



---

## Jana Customer Data Dictionary

**Based on generated dictionary:** 2026-03-31T07:35:37Z

**Tables (this document):** 26

**Domains:** air\_quality, emissions, entity\_data, carbon\_budget, severe\_weather

*Operational, ingestion, and job-management tables are omitted from this customer document. OpenAQ ingestion queues and S3 manifest tables are also omitted.*

For record counts and ingestion completeness, see the operations document *inventory\_data\_ingestion\_status\_all\_data\_sets.md* under `jana/docs/operations_docs/` in the **engineering-documents** repository (not bundled here).

---

### Table of Contents

- [Air Quality \(OpenAQ\)](#)
    - [datasets](#)
    - [openaq\\_locations](#)
    - [openaq\\_measurements](#)
    - [openaq\\_parameters](#)
    - [openaq\\_sensors](#)
  - [Emissions \(Climate TRACE, EDGAR\)](#)
    - [climatetrace\\_annual\\_country\\_emissions](#)
    - [climatetrace\\_assets](#)
    - [climatetrace\\_company\\_asset\\_matches](#)
    - [climatetrace\\_countries](#)
    - [climatetrace\\_emission\\_violations](#)
    - [climatetrace\\_emissions](#)
    - [climatetrace\\_sectors](#)
    - [edgar\\_air\\_pollutant\\_grid](#)
    - [edgar\\_air\\_pollutant\\_totals](#)
    - [edgar\\_country\\_totals](#)
    - [edgar\\_fasttrack\\_emissions](#)
    - [edgar\\_grid\\_emissions](#)
    - [edgar\\_temporal\\_profiles](#)
  - [Entity Data \(GLEIF\)](#)
    - [gleif\\_legal\\_entities](#)
    - [gleif\\_relationships](#)
    - [gleif\\_reporting\\_exceptions](#)
  - [Carbon & Methane Budgets \(GCP\)](#)
    - [gcp\\_gcpfossilemissionsbyfuel](#)
    - [gcp\\_gcpglobalcarbonbudget](#)
    - [gcp\\_gcpmethanebudget](#)
    - [gcp\\_gcpnationalemissions](#)
  - [Severe Weather \(NOAA Storm Events\)](#)
    - [noaa\\_storm\\_events\\_noaastormevent](#)
- 

### Air Quality (OpenAQ)

## datasets

Represents a specific dataset that our system maintains.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
name	varchar(255)	No	UNIQUE. Dataset name
dataset_type	varchar(50)	No	Type of dataset
description	text	No	Dataset description and purpose
geographic_scope	jsonb	No	Geographic boundaries (countries, states, bounding boxes, facility locations)
parameter_filters	jsonb	No	Environmental parameters to monitor (pm25, no2, o3, etc.)
temporal_scope	jsonb	No	Time range and frequency requirements
data_owner	varchar(255)	No	Who owns/manages this dataset
is_active	boolean	No	Whether dataset is actively maintained
created_at	timestamptz	No	Record creation timestamp
updated_at	timestamptz	No	Record last update timestamp

## openaq\_locations

OpenAQ monitoring location model.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
openaq_id	integer	No	UNIQUE. OpenAQ location ID
name	varchar(255)	No	Location name
country_code	varchar(3)	No	ISO 3166-1 alpha-3 country code
geometry	geometry(Point,4326)	Yes	PostGIS Point geometry (longitude, latitude) for spatial queries
latitude	numeric(10,7)	No	Latitude in decimal degrees (synced with geometry field)
longitude	numeric(10,7)	No	Longitude in decimal degrees (synced with geometry field)
timezone	varchar(50)	No	Location timezone
is_mobile	boolean	No	Is location mobile
is_monitor	boolean	No	Is location a monitor
owner_name	varchar(255)	No	Data owner name
provider_name	varchar(255)	No	Data provider name
datetime_first	timestamptz	Yes	First measurement timestamp
datetime_last	timestamptz	Yes	Last measurement timestamp
ingested_at	timestamptz	No	When location was ingested
updated_at	timestamptz	No	Last update timestamp

## Indexes:

- openaq\_loca\_country\_221bed\_idx (INDEX) on (country\_code)
- openaq\_loca\_latitud\_99403d\_idx (INDEX) on (latitude, longitude)
- openaq\_loca\_is\_moni\_fa1da5\_idx (INDEX) on (is\_monitor)
- openaq\_loca\_updated\_5569eb\_idx (INDEX) on (updated\_at)
- openaq\_locs\_dt\_first\_idx (INDEX) on (datetime\_first)
- openaq\_locs\_dt\_last\_idx (INDEX) on (datetime\_last)
- openaq\_locs\_provider\_idx (INDEX) on (provider\_name)
- openaq\_locs\_geom\_idx (INDEX) on (geometry)

## openaq\_measurements

OpenAQ measurement model for time-series air quality data.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
sensor_id	bigint	No	Associated sensor
location_id	bigint	No	Associated location (denormalized for query performance)
parameter_name	varchar(50)	No	Parameter name (denormalized)
value	numeric(15,6)	No	Measurement value
unit	varchar(20)	No	Measurement unit
latitude	numeric(10,7)	Yes	Latitude at measurement time (denormalized from location)
longitude	numeric(10,7)	Yes	Longitude at measurement time (denormalized from location)
measured_at	timestamptz	No	When measurement was taken
period_label	varchar(20)	No	Measurement period label
period_interval	varchar(20)	No	Measurement interval
flag	varchar(50)	No	Quality flag
flag_description	text	No	Quality flag description
ingested_at	timestamptz	No	Timestamp when record was ingested into the platform

### Foreign Keys:

- sensor\_id -> openaq\_sensors.id (ON DELETE CASCADE)
- location\_id -> openaq\_locations.id (ON DELETE CASCADE)

### Indexes:

- openaq\_meas\_sensor\_c63b78\_idx (INDEX) on (sensor, measured\_at)
- openaq\_meas\_locatio\_88ade6\_idx (INDEX) on (location, measured\_at)
- openaq\_meas\_paramet\_734403\_idx (INDEX) on (parameter\_name, measured\_at)
- openaq\_meas\_measure\_e065f4\_idx (INDEX) on (measured\_at)
- openaq\_meas\_latitud\_c2fea5\_idx (INDEX) on (latitude, longitude)
- openaq\_meas\_ingested\_idx (INDEX) on (ingested\_at)
- openaq\_meas\_loc\_param\_dt\_idx (INDEX) on (location, parameter\_name, measured\_at)
- openaq\_meas\_param\_dt\_val\_idx (INDEX) on (parameter\_name, measured\_at, value)
- openaq\_measurements\_unique\_sensor\_measured\_at\_parameter\_name (UNIQUE) on (sensor, measured\_at, parameter\_name)

## openaq\_parameters

Air quality parameter reference model.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
openaq_id	integer	No	UNIQUE. OpenAQ parameter ID
name	varchar(50)	No	UNIQUE. Parameter name (e.g., pm25, no2)
display_name	varchar(100)	No	Human readable parameter name
description	text	No	Parameter description
units	varchar(20)	No	Measurement units
preferred_unit	varchar(20)	No	Preferred display unit
created_at	timestamptz	No	Record creation timestamp
updated_at	timestamptz	No	Record last update timestamp

## openaq\_sensors

OpenAQ sensor model representing parameter measurements at a location.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
openaq_id	integer	No	UNIQUE. OpenAQ sensor ID
location_id	bigint	No	Associated location
parameter_id	bigint	No	Measured parameter
datetime_first	timestamptz	Yes	First measurement timestamp
datetime_last	timestamptz	Yes	Last measurement timestamp
latest_value	numeric(15,6)	Yes	Most recent measurement value
latest_datetime	timestamptz	Yes	Most recent measurement timestamp
ingested_at	timestamptz	No	Timestamp when sensor was first ingested
updated_at	timestamptz	No	Record last update timestamp

### Foreign Keys:

- location\_id -> openaq\_locations.id (ON DELETE CASCADE)
- parameter\_id -> openaq\_parameters.id (ON DELETE CASCADE)

### Indexes:

- openaq\_sens\_locatio\_ffef39\_idx (INDEX) on (location)
- openaq\_sens\_paramet\_51debe\_idx (INDEX) on (parameter)
- openaq\_sens\_latest\_9423b9\_idx (INDEX) on (latest\_datetime)
- openaq\_sens\_loc\_param\_dt\_idx (INDEX) on (location, parameter, latest\_datetime)
- openaq\_sensors\_unique\_location\_parameter (UNIQUE) on (location, parameter)

---

## Emissions (Climate TRACE, EDGAR)

### climatetrace\_annual\_country\_emissions

Climate TRACE annual country-level emissions aggregate.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
country_iso3	varchar(3)	No	ISO 3166-1 alpha-3 country code
year	integer	No	Emission year (e.g., 2023)
sector	varchar(200)	No	Climate TRACE sector name
subsector	varchar(200)	No	Climate TRACE subsector name
gas_type	varchar(50)	No	Gas type (co2, ch4, n2o, co2e_100yr, co2e_20yr)
emissions_quantity	numeric(20,6)	No	Emissions value in tonnes
uncertainty_lower	numeric(20,6)	Yes	Lower bound of uncertainty range
uncertainty_upper	numeric(20,6)	Yes	Upper bound of uncertainty range
country_name	varchar(200)	No	Country name (denormalized for convenience)
methodology	varchar(200)	No	Emission calculation methodology
data_source	varchar(100)	No	Data source: climatetrace_api, climatetrace_bulk_csv
source_version	varchar(50)	No	Climate TRACE data version (e.g., v5.2.0)
ingested_at	timestamptz	No	Timestamp when record was ingested into the platform
updated_at	timestamptz	No	Record last update timestamp

## Indexes:

- ct\_ace\_country\_year\_idx (INDEX) on (country\_iso3, year)
- ct\_ace\_year\_idx (INDEX) on (year)
- ct\_ace\_sector\_idx (INDEX) on (sector)
- ct\_ace\_gas\_type\_idx (INDEX) on (gas\_type)
- ct\_ace\_country\_gas\_year\_idx (INDEX) on (country\_iso3, gas\_type, year)
- ct\_ace\_sector\_year\_idx (INDEX) on (sector, year)
- climatetrace\_annual\_country\_emissions\_unique\_country\_iso3\_year\_sector\_gas\_type (UNIQUE) on (country\_iso3, year, sector, gas\_type)

## climatetrace\_assets

### Climate TRACE emitting asset model.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
climatetrace_id	varchar(200)	No	UNIQUE. Climate TRACE asset ID
asset_name	varchar(500)	No	Asset name
asset_type	varchar(200)	No	Asset type classification
sector_id	bigint	No	Associated emission sector
country_id	bigint	No	Asset country location
geometry	geometry(Point,4326)	Yes	PostGIS Point geometry (longitude, latitude) for spatial queries
latitude	numeric(10,7)	Yes	Latitude in decimal degrees (synced with geometry field)
longitude	numeric(10,7)	Yes	Longitude in decimal degrees (synced with geometry field)
owner_name	varchar(500)	No	Asset owner name
parent_company	varchar(500)	No	Parent company name
capacity	numeric(15,3)	Yes	Asset capacity
capacity_unit	varchar(50)	No	Capacity unit
capacity_mw	numeric(15,3)	Yes	Asset capacity in megawatts
start_time	timestamptz	Yes	Asset start time
end_time	timestamptz	Yes	Asset end time
created_date	timestamptz	Yes	Asset creation or first operation date
last_reported_date	timestamptz	Yes	Last emission data reporting date
latest_co2e_tonnes	numeric(18,6)	Yes	Most recent CO2e emission value
latest_emission_date	timestamptz	Yes	Date of most recent emission measurement
ingested_at	timestamptz	No	Timestamp when record was ingested into the platform
updated_at	timestamptz	No	Record last update timestamp

## Foreign Keys:

- sector\_id -> climatetrace\_sectors.id (ON DELETE CASCADE)
- country\_id -> climatetrace\_countries.id (ON DELETE CASCADE)

## Indexes:

- climatetrac\_sector\_\_b13252\_idx (INDEX) on (sector)
- climatetrac\_country\_9f1ab5\_idx (INDEX) on (country)
- climatetrac\_latitud\_2da526\_idx (INDEX) on (latitude, longitude)
- climatetrac\_owner\_n\_4ba73d\_idx (INDEX) on (owner\_name)

- climatetrac\_latest\_\_ab629d\_idx (INDEX) on (latest\_emission\_date)
- climatetrac\_updated\_02f8d7\_idx (INDEX) on (updated\_at)
- ct\_assets\_type\_idx (INDEX) on (asset\_type)
- ct\_assets\_id\_idx (INDEX) on (climatetrace\_id)
- ct\_assets\_country\_sector\_idx (INDEX) on (country, sector)
- ct\_assets\_sector\_type\_idx (INDEX) on (sector, asset\_type)
- ct\_assets\_start\_time\_idx (INDEX) on (start\_time)
- ct\_assets\_end\_time\_idx (INDEX) on (end\_time)
- ct\_assets\_geom\_idx (INDEX) on (geometry)
- ct\_assets\_country\_latest\_co2e\_idx (INDEX) on ()
- ct\_assets\_parent\_company\_idx (INDEX) on (parent\_company)

### climatetrace\_company\_asset\_matches

Links Climate TRACE assets to companies with confidence scoring.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
company_id	integer	No	Legacy company ID (retained for backward compatibility)
legal_entity_id	bigint	Yes	GLEIF legal entity matched to this asset (LEI-based)
asset_id	bigint	No	Associated Climate TRACE asset
matching_method	varchar(20)	No	Method used to establish company-asset relationship
matching_confidence	numeric(3,2)	No	Confidence score for the match (0-1)
relationship_type	varchar(20)	No	Type of company-asset relationship
ownership_percentage	numeric(5,2)	Yes	Ownership percentage if applicable
verified	boolean	No	Human verified match
verified_by	varchar(200)	No	Verifier name or system
verified_at	timestamptz	Yes	Verification timestamp
notes	text	No	Additional context or notes
created_at	timestamptz	No	Record creation timestamp
updated_at	timestamptz	No	Record last update timestamp

#### Foreign Keys:

- legal\_entity\_id -> gleif\_legal\_entities.id (ON DELETE SET\_NULL)
- asset\_id -> climatetrace\_assets.id (ON DELETE CASCADE)

#### Indexes:

- climatetrac\_company\_68fd4a\_idx (INDEX) on (company\_id)
- climatetrac\_asset\_i\_01a876\_idx (INDEX) on (asset)
- climatetrac\_matchin\_06009a\_idx (INDEX) on (matching\_confidence)
- climatetrac\_verifie\_e26c88\_idx (INDEX) on (verified)
- climatetrac\_created\_7c15f2\_idx (INDEX) on (created\_at)
- ct\_cam\_legal\_entity\_idx (INDEX) on (legal\_entity)
- climatetrace\_company\_asset\_matches\_unique\_company\_id\_asset (UNIQUE) on (company\_id, asset)

### climatetrace\_countries

Climate TRACE country reference model.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
iso3	varchar(3)	No	UNIQUE. ISO 3166-1 alpha-3 country code

Column	Type	Nullable	Description
name	varchar(200)	No	Country name
region	varchar(100)	No	Geographic region
income_group	varchar(100)	No	World Bank income group
total_assets_count	integer	No	Total number of tracked assets in country
last_updated	timestamptz	No	Timestamp of last data update from Climate TRACE

**Indexes:**

- climatetrac\_iso3\_3be133\_idx (INDEX) on (iso3)
- climatetrac\_region\_4761a6\_idx (INDEX) on (region)

**climatetrace\_emission\_violations**

Tracks emission violations and anomalies for ESG risk assessment.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
asset_id	bigint	No	Asset with violation
company_id	integer	Yes	Associated company ID if matched
violation_type	varchar(30)	No	Type of emission violation
gas_type	varchar(10)	No	Gas type involved in violation
threshold_value	numeric(15,3)	Yes	Threshold or expected value
actual_value	numeric(15,3)	No	Actual emission value that triggered violation
violation_severity	varchar(10)	No	Severity level of the violation
detection_date	timestamptz	No	When violation was detected
period_start	timestamptz	No	Start of emission period with violation
period_end	timestamptz	No	End of emission period with violation
description	text	No	Detailed violation description
resolved	boolean	No	Whether violation has been resolved
resolved_at	timestamptz	Yes	Resolution timestamp
resolution_notes	text	No	Resolution details

**Foreign Keys:**

- asset\_id -> climatetrace\_assets.id (ON DELETE CASCADE)

**Indexes:**

- climatetrac\_asset\_i\_1221c6\_idx (INDEX) on (asset, detection\_date)
- climatetrac\_company\_a1a1fa\_idx (INDEX) on (company\_id, detection\_date)
- climatetrac\_violati\_c0a90f\_idx (INDEX) on (violation\_severity, resolved)
- climatetrac\_detecti\_55acb9\_idx (INDEX) on (detection\_date)

**climatetrace\_emissions**

Climate TRACE emission measurement model for time-series data.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
asset_id	bigint	No	Associated emitting asset
climatetrace_id	varchar(200)	No	UNIQUE. Climate TRACE emission ID

Column	Type	Nullable	Description
climatetrace_asset_id	varchar(200)	No	Asset ID (denormalized for performance)
sector_name	varchar(200)	No	Sector name (denormalized for performance)
country_iso3	varchar(3)	No	Country code (denormalized for performance)
temporal_granularity	varchar(20)	No	Time period granularity
start_time	timestamptz	No	Emission period start time
end_time	timestamptz	Yes	Emission period end time
gas	varchar(10)	No	Gas type
co2_tonnes	numeric(18,6)	Yes	CO2 emissions in tonnes
ch4_tonnes	numeric(18,6)	Yes	CH4 emissions in tonnes
n2o_tonnes	numeric(18,6)	Yes	N2O emissions in tonnes
co2e_tonnes	numeric(18,6)	No	CO2 equivalent emissions in tonnes
emissions_factor	numeric(15,6)	Yes	Emission factor used in calculation
emissions_factor_unit	varchar(100)	No	Emission factor unit
activity_data	numeric(18,6)	Yes	Activity data used in calculation
activity_data_unit	varchar(100)	No	Activity data unit
uncertainty_lower	numeric(18,6)	Yes	Lower bound of uncertainty range
uncertainty_upper	numeric(18,6)	Yes	Upper bound of uncertainty range
data_quality_score	numeric(3,2)	Yes	Data quality score (0-1)
ingested_at	timestamptz	No	Timestamp when record was ingested into the platform

#### Foreign Keys:

- asset\_id -> climatetrace\_assets.id (ON DELETE CASCADE)

#### Indexes:

- climatetrac\_asset\_i\_3cbf5b\_idx (INDEX) on (asset, start\_time)
- climatetrac\_climate\_86797a\_idx (INDEX) on (climatetrace\_asset\_id, start\_time)
- climatetrac\_sector\_\_06ebf4\_idx (INDEX) on (sector\_name, start\_time)
- climatetrac\_country\_1df736\_idx (INDEX) on (country\_iso3, start\_time)
- climatetrac\_tempora\_d10f65\_idx (INDEX) on (temporal\_granularity, start\_time)
- climatetrac\_co2e\_to\_8d3755\_idx (INDEX) on (co2e\_tonnes)
- climatetrac\_start\_t\_42d1d6\_idx (INDEX) on (start\_time)
- ct\_emissions\_gas\_idx (INDEX) on (gas)
- ct\_em\_country\_gran\_dt\_idx (INDEX) on (country\_iso3, temporal\_granularity, start\_time)
- ct\_emissions\_quality\_idx (INDEX) on (data\_quality\_score)
- ct\_emissions\_sector\_gas\_dt\_idx (INDEX) on (sector\_name, gas, start\_time)
- climatetrace\_emissions\_unique\_asset\_temporal\_granularity\_start\_time (UNIQUE) on (asset, temporal\_granularity, start\_time)

#### climatetrace\_sectors

##### Climate TRACE emission sector reference model.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
climatetrace_id	varchar(100)	No	UNIQUE. Climate TRACE sector ID
name	varchar(200)	No	Sector name
display_name	varchar(200)	No	Human readable sector name

Column	Type	Nullable	Description
description	text	No	Sector description
parent_sector_id	bigint	Yes	Parent sector for hierarchical organization
methodology_version	varchar(50)	No	Emission calculation methodology version
emission_factors	jsonb	No	Emission factors by gas type
created_at	timestamptz	No	Record creation timestamp
updated_at	timestamptz	No	Record last update timestamp

#### Foreign Keys:

- parent\_sector\_id -> climatetrace\_sectors.id (ON DELETE SET\_NULL)

#### Indexes:

- climatetrac\_name\_58a810\_idx (INDEX) on (name)
- climatetrac\_parent\_\_9010c0\_idx (INDEX) on (parent\_sector)
- ct\_sectors\_display\_name\_idx (INDEX) on (display\_name)
- ct\_sectors\_id\_idx (INDEX) on (climatetrace\_id)

#### edgar\_air\_pollutant\_grid

##### Air pollutant gridded emissions at 0.1 x 0.1 degree resolution from EDGAR.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
created_at	timestamptz	No	Record creation timestamp
updated_at	timestamptz	No	Record last update timestamp
grid_id	varchar(64)	No	Grid cell identifier (lon:lat format)
year	integer	No	Reporting year for emissions
gas	varchar(50)	No	Air pollutant type (NOx, SO2, PM25, PM10, CO, NMVOC, NH3, BC, OC)
sector	varchar(120)	No	EDGAR economic sector classification code
value	numeric(20,6)	No	Emission value in tonnes for the grid cell
lat	numeric(10,6)	No	Latitude in decimal degrees (synced with geometry field)
lon	numeric(10,6)	No	Longitude in decimal degrees (synced with geometry field)
geometry	geometry(Point,4326)	Yes	PostGIS Point geometry (longitude, latitude) for spatial queries
source_version	varchar(50)	No	EDGAR release version identifier
ingested_at	timestamptz	No	Timestamp when record was ingested into the platform

#### Indexes:

- edgar\_ap\_grid\_year\_idx (INDEX) on (year)
- edgar\_ap\_grid\_gas\_idx (INDEX) on (gas)
- edgar\_ap\_grid\_gas\_yr\_sec\_idx (INDEX) on (gas, year, sector)
- edgar\_ap\_grid\_geom\_idx (INDEX) on (geometry)
- edgar\_air\_pollutant\_grid\_unique\_grid\_id\_year\_gas\_sector (UNIQUE) on (grid\_id, year, gas, sector)

#### edgar\_air\_pollutant\_totals

##### Annual national air pollutant emission totals (non-GHG) from EDGAR.

Column	Type	Nullable	Description
id	bigserial	No	Primary key

Column	Type	Nullable	Description
created_at	timestampz	No	Record creation timestamp
updated_at	timestampz	No	Record last update timestamp
country_code	varchar(3)	No	ISO 3166-1 alpha-3 country code
year	integer	No	Reporting year for emissions
gas	varchar(50)	No	Air pollutant type (NOx, SO2, PM25, PM10, CO, NMVOC, NH3, BC, OC)
sector	varchar(120)	No	EDGAR economic sector classification code
value	numeric(20,6)	No	Emission value for the given pollutant and sector
unit	varchar(20)	No	Unit of measurement (e.g., Gg)
source_version	varchar(50)	No	EDGAR release version identifier
ingested_at	timestampz	No	Timestamp when record was ingested into the platform

#### Indexes:

- edgar\_ap\_totals\_ct\_idx (INDEX) on (country\_code, year)
- edgar\_ap\_totals\_gas\_idx (INDEX) on (gas)
- edgar\_ap\_totals\_sector\_idx (INDEX) on (sector)
- edgar\_air\_pollutant\_totals\_unique\_country\_code\_year\_gas\_sector (UNIQUE) on (country\_code, year, gas, sector)

#### edgar\_country\_totals

##### National GHG emission totals by country, year, gas, and sector from EDGAR.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
created_at	timestampz	No	Record creation timestamp
updated_at	timestampz	No	Record last update timestamp
country_code	varchar(3)	No	ISO 3166-1 alpha-3 country code
year	integer	No	Reporting year for emissions
gas_type	varchar(50)	No	Actual gas type (CO2, CH4, N2O, GWP_100_AR5_GHG, etc.)
region_name	varchar(200)	Yes	Region/country name (e.g., 'Oceania', 'Japan', 'Aruba')
sector	varchar(120)	No	EDGAR economic sector classification code
value	numeric(20,6)	No	Emission value in tonnes for the given gas and sector
source_version	varchar(50)	No	EDGAR release version identifier (e.g., v8.0_FT2022)
ingested_at	timestampz	No	Timestamp when record was ingested into the platform
provisional	boolean	No	Whether this emission value is a provisional (fast-track) estimate

#### Indexes:

- edgar\_cty\_totals\_ct\_idx (INDEX) on (country\_code, year)
- edgar\_cty\_gas\_type\_idx (INDEX) on (gas\_type)
- edgar\_cty\_region\_idx (INDEX) on (region\_name)
- edgar\_cty\_totals\_s\_idx (INDEX) on (sector)
- edgar\_cty\_totals\_year\_idx (INDEX) on (year)
- edgar\_cty\_ct\_gas\_year\_idx (INDEX) on (country\_code, gas\_type, year)
- edgar\_cty\_sector\_year\_idx (INDEX) on (sector, year)
- edgar\_cty\_ordering\_idx (INDEX) on (year, country\_code, gas\_type)
- edgar\_country\_totals\_unique\_country\_code\_year\_gas\_type\_sector (UNIQUE) on (country\_code, year, gas\_type, sector)

## edgar\_fasttrack\_emissions

Provisional (fast-track) GHG emission estimates for recent years not yet in the official EDGAR release.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
created_at	timestamptz	No	Record creation timestamp
updated_at	timestamptz	No	Record last update timestamp
country_code	varchar(3)	No	ISO 3166-1 alpha-3 country code
year	integer	No	Reporting year for emissions
gas	varchar(50)	No	GHG gas type (CO2, CH4, N2O)
sector	varchar(120)	No	EDGAR economic sector classification code
value	numeric(20,6)	No	Provisional emission value in tonnes
source_version	varchar(50)	No	EDGAR release version identifier
ingested_at	timestamptz	No	Timestamp when record was ingested into the platform
provisional	boolean	No	Always true for fast-track data

### Indexes:

- edgar\_ft\_country\_year\_idx (INDEX) on (country\_code, year)
- edgar\_ft\_gas\_idx (INDEX) on (gas)
- edgar\_ft\_sector\_idx (INDEX) on (sector)
- edgar\_ft\_provisional\_idx (INDEX) on (provisional)
- edgar\_ft\_gas\_year\_idx (INDEX) on (gas, year)
- edgar\_fasttrack\_emissions\_unique\_country\_code\_year\_gas\_sector (UNIQUE) on (country\_code, year, gas, sector)

## edgar\_grid\_emissions

Spatially gridded GHG emissions at 0.1 x 0.1 degree resolution from EDGAR.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
created_at	timestamptz	No	Record creation timestamp
updated_at	timestamptz	No	Record last update timestamp
grid_id	varchar(64)	No	Grid cell identifier (lon:lat format)
year	integer	No	Reporting year for emissions
gas	varchar(50)	No	GHG gas type (CO2, CH4, N2O, F-gases)
sector	varchar(120)	No	EDGAR economic sector classification code
value	numeric(20,6)	No	Emission value in tonnes for the grid cell
lat	numeric(10,6)	No	Latitude in decimal degrees (synced with geometry field)
lon	numeric(10,6)	No	Longitude in decimal degrees (synced with geometry field)
geometry	geometry(Point,4326)	Yes	PostGIS Point geometry (longitude, latitude) for spatial queries
source_version	varchar(50)	No	EDGAR release version identifier
ingested_at	timestamptz	No	Timestamp when record was ingested into the platform

### Indexes:

- edgar\_grid\_year\_idx (INDEX) on (year)
- edgar\_grid\_gas\_idx (INDEX) on (gas)
- edgar\_grid\_sector\_idx (INDEX) on (sector)
- edgar\_grid\_id\_idx (INDEX) on (grid\_id)

- edgar\_grid\_year\_sector\_idx (INDEX) on (year, sector)
- edgar\_grid\_gas\_year\_sector\_idx (INDEX) on (gas, year, sector)
- edgar\_grid\_geom\_idx (INDEX) on (geometry)
- edgar\_grid\_year\_lat\_lon\_idx (INDEX) on (year, lat, lon)
- edgar\_grid\_emissions\_unique\_grid\_id\_year\_gas\_sector (UNIQUE) on (grid\_id, year, gas, sector)

## edgar\_temporal\_profiles

### Temporal disaggregation profiles (monthly/hourly) for EDGAR emissions data.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
created_at	timestamptz	No	Record creation timestamp
updated_at	timestamptz	No	Record last update timestamp
region_country	varchar(50)	Yes	Region or country code from EDGAR
sector	varchar(120)	No	EDGAR economic sector classification code
year	integer	Yes	Year for temporal profile (0 means no time-dependent profiles)
temporal_level	varchar(20)	No	Temporal resolution level (monthly or hourly)
month	smallint	Yes	Month number (1-12) for monthly profiles
hour	smallint	Yes	Hour of day (0-23) for hourly profiles
factor	numeric(12,8)	No	Temporal scaling factor applied to annual emissions
profile_version	varchar(50)	No	Version of the temporal profile dataset
source_version	varchar(50)	No	EDGAR release version identifier
ingested_at	timestamptz	No	Timestamp when record was ingested into the platform

### Indexes:

- edgar\_tp\_sector\_idx (INDEX) on (sector)
- edgar\_tp\_level\_idx (INDEX) on (temporal\_level)
- edgar\_tp\_region\_idx (INDEX) on (region\_country)
- edgar\_tp\_year\_idx (INDEX) on (year)
- edgar\_tp\_sector\_level\_year\_idx (INDEX) on (sector, temporal\_level, year)
- edgar\_tp\_region\_sector\_idx (INDEX) on (region\_country, sector)
- edgar\_temporal\_profiles\_unique\_region\_country\_sector\_year\_temporal\_level\_month\_hour\_profile\_version (UNIQUE) on (region\_country, sector, year, temporal\_level, month, hour, profile\_version)

## Entity Data (GLEIF)

### gleif\_legal\_entities

#### GLEIF Legal Entity Identifier record (Level 1 data).

Column	Type	Nullable	Description
id	bigserial	No	Primary key
lei	varchar(20)	No	UNIQUE. 20-character Legal Entity Identifier (ISO 17442)
legal_name	varchar(500)	No	Official legal name of the entity
legal_name_language	varchar(10)	No	ISO 639-1 language code for the legal name
other_names	jsonb	No	Alternative names [{name, language, type}]
legal_address_lines	jsonb	No	Legal address lines as JSON array
legal_address_city	varchar(200)	No	Legal address city
legal_address_region	varchar(20)	No	Legal address region (ISO 3166-2 subdivision code)

Column	Type	Nullable	Description
legal_address_country	varchar(2)	No	Legal address country (ISO 3166-1 alpha-2)
legal_address_postal_code	varchar(50)	No	Legal address postal code
headquarters_city	varchar(200)	No	Headquarters city
headquarters_country	varchar(2)	No	Headquarters country (ISO 3166-1 alpha-2)
jurisdiction	varchar(10)	No	Jurisdiction of legal formation (ISO 3166-1/2 code)
entity_category	varchar(50)	No	Entity category (GENERAL, FUND, BRANCH, etc.)
entity_sub_category	varchar(50)	No	Entity sub-category (e.g., STATE_GOVERNMENT, CENTRAL_GOVERNMENT)
legal_form_code	varchar(10)	No	ELF (Entity Legal Form) code from GLEIF
legal_form_other	varchar(200)	No	Free-text legal form description when no ELF code applies
entity_status	varchar(20)	No	Entity operating status (ACTIVE, INACTIVE)
entity_creation_date	timestamptz	Yes	Date the legal entity was created/incorporated
registration_status	varchar(20)	No	LEI registration status (ISSUED, LAPSED, RETIRED, etc.)
initial_registration_date	timestamptz	Yes	When the LEI was first registered
last_update_date	timestamptz	Yes	When the LEI record was last updated by the managing LOU
next_renewal_date	timestamptz	Yes	When the LEI registration is next due for renewal
managing_lou	varchar(20)	No	LEI of the Local Operating Unit managing this registration
bic	varchar(20)	Yes	Business Identifier Code (SWIFT/BIC) if mapped
registered_as	varchar(200)	No	Entity identifier in local business registry
successor_lei	varchar(20)	Yes	LEI of successor entity (for INACTIVE entities that merged/were acquired)
ingested_at	timestamptz	No	Timestamp when record was ingested into the platform
updated_at	timestamptz	No	Record last update timestamp

#### Indexes:

- gleif\_le\_legal\_name\_idx (INDEX) on (legal\_name)
- gleif\_le\_status\_idx (INDEX) on (entity\_status)
- gleif\_le\_jurisdiction\_idx (INDEX) on (jurisdiction)
- gleif\_le\_category\_idx (INDEX) on (entity\_category)
- gleif\_le\_last\_update\_idx (INDEX) on (last\_update\_date)
- gleif\_le\_hq\_country\_idx (INDEX) on (headquarters\_country)
- gleif\_le\_country\_status\_idx (INDEX) on (legal\_address\_country, entity\_status)
- gleif\_le\_reg\_status\_idx (INDEX) on (registration\_status)
- gleif\_le\_legal\_name\_trgm (GinIndex) on (legal\_name)

#### gleif\_relationships

##### GLEIF relationship record (Level 2 RR-CDF data).

Column	Type	Nullable	Description
id	bigserial	No	Primary key
start_lei	varchar(20)	No	LEI of the child entity (start node of the relationship)
end_lei	varchar(20)	No	LEI of the parent entity (end node of the relationship)
relationship_type	varchar(50)	No	Type of relationship between the two entities
relationship_status	varchar(20)	No	Current status of the relationship
start_date	timestamptz	Yes	Start date of the relationship period

Column	Type	Nullable	Description
end_date	timestampz	Yes	End date of the relationship period (null if ongoing)
ingested_at	timestampz	No	Timestamp when record was ingested into the platform
updated_at	timestampz	No	Record last update timestamp

**Indexes:**

- gleif\_rel\_start\_lei\_idx (INDEX) on (start\_lei)
- gleif\_rel\_end\_lei\_idx (INDEX) on (end\_lei)
- gleif\_rel\_type\_idx (INDEX) on (relationship\_type)
- gleif\_rel\_start\_type\_idx (INDEX) on (start\_lei, relationship\_type)
- gleif\_rel\_status\_idx (INDEX) on (relationship\_status)
- gleif\_relationships\_unique\_start\_lei\_end\_lei\_relationship\_type (UNIQUE) on (start\_lei, end\_lei, relationship\_type)

**gleif\_reporting\_exceptions**

**GLEIF reporting exception (Level 2 exceptions).**

Column	Type	Nullable	Description
id	bigserial	No	Primary key
lei	varchar(20)	No	LEI of the entity reporting the exception
exception_category	varchar(60)	No	Which parent relationship the exception applies to
exception_reason	varchar(100)	No	Reason the parent relationship cannot be reported
ingested_at	timestampz	No	Timestamp when record was ingested into the platform

**Indexes:**

- gleif\_exc\_lei\_idx (INDEX) on (lei)
- gleif\_exc\_reason\_idx (INDEX) on (exception\_reason)
- gleif\_reporting\_exceptions\_unique\_lei\_exception\_category (UNIQUE) on (lei, exception\_category)

**Carbon & Methane Budgets (GCP)**

**gcp\_gcpfossilemissionsbyfuel**

**Global fossil CO2 emissions by fuel type (GtCO2).**

Column	Type	Nullable	Description
id	bigserial	No	Primary key
year	integer	No	
coal_gtco2	double precision	No	
oil_gtco2	double precision	No	
gas_gtco2	double precision	No	
cement_gtco2	double precision	No	
flaring_gtco2	double precision	No	
other_gtco2	double precision	Yes	
total_gtco2	double precision	No	
budget_version	varchar(10)	No	
created_at	timestampz	No	
updated_at	timestampz	No	

**Indexes:**

- gcp\_gcpfossilmissionsbyfuel\_unique\_year\_budget\_version (UNIQUE) on (year, budget\_version)

### **gcp\_gcpglobalcarbonbudget**

**Annual global carbon budget components (GtCO2/yr).**

Column	Type	Nullable	Description
id	bigserial	No	Primary key
year	integer	No	
fossil_emissions_gtco2	double precision	No	E_fos: fossil fuel + cement
land_use_emissions_gtco2	double precision	No	E_luc: land-use change
atmospheric_growth_gtco2	double precision	No	G_atm: atmospheric CO2 growth
ocean_sink_gtco2	double precision	No	S_ocean: ocean CO2 uptake
land_sink_gtco2	double precision	No	S_land: terrestrial CO2 uptake
budget_imbalance_gtco2	double precision	No	B_im: unaccounted residual
cement_carbonation_gtco2	double precision	Yes	
budget_version	varchar(10)	No	e.g. '2024'
created_at	timestamptz	No	
updated_at	timestamptz	No	

#### **Indexes:**

- gcp\_gcpglob\_budget\_\_6419d0\_idx (INDEX) on (budget\_version, year)
- gcp\_gcpglobalcarbonbudget\_unique\_year\_budget\_version (UNIQUE) on (year, budget\_version)

### **gcp\_gcpmethanebudget**

**Global methane budget components (TgCH4/yr).**

Column	Type	Nullable	Description
id	bigserial	No	Primary key
year	integer	No	
fossil_fuels_tgch4	double precision	Yes	
agriculture_tgch4	double precision	Yes	
waste_tgch4	double precision	Yes	
biomass_burning_tgch4	double precision	Yes	
wetlands_tgch4	double precision	Yes	
other_natural_tgch4	double precision	Yes	
total_sources_tgch4	double precision	No	
total_sinks_tgch4	double precision	No	
atmospheric_growth_tgch4	double precision	No	
budget_version	varchar(10)	No	
created_at	timestamptz	No	
updated_at	timestamptz	No	

#### **Indexes:**

- gcp\_gcpmethanebudget\_unique\_year\_budget\_version (UNIQUE) on (year, budget\_version)

### **gcp\_gcpnationalemissions**

**National territorial + consumption CO2 emissions.**

Column	Type	Nullable	Description
id	bigserial	No	Primary key
country_code	varchar(3)	No	ISO 3166-1 alpha-3
country_name	varchar(200)	No	
year	integer	No	
territorial_mtco2	double precision	No	Production-based emissions (MtCO2)
consumption_mtco2	double precision	Yes	Consumption-based
transfer_mtco2	double precision	Yes	Trade adjustment
per_capita_tco2	double precision	Yes	
data_source_name	varchar(50)	No	CDIAC, GCP, etc.
budget_version	varchar(10)	No	
created_at	timestamptz	No	
updated_at	timestamptz	No	

#### Indexes:

- gcp\_gcpnati\_year\_56dd48\_idx (INDEX) on (year, country\_code)
- gcp\_gcpnationalemissions\_unique\_country\_code\_year\_budget\_version (UNIQUE) on (country\_code, year, budget\_version)

---

## Severe Weather (NOAA Storm Events)

### noaa\_storm\_events\_noaastormevent

#### NOAA Storm Events Database record.

Column	Type	Nullable	Description
id	bigserial	No	Primary key
latitude	numeric(10,6)	Yes	Latitude in decimal degrees
longitude	numeric(10,6)	Yes	Longitude in decimal degrees
geometry	geometry(Point,4326)	Yes	PostGIS Point geometry (lon, lat)
event_id	varchar(20)	No	Unique
episode_id	varchar(20)	No	
event_type	varchar(100)	No	
state	varchar(100)	No	
state_fips	varchar(10)	No	
cz_type	varchar(1)	No	C=County, Z=Zone, M=Marine
cz_fips	varchar(10)	No	
cz_name	varchar(200)	No	
begin_date	timestamptz	No	
end_date	timestamptz	Yes	
injuries_direct	integer	No	
injuries_indirect	integer	No	
deaths_direct	integer	No	
deaths_indirect	integer	No	
damage_property	double precision	Yes	Property damage in USD

Column	Type	Nullable	Description
damage_crops	double precision	Yes	Crop damage in USD
source	varchar(200)	No	
magnitude	double precision	Yes	
magnitude_type	varchar(10)	No	EG, ES, MS, MG
tor_f_scale	varchar(10)	No	Tornado F/EF scale
year	integer	No	
month	integer	No	
created_at	timestamptz	No	
updated_at	timestamptz	No	

#### Indexes:

- noaa\_storm\_\_event\_t\_8926fc\_idx (INDEX) on (event\_type, year)
- noaa\_storm\_\_state\_57e621\_idx (INDEX) on (state, year)
- noaa\_storm\_\_begin\_d\_7cfa47\_idx (INDEX) on (begin\_date)

---

## Naming Conventions

- **Table names:** {source}\_{entity} (e.g., openaq\_measurements)
- **Column names:** snake\_case (e.g., country\_code, measured\_at)
- **Foreign keys:** {related\_table}\_id (e.g., location\_id, sensor\_id)
- **Timestamps:** {action}\_at (e.g., created\_at, ingested\_at, measured\_at)
- **Coordinates:** latitude, longitude or lat, lon (numeric) + geometry (PostGIS)
- **Country codes:** country\_code (ISO-3) or iso3 / country\_iso3 (ISO-3)