Big Data Applications in the Energy Sector: A Review of the Current Status

Oğuz Kıranlar
kiranlaroguz@gmail.com

Instituto Superior Técnico, Universidade de Lisboa, Portugal

December 2017

Abstract

This master thesis provides a comprehensive overview of Big Data analytics applications in the energy sector, with a special focus on the opportunities for Iberdrola. Data volumes have been growing exponentially in recent years. Penetration of new technologies contributes to the generation of data in various structures. Once analyzed, this large and complex data can create business value for organizations. The field that is dedicated to tackling the size and complexity of data is referred to as Big Data. The energy sector holds a vast amount of data received from each part of its value chain. Further, it is undergoing a transformation where the systems are getting more decentralized, customer demands are growing and regulatory pressure on mitigating climate change is larger than ever. These trends make Big Data analytics a promising tool to adapt to changes while creating new business opportunities. To determine the potential of Big Data analytics, value chain of the energy sector was divided into parts according to their relevance to data sources: generation, transmission & distribution network, flexibility management and customer relations. The activities of electricity utilities and value propositions of companies that use Big Data to create products & services were analyzed. Further, other sectors that are similar to energy in terms of data size and structure were investigated. If applicable, business models that are derived from these applications were identified. Lastly, a business case suitable for Iberdrola was elaborated by using business model canvas and SWOT analysis methods. The results show that real-time visibility of the assets together with predictive analysis enables optimal operations both for generation and transmission & distribution network. The operational costs are driven down by better asset performance management and enhanced field worker productivity. For renewable energy generation, weather-related uncertainties are reduced which leads to the increased adoption of these technologies. It was found that applications in flexibility management create various business opportunities. Big Data analytics is used to aggregate and operate distributed energy resources in near-real time. Meanwhile, it is used to provide services for residential, commercial and industrial end-users which leads to involvement of more DER assets in flexibility management programs. These operations are either run by utilities or aggregators. From customer relations standpoint, utilities convert their customer related data into energy insights to provide services for their customers which leads to a better engagement. Based on various applications identified, *Iberdrola Efficiency Shop*, a business idea that converts smart meter data into energy advice to drive customers to purchase more efficient home appliances, was elaborated. The case study showed the potential of Big Data analytics to create a new line of business for utilities.
Keywords: Big Data analytics, Electric utilities, Energy sector

1. Introduction

Penetration of technologies such as sensors, smart phones, online shopping and social media resulted in an exponential growth in data generation. A simple daily life activity such as a post on social media or credit card use is enough to produce data about location, finance, sales, etc. It is estimated that 2.5 quintillion[1] bytes of data is being generated daily and the growth pace is so rapid that 90% of the data in the world was generated only in the last 2 years [1]. International Data Corporation (IDC) estimates that global data volume will reach 163 zettabytes (ZB)[2] by 2025, nearly 10 times the 16.1 ZB of data generated in 2016 [2]. This vast amount of data, through analysis, becomes valuable information, which can be converted into knowledge to gain new insights. This value has been acknowledged by businesses which led to a high interest

[1] 10^{18}

[2] Equivalant to 10^{21} bytes

towards Big Data[3].

The energy sector is undergoing a tranformation. The declining cost of ownership for distributed energy resources (DER)[4] with technological progress, is enabling a higher market penetration. Incentive-based policies contribute to the adoption of DER's as well. An increasing number of DER's is changing customers into prosumers[5]. This situation challenges utilities to meet a growing number of customer choices and demands while continuing their businesses. Higher number of DER's, together with promoted energy efficiency causes a market erosion for traditional utilities. Moreover, having a large amount of distributed power producing and consuming assets makes the grid unstable, complicating the asset management. From a regulatory standpoint, some innovative regulators see the opportunity to allow new players in the sector, such as aggregators[6], bringing more competition in the sector. To tackle climate change, policy makers put a pressure on the sector to cut the green house gases (e.g. EU Emissions Trading Scheme, COP21, Energiewende) and set a target to increase the share of renewables in the generation mix. Growing environmental concerns and competition make utilities more vulnerable to public opinion. Lastly, emerging technologies such as IoT and smart grids, smart meters and devices enable two way communication, increasing connectivity, controllability and automation of the entire value chain of the sector. Based on the trends above, the traditional utilities who are loyal to business models from the 20$^{\text{th}}$ century, will be disrupted during this transformation. A survey conducted by PwC with 70 utilities shows that 97% of the participants are expecting a disruption in their business by 2020 and 73% anticipate a major business model change by 2030 [3]. Big Data will be a major decisive factor to define the winners of this rapid change in the sector. However, due to its infancy stage, Big Data is not truly embraced by the sector. According to Accenture Technology Vision, only 28% of the businesses think that they are creating value out of their data while 40% accept that they need a plan to implement it [4]. Big Data will play a key role in the following years, and therefore having a comprehensive overview of the Big Data applications in the energy sector related to business intelligence will be very helpful for utilities. From an academic perspective, Big Data applications in electricity sector are still relatively unexplored. There are several overviews, featuring the applications of Big Data in the sector. However, their scope do not go beyond certain sections throughout the value chain (e.g. smart grid, distribution grid). Moreover, they do not present a holistic approach, including the companies, business models and current activities of utilities.

Iberdrola is a Spanish based utility, serving over 32 million customers in Spain, the UK, the USA and Brazil. The company was ranked the world's fourth largest utility in terms of R&D&I[7] investments [5]. Iberdrola Ventures-PERSEO is the venture capital program of the utility. They aim at fostering innovation in the energy sector by giving investor support to entrepreneurs. Therefore, the team is always up to date about new technologies. The opportunities of Big Data analytics for utilities drew attention by Iberdrola Ventures-PERSEO and they wanted to have a comprehensive overview of the technology, specifically its operational and business opportunities for utilities. In light of what have been said, the purpose of the study is to identify the current applications of Big Data analytics for each part of the energy sector value chain.

1.1. Research Methodology

A research methodology was structured as shown in figure 1. It is a general to specific approach, starting with reaching out to as many resources as possible such as academia, reports published by companies and institutes, white papers, press releases, on-site venue visits, direct contact with company representatives, company websites, etc. Analysis of company activities is the essence of research methodology. It reveals the real business value of Big Data analytics since it contains information about latest updates in the sector. It also leads to identification of innovative business models together with the applications. Hence, utilities, companies which leverage data to create products & services in the energy sector and other similar sectors were investigated and their activities were converted into applications. To create a holistic overview, the entire value chain was divided into groups according to their relevance to Big Data analytics: generation, transmission & distribution network, flexibility management and customer relations. Later, these applications were grouped accordingly. If applicable, business models enabled by these applications were revealed. To emphasize the business value of Big Data analytics, a business case which uses Big Data analytics was elaborated by using business model canvas and SWOT analysis techniques.

[3]Big Data is written with capital letters along the paper to emphasize the phenomenon, not only the large amount of data.

[4]It is a comprehensive definition for entire resources connected to distribution grid (e.g. solar PV, small, batteries, CHP and so on.

[5]A consumer who can also produce.

[6]An actor in the sector who aggregates a load capacity from DER's and offers it in the energy market.

[7]R&D&I stands for research, development and innovation.

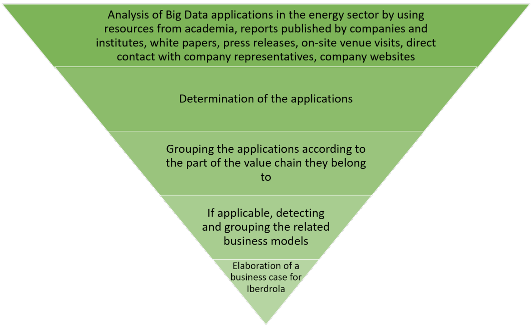

Figure 1: Research methodology: From general to specific approach

2. Big Data

Big data is still in its infancy stage of development, and thus there is not a consensus about the definition of the term. In general, it can be defined as an extremely large amount of structured or unstructured data that is not possible to manage with conventional data management techniques. However, this definition does not capture all the aspects of Big Data. Therefore, a 5V concept which includes volume, velocity, variety, veracity and value was proposed in literature. Volume is associated with the cumulative sum of the data stream which can reach up to zettabytes. Velocity corresponds to the rate of data used to support the interactions and generated by interactions. As the applications which use data become more complex, the velocity of data interaction increases. Consequently, data cannot be stored without processing. For instance, IBM's hospital monitoring application collects and analyzes more than 100.000 data point per patient per second [6]. Variety refers to differences and incompatibility between the data types, sources and structures. The data format can be in the forms of documents, emails, text messages, audio, images, video, graphics data, and others [7]. Veracity describes the degree of reliability of the data. Provenance, quality and noisiness of the data determines the veracity [8]. Provenance is the tracking and recording of the original data. Ideally, using the original data is preferable as it improves the accuracy of processing. Data quality is about objectiveness of data. Many resources such as social media contain subjective data and processing subjective data might result in meaningless conclusions. Lastly, noisiness of data corresponds to uncertainties in the datasets. For instance, the labels of datasets might be inaccurate, and thus degrades the veracity [8]. Value is the monetary worth added to an organization by implementation of Big Data [9].

2.1. Big Data Analytics

Big Data analytics can be divided into four parts: data generation, data acquisition, data storage and data analysis [10]. Data generation and acquisition cover the exploitation of raw material while data storage is needed to keep the data until it is used in the analysis part, where data is converted into value. From the objective perspective, entire Big Data analytics can be categorized as descriptive, predictive and prescriptive analysis [11]. Descriptive analysis aims at determining the reasons of incidents by exploiting historical data. It is commonly used in business intelligence. Predictive analysis refers to capturing future trends and possibilities by predictive modelling. Lastly, prescriptive analysis focuses on decision making by observing a system.

3. Generation

3.1. Wind Energy Generation

In wind industry, Big Data analytics is used for better resource assessment and operational excellence which leads to reduced costs. Precise placement of a wind turbine has a big impact on its performance. The windiest location doesn't necessarily mean the best choice to erect a turbine. Turbulence is a decisive factor for the service life of the turbine since it strains the components. Besides, having the precise information of a location allows virtual layout of the wind farm before the investment is made. This proactive approach allows the right configuration of the turbines to be obtained such as tower height and rotor diameter. Vestas, a global wind turbine producer, combines meteorological data from 35.000 weather stations with the company's historical operational data from 50.000 wind turbines worldwide to obtain a more precise windmap. The size of the data is approximately 2.8 petabytes and includes 160 factors such as temperature, barometric pressure, humidity, precipitation, wind direction and velocity elevation, topography, and satellite images [12, 13]. The company was able to reduce its wind data grids from 27x27km to 10x10m by using the IBM InfoSphere BigInsights Big Data analytics tool [12]. Most of the faults give the early signals before they lead to a breakdown. The signal can be a change in the vibration or a deviation in the temperature. The challenge is to capture these anomalies for each component of every wind turbine in real-time. Big Data analytics allows a better asset management by predictive maintenance. Siemens succeeded detecting a bearing failure one year prior to the deadline of replacing it [14]. It results in a reduction in unplanned downtime of the turbines, and thus an increase in the availability; prevention of costly emergency repairs by detecting problems early; reduction in unnecessary routine maintenance which introduces risks and decreases availability; reduction in inventory costs since decision-making on spare parts storage and logistics is based on failure rates.

Once the operations are more visible in near-real time scale, daily activities in a wind farm, and the work plans can be automatized. The unpredictable nature of wind is limiting its value. The industry has been developing weather forecast models to overcome this uncertainty. Big Data analytics is capable of increasing the accuracy of the forecasts with a longer lead time. HyREF, developed by IBM, is a software which analyses large volumes of sensor-collected weather data and instrument-collected power data in order to forecast power output of a wind farm with high accuracy. It is able to estimate power generation with over 90% accuracy up to four days lead time [15].

3.2. Solar Energy Generation

In solar industry, non-hardware related costs resulting from customer acquisition, installer overhead, financing, contracts, inspection, permitting, interconnection, and installation labor are called soft costs, adding up to 33% of the total cost [16]. Big Data analytics is able to reduce the soft costs by offering smarter marketing strategies, more precise preliminary investigation of the projects and facilitating the financing for solar developers. Being a weather dependent energy source, solar projects hold uncertainty. The risky nature of the investment results in higher interest rates. Moreover, unlike wind, solar projects are mostly developed by smaller size companies which don't have performance guarantees. Gathering data from more than 70.000 solar projects on its platform together with its risk management software kWh Analytics, a US based start-up, aims at being the trustful 3rd party to investigate project performance for the investors. This will decrease the interests on solar project loans by increasing the credibility of the investment. The location of a solar farm determines its profitability due to the distance and conditions of the distribution grid. Kevala, a Big Data analytics company, maps the node and substation locations and ranks them according to their profitability by taking into account the energy value derived from annual hourly solar production through weather data and historical wholesale prices. This helps investors to narrow their site search down to more profitable locations. The company provides these services by gathering various publicly available data such as utility filings, outage reports, permit applications, visual inspection and historical wholesale pricing data [17]. Another company, Geostellar, decreases inspection costs by eliminating the need of preliminary site investigation. It automatizes the process taking into account weather data, the size, orientation and slope of a roof; and shadows from trees, structures and terrain. Later, a financial calculation is conducted by using utility rates, energy usage profiles and energy incentives [18]. The accuracy of the method was verified by National Renewable Energy Laboratory (NREL). 76 % of the calculations were matching with the on-site performance tests results [19]. PickMySolar is a web based market place for solar panel installers. The company allows the home owners to compare and choose the best contractors only by managing the platform. This allows the company to access age, gender, household income, financial status, credit scores, building details data. Later, the application PVimpact leverages data obtained from PickMySolar and combines it with smart meter data from the households. The application enables home owners to have an easy access to their savings, receive alerts and recommendations about their system. PowerScout, a startup from the US, uses light detection and ranging (LIDAR) data with detailed consumer data and advanced image recognition technology to determine households that would adopt solar PV systems in the close future. Each home is tagged with 1.200 data points, including income and education levels, political affiliation, the type of the resident and the car [20]. This insight cuts off the marketing expenses.

Big Data analytics is also used to enhance the operational performance. NexTracker, a US based company, proposes a tracking system, leveraging cloud positioning, fog or haze data to correct the panel direction by using machine learning algorithms. This quick change in the panel position increases the annual energy output up to 6 % by preventing the unnecessary losses [21]. Alextrics, a Greek based company, combines the monitoring data with the financial and business operation data of the operator to deliver actionable information. For instance, the software can observe a low internal rate of return and connect this observation to a lack of performance due to initial bad design of an asset [22].

4. Transmission & Distribution Network

The biggest responsibility of a utility is to ensure that the grid infrastructure is functional at all times. However, this is a major operational challenge since the grid consists of numerous distributed components[8]. Monitoring the assets in real-time gives information about the current status of the grid. This information is then converted to optimized work plans by predictive maintenance, a digital worker concept and asset planning. For instance, Enel, the biggest Italian utility, predicts the failures in feeders in 16.000 substations by collecting data from SCADA, maintenance work orders, fault protection, asset management, historical equipment failures, known network issues, power quality, light-

[8] transformers, cables, towers, circuit breakers, etc.

ning, terrain and vegetation, and weather [23]. This actionable knowledge can be used to empower field workers with mobile devices in which optimized work plans are present. At this point, natural language generation is used to increase productivity. The application enables complex datasets to be converted into easy-to-understand texts. A US based utility implemented the application offered by Arria[9] to improve the response time to an outage. The technology automatically creates work orders to repair the assets in near-real time by learning from previous outages [24]. Further, analysis of the grid can optimize the appropriate size of a component for a certain area to achieve a better asset planning. Big Data analytics can also be used to manage outages. Weather Company, a subsidiary of IBM, can detect weather related outages with a 72 hours lead time by analysing data about wind speed, precipitation, temperature, atmospheric pressure, humidity, soil moisture, and foliage vegetation together with asset condition (location and/or health)[25]. Lastly, electricity theft can be detected with grid analysis. Light S.A., a Brazilian utility providing electricity to more than 10 million customers in Rio de Janeiro, used Choice's[10] *Revenue Intelligence* service to tackle their theft problem. As a result, 160 GWh of energy, which corresponds to 40 million dollars were saved annually [26].

5. Flexibility Management

Flexibility of the grid refers to the degree of balancing the generation and consumption ability at all times. There are two approaches to increase the flexibility capacity. The first approach is demand response (DR), which is a set of programs that motivate residential, commercial and industrial (C&I) consumers to reduce or shift their electricity consumption during peak periods [27]. However, Big Data analytics takes it to the next level by enhancing predictability and controllability. The second approach comprises of aggregating DER's in order to create a load capacity and offer it to energy markets.

5.1. Utility Focused Big Data Applications: DRMS

Big Data analytics, in combination with IoT, enables aggregation and visibility of data. Data gathered a from utility's advanced metering infrastructure (AMI), meter data management system (MDMS), and legacy systems such as SCADA, customer information systems (CIS), customer billing data, weather data, and geographic information systems (GIS) can be used to optimize and manage DR programs from one center. Such systems are called demand response management systems (DRMS).

[9] A British based natural language generation company.

[10] Luxemburg based analytics company.

5.2. End-user Focused Big Data Applications: EMS

Energy Management Systems (EMS) focuses on cost effective energy efficiency solutions for end-users. Intelligent efficiency, a term highly linked to EMS, is defined as the ability to save energy by gathering large volumes of data [28]. It consists of all products and services that monitor, control or analyze energy consumption. These solutions leverage data from traditional control and automation systems, smart meter interval electricity consumption data, supplemental submeters and advanced sensors. The energy efficiency applications are provided by controlling end-user's applications such as heating, ventilation and air conditioning (HVAC), lighting and plug loads citation as needed.

5.3. Findings from Related Companies and Case Studies

Based on the applications above, two opportunities become prominent. Firstly, end-users receive greater information about their energy consumption, as well as tips to reduce their energy use. Secondly, the actionable knowledge received by analytics can be used by automated systems to optimize energy use. Needless to say, having an intelligence control over the assets makes for smooth integration into DRMS systems, and allows them to be valuable DR assets. Therefore, many companies in this field offer energy management solutions to residential and non-residential end-users to engage them into a DR program. Figure 2 presents a comprehensive overview of the business relationship between the DR providers, end-users and utilities. The term DR provider refers to companies who use Big Data analytics in order to provide DR services to utilities either by turnkey sales of firm capacity and ancillary services, implementation of DRMS or an energy efficiency SaaS which connects end-users to utilities. To better understand the subject, the providers are grouped into two categories depending on their customer segment (Industrial and commercial focused and residential/behavioural focused DR providers). This categorisation is necessary since not all the DR providers serve entire electricity end users.

The business models of all DR providers have similarities to a certain degree. They aim at the highest possible end-user acquisition rate which would turn into load reduction for utilities. They are positioned as a mediator between utilities and the end-users, controlling the data/information and monetary flow. However, there are fundamental differences in the way they achieve it, particularly in the methods of customer engagement. Moreover, the improving Big Data analytics abilities bring new players to the market with different business models and value proposals. The first group, C&I focused DR providers are the largest and the most mature

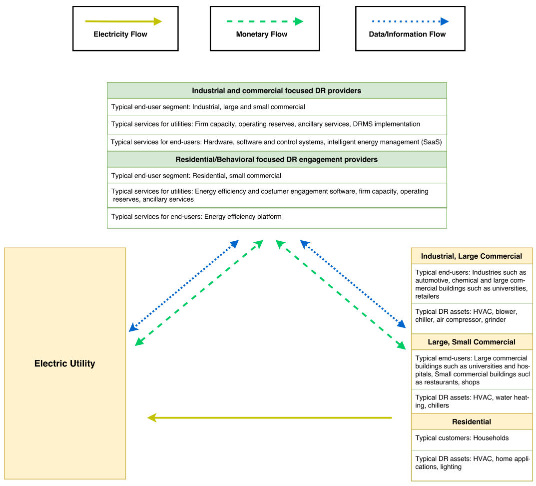

Figure 2: Business relation between the DR providers, electric utilities and end-users

group, consisting of some companies traced back to early 2000s. There are mainly three different business models inside this group. The first business model provides intelligent energy management systems for industrial and large commercial end-users by software that optimizes their businesses energy consumption. This software gives the businesses the option to participate in DR programs and the provider sells the firm capacity and ancillary services to utilities as a turnkey solution without any further interference. Their revenue streams are subscription fees for their intelligent EMS and the brokerage fee that they agree with the utility for the load reduction. The second business model is provided by DRMS providing companies. Their revenues are directly linked to utilities since their service does not involve end-users like the previous group. The revenue streams are either software sales or subscription fees. The companies who produce hardware such as HVAC systems are also active in the field. DR is a very advantageous business opportunity to them as DR participation of a HVAC system increases its value proposal, and this works both ways. Therefore, many companies offer an intelligent commercial/residential EMS system which can be embedded in utilities DRMS to increase the DR participation and customer engagement rates. These companies make revenue with software sales of either energy efficiency SaaS, or entire DRMS on top the hardware sales.

The second group, residential/behavioural DR providers, are the new players in the market. Their value proposal is similar. They provide an energy efficiency platform for residential end-users which gives actionable information about their energy consumption and tips to reduce it. It encourages end-users to participate in DR programs through behavioural change. The reliability of DR is provided with learning algorithms which conduct an analysis to predict the participation of every house. They receive revenues through software sales to utilities.

5.4. Utility Focused Big Data Applications: DERMS and VPP

A distributed energy resource management system (DERMS) is a software-based solution that enhances utility's real-time visibility into the DER assets. It is highly linked to DRMS in terms of functioning. The main difference is that DERMS has a higher scale since it deals with more assets connected. A virtual power plant (VPP) is a system which includes the smart aggregation and optimization of DERs in order to create a dispatch capacity that is comparable to a conventional centralised power plant. VPPs, leveraging Big Data analytics and smart grid, can accommodate the complexity of DERs, providing regulation service, voltage management, fast DR, contingency reserve, peak demand management, and renewable integration [29].

5.5. End-user Focused Big Data Applications: Intelligent Batteries

End-user focused Big Data applications include activities which monitor, optimize and control the DER assets in order to reduce energy related costs. The DER assets suitable for these applications are intelligent batteries and solar-plus-storage co-optimization solutions. There is a strong similarity between the energy management systems and these applications. They both aim at providing solutions to end-users while they are used, as a tool to increase the customer engagement for utilities. The main difference is that when EMS optimize the assets with an switch on/off feature, end-user focused applications optimize batteries by charging/discharging. This allows end-users to avoid electricity consumption during the high electricity prices, and therefore brings about savings.

5.6. Findings from Related Companies and Case Studies

The value proposition of DER related business models are similar to DR. They both help utilities manage the flexibility of the grid to avoid grid related problems, reduce the generation costs, avoid unnecessary investments and engage with prosumers. However, as mentioned earlier, DR assets only reduce the load simply by switching off when necessary while other DER assets actually provide power to the grid. This fundamental difference results in different business models to evolve in time. There are mainly three groups of models who provide DER management services. The first group includes the companies who provide DERMS and VPP for utilities by a SaaS or a on-premises solution. The second group of companies dealing with DER management by using Big Data analytics are firm capacity providers for wholesale electricity markets. Instead of offering their digital solutions to utilities as a key partner, they aggregate a firm capacity (i.e. VPP) and participate in power markets. The third group consists of companies which provide end-user focused Big Data applications through intelligent batteries. They use the relation that they have with the end-users to create a portfolio of assets which can be sold to utilities as firm capacity, operating reserves and ancillary services. The companies, being the mediator between the utility and the end-users receive revenues from both parties. They don't necessarily sell the products to end-users but they may also finance them through leasing and loaning. This decreases the initial investment and makes the assets more affordable. Depending on the agreement, they also receive revenues through the bill reduction their system manages. The revenues related to utilities come from direct sales of firm capacity, operating reserves and ancillary services.

6. Customer Relations

Big Data analytics converts electricity into a bundled service which consolidates the bond between the utilities and customer through personalization. Further, the analysis of the data brings more customer insights to utilities and makes customer operations easier. The starting point of reaching customers through bundled services is accurately estimating the characteristics of end-users' buildings. The algorithms relate the energy consumption of buildings to the interactions between weather, resident behavior and schedules, building characteristics, and individual devices present in the building [30]. Detecting the individual devices in the building through consumption data is provided by disaggregation[11]. This interconnected calculation of energy consumption captures the physical processes that drive the energy use and enables the actual suggestions that would have an influence on consumption. For instance, these algorithms can associate the energy saving potential of a building by insulation, or changing the windows, changing the thermostat setting and so on. Moreover, as one of the data inputs is generated by end-users, the more they are used the more accurate they get with machine learning algorithms. The algorithms are able to accomplish a task such as *"Identify residential consumers who spend more than 40 percent of their energy on HVAC on days when the temperature is above 30 degrees; who own their homes; have incomes below 30.000; and live in the North Eastern United States."*[30]. Having such a specific insight about the customers, their buildings and energy use enables opportunities for utilities. Firstly, customers who are most receptive to participating in utility programs or reacting to efficiency tips when they are offered can be identified. These "better recognised" customers are then proposed the most relevant automated tips through recommendation engine which can take place visually either on a web portal, mobile app or the bills. The same technique can be applied to promote the best utility program to the customers as well. Home energy reports are another value that can be added to individual building insights. They are detailed energy consumption reports which empowers customers to make decisions on their energy use. Utilities are also able to increase the customer satisfaction by providing notifications about unexpected high bills before the issuing date. A notification to a customer who is on track to a high bill can also be supported by personalized tips to decrease it.

[11] It is a set of algorithms which separate the use of energy into appliances.

6.1. Findings from Telecom and Banking Sectors

Customer churn rates are considerably high in the telecom sector, ranging between 20-40% [31]. From a customer experience standpoint, the period of time which leads to the churn point is unpleasant, and thus a prediction this dissatisfaction is valuable. To achieve this, data collected on customer usage, complaints, transactions, social media is analysed and the customers who are on track to leave the company are determined. As the analysis gives insights about the reasons of dissatisfaction, personalized offers can be made to the customers. For instance, a customer who considers leaving the company due to his old and expensive contract can be offered a new contract with more advantageous tariffs.

Banking sector is using customer data to create personalized loyalty programs to retain customers. The method is based on bringing more benefits to the customer as long as they use the banking services. Analysing the transactional data, several banks determine the spending habits of their customers and create personalized loyalty programs. For instance, The Royal Bank of Canada offers promotions in pet shops to their customers who have pet related transactional data in their systems. Nowadays, the sector is experiencing a new concept: Open Banking, which is based on banks sharing their customer data with third parties in order for them to create new services. Such a partnership enables development of Big Data applications by not only the banks themselves but any developer who has a good idea. Some highlighted services created by third parties are personal finance management platforms, which help users to track their income and expenditure. The data obtained from different banks can be combined to give insights about customer's spendings. Further automated advice that rely on Big Data analytics can guide the users about investment products available in the market. Another customer related application is transparent credit scoring. Traditional credit scoring is limited in terms of capturing the real performance of a customer due to the limited data. A better analysis of individuals is possible by combining bank-owned data with third party data. A more transparent credit scoring enables a better match between the customers and the available products for them. Lastly, shared customer data allows the banking services to be accessible in different platforms. For instance Amazon Echo, which is a hands-free speaker that connects to internet to access Amazon enabled services and more. The partnership between Capital One, a bank operating in the US, and Amazon allows customers to access banking services from their device with their own voice [32].

7. Elaboration of a Business Case: Iberdrola Efficiency Shop

The analysis of Big Data applications in the electricity sector shows that customer oriented applications are the most suitable for creating a new line of business. *Iberdrola Efficiency Shop* is an idea that positions Iberdrola as a trusted energy advisor for customers' energy consumption behaviour and efficient home appliances choices. The idea provides two main services to the customers. The first service is called *Iberdrola Energy Advisor*. It makes the energy consumption of the customers more transparent to them. It is based on the applications mentioned in the 6^{th} chapter. A disaggregation of energy use data is conducted to identify the electricity consumption at an appliance level. This information in real-time is then utilized to create the platform which is available as a web portal and a mobile app. The platform provides many features that increase the engagement such as energy home reports, pro-active high bill notifications, personalized tips on energy reduction, utility offers according to customers consumption habits, a benchmarking with the similar buildings' energy consumption and connection to social media accounts that allows customers to share their energy related achievements. The platform also includes *Efficiency points*, which is a loyalty program that motivates customers to interact with the platform. Lastly, *Ask your personal energy advisor* tab, a search engine located in the upside of the screen at all times. It simplifies the app which is heavily filled with information. For instance the customer can type "How is my TV doing?" and a energy consumption comparison of the TV with the best available models in the market will show up. The pulling point of the idea is to connect this platform to an online marketplace, *Iberdrola Efficiency Shop*, to motivate its customers to buy energy efficient products. To achieve this connection successfully, the platform reaches out to the customer at the right moment, it is easy-to-use and it incentivizes the customers to use it more. Once, the customer is directed to the *Iberdrola Efficiency Shop*, it is possible to compare the the products according to their energy efficiency, price and the customer reviews collected from different retailers.

7.1. Business Model Canvas of Iberdrola Efficiency Shop

The value proposition of the idea targets both customers and Iberdrola. It empowers the customers by helping them lowering their monetary losses related to inefficient use of energy without any discomfort. It acts as an advisor for environmentally concerned customers who need guidance to lower their carbon footprint. Moreover, it creates an alternative to the channel of purchasing new home appliances. Lastly, the loyalty program creates tangible benefits such as discounts and special offers as long as they are interacting with the platform and the marketplace. The most tangible outcome of the idea for Iberdrola is the new revenue stream created by *Iberdrola Efficiency Shop*. Secondly, being the energy advisor boosts the customer engagement and makes Iberdrola the favourable energy provider in an open market. The loyalty program, *Efficiency Points*, acts as a tool for Iberdrola to drive people to achieve company objectives such as efficiency targets, promoting new utility programs, creating company awareness through social media posts, reducing customer churn and peer-to-peer engagement. Lastly, being the operator of such a platform, Iberdrola pioneers a pro-environment movement which brings about a positive public opinion about the company.

To realize the idea, a partnership with white label marketplace and energy insights platform providers is necessary. The key activities are launching of the platform and the marketplace. Once the technology is ready, a marketing campaign to create awareness must be run. Lastly, a continuous improvement of the platform and the marketplace according to objectives of Iberdrola must take place. A in-house team dedicated to operate the business must be constituted. The team should have competences in IT technologies, business development and marketing. The realization of the business requires several fixed and variable costs. The initial fixed costs are salaries of the team members and marketing. Depending on the key partner's business model, a subscription or a commission fee might be paid proportional to the volume of sales on the marketplace. The direct revenue stream is the commission fees for the products sold in the marketplace. An indirect revenue is created by increased utility program participation rates.

7.2. Operational & Financial Planning

Operational and financial planning is made for 2 years and divided into 3 stages. These stages differ from each other according to the targeted market and actions to take. Therefore, revenues and costs change in every stage. The first stage aims at implementing the idea in Spain, and lasts for 3 months. The objective of this stage is to expand *Iberdrola Efficiency Shop* to the UK, which is the second country Iberdrola is operating in Europe. This stage is planned as 9 months. Lastly, an expansion to the US market is planned during the 3rd stage, which takes 12 months. For Spain and the UK markets the target is to reach 3% of the Iberdrola customers while 1.5% is expected in the US.

7.3. Cumulated Cash Flow Diagram

Cumulated cash flow diagram shown in figure 3 is a summary of financial projections. The anticipated break-even point is around 20 month, which proves that the idea is viable. It must be noted that the most striking value proposition of the idea is qualitative benefits for both customers and Iberdrola as mentioned above. Thus, a larger marketing campaign which would increase the customer engagement could be planned even though this would shift the break-even point to later months.

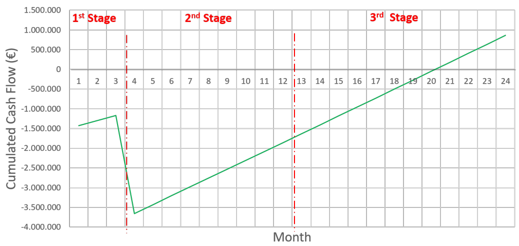

Figure 3: Cumulated cash flow diagram

8. Conclusions and Recommendations

The results show that Big Data analytics applications either to improve their operations or to create a new line of business. From the generation perspective, the applications enhance the adoption and performance of renewables. The applications related to wind energy generation are focused on reducing the costs by providing better forecast and operational excellence. Big Data analytics is either implemented by turbine manufacturers as a turnkey service directly, or utilities develop it in-house or with an external partner. In the meantime, solar energy is highly linked to customer collaboration and adoption. Therefore, many applications are derived to promote its adoption by decreasing financing and installation costs. This creates new business model opportunities for utilities. There is a strong link between smart grid and transmission & distribution network applications. In this section, all of the applications aim at reducing the costs related to operation of these assets. Better monitoring and control over the grid also allows a more optimized asset planning. Further, outages related to severe weather conditions can be foreseen to plan a proactive management. Lastly, a more precise electricity theft investigation to protect revenues is possible for the regions of the world where theft is a major issue. Flexibility management is one of the most important challenges of modern distribution systems. Big Data analytics is transforming this challenge into business opportunity by various applications. Having the highest number of related business models detected, this section involves many actors such as energy intelligence companies, DR providers, VPP operators, energy management companies, smart home appliances manufacturers, etc. The field is very active in North America and some countries in Europe (France, the UK, Germany, the Netherlands, Belgium). Customer relations applications offer one of the highest values for utilities. This is mainly due to change in the customer expectations and customer-centric response to this change by utilities. Having an individual and personalized contact with the customers by using their data provides benefits to both sides. Customers can receive recommendations not only about their energy use but also their home assets. The marketplace link of this personalized knowledge empowers them to optimize their home related purchases. Such service scales up to projects as well. They can decide on an upgrade for their buildings (windows, roof, solar adoption, etc.), according to its energy related benefits. From a utility standpoint, reaching to the customers at the right time with the relevant content positions them as an energy advisor and bonds the customer engagement. Findings from other similar sectors show that creating bundled services with data increases the business volume and ensures the prevention of customer churn.

8.1. Recommendations

Based on the conclusions above, this section provides recommendations for Iberdrola. As elaborated in the 7th chapter, Iberdrola could use its customer data to become an energy advisor and a marketplace operator. The personalized level of communication is a strong motivation for customer engagement and creating business out of it. In parallel, a project marketplace can be launched to provide building upgrades to customers. The open banking concept described in the 6th chapter could be adopted from banking sector. Positioning Iberdrola as an *App-Store* operator would involve independent developers to create services for customers where every interaction in the store returns to Iberdrola as revenue and brand awareness. Iberdrola Smart Solar initiative can be supported by solar energy generation applications. Firstly, tailored marketing services that use demographic and LIDAR data can decrease customer acquisition costs. Further, installation design can be managed off-site, without the need of field inspection. The programs can also involve a contractor marketplace for installation and maintenance to empower the customers. Lastly, partnering with a third party financial risk analyser company (see kWh Analytics) can decrease the financing costs of customers. Flexibility management is essential in the new era of energy. As mentioned in the 6^{th} chapter, utilities in the US and Europe position themselves as energy management partners of C&I and residential customers. In return they involve their customers in demand side management programs. Iberdrola already provides various products and services at energy management. Implementation of a DERMS would position Iberdrola as a VPP provider and help the company balance the grid while creating a new business line. However, this opportunity is highly linked to legislative decisions. Natural language generation is a very promising application. It could cut the costs of grid operations by instantly converting data into simple texts and work orders. Brazil is a problematic field of operation of Iberdrola in terms of electricity theft. A collaboration with a theft analytics company (see Choice) could mitigate the problem.

This study shows that Big Data analytics has already proved itself to be a trustful tool to manage the paradigm shift in the energy sector. It is used by various actors ranging from multinational companies to garage start-ups. However, it can still be considered at its infantry stage. There is still a lot of space for new applications and business models. Therefore, innovative utilities should examine the phenomena carefully and concentrate their R&D&I activities on the topic.

References

[1] Paul Zikopoulos, Dirk deRoos, Christopher Bienko, Rick Buglio, and Marc Andrews. *Big Data Beyond the Hype.* McGraw Hill, 2015.

[2] David Reinsel, John Gantz, and John Rydning. Data Age 2025 - The Evolution of Data to Life-Critical: Don ' t Focus on Big Data ; Focus on the Data That ' s Big, 2017.

[3] PwC. 14th PwC Global Power & Utilities Survey. Technical report. URL http://www.pwc.com/cl/es/publicaciones/assets/14th-pwc-global-p-u-survey-web-12-05-15.pdf.

[4] Accenture. Accenture Technology Vision 2015, 2015.

[5] Iberdrola. Iberdrola, the world's fourth largest utility. URL https://www.iberdrola.com/top-stories/iberdrola-shares-with-you/iberdrola-supports-innovation.

[6] Paul Zikopoulos, Dirk DeRoos, Christopher Bienko, Rick Buglio, and Marc Andrews. *Big Data Beyond the Hype.* New York, 2015.

[7] B. Furht and F. Villanustre. *Big Data Technologies and Applications.* Springer International Publishing, 2016. ISBN 9783319445502. URL https://books.google.com.tr/books?id=5n8XDQAAQBAJ.

[8] Alexandra L'Heureux, Katarina Grolinger, Hany F. ElYamany, and Miriam Capretz. Machine Learning with Big Data: Challenges and Approaches. *IEEE Access*, pages 1–22, 2017. ISSN 2169-3536. doi: 10.1109/ACCESS.2017.2696365. URL http://ieeexplore.ieee.org/document/7906512/.

[9] Hai Wang, Zeshui Xu, Hamido Fujita, and Shousheng Liu. Towards felicitous decision making: An overview on challenges and trends of big data. *Information Sciences*, 367-368:747–765, 2016.

[10] Min Chen, Shiwen Mao, and Yunhao Liu. Big data: A survey. *Mobile Networks and Applications*, (19):171–209, 2014. ISSN 1383469X. doi: 10.1007/s11036-013-0489-0.

[11] Gavin Blackett. The OR Society: Special Interest Groups - Analytics Network. URL https://www.theorsociety.com/Pages/SpecialInterest/AnalyticsNetwork_analytics.aspx.

[12] Vestas Wind. Vestas. *2011.*

[13] Pittman DAvid. Lords of the Data Storm: Vestas and IBM Win Big Data Award — IBM Big Data & Analytics Hub, 2012. URL http://www.ibmbigdatahub.com/blog/lords-data-storm-vestas-and-ibm-win-big-data-award.

[14] Ulrich Kreutzer. From Big Data to Smart Data: Remote Maintenance - Digitalization & Software - Pictures of the Future - Innovation - Home - Siemens Global Website, 2014. URL https://www.siemens.com/innovation/en/home/pictures-of-the-future/digitalization-and-software/from-big-data-to-smart-data-remote-service-wind-turbines.html.

[15] IBM Research - China: Smarter Energy Solution Suite. URL http://research.ibm.com/labs/china/smarter_energy.html.

[16] Kristen Ardani, Daniel Seif, Carolyn Davidson, Jesse Morris, Sarah Truitt, Roy Torbert, and Robert Margolis. Preliminary non-hardware ('soft') cost-reduction Roadmap for residential and small commercial solar photovoltaics, 2013-2020, 2013. ISSN 01608371.

[17] By Ethan Sprague. Three Ways Data Will Accelerate Solar Electricity Development in the Next Five Years. pages 1–10, 2015.

[18] Geostellars Big DataSolar Power Geomatics, 2017. URL https://de.scribd.com/document/53298510/Geostellar-Big-Data-Solar-Geomatics-1-0-041611.

[19] National Renewable Energy Laboratory. Geostellar : Remote Solar Energy Assessments Personalized. Technical report, 2013. URL https://www.nrel.gov/docs/fy16osti/63615.pdf.

[20] PowerScout Secures $5.2 Million In Funding For Machine-Learning Clean Energy — CleanTechnica, 2016. URL https://cleantechnica.com/2016/09/27/powerscout-secures-5-2-million-funding-machine-learning-clean-energy/.

[21] NEXTracker Announces TrueCapture, a Revolutionary Intelligent Control System for Solar Power Plants — NEXTracker, 2017. URL https://www.nextracker.com/.

[22] Alectris Solar Asset Innovation. Puglia , Italy Deeply Distressed PV Plants Revived by Alectris Resulting in Dramatic Increase of the Solar Portfolio's Profitability, 2010.

[23] C3 IoT. Enel and C3 Energy's Machine Learning-Based Predictive Maintenance — C3 IoT. URL https://c3iot.com/enel-and-c3-energys-machine-learning-based-predictive-maintenance/.

[24] Arria. Automate expert reports anytime anywhere.

[25] The Weather Company. Proactive preparation and response to weather-related power outages. URL https://business.weather.com/products/outage-prediction.

[26] Choice. Success Story: Light S.A. URL http://choiceholding.com/light-s-a/.

[27] Demand Response — Department of Energy. URL https://energy.gov/oe/activities/technology-development/grid-modernization-and-smart-grid/demand-response.

[28] Ethan A. Rogers, R. Neal Elliott, Sameer Kwatra, Dan Trombley, and Vasanth Nadadur. Intelligent Efficiency: Opportunities, Barriers, and Solutions, 2013. URL http://www.aceee.org/sites/default/files/publications/researchreports/e13j.pdf.

[29] Peter Asmus. Stacking Value with Virtual Power Plants in Today's Digital Power Grid Moving Distributed Networked Energy into the Mainstream, 2017.

[30] Tendril. The Tendril Physics-based Building Model White Paper, 2014.

[31] Deloitte. Opportunities in Telecom Sector : Arising from Big Data Descriptive, 2015.

[32] Capital One. Capital One is on Amazon Echo. URL https://www.capitalone.com/applications/alexa/.