Application and Development of Energy Big Data

Cai Zexiang[1], Li Licheng[1,2], Liu Ping[1], Xu Min[2], Chen Zexing[1], Zhang Yongjun[1], Han Yongxia[1], Xu Aidong[2]

1. School of Electric Power, South China University of Technology, Guangzhou 510641, China
2. Electric Power Research Institute, China Southern Power Grid, Guangzhou 510080, China

Abstract: Energy big data, which is a significant part of "Internet Plus" smart energy, plays a critical role in promoting China's energy revolution, facilitating the transformation of the country's energy structure, and stimulating innovative development in energy industries. In the context of "Internet Plus" smart energy, the basic framework and key features of energy big data are first discussed, followed by the major applications of energy big data in energy industries. This paper also analyzes significant obstacles to energy big data development based on the status quo of energy big data in China. Finally, several suggestions are made for energy big data development to overcome these obstacles and promote the construction of energy big data and applications of "Internet Plus" smart energy in China.

Keywords: energy big data; Internet plus; smart energy; information and communication infrastructure; energy industries

1 Introduction

Based on the urgent demand for energy transition around the globe, the concept of "Internet Plus" smart energy has attracted significant attention from China's energy industry. The integration of energy networks with the Internet, which is an essential representation of "Internet Plus" smart energy, is the central driving force for resolving current energy issues and reshaping the global energy structure [1]. Energy big data, which combines a massive volume of energy data with big data technologies [2], is a key concept that contributes to the construction of "Internet Plus" smart energy systems. Data on various types of energy, including electricity, coal, petroleum, natural gas, and heat are collected during the processes of energy generation, transmission, storage, consumption, and trade, and are integrated using various energy big data techniques. In this manner, governments can supervise energy industries, societies can share energy information resources, and energy industries can conduct in-depth market reform [3,4]. In addition, open data-sharing forms the core of energy big data and is an efficient approach to facilitate the smart transformation and enhancement of energy systems. In

other words, it provides an excellent opportunity to restructure conventional energy systems through the application of Internet mechanisms and technologies. Energy big data is not only a force that breaks barriers to entry, promotes the integration of multiple energy systems, and creates new business models for smart energy systems, but it is also the key technology for transforming and upgrading energy systems and power sources into future economic growth engines [5–7].

Therefore, this paper first discusses the basic concepts of energy big data technologies and presents various frameworks and attributes associated with such technologies. Significant issues that hinder the development of energy big data in China are then discussed. Finally, proposals for the future layout of China's energy big data industry are provided to aid in the development of "Internet Plus" smart energy strategies.

2 Basic concepts of energy big data

Big data is a synthesis technique that considers an entire dataset as a research object, and combines sensor information and communication, computer science, data analysis, and do-

main-specific technologies. It effectively expands upon traditional data mining and analysis technologies [8–10]. Because the concept of "Internet Plus" has completely penetrated China's energy industry, the resulting emergence of "Internet Plus" smart energy combines ideas and technologies from the Internet in order to reshape the entire industrial chain of conventional energy industries, including production, transmission, consumption, transformation, and transactions. The implementation of energy big data facilitates the establishment of novel, interconnected, transparent, open, mutually beneficial, and sharable energy systems based on the integration of energy and information. The energy big data framework for "Internet Plus" smart energy consists of application, platform, data, and physical layers, as shown in Fig. 1.

The physical layer of energy big data contains all aspects of energy production, transmission, and consumption as well as different types of associated energy equipment. System operation information and health condition data are accessible because of the installation of nonintrusive sensors and energy meters for energy networks and energy equipment. These sensors and energy meters transmit relevant data to the smart operation and maintenance with situational awareness (SOMSA) system to implement functions, including data visualization, state monitoring, intelligent warnings, and fault location. The information communication and smart control (ICSC) system is responsible for the communication and control tasks of each aspect and device in energy systems. The massive quantities of data generated by the ICSC system are stored with data from external systems (e.g., meteorological data) in databases specific to energy big data. The functions in the platform layer such as energy efficiency assessment, risk identification and evaluation, and economic and utilization analysis of energy are dependent on the processing of stored data. From the perspective of energy generation, energy

big data technology facilitates the accurate prediction of power outputs from renewable sources and multi-energy optimal dispatch of electricity, gas, heat, and cooling. From the perspective of energy transmission, it facilitates smart online maintenance that can monitor the operation status of energy networks and perform automatic identification of fault locations. From the perspective of energy consumption, it provides services that enable users to analyze and promote energy efficiency and coordinate various types of loads to create a demand-side response to further facilitate energy utilization efficiency.

We consider the electricity usage data listed in Table 1 as an example to illustrate the basic features of energy big data. Energy big data generally gathers information from multiple sources, such as enterprise reports at the user level, electricity meters at the equipment level, and control, operation, and maintenance systems at the system level. Compared to the structured measurement data in traditional energy systems, each type of data resource for energy big data covers different ranges of physical areas, focuses on different spatiotemporal scales of data, and explicitly presents multisource and heterogeneous features in terms of data diversity [11]. Based on the extensive application of big data technology utilizing sensory devices and energy equipment in energy systems, the amount of energy big data is growing exponentially from the terabyte (TB) to the petabyte (PB) range, and even the exabyte (EB) range. In addition, energy big data values timeliness and comprehensiveness, which require minute-level sampling procedures for data gathering, leading to significantly higher data volumes. Finally, the collected data is processed, analyzed, and utilized to identify patterns and make predictions on energy usage, forecast variations in energy generation and load demand, and aid in energy market trading as well as other value-added services, instead of merely being utilized for statistical analysis and report generation.
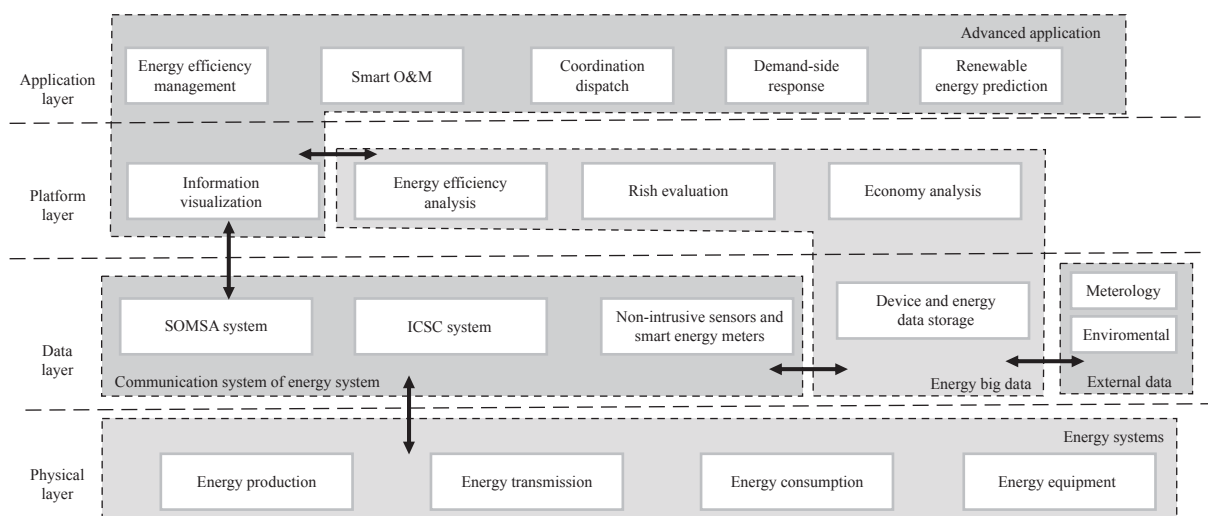

Fig. 1. Basic framework for energy big data. Smart O&M: Smart operation and maintenance; SOMSA: Smart operation and maintenance with situational awareness; ICSC: Information communication and smart control.

Table 1. Major sources that generate electricity usage data.

Data sources	Data volume	Average data gathering cost	Coverage scope	Data granularity	Direct usage	Level of data sharing
Energy report forms from enterprises	Kilobyte (kB)	No direct cost	Large-scale enterprises	Data collected monthly or annually by enterprises	Official statistics	Open access to macro information
Electricity measurements	TB	2000–5000 RMB per data source point	Full coverage	Data collected monthly by private and public transformers	Billing services for electricity utilities	Confidential and inaccessible
Load control systems	PB	>10000 RMB per data source point	Private and public transformers: 100% coverage in developed regions, different coverage scope in other regions	Data collected every 15 minutes by private and public transformers	Large load control for electricity utilities	Confidential and inaccessible
Operation and maintenance monitoring systems /energy efficiency management systems/ electricity sales service systems	EB	2000–5000 RMB per data source point	Low coverage, third-party energy service providers or enterprises install meters voluntarily	Data collected every 1–5 minutes by branch lines, production lines, or process flow	Equipment maintenance, energy efficiency management, and electricity sales service in enterprises	Data owned by different enterprises or third-party energy service providers, low data concentration

3 Application of energy big data

As Internet technologies and "Internet thinking" gradually integrated with energy systems, energy industries have realized the extensive potential of energy big data applications in every aspect of energy systems. In China's new energy era, energy big data has made significant contributions to facilitating the development of renewable energy and stimulating the growth of cross-industry integration and innovative development in energy industries. Energy big data, which helps governments conduct energy supervision and societies share energy information resources, is the foundation for promoting energy market reform and is vital to implementing China's "Internet Plus" smart energy development strategy with intelligent enhancement of existing energy systems. Moreover, multiple-energy-system integration and energy industry innovation are highly dependent on energy big data, which plays an integral role in creating novel business models for smart energy industries and bolstering economic growth engines.

The typical fields where energy big data are applied include the following:

(1) Energy planning and energy policymaking

In the government decision-making domain, energy planning and policymaking are two major fields that implement energy big data. In energy planning, governments utilize energy big data to obtain and analyze information on energy efficiency and energy usage behavior from end-users by gathering energy usage data from regional enterprises and households. In this manner, energy big data can aid governments in planning energy networks and selecting energy stations. Additionally, the energy composition and energy resources in a region can be investigated based on energy usage data as well as geographical and meteorological information, which provide guidance for sustainable exploration

and utilization of energy. In energy policymaking, through the analyses of energy usage levels and behaviors of regional users via energy big data, governments can identify energy consumption issues from local enterprises to further study the rational design of industrial layouts and offer reasonable suggestions for making economic development policies. Furthermore, mining and processing data on energy resources and load demand utilizing energy big data can provide guidance in national and regional policymaking, such as renewable energy and electric vehicle subsidy plans, electricity-price-based incentive mechanisms, optimization of urban planning, development strategies for smart cities, proper development of renewable-energy-powered cars, among others.

(2) Energy production

In the field of energy production, big data technologies are currently applied for the accurate prediction of renewable energy generation and promotion of renewable energy consumption. Based on the stochastic and intermittent nature of renewable energy, it is necessary to deploy flexible solutions, such as energy storage systems, and implement dispatch plans with credible forecasting results for renewable energy generation. In this manner, the impact of renewable energy on power grids can be reduced, renewable power curtailment issues can be mitigated, and the reliability of power grids can be ensured. Currently, Envision Energy, an energy company in China focusing on the smart implementation of wind and solar farms, has integrated technologies from the Internet of Things, big data, and machine learning to construct the EnOSTM platform. This platform processes nearly 1 TB of data every day and has become a leader in accurate generation prediction and flexible control of renewable energy. Likewise, researchers from other countries have conducted considerable research on wind farm location and autogeneration control by processing meteorological and geographical data uti-

lizing big data technologies [12,13]. The trend of applying Internet technologies in energy production facilitates the integration of data from all wind farms operating in the same area, which resolves many issues with the conventional method of separately forecasting data from a single wind farm, and enhances the accuracy of renewable energy prediction [14].

(3) Energy consumption

With the increasing penetration of renewable energy into the demand side of the industry and the gradual maturity of microgrid systems, the role of energy users is changing from conventional consumers to prosumers. In terms of application on the demand side, utilizing energy big data to incorporate end-user renewable generation resources and flexible loads with controllable features (e.g., electric vehicles), and maximizing the profits of end-user energy systems by enabling them to participate in market trading, are hot research topics [10]. Some pioneering work has been conducted on the practical application of energy big data in energy consumption fields. C3 Energy and Opower, two American companies, have developed an analytical engine platform and energy services platform based on big data technologies, providing users with energy services and supporting demand-side response programs [15]. The E-energy program in Germany has implemented various big data-based solutions to facilitate renewable generation prediction, develop new business models for energy services, and promote energy markets [16,17]. In China, the National Cloud Platform of Smart Energy Public Services was initiated in February 2015. To date, 14 provincial and municipal departments have approved the establishment of regional subplatforms for the national cloud platform. The national cloud platform offers services for gathering and analyzing energy data. It can also develop data subplatforms for real-time equipment management and foster new sales models to obtain cost-effective products and energy solutions. This national platform is expected to reduce energy costs and improve energy efficiency. In the future, with the help of the national cloud platform, the "information island" issues of governments and financial institutions can be addressed, transparent and real data can be obtained, and effective supervision and management programs can be implemented.

(4) New business models for smart energy

Energy big data, with its in-depth expansion into various energy systems, will likely foster a batch of new business models for smart energy services such as monitoring and maintenance of energy networks and energy market transactions. In terms of energy system maintenance, situational awareness technology for data measured over a wide area has been applied to the online operation and maintenance of transmission and distribution stations in smart grids, providing functions such as real-time contingency alerts, fault location, and oscillation detection [18,19]. In addition, as a result of the complex hardware and scattered locations of renewable-energy-based power stations, the generation units inside these power stations require big data

technology to monitor the fatigue and wear of components in accordance with data from sensors. In this manner, equipment operation status can be predicted and identified online to avoid the large-scale deployment of monitoring systems and prevent potential malfunctions as early as possible [20]. In the foreseeable future, the integration of energy systems will expand the scale of equipment and deepen the complexity of energy networks. Moreover, along with the gradual improvement of an open and sound electricity market, multiple electricity retailers are expected to emerge. This trend will inevitably result in the dispersal of regions of operation and electricity assets, which means that it will be less economical to retain professional operation and maintenance (O&M) teams. This indicates that conventional collective approaches to O&M are incompatible with development of energy integration. Motivated by sharing concepts from the Internet, it is feasible to implement distributed systems with the aid of the Internet and big data technologies. Professional O&M teams can be geographically deployed to cater to the specific needs of users. O&M requirements will also form or serve as the backbone of new business models based on energy big data in the future.

Energy big data plays a significant role in the development and improvement of energy markets. Currently, the application of energy big data around the world for energy trading is still in its infancy. In America, a microsolar electricity trading market at the Buffalo Niagara Medical Campus was established by the National Grid of the United Kingdom. Based on big data technologies, this company aims to optimize the dispatch of solar generation, storage systems, and load demand on the campus and provide pricing services for generation resources [21]. The development of energy big data integrated with energy production, transmission, and consumption offers energy industries an open and shared energy information platform and promotes free and flexible energy trading. Such a system would enable energy prices to directly reflect the supply–demand relationship and lead to the optimal allocation of resources, thereby facilitating the creation of a fair, open, and shared energy market. Furthermore, energy big data can facilitate the optimal combination of various types of energy technologies in accordance with specific requirements and technological characteristics, and subsequently bring about new business models for integrated energy services, including trading of different types of energy and other added-value services.

4 Status quo analysis of China's energy big data technology

The communication quality of information systems is rapidly improving owing to the development of new sensors, information transmission mechanisms (e.g., multiple access technology and spread spectrum technology), fiber transmission technology, data preprocessing technology, among others. Energy sys-

tem decision-making based on the processing of energy data, whose preliminary applications have been energy generation, transmission, and consumption, is prompting energy systems to adapt and grow. However, as a result of the imperfection in the mechanisms of information management and construction of information system infrastructure, the slow construction and relatively low-level applications of energy big data have become bottlenecks that hinder "Internet Plus" upgrades for energy systems. The main issues with the status quo of energy big data applications are summarized below.

(1) Prevalence of information island issues in energy systems

Access to massive energy data is a prerequisite for the construction of energy big data. However, information island issues, which are prevalent in energy systems, represent a major obstacle to the integration of energy information resources.

During the informatization of energy enterprises, such as electricity, coal, oil, gas, cooling, and heating, the lack of a sufficient collective management mechanism results in each energy enterprise having an exclusive energy management and control system, where energy data are separately captured by the sensors used by each enterprise. The discordance between information protocols and incompatibility of energy frameworks mean that the data collected by each energy enterprise cannot be shared between enterprises, thus prohibiting the use of energy big data to further analyze the collected data. Additionally, it is a long-standing issue that conventional electrical systems and other energy systems are highly fragmented and carry out plans and operations independently. Large industrial barriers in multiple energy systems prevent information connections between energy systems. This phenomenon exacerbates information island issues and restricts the development of energy big data.

(2) Shortcomings of basic infrastructure in terms of supporting energy big data

Big data technology requires the universal support of information supply chains ranging from low-level chips and basic software to advanced applications and analysis software. However, this is currently a low-priority issue for basic energy information infrastructures in China.

In the field of sensory technology, new computing platforms, distributed computing frameworks, and data processing, analysis, and display methods for big data, the energy information technology in China lags behind similar technologies in many other countries. The current level of China's energy information technology is inadequate for wide-area gathering, effective storage, and fast processing of polymorphic and heterogeneous data generated by multiple sources such as electrical systems and other energy networks. Again, we consider the electricity usage data in Table 1 as an example. This electricity usage data was captured from various sources, including enterprise statistics, measurement meters, electricity utilities, and third-party energy companies. These sources differ in terms of data volume, range of coverage, data granularity, data obtainability, among others.

Furthermore, there is a lack of awareness regarding the exploitation and application of energy information. Considering the massive amount of data collected by integrated systems, the conversion of collected data into resource advantages that can be utilized to optimize the operation of energy systems is vital. As shown in Table 1, the volume and concentration of electricity usage data collected by electrical utility systems is relatively high compared to other data sources. However, the collected data is currently utilized only at the level of individual businesses; its full potential remains unexplored.

(3) Information security issues are prominent in energy systems

An open, compatible, and interconnected energy system inevitably includes security risks. The security of an entire energy system is difficult to manage, especially considering the fact that the application of Internet technology to energy systems, which are open and interconnected networks, as well as the interaction between information and physical components represents a significant threat to the security of energy systems [22,23]. Given that energy big data depends on an open and shared environment for energy data, the construction and application of energy big data demands the consolidation of protection capabilities for energy information security. Additionally, energy big data gathers a large amount of energy usage data from users, which can potentially lead to privacy issues. In the process of constructing energy big data, the tradeoff between sharing and securing energy data is a top priority.

5 Suggestions regarding the development of China's energy big data industry

(1) Make big data support the intellectualization of energy systems through in-depth integration of energy systems and big data technologies

The integration of big data should be facilitated by energy systems in various fields, including hardware sources, fundamental software, network communications, data integration, computing support, application support, and security control. To address information island issues caused by heterogeneous data from multiple sources, adopting uniform energy information standards for the processes of data gathering, integration, and storage is necessary. Enhancing the processes of collecting, transferring, and storing wide-area multi-timescale energy data as well as other field-related data is required. In addition, organizations should focus on deep knowledge refinement and the value of practical applications of heterogeneous multisource data while strengthening the applications of visualization and human–machine interaction technologies, thereby enhancing the usability of energy big data. For the purpose of ensuring energy information security, it is also vital to facilitate the construction of security systems to implement measures for information security protection as well as management techniques.

(2) Improve basic infrastructure construction to overcome the weaknesses of physical and information connections between multiple energy systems

The structural reform on the supply side should be consolidated for multi-energy networks, specifically addressing the weaknesses of physical and information connections, and prompting harmonious development between energy physical networks and corresponding information frameworks. Further improvement of the fundamental information structure of energy big data should accelerate the development of transparent power grid/transparent energy networks. For information transparency and the intellectualization of energy production, transmission, and consumption processes, certain critical technologies, particularly those for intelligent decision making for energy big data, must be investigated. Such technologies can ensure the utilization of efficient renewable energy in energy systems and promotion of end-user energy efficiency.

(3) Intensify energy structural reform and eliminate barriers to information sharing between different energy industries

The further opening of energy markets from the user end, distribution end, and supply end is strongly recommended. Electricity distribution and retail markets should be expanded from incremental markets into stock markets. The oil and gas markets should allow the injection of private and overseas capital. A batch of new energy business models, such as energy finance and third-party energy service providers, may be fostered by end-user market mechanisms, which will force supply-end energy reform, accelerate the speed of energy production and utilization, and promote comprehensive energy efficiency in the entire society. With the goal of energy market reform, it is possible to eliminate barriers between different energy industries and enable the construction of energy big data.

(4) Improve the top-level design of energy big data industries and foster smart energy business models

Sound laws and regulations for energy markets and sustainable market mechanisms can be obtained through government policies, subsidy support, platform construction, and market design. In particular, laws and regulations for energy markets should be enacted and improved at a faster pace. Various pilot projects should be implemented to establish rankings for energy industries. An entirely new supply chain, including energy production, storage, transmission, exchange, consumption, and added-value services, must be developed. Furthermore, governments should improve the fundamental structure of energy industries and build energy big data platforms to cover different geographical scopes, including national, regional, and end-user platforms. Based on the utilization of open and shared data resources, different industries can be upgraded and various economic growth engines can be established in smart energy industries.

6 Conclusions

The construction of energy big data plays a significant role in implementing China's "Internet Plus" smart energy strategies, promoting in-depth integration of multiple energy systems, and facilitating energy market reform. After investigating the status of energy big data in China, we identified three major issues:

(1) Owing to the lack of information connections between energy systems as a result of cross-industry barriers, energy big data is still unable to facilitate information interconnections between energy systems.

(2) Based on the specificity and complexity of energy systems, the existing fundamental infrastructure and technologies of big data are inadequate for the requirements of energy systems.

(3) The operation security of energy systems is closely tied to national interests and individual livelihoods, which creates challenging information security problems for energy big data.

To solve these problems, the following actions for the development of China's energy big data industries were proposed:

(1) At the level of information and physical systems, energy big data should comprehensively scrutinize the operational characteristics and data features of energy systems in order to integrate big data technology and energy systems.

(2) At the level of basic infrastructures, a major goal should be the construction of transparent power grid/transparent energy networks to facilitate the improvement of basic infrastructures for energy information frameworks.

(3) At the level of energy systems, organizations should accelerate energy market reform and promote information interconnections between multiple energy systems by removing barriers between energy industries.

(4) At the level of top-level design for energy industries, it is critical to implement new business models for energy systems and facilitate the development of energy big data according to sound policies and regulations.

References

[1] Li L C, Zhang Y J, Chen Z X, et al. Merger between smart grid and energy-net: Mode and development prospects [J]. Automation of Electric Power Systems, 2016, 40(11):1–9. Chinese.

[2] Jiang H, Wang K, Wang Y, et al. Energy big data: A survey [J]. IEEE Access, 2016 (4): 3844–3861.

[3] Liu D N, Tang T Q, Zhao J W, et al. Big energy data information service pricing and its application in electricity market [J]. Electric Power Construction, 2017, 38(2):52–59. Chinese.

[4] Huang X Q, Chen J, Tian S M, et al. Big data integration for optimal planning and operation of electric vehicle charging stations [J]. Power System Technology, 2016, 40(3):762–767. Chinese.

[5] Zhang Y J, Chen Z X, Cai Z X, et al. New generation of cyber-energy system: Energy Internet [J]. Electric Power Automation Equipment, 2016, 36(9): 1–7. Chinese.

[6] Tian S M, Luan W P, Zhang D X, et al. Technical forms and key technologies on energy Internet [J]. Proceedings of the CSEE, 2015, 35(14): 3842–3894. Chinese.

[7] Chen Q X, Liu D N, Lin J, et al. Business models and market

mechanisms of Energy Internet (I) [J]. Power System Technology, 2015, 39(11): 3050–3056. Chinese.

[8] Naimi A I, Westreich D J. Big data: A revolution that will transform how we live, work, and think [J]. Mathematics & Computer Education, 2013, 47(17):181–183.

[9] McAfee, A, Brynjolfsson, E, Davenport, T H, et al. Big data: the management revolution [J]. Harvard Business Review, 2012, 90(10):60–68.

[10] Zhou K, Fu C, Yang S. Big data driven smart energy management: From big data to big insights [J]. Renewable and Sustainable Energy Reviews, 2016 (56): 215–225.

[11] Zhang D X, Miao X, Liu L P, et al. Research on development strategy for smart grid big data [J]. Proceedings of the CSEE, 2015, 35(1): 2–12. Chinese.

[12] Billinton R, Gao Y. Multistate wind energy conversion system models for adequacy assessment of generating systems incorporating wind energy [J]. IEEE Transaction on Energy Converstion, 2008, 23(1): 163–170.

[13] Kaldellis J. Optimum autonomous wind-power system sizing for remote consumers, using long-term wind speed data [J]. Applied Energy, 2002, 71(3): 215–233.

[14] Lu Z X, Xu M, Qiao Y, et al. New Internet based operation pattern design of wind power forecasting system [J]. Power System Technology, 2016, 40(1): 125–131. Chinese.

[15] Liu S C, Zhang D X, Zhu Z Y, et al. A view on big data in energy internet [J]. Automation of Electric Power Systems, 2016,

40(8):14–21. Chinese.

[16] B.A.U.M. Consult Gmbh. Smart Energy made in Germany [R]. Munich: B.A.U.M. Consult Gmbh, 2012. 17–33.

[17] Wang X W, Wang Y Z. Introduction of German smart grid "E-Energy" project promotion [J]. Power Demand Side Management, 2011, 13(4):75–76. Chinese.

[18] Liu Y, Zhan L, Zhang Y, et al. Wide-area measurement system development at the distribution level: An FNET/GridEye example [J]. IEEE Transactions on Power Delivery, 2015, 31(2): 721–731.

[19] Zhou D, Guo J, Zhang Y, et al. Distributed data analytics platform for wide-area synchrophasor measurement systems [J]. IEEE Transactions on Smart Grid, 2016, 7(5): 2397–2405.

[20] Peng X S, Deng D Y, Cheng S J, et al. Key technologies of electric power big data and its application prospects in smart grid [J]. Proceedings of the CSEE, 2015, 35(3): 503–511. Chinese.

[21] Magavern S. Greening the buffalo niagara medical campus [R]. New York: The Open Buffalo Innovation Lab, 2015.

[22] Becker T, Curry E, Jentzsch A, et al. New horizons for a data-driven economy: Roadmaps and action plans for technology, businesses, policy, and society[M]. London: Springer International Publishing, 2016.

[23] Li C B, Li X P, Tian S M, et al. Challenges and prospects of risk transmission in deep fusion of electric power and information for Energy Internet [J]. Automation of Electric Power Systems, 2017, 41(11): 17–25. Chinese.