SUBMITTING DATA FOR USE IN THE RENEWABLE ENERGY DATA EXPLORER

Data Standard and Guidance to Third-Party Data Providers

Introduction

This document seeks to provide a standard set of specifications and guidance to ensure an efficient process for sharing and integrating data for use in Renewable Energy (RE) Data Explorer, a web-based application developed and maintained by the United States Department of Energy's National Renewable Energy Laboratory (NREL). The RE Data Explorer contains country-specific geospatial data that describe RE resource distribution as well as other geographic features such as land use, terrain, infrastructure, and energy sector and/or national development indicators. NREL welcomes and relies upon third-party data providers to suggest new and/or updated data for inclusion in the RE Data Explorer.[1] To support data sharing, NREL has developed these guidelines for formatting, describing, and recording data for inclusion in the RE Data Explorer.

Purpose and Scope

The purpose of this document is to provide detailed specifications for geospatial datasets that will be uploaded to the RE Data Explorer for visualization, analysis, and/or download. When NREL receives new or updated datasets for inclusion RE Data Explorer, laboratory staff process these datasets into layers in the RE Data Explorer. This document intends to provide a common set of guidance to third-party data providers to ensure NREL can efficiently process and upload new data into the RE Data Explorer.

The RE Data Explorer includes data that characterize the distribution RE resources (such as solar, wind, biomass, and geothermal), environmental and terrain features (such as land cover and elevation), infrastructure (such as transmission lines, substations, and power plants), and administrative and other factors (such as jurisdictional boundaries, protected areas, population distribution, and special economic development zones). Because of the breadth of information in the application, the data in RE Data Explorer come from a variety of providers, each of which may have its own methods and standards for producing data. This document does not attempt to provide any information about how to develop or procure any particular type of data, such as land cover, RE resource, transmission infrastructure, etc. Instead, it focuses on ensuring existing datasets of any type are provided in a format and with sufficient accompanying information so as to be compatible with the RE Data Explorer.

Applicability

The specifications in this document are applicable to any dataset(s) a third-party data provider wishes to include in the RE Data Explorer application. As of the writing of this document, RE Data Explorer applications are available for Afghanistan, Bangladesh, Ghana, India, Indonesia, Kazakhstan, Kenya, the Lower Mekong region of Asia (including Myanmar, Cambodia,

[1] In this document, "third-party" refers to entities other than NREL.

Lao PDR, Thailand, and Vietnam), Mexico, Nepal, Pakistan, and the Philippines. These specifications will also apply to any new country applications included in the RE Data Explorer in the future.

Background on the RE Data Explorer

The RE Data Explorer is a no-cost, web-based application that uses geospatial and spatiotemporal RE data to visualize, execute, and support analysis of RE potential under various user-defined scenarios. Developed by NREL with support from United States Agency for International Development (USAID), the application enables users to visually explore RE resources such as solar and wind relative to other geographic features (e.g., infrastructure, terrain, and environmental and administrative areas) that might impact the development of these resources. It also provides the capability to quantitatively estimate resource and technical potential of some RE resources based on user-defined inputs for technology type and land area exclusions. For example, RE Data Explorer users can specify whether to filter protected areas, certain types of land uses or land covers, and lands with steep slopes from the estimation of the technical potential for solar or wind energy generation in a country of province. Though this type of analysis can also be conducted by analysts with other tools, the RE Data Explorer wraps complex spatial analysis techniques into an easy-to-use interface targeted at non-specialists such as policymakers, energy planners, and RE project developers, among others.

The RE Data Explorer, tutorials, and additional documentation are accessible at https://www.re-explorer.org.

Specifications

Overview

Box 1 provides a summary checklist of requirements for third-party data submissions for inclusion in the RE Data Explorer. The sections that follow provide additional detail and guidance on each item in this list.

Box 1. Checklist of Requirements for Submitting a Dataset for Inclusion in the RE Data Explorer

- ✓ The dataset is geospatial.
- ✓ The dataset is in a preferred file format.
- ✓ The dataset is accompanied by documentation in the form of a metadata text file that includes (at a minimum) descriptions of attribute data and the projection or coordinate system in which the data are provided.
- ✓ The dataset provider has documented the dataset with an entry in NREL's Data Catalog.
- ✓ The dataset and its associated metadata are written in English, using the Latin (Roman) alphabet.
- ✓ If the dataset is referenced to administrative boundaries (country, province, district, etc.), the data provider has ensured it uses the same naming conventions and spelling as applied in the boundary layer(s) in use for the same country within the RE Data Explorer. Unless otherwise specified in RE Data Explorer, the default naming convention is that used by the Global Administrative Areas, or GADM, database.
- ✓ The data provider has supplied written permission to use the dataset in the RE Data Explorer.
- ✓ The data provider has performed a quality check on the dataset.

Geospatial Data

All data in the RE Data Explorer must be geospatial data. Simply put, geospatial data have a geographic or locational component; they associate a particular record or variable of interest (e.g., wind speed, electrification rate) with a location on

Earth.[2] Thus, any data submitted by third-parties for inclusion in the RE Data Explorer must include the following fundamental components:

- *Spatial reference*: the locational component of the data. Spatial reference can be expressed in a variety of ways, for instance, coordinates (e.g., latitude and longitude or other unit depending on the spatial projection used to create the data) or administrative area (e.g., country, province, district, and/or other area with established spatial boundaries). Spatially-referenced transmission line data, for instance, would include data on the actual location of the transmission lines (e.g., the latitude and longitude of its start and end nodes).
- *Attribute data*: additional information about the spatial features. For example, attribute data for transmission lines might include the name of the line and its voltage, operational year, and length.

Appendix A provides a geospatial data "wish list" for the Renewable Energy Data Explorer. Spatial data can be provided in a variety of formats, as discussed in the following section. In general, NREL endeavors to utilize data that has already been created by data providers and does not provide prescriptive templates, as each country will have its own priorities for data creation, and templates might discourage the provision of complete datasets. To view examples of different types of geospatial data, NREL advises data providers to visit the RE Data Explorer and view the data currently in the application. Many of the data in the RE Data Explorer can be downloaded in several formats (including csv, shapefile, KML, and GeoJSON formats). The data files available in RE Data Explorer can provide references and examples for different types and formats of data.

The tables below provide tabular examples intended to highlight difference between geospatial and non-geospatial data to illustrate the types of information that are and are not appropriate for inclusion in the RE Data Explorer. ***These examples are purely illustrative and should not be treated as templates for transmission or electricity statistical data, since geospatial data can be provided in a variety of formats (discussed in the next section) and units and can include a variety of attributes not shown in these examples.*** Tables 1 and 2 provide examples of transmission line data that are not and that are geospatial, respectively. Table 2 is considered geospatial data because it includes the latitude and longitude of start and end nodes of each transmission line. Similarly, Table 3 and Table 4 provide examples of electricity statistical data that are not and that are geospatial, respectively. Table 4 is considered geospatial data because each attribute is referenced to a specific province. In these of these cases, the data are purely illustrative and do not represent an actual power system or country.

Example 1: Transmission Line Data

Table 1. Example of Transmission Data That Are Not Geospatial (Not Appropriate for the RE Data Explorer)

Transmission Line ID	Voltage (kV)	Line Length (km)	Operation Year
TX-1	115	42	2012
TX-2	230	107	2015
TX-3	115	68	1998

Table 2. Example of Transmission Data That Are Geospatial (Appropriate for the RE Data Explorer)

Transmission Line ID	Voltage (kV)	Line Length (km)	Operation Year	Start Node (Lat, Long)	End Node (Lat, Long)
TX-1	115	42	2012	(-5.724807, -124.143573)	(-8.515334, -124.445696)

[2] Geospatial data are often created in a geographic information system (GIS) environment, which allows for the ingestion and analysis of data sourced through different methods, such as remote sensing and field survey. Data created through GIS processes share key aspects such as spatial location and projection. The data layers in the RE Data Explorer were produced through GIS processes.

TX-2	230	107	2015	(-8.515334, -124.445696)	(-8.384929, -124.500628)
TX-3	230	250	2022	(-8.384929, -124.500628)	(-8.450137, - 124.522601)

Example 2: Electricity Statistical Data

Table 3. Example of Electricity Statistical Data That Are Not Geospatial (Not Appropriate for the RE Data Explorer)

Electricity Generation Statistics for Country X		
Installed Electricity Generation Capacity (GW)	**Electricity Generation (TWh)**	**Electricity Demand (GWh)**
230	813	813

Table 4. Example of Electricity Statistical Data That Are Geospatial (Appropriate for the RE Data Explorer)

Province	Installed Electricity Generation Capacity (GW)	Electricity Generation (TWh)	Electricity Demand (GWh)
Province A	43	159	179
Province B	6	84	242
Province C	91	228	145
Province D	84	220	112
Province E	6	122	135

File Formats

Preferred File Formats

NREL can process geospatial data provided in variety of formats for incorporation into the RE Data Explorer. NREL prefers to receive data in some formats over others because they are easier to process and/or because they provide more accurate and usable data for use in the RE Data Explorer. Table 5 provides a list and description of NREL's preferred formats for data files provided by third-parties for use in the RE Data Explorer, in order of most to least preferred. The table also identifies special considerations related to particular file formats.

In general, NREL accepts both vector and raster GIS data but prefers vector data due to its ease of use.

Table 5. Preferred file formats for data to be incorporated into the RE Data Explorer, in order of most to least preferred

FILE TYPE	DESCRIPTION	SPECIAL CONSIDERATIONS
POSTGRESQL (POSTGRES) SCHEMA	PostgreSQL (Postgres) is an open-source object-relational database system. NREL uses this system to process and organize data layers for the RE Data Explorer. If third-party data providers also use this system, the schema (i.e., collection of data tables)	Providers submitting data in this format should provide all tables that are part of the particular schema they wish to upload to the RE Data Explorer.

FILE TYPE	DESCRIPTION	SPECIAL CONSIDERATIONS
	can be exported from the third-party's database and easily imported to NREL's database.	
SHAPEFILE	A geospatial vector data format developed by ESRI and operable by other GIS software. A "shapefile" consists of a collection of files with a common filename prefix, stored in the same directory.	Providers submitting in this format should include *at a minimum* the files with the following extensions: .shp, .shx, .dbf, and .prj.
TAB	A geospatial vector format developed by MapInfo Corporation and operable by other GIS software. Like shapefile, several file extensions with the same prefix are needed to open TAB datasets.	Providers submitting in this format should include *at a minimum* the files with the following extensions: .tab, .dat, .map, and .id.
CSV, TSV, AND ANY OTHER TABULAR FILE	Files that store tabular data as plain text. Tabular files usually include a series of named columns and associated values.	Data created or saved in Microsoft Excel can be saved as a .csv file. Data providers should ensure tabular data include a spatial component, as these file formats are not inherently spatial.
GEOJSON	An open standard format for representing geographical features, based on JavaScript Object Notation (JSON). Unlike Shapefile and TAB, GEOJSON formats include only a single file ending in the extension .geojson.	None.
GEOTIFF OR ARCGRID	File formats for gridded geospatial data.[3]	Providers who submit in these formats should ensure metadata explicitly define the attributes, projection, and units of the data, as this information cannot be inferred from the file itself.
KML, GML	Spatial variants of markup-like formats.[4]	None
XML, JSON , AND OTHER NON-TABULAR FILES	Files that do not follow a tabular structure but still present a logical internal structure.	Data providers should ensure that data include a spatial component, as these file formats are not inherently spatial.

[3] Gridded data are a type of continuous geospatial data, which show how a certain attribute (e.g. elevation, slope, and wind speed) varies over a surface. Gridded data are associated with a spatial resolution that describes the size of the grid cell. For example, 1 x 1 km resolution solar data assign a solar resource value to each 1km by 1km cell over a prescribed geographic area. In contrast, discrete features (e.g., roads, parks, transmission lines, or ground measurement stations) have a definable beginning and end and are not represented on a continuous grid (Source: Cox, Sadie, Andrea Watson, and Anthony Lopez. Forthcoming. "From Data to Renewable Energy Decisions: An Introductory Guide." Golden, CO: National Renewable Energy Laboratory.)

[4] Markup formats are written in a markup programming language. For more information, please refer to the KML and GML standards provided by the Open Geospatial Consortium (http://www.opengeospatial.org/docs/is).

FILE TYPE	DESCRIPTION	SPECIAL CONSIDERATIONS
WEB MAP SERVICES (WMS) AND WEB FEATURE SERVICES (WFS)	Web services provide access to geospatial data, either directly (WFS) or by visualizing the data as maps (WMS), independently of the format of the base dataset. WFS and WMS require access to the server on which the data are stored.	NREL can incorporate this format into the RE Data Explorer for *visualization only*. Data in this format cannot be used in analysis (e.g., dynamic technical potential) or downloaded from RE Data Explorer.

File Formats to Avoid

Some file formats are cost-prohibitive for NREL to process into data layers in the RE Data Explorer, and in some cases, processing these files may yield less accurate information than would be achieved through the use of other file types. NREL requests third-parties to avoid providing data in the following formats:

- Image files, including but not limited to JPEG, TIFF, and PNG
- Satellite imagery data
- Portable document format (PDF)
- PowerPoint and Word documents
- Database formats other than PostgreSQL
- Others not listed in Table 5

In many cases, data in one of the formats above can be extracted or reconfigured into an alternative, preferred format. For example, tabular data provided in PDF, PowerPoint, or Word formats can be transcribed into a CSV or other tabular file format. Maps and other data provided in an image file are frequently based on an underlying data source that can be exported in a preferred file format. Finally, for database formats other than PostgreSQL (e.g., MySQL), individual tables can be exported into CSV files and provided to NREL in that format.

In rare cases, data may *only* be available in one of the non-preferred formats listed above. NREL will work with data providers to handle such instances on a case-by-case basis, depending on the importance of the data to a particular country's RE Data Explorer objectives, and the availability of resources for NREL to undertake more intensive data processing efforts. NREL requests data providers to make reasonable efforts to provide data in a preferred format and can provide technical guidance before exploring this alternative.

Metadata (Documentation)

Data providers must include metadata that provide documentation of any data they are submitting for inclusion in the RE Data Explorer. Metadata refers to data and information that describe other data; more simply, metadata provide documentation of the data itself. Metadata are crucial to understanding and interpreting data. Data providers should include with their data submission **a separate text file** (e.g., XML or TXT extensions or Microsoft Word) that provides metadata. This file is sometimes referred to as a "data dictionary." NREL can provide examples of metadata files upon request; however, NREL does not require metadata to be in any particular format.

This section describes some of the most important metadata for third-parties to provide along with their datasets. Appendix B provides a template for metadata submissions that include these items.

Attribute Data Description

Metadata should include a qualitative definition of each variable or type of value included in the dataset. For example, if a data file is organized into a series of columns, the metadata should provide a brief, qualitative description of the information included in each column. Metadata for attribute data should include, but are not limited to:

- A brief description of the what the value in a particular column or file represents (e.g., “Annual average wind speed at 100-meter hub height”)
- The units of the value in a particular column or file (e.g., “m/s”) and any scalar multipliers applied to the data (e.g., “10^6 Btu”). NREL can accept data in any unit, so long as the metadata specify the unit of the data.
- Definition of any acronyms used in the data file

Spatial Reference Information

Metadata files should include a qualitative description the spatial reference system associated with the data. For GIS data, the metadata should define the projection of the data.[5] NREL can accept data in any projection, so long as the metadata specify the projection.

Occasionally, datasets provided in formats that are not spatial, such as tabular formats, might use an (x, y) coordinate system other than latitude and longitude (e.g. meters or feet) to define the location of one or more attributes. In such cases, NREL encourages data providers to transform their data into latitude/longitude coordinates. If such transformation is not feasible, the metadata should provide sufficient information about the coordinate system for the data so that NREL can perform this transformation.

Documentation in the Data Catalog

To support metadata documentation, NREL requires third-party data providers to document their data submission in NREL’s Data Catalog. The Data Catalog serves as a central location for research and discovery of the RE resource and GIS data that power the RE Data Explorer. The Data Catalog provides a platform for researchers to contribute additional data and metadata in support of the RE Explorer project. The Data Catalog’s functions include:

1. Allowing users to identify potentially useful datasets based on certain identifiers (e.g., author, title, keyword, date of publication);

2. Showing what datasets are available and enabling co-locating based on author, subject, geographic coverage, etc.

3. Assisting in the choice of an appropriate dataset by providing information on evaluating objectives such as spatial and temporal resolution and/or licensing restrictions.

The Data Catalog is primarily a metadata repository; while users can attach publicly-available datasets to submission, its main purpose is to provide a platform to document the existence of both public and proprietary datasets so that potential data users can efficiently identify data availability and, if necessary, acquire relevant data through the appropriate channels. Submissions to the Data Catalog capture the following information about datasets:

Qualitative description	An explanation of the information included in the dataset, any details the data provider wishes to share about how the dataset was developed, and any details the data provider wishes to share about how the dataset should be used.
Data source (and, if different, data author)	Organization(s) that creates, maintains, and/or disseminates the dataset.

[5] Projection refers to the system used to represent the three-dimensional surface of the Earth on a 2-dimensional plane or map. Common projections include WGS84, Universal Transverse Mercator, and Lambert Conformal Conic.

Geographic extent	The administrative area that the dataset covers (e.g., Lower Mekong sub-region; Thailand; Thanh Hoa Province).
Spatial resolution	The administrative level at which the data are enumerated, or, alternatively, the number of data points within a given area (e.g., district-level data; or 1 x 1 km grid cells over a continuous surface or discrete sites).
Temporal range	The total period of record observed in the data (e.g., one year, multiple years)
Temporal resolution	The time steps taken in the dataset (e.g., annual; monthly; hourly).
Vintage	The year in which the dataset was published.
Format	The file type in which the dataset is available (e.g., tabular, shapefile)
Accessibility	Any limitations (such as licenses) on who may access or utilize this dataset, and/or if there is a cost associated with the dataset (e.g., publicly available; available for purchase; for official use only)
Language	The language(s) in which the data is available.

This information can also be included in the metadata text file. NREL will use this information to describe the dataset in the RE Data Explorer.

The Data Catalog and example entries are accessible at https://data.re-explorer.org.

Language

At this time, the RE Data Explorer supports only data written in Latin (Roman) characters. Other alphabets, special characters, and diacritical marks should be replaced with their Latin character translations. For example, the Vietnamese province “Đắk Lắk” should be written as “Dak Lak.”

NREL also requests an English language copy of all metadata documentation, though data providers can supply metadata in additional languages if they so desire.

Administrative Boundary Names

If a dataset references administrative boundaries such as country, province, district, etc., data providers should use the naming and spelling convention applied for the same boundaries in the RE Data Explorer. Data providers can find this information in the RE Data Explorer.[6] Additionally, NREL can provide a list of administrative boundary names for any particular country upon request.

[6] To view administrative boundary layers, navigate to https://www.re-explorer.org, click on the country or region of interest, and open the “Administrative” section of the layer tree. Administrative boundary layers might include country, province, district, and/or other jurisdictional areas defined in use within a particular country. RE Data Explorer users can find the data source for each administrative boundary data layer by clicking on the metadata (question mark) icon associated with the layer.

NREL welcomes data providers to submit official administrative boundary data defined by a country's national government, if available. The data provider must translate country-specific administrative boundary data into Latin (Roman) characters if its original format is in a different alphabet.

For countries for which official, country-specific administrative boundary source has not been provided, NREL uses administrative boundary data and naming conventions from the GADM database as a default.[7] In these cases, NREL requests data providers to ensure their datasets use the same naming and spelling conventions for administrative boundaries as are applied in GADM. GADM is a high-resolution spatial database with global coverage of the world's administrative boundaries. It is publicly-available and may be used for free for academic or other non-commercial use. The administrative boundary data in GADM is based on the contributions of numerous individuals, academic organizations, public agencies, and other entities and is periodically updated. While the GADM dataset it not necessarily formally endorsed by government officials from any given country or intergovernmental organization, NREL uses this dataset as a default because it is well-utilized around the world, and many available datasets are already normalized to the GADM administrative boundaries.

Permissions

Data submissions must be approved by the data owner for dissemination via the RE Data Explorer. The RE Data Explorer application primarily includes only those datasets that can be shared publicly. The RE Data Explorer enables two levels of data access:

- ***Visualization only***: Data will be included as a layer that any user may view within the RE Data Explorer's web-based interface. Visualization-only data are not available for download from the RE Data Explorer. If requested by the data provider, NREL can provide a link to the data provider's website so that users can download data directly from the data source.
- ***Visualization and download:*** Download functionality can be enabled for distributable public datasets. In this case, data layers can be downloaded in a variety of formats (shapefile, CSV, KML, GeoJSON) directly from the RE Data Explorer interface. This option enables easier access to the data within the application.

When submitting data for incorporation into the RE Data Explorer, data providers must include written permission (in the form of email or scanned letter) from the data owner to publish the data in the RE Data Explorer either for visualization only, or for both visualization and download. Appendix C providers a template letter of permission that data providers may adapt for this purpose.

Quality Assurance and Control

NREL does not independently validate data provided by third-parties, though laboratory staff do perform a general check for completeness and legibility. To maximize the usefulness and sustainability of the data in the RE Data Explorer, NREL encourages data providers to provide their own review to ensure data submissions are of high quality. The checklist below provides a list of considerations that third-party providers can use to review any given dataset for quality.

- ✓ **General:**
  - Does the dataset meet all the requirements listed in Box 1?
- ✓ **Relevance:**

[7] See http://gadm.org for more information and to download GADM administrative boundary layers.

- Will the dataset support the purpose of the RE Data Explorer, which is to inform RE planning, investment, and deployment?
- Do metadata provide clear description of the intended use of the dataset as it relates to RE planning, investment, and deployment?

✓ **Completeness:**

- Have all required submission fields of the Data Catalog for the dataset?
- Do any fields or values appear to be missing from the dataset?

✓ **Legibility:**

- Are the metadata sufficient to ensure the data is readily understandable to a user who has not previously been exposed to the data source?
- Are metadata free of spelling and grammatical errors?

✓ **Other quality concerns:**

- Is the submission free of personally identifiable information; business proprietary information; and/or copyrighted material?[8]
- Are sufficient contact information and links provided for an end-user to independently verify the dataset?
- Do all hyperlinks included in the metadata work, and do linked files reference a permanent URL?
- Does the metadata capture when, where, why, and how the data was collected?
- Does the metadata clearly state any assumptions, proprietary software requirements, or other prerequisites to using the data?

Good Practices

In addition to following the requirements outlined in the preceding sections, NREL encourages data providers to implement the following good practices in their data submissions. While not mandatory, implementation of these practices will further improve the efficiency of data processing and the overall quality of the RE Data Explorer.

- For data in tabular format, provide descriptive columns names, using all lowercase letters with underscores (no spaces). Descriptive columns names should have the all words written out, with no abbreviations (for example, “capacity_factor” rather than “cf”).
- For data in tabular format, if multiple attributes correspond to the same location or spatial reference, include more than one column per data file. For example, district-level data for electrification rate and energy consumption can be provided in one file with three columns (district, electrification rate, and energy consumption) rather than two since these data are both linked to district.
- Provide data in the highest-resolution format possible. For example, if available, provide provincial data instead of country data.

[8] Personally identifiable information refers to any piece of information or combination of pieces that could be used to compromise the identity of an individual.

Maintenance

RE resource data, complementary data, and the barriers to and priorities for energy development are dynamic and may change as new information becomes available or policies are enacted. NREL periodically surveys available global datasets and reflects updates within the RE Data Explorer. However, NREL relies on third-party data providers to inform the RE Data Explorer team when country- or region-specific data have been updated, and to submit updates whenever possible. When maintained according to the processes described in this document, the RE Data Explorer provides a platform for visualizing and analyzing the most up-to-date publicly available RE resource and complementary data so that users can develop and compare different scenarios in a dynamic way.

Additional Information and Contacts

NREL's RE Data Explorer team welcomes questions and can provide guidance and support to third-party data providers in evaluating and preparing their data for incorporation into the RE Data Explorer. For more information, please refer to the following contacts:

- Sadie Cox: Sadie.Cox@nrel.gov
- July Tran: July.Tran@nrel.gov
- RE Explorer "Ask an Expert" service: https://www.re-explorer.org/expert.html

Appendix A: Generic Geospatial Data "Wish List" for the RE Data Explorer

The list below represents a "wish list" of datasets that can support a wide variety of RE analyses in the RE Data Explorer. **NREL does not expect that all of these data will be available for a given country or region;** this list merely seeks to be comprehensive. **NREL also welcomes insights from data providers on additional relevant datasets that are not captured in this list**.

The data wish list is organized in order of highest to lowest priority, though **NREL also welcomes insights from data providers on the relative priority of these datasets in a particular country or region**. The prioritization is primarily based on the usefulness of the data to visualize RE resources relative to crucial development considerations and to conduct technical and economic potential analysis in the RE Data Explorer. In some cases, NREL can use publicly accessible "default" datasets to populate the RE Data Explorer.[9] However, even in these cases, data from in-country sources is preferred if it is higher quality—e.g., more recent, higher resolution—than the default dataset.

Please follow the guidance in the main body of this document regarding data formatting, language, etc. As a general rule, higher resolution spatial data is most valuable for use in the RE Data Explorer; for example, NREL prefers provincial data over national data, and point (i.e., site-specific) data over provincial data.

Highest Priority (Most Valuable to RE Data Explorer Core Functionality)

Grid Data

- Electricity transmission and/or distribution line locations (including interconnections with other power systems), capacity, and other relevant characteristics
- Electricity substations locations, voltages, and other characteristics

RE Resource Data

- Solar energy resource data (e.g., irradiance, ground measurements)
- Wind energy resource data (e.g., wind speeds, power density, ground measurements)
- Biopower resource data (e.g., crop or forestry residues, biogas availability, municipal solid waste)
- Hydropower resource data
- Geothermal resource data
- Marine hydrokinetic resource data

Medium Priority (Enables RE Data Explorer Analysis Functionality and Key Visualizations)

Administrative Data

- Boundaries (province, district, etc.)

Environmental

- Designated protected areas
- Land cover or land use

---

[9] Please refer to the RE Data Explorer to view the datasets already available for particular countries or regions.

- Elevation

Infrastructure

- Roads

Lower priority (Additional data that inform the identification of tradeoffs and synergies between RE development and other national development goals)

Administrative Data

- Population density or distribution
- Land ownership (e.g., delineation of privately owned and publicly owned lands)

Infrastructure Data

- Other transportation infrastructure (e.g., railroads, airports)
- Existing and/or planned power plants (e.g., location, fuel type, capacity, operational status)
- Other energy infrastructure (e.g., natural gas pipelines)

National Development Goals or Concerns

- Priority economic development areas
- Electrification rates
- Poverty rates
- Vulnerable communities (e.g., indigenous populations; populations relying on subsistence farming)
- Economically important and/or sensitive fisheries (marine or freshwater)
- Cropland type and productivity
- Any other data set that can help inform the identification of tradeoffs and synergies between RE development and other national development goals

Energy demand and costs

- Electricity and/or heating demand (e.g., by end-use sector, primary fuel)
- Electricity and/or heating price (e.g., marginal generation or wholesale energy market price; retail rates; projected energy price)
- Building inventory (e.g., building count, building type/area, occupancy rates, roof area, roof suitability)
- Any other data set that allows for the characterization of energy usage/demand in a given region.

Resilience

- Critical electricity loads
- Fault lines
- Tsunami frequency
- Landslide frequency
- Fire risk
- Earthquake frequency

- Drought events
- Flood risk
- Heatwave risk
- Tornadoes risk
- Other natural or man-made hazards

Appendix B: Metadata Template

Data providers should provide a text or word processor file that includes the information below. Fields marked with an asterisk (*) may be provided via the RE Exlorer Data Catalog instead of the metadata text file.

Qualitative description	*A general explanation of the information included in the dataset, any details the data provider wishes to share about how the dataset was developed, and any details the data provider wishes to share about how the dataset should be used. Please include the following information, at a minimum:* • *A brief description of the what the value in a particular column or file represents* • *The units of the value in a particular column or file and any scalar multipliers applied to the data.* • *Definition of any acronyms used in the data file*
Spatial reference information	*Qualitative description the spatial reference system associated with the data, including the projection of GIS data, coordinate system, and/or administrative boundary reference. .*
Data source (and, if different, data author)*	*Organization(s) that creates, maintains, and/or disseminates the dataset.*
Geographic extent*	*The administrative area that the dataset covers (e.g., Lower Mekong sub-region; Thailand; Thanh Hoa Province).*
Spatial resolution*	*The administrative level at which the data are enumerated, or, alternatively, the number of data points within a given area (e.g., district-level data; or 1 x 1 km grid cells over a continuous surface or discrete sites).*
Temporal range*	*The total period of record observed in the data (e.g., one year, multiple years)*
Temporal resolution*	*The time steps taken in the dataset (e.g., annual; monthly; hourly).*
Vintage*	*The year in which the dataset was published.*
Format*	*The file type in which the dataset is available (e.g., tabular, shapefile)*
Accessibility*	*Any limitations (such as licenses) on who may access or utilize this dataset, and/or if there is a cost associated with the dataset (e.g., publicly available; available for purchase; for official use only)*
Language*	*The language(s) in which the data is available.*

Appendix C: Data Permissions Letter Template

The following page provides a template for a letter from data providers to NREL providing permission for NREL to incorporate one or more datasets into the RE Data Explorer. NREL welcomes data providers to adapt this letter (for example, by replacing the highlighted text in brackets with information specific to the data and its provider). Alternatively, data providers may provide written permission in their own format. NREL welcomes scanned or mailed letters and/or emails containing the information included in this template.

NREL requests that data permissions be provided on letterhead of the data provider's organization, and/or from an official email account associated with the data provider's organization.

Data permissions letters may be sent to the RE Data Explorer contacts listed in this document.

[Name of Data provider's organization]
[Street Address of Data provider's organization]

[Date]

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401

To Whom It May Concern:

With this letter, [Data provider's organization] hereby grants permission for the United States Department of Energy's National Renewable Energy Laboratory (NREL) to incorporate the following dataset(s) into its Renewable Energy (RE) Data Explorer application for [country or region]:

1. [List of the datasets that the Data Provider wishes to include in the RE Data Explorer
2.
3.
…]

Further, [Data provider's organization] acknowledges that incorporating the dataset(s) listed above into the RE Data Explorer will enable public access these data. [Data provider's organization] authorizes the following level of access for the dataset(s) listed above [***Please check one of the boxes below***]:

☐ **Visualization only.** The dataset(s) listed above will be included as data layers that any user may view within the RE Data Explorer's web-based interface. Visualization-only data will not be available for download directly from the RE Data Explorer. Instead, [Data provider's organization] requests NREL to provide RE Data Explorer users with the following link so that users can request data directly from the data source: [insert website URL or contact information that members of the public should use to access or request access to the dataset(s) listed above].

☐ **Visualization and download.** In addition to including the dataset(s) above in the RE Data Explorer's web-based interface for users to view, NREL may also enable download functionality to allow RE Data Explorer users to download the dataset(s) listed above in a variety of formats (including shapefile, CSV, KML, GeoJSON) directly from the RE Data Explorer interface.

Finally, this letter certifies that [Data provider's organization] is the owner of these data and has the authority to grant NREL permission to incorporate them into the RE Data Explorer.

Thank you for your consideration.

Sincerely,

[Signature]