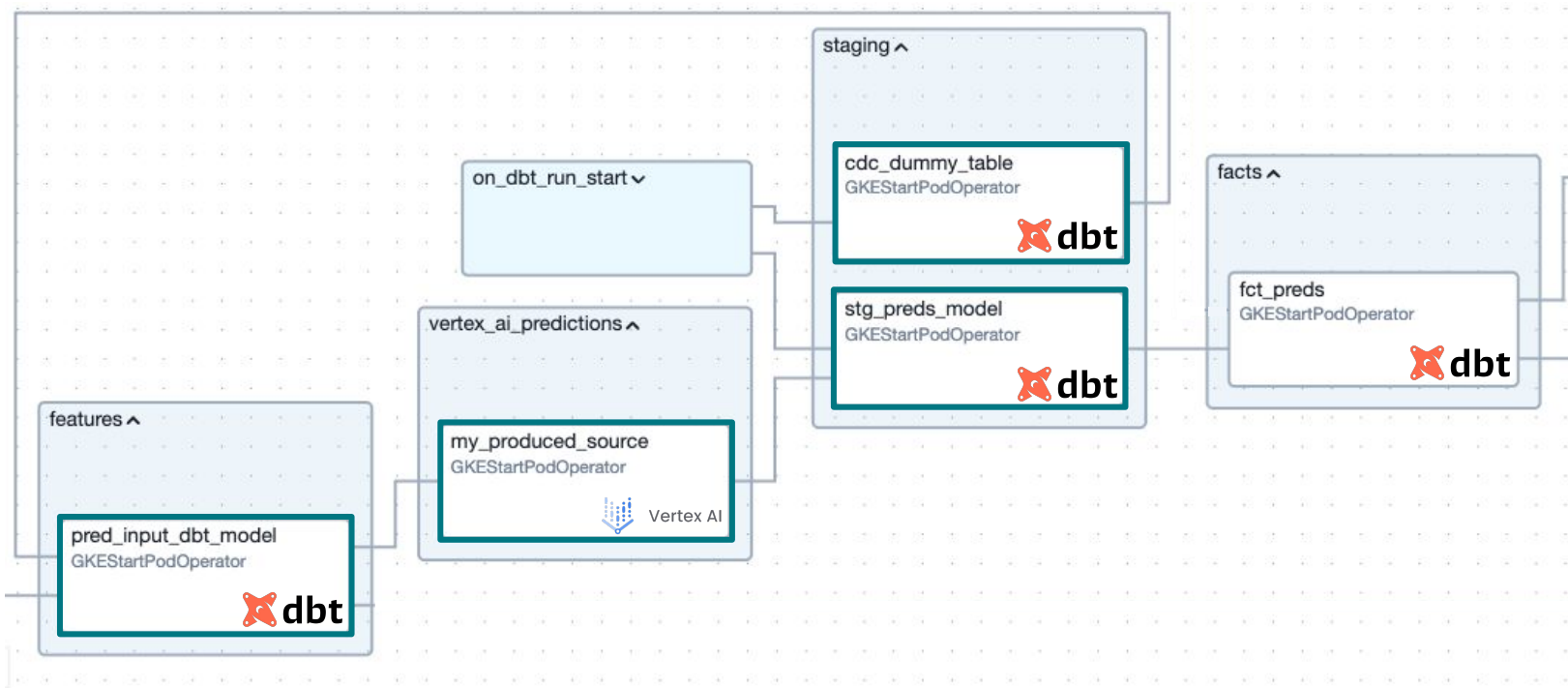


DAG generator design: User interface

```
1 version: 0.17
2 tags:
3   - tag_1
4   - long_tag
5
6 alerting:
7   channel: "#some-alerts-channel"
8   tag:
9     - "@amazing.person" # a person
10    - "!subteam^S05VAFP4A6N" # a Team
```

```
12 dbt_settings:
13   enabled: true
14   on_run_start:
15     - stage_external_sources
16   on_run_end:
17     - send_warnings_message
```

```
19 tasks:
20   - kind: universal
21     group: some_task_group_name
22     conditions:
23       - some_source
24     assets:
25       - produced_output_table
26     timeout:
27       minutes: 30
28     image: image_name:tag
29     command: poetry run python -m my_module.main
30
31 exposed_assets:
32   - regex: "mrt_daily_exposed"
33     slo:
34       daily:
35         hours: 4
```



Connecting Tasks

dbt_models.yml

```
1 models:
2   - name: input_dbt_model
3     config:
4       meta:
5         scheduler:
6           pipelines:
7             - hourly
```

scheduler.yaml

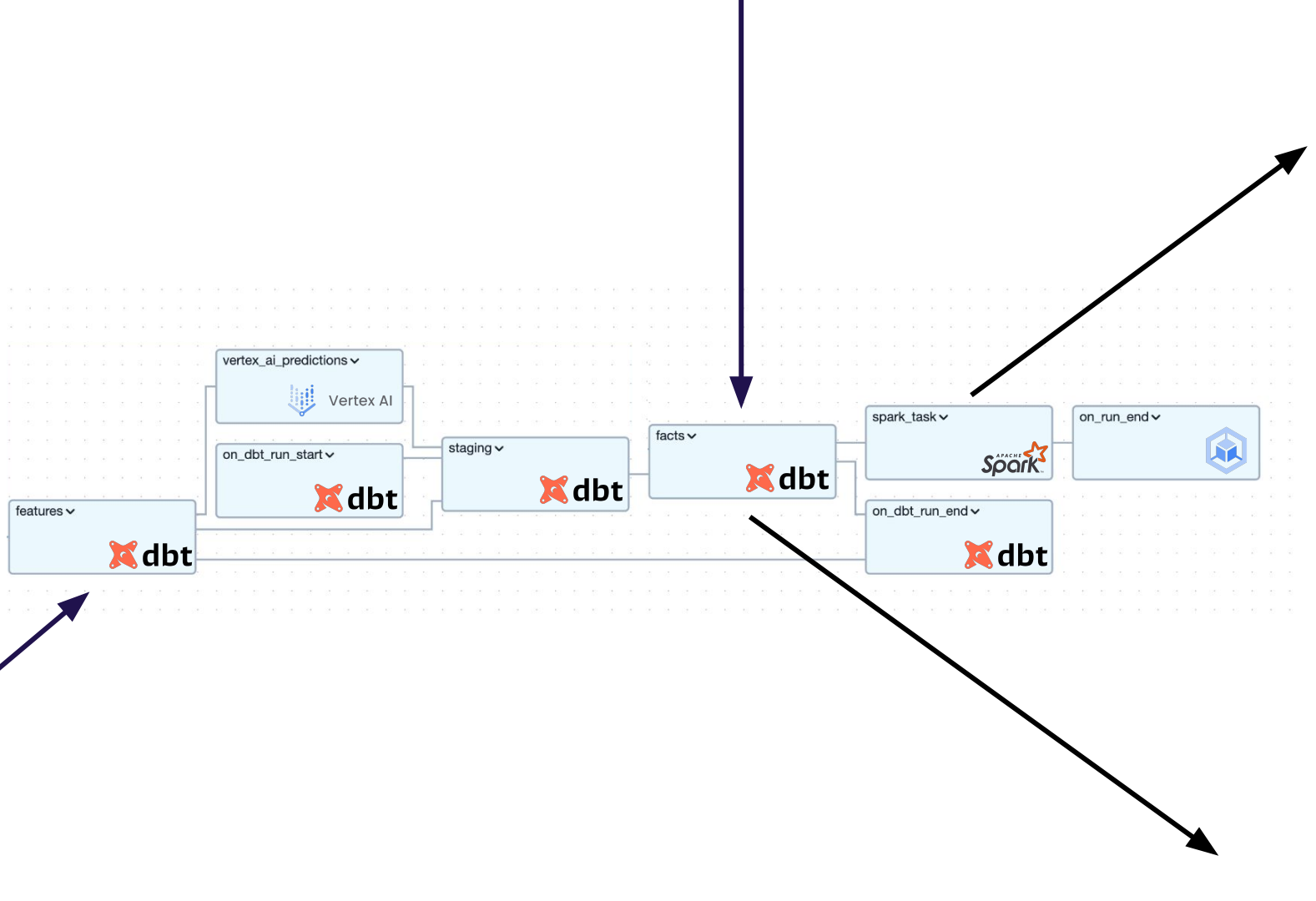
```
1 tasks:
2   - kind: vertex-ai
3     group: predictions
4     conditions:
5       - input_dbt_model
6     assets:
7       - my_produced_source
```

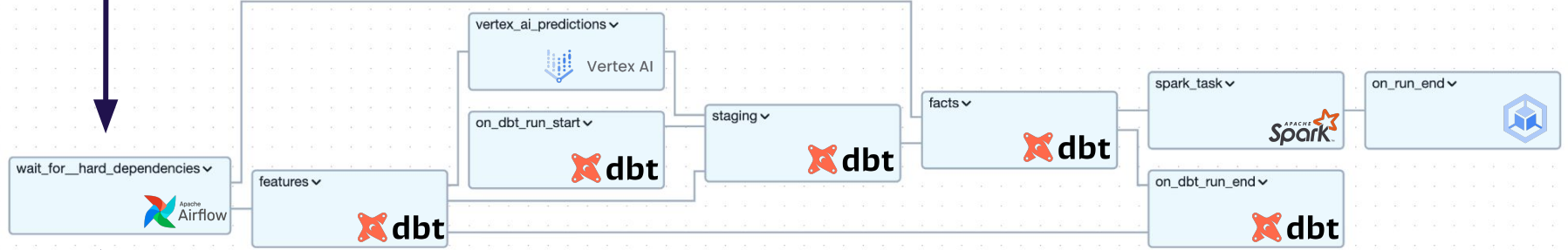
predictions_sources.yml

```
1 sources:
2   - name: predictions
3     tables:
4       - name: my_produced_source
5         columns:
6           ...
```

some_model.sql

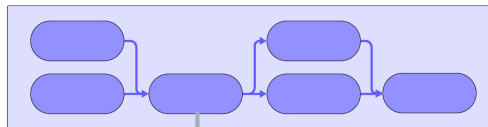
```
1 ...
2 select customer_id, predicted_value
3 from {{ source("predictions", "my_produced_source") }}
4 ...
```





Using other team's assets

Other team's DAG



wait_for__hard_dependencies ^

dim_users

ExternalTaskSensor

facts v

Using other team's assets

scheduler.yaml

```
1 exposed_assets:
2   # All dimensions
3   - regex: "dim_.*"
4     slo:
5       daily:
6         hours: 6
```

Asset registry object

```
1 {
2   "some_team": {
3     "produced": {
4       "dim_customers": {
5         "dag_prefix": "some_team",
6         "task_group": "dimensions",
7         "task_id": "dim_customers",
8         "pipeline": "daily",
9         "slo_seconds": 21600
10      }
11    }
12  }
```

CI/CD validation

```
1 sources:
2   - name: global_assets
3   tables:
4     - name: dim_customers
```

```
1 ...
2 select customer_id, country_id
3 from {{ source("predictions", "dim_customers") }}
4 ...
```

Developer Flow

▼ playground/models/global_assets/global_sources.yml 

+4 

☐ Viewed



... @@ -0,0 +1,4 @@

1 + sources:

2 + - name: shipping

3 + tables:

4 + - name: fct_shipments

▼ playground/models/marts/my_new_model.sql 

+3 

☐ Viewed



... @@ -0,0 +1,3 @@

1 + select *

2 + from {{ ref("fct_preds") }}

3 + left join {{ source("shipping", "fct_shipments") }} as s on fct_preds.shipment_id = s.shipment_id



Scheduler report: playground

✓ No breaking changes found in project registry.

Scheduler state diff

Dags: 0 added, 0 removed

Tasks: 1 added, 0 removed

Global dependencies: 1 added, 0 removed

Tasks and global dependencies

playground_daily

+ global_dependency.fct_shipments

+ marts.my_new_model



Scheduler report: playground

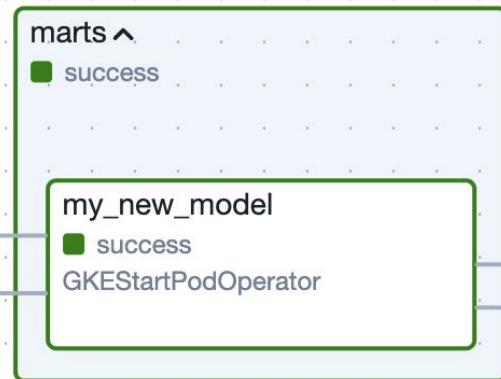
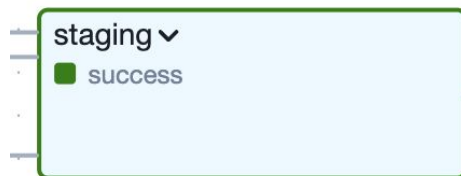
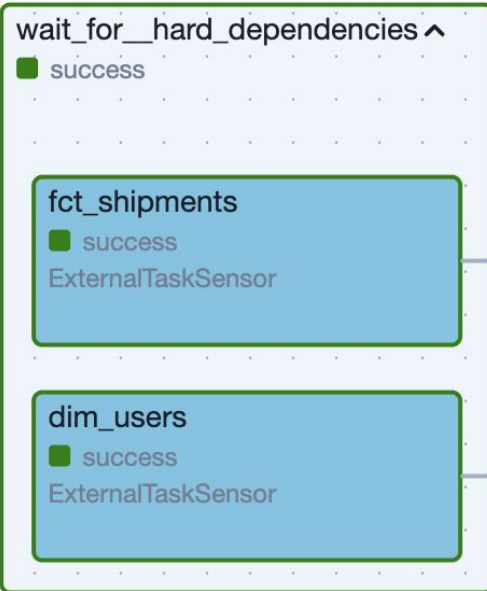
 **Breaking changes found in project registry**

Asset `test_users`

- Pipeline changed from hourly to daily.
- SLO increased from 2700 to 162000.

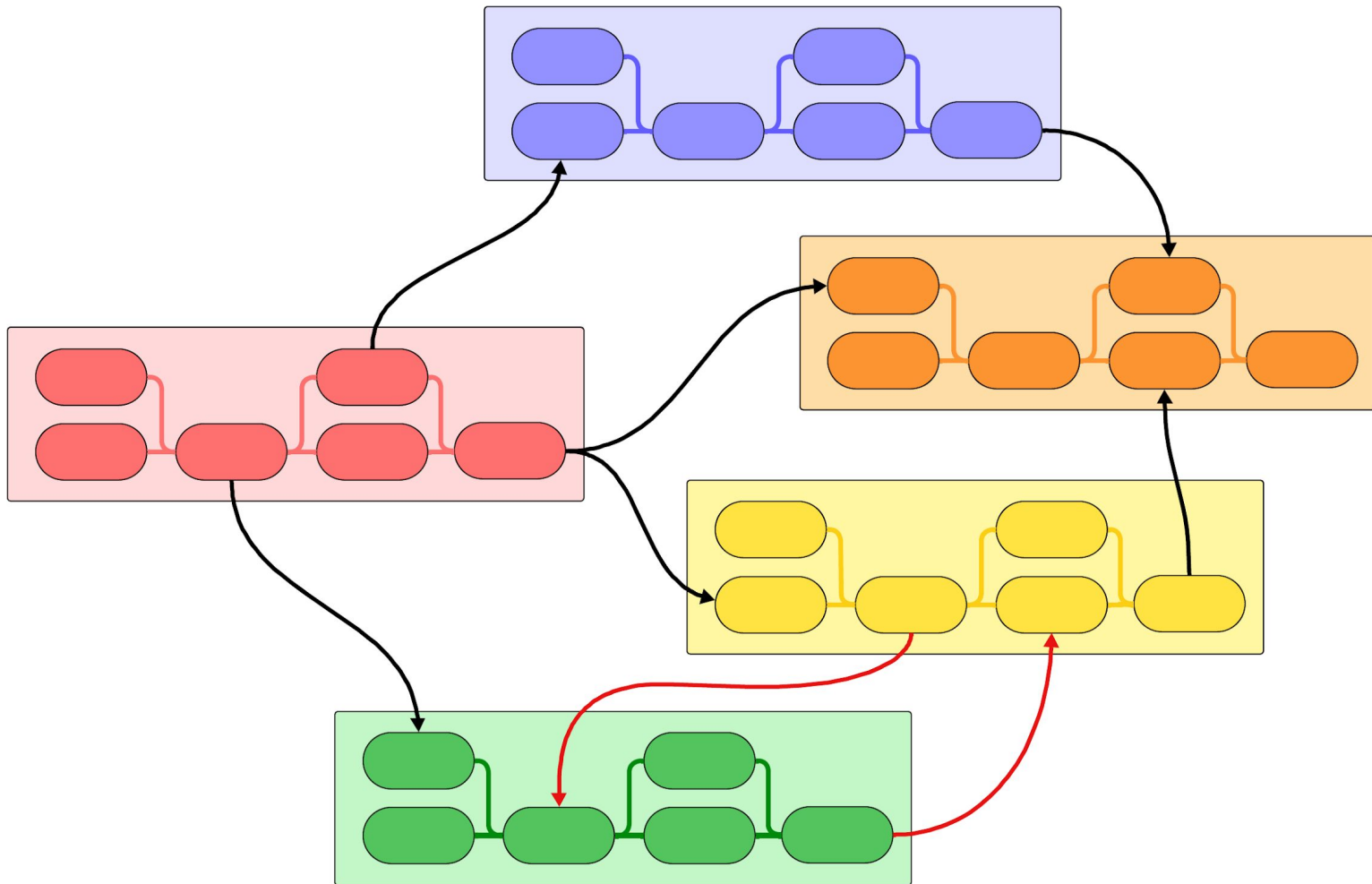
Consumers:

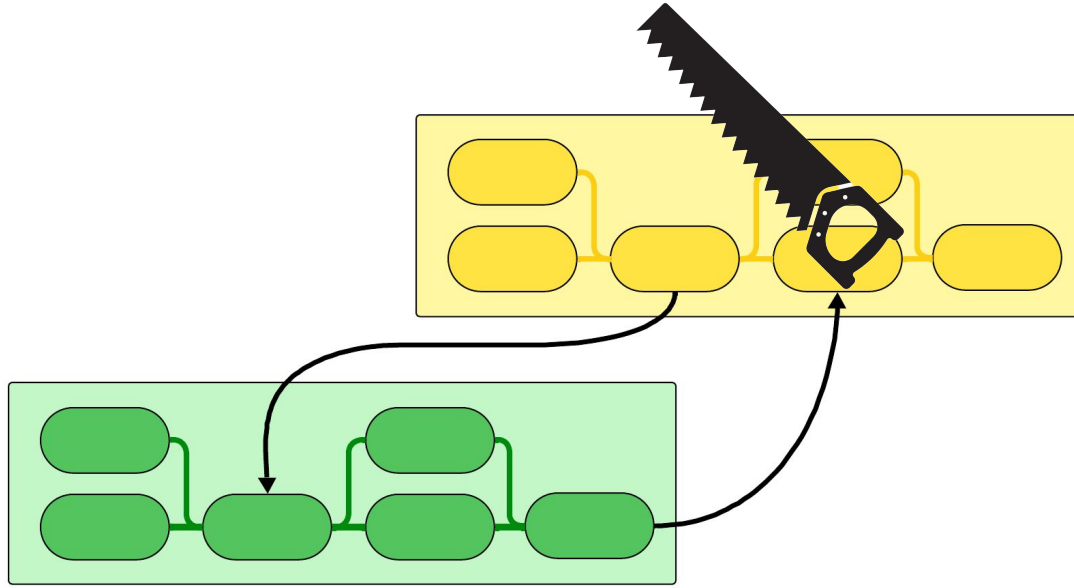
- `some_team:#some-team-contact-channel`

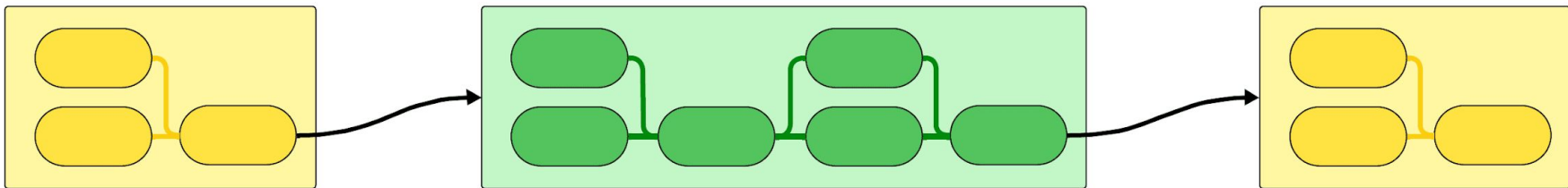


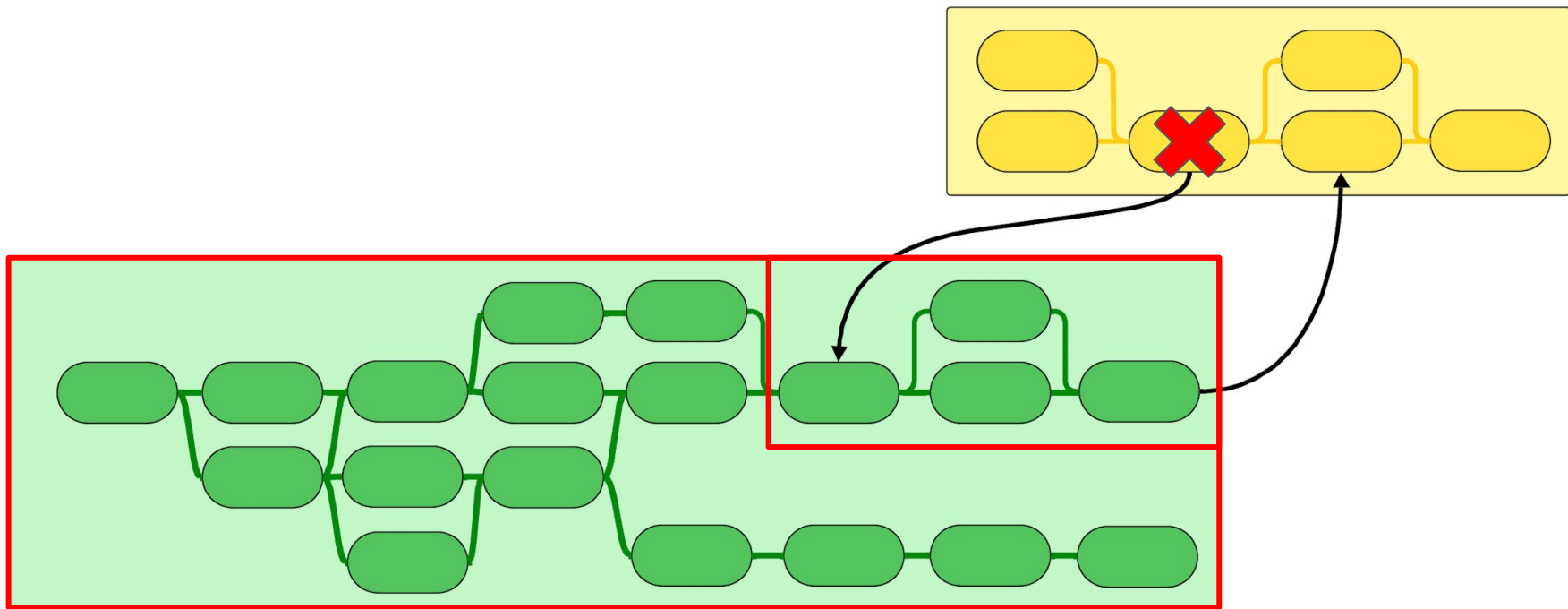
Sensor Setup













Supply



supply_daily task failure 🐛

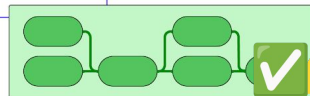
Date interval start: 2025-09-09 00:10:00+00:00

[marts.mrt_my_model - view logs](#)

@some.owner.team, @some.owner.person



Transactions



transactions_daily dependency failure 🐛

Date interval start: 2025-09-09 00:10:00+00:00

[wait_for__hard_dependencies.mrt_my_model - view logs](#)

When the task finishes successfully on the producers' side, your DAG will resume automatically.

@some.owner.team, @some.owner.person



Safety



safety_daily dependency failure 🐛

Date interval start: 2025-09-09 00:10:00+00:00

[wait_for__hard_dependencies.mrt_my_model - view logs](#)

When the task finishes successfully on the producers' side, your DAG will resume automatically.

@some.owner.team, @some.owner.person



The New Normal



HOTEL

GENERAL
STORE



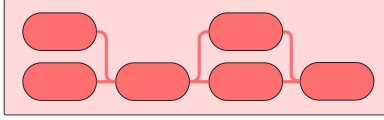
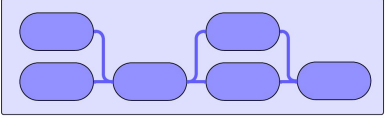
Supply

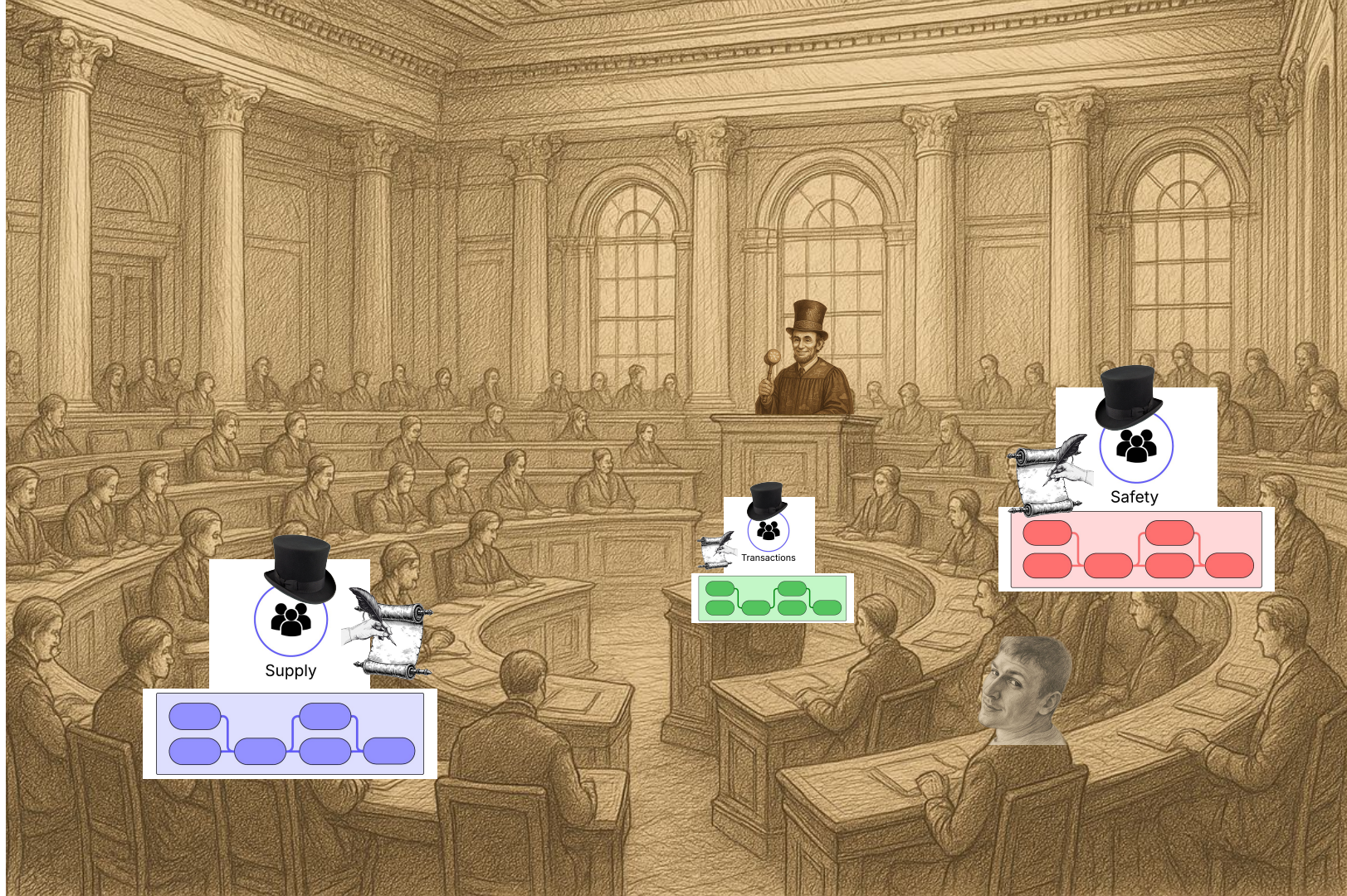


Transactions

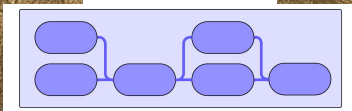


Safety

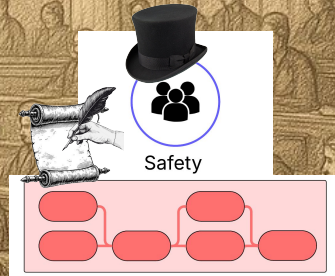




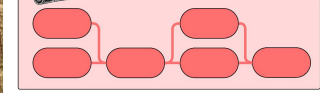
Supply



Transactions



Safety





Thank you!
Questions?