

Ensuring Data Integrity for Campus Sustainability



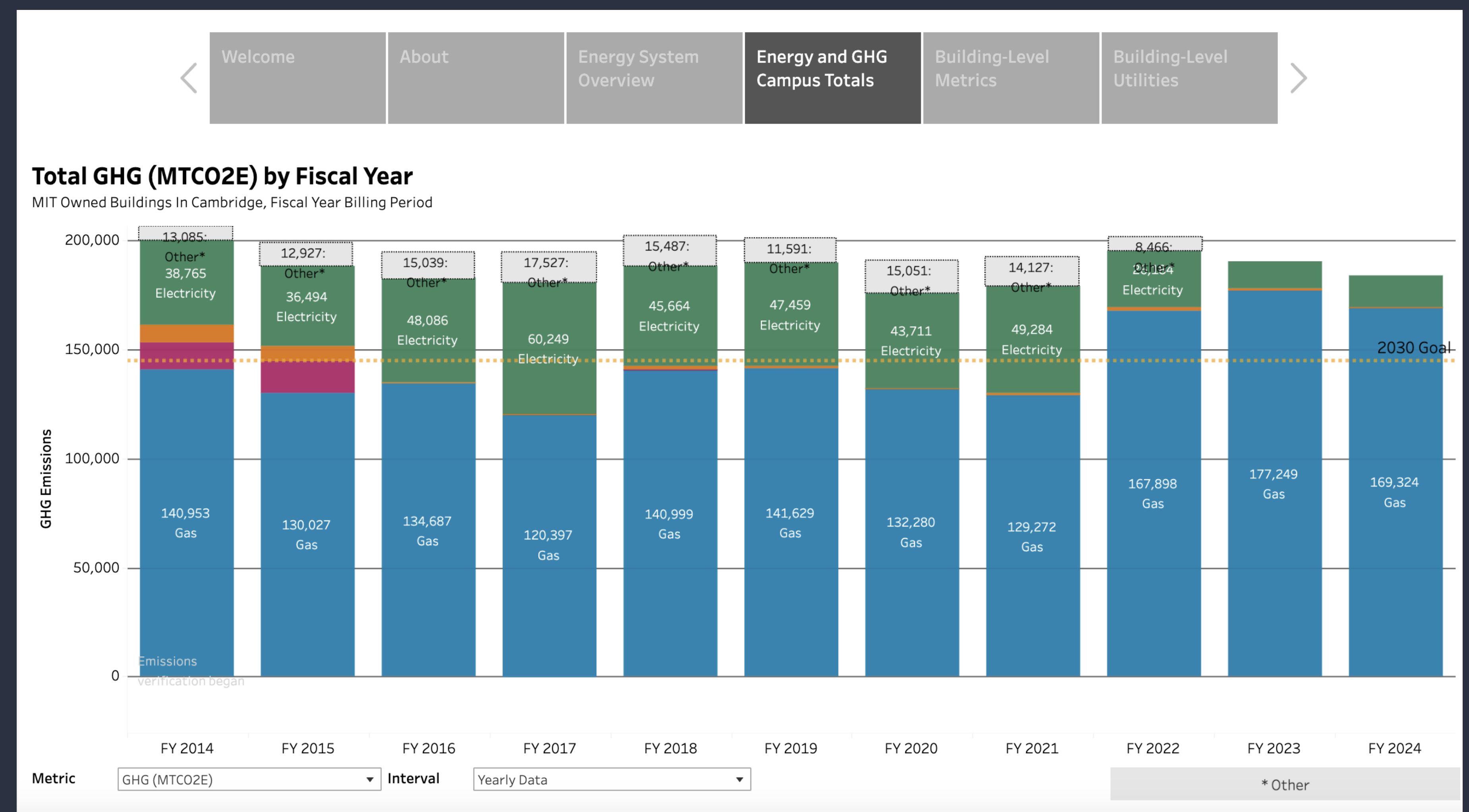
Yu Cheng

MITOS Data Mission

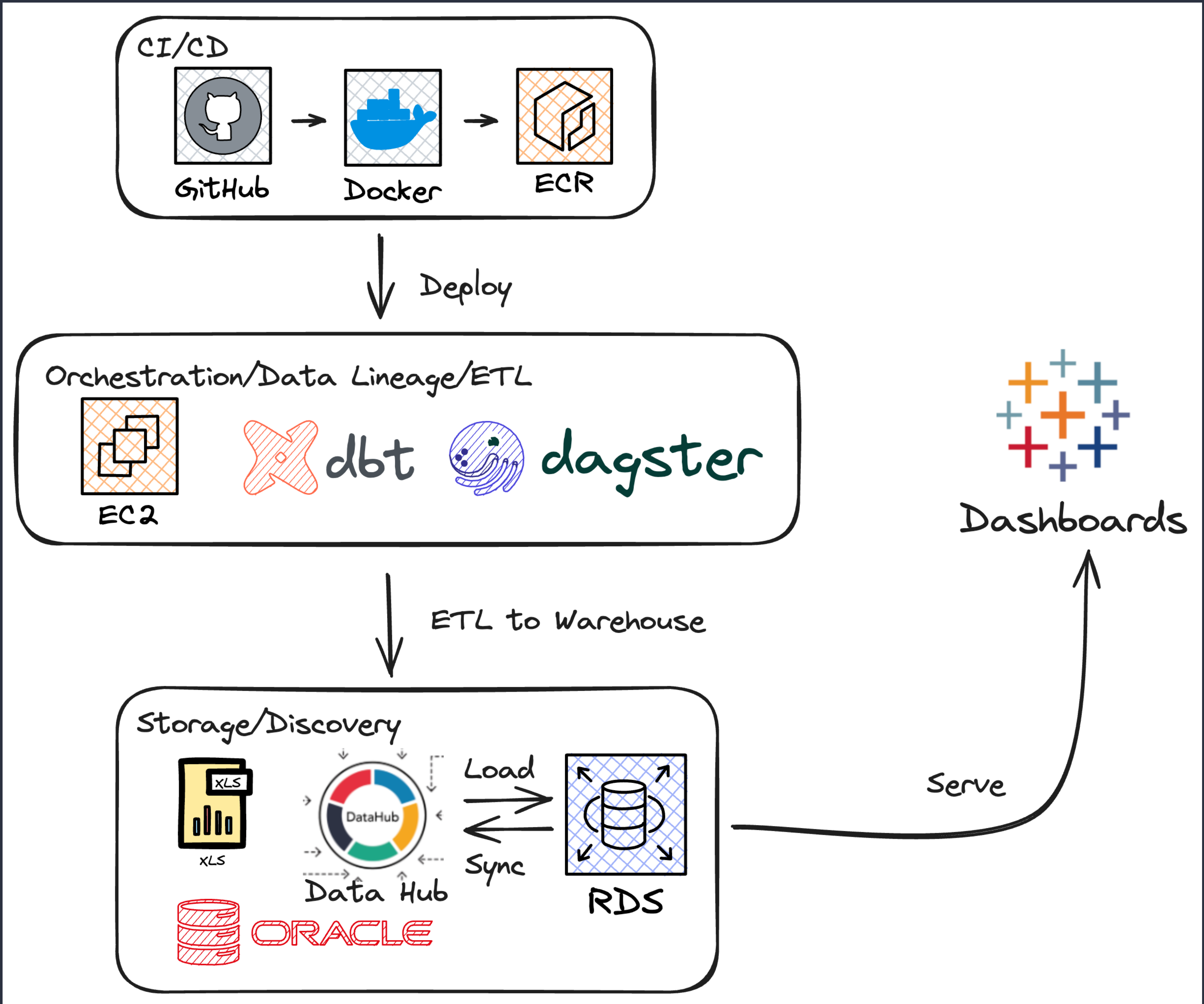


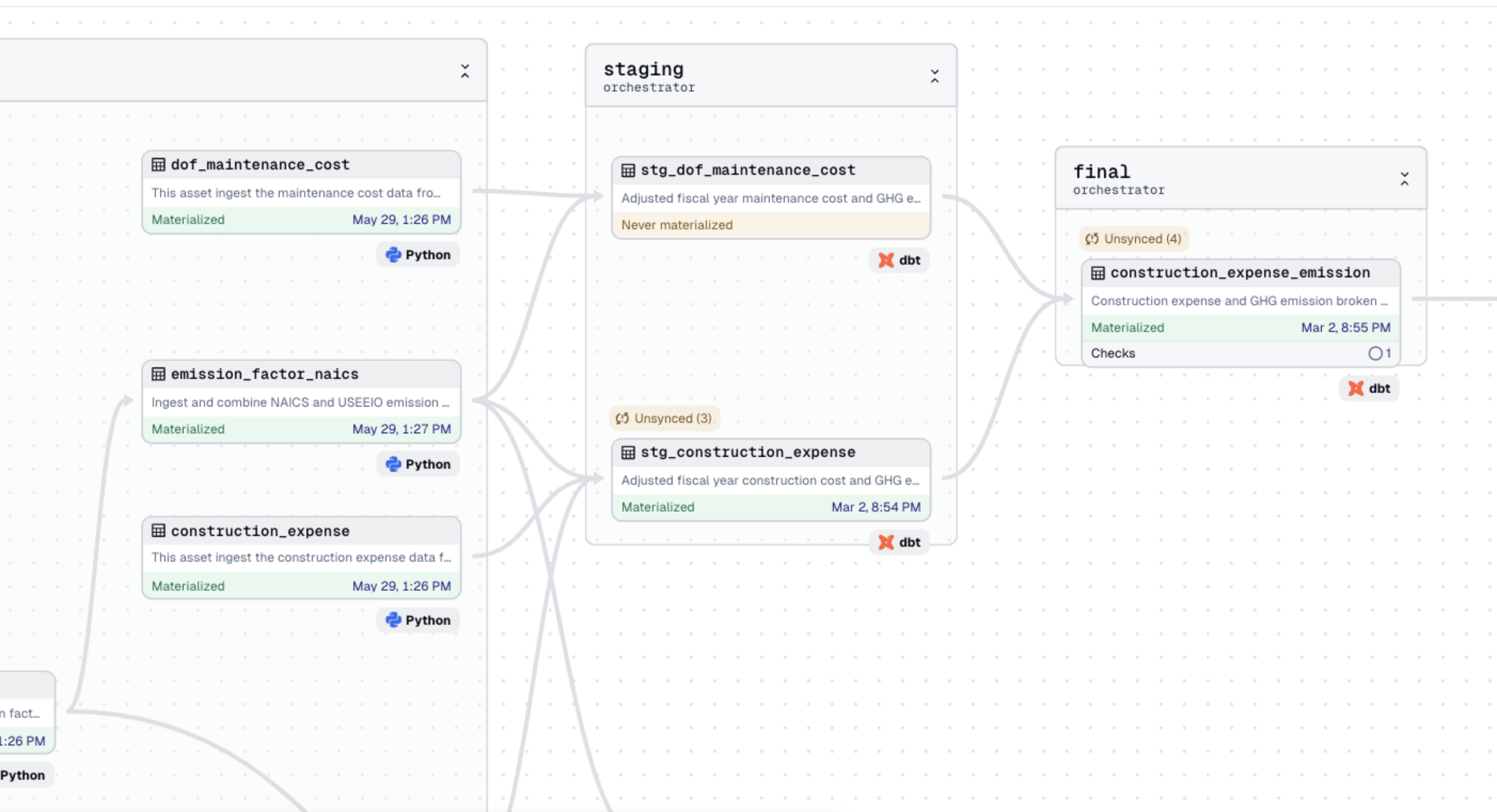
Sustainability Datapool

Since 2018, datapool reports campus sustainability data, including Energy, GHG emissions, flooding risk, waste recycling, and more.



Basin: the data platform





View global asset lineage ↻ Reload

0:02 ↻ Materialize selected

Status	Timestamp
● Materialized	Feb 1, 11:56 AM
● Materialized	Jan 27, 7:30 AM
● Materialized	Apr 16, 8:34 PM
● Materialized	Apr 16, 8:36 PM
● Materialized	Apr 16, 8:34 PM
● Materialized	Apr 16, 8:35 PM
● Materialized	May 3, 11:02 AM
● Materialized	Feb 21, 12:04 PM

```

-- set static variables using jinja2 syntax
{% set reply_rate = 0.33 %}
{% set work_week = 50 %}
{% set remote_rate = 0.21 %}

WITH distance AS (
  SELECT
    "drove alone" * {{ var('car_speed') }}
    "carpooled(2-6)"
    * {{ var('car_speed') }}
    * {{ var('car_share_ratio') }}
    * commute_time_average_hours AS c
    "vanpooled(7+)"
    * {{ var('car_speed') }}
    * {{ var('van_share_ratio') }}
    * commute_time_average_hours AS v
    shuttle * {{ var('bus_speed') }}
    "public transportation"
    * {{ var('t_ratio') }}
    * {{ var('t_speed') }}
    * commute_time_average_hours AS s
  
```

Cloud-based ETL

- Data Lineage and Observability
- Task scheduling
- Run History
- SQL data modeling
- Checks and tests

Search for models...

Overview

raw.purchased_goods_invoice source table

Details Description Columns Referenced By SQL

Columns

COLUMN	TYPE	DESCRIPTION	CONSTRAINTS	DATA TESTS
sap_invoice_number		Unique identifier for SAP-g...		
invoice_number		General identifier for invoic...		
invoice_date		The date on which the invo...		
header_status		Status of the invoice head...		
po_number		Purchase order number rel...		
po_order_date		The date the purchase ord...		
po_status		Current status of the purc...		
commodity		Type of commodity or serv...		
po_line_commodity		Type of commodity or serv...		

Tables and Views

- dbt
- final
- raw
- annual_cpi_index
- commuting_emission_factors_EPA
- commuting_survey_2018
- commuting_survey_2021
- commuting_survey_2023
- commuting_survey_modes
- construction_expense
- cost_object_warehouse
- dof_maintenance_cost
- emission_factor_naics
- emission_factor_useeio_v2
- expense_emission_mapper
- food_emission_factor

MIT Data@MIT DataHub

Search DataHub...

6 resources | Page 1 of 1

clear all filters

MY FILES

- For private use
- For public use

SHARED PROJECTS

- Bluebikes
- campus-hazards
- Climate Resiliency
- Commuter MBTA
- Commuter Parkin...
- commuting

Processed Waste Data

Processed waste data including historical, Casella and small stream - 2024-10-1

Project: Material Matters | Owner: Yu Cheng | Org: MITOS | Size: 1.3 MB | Last Modified: 10/7:34:05 PM

waste_tonnage_newbatch.xlsx

Data Pulled from Casella PowerBI dashboard till August 2024

Project: Material Matters | Owner: Yu Cheng | Org: MITOS | Size: 107.8 KB | Last Modified: 11:35:59 PM

recycle-waste

waste_tonnage_june2023

Historical waste recycle till June 2023

Project: Material Matters | Owner: Yu Cheng | Org: MITOS | Size: 599.1 KB | Last Modified: 11/2:53:38 AM

Automated Documentation

- Single source of Truth for Dashboards and Download links.
- CI/CD includes ETL code deployments and data catalog.
- Integration with DataHub