Methodology Note: Database of Chinese Overseas Renewable Energy Investment (2022 – 2025)

1. Introduction and Purpose

The Database of Chinese Overseas Renewable Energy was developed to provide a structured and verifiable empirical basis for analysing Chinese enterprises' participation in overseas renewable and low-carbon energy projects. Public information on China's overseas energy activities is widely disseminated across corporate announcements, media reports, and secondary sources. As a result, project details are often reported with inconsistent conventions, geographic references, descriptions of participating organisations, and financial information. These features limit the comparability of existing information and constrain its usefulness for policy analysis, market assessment, and cross-country comparison.

The database responds to these challenges by consolidating publicly available information into a project-level dataset governed by unified definitions, coding rules, and verification procedures. Its purpose is not to replace official statistics or corporate disclosures, but to systematise the public information environment in which policy and investment decisions are made. By making both verified information and remaining uncertainties explicit, the database provides an updated and transparent foundation for descriptive analysis and supports in-depth analysis on portfolio exposure, climate-related risk assessment, and policy evaluation.

2. Scope, Definitions, and Coverage

The database covers overseas renewable and low-carbon energy projects in economies associated with the Belt and Road Initiative from 2022 to 2025. It includes project activities in solar photovoltaic power, wind power, hydropower, biomass, geothermal, and waste heat utilisation; integrated energy systems; and energy storage facilities. Integrated energy projects are configurations that combine multiple technologies, such as wind and solar generation, renewable generation with storage, or renewable generation with 'green' hydrogen or ammonia-related facilities.

Fossil fuel power generation, including coal, oil, and gas-fired power plants, is explicitly excluded from the database. In addition, energy-efficiency retrofits, consultancy services, transmission infrastructure, and grid projects are excluded. These activities are excluded unless they are clearly reported as an integral component of a renewable electricity generation project and share the same project boundary. Standalone efficiency upgrades, advisory services, or grid-related investments are therefore not treated as eligible entries in the database.

The project stage is treated as a descriptive attribute rather than a strict inclusion criterion. Projects reported at early, non-binding stages, such as memoranda of understanding, preliminary cooperation intentions, or due diligence phases, are retained in the database and explicitly flagged. These entries are not interpreted as realised investment outcomes, but their inclusion allows the database to reflect the evolving project pipeline and the information context in which

expectations and decisions are formed. This approach helps reduce survivorship bias and supports analysis of both realised and unrealised project trajectories.

3. Unit of Analysis and Database Structure

The primary unit of analysis in the database is the project activity, which is defined as a discrete renewable or low-carbon energy development with a specific location, technology configuration, legal framework, contractual value, and operational boundary, regardless of the number of Chinese (or foreign) enterprises involved in its development or financing. This definition enables the analysis of installed capacity, geographic distribution, and technology composition at a consistent level.

Meanwhile, because many project activities involve multiple Chinese enterprises in different roles, the database employs a project-based, actor-disaggregated coding approach to capture this complex structure. When more than one Chinese enterprise or entity participates in a single project activity, each enterprise is recorded in a separate row, while all core project attributes remain identical across those entries. This structure preserves project-level aggregation while enabling enterprise-level analysis regarding their participation modes, market share, major strengths, and financial exposure.

To operationalise this design, the database entries are organised into three main groups, which are summarised in Table 1. The first group, corresponding to columns A to G, captures core attributes of the project activity, including project identification, location, energy technology, and installed capacity. These fields describe the physical characteristics of the project asset and form the basis for aggregation by geography, technology type, and deployment scale.

The second group of data, corresponding to columns H to Q, records information on project ownership, participating enterprises and their roles, the timeframe of project construction commission, and operation. This group distinguishes between different forms of corporate participation, such as engineering and construction, equipment supply, and equity investment, which somehow illustrate the organisational context that helps to interpret how Chinese enterprises' preferences and strategies in a given market.

The third group of data, corresponding to columns R to W, captures publicly disclosed financial information, including total project investment, financing timelines, the involvement of Chinese financial institutions, and financing structures where explicitly reported. These fields are designed to support analysis of funding structures and transactional patterns, while applying conservative coding rules when public information is incomplete, ambiguous, or contradictory.

Together, these three-part data structures separate the physical characteristics of renewable energy projects from corporate involvement and financing arrangements, allowing each dimension to be analysed independently while remaining interlinked at the project activity level. Table 1 summarises the workflow used to construct the database, transforming dispersed public information into a consistent project-level dataset.

Field group	Column range	Description of fields	Analytical purpose
Physical information	A – G	Project name, host country and ISO country code, subnational location where available, energy technology type, installed capacity and unit	Forms the basis for aggregation and comparison by geography, technology type, and deployment scale, and supports analysis of spatial and technological distribution of projects
Corporate Information	H – Q	Project owner or developer, Chinese participating enterprise(s), identification of parent – subsidiary relationships, mode of participation, and construction and commissioning timelines	Enables analysis of how Chinese enterprises participate in overseas renewable energy projects, including implementation roles, ownership structures, and project progress
Financial Information	R – W	Total project investment, subcontract value where applicable, investment or financial close date, involvement of Chinese financial institutions, and financing structure where explicitly disclosed	Supports analysis of financing patterns, the role of Chinese financial institutions, and potential financial exposure associated with overseas renewable energy projects

Table 1: Field Structure of the Database of Chinese Overseas Renewable Energy

4. Data Sources and Collection Strategy

Data processing begins with the systematic screening of Chinese-language public information channels using predefined keywords related to renewable energy to identify overseas project announcements. Identified items are then converted into preliminary project entries and consolidated at the project activity level through deduplication and reconciliation of duplicate mentions across sources and time periods.
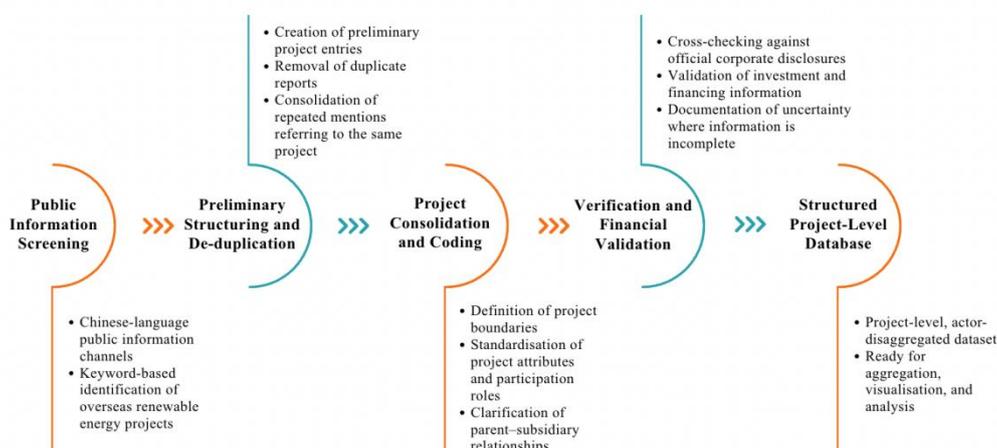

Figure 1. Data Collection and Processing Workflow

Following initial consolidation, project information is coded in a standardised manner. Project boundaries are confirmed using key physical descriptors, including location, technology type, and installed capacity, and participating enterprises are identified with attention to parent – subsidiary relationships where relevant. Each Chinese enterprise has demonstrated a participatory role, based on verifiable evidence.

The final stage focuses on verification and financial validation. Key project attributes are cross-checked against disclosures and formal announcements in official corporate or mainstream media, whenever available, and any unspecified information is also documented. Information is formally registered only when it is backed by public evidence and left blank when it is not. Through this workflow, the database maintains transparency on both the sources used and the confidence level of recorded information.

We constructed the dataset using a two-stage automated pipeline. First, a custom-built Python crawler scraped publicly available news articles and applied keyword filters (e.g., "solar," "wind," "MW") to identify relevant energy projects. Second, we employed a large language model (Google Gemini) to process the unstructured text. The AI acted as a supporting analyst, reading each article to extract specific variables, such as installed capacity, location, and developers, which were then compiled into a standardised database structure. This approach combined the efficiency of automated collection with the semantic understanding capabilities of AI to ensure data accuracy.

5. Geographic Coding and Project Consolidation

Geographic location is treated as a core methodological requirement of the database. Where public information is available, projects are coded at the subnational level, typically corresponding to the provincial administrative unit. Subnational coding is based on explicit references in corporate disclosures or credible reporting. Where only vague information is available, the subnational field is left blank rather than inferred.

When specific localities, ports, or towns are reported, these are mapped to the corresponding province-level units using verifiable geographic references. This conservative approach avoids introducing false precision and ensures that future integration with other GIP datasets, such as hazard exposure or stress-testing frameworks, rests on defensible geographic information.

Project consolidation decisions are made prior to geographic coding to avoid fragmenting a single project into multiple entries due to minor variations in reported location or naming. The consolidation logic prioritises physical and operational characteristics over reporting-source differences, ensuring that geographic aggregation reflects the actual project distribution rather than reporting patterns. The proofreading process can ensure the consistency of project names. When project names across media are inconsistent, the inconsistent information will be corrected to ensure that the same project doesn't have multiple names.

6. Actor Identification and Participation Roles

For each project activity, the database identifies the project owner or developer, which may be a

local entity, an international developer, or a project company or SPV (Special Purpose Vehicle). Chinese participating enterprises are then identified, and where possible, parent - subsidiary relationships are clarified to reduce misclassification arising from complex corporate group structures.

Enterprise participation is classified according to a standardised taxonomy that distinguishes engineering, procurement, and construction (EPC) activities from equipment supply, design services, equity investment, debt financing, and concession-based arrangements such as BOT or BOOT. Long-term ownership structures and offtake arrangements are recorded where explicitly disclosed. Where a single enterprise performs multiple functions within the same project, combined participation categories are used in a controlled manner to reflect substantive involvement without excessive fragmentation.

This role-based classification supports the analysis of corporations' comparative advantages across the value chain, market exposure, and risk appetite. Different participation patterns also imply distinct revenue structures and portfolios, thereby enabling in-depth analysis.

7. Financial Information and Verification Logic

Financial information is handled with caution due to uneven and limited disclosure across countries, enterprises, and project stages. Where total project investment figures are reported, values are standardised to USD million using documented currency conversion conventions. The database distinguishes total project investment from subcontract or EPC values when such information is disclosed, recognising that subcontract or EPC amounts represent partial transactional exposure.

Financing information is recorded only when explicitly disclosed by public evidence. Where financiers are identified, the database records whether Chinese financial institutions are involved and lists the relevant institutions. Financing mode categories, such as project finance, export credit facilities, syndicated loans, equity finance, or mergers and acquisitions loans, are assigned only when clearly indicated by the underlying sources. In the absence of sufficient evidence, these fields are left blank rather than inferred.

Verification follows a minimum two-source cross-checking principle that prioritises official disclosure. Key project attributes are verified through corporate websites, formal announcements, and other official public channels, when available, and then cross-checked against reputable secondary sources. When information conflicts across sources, official corporate disclosure takes precedence. When discrepancies cannot be resolved using publicly available information, the database records uncertainty via notes rather than imposing a definitive value.

8. Updates, Revisions, and Limitations

The database is updated regularly, with quarterly analyses based on the latest versions. Updates may include newly identified projects, revisions to existing entries, status changes, and additional verification information as new disclosures emerge. Historical entries may be revised if more accurate or authoritative information becomes available, and revisions are documented to maintain transparency.

In any database, delays and inconsistencies in information can occur. These limitations are addressed through conservative coding principles, explicit verification markers, and transparent documentation rather than speculative data compilation.

Accordingly, the Database of Chinese Overseas Renewable Energy should be understood as a curated and evolving representation of publicly verifiable overseas renewable and low-carbon energy projects involving Chinese enterprises from 2022 to 2025. It is intended to support systematic descriptive analysis and to provide a stable foundation for future policy, financial, and climate-related research.

Definitions:

BOT (Build-Operate-Transfer): A project delivery model where a private entity receives a concession from the public sector to finance, design, and build a project. The entity then operates the facility for a specific period to recover its investment before transferring ownership back to the government or local authority.

BOOT (Build-Own-Operate-Transfer): A variation of BOT where the private entity explicitly owns the project during the concession period. After the agreed-upon period of operation and ownership, the asset is transferred to the public authority.

EPC (Engineering, Procurement, and Construction): Comprehensive agreements under which a single contractor is responsible for the entire project lifecycle, from engineering through procurement to construction.

Equipment Supply: The provision of core technical components and machinery.

Equity Investment: Direct ownership through shareholding in the project asset.

Please cite as:

International Institution of Green Finance. (2026). Methodology Note: Database of Chinese Overseas Renewable Energy Investment (2022–2025): Methodology Note. Available at: https://iigf.cufe.edu.cn/info/1014/11031.htm

Contact:

For inquiries, please contact iigf-china@iigf.com.cn