1) Item-by-item check and overview (grouped by theme)

1) Emissions, pollution & trading data

Dataset	Source / Maintainer	Geography	Key fields	Access	Analytical value
EU Pollutant Release and Transfer Registry (E-PRTR)	European Commission / EEA	EU/ EEA	Facility emissions & transfers, pollutant types, industry, geo-coords	Web downloads; some APIs by year/ country	High cross-country comparability; facility-level hotspot analysis, trends, compliance checks.
EU Emissions Trading System (EU ETS) records	European Commission	EU	Allocation, transactions, registry accounts, compliance	Bulk files; registry search	Carbon asset dynamics & compliance risk; firm-level carbon exposure and event studies.
European Industrial Emissions Portal	EEA	EU	Permits, monitoring, compliance, volumes	Portal download	Complements E-PRTR with permit/monitoring context; regulation & operations risk.
PRTRs (Japan, Mexico, North America)	National environment agencies	Country-specific	PRTR-reported releases & transfers	National portals	Cross-country PRTR benchmarking; upstream/ downstream pollution risk in supply chains.
IPE Blue Map (China)	Institute of Public & Environmental Affairs	China	Point-level violators, fines, live monitoring	Website / partial API	Enforcement + social disclosure; supply-chain screening, violation alerts, regional risk maps.
EDGAR	EU JRC, partners	Global	National/grid GHG & pollutants inventories	Site downloads	Long-run, method-consistent; national/ regional trends and scenario baselines.

2) Energy assets & industry

Dataset	Source / Maintainer	Geography	Content	Access	Value
Global Energy Monitor (GEM)	Global Energy Monitor (NGO)	Global	Project-level attributes for power plants, pipelines, coal mines, etc.	Web / open sheets	Among the best for project detail; due diligence, supply risk, energy transition tracking.
Hydrogen facilities (GEM subset)	GEM	Global	Hydrogen production & infrastructure	Same as above	Transition readiness and policy impact assessment.
Open global coal & metal mine production (labelled “Open…”)	Open data / academic or NGO (needs maintainer confirmation)	Global	Mines, production, operators	Project page downloads	Mineral supply chains, local impacts, biodiversity risks; **confirm maintainer before use**.

3) Climate, weather & air quality (remote sensing / re-analysis)

Dataset	Source / Maintainer	Geography	Content	Access	Value
Copernicus CAMS Global Near-Real-Time	ECMWF / Copernicus	Global	Near-real-time PM2.5/PM10 & AQ fields	API / downloads	Event detection, exposure & health risk, ground-station cross-validation.
Sentinel-5 (e.g., S5P/TROPOMI)	ESA	Global	NO_2, SO_2, O_3, CH_4 columns/ concentrations	Scientific repositories	High spatio-temporal resolution; source attribution & transboundary transport.
ERA (e.g., ERA5)	ECMWF	Global	Reanalysis of wind, temp, precip, etc.	API / downloads	Long-term baselines; extremes, wind resource, climate-risk quantification.
Global Flood Database	Academic/industry collaboration (variant by release)	Global	Historical inundation extent, timing, area	Platform downloads	Zonation, (re)insurance pricing, asset resilience; **confirm version & curator**.

4) Land cover, forests & water

Dataset	Source / Maintainer	Geography	Content	Access	Value
Dynamic World	WRI × Google	Global	Near-real-time 10-class land cover	API / map	Urban expansion, forest/agri change, site screening & impact ID.
Global Forest Watch (GFW)	WRI coalition	Global	Forest loss/gain, fire alerts, protected areas	Portal / API / downloads	Ecological impact & deforestation risk in supply chains; compliance & negative screens.
Global Flood Database (repeat)	Same as above	Global	Flood history & zones	Same as above	Combine with land cover for vulnerability metrics.

5) Material flows & resources

Dataset	Source / Maintainer	Geography	Content	Access	Value
Global Material Flows Database	UNEP IRP + research partners	Global/ national	Extraction, consumption, footprints	Reports / data files	Circular-economy metrics, national material intensity, policy scenarios.

6) Aggregators & curated lists

Portal / List	Maintainer	Content	Access	Value
Resource Watch	WRI	300+ climate, air, forest, urban	Web/API	Fast discovery and map fusion; problem-led
“Awesome Data” lists (Agriculture,	GitHub community	Curated links to open datasets	GitHub	Flexible but variable quality; scoping & ideation.
Topic repositories: “Awesome Forest”,	Community / FAO	Forestry & water,	Official sites/APIs	Long-series agri-water baselines; join with climate/
Energy & electricity market data (incl. South	Market operators/	Prices, demand/ peak, units	Portals/ CSV/API	Power-market research, peak load & carbon-intensity

University library portals: Princeton,	University libraries	Curated environmental/	Index pages	Systematic discovery with quality gatekeeping.
US EPA Datasets	US EPA	Air/water/soil monitoring,	Data.gov / EPA	Authoritative US environmental & regulatory
EEA Datahub	European Environment	European environmental	Datahub	EU regional baselines & cross-country comparability.

2) Summary prioritisation by "value" for analytics

- **Baselines & trend models (long, comparable time series):** *EDGAR, ERA5, FAOSTAT, AquaStat, Global Material Flows*.
    - Value: consistent methodology; solid for regression, attribution, and policy scenarios.
- **Facility / project-level risk:** *E-PRTR, Industrial Emissions Portal, IPE Blue Map, GEM, GFW.*
    - Value: precise geocoding of hotspots and supply-chain nodes; due diligence, screening.
- **Near-real-time monitoring & events:** *CAMS, Sentinel-5, Global Flood Database, Dynamic World*.
    - Value: event detection, exposure/vulnerability, early-warning, parametric insurance.
- **Markets & carbon assets:** *EU ETS* (plus vetted power-market feeds).
    - Value: connects environmental drivers to financial risk and carbon-cost pass-through.
- **Discovery hubs:** *Resource Watch, university portals, Awesome lists*.
    - Value: rapid scoping; requires extra validation and version control.

3) Data governance & implementation tips (practical)

- **Source registry:** For each dataset store maintainer, version/release date, license, API/ endpoints, update cadence, and citation text.
- **Comparability controls:** Before joins, standardise **geographies, temporal granularity, and units**; keep conversion functions (e.g., kg→t; ppbv→μg/m^3 under conditions).
- **Versioning & lineage:** Use a **bronze/silver/gold** data-lake pattern; retain raw, cleaned, and feature-engineered layers with checksums and full transformation logs.
- **Quality checks:** Track **missingness, outliers, logical constraints** (e.g., emissions ≥ 0; mass balance), and **geospatial topology** validity.
- **Access & licensing:** Confirm commercial vs non-commercial terms; manage API keys via a secrets vault.

4) Items to confirm or strengthen

- **“Open global coal & metal mine production”** and **“Energy & electricity markets (incl. South Africa)”** are umbrella labels in your list; the *actual* primary maintainers vary. Lock down the authoritative sources per country/market before ingestion to preserve comparability and licensing clarity.

- **Global Flood Database:** Verify which edition and curator you will use (different academic/industry consortia exist) to keep methods and citations consistent.