Renewable and Sustainable Energy Reviews 52 (2015) 937–947

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

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

Big Data issues and opportunities for electric utilities

Beth-Anne Schuelke-Leech [a], Betsy Barry [b], Matteo Muratori [c], B.J. Yurkovich [d]

[a] The John Glenn College of Public Affairs and Center for Automotive Research (CAR), The Ohio State University, Columbus, OH 43212, USA
[b] Illocution Inc., Athens, GA 30605, USA
[c] Pacific Northwest National Laboratory – Joint Global Change Research Institute, College Park, MD 20740, USA
[d] Center for Automotive Research (CAR), The Ohio State University, Columbus, OH 43212, USA

ARTICLE INFO

Article history:
Received 14 March 2015
Received in revised form
23 May 2015
Accepted 27 July 2015
Available online 25 August 2015

Keywords:
Electric utilities
Big Data Analytics
Smart grid
Data management
Policy discourse

ABSTRACT

Advances and innovations are crucial for a sustainable electricity system that includes smart grid technologies, renewable energy sources, and greater energy efficiency. These technologies are often layered on top of the existing infrastructure and legacy information systems. The management and utilization of the data generated from the different components of the electrical system are critical for the successful deployment and operation of this system. This paper reviews the issues and opportunities of the use of Big Data for electric utilities. Big Data provides the opportunity to better monitor, correct, and integrate smart grid technologies and renewable energy. At the same time, data management and utilization must be integrated into organizational operations if the potentials are to be realized. Electric utilities are conservative, heavily-regulated, and concerned with both system reliability and overall profitability. Thus, technological, economic, institutional, and policy constraints must all be addressed. After reviewing these issues and opportunities, we empirically analyze whether these are part of the discussions about electric utilities with federal policymakers. The results show that while conversations about electric utilities overall are plentiful, conversations about data in the context of electric utilities are relatively rare.

Contents

1. Introduction . . . 937
2. Theory . . . 938
 2.1. Electric utilities . . . 938
 2.2. Big Data . . . 938
 2.3. Smart grid and operational analytics . . . 939
 2.4. Data management . . . 940
 2.5. Privacy and security . . . 941
 2.6. Organizational and business analytics . . . 942
 2.7. External analytics: legal, regulatory, socio-techno-economic, and political . . . 942
3. Methods . . . 943
4. Results . . . 944
5. Discussion . . . 945
6. Conclusions . . . 946
References . . . 946

1. Introduction

Future sustainable energy systems will be smart and integrated, including technologies currently under deployment such as “smart grids”, renewable sources, storage, and energy management and monitoring systems. Proper operation of such a complex system depends on coordination of all the components, which

E-mail addresses: schuelke-leech.1@osu.edu (B.-A. Schuelke-Leech), bbarry@illocutioninc.com (B. Barry), matteo.muratori@pnnl.gov (M. Muratori), yurkovich.7@osu.edu (B.J. Yurkovich).

http://dx.doi.org/10.1016/j.rser.2015.07.128
1364-0321/

results in significant challenges in terms of data management. Electronic sensors and communications technologies provide data on energy demand, supply, system performance, and operations. This information is essential for the deployment of smart grid technologies [1,2], effective operation of demand response programs [3] and the integration of renewable energy [4]. It is also necessary to improve fuel efficiency, reduce emissions and energy consumption [5]. On the other hand, inappropriate use of technology and information can actually be detrimental to supply and demand [6]. With the potentials and problems associated with the proper integration and coordination of technologies and information, the management and understanding of the data coming from all these sources is crucial.

This paper explores the potential impacts of Big Data on electric utilities. Electric utilities are confronted with both common and unique opportunities and challenges with Big Data. In the United States, smart grid technologies and advanced communications are layered on top of an aging infrastructure. While organizations typically understand their current operations and markets, many are not prepared to manage and utilize the data that is now becoming available. Big Data Analytics and related modeling techniques are currently being developed and employed with the intention of managing large volumes of data to improve system and company performance. These are not useful unless organizations have the capabilities, capacity, infrastructure, and technologies to incorporate them into daily operations. These capabilities vary across industries and organizations [for example see [7] for a discussion about Big Data and the oil and gas industry].

There is a growing body of the literature about Big Data, Data Analytics, smart grid technologies, and data management. This literature is targeted to both practitioners and scholars and covers both the technical and business aspects. However, the literature is fragmented. Sustainable and renewable energy systems require smart grid technologies, which in turn require Data Analytics. The data from all the sensors and sources in the electrical system can only deliver on the promises of increased reliability, greater efficiency, and improved sustainability if the data is utilized. This requires utilities to take an integrated and holistic view of the data.

In the first section of the paper, we discuss Big Data in general, defining what Big Data is, and discussing the different issues of Big Data for electric utilities. In the next section, we discuss the different challenges and opportunities they each present, as well as the pertinent types of Data Analytics. Specifically, we look at: smart grid and operational analytics, data management, privacy and security, organizational and business analytics, and external analytics. Then, we analyze the way that data is actually used and discussed in the context of electric utilities by employing corpus linguistic analysis on a corpus of U.S. Congressional transcripts and reports from 1981 through 2014. This corpus is used because it an important aspect of the communication between electric utilities and the policymakers that regulate them. In the last section, we discuss the implications of our findings and the role of policymakers and regulators in promoting the use of Big Data Analytics.

2. Theory

2.1. Electric utilities

The US electric utility industry is one of the more mature and stable industries in the US economy. Electric utilities are responsible for the generation, transmission, and distribution of electrical power via the electrical grid. Traditionally, this process has been mostly electro-mechanical, with relatively little need for sophisticated data management. Within the last 30 years, however, the industry has seen massive changes in regards to deregulation, distributed generation, and cogeneration [8], as well as legislative and public pressures to incorporate renewable energy sources and improve energy efficiency and conservation [see for example, [9,10]]. With these changes have come the needs to continually update information technology systems to meet the growing requirements of demand, access, and reliability. This has been accompanied by a push to use the data generated by these technologies.

Electrical systems are extremely complex with an instantaneous need of matching millions of demand requirements with supply. Big Data Analytics and advanced information technologies hold the promise of improved system reliability, greater energy efficiency, and lower costs for consumers. Big Data Analytics allow the massive amounts of data generated by electronic sensors, smart grid technologies, electricity supply, grid operations, and customer demands to be coordinated, analyzed, understood, and effectively utilized.

Specifically, Big Data Analytics can be used to:

- Develop models and simulations of the electrical grid and infrastructure to improve their reliability, resilience, technology adoption, and energy demand management.
- Predict equipment failures and power outages, allowing utilities to optimize their maintenance budgets.
- Improve the operating efficiency of electrical generation, transmission, and distribution.
- Integrate intermittent power sources (i.e., renewables) more efficiently and effectively.
- Help managers, employees, and consumers to make better decisions, founded on data and empirical investigation, rather than on intuition or past-practice.
- Better target and tailor services to different customers.

2.2. Big Data

The terms “Big Data” and “Data Analytics” have entered the business lexicon. Generally, *Big Data* is the term for a collection of data or documents so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications [11]. “Big” encompasses both the volume, the diversity of data, and the speed with which the data is generated (also known as the volume, variety, and velocity of data) [12,13].

Much of these data are readily available for perusal and analysis and can provide substantial insights on electricity generation, grid operations, technological problems, organizational behavior, regulatory compliance, employee morale and behavior, customers concerns and the changing business environment. However, these insights remain unrealized without a methodology for collecting, organizing, and analyzing substantial amount of data. The sheer volume of data available has made it virtually impossible for a single person to be able to analyze and process it effectively.

Data Analytics and data mining techniques have been developed to analyze large repositories of information and data in order to find patterns and trends that are not necessarily obvious. However, there are substantial challenges to discerning actionable knowledge from the data, including: capturing, curating, storing, searching, sharing, transferring, analyzing and visualizing data. In order to achieve their objectives, Big Data methodologies need to be applied appropriately and with an understanding of both the technical issues and the broader implications.

In order to understand the potential role of Big Data for electric utilities, it is first necessary to define “smart grid” and “soft grid.” The term *smart grid* was first proposed by Vu and colleagues [14] who suggested that to maintain the reliability of the electrical grid, an advanced and “smart” protection and control scheme is needed, which integrates local measurement devices with a centralized