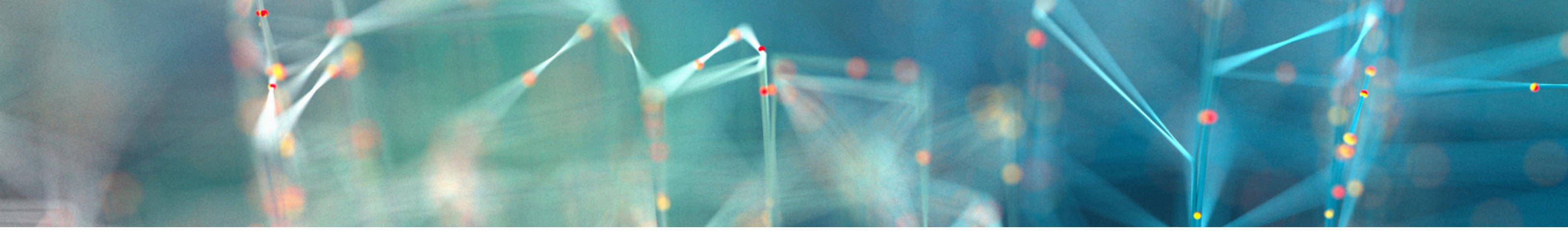
TEA

THE EnergyAuthority®

A SCALABLE DATA MODEL FOR PUBLIC POWER'S DATA-DRIVEN DECISION MAKING



TEA DATA DYNAMICS
ELEVATE • ENERGIZE • ENRICH

ELEVATE YOUR COMMUNITY.
ENERGIZE YOUR DATA.
ENRICH YOUR GRID.

INTRODUCTION – Utilities and Big Data

Are you a utility company looking to take your decision-making abilities to the next level? If you're like others, you are collecting rich datasets from your Advanced Metering Infrastructure (AMI), but converting that data into truly transformational information is proving to be elusive. One reason why is that traditional IT infrastructure and database management tools fall short in their ability to store and organize this data efficiently.

That's where The Energy Authority's (TEA) Data DynAMIcs™ advanced analytics solution comes into play. By leveraging modern data engineering techniques, TEA is able to unleash valuable insights from AMI data that would otherwise remain undiscovered. This white paper describes how modern data engineering supports TEA's Data DynAMIcs advanced analytics solution to unleash valuable insights from AMI data, giving you the edge you need in a competitive market.

Asking Questions of the Data

The Needle in the Haystack

In the typical use case for AMI, meter reads go to the meter data management system (MDMS), usage data moves to the billing system, and then the process repeats with each billing cycle. Meanwhile, raw meter data continues to pile up in the MDMS or, eventually, an archive.

However, when decision makers need to glean key insights from this wealth of data, they find that it is virtually impossible. Scouring data in its rawest form to perform deep analytics can be an exercise in futility if the data is not organized appropriately. Like manually searching for a record in a room of unlabeled file cabinets, querying raw data requires much effort but yields little insight. Often, the effort outweighs the expected benefit of the insight, and the question goes unanswered.

The Improvised Approach

In some cases, the answer seems important enough to form a team of data engineers and analysts to find it. They use database applications to cleanse and curate the raw data into a queryable format known as a data model. The data relevant to the question can be filtered and compiled with the model. Statistical analyses and algorithmic processing can be applied as needed, eventually yielding an answer.

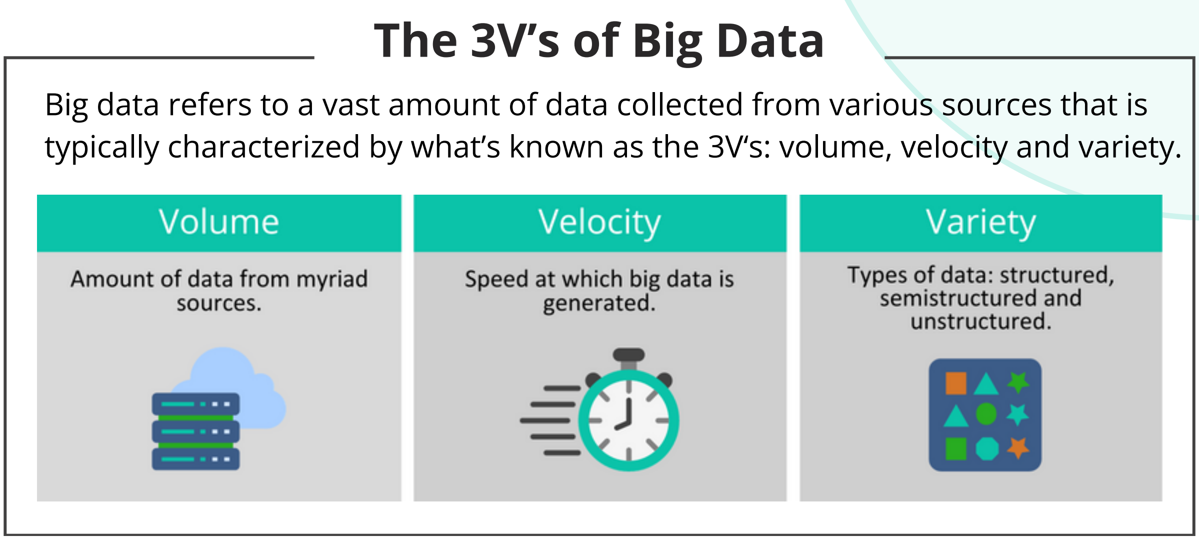
However, because the model was developed for a particular question, it may not be helpful for the next project. Moreover, the dataset becomes stale over time without a process to continuously update and integrate new raw data into the model. For both reasons, the analytics team often must start from scratch on organizing the relevant data when presented with a new business problem.

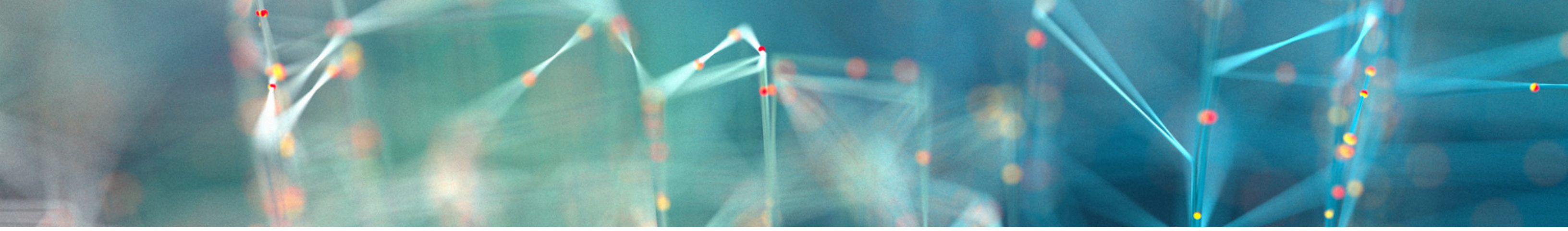
Facing the 3Vs - Volume, Velocity, Variety

When the data sets become larger and trend towards Big Data sets, it becomes more difficult for a utility's analytics team to develop data models and pipelines capable of scaling up and performing reliably when deployed into a production process.

Additionally, executing several analytics projects with unique data models and integration code leads to redundant data discovery, cleansing, and curation work. This redundancy can lead to discrepancies across models as each model may have its own data-cleansing assumptions and errors.

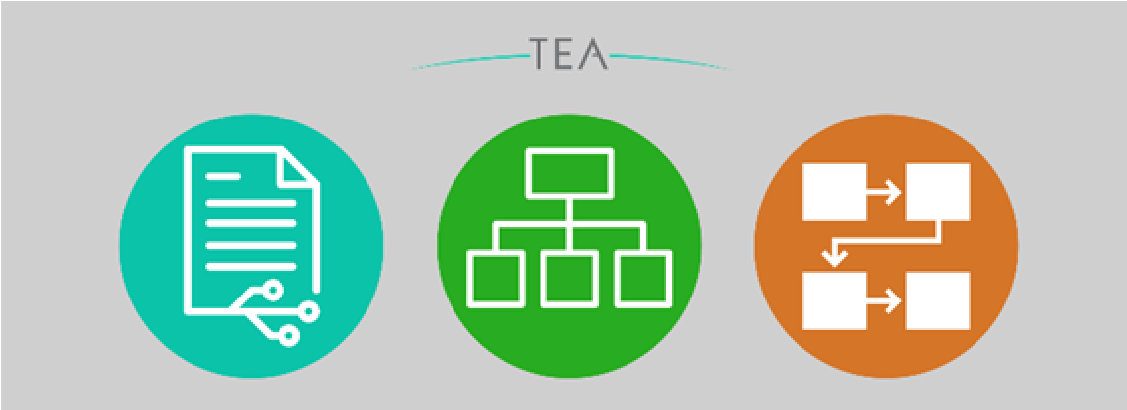
Ultimately, this approach fails under the strains of the three Vs (3Vs): the data's volume, variety, and velocity—a central challenge of Big Data. Tools and approaches that were adequate only a few years ago are now rendered obsolete. Successfully managing the 3Vs requires a more robust and efficient framework to be reliable and cost-effective.





COMMON DATA PLATFORM

The Data DynAMIcs team designed the Common Data Platform, a cloud-based, state-of-the-art data lake architecture to store and organize AMI data. It provides a reusable platform for data science and analysis. It also serves as a scalable, flexible, and efficient repository for the ever-increasing amount of data and includes reporting functionality in an easy-to-use online dashboard.

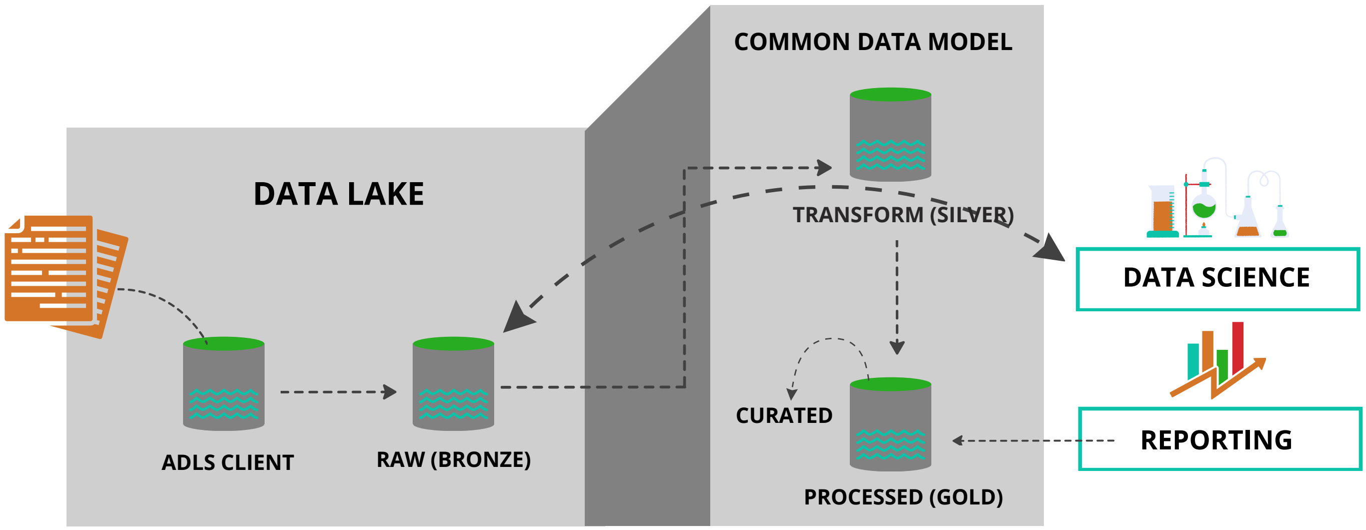


Specializing in Data

Data DynAMIcs enables enterprise-level analytics using a dimensional data model based on utility-specific process measurements and their context, e.g., kVA per hour for transformers. TEA's specialized team of data architects and engineers have developed a data model within the Common Data Platform that take these important characteristics into account, along with the source data pipelines required to populate it.

This specialization separates data discovery, cleansing, modeling, and curation from the statistical analysis and data science performed on the data sets. TEA's Common Data Platform framework provides a distinct advantage over traditional models by continuously reusing one data model for all data analysis projects, reducing the amount of rework the analytics team performs.

How does this work? The process relies in part on extensive automation to continuously and efficiently bring new data into a scalable data lake environment. The data pipelines do the heavy lifting via extract, load, and transform (ELT) scripts.



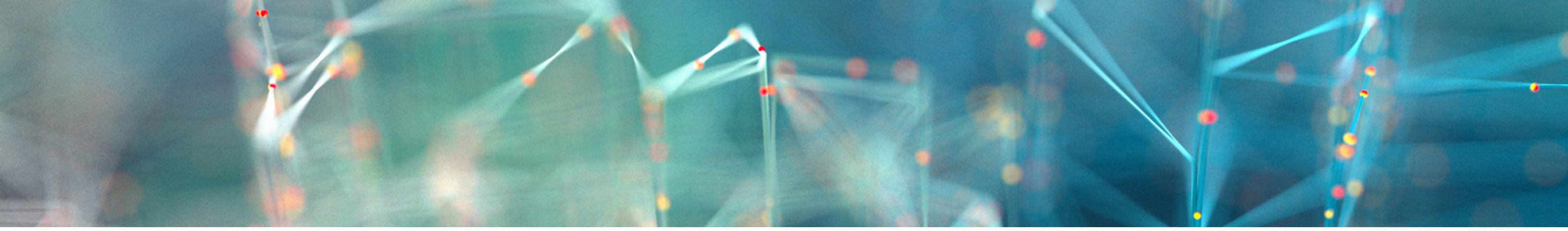
Scaling in the Cloud

As cloud computing has evolved, the value proposition for customer data-hosting in the cloud has also improved. TEA's Common Data Platform provides value and security to its utility partners by eliminating the requirement to maintain hardware, virtualization, disaster recovery, and system configurations. Cloud-specific technologies further eliminate the need to host equipment on-site to provide the computing power needed to run modern enterprise analysis software. Cloud providers offer virtually effortless reconfigurability and expansion as data sets grow and evolve. This platform, combined with TEA's Data DynAMICs service, has revolutionized decision-making for some of the nation's largest Public Power utilities.

Many Utilities, One Platform

Despite public power's shared mission to provide power, water, and gas to their customers and communities, IT departments vary in size, budget, and database management. Thus, Data DynAMIcs engineers write customized code for each utility to populate data at its lowest level of granularity.

This lets TEA's utility partners focus on their business without having to make system changes for compatibility with the Common Data Platform. This effort also ensures each utility's data are organized the same way—the *Common* in Common Data Platform. An algorithm created for one utility can be applied easily to any Data DynAMIcs utility partner. In short, collaboration with one utility often makes a new tool available to all.



CONCLUSION

With Data DynAMIcs, TEA takes its core mission to the cloud to maximize the value of utilities' data. The Common Data Platform performs seamlessly for utility partners ranging in size from 12,000 to 500,000 customers as a hosted, state-of-the-art data model purpose-built for public power. Its scalability makes it an economical yet robust platform that grows with utilities as they adopt new techniques to deliver power, water, and gas, manage their distribution systems, and better understand their customers. Now, public power has a partner to tackle its Big Data opportunities.

To learn more, contact Jamie Mahne, VP, Client Services & Chief Client Officer at jmahne@teainc.org or 904-356-3900.