

Proof-of-concept Energy Information Portal – Final Report

Prayas (Energy Group), Pune

May 2017

This is the final report of the research study on “Development of an Energy Information Portal for India” sanctioned through letter no. O-15012/3/15-Research dated 28th December 2015 from NITI Aayog.

This study was sponsored by the Research Vertical, NITI Aayog, Government of India, and conducted by Prayas, Pune. Prayas received the grant under the Research and Study Scheme 2013 of the NITI Aayog to produce this document. NITI Aayog is not responsible for findings or opinions expressed in the document prepared. This responsibility rests with Prayas.

Project duration

April 2016 to March 2017

Timeline of deliverables

August 3 rd 2016	Submitted progress report indicating initial design of the portal
February 17 th 2017	Workshop held at NITI Aayog to present the portal and to get feedback
March 24 th 2017	Deployment of the portal on NITI Aayog energy division website
May 8 th 2017	Submission of draft final report
May 11 th 2017	Handover and technical training to be conducted for Energy Division, NITI Aayog
May 12 th 2017	Portal to be officially launched by Vice Chairman, NITI Aayog
May 26 th 2017	Submission of revised final report



Prayas (Energy Group)

Prayas (Energy Group), III A&B, Devgiri, Kothrud Industrial Area, Kothrud, Pune – 411038.

Introduction and purpose

Availability of official data in the public domain is crucial for formulation, analysis and research of public policy. In addition to policy making, data availability is crucial for decision making in businesses and investment firms.

This is particularly true of the energy sector, which is a key sector for the economy and where technical complexity is high. In addition, availability of energy data in the public domain helps democratise the sector by making information accessible to the general public. Such data includes information such as available reserves, trends in production and prices, patterns of consumption, and statistics related to energy access.

The energy sector in India is administered by multiple ministries and departments at central and state governments. At the central government, there are different ministries responsible for coal, oil and gas, power, renewable energy and nuclear energy. There are inter-linkages between the different sub-sectors and this has a bearing on how data is collected and organised. At the same time, some of the administration is concurrent in nature between Centre and the states, most prominently in the case of electricity. Given these considerations, collection and dissemination of energy data is a rather complex activity and requires coordination across multiple levels. On the consumption side, data is available from supply side ministries, household and enterprise consumption surveys and from various ministries dealing with various consuming sectors such as transport and industry.

Although there is a strong mandate to collect data for administrative purposes, official agencies are generally not mandated to disseminate that data¹. Consequently, data available through official statistics is often fragmented, incomplete, inconsistent and inaccessible to a wider audience². As a result, there are large variations in how energy data is managed in the country, leading to several data gaps. Filling these gaps will require greater inter-ministerial coordination, institutional capacity and interaction with data users.

One low hanging fruit in this space is a central energy information portal where data collected by different energy ministries is collated and made available to the general public in an easy to understand manner. To help demonstrate this, development of a proof-of-concept energy information portal (henceforth referred to as ‘the portal’) was taken up by Prayas (Energy Group) with support of a grant from NITI Aayog under the socio-economic research scheme. The purpose of this effort is to show official Indian energy statistics in one place, to make the statistics accessible to the general public and to present the data in visually interesting and easy-to-understand ways. The effort is intended to highlight the effectiveness of presenting data in a unified portal, and can also help in identifying some candidates for greater inter-ministerial coordination with respect to ensuring consistency³ and quality across data sets. In this regard, this exercise is intended to be useful to policy makers and analysts in the energy sector as well.

¹ A notable exception to this is the electricity sector, with the Central Electricity Authority being mandated by The Electricity Act (2003) to disseminate data on a periodic basis.

² These issues are elaborated in a couple of studies undertaken by Prayas (Energy Group) titled [An Assessment of Energy Data Management in India](#) and [Data Gaps in India's Energy Sector](#).

³ For example, data on consumption of imported coal in the power sector is reported by CCO and CEA. However, these numbers do not match.

The period of the research grant is from April 1, 2016 to March 31, 2017. This is the final report under this grant, describing the structure of the portal, the various visualisations embedded in it, the underlying software architecture, some insights we got while developing the portal, and some pointers towards future work.

Features of the portal

The proof-of-concept portal displays a subset of official energy statistics, chosen to demonstrate the range of possible data and features that such a portal can host and present. Following are some of the high level features of the portal.

- Data shown on the portal is limited to **commercial energy** and is more focused on the **supply side**, reflecting the nature of available official data.
- The portal is **visual** in nature. Appropriate charts have been used to bring out interesting trends and insights. Data has been mapped spatially where such a representation can provide some insights into how the energy is produced and consumed across the country.
- These visual elements are **interactive** in nature and the user can drill down deeper into a specific sub-sector or compare energy production and consumption across states and over the years through a series of clicks.
- Effort has been made to provide as much data as possible at the **state level**. As a demonstration of further drill-down, electricity generation and household access are provided at a district level for 3 states, namely Maharashtra, Odisha and Rajasthan.
- All data on the portal is **yearly** for the financial years **2005-06 to 2015-16**. Some 2015-16 data such as for electricity consumption and captive electricity generation is not yet available in the public domain.
- A national **energy balance** and a **Sankey** (energy flow) diagram are constructed from the data accessible through the portal.
- While the dashboards are visual in nature, all of the underlying data is available for **download** in easy-to-use formats such as csv and pdf.
- Data is accompanied by **metadata** such as the data sources, clarifications and caveats where applicable.
- Data is **publicly accessible**, free of charge and without a login.

Structure of the portal

One of the primary aims of the portal is to represent national energy statistics in an easy to understand format even for non-experts. In order to achieve this, data is shown in dashboards consisting of a variety of interactive visualisations such as charts, maps, and tables on a particular topic.

The dashboards are organised in a tiered manner. At the top level are the commercial energy balance tables and Sankey diagrams showing annual energy balance and flows at the national level. These numbers are broken down in the next level by sub-sectors, namely electricity, coal, oil and gas, listed in the left hand side pane of the portal. Within each sub-sector menu, there are sub-menus dealing with different aspects of that sector, such as reserves, production, imports/exports, consumption and prices. Selecting a sub-menu opens up a tabbed page, where each tab represents one dashboard consisting of a few charts. Table 1 describes the menus and dashboards on the portal in a bit more detail. [Annexure](#) has a few snapshots from the portal.

Table 1: List of dashboards in the proof-of-concept portal

Menu	Sub-menus	Highlights
Energy Flows		- Sankey diagram showing energy flow from resource extraction to energy production and conversion to sector-wise consumption
Energy Balance		- Yearly energy balance tables showing the primary energy sources, transformation from one energy source to another and final consumption of those energy sources
Electricity	Overview, Potential, Capacity, Generation, T&D Losses, Consumption, Supply & Deficit, Access, Tariff	- Electricity generation capacity potential map for large hydro and renewable electricity sources - solar, wind, small hydro and biomass - Capacity and generation of electricity by state, fuel source and ownership - State-wise transmission and distribution losses over the years - State-wise, final consumer category-wise consumption of electricity along with regional trends - National and regional power and energy deficit in electricity supply - Household and village electrification trends - Power plant-wise electricity generation tariffs
Coal	Reserves, Supply, Import, Transport, Consumption, Prices	- State-wise coal and lignite reserves and production by producing company - Quantity and price of imported coal by exporting country, Quantity of imported coal by consuming sector - Modal share of coal transport - Consumption by state and industry sector - Trajectory of CIL notified prices over the years
Oil	Reserves, Crude Supply, Products Supply, Pipeline, Consumption, Prices	- Basin-wise and state-wise oil reserves - Domestic production of crude oil and petroleum products - Quantity and value of import and export of crude oil and petroleum products - Map of existing and proposed petroleum product pipelines - State-wise, sector-wise consumption of petroleum products - Domestic crude oil prices and petroleum product price trends
Gas	Reserves, Supply, Pipelines, Consumption, Prices	- Basin-wise and state-wise reserves and production of natural gas - Quantity and value of import and export of natural gas - Map of existing and proposed natural gas pipelines - Sector-wise gas consumption and city gas distribution sales - Gol notified gas prices
States Overview		- State-wise per-capita production and consumption of energy commodities

Software design

The portal focuses on providing interesting, accurate and current information on the Indian energy sector through dashboards that illustrate the data in an accessible manner.

The following software tools were used to develop the portal:

- Tableau®
- MySQL® database
- Google’s Sankey library

- CARTO®, a free GIS mapping interface

The dashboards were created using the Tableau Desktop® software (henceforth referred to as Tableau) – a desktop software tool that is used for data visualisation and analysis. It provides a rich visualisation toolkit along with the ability to quickly pilot and publish dashboards resulting in significantly lower effort than the use of comparable visualisation libraries such as d3.js⁴. This will make it easier for NITI Aayog to maintain and enrich the portal in the years to come. Tableau was also chosen due to the following features it possesses:

- rapid analysis of data
- easy creation of visualisations using built-in charting functions
- highly interactive charts
- custom coding for tailoring visualisations to specific needs
- mapping platform for spatial representation of data
- basic statistical analysis (for future needs)
- ability to import data from a wide variety of sources

A Tableau license needs to be purchased to create new dashboards and modify existing dashboards using Tableau Desktop. Prayas possesses a license for Tableau software, which it used for the purposes of developing the portal. Tableau training resources can be accessed from <https://www.tableau.com/learn>.

Data for the portal is sourced from a variety of public sources, primarily from annual statistics published by the Central Electricity Authority (CEA), Coal Controller's Organization (CCO), Ministry of Petroleum and Natural Gas (MoPNG) and Ministry of Statistics and Programme Implementation (MoSPI). These sources are listed on the References page on the portal. Some of the references are available online in soft copy/pdf format and some of these are available only in hard copy form.

The data resides in a MySQL® database, organised in tables that align with the sectoral and sub-sectoral organisation in the portal. Data is pulled directly from the database into Tableau.

After visualisations are created, they are published through Tableau Online®⁵. Tableau Online is a cloud based service that abstracts the complexity of maintaining a highly visual data analytics web platform and thus enables publishing Tableau dashboards to the wider public. The visualisations can also be published on a different web server with the help of the Tableau Server®⁶.

Dashboards published on Tableau Online are embedded in a web application. The web application's main purpose is to provide an interface for accessing the dashboards. In addition to this, the application also provides user interactivity features such as propagation of year and state filters across all dashboards (where those filters are applicable). The application is lightweight and written in html and javascript. Following are some of the components/features of the web application:

- Landing page explaining the portal along with some live data
- Navigation between various dashboards, charts and maps
- Propagation of year and state filters from one dashboard to another

⁴ Refer to <https://d3js.org/>.

⁵ Tableau Online service is at <https://public.tableau.com>.

⁶ More details regarding the Tableau Server are available at <https://www.tableau.com/products/server>.

Two of the visualisations on the portal were created outside Tableau – Sankey diagrams and oil and gas pipeline maps.

The portal comprises of energy flow diagrams (called Sankey diagrams) showing the flow of energy from resource extraction to transformation from one energy source to another to consumption by end users. These annual energy flow diagrams are created using the Sankey library provided free by Google⁷ and the same data that is used to create the visualisations. Google’s Sankey library was chosen since, as of version 10.1, Tableau does not provide an easy way to create Sankey diagrams.

For the same reason, i.e., owing to Tableau limitations, oil and gas pipeline maps were developed outside Tableau. PNGRB publishes oil and gas pipeline maps in image format. Such maps are not interactive; hence they need to be converted to a digital format. This was done using a GIS software tool, published through a free map service called CARTO® (formerly CartoDB) and embedded in the portal.

Details of the database, its interface with Tableau, and usage of Tableau, Google’s Sankey library and CARTO are provided in a separate document to NITI Aayog in addition to a short training session on using these.

Deployment and future support

The portal is published on the web and will be publicly accessible at <http://indiaenergy.gov.in/edm> and is accessible from NITI Aayog’s Energy division portal. The portal will be maintained by NITI Aayog going forward with Prayas (Energy Group) assisting with technical queries as needed.

Future work

The portal should be continuously updated with new data as and when it becomes available to continue to be useful. Some of the data that appears on the portal, such as consumption of imported coal in the power sector, may need to be reconciled with other data sources publishing the same data. Also, the energy balance tables displayed on the portal need to be reconciled with the energy balance published by Central Statistical Office in the annual Energy Statistics publication.

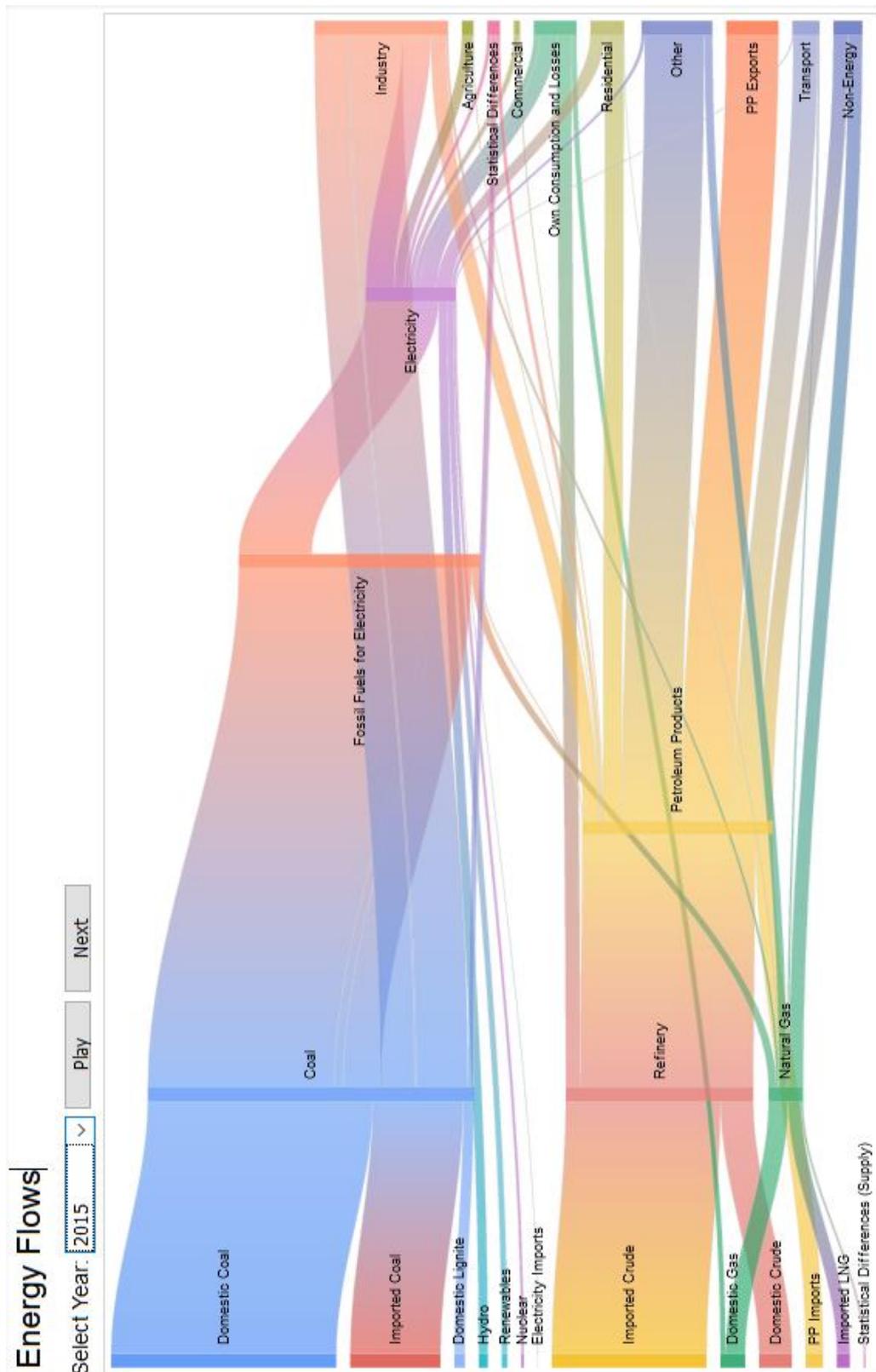
As indicated earlier, the purpose of this project is to develop a proof-of-concept portal to demonstrate how existing data from government sources can be shown in a more accessible way with interesting visualisations. Hence, the current dataset can be expanded further in order to bring all energy sector data in one place. Following are some suggestions for the same:

- Add state-level dashboards with data relating to different energy sources
- Expand the list of data sources to other public sources such as regulatory bodies and state government agencies.
- Add sub-state data similar to district household electrification data shown for 3 states
- Add sub-yearly data, particularly for electricity and oil-gas sectors
- Financial data such as capital costs, average cost of supply and consumer tariffs
- Add historic data prior to 2005-06
- Include data available in the public domain from trusted non-government sources
- Greater interactivity such as navigating from one dashboard to another through hyperlinks
- More live data on the landing page or other sectoral pages

⁷ More details can be found at <https://developers.google.com/chart/interactive/docs/gallery/Sankey>.

Annexure: Snapshots from the portal

1. Sankey Diagram: Commercial energy flows during the fiscal year 2015-16



2. Energy Balance: Commercial energy balance for the fiscal year 2015-16

Fiscal Year 2015

All values in MTOE

Energy Balance

	Coal	Lignite	Crude	Petroleum Products	Natural Gas	Hydro Electricity	Nuclear Electricity	Renewable Electricity	Electricity	Total
Domestic Production	275.8	13.0	40.1	0.0	24.1	11.1	3.1	7.0	0.0	374.3
Imports	110.5	0.0	189.4	21.3	15.7	0.0	0.0	0.0	0.4	337.4
Exports	0.0	0.0	0.0	63.9	0.0	0.0	0.0	0.0	0.0	63.9
Primary Energy Supply	386.3	13.0	229.6	-42.6	39.8	11.1	3.1	7.0	0.4	647.8
Non-Energy Use	0.0	0.0	0.0	-18.1	-17.3	0.0	0.0	0.0	0.0	-35.5
Statistical Differences	-6.6	-0.4	0.0	-4.2	2.4	0.0	0.0	0.0	-4.5	-13.1
Electricity Plants	-274.2	-10.7	0.0	-1.2	-9.6	-11.1	-3.1	-7.0	110.5	-206.5
Refineries	0.0	0.0	-211.9	211.9	0.0	0.0	0.0	0.0	0.0	0.0
Transformation Losses	0.0	0.0	-17.7	0.0	0.0	0.0	0.0	0.0	-7.8	-25.5
Transport Losses	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-20.6	-20.6
Final Energy Consumption	105.6	2.0	0.0	145.7	15.3	0.0	0.0	0.0	78.0	346.6
Residential	0.0	0.0	0.0	23.0	0.3	0.0	0.0	0.0	18.0	41.3
Industry	105.6	2.0	0.0	18.4	2.9	0.0	0.0	0.0	34.8	163.6
Commercial	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	6.4	7.5
Agriculture	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	13.4	14.1
Transport	0.0	0.0	0.0	29.8	2.3	0.0	0.0	0.0	1.3	33.4
Other	0.0	0.0	0.0	72.9	9.8	0.0	0.0	0.0	4.1	86.7

3. Example dashboard – Natural gas overview dashboard

Dashboard shows reserves and production by state, and consumption by consuming sector for the chosen year (2016) along with nation-wide trends over the years.

