ESMAP
Energy Sector Management Assistance Program

Knowledge Series 027/17

Energy Analytics for Development

Big Data for Energy Access, Energy Efficiency, and Renewable Energy

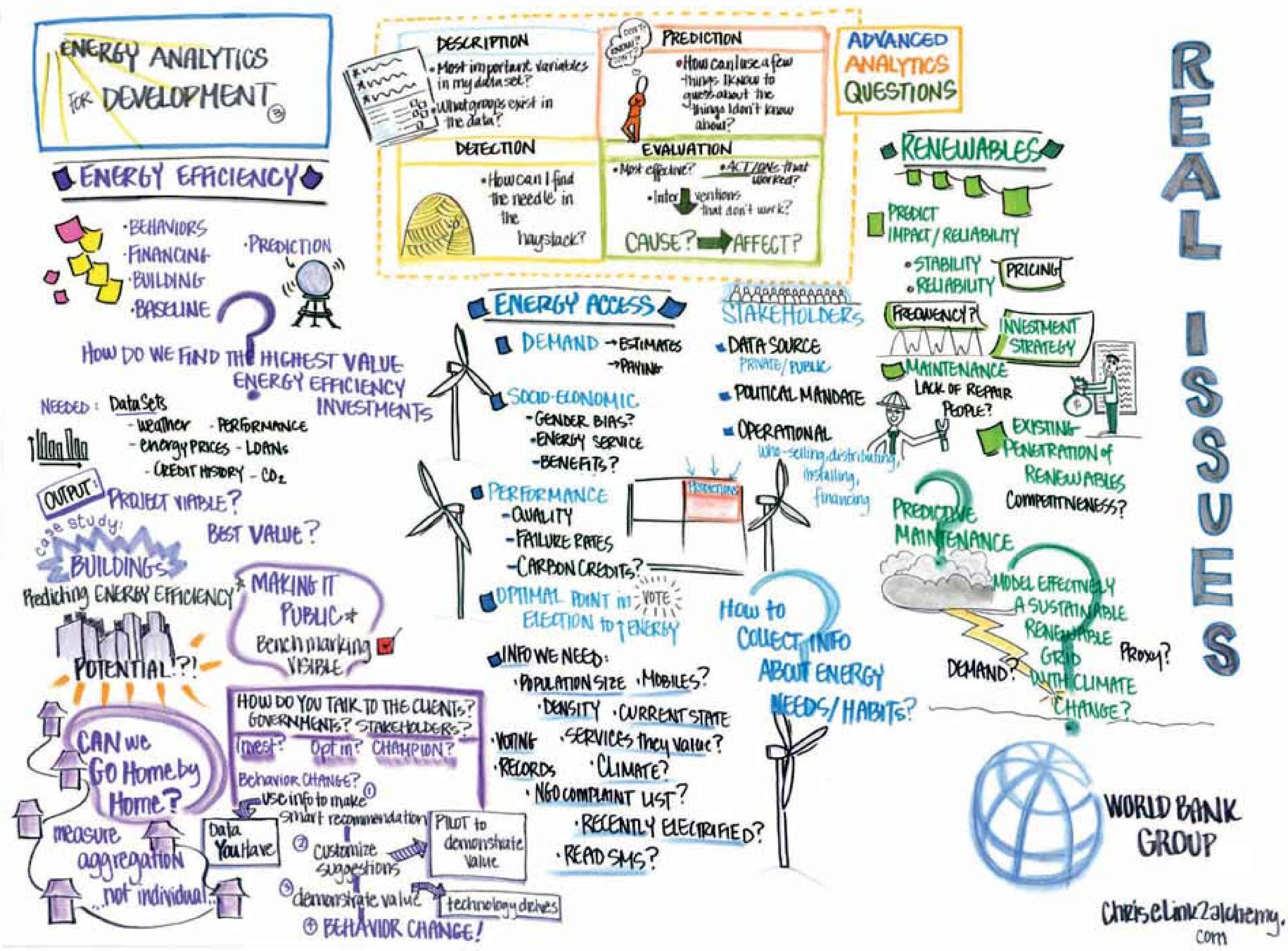


TABLE OF CONTENTS

WHAT IS BIG DATA?

With unprecedented speed and scale, digital transformation is having significant impact on multiple industries, including the energy sector. The digital agenda in the energy sector is being driven by a combination of technologies, collection and use of data, and a more complex world demanding greater agility, speed, and digital competences. It is expected to impact all aspects of the energy sector, including changing patterns of consumption, new ways of optimizing assets, cross-industry partnerships, and the greater use of industrial platforms. Digital technologies and 'smart solutions' are being placed at the center of new business models and data is seen as an increasingly valuable asset.

Reliable and affordable power is a fundamental necessity in achieving the twin goals of the World Bank Group—shared prosperity and elimination of extreme poverty by 2030. Recognizing this, the World Bank Group co-chairs the Sustainable Energy for All (SEforALL) initiative, and is committed to working toward the Sustainable Development Goal 7 on energy:

1 | Universal access to electricity and clean cooking fuels;

2 | Doubling the rate of improvement in energy efficiency; and

3 | Doubling the share of the world's energy supplied by renewables from 18 to 36 percent.

Efforts to support the implementation of the Sustainable Development Goals will require increased investment, new technologies, and, in some cases, policy and institutional reform. Broad applications of digital technologies will be key given the transformational ability of digital services and systems.[1] Specifically, as outlined in the *World Development Report 2016: Digital Dividends,* more targeted use of data and advanced analytics can play an essential role in helping World Bank Group (WBG) client countries meet their energy targets.[2]

The transformational impact of digital technologies will not be covered in detail in this solutions brief on Big Data. However, given the interrelation between data and Internet-of-Things (IoT), the technology merits a word of introduction. IoT is the inter-networking of physical devices and items embedded with electronics, software, sensors, actuators, and network connectivity that enable these objects to collect and exchange data. Thanks to the lowering costs and increasing ability of computing power, data storage, and connectivity bandwidth utilization, IoT is becoming more mainstream. The benefits of IoT include applications to collect, track, and communicate data in ways previously not possible and this is a core source of Big Data.

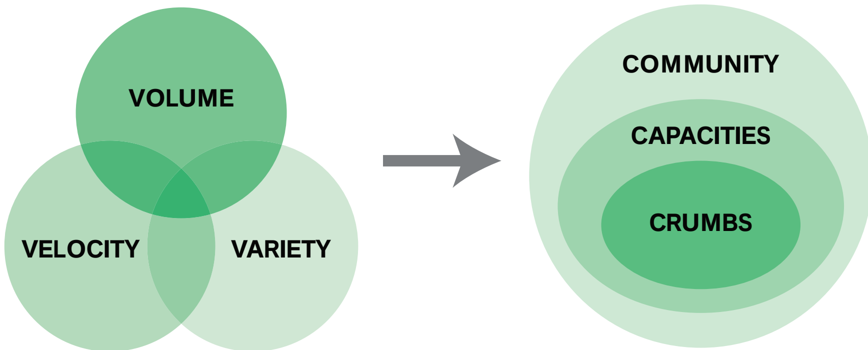
The shorthand "Big Data Analytics" refers to the attempt to extract actionable insights from streams of information that are large, heterogeneous, and fast-moving (Figure 1). During the last decade, the tools for dealing with such data—from distributed computing to machine learning and data mining—have matured to the point where they are readily usable in commercial settings.[3] Machine learning refers to the study of pattern recognition and computational learning theory, and explores the construction and study of algorithms that can learn from and make predictions on data.[4]

To be forewarned, as digital technologies and data become more integrated in energy infrastructure (used to control energy production, transmit information about consumption, and monitor demand), the energy sector will grow increasingly vulnerable to cyber

Figure 1 | What Is Big Data?

In 2010, Gartner, an information technology research and advisory company, described Big Data as high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery, and process optimization. In more recent descriptions, the Big Data ecosystem is characterized as the union of 3 Cs: Big Data Crumbs, Capacities, and Communities.[i]

EVOLUTION OF THE DEFINITION OF BIG DATA



Circa 2010: the 3 Vs of Big Data

Now: the 3 Cs of Big Data

The importance and prevalence of Big Data has risen as more activity has moved online, where Big Data is a natural by-product of each interaction. Big Data is often labelled as exhaust (data generated as trails or information by-products resulting from all digital or online activities), web-based (data generated from the Internet), or sensing (data collected by a sensor that is not in direct contact with the area being mapped) data. Crowdsourcing is a fourth division that can be considered as a category of its own. In the energy industry,[ii] some relevant sources of Big Data include:

- Smart meters
- Sensor networks
- Customer payment information and credit histories
- Web server logs
- Cell phone call detail records
- Satellite imagery

[i] Letouze, Emmanuel. *Big Data and Development: An Overview.* World Bank Group, 2014.
[ii] Ibid.

threats. To ensure reliable energy delivery and services, a cyber-resilient energy delivery system is needed.[5] As such, cybersecurity functions (hardware, software, institutional organization, and human capacity) must be carefully engineered as to not interfere with energy delivery and service functions.[6] Cybersecurity is often considered an IT/software issue, but security is only as strong as the weakest link, which is often employees and executives who are not adequately trained in security threats.

The purpose of this solutions brief is to explore ways in which Big Data Analytics can be better integrated in the energy sector in emerging and developing markets. As stakeholders consider applications of Big Data Analytics throughout the project cycle, it is important to understand what type of questions are amenable to this approach, as well as to identify what type of questions are better addressed through traditional means of data analysis and project design. This solutions brief presents overviews and case studies of four general categories of problems where advanced analytics techniques are especially useful: *Description, Prediction, Detection,* and *Dynamic Evaluation* (Figure 2). These four categories of analytics techniques lead to the identification of classes of energy sector problems where big data analytics can be especially helpful. This solutions brief is structured around three categories of energy projects:

1 | **Energy Access.** Supporting efforts to broaden electricity access, and ensuring the economic and operational sustainability of these efforts for energy service providers and their customers

2 | **Energy Efficiency.** Incentivizing and scaling energy efficiency investment

3 | **Renewable Energy.** Adapting the electrical grid and the energy market to the demands of distributed storage and renewable generation

Figure 2 | Classes of Energy Sector Problems Paired with Big Data Analytics Techniques

Description *describes patterns or characteristics in a sample group* • What patterns or trends can be detected in energy use data from a city's buildings? • What community attributes relate most strongly with program outcomes? • Are nighttime satellite images useful in identifying gaps in energy access? • What are the patterns or trends in power quality throughout the grid? • What energy usage patterns are observed within a community or test group?	**Prediction** *predicts a pattern or an impact within a sample group* • Which appliances will be used in a specific set of homes if we offer incentive x? • How much energy will this city use tomorrow? • Which buildings have the biggest energy efficiency potential? (prescription question) • How much sun or wind will these buildings get in the next year? • How will energy consumption change in this community given an external climate shock?
Detection *detects specific patterns or attributes within a sample group* • Where are the biggest power quality issues on the grid? • Which building's energy use does not fit the norm? • Where are there unexpected changes in use that might indicate power theft? • Where and when is something unexpected happening?	**Dynamic Evaluation** *evaluates which interventions contributed to change* • Which information intervention or campaign works best in reducing energy use? • Which educational campaigns drive adoption of off-grid lighting? • Were the savings predictions made during planning phase accurate (and not) and how can they be more accurate next time? • What survey format is most effective in getting people to respond?

Source | Presentation by Matt Gee, University of Chicago at "Energy Analytics for Development" workshop at the World Bank Group in May, 2015.

In each category, we discuss how the outputs of Big Data Analytics can be used to improve the development and implementation of policy. This solutions brief focuses less on the particular technology used to collect and analyze data, and more on identifying the relevant problems and the value of the potential solutions.

ENERGY ACCESS

Over a billion people still lack proper access to electricity and almost three billion people use solid fuels—wood, charcoal, coal, dung—for cooking and heating. Lack of energy access severely impedes quality of life, delivery of social services, and economic development. The recently adopted Sustainable Development Goal 7 proposes to ensure access to affordable, reliable, sustainable, and modern energy for all by 2030.[7] To achieve this ambitious target, a range of technology and business models must be deployed, including improved collection, storage, and analysis of data.[8] Big Data Analytics has a supporting role to extend access, both in evaluating and optimizing the deployments of new technologies, and in designing incentive and pricing structures that are informed by the needs and behaviors of consumers and energy service providers.

Least-Cost Planning with Energy Analytics

To reach universal access to electrical power, planning and evaluating technology options for meeting present and future electricity demand is crucial. Least-cost planning methodology (sometimes referred to as integrated resource planning) is a process by which energy ministries or electric utilities evaluate the costs, benefits, and risks of different resources for expanding energy access and meeting electric power demand (including traditional power plants and energy efficiency) to arrive at the mix of resources that will meet future demand at the lowest cost while still providing reliable electric service. A national least-cost electricity plan must deal with some of the main challenges

Energy Access Highlights

- Energy access can be expanded in a faster and more cost-effective manner with more detailed understanding of energy end-use needs related to specific communities, including potential impact of innovative business models and incentive structures.
- Unconventional data sources such as satellite photography, money transfers, and mobile telecommunications use can assist in predicting demand, assess creditworthiness of underserved customers, and can be combined with climate and weather data to assess competing technologies.
- Least-cost planning exercises can be increasingly data driven, benefiting from the growing amount of interrelated data generated by digital technologies in the energy industry. New data sets and improved analytics can give better predictions of demand growth in regions targeted for electricity access expansion, including visualizing urbanization trends and offer more detailed socioeconomic customer data by triangulating various data sources. Leveraging synergies with connectivity access programs would also benefit electricity planning.
- Working to regularize customers (reducing nontechnical losses from electricity theft), advanced analytics can help set up systematic real-time experimentation and monitoring. This can improve the design of incentives and information campaigns to make them more effective tools to encourage changes in consumer behavior and detect—in real time—how customers are responding to different information campaigns and behavioral nudges.

to expanding energy access, including the difficulty of building distribution infrastructure for remote and sparsely populated regions, the limited ability of utilities and potential customers to pay up-front costs for conventional and off-grid solutions, and the necessity of finding a sustainable business model to serve low-income areas. Detailed energy consumption and socioeconomic data are often scarce in remote or low-income areas. Today, the expansion of information and communication technologies (like satellite imagery, mobile phones, and mobile payment systems), has made available an increasing amount of interrelated data that can be integrated in energy planning models. These data sets can help *describe* characteristics of regions targeted for energy access expansion—for example, by offering more detailed energy access data from satellite images, deriving alternative creditworthiness scores and customer stickiness indicators from mobile payment history, and indicating user mobility and urbanization trends from call detail records, but also to help make *predictions* about future energy demand that, in turn, will help select appropriate energy service alternatives and tailored pricing models for these regions.

Description of the electricity access problem has two components: identifying underserved populations, and estimating their connectability via existing assets. Once the description is in place, *predictions* of cost and likelihood of project success can be used to help compare competing technologies. In particular, the choice often comes down to main grid versus mini-grid or off-grid solutions, the latter of which the International Energy Agency estimates will reach 70 percent of those that currently lack access to power.[9] Wind and solar modeling can quantify the potential renewable capacity in a given area, external data sources such as satellite imagery and mobile phone records can be combined with historical project performance data to more accurately forecast demand, and behavioral economics models can help assess the financial sustainability of proposed payment and incentive structures.

Interestingly, many of the technology and business model alternatives to main grid expansion have integrated Big Data Analytics and Internet-of-Things as part of their core business offerings. For example, distributed solar companies in East Africa are leveraging increased mobile telecommunications penetration to establish a variety of pay-as-you-go business models, such as lease-to-own financing or solar-as-a-service. These products have generated significant energy access in a relatively short time, reaching around 800,000 households in about 5 years. The World Bank's Lighting Global Program, together with Bloomberg New Energy Finance, estimate in their *Market Trends Report 2016* that about one in three off-grid households will use off-grid solar power by 2020.[10] The core drivers of this quick uptake are the integration of data and embedded sensors to offer real-time energy consumption and predict maintenance needs, digital payment options (predominantly mobile money but some companies are also experimenting with blockchain) as an alternative to costly cash collections, globally available and affordable battery storage capacity, and the advancement and increasing affordability super-efficient appliances that enable off-grid PV systems to offer more energy services at lower prices.[11]

Incentives and Pricing Structure

Sufficient access to energy can remain a challenge even for those that have a grid connection. The two most pressing concerns are affordability and reliability, especially at peak hours when the grid may be overwhelmed. Fortunately, the use of Big Data and advanced analytics can transform the existing energy distribution market and ease these challenges and concerns, specifically through improved modeling of incentives and pricing.

Using advanced analytics, utilities can afford to customize service for each customer by modeling which pricing scheme best meets their needs and ability to pay, size of a subsidy, how to price expected future consumption, and which response to expect for a given price differential. The payment and service plans that emerge as optimal, in some cases, may have more in common with mobile telecommunications than traditional monthly utility bills.[12]

Prepaid service is another area of innovation. It offers flexibility in payment schedules, allowing customers with irregular income patterns (e.g., participants in the informal economy) to pay for what they use when they have the resources. In addition, a prepaid service can also increase energy conservation by making people more aware of the actual cost of energy and of the pace of electricity consumption.

The cost and technical expertise required to implement flexible payment plans, however, can be a constraint for many emerging market players. Support from international development partners can be crucial in overcoming this. For example, financing the needed hardware and software upgrades can make an impact, as well as analytical support to select the most appropriate geography and customer mix for pilot projects, engaging with and evaluating the effects on customers after the testing phase. Using

Box 1 | Using Detection and Prediction Analytics to Expose Electricity Theft in Jamaica

Nontechnical losses are a common problem for utilities in developing countries. In some parts of Sub-Saharan Africa, particularly Nigeria, utilities capture less than 25 percent of revenue owed.[i] A large portion of these losses can be caused by informal grid connections by customers with a limited ability to pay. However, bigger consumers, like large commercial customers or industries, can also find workarounds when enforcement and societal norms are weak.

This is also the case in Jamaica. In an effort to stop power theft by large commercial customers, Jamaica Public Service Company, Ltd. (JPS), power provider, has installed advanced metering infrastructure (AMI) meters for approximately 6,000 of its largest customers. However, JPS's system for monitoring these AMI meters for irregularities was highly labor-intensive.

In order to address this problem, the World Bank Group partnered with JPS and a data science consultancy called The Impact Lab to develop an automated detection system for loss-impacting irregularities in large accounts. The Impact Lab team worked with JPS to obtain a set of historical data and inspection results, created a set of predictive variables from the AMI meter output, and built a machine-learning model that assigned a risk score to each account based on the likelihood of an inspection uncovering nontechnical losses.

Before, JPS inspections uncovered irregularities of 5 to 7 percent. With the predictive model, inspections were able to achieve more than 50 percent accuracy on cross-validation. When tested on the ground by the JPS detection team, the approach more than doubled the accuracy in uncovering irregularities. A very positive development for the utility.[ii]

[i] Antmann, Pedro. *Reducing Technical and Non-Technical Losses in the Power Sector.* World Bank Group, 2009.
[ii] JPS targets electricity-stealing commercial customers. *Jamaica Observer,* June 17, 2015. Accessed on June 29, 2015.

advanced analytics to identify and respond to problems like *detecting* nontechnical losses or electricity theft can be another impactful strategy.

Deploying and Sharing Data

While alternative service models and new technologies may improve reliability and affordability, data sharing can further improve energy access. For example, advanced metering infrastructure (AMI) provides a technical means to share information between consumers and utilities. As developing countries continue with grid modernization and utility reform initiatives, upgrading their systems with digital technologies as well as introducing smart meters and other sensors, the volume of data and the opportunities they present will increase. For example, IBM Research-Africa is developing software applications to model rural electrification strategies and to *predict* potential economic and social benefits. Such data-driven tools, built on a mix of open and big data, could assist local governments and donors when designing energy access expansion programs.[13] The Earth Institute at Columbia University has also undertaken significant geospatial planning exercises over the past 10 years in several countries of Africa, Asia, and Oceania, many of them related to World Bank activities.[14] The Earth Institute recently implemented a World Bank-supported program in Nigeria, piloting geospatial least-cost planning in selected distribution zones while ensuring that the sector master plan is developed and updated in coordination with the distribution plan.

These types of products require institutional and regulatory structures to allow third-party operators—as well as researchers—to obtain access to the data. Regulators must also have access to the data necessary to optimize tariffs and subsidies, which they may choose to *evaluate dynamically* with experimental evidence to ensure that policy priorities are being met in a cost-effective manner. Drawing from experience in OECD countries, the most effective frameworks are justified with analytics, and both the data and the methodology are publicly available. Public data releases must respect the privacy of citizens while still being sufficiently detailed to enable accountability.

The World Bank recently launched an open data platform providing access to datasets and data analytics that are relevant to the energy sector. ENERGYDATA.INFO has been developed as a public good available to governments, development organizations, private sector, nongovernmental organizations, academia, civil society, and individuals to share data and analytics that can help achieve the Sustainable Development Goal 7. In addition, building on the Open Data Readiness Assessment (ODRA),[15] the World Bank Group has developed a dedicated framework to collect, share, and use open energy data. Primarily aimed for national or subnational institutions, the Open Energy Data Assessment[16] tool offers a list of quick wins to establish an open energy data ecosystem. To date, it has been piloted in Nairobi and Accra.

ENERGY EFFICIENCY

Energy efficiency saves money, reduces emissions, and can be improved incrementally with modest up-front investments. For this reason, efficiency is often called the "first fuel."[17] Major economic powerhouses such as China, India, and the Middle East are expected to account for 36, 15, and 8 percent (respectively) of energy usage growth over the next two decades, while the United States will account for just 2 percent.[18] As such, it might be more beneficial to facilitate energy efficiency investments through digital technologies and data in high-growth markets. Of the three major areas discussed in this solutions brief, energy efficiency is perhaps the one area with the greatest need for

Energy Efficiency Highlights

- Energy efficiency investments have very different paybacks across diverse environments and cultures. Using analytics, energy and cost savings can be calculated systematically and incorporated into program design, so that the market can determine the optimal investment strategy.
- Consumers are increasingly requesting actionable data on their energy use. Data-driven outreach campaigns can be used to empower consumers to make smarter energy decisions, and new technologies can be used to respond to these signals automatically.
- For larger-scale institutional energy users, large arrays of networked sensors and devices ("smart buildings" or "smart cities") can generate the data and take the actions necessary to achieve very significant savings while having low impact on operations.

advanced analytics. It requires coordination between utilities, devices, and customers, and payoffs cannot be measured directly. Instead, increasing efficiency is an intrinsically data-intensive proposition, where results must be predicted and assessed using statistical models.

Lighting often offers the highest payoff when it comes to energy efficiency interventions.[19] However, the best efficiency project designs vary from region to region based on climate, economy, and culture. Past attempts at mandating or incentivizing particular solutions are gradually giving way to the notion of setting up a marketplace that rewards results. New alternatives have emerged to the energy service company (ESCO) model—the traditional approach of investing through capital expenditures by enlisting third parties to deliver efficiency via operational expenditures. Parties can range from traditional efficiency utilities to technology providers like Apple, Nest, or FirstFuel. For example, India has seen energy efficiency software companies like Ecolibrium Energy that fund high-technology improvements via operational expenditures rather than capital expenditures (Figure 3).

As efficiency moves to a new outcome-based *dynamic evaluation* paradigm, advanced analytics play a critical role in unlocking its full potential. This is primarily because efficiency potential and efficiency gains cannot be directly measured, only calculated. As a result, data analytics is essential to *describing* efficiency opportunities, *predicting* realized savings, and dynamically *evaluating* technical and behavioral interventions for relative cost effectiveness at scale.

Finding efficiency potential across homes and businesses has historically been an expensive endeavor, requiring lengthy audits performed by specialized technicians. However, new applications of machine learning that combine building energy use data with building characteristic data are beginning to reduce the costs of efficiency targeting. Several companies in North America and Europe have made strides toward providing actionable descriptions of energy usage to consumers. Opower (USA), for example, offers personalized information about usage, relevant benchmarks, and advice on saving energy. Rainforest Automation (Canada), Bidgely (USA), and Powerley (USA) all sell devices to help customers monitor their usage, identify high usage appliances, and take individual actions or offer automated solutions to increase energy efficiency or reduce consumption and costs in other ways. In fact, combining smart meters with home energy monitors, or other tools to visualize energy savings, offers 11 to 14 percent additional energy savings compared to a smart meter alone.[20]

Figure 3 | Intelligent Energy in India

Ecolibrium Energy is a leading provider of Energy Analytics for industrial and commercial enterprises in India. Their flagship product, SmartSense, converts energy data into intelligence using advanced machine learning algorithms. The Indian company uses ICT to monitor, analyze and optimize energy consumption in industrial and commercial enterprises. Apart from the industry, the company also has indigenous smart-grid solutions for government and private utilities. With more than 450 active customers and 1,400 nodes being monitored in real time, Ecolibrium's data analytics platform analyses over 100 million data points on a daily basis. This has helped save, on an average, of 5 to 25 percent of power consumption, with a payback period of under 8 months across multiple industry verticals such as iron and steel, pharmaceuticals, textiles, fast-moving consumer goods (FMCGs), automobiles, etc. Ecolibrium Energy was recognized as one of the winners of the 2017 Ashden Awards.

Source | Ecolibrium Energy. http://www.ecolibriumenergy.com

Companies like Opower (USA) and Agentis Energy (USA) have worked with utilities to identify efficiency opportunities for targeting utility-run efficiency programs. Cities across the globe have begun to use open data portals to release community-level energy intensity and load profile data, as well as building-level energy benchmarking data that allow private sector actors to develop sophisticated targeting models for marketing energy efficiency products and services (Figure 4).[21] The emergence of a set of standard methods, data formats, and means of third-party access would accelerate this process; organizations like the World Bank Group could use their leverage as funders to incentivize standards adoption.

Figure 4 | Map of Open Data Portals in Chicago

Cities such as Chicago have begun to aggregate and release energy usage data, enabling companies and nonprofits to zero in on the areas with the most efficient investment potential.

Source | Chicago Energy Data Map. Accessed on June 25, 2015. http://energymap.cityofchicago.org

If efficiency gains are calculated and distributed in a transparent way, allowing utility-level data to be combined with attempts to anonymize individual-level usage data, technical solutions and location-specific interventions will compete and win on a level playing field. Once these rewards and incentives are coupled with financial innovations that help customers make up-front investments and share in the savings, it will no longer be up to governments and international organizations to engage in the complex optimization problem of efficiency investment. Moreover, behavioral science insights will be needed to get consumer buy-in and help customers take an active role in managing their own energy use.

Finally, smart and connected devices can help take the day-to-day burden off consumers. Much of the relevant work in this area is marketed under the banners of “smart buildings” and “smart cities.” Major industry players like IBM and Siemens have spent the last two decades building a framework for buildings and campuses to optimize their energy usage. For example, a “smart building” can adjust its lighting and heating, ventilation, and air conditioning (HVAC) systems in response to occupancy, and can shift energy-intensive tasks in accordance with price signals. These automation frameworks have mostly been deployed in large commercial and industrial settings, where ideas like real-time pricing, demand response, and peak time pricing incentives can work especially well. For example, FirstFuel (USA) offers remote large-scale building assessments that helped the U.S. Department of Defense identify approximately $4 billion in potential energy savings from assessments of around 100 buildings. This data driven approach offered 16 to 37 percent more savings compared to the leading on-site audit performed by a third-party independent building auditor. Along the same lines, Warwick Analytics (UK) has spent the last 10 years improving the ability to perform predictive analytics on production processes, offering an automated plug-and-play solution that makes sense of big data analytics.

Cities such as Barcelona, London, New York, and Dubai are moving to implement energy-saving “smart” technologies at the neighborhood or municipality level—ideas like connected street lighting, district heating and cooling, and electric car infrastructure. As the developing world urbanizes, there are innumerable opportunities to deploy connected infrastructure right from the beginning—raising the up-front cost, but reducing the need for expensive retrofits down the road. The technologies that work for customers in the developed world will need to be adapted to local conditions, made more affordable, and evaluated for performance and return on investment.

RENEWABLE GENERATION

One of the most encouraging developments in the energy industry over the past decade has been the emergence of cost-competitive renewable generation. Along with widespread recognition of the problem of climate change, the new economic viability of renewables has led to a sea change in development funding. Mass renewable generation is coming, and the question is whether the grid will be ready for it.

Prediction is of particular importance in optimizing renewable deployments,[22] as generation technologies vary dramatically in viability across diverse environments and place large demands on the grid. Utilities want to predict which technologies—wind, solar, hydropower, geothermal—work the best in which climates and geographies. They also must predict the outcomes of new distributed inputs with variable capacity, in terms of both overall capacity and grid stability. Fortunately, the granular nature of most renewable generation allows predictions to be refined using *dynamic evaluation*.

Renewable Generation Highlights

- Renewable generation must usually be optimized for its environment, taking into account factors like sun and wind exposure, but also consumer behavior so that the grid is reliable and stable in the face of rapid changes in supply and demand.
- Big data from sensor networks and simulations can be used to systematically assess the viability of generation technologies in a given environment, and real-time monitoring and experimentation can be used to optimize pricing signals to ensure demand, supply, and storage capacity are aligned.
- The widely distributed infrastructure inherent in most renewable deployments also increases the need for predictive analytics in equipment maintenance.

Renewables will drive the energy industry to move from day-ahead markets to five minutes-ahead markets, creating a corresponding need for real-time *prediction* and *evaluation*.[23] This shift to anticipate supply and demand will become the primary energy analytical challenge over the long term. In order to plan investments and maintain reliability, utilities will need to know how much capacity will be provided as a function of time, which will mean *predicting* wind and solar exposure on different parts of the grid. On the demand side, utilities will need to anticipate consumer responses to price signals. Especially in a distributed network with small generation assets, *predictions* will need to be made locally and bubbled up to the macro level, creating a "responsive grid."

Similar challenges will arise for distributed storage, which is a natural complement to most renewable generation. Already, companies like Totus Power (India) and FreeWire Technologies (USA) are deploying second life electric car batteries to provide backup storage at a lower cost per kilowatt than kerosene generators. Modeling will be necessary to understand how, where, and what kinds of storage to incentivize in order to most effectively smooth over peak demand and ensure reliability. Responsive grid technology will also be necessary to determine when these storage assets should charge and discharge.

An additional complication arises from the fact that many of these new generation and storage assets will be decoupled from the grid to some degree.[24] That creates a new need for local power management with coordination across networks. Increased renewable penetration will also create the need for higher pricing flexibility lower down in transmission lines, at the industrial, commercial, and residential levels. Generation and storage assets will need to respond to real-time pricing. All of these are huge challenges for the traditional grid.

It is in some ways fortunate, however, that this problem has arisen concurrently with the push to extend access: renewable deployment and operation at scale will require a grid powered by Big Data. Because they will essentially be building out new systems from scratch, developing world utilities can create a flexible and distributed grid from the beginning, can intelligently deploy generation and storage assets, and can take advantage of connected devices to orchestrate the system. This opens interesting opportunities to network urban infrastructure and improve information flow between water, transport, telecom, and energy sectors.

Box 2 | Predicting Wind Speeds in Rural India

Indian company REConnect Energy is a trading venture in the area of Renewable Energy Certificate (REC), Energy Portfolio Management, and Wind Forecasting and Scheduling. In 2013, they developed a data-powered analytics tool for wind and solar power forecasting and scheduling. This tool integrates high-frequency, large-scale data from a variety of intermittent generation assets, processes this streaming data, and uses a set of machine learning algorithms to make predictions of future power output for each generation asset. Accurate predictions of power output from renewables are critical to increase renewables integration on the grid while simultaneously improving power quality. By improving power predictions of renewable assets, utilities can better anticipate changes to generation, balance with load, and reduce the overall costs of integrating renewables onto the grid. With 36 percent market share and more than 400 projects over 16 states it is the largest REC trading company in India.

Source | REConnect Energy. http://reconnectenergy.com

ENERGY ANALYTICS: GOING FORWARD

To end extreme poverty and expand shared prosperity, universal access to clean, affordable energy is required. Sustainable energy solutions require: improved *description* and *detection* of human needs; appropriate deployment of technology to help meet these needs; the ability to *predict* behavior and design appropriate incentives to influence preferences and decisions; more targeted interventions; and be capable of dynamically evaluating interventions in real time to make sure programs are actually delivering on their promises. Each of these components benefits from the tools offered by Big Data Analytics.

Despite many of the potential benefits described above, significant challenges remain in order to scale up the use of Big Data Analytics in the energy sector globally. Several technical challenges have been discussed, including the cost of data infrastructure upgrades, lack of human capacity to perform analytics, poor connectivity limiting capture and dissemination of data, and human behavior—reluctance to adopt new technologies. As integration of digital technologies and data will increase in energy infrastructure, the energy sector will be more at risk of cyber threats: potential attacks on the computer systems used for these purposes. Strategic challenges and specific needs for cyber security in the energy sector are particularly relevant in the following four areas: (i) management of risk and threats, (ii) cyber defense, (iii) cyber resilience, and (iv) the capacity and competences needed to take action.[25]

Some of the more pressing policy challenges include broader governance issues, including digital privacy and security, outdated regulatory models, as well as data ownership and sharing agreements.

Given the opportunities and challenges presented in this solutions brief, and others,[26] a number of stakeholders could benefit from deploying advanced analytics in emerging and developing markets. However, the uptake of Big Data Analytics practices will not happen by itself. Lessons learned from early pilots and good practices need to be developed and disseminated in order to support energy sector stakeholders in taking advantage of the benefits offered by Big Data Analytics in developing and emerging markets. Some specific considerations include:

- **Governments** could contribute by developing regulations around data ownership, best practices for data management and access rights to data, as well as explore ways of deploying connected devices to increase access and efficiency on regional scales. Championing open data principles on the national level would also be beneficial.
- **Regulatory agencies** need to build capacity among staff to make sure they are trained to understand and enforce new regulations that enable innovative business models. They need to balance the importance of ensuring privacy and customer protection while not overburdening the industries with unnecessary requirements.
- **Utilities** could invest in data infrastructure and analytics capacity and improve their operational efficiency by exploring predictive maintenance, nontechnical loss detection, and alternative payment options.
- **Large energy consumers** could explore investments in renewable generation, efficiency upgrades, and smart building technology to reduce their usage and increase their competitiveness.
- **Academic institutions** can continue to add researchers and degree programs in data science and could take on research questions such as grid stability modeling, predictive maintenance, and payment plan optimization using open data and proprietary data from partners in developing countries.
- **International organizations** could be both funders and conveners to help shape the dialogue. Since local technical capacity will be a key bottleneck in deploying each of the ideas discussed in this solutions brief, international organizations could help supplement and grow that capacity, specifically to improve the ability of governments and electricity utilities to make informed policy decisions and improve efficiency and effectiveness across energy-related operations employing data-driven solutions. This would include assisting energy stakeholders in building digital infrastructure (software and connectivity infrastructure, such as integrated networks and sensors) and human capacity to collect, store, analyze, and share data through targeted trainings and skills building programs.

It is important to note, however, that applying Big Data Analytics in the energy sphere is not an end goal. This solutions brief has highlighted areas in which data science can be usefully deployed, but in no case does the application of data science solve a problem on its own. Instead, each solution will require a broad and cross-disciplinary effort that also includes technologists, development specialists, energy sector players, and social scientists.

International organizations can build this coalition. It will be possible—and desirable—to leverage work in the developed world, but the specific challenges in developing and emerging markets will require adaptation and further innovation. Development banks and international organizations in particular are well placed to leverage their experience working on the ground, mandate to take a long-term time horizon thinking on scales of decades, and their commitment to reduce poverty and expand shared prosperity.

ENDNOTES

1 *ICT & SDGs—How Information and Communication Technology Can Achieve the Sustainable Development Goals*. Ericsson and the Earth Institute at Columbia University.

2 *World Development Report 2016: Digital Dividends*. Sector Focus 5. World Bank Group, 2016.

3 Gee, Matt. Workshop: "Energy Analytics for Development." World Bank Group, May 2015.

4 Ng, Andrew. "Machine Learning." Coursera, Stanford University, January 2015.

5 https://ec.europa.eu/energy/sites/ener/files/documents/eecsp_report_final.pdf

6 http://www.pwc.com/gx/en/oil-gas-energy/publications/assets/pwc-embedding-cyber-security-into-the-energy-ecosystem-pdf

7 "Proposed Targets and Goals." Sustainable Development Solutions Network. Accessed on June 24, 2015. http://unsdsn.org/resources/goals-and-targets/

8 Crousillat, Enrique, Richard Hamilton, and Pedro Antmann. *Addressing the Electricity Access Gap*. World Bank Group, 2010.

9 *World Energy Outlook 2010*. International Energy Agency, 2010.

10 *Off-Grid Solar Market Trends Report 2016,* Bloomberg New Energy Finance, 2016.

11 Mockel, Peter. Senior Industry Specialist. IFC, 2015.

12 Conference Proceedings from the University of California, Berkeley, "Revealing Demand for Pro-Poor Innovations." March 2014.

13 Taneja, Jayant, IBM Research-Africa.

14 Sahai, Dhruva. *Toward universal electricity access: renewable energy-based geospatial least-cost electrification planning*. World Bank Group, 2013.

15 Readiness Assessment Tool. http://opendata-toolkit.worldbank.org/en/odra.html

16 Ibid. See specifically Open Energy Data Assessment.

17 *Energy Efficiency Market Report 2014*. International Energy Agency, 2014.

18 *Unlocking American Efficiency: Economic and Commercial Power of Investing in Energy Efficient Buildings.* Rhodium Group on behalf of United Technologies, May 2013.

19 *Capturing the Multiple Benefits of Energy Efficiency*. International Energy Agency, 2014.

20 Jessoe, Katrina, and David Rapson. "Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review*, 2014.

21 For more information, see C40 Cities Climate Leadership Group or IBM Smart Cities Challenge.

22 Gibbles, Rolf, and Futch, Matt. Smarter Energy: optimizing and integrating renewable energy resources. IBM Thought Leadership White Paper. November 2012.

23 Smith, J. C., et al. Impact of Variable Renewable Energy on U.S. Electricity Markets. 2010.

24 Kempener, Ruud, Komor, Paul, and Hoke, Anderson. *Smart Grids and Renewables: A Guide for Effective Deployment*. International Renewable Energy Agency, November 2013.

25 *Cyber Security in the Energy Sector, Recommendations to the European Commission on a European Strategic Framework,* EECSP, 2017.

26 See specifically: Letouze, Emmanuel. "The Law, Politics and Ethics of Cell Phone Data Analytics," World Bank Group, April 2015.

ACRONYMS AND ABBREVIATIONS

AMI	Advanced metering infrastructure
ICT	Information and communication technology
IoT	Internet of Things
IT	Information technology
JPS	Jamaica Public Service Company, Ltd.
OECD	Organisation for Economic Co-operation and Development
PV	photovoltaics
WBG	World Bank Group

All currency in United States dollars (USD, US$, $), unless otherwise indicated.

Written by | Anna Lerner and Zhenia Viatchaninova Dalphond, World Bank, and Matthew Gee and Thomas Plagge, Impact Lab

Photo Credits

Cover: © Chris Chopyak | Alchemy
Table of Contents: © Anna Lerner | World Bank
Page ii: © Asian Development Bank (via Flickr CC BY_NC_ND 2.0)

Production Credits

Production Editor | Heather H. Austin

Typesetting | Shepherd, Inc.

The International Bank for Reconstruction

And Development / THE WORLD BANK GROUP

1818 H Street, NW | Washington DC 20433 | USA

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

ESMAP
Energy Sector Management Assistance Program

The Energy Sector Management Assistance Program (ESMAP) is a global knowledge and technical assistance program administered by the World Bank. It provides analytical and advisory services to low- and middle-income countries to increase know-how and institutional capacity to achieve environmentally sustainable energy solutions for poverty reduction and economic growth. ESMAP is funded by Australia, Austria, Denmark, the European Commission, Finland, France, Germany, Iceland, Japan, Lithuania, the Netherlands, Norway, the Rockefeller Foundation, Sweden, Switzerland, and the United Kingdom, as well as The World Bank.

Energy Sector Management Assistance Program
The World Bank
1818 H Street, NW
Washington, DC 20433 USA
email: esmap@worldbank.org
web: www.esmap.org