BONSAI Climate Footprint Analyser

Valentin Starlinger[1], Bo Weidema[2], Jannick Schmidt[2], Albert Osei-Owusu[2], Fan Yang[2], Maik Budzinski[2], Quanliang Ye[2], Sofia Topcu Madsen[2], Rahul Nallithodi[2], Jesper Dalgaard Pøhler[2], Stefano Merciai[1], Miguel Astudillo[1], Joao Rodrigues[1], Bertram de Boer[1], Angelica Mendoza Beltran[1], José Mogollón[3], Franco Donati[3], Anniek Kortleve[3], Glen Peters[4], Benjamin Sanderson[4]

[1] 2.-0 LCA consultants, [2]Aalborg University (lead), [3]CML, Leiden University, [4]CICERO, Oslo

2.0 LCA consultants

Overview

- BONSAI and the BONSAI Climate Footprint Analyser
- Database workflow and structure
- Integration with Apache Airflow
- Where we are right now
- Our API
- The road ahead & how you can help!
- Questions and API demo

2.0 LCA consultants

What's BONSAI?

- Non-profit organization

- Goal is to tackle the challenge of
  - Incomplete
  - Lacking in detail
  - Out-of-date
  - Difficult to access

"product footprints"

- How?
  - Completely open source and open data
  - Automated data harvesting
  - Use of national and international statistic data that can be disaggregated using science based algorithms
  - Community involvement

2.0 LCA consultants

BONSAI Climate Footprint Analyser

- Funded by KR foundation through ‘getting the data right’ project

- Project partners:
  - Aalborg University (lead)
  - CML, Leiden University
  - 2.-0 LCA consultants
  - CICERO, Oslo

- Heavily inspired by the hybrid version of Exiobase

2.0 LCA consultants

Goals

- Provide up-to-date climate footprint generator
- Minimize effort for continuous updates
- Enhance internal models:
  - Household consumption patterns, carbon flows in ecosystems, uncertainty handling, …
- Make it accessible!

Database key stats

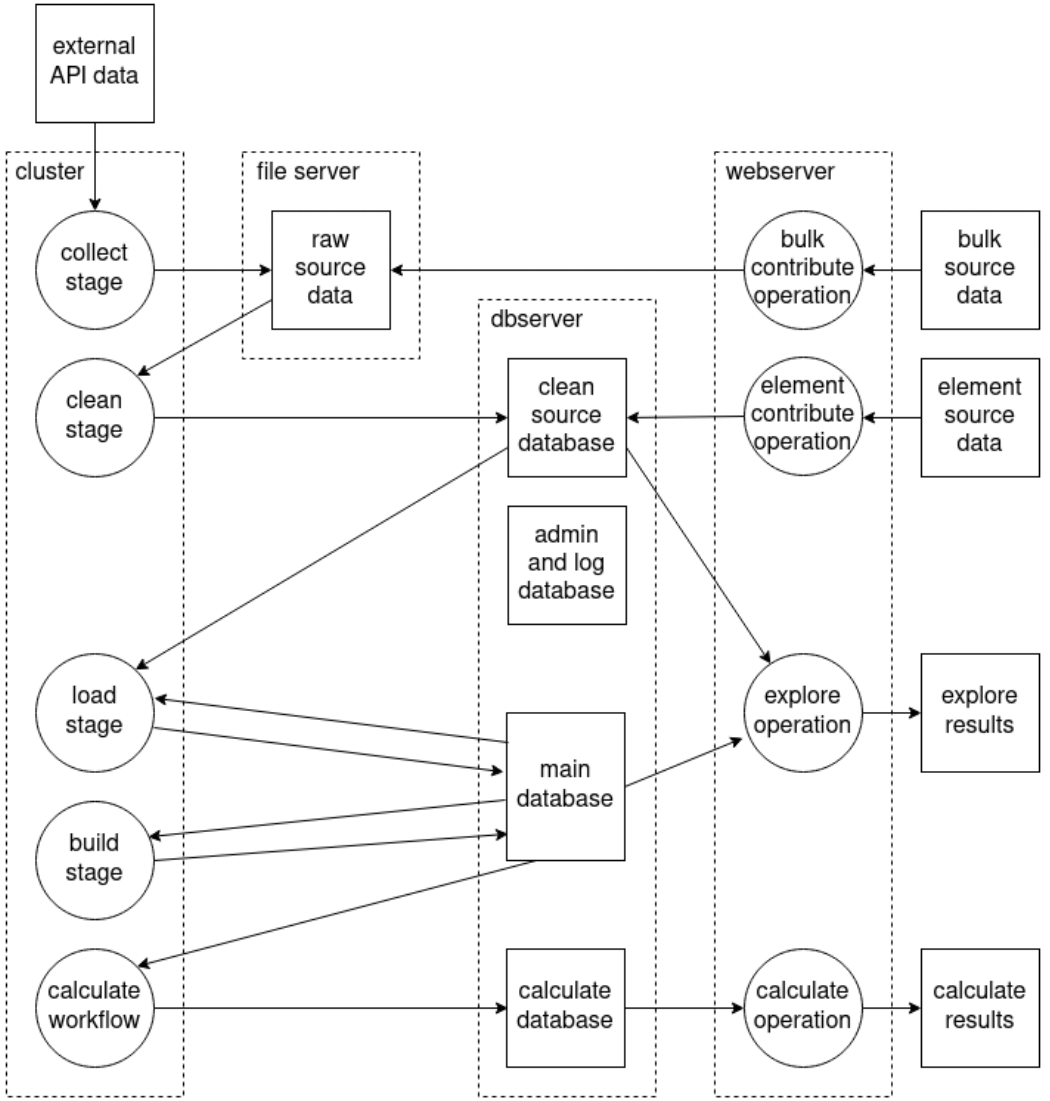
- Based on the hybrid version of Exiobase (v4)

Our minimum goal for data coverage:

- > 55 countries
- ~ 2000 product flows

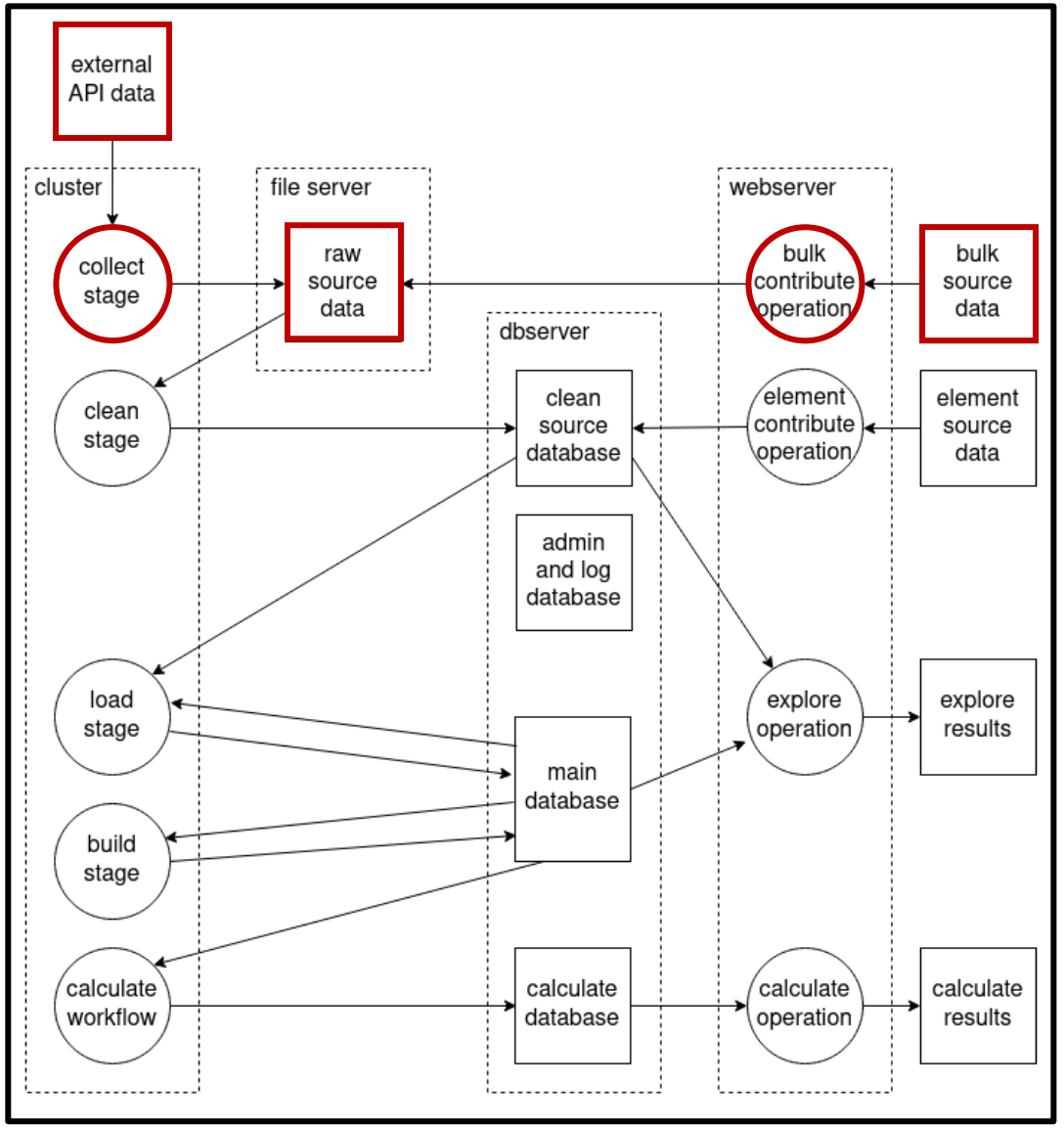
Database Workflow

- Webserver
  - React + Django
- DBServer
  - PostgreSQL
- Cluster workflow
  - Python
  - Apache Airflow
  - Kubernetes + Docker



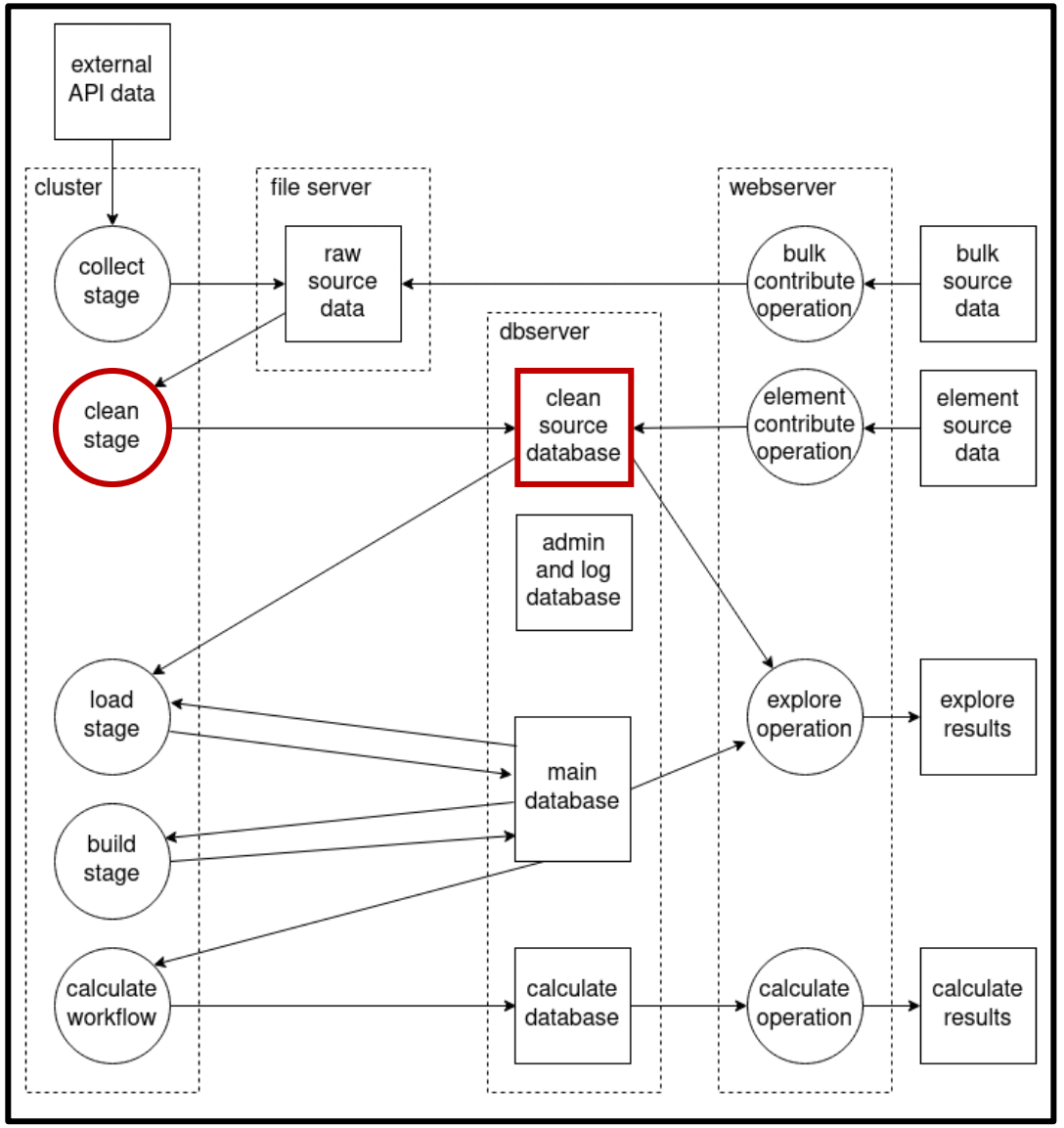
Collect and Clean Stage

- Input from external APIs and scraping websites
- Data stored as it comes
- Bulk contribution:
  - E.g.: Higher detail national SUT
  - Requires code contribution



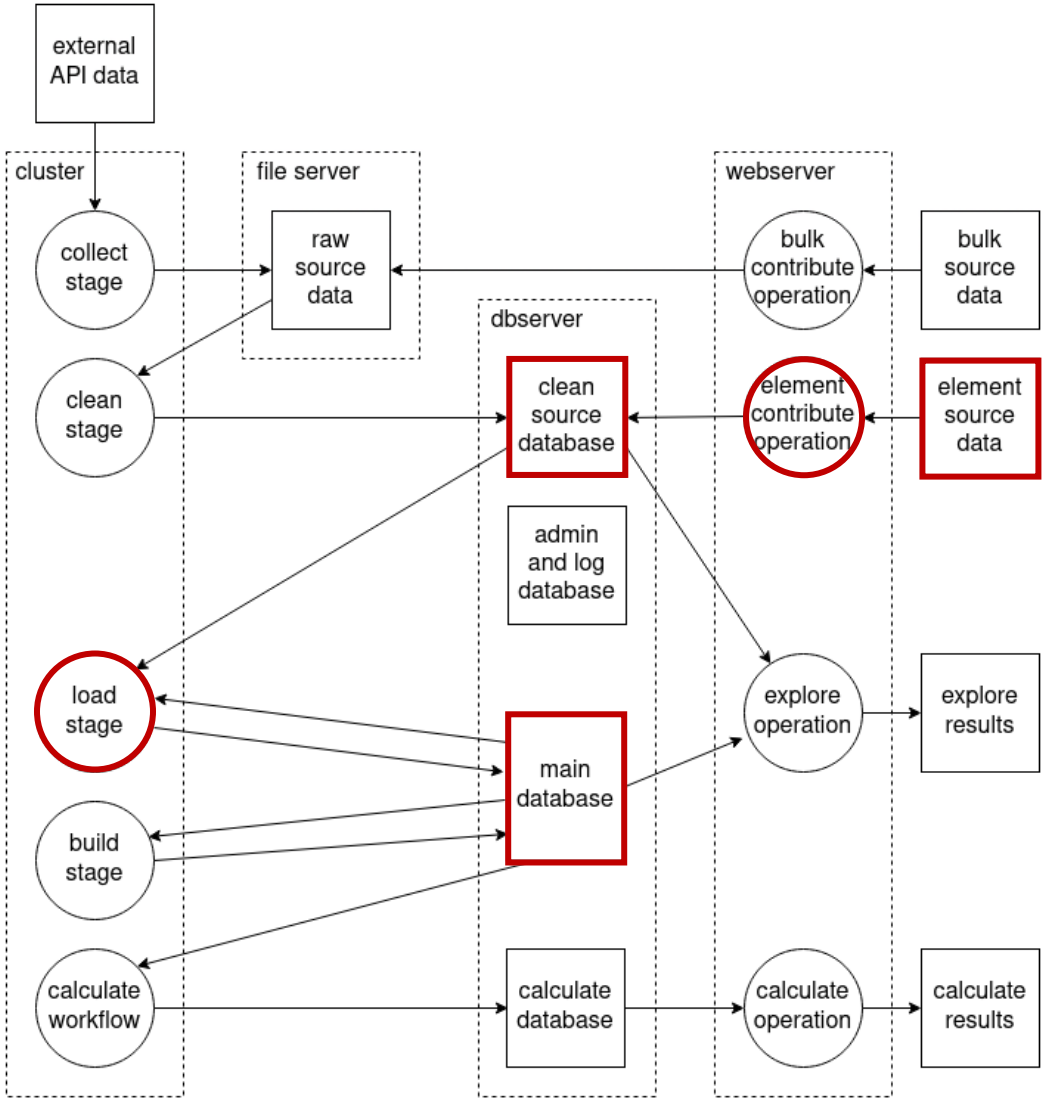
Clean Stage

- Data transformed to relational form
- Cleaning strategies applied
- Database and .csv files



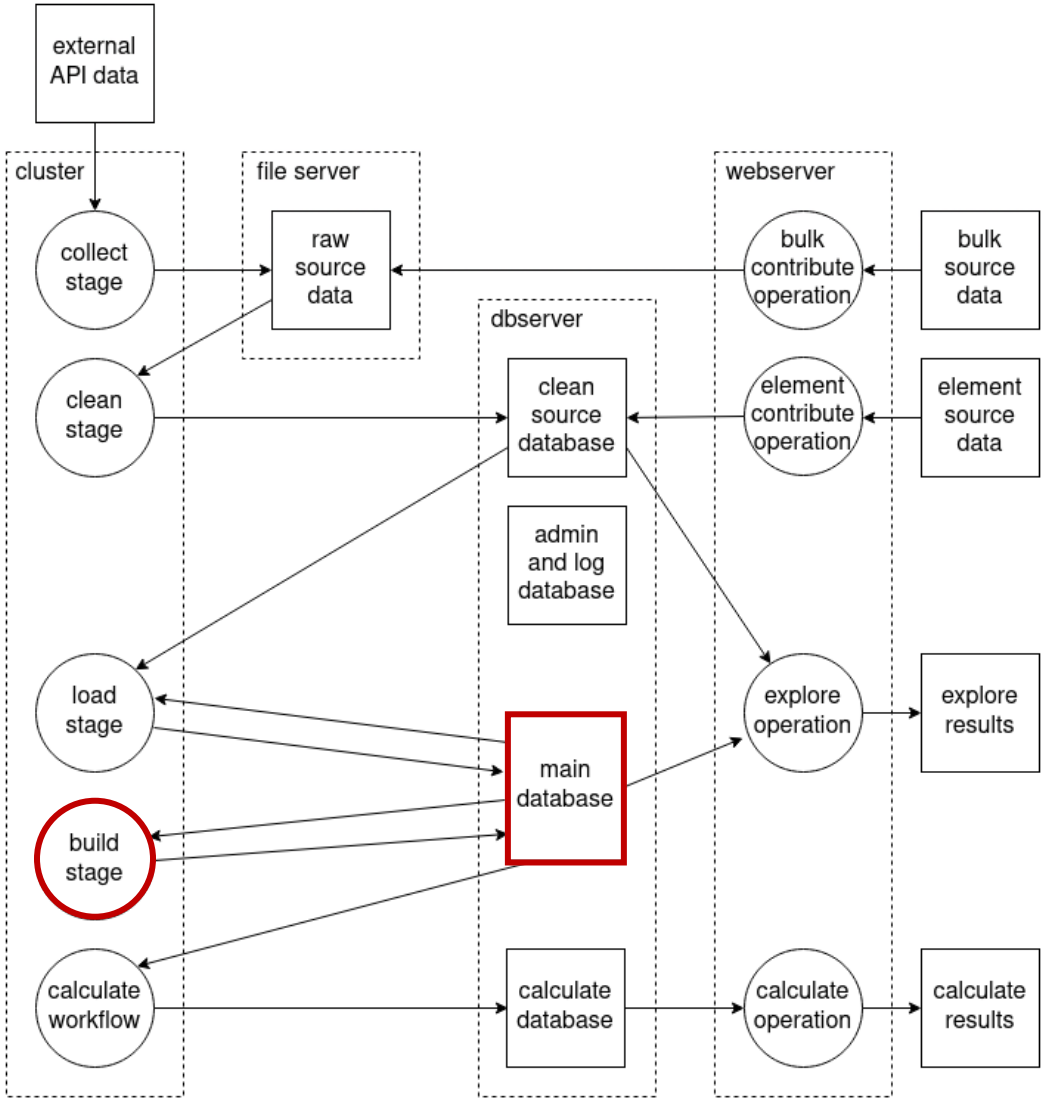
Load Stage

- Element contribution:
  - Highlight/edit erroneous data
- Reads and harmonizes clean data



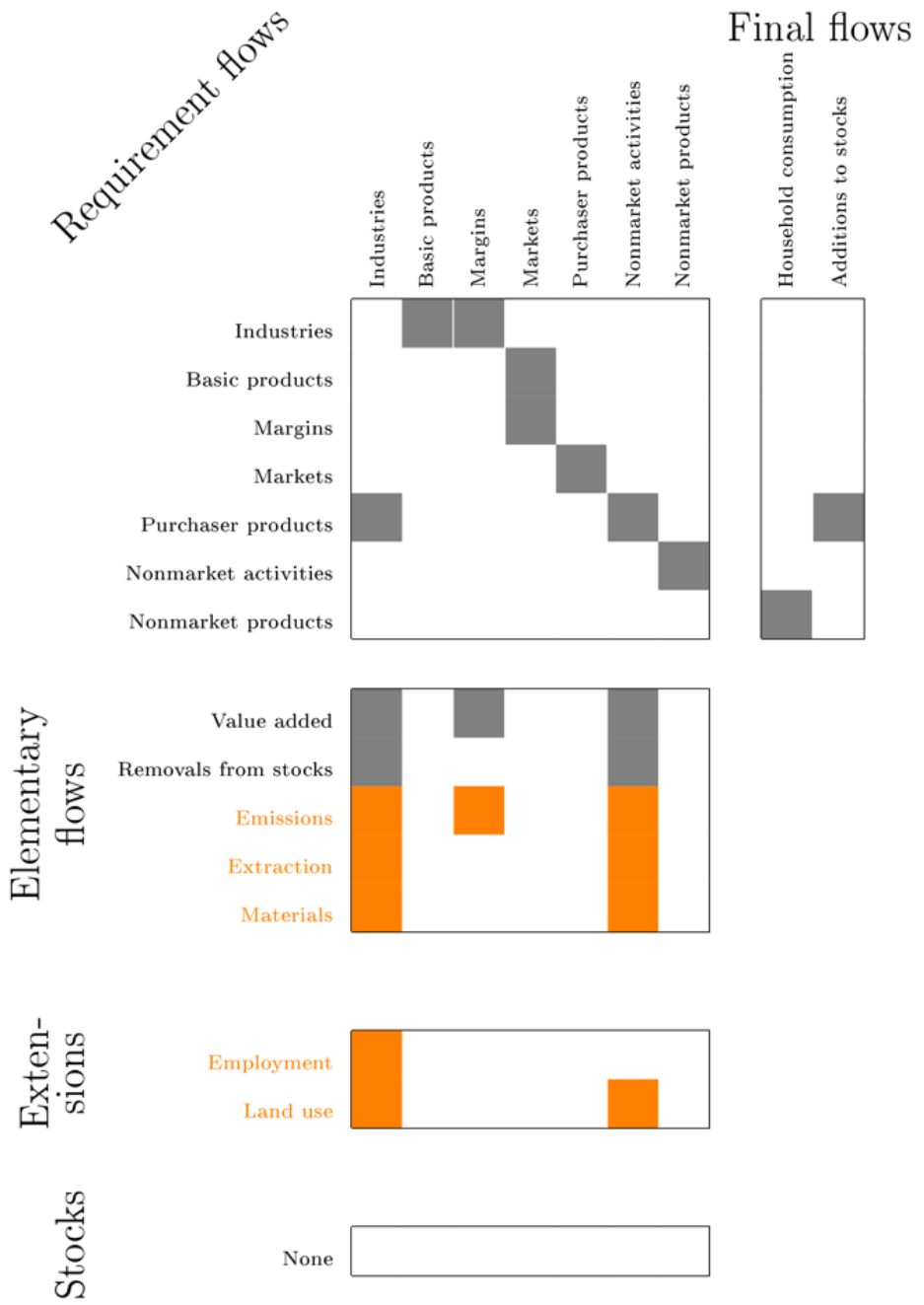
Build stage

- Transforms data
- Clean data to IO system
- Only interacts with main database



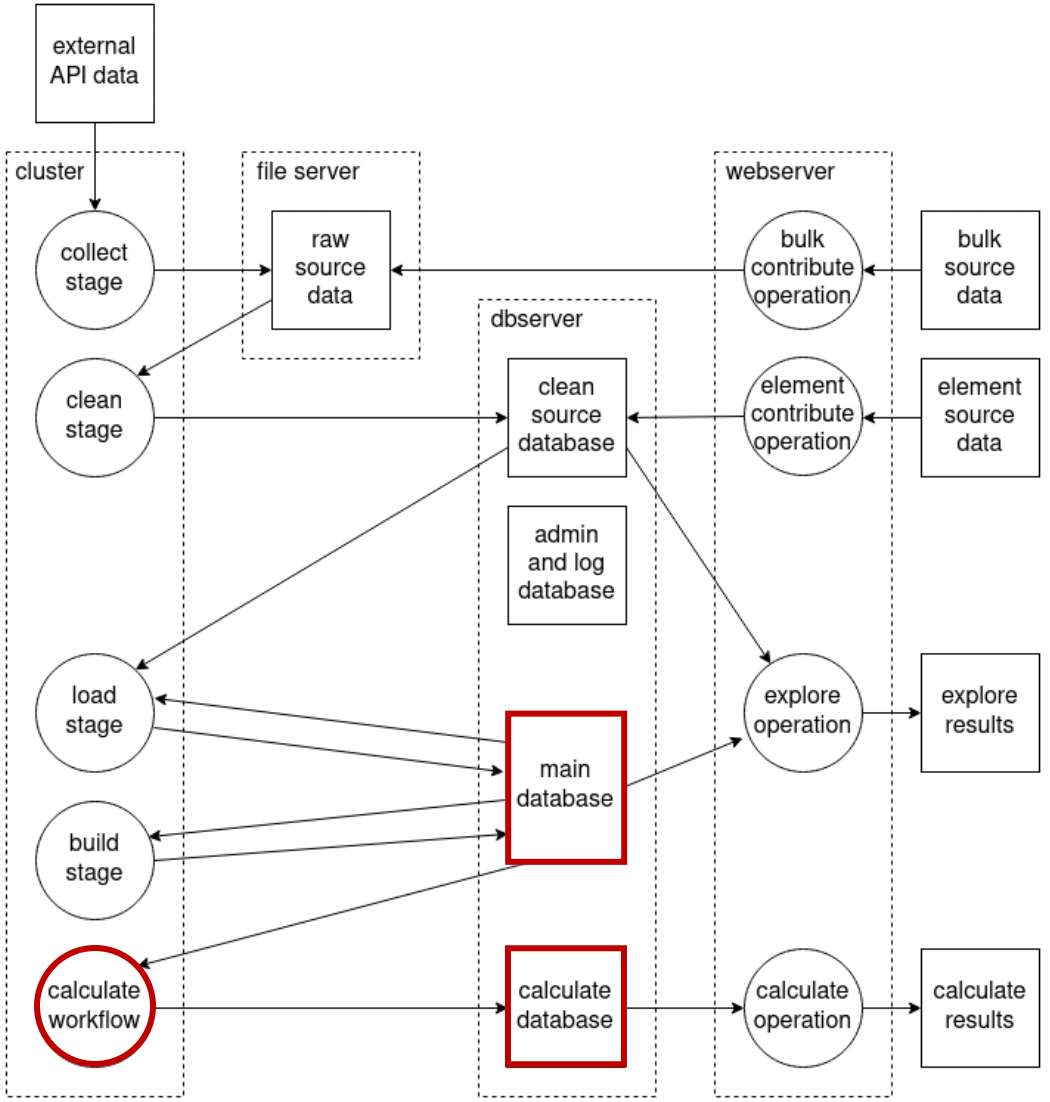
Build Stage

- Gap filling
- Decomposition of industries
- Parametrized production functions
- Modelling of sub-regional data
- Balancing



Calculate Stage

- Generates footprint results
- Generates hotspot/contribution analysis
- Results cached but not persistent

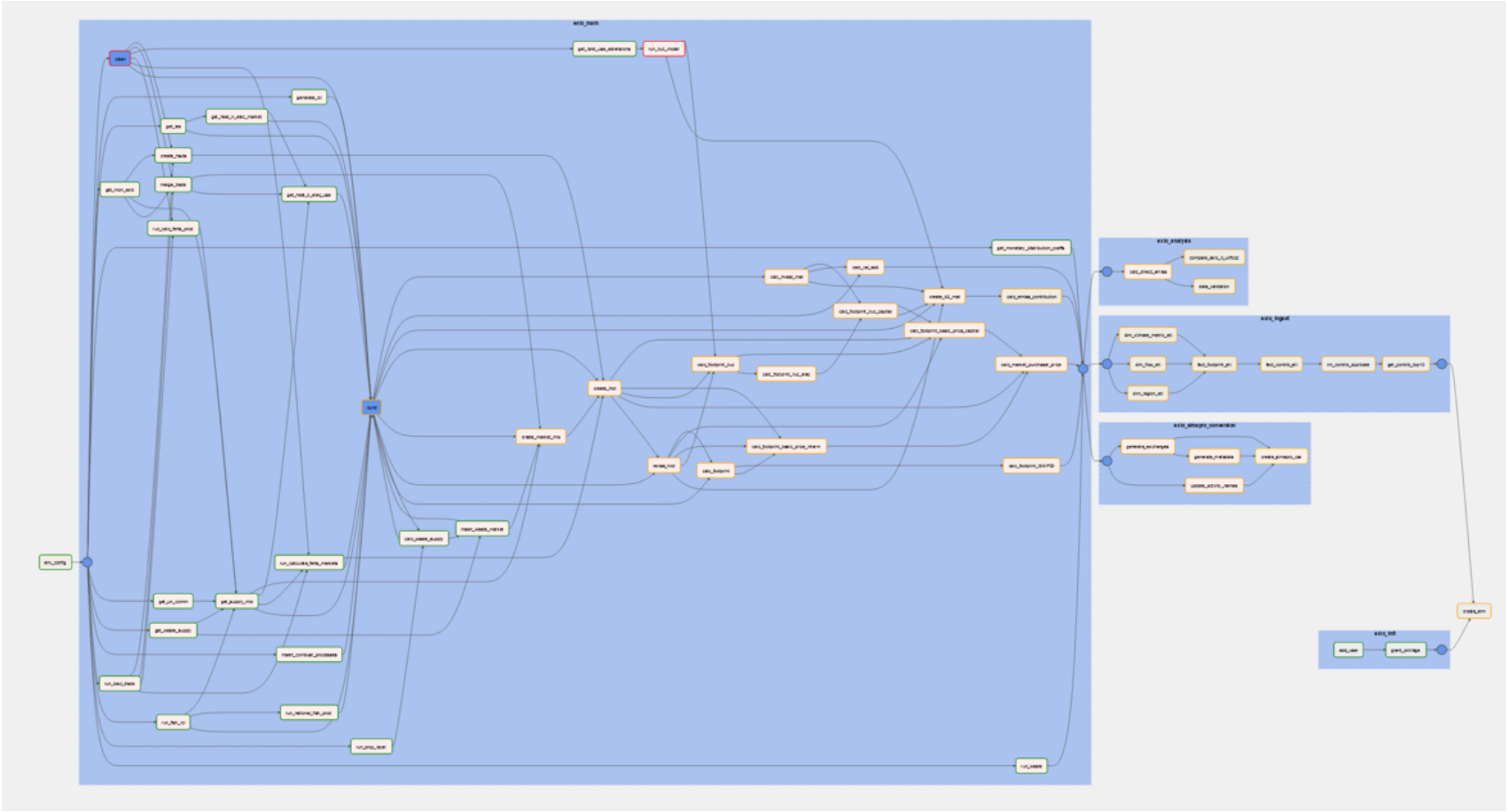


LCA consultants

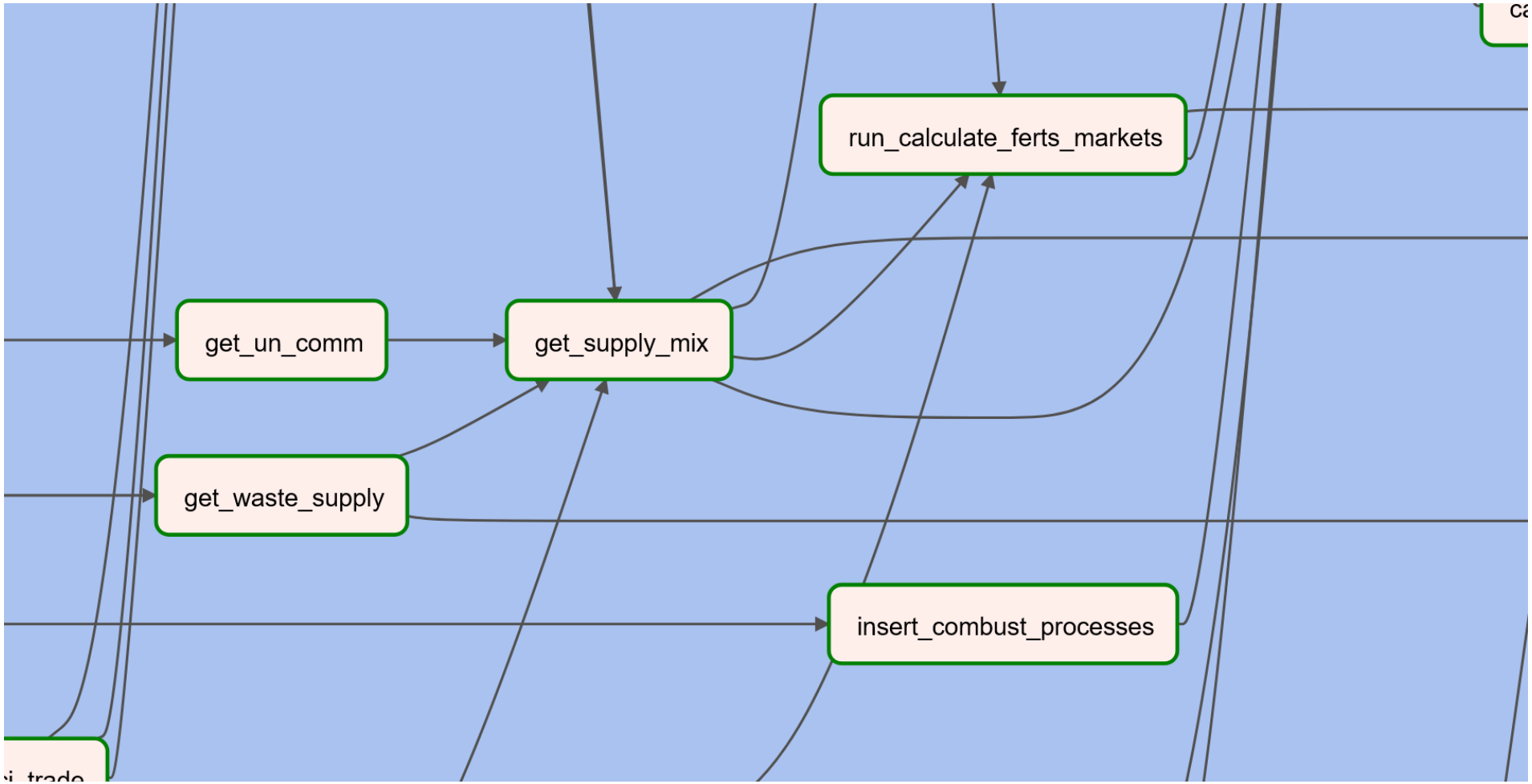
Integration with Apache Airflow

- Workflows can be stored under version control
- Allows to build a Directed Acyclic Graph (DAG)
- Manages task execution and load balancing
- Self-contained tasks for easier collaboration
- Airflow allows automatic scheduling of tasks
  - Triggers are automatic or manual
- Only necessary parts of workflow are executed

Integration with Apache Airflow



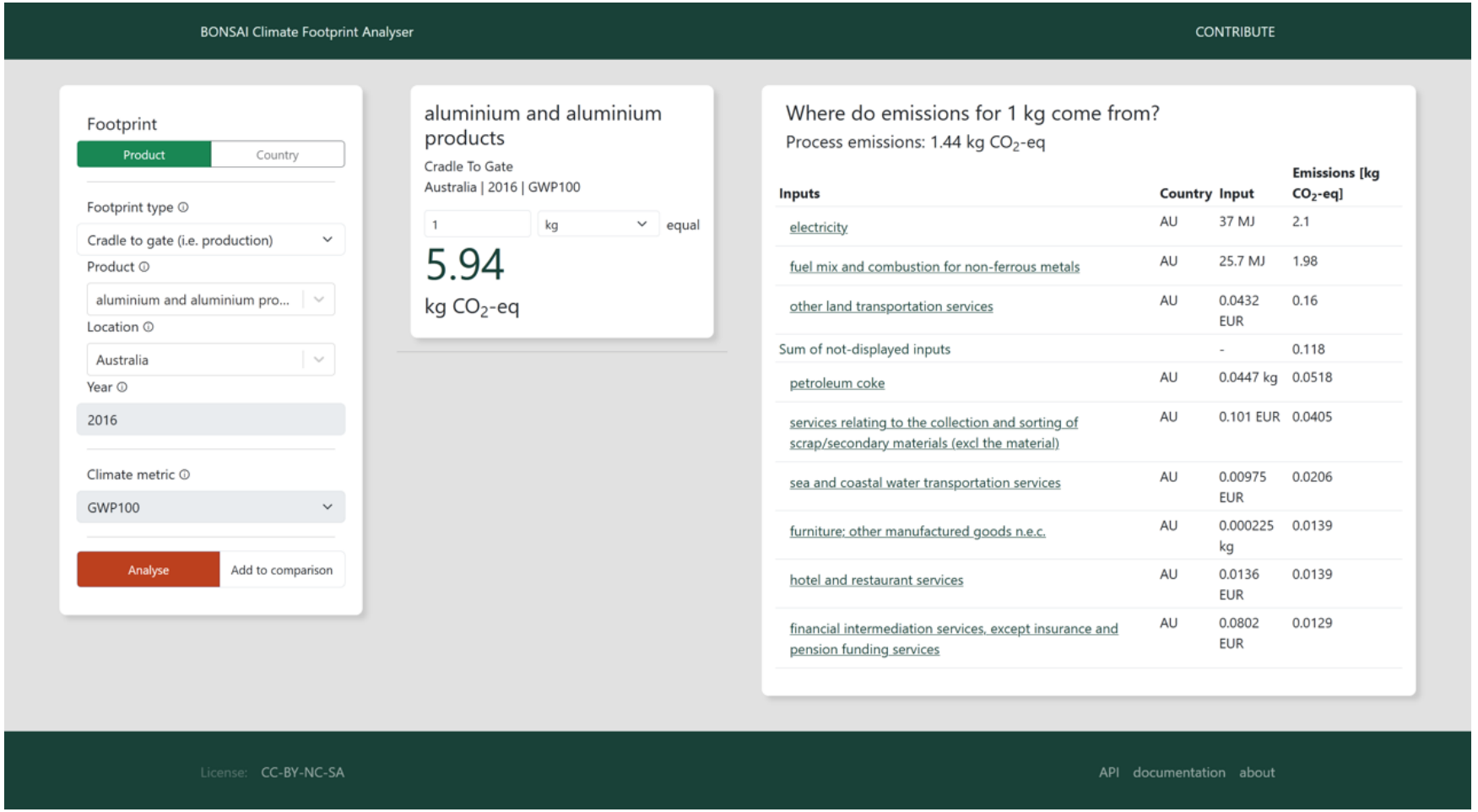
Integration with Apache Airflow



Parallelization using Docker + Kubernetes

- Each task can run on its own
- Tasks are launched as docker containers
- Distributed over several worker nodes
- Worker orchestration is done via a Kubernetes cluster

Where are we right now?



https://lca.aau.dk

Our current API

- Currently at: https://lca.aau.dk/api/docs/

- Available data:
  - Footprints for product/market flows
  - Production recipes for product flows
  - Market composition for market flows

  - Endpoints for available flows/countries/units

The Road Ahead & How You Can Help!

- Contact us if you have data!

- We want to start open-sourcing our pipeline start of next year
  - Frontend + API
  - IPCC package
  - Parametrized Models

2.0 LCA consultants

Questions + Live Demo

2.0 LCA consultants