The Data Challenge
In recent years, electric utilities have been tasked with operating the distribution grid with thousands of dynamic, ever-changing, loads in addition to an ever-increasing number of distributed energy resources (DER) per distribution circuit. This complicated load and generation mix presents many challenges to the utility, such as generation that depends on weather conditions, reverse power flow, and relay setting to protect the grid for all operating scenarios. To counteract the challenges, it is becoming more critical to know the types and operating characteristics of both the loads and power-producing devices.

Solution Overview
Disparate data sets are leveraged by recognition algorithms to determine load or generation profiles and to predict the energy consumption of dynamic loads and the performance characteristics of the DER based on external influences, such as weather, time of day, and season of year. Once the presence of a certain type load or DER is determined, a predictive solution can be used to understand the operational characteristics of the load or DER and its impact on the grid.

Potential Methods for Solving the Problem
The most straightforward method would be to develop specific load and DER output profiles of equipment and devices of interest in order to compare to the metering and SCADA data. The correlation analysis or algorithm would be able to scan individual smart meter and SCADA device data and pick out metering points that serve equipment and devices of interest.

Available Data Sets
The data sets highlighted in the following figure are available in the EPRI Data Repository to solve this data analytics case.
Classifications of Data:
- Traditional Data Set
- New Data Set
- Structured Data
- Un-structured Data
- Format of Data Varies

Frequency of Measurement:
- Cycles (C)
- Seconds (S)
- Minutes (M)
- Hours (H)
- Days (D)
- Months to Years (Y)
- Event Driven (E)

Denotes a primary data set used to solve this data analytics case.