

BONSAI

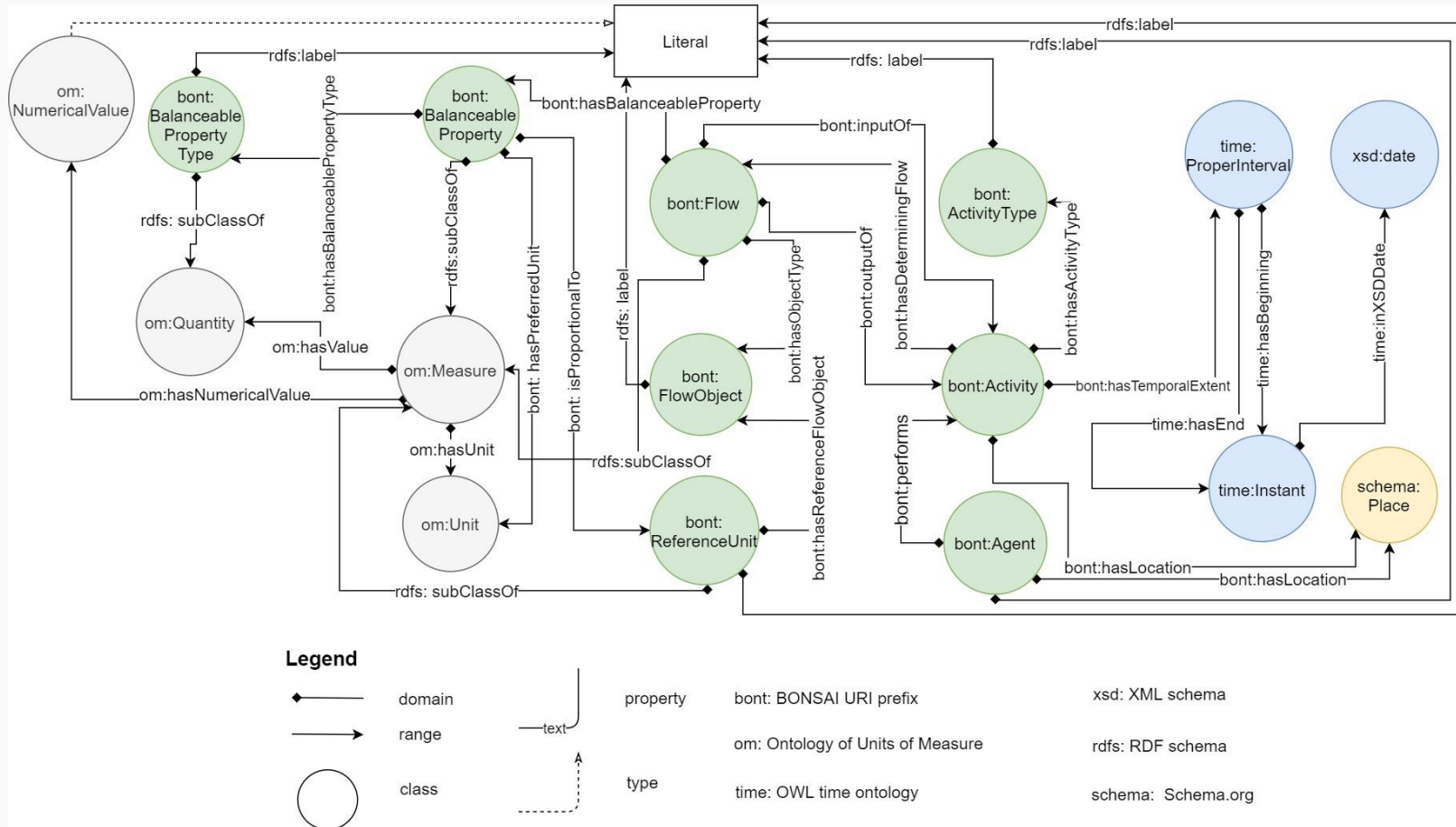
Getting the data right and 5 years of building a background database

`bonsai.uno`

What is BONSAI?

1. The ontology
2. The non-profit
3. The database
4. Getting the data right

The ontology



What is BONSAI?

1. The ontology
2. The non-profit
3. The database
4. Getting the data right

Part of the team



Our goal

1. Transparency
2. Traceability of data
3. Open source
4. Easy to update
5. Open to contribution from the community

Where are we right now?

- "Getting the data right" finished in June 2025
- We have data ready on bonsai.uno
- Code and development out in the open on [gitlab](https://gitlab.com/bonsai)
- Further funding for ~1 year of maintenance



Input-Output database?

- Built on data from global monetary flows
- Refined using physical flow data
- ~900 products in 43 countries and 5 Rest of World regions

What's included?

extraction



transformation



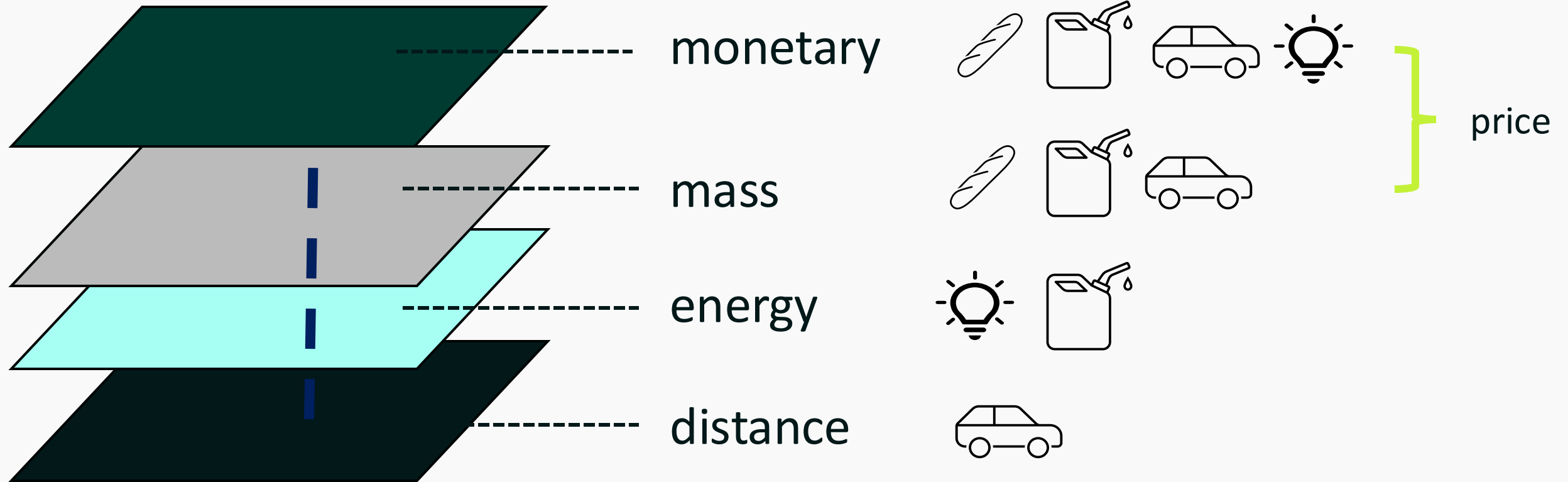
Consumption
/accumulation



emissions/waste



Multi-layer approach



How do we set things up?

- Development mostly in python (and a little R)
- Separation in 4 main steps
 - Collection/cleaning of data
 - Harmonization
 - Main modelling pipeline
 - Balancing and calculations
- Task orchestration via Apache Airflow
- Library for data handling between tasks
- Individual sector models for refining models

Task orchestration

- Currently 442 tasks in the workflow
 - Collect USGS data
 - Redistribute Energy to Incineration
 - Merge supply tables before balancing
- Workflow/Task classes
- Task_registration.py
- Airflow runs tasks in isolated environment

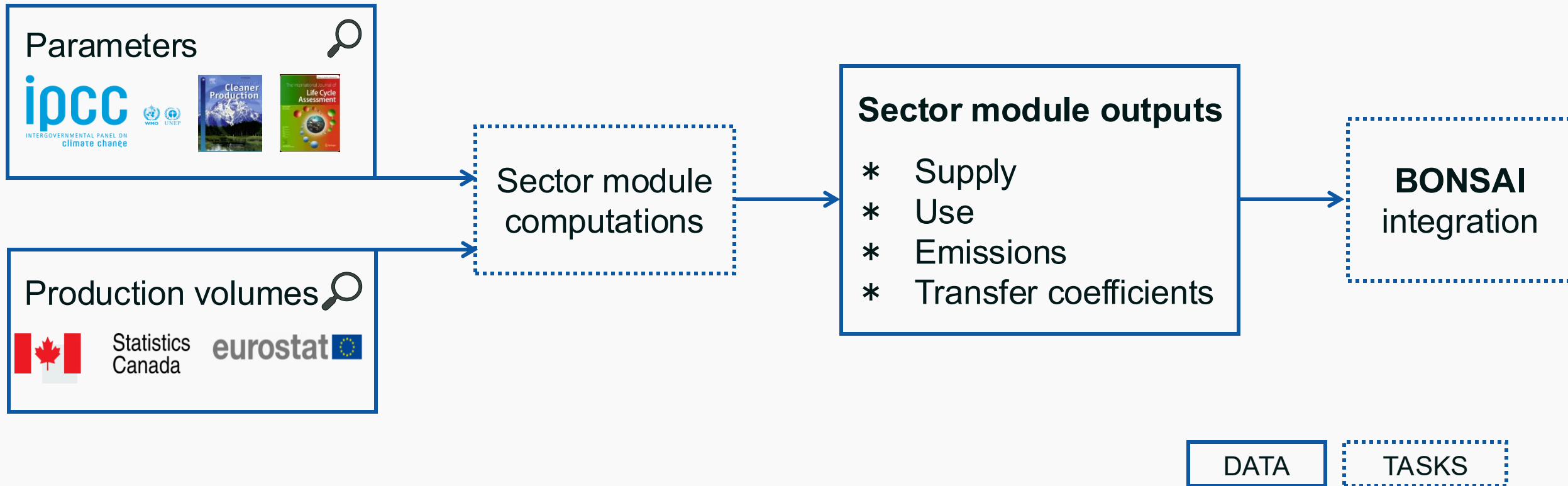
Data storage / schemas

- dataio package
 - Keep track of every read/write
 - Enforce schemas and vocabulary
 - Common interface for all tasks
- PyST for publishing
 - Hopefully soon also for maintaining

Uncertainty

- Bottleneck of Balancing (Using Markov Chain Monte Carlo (MCMC))
- Lack of data availability
- Currently not available

Refining data with models



Whats next?

- Core team keeps developing
- Next big update end of October
- It would be great if you start using data to report back and contribute

Closing remarks

Contact us through bonsai.uno or [gitlab](#)

