Policy Brief: Checking What a Data Centre Is Actually Costing Your Community

A practical guide to the open dataset, for journalists and local advocates

Sinar Project PLT | Global South Alliance | For journalists and civil society | July 2026

What's now available

This project has independently measured 71 data centre buildings across Malaysia using satellite imagery, and published the results as an open dataset. For each building, you can look up its approximate location, physical size, estimated power demand, estimated water use, estimated carbon emissions, and, where a company has actually disclosed its real capacity, that disclosed figure too. You can also see how confident the measurement is, and whether the facility was excluded from our totals (and why).

> *You no longer have to take a company's press release, or its silence, as the only available answer.*

How to look up a specific facility

1. Open GSA_Master_Facility_Database.csv (a spreadsheet file) in Excel, Google Sheets, or any spreadsheet program.
2. Use the Tenant_Names column to search for the company or facility name you're interested in. Some buildings house more than one company; all names sharing one building are listed together, separated by semicolons.
3. Check the Status column first. If it says EXCLUDED, we could not confidently measure or confirm the location of that facility, and none of the resource figures for that row should be used or cited.
4. Check the Power_Source column. DISCLOSED means the power figure came directly from the operator; MODELED means it is our estimate from the building's physical size, and should be treated as an order-of-magnitude figure, not a precise number.
5. Check the CV_Confidence and CV_Flags columns for any quality warnings on that specific row before quoting a figure publicly.

Key columns explained

Column	What it means
Building_ID	A unique identifier for the physical building (its coordinates).
Extracted_Area_m2	The building's footprint area, as measured from the satellite image, in square metres.
Est_Facility_Power_kW	Estimated total power draw, in kilowatts. Divide by 1,000 for megawatts.
Est_Annual_Carbon_tCO2e	Estimated annual carbon emissions, in tonnes of CO2 equivalent.
Est_Water_Optimized_L_yr / Est_Water_Regional_L_yr	Estimated annual water use, under a best-case and a more realistic tropical-climate scenario respectively. Use the regional figure for a more conservative (higher) estimate.

QA_Notes	Any specific data-quality issue we found for that row during independent verification.

What you can and can't say with this data

- You can say: "an independent satellite-based estimate suggests this facility's power demand is approximately X MW," citing this project as the source.
- You should not say: this is the facility's actual, confirmed power bill or usage. It is an estimate, and we have shown that estimates like this can be wrong by a large margin for any single facility, even though they are useful in aggregate.
- For deforestation figures, you can say a facility sits within a corridor that has lost a certain amount of tree cover since 2001. You should not say the facility itself is confirmed to have caused that specific clearance; the corridor-level figure includes surrounding land, not just the facility footprint.
- Always check the Status and CV_Flags columns before publishing a figure for a specific named facility; a small number of rows are known to have location or detection issues and are explicitly marked as such.

Where to get help

The full technical methodology, the Policy Audit Report, and the underlying pipeline code are all published alongside this dataset. If you are working on a specific story and want to verify a finding, or need help interpreting a particular row, the project team is reachable via the contact details on the main public report.