The Electricity Journal 30 (2017) 4–9

Contents lists available at ScienceDirect

The Electricity Journal

journal homepage: www.elsevier.com/locate/electr

Energy finance data warehouse: Tracking revenues through the power sector

Claire Zeng[a], Stephen Hendrickson[b], Sangkeun Matt Lee[c], Supriya Chinthavali[c], Jessica Lin[a,*], Eric Hsieh[b], Mallikarjun Shankar[c]

[a] ORISE Fellow to the US Department of Energy, United States
[b] U.S. Department of Energy, United States
[c] Oak Ridge National Lab, United States

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Energy
Financial data
Power sector revenue
Sankey
Visualization

ABSTRACT

Reliable data is needed to understand financial relationships in the power sector. However, relevant data acquisition and visualization can be a challenge due to the fragmented nature of the power sector. The US DOE and ORNL leveraged a Sankey prototype to elucidate the 'big picture' of financial flows to understand the complex relationships between specific actors within the power sector. The continued incorporation of high quality data can improve the fidelity of such an approach and lead to an increasingly detailed understanding of financial relationships in the power sector and their implications for policymakers.

1. Introduction

The finance team in the Office of Energy Policy and Systems Analysis (EPSA-50) provides data-driven policy analysis to inform energy policy deliberations. To that end, EPSA develops and maintains analytical tools and capabilities for understanding the role of finance and financial policy in the energy sector.

In the context of federal policy deliberations, quantitative analysis requires data at the company and asset level, collected nationally. Financial data that is both granular and uniform rarely exists in the highly fragmented power sector industry. Multiple entities collect power sector financial data, ranging from the Federal Energy Regulatory Commission (FERC), to regional transmission organizations (RTOs) and independent system operators (ISOs), to state regulatory bodies. Each of these actors plays a different role in the overall U.S. power system, which can complicate the exchange of relevant information.

The project team, led by EPSA, collected revenue information from relevant sources to piece together revenue flows within the U.S. electricity system to understand the complex relationships among specific actors within the market. Data scientists at Oak Ridge National Lab used this data to map these revenue flows into Sankey visualizations.

2. Theory

Sankeys are flow diagrams that depict a set of dynamic relationships in a system. The Irish captain Matthew Henry Phineas Riall Sankey first used this diagram in 1898 to depict the energy efficiency of a steam engine (EIA, August, 2016). Research organizations commonly use Sankeys to describe physical resource flows. For example, domestically, the Lawrence Livermore National Laboratory maintains energy, carbon, and water flow charts at the international, state, municipal, and organizational level (EIA, 2016a). Internationally, Eurostat, the statistical office of the European Union (EIA, 2016b), and the International Energy Agency (FERC, 2015) maintain Sankeys depicting European and global energy consumption, respectively. This project extends the Sankey concept to describe the financial transactions that occur in

* Corresponding author.
E-mail address: jessica.lin@alum.mit.edu (J. Lin).

http://dx.doi.org/10.1016/j.tej.2017.03.001

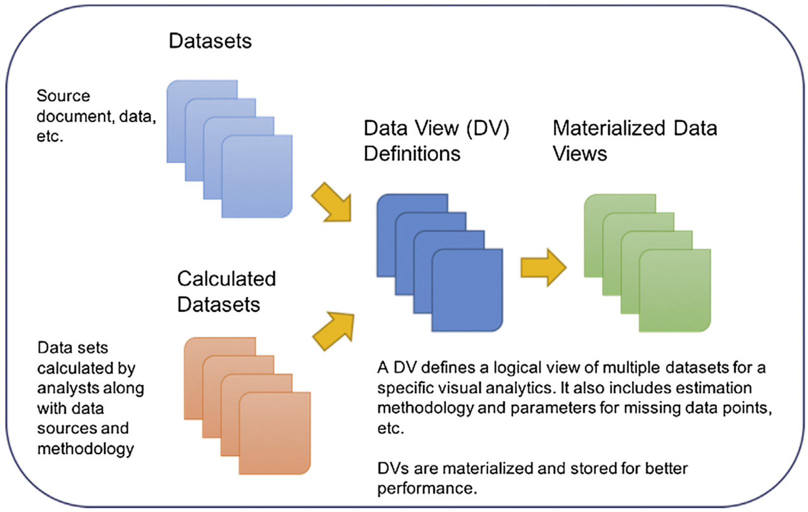

Fig. 1. Data stored in the EFDW repository.

exchange for physical flows of electricity through a Sankey diagram as a use case for the Energy Finance Data Warehouse.

3. Methodology

3.1. Energy Finance Data Warehouse

EPSA-50 commissioned the Energy Finance Data Warehouse (EFDW) project to aggregate financial information from a variety of data sources. EFDW is composed of three parts: the EFDW repository, EFDW data view interpreter,[1] and EFDW visualization dashboards. The EFDW is currently stored in an Excel database, but may change platforms; the purpose of the EFDW is to be the foundation for data collection in the industry for future analysts and the public.

Information stored in the EFDW repository includes data, metrics, and projections associated with company financials and physical assets. The data stored in EFDW can be grouped into the following categories: generic datasets, calculated datasets, data views,[2] and materialized data views. The datasets are organized and indexed so that EPSA analyst can quickly identify what datasets are available for analyses.

The first use case chosen for the EFDW was the Sankey; the team is currently exploring other visualization use cases.

3.2. EFDW visualization dashboard

The analyst team determined that no single existing software tool could provide all desired analytical and visualization capabilities. Consequently, the team developed a set of front-end graphical user interfaces known as EFDW visualization dashboards. These dashboards let analysts view illustrations of the data from the EFDW repository. As a specific use case for the EFDW, Sankey diagram dashboards allow analysts to interact with diagrams displaying the context around the data points in which they are interested. The EFDW visualization dashboards incorporates existing web tools, including Infocaptor (which provides access to Tableau), SankeyMATIC, and Sankey Diagram Generator visualizations.

4. Discussion

The Sankey visualization assumes that every unit of delivered electrical energy corresponds to a payment for power, where the payment flows run opposite to electricity flows. The visualization depicts the net value of transactions between two entities and excludes trades or hedges not tied to physical delivery of electricity.[3] The visualization also excludes electricity consumption from distributed energy resources (Figs. 1 and 2).

The Sankey follows the movement of revenues from the initial customer, the end user, through the distributor, transmitter, generator, and fuel and operations and maintenance (O&M) companies. The Sankey given as an example (Fig. 3) uses data from 2014, when end users spent approximately $390 billion on electricity.

The Sankey simplifies relationships to graphically represent the flow of funds. For instance, electricity is bought and sold in either vertically integrated or centrally organized markets. The calculations for revenues are slightly different for the two categories, but revenues are combined in the final Sankey; end users may pay retail marketers directly or through the incumbent utility, depending on state regulations. In addition, functions such as generation, transmission, and distribution are depicted separately even if an IOU is vertically integrated.

Revenue payments start in Fig. 3 from the leftmost column depicting end users. Residential, commercial, and industrial consumers account for about 40%, 40%, and 20%, respectively. End users pay the load-serving entity (LSE), which owns and operates the distribution system, shown in column 2. LSEs include large investor-owned utilities (IOUs) as well as public power (municipal, state-owned, federally owned, and political subdivision power), cooperatives, and retail marketers (which do not own

[1] EFDW Data View Interpreter is a software module that is responsible for parsing data views and materializing them in physical datasets so that data views can be exploited for various analyses. The materialized datasets are also stored in the EFDW repository. For a proof-of-concept, the team developed EFDW Data View Interpreter for EFDW's Sankey diagram visualization dashboards.

[2] A data view is a definition of a virtual dataset constructed from other datasets in the repository. Analysts can describe their needs in their data view definitions. For instance, a data view can describe an input dataset for EFDW's Sankey diagram visualization dashboards, where the data view includes the related available data, structure of Sankey diagram, and proportions for computing unavailable data points. Data views themselves cannot be utilized directly for analyses. Data views describe how to construct datasets as structured JSON documents.

[3] The exception is RTO/ISO retained expenses, which are accounted for as a small outflow from the RTO/ISOs.