REVIEW ARTICLE

Big data and artificial intelligence application in energy field: a bibliometric analysis

Yali Hou[1] · Qunwei Wang[2]

Abstract

This paper uses bibliometrics to characterize the knowledge systems of big data, artificial intelligence (AI), and energy based on the Science Citation Index Extension (SCI-E) and Social Science Citation Index (SSCI) of the Web of Science from 2001 to 2020. Results show that China is the country with the highest number of publications (1115), accounting for 29% of the total; however, the most influential country in the field is the USA, with an h-index of 75. The Chinese Academy of Sciences publishes the largest number of papers (104) and plays a vital role in the collaboration network. The study also reveals that the IEEE Access is the most productive journal (195) in terms of the number of publications, and engineering is the most popular discipline (1526). The key theoretical foundation includes deep learning (293), big data (105), energy consumption (79), and reinforcement learning (40). The application of big data and AI in the field of energy focuses on smart grid, energy consumption, and renewable energy. Early research frontiers involve optimization and prediction of energy-related problems using the genetic algorithm and neural networks. Since 2013, energy big data have gained prominence. At present, machine learning, deep learning, and fog computing are frequently combined with energy saving. In the future, big data and AI will be utilized to promote the application of renewable energy and energy-saving renovation of buildings. These findings can help researchers understand the developmental trends and correctly grasp the research direction and method of the emerging interdisciplinary field.

Keywords Renewable energy · Smart grid · Big data · Artificial intelligence · Bibliometrics

Introduction

An energy system provides strong support for economic growth and social development. However, there is a contradiction between the current energy system and the climate change mitigation and sustainable development goals. The International Energy Agency (IEA) assesses the climate commitments and intentions submitted by different countries, as well as their impact on the energy sector. IEA urges different countries to strengthen their emission-reduction commitments in order to meet the 2 °C target. The energy sector must play a key role in reducing global emissions because energy production and usage contribute to two-thirds of global greenhouse gas emissions. The development of emerging technologies is the key to ultimately transforming the energy system and meeting the climate goals (IEA 2015).

Big data and artificial intelligence (AI) are emerging technologies, and their applications in the energy field are supported by governments and recognized by the industry. In the past 20 years, research on the combination of big data, AI, and energy that covers a wide range of disciplines has shown a rapid increase. AI trains data through algorithms that allow computers to mimic human thought and consciousness. Therefore, big data is the foundation of AI. Big data cannot be completely stored in a single machine and must rely on distributed processing, distributed database, cloud storage, virtualization and other cloud computing-processing technologies. The Internet of Things has become an important source of big data by connecting a collection

✉ Qunwei Wang
  wqw0305@126.com

1 College of Information Engineering, Nanjing Xiaozhuang University, Nanjing 211171, China

2 College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

of objects that need to be monitored in real-time, as shown in Fig. 1.

Large amounts of data are accumulating in the energy field due to the continuous application of technologies such as sensors and cloud computing (Zhou et al. 2016). Energy big data have permeated different energy-related fields, including energy production (Ifaei et al. 2018), energy consumption (Abid et al. 2017), energy planning, and policy (Giest et al. 2018). The development of big data has promoted the advancement of AI technology to a large extent. The most common application of AI in the field of energy is prediction, including load forecasting (Chou et al. 2014), power generation output forecast (Puri et al. 2019), and energy consumption forecast for agricultural production (Nabavi-Pelesaraei et al. 2018; Kaab et al. 2019). Fault detection and diagnosis AI technology also play a key role in power system fault diagnosis (Zhao et al. 2019).

There is a continuous and rapid increase in the amount of literature published on the combination of big data, AI, and energy. Previous studies reviewed the application of big data and AI in building energy systems (Daut et al. 2017, Wang et al. 2017, Mehmood et al. 2019, Zhao et al. 2019) and renewable energy (Jha et al. 2017). A systematic review of big data analytics for smart energy management was conducted from a methodological perspective (Zhou et al. 2016). In addition, applications of machine learning in smart grid (Cheng et al. 2019, Kumar et al. 2020) and potentials of AI in the energy system (Sun et al. 2019) have been discussed. However, existing literature reviews are limited in their scope, and traditional analysis methods have not comprehensively analyzed the research hotspots and evolutionary track of the research frontier in interdisciplinary research. It is necessary to make a timely quantitative and qualitative analysis of the research papers in this field for

evaluating the growing body of knowledge. Bibliometrics technology provides a useful tool to achieve this aim.

This paper uses bibliometrics to characterize the knowledge systems of big data, AI, and energy based on the Science Citation Index Extension (SCI-E) and Social Science Citation Index (SSCI) of the Web of Science from 2001 to 2020. To be specific, this study shows that the interdisciplinary research on big data, AI, and energy is gradually becoming a popular research topic. Then, this study illustrates cross-disciplines and collaborative networks of institutions. Finally, this study presents the research foundation, research hotspots, and the evolutionary track of the research frontier.

This study provides a comprehensive overview of research for scholars and practitioners interested in big data and AI research in energy. It can help researchers understand the developmental trends and correctly grasp the research direction and method of the emerging interdisciplinary field. It can also assist energy enterprises to find the right direction for technology investment on big data and AI.

Methodology

Bibliometric methods apply mathematics, statistics, and philology to comprehensively and quantitatively analyze the number of articles, authors, institutions, and discipline words across all knowledge carriers. This analysis enables data mining on the research status quo, research hotspots, and frontiers of the discipline; and supports their accurate and intuitive display using a knowledge graph. The knowledge graph is a new field of scientometrics that can illustrate the process and track knowledge development in a graphical way.

Fig. 1 Interrelation of big data and AI
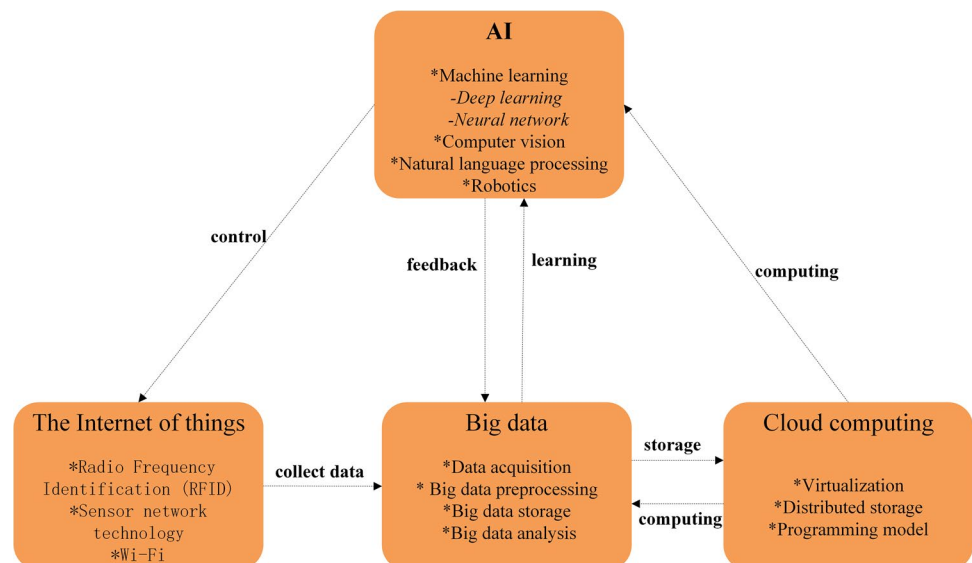
H-index

The h-index, where *h* stands for high citations and evaluates the academic achievements of individuals or institutions. The index was first proposed by the physicist Jorge Hirsch in 2005 according to the following definition: "A scientist has index h if *h* of his or her Np papers have at least *h* citations each, and the other (Np-h) papers have no more than h citations each" (Hirsch 2005). The goal of the h-index is to quantify the research results of researchers as independent individuals. A higher h-index of an individual demonstrates a greater influence of his/her paper. In this study, we applied the h-index to measure the comprehensive influence of an institution's or a journal's academic achievements in the field of big data, AI, and energy.

Social network analysis

Social network analysis (SNA) is a quantitative analysis method based on mathematical methods and graph theory. In this study, we use three concepts of SNA, including co-occurrence analysis, centrality, and burst detection. In order to visually show the conclusions of SNA, knowledge graph is applied. In this study, CiteSpace5.6.R5 is used to obtain the knowledge graph in the fields of big data, AI, and energy.

Co-occurrence analysis

In bibliometrics, commonly featured items in the same article must be correlated. The degree of correlation is measured by the co-occurrence frequency. A co-occurrence network can show the results of co-occurrence analysis more vividly in the form of a co-occurrence graph.

Centrality

The importance of a node in a co-occurrence graph is quantified according to its centrality. Centrality is used in a social network to express the degree to which a node is located in the center of the entire network. According to the calculation method, centrality can be divided into three types: degree centrality, closeness centrality, and betweenness centrality. The degree centrality represents the sum of a node and other nodes that are directly connected. The betweenness centrality represents the number of shortest paths through a node. A higher count of shortest paths through a node results in a higher betweenness centrality values. Nodes with a betweenness centrality greater than 0.1 are considered to be key nodes in CiteSpace (Chen et al. 2014a, b). In this study, degree centrality and betweenness centrality are used to analyze the scientific research cooperation and interdisciplinary integration.

Burst detection

According to the changing trends in the frequency of specific words over a period of time, burst detection was used to discover the words with a high change in word frequency in the social network. In CiteSpace, the burst words were detected using an algorithm proposed by Kleinberg in 2002 (Kleinberg 2003). The number of burst words can be adjusted by varying the duration of appearance. In this study, key words were detected and burst time was set at 2 years.

Unlike traditional literature review, this paper uses bibliometrics to study the application of big data and AI in the energy field from the perspectives of inter-discipline integration, research foundation, research hotspots, and the trajectory of the research frontier, as shown in Fig. 2. As

Fig. 2 Highlights of this study
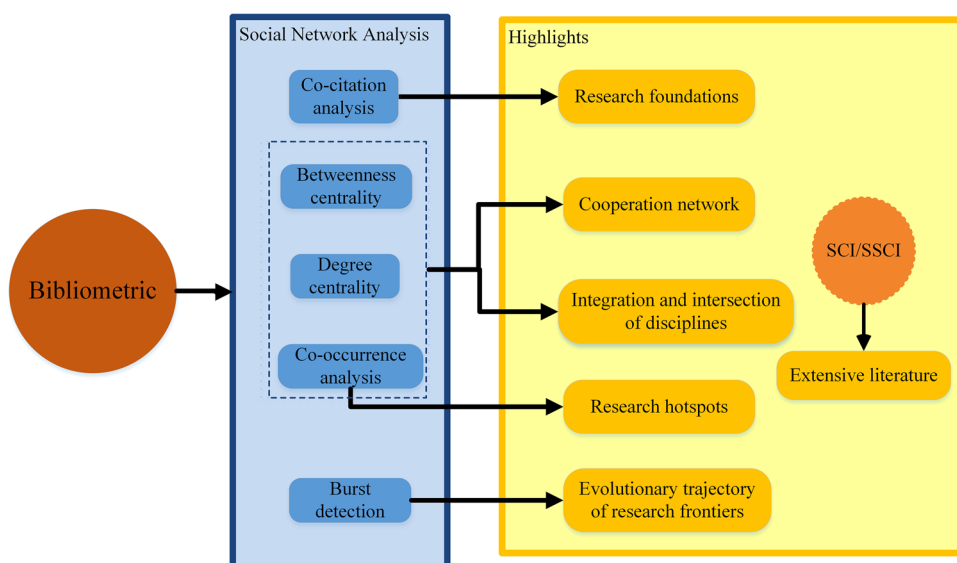
this research area involves many professional disciplines, the conclusion of this research is helpful to correctly grasp the research direction and method of the emerging interdisciplinary field with many"cross-industry professional concepts.

Results

The data for this study were collected from Science Citation Index Expanded (SCI-E) and Social Sciences Citation Index (SSCI). The data were retrieved from 2001 to 2020 to ensure the timeliness and accuracy of the research content. The retrieval date was May 5, 2021. The recall rate of the literature was improved by using the following phrases to search titles, keywords and abstracts: TS = (("artificial intelligence" OR "artificial intelligent" OR "big data" OR "large data") AND energy). The literature type was set as "article" and the language was set as "English". A total of 3842 articles were obtained.

Publication distribution of countries

From 2001 to 2020, a total of 108 countries contributed to the research relating to big data, AI, and energy by means of publishing academic articles in the SCI and SSCI databases. There were 15 countries, where each country published more than 2% of the articles. Among the three main publishing countries, the most influential country in the field was the USA, with an h-index of 75. Although China ranked first in the number of published papers, its citation rate was lower than that of the USA, with an h-index of 68.

Countries or institutions that appeared together in an article were considered to have research partnerships. The co-country network or co-institute network is obtained when the node of the co-occurrence is a country or an organization. The centrality obtained based on the co-country network are shown in Table 1. Degree centrality reflects the breadth of a country's cooperation. France has the largest number of cooperation countries. France, Spain, and Italy show betweenness centralities greater than 0.1 and play a key role in the process of forming national cooperation.

Figure 3 shows the variations in the number of annual publications in the top three countries according to the number of publications. The top two countries were China (1115) and the USA (927), which published significantly higher number of papers than any other country. Starting from 2016, the number of publications in China surpassed that in the USA. The USA always ranked first prior to 2016. From 2001 to 2020, the country with the most total publications was China (1115 articles), followed by the USA at 927 articles and the UK at 290 articles.

Table 1 Characteristics of the 15 most productive countries

Country	TP/R (%)	h-index	DC	BC
China	1115/29.021	68	30	0.03
USA	927/24.128	75	36	0.04
UK	290/7.548	38	37	0.08
India	232/6.039	29	25	0.04
South Korea	229/5.96	24	19	0.01
Spain	213/5.544	31	37	0.13
Germany	192/4.997	34	40	0.06
Italy	188/4.893	30	35	0.1
Canada	179/4.659	33	25	0.05
France	166/4.321	28	42	0.21
Australia	160/4.164	35	40	0.07
Iran	146/3.8	30	21	0.01
Japan	117/3.045	28	25	0.01
Saudi Arabia	102/2.655	19	22	0.04
Switzerland	79/2.056	27	27	0.01

TP/R (%), total number and proportion of publications of a country; *DC*, degree centrality of a country in a co-country network; *BC*, betweenness centrality of a country in a co-country network

Sources: the authors developed this table based on results

Publication distribution of institutions

From 2001 to 2020, a total of 3963 institutions contributed to the research on big data, AI, and energy by publishing academic articles in the SCI and SSCI databases. The co-institution network was built based on the top 100 levels of most cited each year, and the degree centrality was obtained. Table 2 shows that there were 30 institutions that published 20 or more papers.

The top 30 productive institutions are mainly from China and the USA: 14 are from China and eight are from the USA. The other institutions are from South Korea, Iran, UK, Singapore and Vietnam. Some of the top 20 productive countries do not appear in Table 2. These countries include India, Australia, Japan, Taiwan, and several European countries. The institutions in these countries that publish papers on big data, AI, and energy are relatively scattered. The Chinese Academy of Sciences (Chinese Acad. Sci.) publishes more papers and collaborates with more institutions than any other institution. In the top five productive institutions, one of them belongs to the USA and the rest belong to China. According to degree centrality, Chinese Acad. Sci. has the largest number of partners.

Figure 4 graphically illustrates the results of the co-institution analysis. The nodes in the purple circle indicate the key nodes, whose betweenness centrality is higher than 0.1. A line indicates a cooperative relationship between institutions. The thickness of the line indicates how closely the institutions work together. The thicker the line, the closer

Fig. 3 Annual number of articles of the three most productive countries
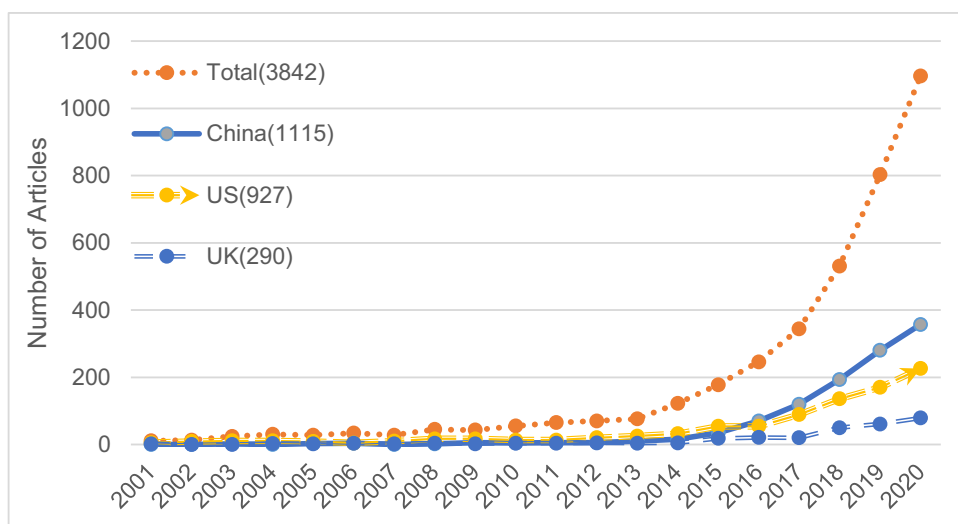
Table 2 Characteristic of the 30 most productive institutions

Institution	Freq	DC	Institution	Freq	DC
Chinese Acad. Sci., China	104	61	Beijing Univ. Posts Telecommun., China	24	13
Tsinghua Univ., China	63	34	Univ. Michigan, USA	23	34
Huazhong Univ. Sci. Technol., China	48	25	Stanford Univ., USA	23	23
Georgia Inst. Techno., USA	46	22	Northeastern Univ., China	23	10
Univ. Chinese Acad. Sci., China	40	21	Univ. Tehran, Iran	22	7
King Saud Univ., Saudi Arabia	34	21	Tianjin Univ., China	21	9
Islamic Azad Univ., Iran	34	27	Purdue Univ., USA	21	12
Hong Kong Polytech Univ., China	34	19	MIT, USA	21	48
Univ. Elect. Sci. Technol. China, China	29	14	Duy Tan Univ., Vietnam	21	11
Southeast Univ., China	29	19	Beihang Univ., China	21	27
Shanghai Jiao Tong Univ., China	29	22	Univ. Cambridge, UK	20	25
Univ. Oxford, UK	27	21	Univ. Calif. Berkeley, USA	20	22
Univ. Calif. Los Angeles, USA	26	14	Sejong Univ., Korea	20	8
Nanyang Technol. Univ., Singapore	26	19	Korea Adv. Inst. Sci. Technol., Korea	20	3
Zhejiang Univ., China	25	10	Dalian Univ. Technol., China	20	17

Freq, frequency with which an institution appears in a co-institution network; *DC*, degree centrality of an institution in a co-institution network

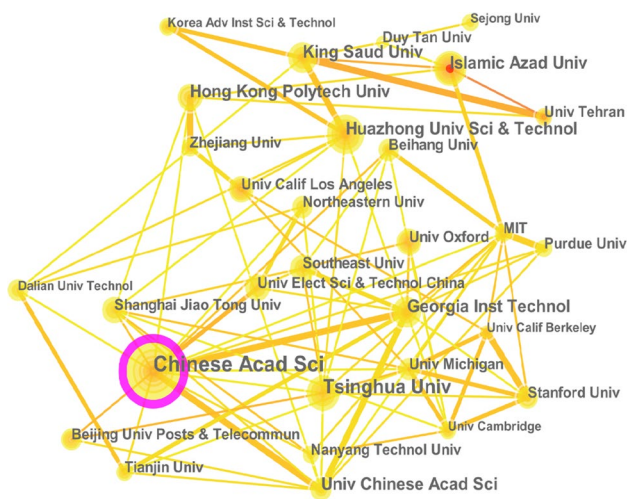Sources: The authors developed this table based on results

the cooperation between the institutions. The color of the line indicates the year of the first collaboration between the institutions. The darker the color, the earlier the first collaboration between the institutions.

As Fig. 4 shows, the Chinese Acad. Sci. is the only key node. The University of Chinese Academy of Sciences (Univ. Chinese Acad. Sci.) is directly under the Chinese Acad. Sci.. The two institutions have a strong cooperation relationship. The University of Electronic Science and Technology of China (Univ. Elect. Sci. & Technol. China) is another important member of the Chinese Acad. Sci.'s domestic publishing partner. As the key node, the Chinese Acad. Sci. has established partnerships with overseas

institutes. It has the closest relationship with Georgia Inst. Technol. in the USA. In recent years, the number of cooperative publications between the Chinese Acad. Sci. and Georgia Inst. Technol. has also increased gradually.

Huazhong University of Science and Technology is another Chinese institution that collaborates closely with the overseas institutions. It has produced many co-institution publications with King Saud University (King Saud Univ.) in Saudi Arabia and Korea Advanced Institute of Science and Technology (Korea Adv. Inst. Sci. & Technol.) in Korea. Other Chinese research institutions mainly collaborate with domestic institutions to publish papers on the application of big data and AI in the energy field. Another obvious

Fig. 4 Cooperation network diagram of the 30 most productive institutions

Table 3 Fifteen most productive disciplines during 2001–2020

Disciplines	TP	Ratio (%)	DC	BC
Engineering	1526	39.688	46	0.31
Computer Science	1122	29.181	45	0.36
Energy fuels	592	15.397	17	0.02
Telecommunications	515	13.394	5	0.00
Science Technology Other Topics	463	12.042	23	0.18
Physics	389	10.117	23	0.07
Chemistry	383	9.961	34	0.22
Environmental Sciences & Ecology	302	7.854	42	0.18
Materials Science	275	7.152	16	0.02
Mathematics	169	4.395	25	0.08
Thermodynamics	136	3.537	8	0.00
Instruments instrumentation	134	3.485	11	0.00
Construction building technology	129	3.355	6	0.00
Automation control systems	70	1.821	8	0.00
Business Economics	64	1.664	22	0.17

TP, total number of papers on a discipline; *r*atio (%): proportion of the number of papers on a discipline; *DC* degree centrality of a discipline in a co-discipline network; *BC*, betweenness centrality of a discipline in a co-discipline network

Sources: The authors developed this table based on results

international collaboration is between King Saud Univ. in Saudi Arabia and the University of Tehran (Univ. Tehran) in Iran.

There are three obvious core institutions in the USA: (1) Georgia Inst. Technol. is the largest publishing institution in the USA and has the highest number of collaborations with Chinese research institutions. (2) The collaborative network with the University of California (Univ. Calif. Berkeley) at Berkeley at its core includes Stanford University, the University of Michigan (Univ. Michigan) in the USA, and the University of Cambridge (Univ. Cambridge) in the UK. These institutions maintain a close and stable research cooperation with one another. (3) The third cooperation network with Massachusetts Institute of Technology (MIT) at its core includes Purdue University (Purdue Univ.) in the USA, as well as Beihang University (Beihang Univ.) in China and the Islamic Azad University (Islamic Azad Univ.) in Iran.

Characteristic of disciplines/journals

A total of 3842 articles on big data, AI, and energy in SCI-E and SSCI databases have been covered by 108 discipline categories. Table 3 shows the 15 most productive disciplines. A co-discipline network was built based on the top 200 levels of most cited each year, and the betweenness centrality values were obtained.

Degree centrality refers to the number of disciplines related to a specific discipline. A higher value indicates a higher proportion of cross-disciplinary research offered by that discipline. The top two disciplines are engineering and computer science, which are also key nodes in the co-discipline network. These two disciplines respectively encompass the research contents of big data, AI, and energy from two aspects: theoretical basis, and engineering practice and

environmental benefits. The other four key disciplines open up new research avenues for ecology, chemistry, science technology, and business on big data, AI, and energy.

Articles on big data, AI, and energy were published in a total of 1235 journals belonging to the SCI-E and SSCI databases during 2001–2020. Table 4 shows the top 20 productive journals, which published 26.6% of all articles. The IEEE Access is the most productive journal, followed by Energies and Applied Energy. Applied Energy publishes less articles than the IEEE Access, but has the highest h-index among these 20 journals.

Most highly co-citation articles

Co-citation analysis is a kind of co-occurrence analysis and refers to the fact that two articles appear together in the references of a third citing article, thus indicating co-citation relationship between two references. Co-citation analysis was used to explore the research foundation with respect to the application of big data and AI in the energy field. The ten most frequently co-cited articles since 2012 are shown in Table 5.

These articles are mainly related to deep learning, big data, energy consumption, and reinforcement learning. Deep learning in AI is a machine-learning technique that uses neural network as a parameter structure for optimization. It has laid a solid theoretical foundation for forecasting problems (LeCun et al. 2015; Schmidhuber 2015), fault detection (He

Table 4 Most productive journals during 2001–2020

Journal	TP	Ratio (%)	h-index
IEEE Access	195	5.072	26
Energies	123	3.199	15
Applied Energy	75	1.951	30
Sensors	68	1.769	11
Energy	63	1.638	20
Sustainability	63	1.638	9
Journal of Cleaner Production	59	1.534	23
Energy and Buildings	47	1.222	22
Future Generation Computer Systems The International Journal of eScience	43	1.118	17
Applied Sciences Basel	35	0.91	8
IEEE Internet of Things Journal	31	0.806	12
Nano Energy	28	0.728	17
Sustainable Cities and Society	28	0.728	14
Cluster Computing The Journal of Networks Software Tools and Applications	24	0.624	6
Electronics	24	0.624	4
Energy Conversion and Management	24	0.624	15
IEEE Transactions on Industrial Informatics	24	0.624	19
Renewable Energy	24	0.624	12
IEEE Communications Magazine	23	0.598	15
ACS Applied Materials Interfaces	22	0.572	13

TP, total number of papers published in this field by a journal; *Ratio (%)*, ratio of papers published on the field in a journal; *h-index*, h papers of a journal are cited at least h times

Sources: The authors developed this table based on results

et al. 2016; Krizhevsky et al. 2017), and optimal solution (Silver et al. 2016) in energy systems.

The development of AI cannot be separated from the support of big data technology. Big data-driven intelligent energy management places higher demands on IT infrastructure, data collection and sharing, processing and analysis, security, and privacy (Chen et al. 2014a, b; Zhou et al. 2016).

Building energy consumption and data center energy consumption are the focus of energy consumption. Building energy consumption is the main part of energy consumption. Due to its complexity and uncertainty of influencing factors, the review of building energy consumption prediction using the AI algorithms has become the basis of related research (Zhao et al. 2012). With the development of cloud computing technology, the problem of energy consumption in data centers is becoming more and more serious. An architectural framework and principles for energy-efficient cloud computing are defined. Energy-aware resource provisioning and allocation algorithms are presented (Beloglazov et al. 2012).

Reinforcement learning can not only utilize the existing data, but also obtain new data by exploring the environment. The new data are used to repeatedly update and iterate the machine learning algorithm of the existing model. Deep reinforcement learning consists of all reinforcement learning algorithms that use neural network as the parameter structure to be optimized. It provides a theoretical basis for the realization of intelligent energy systems (Mnih et al. 2015).

Research hotspots

Keywords provide a high-level abstraction of the core content of an article. When two keywords appear in the same article with a high frequency, the research direction represented by the keywords is considered a current research hotspot. In the keyword co-occurrence graph, a total of 34 high-frequency keywords were obtained by cleaning and merging keywords, and reserving nodes having word frequencies higher than 20 (Table 6). The titles, keywords, and

Table 5 Most frequently co-cited articles during 2001–2020

Category	Frequency	Author	Year	Title
Deep learning (293)	97	LeCun Y	2015	Deep learning
	61	He KM	2016	Deep residual learning for image recognition
	56	Krizhevsky A	2012	ImageNet classification with deep convolutional neural networks
	44	Silver D	2016	Mastering the game of Go with deep neural networks and tree search
	35	Schmidhuber J	2015	Deep learning in neural networks: an overview
Big data (105)	62	Zhou KL	2016	Big data driven smart energy management: from big data to big insights
	43	Chen M	2014	Big data: a survey
Energy consumption (79)	43	Zhao HX	2012	A review on the prediction of building energy consumption
	36	Beloglazov A	2012	Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing
Reinforcement learning (40)	40	Mnih V	2015	Human-level control through deep reinforcement learning

Frequency, frequency with which an article appears in a co-citation network

Sources: The authors developed this table based on results

Table 6 High frequency keywords on big data, AI, and energy

Frequency	Keyword	Frequency	Keyword
484	Artificial intelligence	43	Renewable energy
431	Big data	42	Energy management
293	Machine learning	36	Edge computing
210	Neural network	31	Sustainability
183	Energy efficiency	31	Clustering
171	Internet of thing	28	Feature extraction
104	Cloud computing	28	Data center
90	Smart grid	27	sensor
88	Deep learning	27	Smart meter
86	Energy consumption	26	Resource allocation
69	Optimization	25	Reinforcement learning
65	Data mining	25	Energy saving
64	Prediction	25	Security
60	Smart city	22	Electric vehicle
53	Wireless sensor network	22	Fog computing
52	genetic algorithm	22	Fuzzy logic
46	energy	22	Mapreduce

Sources: The authors developed this table based on results

abstracts usually contain terms that are typical in biomass energy research. The keywords "Biomass Energy" and "Renewable Energy" appeared more frequently than other keywords. By analyzing the highly cited articles with these high frequency keywords, it was found that the application of big data and AI in the field of energy included the following three aspects: smart grid, energy consumption, and renewable energy.

Smart grid

A smart grid is a complicated, interconnected power grid that is one of the important applications of big data and AI technology in the field of energy. The smart gird employs sensors, deployment strategies, smart meters, and real-time data processing to deliver secure energy supply. Research areas relating to smart grid include integrated architecture, security issues, and key-enabling technologies.

The main challenges of smart grids include efficiently managing different types of front-end intelligent devices and processing a huge amount of data received from these devices. The service and technology patterns of cloud computing can provide a computing platform for different types of computing services (Baek et al. 2015, Munshi et al. 2017). Although energy planning can be rendered more efficient, accurate, and intelligent using big data based on cloud computing, different issues involving the privacy of

end users and the safe operation of critical infrastructure must be addressed (Hu et al. 2016).

AI will become a core part of the development of smart grid with the support of cloud computing and big data. Based on the big data cloud platform, a deep neural network can forecast the price and energy demand for improving the profitability of service providers and customers (Lu et al. 2019). Clustering algorithms can be applied to household consumption data and help in distinguishing user types for tariff design and switching, fault, and fraud detection (Granell et al. 2015). The analysis of smart meter data (Koutitas et al. 2016) and detection and dealing with missing or inaccurate data (Peppanen et al. 2015) is another research hotspot around smart grid.

Energy consumption

Research on energy consumption mainly includes the forecasting of energy consumption and energy saving. Big data technology can centrally manage a decentralized network data. On this basis, improved machine learning algorithms can accurately predict and reduce energy consumption.

Energy consumption forecasting is mainly divided into methodological oriented research and energy consumption-related issue-oriented research. One of the most commonly used methods for energy consumption forecasting is the time series analysis, which uses the characteristics of a certain event in the past to predict its characteristics in the future. Time series forecasting models commonly used in energy consumption forecasting problems include traditional forecasting methods, such as seasonal autoregressive integrated moving average (SARIMA), and deep learning algorithms, such as recurrent neural networks (RNN) and long short-term memory (LSTM). By integrating a SARIMA model and metaheuristic firefly algorithm-based least squares support vector regression model, a novel time-series sliding window metaheuristic optimization-based machine learning system can use data collected by a smart grid to effectively predict building energy consumption (Chou et al. 2016). Long short-term memory is a kind of time cycling neural network that allows the analysis of ordered data with long-term dependence. An energy consumption forecasting model that employs LSTM and improved sine cosine optimization algorithm shows higher accuracy and stability for building energy consumption forecasting (Somu et al. 2020).

Problem-oriented research analyzes the characteristics of energy consumption-related problems and subsequently designs appropriate forecasting models. An adaptive neuro-fuzzy reasoning system can be used to predict the wheat yield according to the energy input data (Naderloo et al. 2012). The impacts of bike sharing on energy use are related to the emission of greenhouse gases, and using big data technology to evaluate the impact of energy consumption on the

environment is also a topic of academic interest (Zhang et al. 2018). In many cases, data acquisition is a challenge in the way of using AI for energy demand forecasting. For a limited data set, a novel hybrid dynamic approach that combines a dynamic grey model with genetic programming is more suitable for energy consumption prediction (Lee et al. 2012).

Buildings are key players of end-use energy demand and, therefore, the Internet of Things (IoT) has been used to monitor and control a large variety of energy-related agents in the buildings for energy reduction (Terroso-Saenz et al. 2019). A sensor is an enabler for smart industrial IoT. However, the computation ability and energy consumption of sensor nodes are challenged due to complex intelligent algorithms. Research on edge computing in the energy field has been increasing since 2018. Mobile edge nodes with relatively strong computation and storage ability can provide intelligent trust evaluation for sensor nodes. Therefore, mobile edge computing can help in scheduling of the moving path for the edge nodes to decrease the moving distance and energy consumption (Wang et al. 2020). At the same time, edge computing can cooperate with cloud computing to transfer part of intensive computing and storage resources to edge devices for energy consumption optimization (Huang et al. 2020).

Renewable energy

Long-term forecasting of renewable energy sources such as wind, water, and solar has become a research hotspot in academia because these sources have the characteristics of instability. However, the construction and stability of renewable energy generation systems significantly impacts the application of renewable energy. Thus, the optimal configuration and maintenance of generation systems using big data and AI are getting more and more attention.

A data fusion algorithm based on several neural networks can provide long-term wind speed forecast and avoid inefficient and less reliable results (Azad et al. 2014). The support vector machine (SVM) with radial basis function (RBF) as a kernel function can effectively predict global solar radiation (Ramedani et al. 2014).

Suitability analysis of a renewable energy power generation system is the basis of the normal system operation. The first suitability map of potential geothermal sites at a global scale based on a maximum entropy model is presented in (Coro et al. 2020). At the same time, the generation system scale also directly affects the later operation. The simulated annealing and a combination of simulated annealing with harmony search and chaotic search have proved to be effective in determining the most suitable size of autonomous hybrid photovoltaic/wind turbine/fuel cell system for electrification (Maleki et al. 2016).

Energy storage system is an important component of a renewable energy system due to its inherent instability. Research on the application of AI in the energy storage system focuses on remote monitoring and battery maintenance. An echo state network (ESN)-based Q-learning method has been developed for optimal energy management in a solar power supply system with a battery containing a control unit (Shi et al. 2017). For micro power plants, a decentralized cloud system based on general systems and remote telemetry units (RTUs) can be used to gather data for tele-monitoring because only satellite or sparse GSM radio signals are available (Suciu et al. 2016).
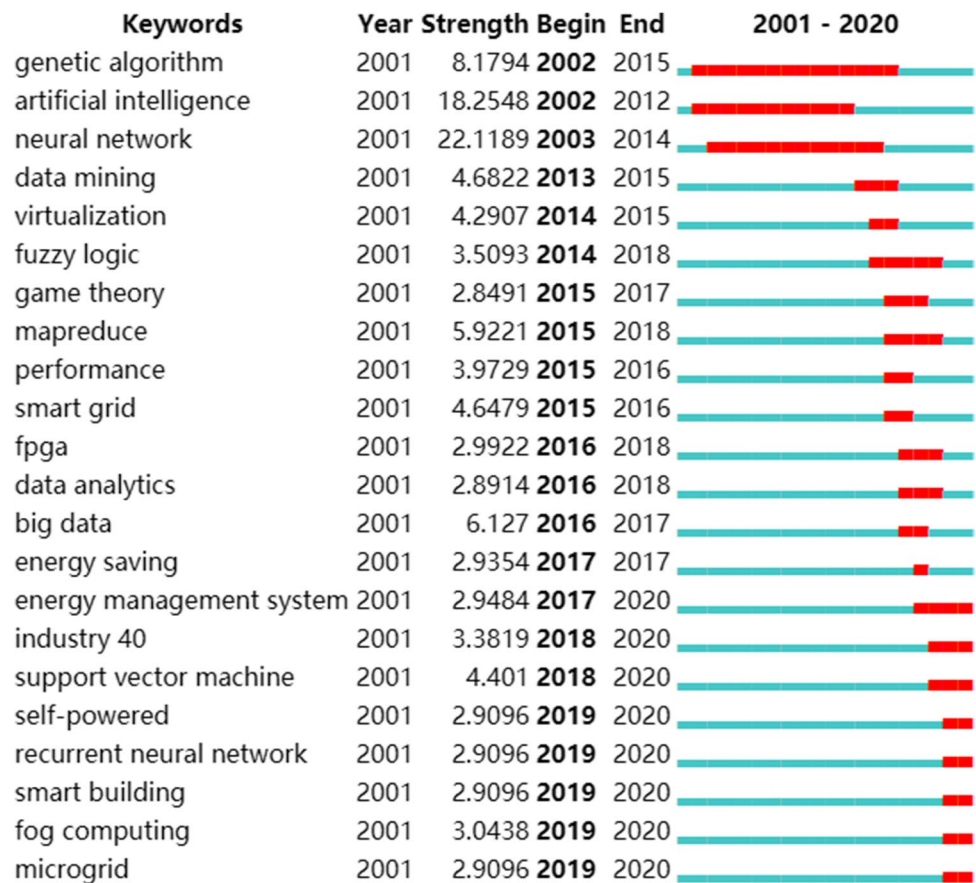
Evolutionary trajectory of the research frontier

The burst words in the field of big data, AI, and energy were detected using the keyword burst detection in CiteSpace. These words show a significant change in the number of keyword citations from 2001 to 2020. The burst and decline of the research direction represented by the keywords and the evolutionary path of the research frontier were then identified, which are shown in Fig. 5. The "strength" reflects the burst intensity of each keyword. The terms "begin" and "end" represent the beginning and ending years of the burst of the keywords, respectively, corresponding to the red area in the figure. According to the beginning time of burst words, the research on big data and AI in the energy field can be divided into three stages, which are described in the following sub-sections.

First stage (2001–2012)

The early research frontiers involved the genetic algorithm, AI, and neural networks, which have the longest duration of burst. The genetic algorithm and neural networks belong to the intelligent computing method and the category of AI. The genetic algorithm simulates the process of biological evolution in nature (Mukherjee 2002), where an individual entity evolves until the global optimal value is found. It can effectively solve the highly nonlinear optimization problems that the conventional optimization algorithms cannot solve. These problems include system optimization for improving energy efficiency (Caldas et al. 2002, Chen et al. 2002), and improving the energy system stability (Soliman et al. 2018). In recent years, the genetic algorithm is increasingly being combined with other AI algorithms to solve energy-related optimization problems (Krzywanski 2019; Su et al. 2019).

Compared with the genetic algorithm, neural networks were often used in the prediction problems. One of the main issues was forecasting the potential of renewable energy. Due to the unstable and intermittent nature of wind power, neural networks were often used to predict wind speed in order to avoid inefficient and unreliable results (Mabel et al.,

Fig. 5 Top 22 keywords with the strongest citation bursts

Keywords	Year	Strength	Begin	End	2001 - 2020
genetic algorithm	2001	8.1794	**2002**	2015	
artificial intelligence	2001	18.2548	**2002**	2012	
neural network	2001	22.1189	**2003**	2014	
data mining	2001	4.6822	**2013**	2015	
virtualization	2001	4.2907	**2014**	2015	
fuzzy logic	2001	3.5093	**2014**	2018	
game theory	2001	2.8491	**2015**	2017	
mapreduce	2001	5.9221	**2015**	2018	
performance	2001	3.9729	**2015**	2016	
smart grid	2001	4.6479	**2015**	2016	
fpga	2001	2.9922	**2016**	2018	
data analytics	2001	2.8914	**2016**	2018	
big data	2001	6.127	**2016**	2017	
energy saving	2001	2.9354	**2017**	2017	
energy management system	2001	2.9484	**2017**	2020	
industry 40	2001	3.3819	**2018**	2020	
support vector machine	2001	4.401	**2018**	2020	
self-powered	2001	2.9096	**2019**	2020	
recurrent neural network	2001	2.9096	**2019**	2020	
smart building	2001	2.9096	**2019**	2020	
fog computing	2001	3.0438	**2019**	2020	
microgrid	2001	2.9096	**2019**	2020	

2008, Azad et al. 2014). Neural networks also proved to be an efficient and easy methodology for solar radiation measurement (Bosch et al. 2008; Li et al. 2016). Building energy consumption is an important part of energy consumption. Therefore, neural networks have emerged as a key method to address the nonlinearity of building energy data and obtain robust calculation of large and dynamic data (Yezioro et al. 2008; Platon et al. 2015; Biswas et al. 2016).

Second stage (2013–2017)

The application of big data technology in the energy field has been gaining prominence since 2013. A big data technology system is huge and complex, and related technologies emerging in the second phase include data mining (Chou and Bui 2014; Pan et al. 2015) and analysis, virtualization, and MapReduce. The concept of energy big data refers to the relevant technologies and ideas for comprehensive collection, processing, analysis, and application of data in energy fields, such as electric power (Yu et al. 2016), energy-intensive industries (Zhang et al. 2018), and renewable energy (Buffat et al. 2018; Sharifzadeh et al. 2019). The most popular research directions on energy big data include smart power grid (Baek et al. 2015; He et al. 2017) and building energy consumption (Plageras et al. 2018) aimed at energy saving.

The collection of energy big data is the basis of data mining and analysis in the energy field. The cloud computing environments and the MapReduce parallel programming model provide solutions for the mining and analysis of massive data (Mashayekhy et al. 2015) (Feller et al. 2015). Hadoop is an open source implementation of MapReduce (Feller et al. 2015). However, not all enterprises have sufficient resources and ability to deal with the challenges brought by Hadoop deployment. In this case, the introduction of virtualization provides a feasible solution. Virtualization can be utilized to reduce the energy consumption of clouds by using server consolidation without degrading the performance of users applications (Kansal et al. 2015, Shen et al. 2015).

Third stage (2018–2020)

Research on big data and AI in the field of energy has been expanding since 2018. From the technical point of view, SVM, convolutional neural network (CNN), and fog computing are emerging frequently. Support vector machine is one of the most popular machine-learning models. This algorithm performs remarkably for solving the problems in the energy field, including prediction of the heat load in heating systems (Gu et al. 2018), classification of diode rectifier

faults in brushless synchronous generator (Rahnama et al. 2019), and prediction of building energy usage for heating (D'Oca et al. 2015).

As one of the representative algorithms of deep learning, CNN have achieved outstanding results in computer vision, classification, and other fields at the expense of a huge amount of parametric modeling and computations. Therefore, low power and high-efficiency design is a necessary condition for productization when deploying these algorithms in actual applications (Ha et al. 2019; Pang et al. 2020). The CNN have also been applied for classification and prediction in the energy field, including short-term load forecasting of power systems (Bendaoud et al. 2020), prediction of wind speed (Ehsan et al. 2020), and nuclear fault diagnosis system (Yao et al. 2020).

Cloud computing plays a vital role in processing a large amount of data. Compared to cloud computing, fog computing can support delay-sensitive service requests from end-users, which reduces energy consumption and traffic congestion (Fu et al. 2018; Mukherjee et al. 2018).

At present, the research on big data and AI in the field of energy mainly aims to reduce energy consumption. This research can be divided into three categories. First, microgrid is regarded as an important part of the energy Internet in the industry. It has become a strategic choice for countries around the world to realize energy transformation and achieve optimal integration of distributed generation (Mahmoud et al. 2019; Zhou et al. 2019). Second, building energy consumption contributes the most to energy consumption. Smart buildings integrate AI technologies with energy consumption management systems, which can simultaneously manage the energy consumption effectively and also increase the user comfort (Gonzalez-Vidal et al. 2019; Wahid et al. 2019). Third, with the development of the IoT, the sensor industry will also undergo significant expansion, further promoting the application of self-powered systems in various industries (Chen et al. 2019; Zhong et al. 2019).

Conclusion

The characteristics of big data, AI, and energy-related literature from 2001 to 2020 were examined based on the SCI-E and SSCI databases using bibliometric methods. This paper revealed that the literature in this field had become more extensive over the past 20 years. The pace of publishing in this field increased rapidly, especially starting from 2014.

Based on the country analysis, China was shown to be one of the most important contributors to the literature on big data, AI, and energy with the highest number of publications (1115), followed by the USA and the UK. However, the USA had the highest h-index. The institution analysis showed that the Chinese Acad. Sci. made the most

significant contributions in the research field. Similarly, it played a vital role in the collaborative network of 30 productive institutions. It developed strong cooperation relationships with other institutes, especially with the Univ. Chinese Acad. Sci. in China and Georgia Inst. Techno. in the USA.

The study also revealed that engineering was the most popular discipline with the highest number of publications on big data, AI, and energy. Furthermore, 20 core journals contributed about 26.6% of the total number of journal publications, with a significantly higher number of papers published in the "IEEE Access" and "Energies" journals compared to other journals.

The results of co-citation analysis showed that deep learning, deep reinforcement learning, energy big data, and prediction of energy consumption had laid a solid theoretical foundation in this field. Keyword co-occurrence analysis showed a popular research topic that the application of big data and AI on energy focused on smart grid, energy consumption, and renewable energy. The burst detection of keywords revealed that the research evolved through three stages. Early research frontiers involved optimization and prediction of energy-related problems solved using the genetic algorithm and neural networks. Energy big data involving data mining, virtualization, and MapReduce gained prominence since 2013. At present, machine learning, deep learning, and fog computing are frequently combined with energy, including microgrid, smart building, and self-powered systems. These studies will lay a solid foundation for application of big data and AI to renewable energy and energy-saving renovation of buildings in the future.

The findings presented in this paper can help scholars understand the developmental trends for big data, AI, and energy, and correctly grasp the research direction and method of the emerging interdisciplinary field with many cross-industry professional concepts. They can also assist energy enterprises in finding the right direction for technology investment.

Data availability The datasets used and analyzed during the current study can be provided on reasonable request.

Declarations

Ethical approval The authors declare that the work described was original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part.

Consent to participate There are no individual participants included in the study.

Consent for publication The manuscript is approved by all authors for publication.

Competing interests The authors declare no competing interests.

References

Abid MR, Lghoul R, Benhaddou D (2017) ICT for renewable energy integration into smart buildings: IoT and Big Data Approach. 2017 Ieee Africon. D. R. Cornish**:** 856–861

Azad HB, Mekhilef S, Ganapathy VG (2014) Long-term wind speed forecasting and general pattern recognition using neural networks. IEEE Trans Sustain Energy 5(2):546–553

Baek J, Vu QH, Liu JK et al (2015) A secure cloud computing based framework for big data information management of smart grid. Ieee Trans Cloud Comput 3(2):233–244

Beloglazov A, Abawajy J, Buyya R (2012) Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Gen Comput Syst Intl J Esci 28(5):755–768

Bendaoud NMM, Farah N (2020) Using deep learning for short-term load forecasting. Neural Comput Appl 32(18):15029–15041

Biswas MAR, Robinson MD, Fumo N (2016) Prediction of residential building energy consumption: a neural network approach. Energy 117:84–92

Bosch JL, Lopez G, Batlles FJ (2008) Daily solar irradiation estimation over a mountainous area using artificial neural networks. Renewable Energy 33(7):1622–1628

Buffat R, Grassi S, Raubal M (2018) A scalable method for estimating rooftop solar irradiation potential over large regions. Appl Energy 216:389–401

Caldas LG, Norford LK (2002) A design optimization tool based on a genetic algorithm. Autom Constr 11(2):173–184

Chen HT, Song Y, Cheng XL et al (2019) Self-powered electronic skin based on the triboelectric generator. Nano Energy 56:252–268

Chen K, Lin GT (2002) Optimization of multiple-module thermoelectric coolers using artificial-intelligence techniques. Int J Energy Res 26(14):1269–1283

Chen M, Mao S, Liu Y (2014a) Big data: a survey. Mob Netw Appl 19(2):171–209

Chen Y, Chen C, Hu Z et al (2014b) Principles and applications of analyzing a citation space

Cheng LF, Yu T (2019) A new generation of AI: a review and perspective on machine learning technologies applied to smart energy and electric power systems. Int J Energy Res 43(6):1928–1973

Chou JS, Bui DK (2014) Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. Energy Build 82:437–446

Chou JS, Ngo NT (2016) Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns. Appl Energy 177:751–770

Coro G, Trumpy E (2020) Predicting geographical suitability of geothermal power plants. J Clean Prod 267:11

D'Oca S, Hong TZ (2015) Occupancy schedules learning process through a data mining framework. Energy Build 88:395–408

Daut MAM, Hassan MY, Abdullah H et al (2017) Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: a review. Renew Sustain Energy Rev 70:1108–1118

Ehsan MA, Shahirinia A, Zhang N et al (2020) Investigation of data size variability in wind speed prediction using AI algorithms. Cybern Syst 52(1):105–126

Feller E, Ramakrishnan L, Morin C (2015) Performance and energy efficiency of big data applications in cloud environments: a Hadoop case study. J Parallel Distrib Comput 79–80:80–89

Fu JS, Liu Y, Chao HC et al (2018) Secure data storage and searching for industrial IoT by integrating fog computing and cloud computing. IEEE Trans Industr Inf 14(10):4519–4528

Giest S, Mukherjee I (2018) Behavioral instruments in renewable energy and the role of big data: a policy perspective. Energy Policy 123:360–366

Gonzalez-Vidal A, Jimenez F, Gomez-Skarmeta AF (2019) A methodology for energy multivariate time series forecasting in smart buildings based on feature selection. Energy Build 196:71–82

Granell R, Axon CJ, Wallom DCH (2015) Impacts of raw data temporal resolution using selected clustering methods on residential electricity load profiles. IEEE Trans Power Syst 30(6):3217–3224

Gu JH, Wang J, Qi CY et al (2018) Medium-term heat load prediction for an existing residential building based on a wireless on-off control system. Energy 152:709–718

Ha M, Byeon Y, Kim J et al (2019) Selective deep convolutional neural network for low cost distorted image classification. IEEE Access 7:133030–133042

He K, Zhang X, Ren S et al (2016) Deep residual learning for image recognition. IEEE Conf Comput Vis Pattern Recognit 2016:770–778

He X, Ai Q, Qiu RC et al (2017) A big data architecture design for smart grids based on random matrix theory. IEEE Trans Smart Grid 8(2):674–686

Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA 102(46):16569–16572

Hu JK, Vasilakos AV (2016) Energy big data analytics and security: challenges and opportunities. IEEE Trans Smart Grid 7(5):2423–2436

Huang MF, Liu W, Wang T et al (2020) A cloud-MEC collaborative task offloading scheme with service orchestration. IEEE Internet Things J 7(7):5792–5805

IEA (2015). Energy and climate change: world energy outlook special report 2015. International Energy Agency

Ifaei P, Farid A, Yoo C (2018) An optimal renewable energy management strategy with and without hydropower using a factor weighted multi-criteria decision making analysis and nationwide big data - Case study in Iran. Energy 158:357–372

Jha SK, Bilalovic J, Jha A et al (2017) Renewable energy: present research and future scope of artificial intelligence. Renew Sustain Energy Rev 77:297–317

Kaab A, Sharifi M, Mobli H et al (2019) Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. Sci Total Environ 664:1005–1019

Kansal NJ, Chana I (2015) Artificial bee colony based energy-aware resource utilization technique for cloud computing. Concurr Comput Pract Exper 27(5):1207–1225

Kleinberg J (2003) Bursty and hierarchical structure in streams. Data Min Knowl Disc 7(4):373–397

Koutitas GC, Tassiulas L (2016) Low cost disaggregation of smart meter sensor data. IEEE Sens J 16(6):1665–1673

Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Commun ACM 60(6):84–90

Krzywanski J (2019) A general approach in optimization of heat exchangers by bio-inspired artificial intelligence methods. Energies 12(23):32

Kumar NM, Chand AA, Malvoni M et al (2020) Distributed energy resources and the application of AI, IoT, and blockchain in smart grids. Energies 13(21)

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444

Lee YS, Tong LI (2012) Forecasting nonlinear time series of energy consumption using a hybrid dynamic model. Appl Energy 94:251–256

Li JM, Ward JK, Tong JN et al (2016) Machine learning for solar irradiance forecasting of photovoltaic system. Renewable Energy 90:542–553

Lu RZ, Hong SH (2019) Incentive-based demand response for smart grid with reinforcement learning and deep neural network. Appl Energy 236:937–949

Mabel MC, Fernandez E (2008) Analysis of wind power generation and prediction using ANN: a case study. Renewable Energy 33(5):986–992

Mahmoud TS, Ahmed BS, Hassan MY (2019) The role of intelligent generation control algorithms in optimizing battery energy storage systems size in microgrids: a case study from Western Australia. Energy Convers Manage 196:1335–1352

Maleki A, Pourfayaz F, Rosen MA (2016) A novel framework for optimal design of hybrid renewable energy-based autonomous energy systems: a case study for Namin. Iran Energy 98:168–180

Mashayekhy L, Nejad MM, Grosu D et al (2015) Energy-aware scheduling of MapReduce jobs for big data applications. IEEE Trans Parallel Distrib Syst 26(10):2720–2733

Mehmood MU, Chun D, Zeeshan, et al (2019) A review of the applications of artificial intelligence and big data to buildings for energy-efficiency and a comfortable indoor living environment. Energy and Buildings 202:13

Mnih V, Kavukcuoglu K, Silver D et al (2015) Human-level control through deep reinforcement learning. Nature 518(7540):529–533

Mukherjee B (2002) A high-resolution neutron spectra unfolding method using the genetic algorithm technique. Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment 476(1–2):247–251

Mukherjee M, Shu L, Wang D (2018) Survey of fog computing: fundamental, network applications, and research challenges. IEEE Commun Surv Tutor 20(3):1826–1857

Munshi AA, Mohamed Y (2017) Big data framework for analytics in smart grids. Electr Power Syst Res 151:369–380

Nabavi-Pelesaraei A, Rafiee S, Mohtasebi SS et al (2018) Integration of artificial intelligence methods and life cycle assessment to predict energy output and environmental impacts of paddy production. Sci Total Environ 631–632:1279–1294

Naderloo L, Alimardani R, Omid M et al (2012) Application of ANFIS to predict crop yield based on different energy inputs. Measurement 45(6):1406–1413

Pan SY, Morris T, Adhikari U (2015) Developing a hybrid intrusion detection system using data mining for power systems. IEEE Trans Smart Grid 6(6):3104–3113

Pang W, Wu CL, Lu SL (2020) An energy-efficient implementation of group pruned CNNs on FPGA. IEEE Access 8:217033–217044

Peppanen J, Reno MJ, Thakkar M et al (2015) Leveraging AMI data for distribution system model calibration and situational awareness. IEEE Trans Smart Grid 6(4):2050–2059

Plageras AP, Psannis KE, Stergiou C et al (2018) Efficient IoT-based sensor BIG data collection-processing and analysis in smart buildings. Future Gen Comput Syst Intl J Esci 82:349–357

Platon R, Dehkordi VR, Martel J (2015) Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis. Energy and Buildings 92:10–18

Puri V, Jha S, Kumar R et al (2019) A hybrid artificial intelligence and Internet of Things model for generation of renewable resource of energy. IEEE Access 7:111181–111191

Rahnama M, Vahedi A, Alikhani AM et al (2019) Machine-learning approach for fault detection in brushless synchronous generator using vibration signals. IET Sci Meas Technol 13(6):852–861

Ramedani Z, Omid M, Keyhani A et al (2014) A comparative study between fuzzy linear regression and support vector regression for global solar radiation prediction in Iran. Sol Energy 109:135–143

Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117

Sharifzadeh M, Sikinioti-Lock A, Shah N (2019) Machine-learning methods for integrated renewable power generation: a comparative study of artificial neural networks, support vector regression, and Gaussian process regression. Renew Sustain Energy Rev 108:513–538

Shen D, Luo JZ, Dong F et al (2015) Stochastic modeling of dynamic right-sizing for energy-efficiency in cloud data centers. Future Gen Comput Syst Intl J Esci 48:82–95

Shi G, Liu DR, Wei QL (2017) Echo state network-based Q-learning method for optimal battery control of offices combined with renewable energy. IET Control Theory Appl 11(7):915–922

Silver D, Huang A, Maddison CJ et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature 529(7587): 484-+

Soliman MA, Hasanien HM, Azazi HZ et al (2018) Hybrid ANFIS-GA-based control scheme for performance enhancement of a grid-connected wind generator. IET Renew Power Gener 12(7):832–843

Somu N, Raman MRG, Ramamritham K (2020) A hybrid model for building energy consumption forecasting using long short term memory networks. Appl Energy 261:20

Su H, Zio E, Zhang JJ et al (2019) A hybrid hourly natural gas demand forecasting method based on the integration of wavelet transform and enhanced Deep-RNN model. Energy 178:585–597

Suciu G, Vulpe A, Martian A et al (2016) Big data processing for renewable energy telemetry using a decentralized cloud M2M system. Wireless Pers Commun 87(3):1113–1128

Sun QY, Yang LX (2019) From independence to interconnection-a review of AI technology applied in energy systems. CSEE J Power Energy Syst 5(1):21–34

Terroso-Saenz F, Gonzalez-Vidal A, Ramallo-Gonzalez AP et al (2019) An open IoT platform for the management and analysis of energy data. Future Gen Comput Syst Intl J Esci 92:1066–1079

Wahid F, Ismail LH, Ghazali R et al (2019) An efficient artificial intelligence hybrid approach for energy management in intelligent buildings. KSII Trans Internet Inf Syst 13(12):5904–5927

Wang T, Luo H, Jia WJ et al (2020) MTES: An intelligent trust evaluation scheme in sensor-cloud-enabled industrial Internet of Things. IEEE Trans Industr Inf 16(3):2054–2062

Wang ZY, Srinivasan RS (2017) A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. Renew Sustain Energy Rev 75:796–808

Yao YT, Wang J, Long PC et al (2020) Small-batch-size convolutional neural network based fault diagnosis system for nuclear energy production safety with big-data environment. Int J Energy Res 44(7):5841–5855

Yezioro A, Dong B, Leite F (2008) An applied artificial intelligence approach towards assessing building performance simulation tools. Energy and Buildings 40(4):612–620

Yu XH, Xue YS (2016) Smart grids: a cyber-physical systems perspective. Proc IEEE 104(5):1058–1070

Zhang YF, Ma SY, Yang HD et al (2018) A big data driven analytical framework for energy-intensive manufacturing industries. J Clean Prod 197:57–72

Zhang YP, Mi ZF (2018) Environmental benefits of bike sharing: a big data-based analysis. Appl Energy 220:296–301

Zhao HX, Magoules F (2012) A review on the prediction of building energy consumption. Renew Sustain Energy Rev 16(6):3586–3592

Zhao Y, Li TT, Zhang XJ et al (2019) Artificial intelligence-based fault detection and diagnosis methods for building energy systems: advantages, challenges and the future. Renew Sustain Energy Rev 109:85–101

Zhong TY, Guan HY, Dai YT et al (2019) A self-powered flexibly-arranged gas monitoring system with evaporating rainwater as fuel for building atmosphere big data. Nano Energy 60:52–60

Zhou K, Fu C, Yang S (2016) Big data driven smart energy management: from big data to big insights. Renew Sustain Energy Rev 56:215–225

Zhou SY, Hu ZJ, Gu W et al (2019) Artificial intelligence based smart energy community management: a reinforcement learning approach. CSEE J Power Energy Syst 5(1):1–10