21, rue d'Artois, F-75008 PARIS
http : //www.cigre.org

CIGRE US National Committee
2018 Grid of the Future Symposium

Data Governance Driven Visualizations with PI Data Historian in Dominion Energy

X. LI
Virginia Tech
USA

H. LIU, S. P. ADAMS, H. HUANG, K. THOMAS
Dominion Energy, Inc.
USA

SUMMARY

Data visualization is a powerful tool to monitor power system status and analyse system operating conditions in a quick and intuitive manner. However, the amount of time and labour to maintain the data is rapidly increasing with the scale and complexity of the enterprise. For a large power system, challenges occur when the data are collected and stored in incompatible formats, types and structures. To tackle this challenge, a solution of data governance by implementing OSIsoft® PI System to visualize data is proposed. The key contribution of the PI System is the proposal of the Asset Framework. In Asset Framework, assets are organized in a hierarchical structure, which makes data access and management efficient. Another most powerful feature of Asset Framework is the capability to create templates which can easily standardize the same types of operations. As a result, the measurement data of the monitored objects are able to be retrieved and maintained in a uniformed way if they come from the same manufactures or the same types of assets. In this paper, the build of the OSIsoft® PI Data with two equipment dashboards which are transformer monitoring map and relay failure alarm map are demonstrated, respectively. To access the real-time PI data, AF SDK and PI WEB API are implemented and multi-language hybrid programming is involved. The methods proposed here, as we envision, can serve as a guideline for enterprise to set up real-time data visualization systems, benefit data governance, and encourage collaborations across the enterprise.

KEYWORDS

Data Governance, PI System, Asset Framework, Data Visualization

xiawenli@vt.edu

I. BACKGROUND AND PROJECT INTRODUCTION

Nowadays, the power of data in driving businesses has been well-recognized. In power systems, data collection and processing have shown more significance in revealing and resolving ongoing and potential problems like system stability realization, cyber-security awareness, etc. Meanwhile, it is becoming even more challenging since every single moment and a large volume of data is being generated at a rapid rate across the whole power system by people, machines, Internet-enabled devices, and other sources. To be fully utilized, the data must be gathered, organized, made interpretable, and then analysed to provide meaningful insights. This is where data-visualization steps in, allowing the organizational leaders to have end-to-end accessibility to the data as well as the meaning behind them in real-time before a highly informed decision can be made. Data visualization converts the raw data or metadata to visuals and images which are easily grasped and comprehended by humans. It enables users to more effectively and intuitively see the correlation between operating conditions and measurements. Thus, even a new system operator with little experience in power system operation is able to notice the possible problem with the help of the data visualization tools [1, 2].

Data visualization can be made more efficient with data governance, which is the set of processes that ensures only important and dependable data assets are formally managed throughout the enterprise and dictate the most critical decision making, accounting and business processes. However, without effective data governance in a company, data visualization is almost impossible. Data governance is defined as the management of data to validate its accuracy and ensure its quality as per requirement, standard, and rules that the organization needs to their individual business.

The OSIsoft® PI System is a powerful tool and it is applied to address the bottleneck of data governance for large-scale power systems. In such cases, critical data can often reside in multiple incompatible systems across the enterprise. Accessing the data may involve the use of various software platforms and through complicated paths. The PI System is a server-based intelligent system to access, analyze and visualize data from multiple sources to people and systems across all operations. It breaks the boundary between different incompatible systems, formats, processes and provides a unifying solution to achieve end-to-end visibility across the enterprise. With its real-time acquisition and the high-fidelity of the data, the PI System can accurately and continuously diagnose the system. By combining historical time-series and event-based real-time data, it can also identify how and where to improve operations and provide insights for operational improvements and breakthroughs, largely helping with the streamlining operations for companies [3].

The key to unlock the value of PI system is PI Asset Framework (AF). The implication of AF to data governance is also thoroughly explained. To read and write to the PI points, two major data access methods are available, which are utilizing AF SDK or PI WEB API. To demonstrate the design and implementation of such process with PI System, we build a transformer monitoring map as well as a relay failure alarms map using Dominion Energy data. Furthermore, we explore detailed functions of the dashboards and discuss future possible improvements.

II. DATA GOVERNANCE

Data governance refers to processes that filters out duplicated, broken, badly structured information in data processing as an effort to improve its efficiency, reducing risks and provides clarity. In a power system, the generated data is much more than the effective data. The inception of a data governance project is a golden opportunity to get everyone on the same page and force a business to have clear definition about the core data. The lacks of data governance will guarantee the existence of poor data as shown in figure 1. Poor data includes inconsistent definitions, duplicate data points, missing fields and other data problems. This will lead to the difficulty in the future data processing.

A good data governance is constituted by three key factors - the rules for inputting and maintaining data, the enforcement of these rules, and management of the data. Rules for inputting and maintaining data include accepted definitions, stylistic rules, etc. Enforcement of these rules requires pointing the users to operate under the rules while having minor tolerance in input inconsistance. Management of the data means data stewards and users within the company working on the data in accordance with the rules. For instance, in Dominion Energy, renaming is the enforcement and management to address the poor data issue. These three factors clearly reveal two requirements to achieve a good data governance. One is the tool to allow this governance to happen, and the other is the efforts that carry this out. Dominion Energy is making effort to the naming convention based on one uniformed rule of all the assets in the system to achieve better management.

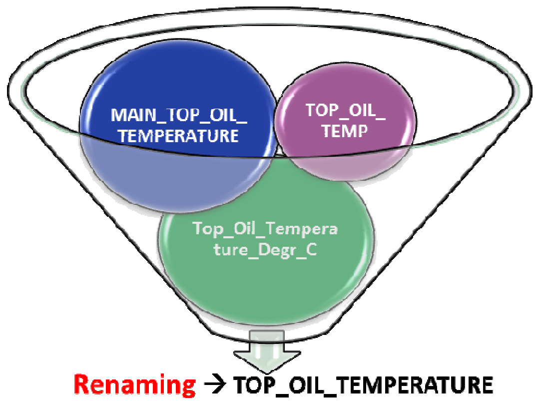

Figure 1. Example of poor data

III. PI SYSTEM INFRASTRUCTURE

A. PI System Infrastructure

In order to organize the data, PI system has been utilized in this project. Understanding the PI System Infrastructure is the cornerstone to build a PI System that works well with a real power system. Data infrastructure is what we connect the data source to its destination. As a utility company, Dominion Energy owns various data sources ranging from pieces of equipment like pump storages, battery units, transformers, relays, to different vendors, substations, or even power plants. Meanwhile, there are system operators and engineers who want the data from those data sources. These are the sources and destinations of the data, respectively. The data sources may be stored in incompatible databases and their destinations might be of different types of assets as well. Hence, accessing them may require usages of multiple software platforms. Besides, the end-to-end communication through the conventional infrastructure can become complicated due to the complexity in system topology. For example, a certain key connection may be heavily queried and it has to be carefully maintained. In the end, the problem remains as of how should we integrate the whole dataset from the use across the whole enterprise.

The PI System infrastructure is a scalable, high performance, reliable, manageable and secure infrastructure which can be sketched in figure 2 and it includes several core elements. The PI System collects all the sensor-based real-time data from various data sources into the PI Data Archive, models data in the PI AF, and passes the outcome to end users in a one-side manner. In this chain, PI AF is the key to unlock the values of PI System. Assets are generally thought of as pieces of equipment like a turbine or meter. But they can also be more abstract like an individual building, an entire power plant or a city where you maintain multiple facilities. PI AF integrates, contextualizes, refines, references, and further analyses data from multiple sources, including one or more PI Data Archives and non-PI sources such as external relational databases. Together, these metadata and time series data provide detailed descriptions of equipment. AF has many featured functionalities, such as templates, analytics, notifications, etc. By using this uniform platform of PI System infrastructure, the enterprise-wide operation can be achieved [3].

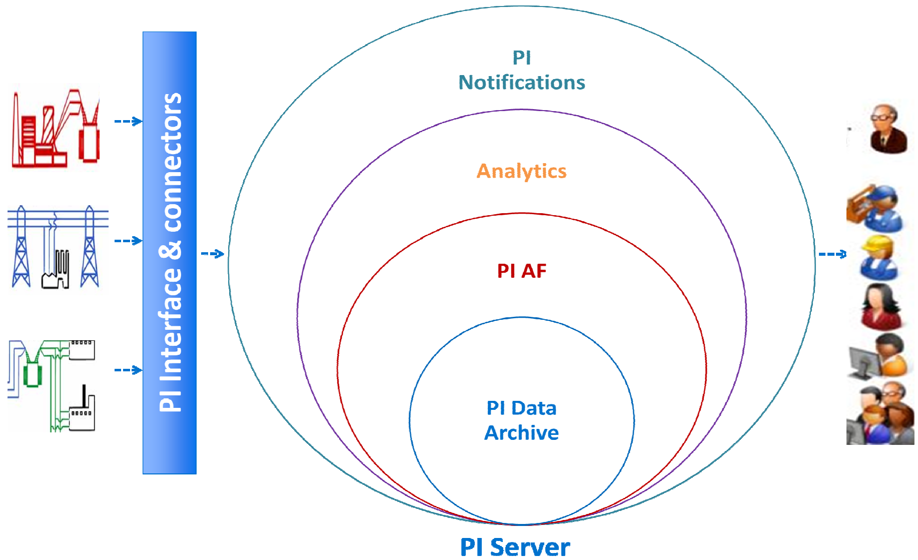

Figure 2. PI system infrastructure [3]

B. Asset Framework

How does AF fulfill these factors and execute data governance through its various functionalities? PI AF is a single repository to organize the data and help with scaling, analyzing and tracking processes. Particularly, power grid data architecture can benefit from this well-organized framework [4, 5]. With multiple views and its hierarchical structure given in figure 3, AF provides an intuitive way for a maintenance engineer to find a Transformer 1 that has overheated or a plant manager to see all the assets he is responsible for. AF also offers the ability to logically map out technical and business operations and those are restricted to one map. For each asset, AF can actually provide context to your data by collecting all of the sensor data from traditional PI points alongside non-real-time metadata. Items like last service time, equipment model or manufacturer, which are vital to keeping healthy assets, are also readily available. It's also easy to create tables with your asset metadata. AF is easy to maintain since it can import and link with tables already maintained to ensure that all asset data is synced to date. Asset data can also be grouped together to create a holistic view of the PI System while eliminating issues with finding data through esoteric tag naming conventions [4].

Templates provide convenience to create or modify the same types of assets in the system. It is arguably the most powerful feature of AF. It can easily standardize operations and will streamline the build-out of existing or new systems. As an advanced option of template design, one is allowed to create layouts that can apply to several types of equipment, or even model out the entire process from start to finish.

Once the assets are organized in an intuitive way, extracting more useful information out of the PI Points becomes simple. The asset Analytics is a built-in engine in AF. It allows users to configure predictive models to set targets for processes, write equations, increase understanding or compare roll-up analysis to investigate which equipment or site is succeeding more than the rest. Since the configuration is straight-forward, users can leverage the metadata to see trends that they might not have insight into before [3, 4].

Another merit of AF is the ability to capture events that happen throughout the operations and raise an alert if needed. Live troubleshooting writing repeatable reports or sifting through huge amounts of data becomes much simpler with AF. Event Frames allow one to track individual shifts, batches or process excursions while notifications can alert an engineer prior to a failure. This guarantees continuous monitoring of the operations and ensures the problems can be identified the moment the system fails. With proper customization, the notification feature can also be exploited as a stage summary report on a scheduled basis.

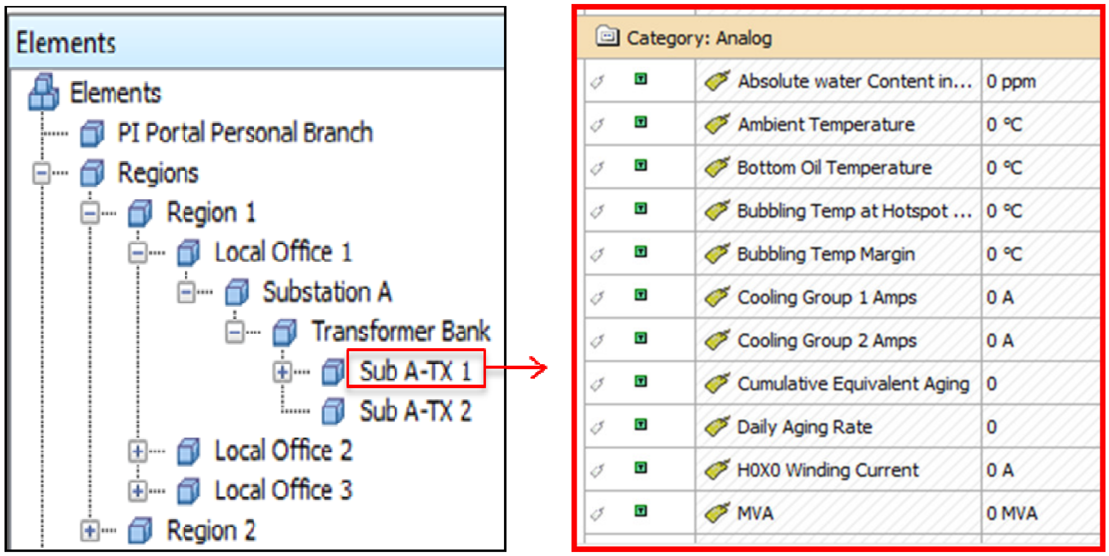

Figure 3. Asset Framework structure

IV. PI DATA ACCESS METHODS

For the purpose of accessing (reading and writing) to the data, AF SDK and PI Web API have been utilized and will be discussed as follows.

A. AF SDK

AF SDK is a library that provides access to data which includes the assets in AF as well as their properties and relationships. The data also contains the time series data from the PI Data Archive and other data sources. In addition, AF SDK is able to access the event frames that record and contextualize process events in AF. AF SDK can be easily accessed from Microsoft .NET languages like C#, Visual Basic.NET and Managed C++ as illustrated in figure 4. It is available for .NET 3.5 and .NET 4.5 Frameworks and designed for both 32-bit and 64-bit Windows operating systems [3]. Since AF SDK is an AF-based library, it presents a hierarchical model of objects. One or more objects can be found, modified or removed based on similar existing objects. Another important feature of AF SDK is its interaction between the PI Data Archive, especially for reading a large amount of data.

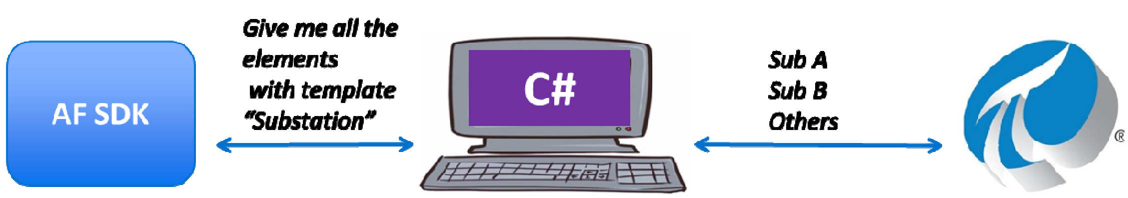

Figure 4. Mechanism of AF SDK

B. PI WEB API

PI Web API is an interface to the PI System in Representational State Transfer (REST) style. REST is an architectural style that defines a set of constraints to be used for creating web services. PI Web API provides basic functionality to retrieve and manipulate time series, asset, and event frame data [3]. In a word, PI Web API allows clients to read and write to their AF and PI data over HTTPS links as shown in figure 5. It includes searching object through its path in AF and querying the object via its WebID. Every primary object in the PI Web API is associated with a WebID, which is a unique identifier. Since the WebIDs are persistent, customers can cache the URLs and save processing time. Likes the other normal Web API, PI Web API supports JavaScript Object Notation (JSON) as the primary media.

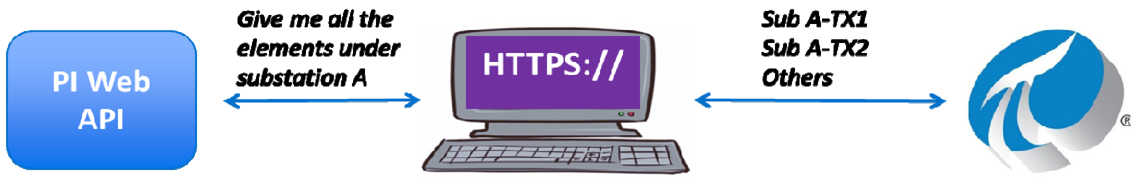

Figure 5. Mechanism of PI Web API

V. DATA VISUALIZATIONS

To demonstrate the applications of the OSIsoft® PI System, two real-time data based maps have been established to visualize the data in Dominion Energy power system.

A. Transformer Monitoring Map

Figure 6 shows the transformer monitoring map. The red question markers represent the abnormal substations which have warnings or alarms. The yellow-dotted markers are the substations which have stayed in normal operating conditions. More details can be found in the right panel. The "Transformer Bank" drop-down menu allows customers to select the transformer which they are interested in. The listed table provides all the information for the attributes of the selected transformer. The red color of the attribute indicates the abnormal condition of the specific attribute. With the aid of this map, the system operator will be able to notify any suspicious operation conditions of the transformers in a system-wide manner. Moreover, the users have an intuitive sense of the status of the whole system.

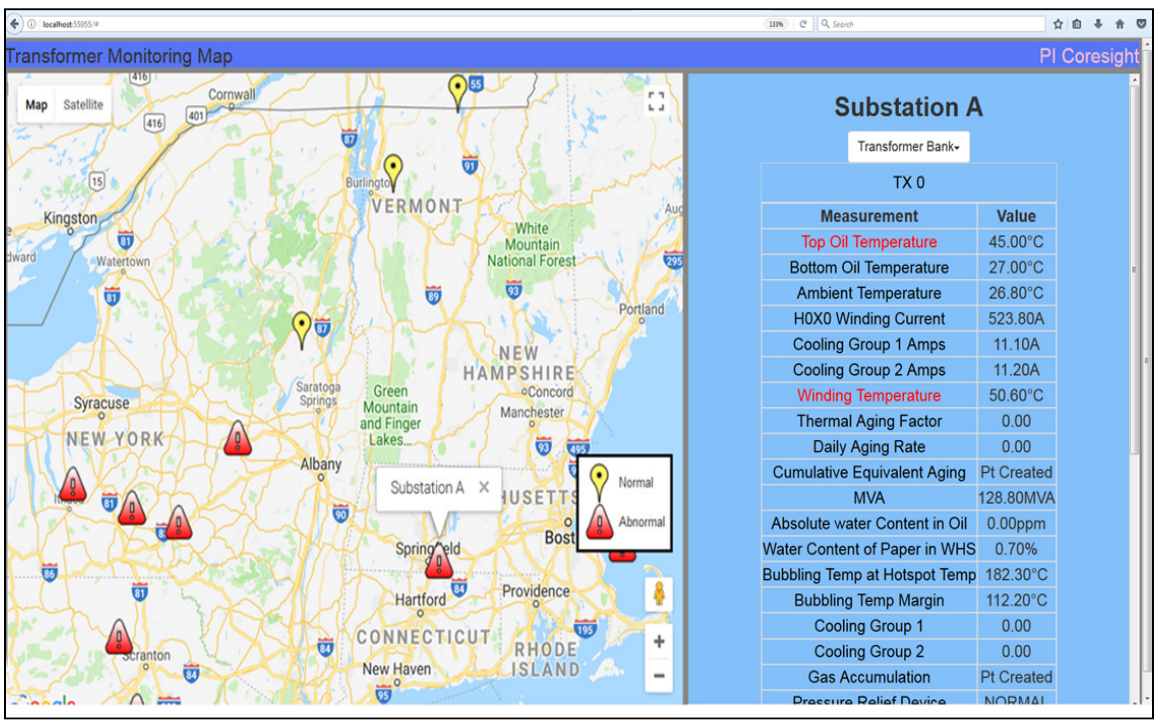

Figure 6. Transformer monitoring map

B. Relay Failure Alarms Map

The relay alarm failure map is shown in figure 7. On this map, only the markers which represent the abnormal substations are shown. Similar to the transformer monitoring map, the right panel gives the users more information about the relay failure alarms. On the top of the panel, two charts are presented. The bar chart shows the hourly increments in the number of the alarms while the line chart shows the total number of the alarms within 24 hours. On the bottom of the panel, a list of the recent alarms is given.

The merit of the map, compared to traditional forms of tables and charts [6], lies in the ability to quickly troubleshooting and estimating system problems based on geological connections. System operators can narrow down and pinpoint the potential fault occurred in the system and identify cyber-attacks by examining the relay map. In addition, based on the acquired data, it is also possible to estimate the extent of damage as well as its spreading rate.

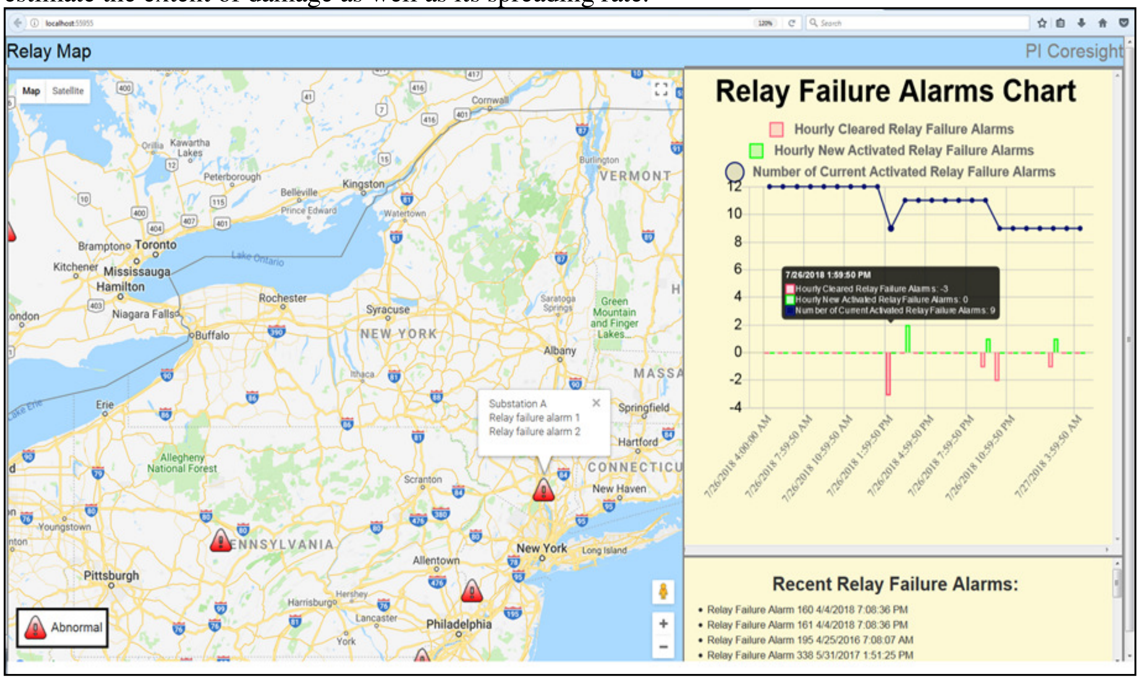

Figure 7. Relay failure alarm map

VI. CONCLUSIONS AND FUTURE WORK

In this paper, based on PI System, two maps have been created with AF SDK and PI Web API using the data stored and organized in AF structure. The transformer monitoring map provides the users the information of the on-line transformers. Users will be able to notify, response and troubleshoot the problems existed in the transformers within several minutes. The relay map is also of great significance since it shows the users an overview of the system status. For instance, if an attempt to attack the power system by switching relays on and off is identified, the system operators will capture this abnormal condition and report to their manager immediately. Besides, different departments have different naming standards for the same equipment. This leads to the difficulty to utilize the data. Therefore, data governance is ongoing. More features will be added on these two maps to achieve customized purpose once the data governance is finished. We expect these efforts can lead to less redundancy, better efficiency, accurate diagnoses and clarified decision for the power system.

BIBLIOGRAPHY

[1] B. Wang, B. Fang, Y. Wang, H. Liu, and Y. Liu, "Power System Transient Stability Assessment Based on Big Data and the Core Vector Machine," *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2561–2570, Sep. 2016.

[2] J. Liu, X. Li, D. Liu, H. Liu, and P. Mao, "Study on Data Management of Fundamental Model in Control Center for Smart Grid Operation," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 573–579, Dec. 2011.

[3] *OSIsoft*, [online] Available: http://www.osisoft.com

[4] Liu, Hesen, et al. "The design and implementation of the enterprise level data platform and big data driven applications and analytics." *Transmission and Distribution Conference and Exposition (T&D), 2016 IEEE/PES*. IEEE, 2016

[5] S. You, L. Zhu, Y. Liu, H. Liu, Y. Liu, M. Shankar, R. Robertson, and T. King, “A survey on next-generation power grid data architecture,” in *2015 IEEE Power & Energy Society General Meeting*, 2015, pp. 1–5.

[6] H. Huang and X. Zou, “Implementation of model-driven visualizations in dominion energy,” in *2017 Grid of the Future Symposium*, (Cleveland, Ohio), CIGRE US National 6 Committee, October 2017.