Indonesian Journal of Computer Science
ISSN 2549-7286 (*online*)
Jln. Khatib Sulaiman Dalam No. 1, Padang, Indonesia
Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

Data Science for Energy Applications: A Bibliometric Analysis

Sello Prince Sekwatlakwatla[1], Vusumuzi Malele[2]
[1]sek.prince@gmail.com , [2]vusi.malele@nwu.ac.za
Unit for Data Science and Computing School of Computer Science and Information Systems
North-West University Vanderbijlpark, South Africa

Article Information

Submitted :24 Feb 2024
Reviewed: 6 Mar 2024
Accepted : 14 Mar 2024

Keywords

Bibliometric analysis
Data science, energy application

Abstract

Global digitalization is altering the energy sector, demanding the adoption of data science applications to improve efficiency and innovation, despite the industry's existing data analytics. Data science is revolutionizing the energy and utilities industries, enabling efficient, sustainable, and innovative decision-making through data analysis and smart grid optimization. In the energy industry, organizations are turning to data science to reduce waste, optimize energy usage, and provide alternative energy sources. With the different parts of Africa facing energy crises, different applications are needed to provide a solution. Data science has the potential to provide good information and knowledge that could be used to contribute to energy solutions. To address these concerns, data science models enable utilities to accurately forecast energy demand, enabling efficient generation, distribution, waste reduction, and informed investment decisions by leveraging historical consumption data, weather patterns, and economic indicators. This article aims to explore data science for energy applications. The findings show tools and techniques that can be utilized to provide energy efficiency and energy sustainability through data science applications.

A. Introduction

Data Science comprises different data related theories, methods, applications, tools, and technologies that generate, extract and analyse large data sets for obtaining meaningful insight that contribute business solutions[1-2]. The theories, methodologies, practices and applications used in Data Science are extracted from different multidisciplinary study areas like computer science, systems engineering, mathematics, statistics, etc. In this regard, Data Science uses models and algorithms emanating from artificial intelligence, machine learning, data analytics, data engineering, etc.

The application of Data Science could be traced in different fields that generate data like engineering, social science, humanities, health, law, business, etc. In this regard, Data Science could be applied in areas that could enhance the life of human beings [3]. Data science can be used to conduct: (i) Descriptive analysis (manipulation of data to get insights into the past and present happening); (ii) Diagnostic analysis (examining data to understand the reasons of why something happened); (iii) Predictive analysis (which uses historical data to make forecasts about the future); and (iv) Prescriptive analysis (uses predictive analysed data to suggests an best solution).

With the different parts of Africa facing the energy crises, different applications are needed to provide a solution. Data Science has a potential to provide good information and knowledge that could be used to contribute energy solution. Of interest is to understand how African researchers are using Data Science for enhancing solutions in energy applications. In this regard, this paper uses Bibliometric analysis to explore and analyzing data sets with an aim of unpacking the use of Data Science in energy studies especially the energy applications and also identify new trends and fields [4-5].

Afrrica faces energy crises, requiring diverse applications. Data science offers potential for energy solutions. Understanding African researchers' use of this technology is crucial [2-4].Casini et al. found that while deep learning methods are still in their early stages, an increasing trend in publications per year indicates their potential in energy applications[1], As artificial intelligence (AI) advances, it is increasingly being utilized to address environmental challenges such as energy, water, biodiversity, and transportation, particularly in the agricultural sector [5-6].

Yussuf and Asfour's study of AI's energy efficiency applications in buildings reveals challenges like user-friendly applications and integration and suggests future research on implementation, training, and security [5]. Data science for energy applications is proposed in the data center's energy consumption system, which primarily uses a redundant operation strategy, which results in a lack of cooperation between energy supply and data business. In order to reduce costs and implement energy management, operators should take into account energy supply advantages and time-of-day tariffs, Waste heat recovery should be adapted to local conditions [5-6].

A multi-energy complementary energy supply system design method accurately predicts airport energy consumption by analyzing in-depth data, determining an ideal system size, calculating economic and environmental benefits, and evaluating photovoltaic and geothermal power [8].

Data Science can be applied across various fields such as engineering, social science, humanities, health, law, and business to enhance human life through data-driven insights,Many industries use real-time big data analytics and complex event detection, including manufacturing, chemicals, food and beverage, automotive, aviation, maritime, drug development, and investment. Machine failures are predicted, critical events are prevented, product quality is optimized, and energy is reduced with these technologies [5-8].

Green energy and environmentally friendly decorative composite materials are proposed for building technologies, with future studies exploring a solution for optimizing energy consumption in post-merger farms [11]. A systematic literature review was conducted to identify techniques and technologies used in data science for energy applications, with results, methods, merits, and demerits highlighted in Table 1.

Table 1. The SLR on Data science for energy applications.

Title	Method	Merits	Demerits
Reducing energy consumption using heterogeneous voltage frequency scaling of data-parallel applications for multicore systems	DVFS (dynamic voltage frequency scaling) [13]	Substantial energy enhancement.	Clock frequency cannot be managed effectively
Energy Efficiency Solutions for Buildings: Automated Fault Diagnosis of Air Handling Units Using Generative Adversarial Networks	hybrid generative adversarial network (GAN) [14]	Improves the ability of the model to recognize traffic flow patterns.	A faulty training sample durin g training
Towards efficient and effective renewable energy prediction via deep learning	Convolutional neural network (CNN) [15]	Predictions for solar power generation and energy consumption.	The energy consumption prediction very high
A Bayesian Deep Learning Technique for Multi-Step Ahead Solar Generation Forecasting	Bayesian bidirectional long-short term memory (BiLSTM)[16]	Improve solar generation forecasting	Less precision
Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine	photovoltaic (PV) power output [17]	Prediction accuracy	Effect on the algorithm's overall performance.

Despite this introduction, the present document encompasses a section on research method, followed by the result, discussion, and conclusion.

B. Research Method

A total of three (3) analysis processes are proposed in order to achieve the study's objective, as indicated in Figure 1.

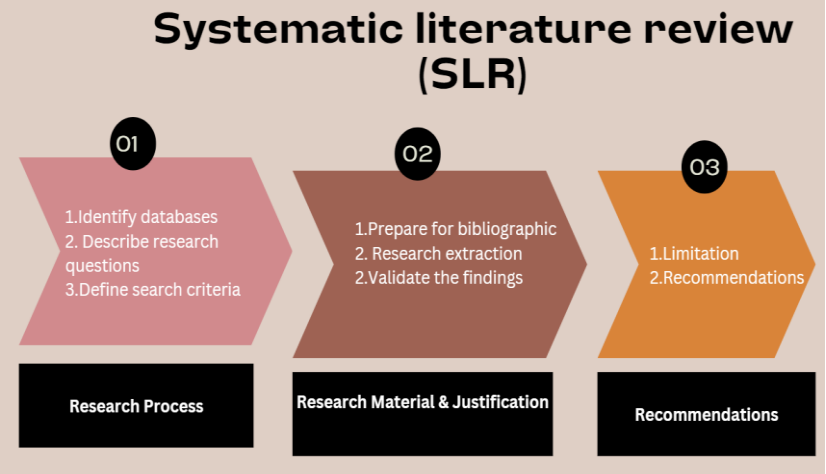

Figure 1. Proposed model

1. Research process

The study utilized databases like the Institute of Electrical and Electronics Engineering (IEEE), Web of Service, and Scopus to explore data science for energy applications. The search criteria used to find out information was "Data science for energy applications. "Databases were accessed on January 21, 2024. The data selected was from January 20, 2021, to December 30, 2023. This includes conferences, journals, early-access articles, and magazines that were selected for this research. In order to conduct this research, the following research question is used: What techniques are used in data science for energy applications?

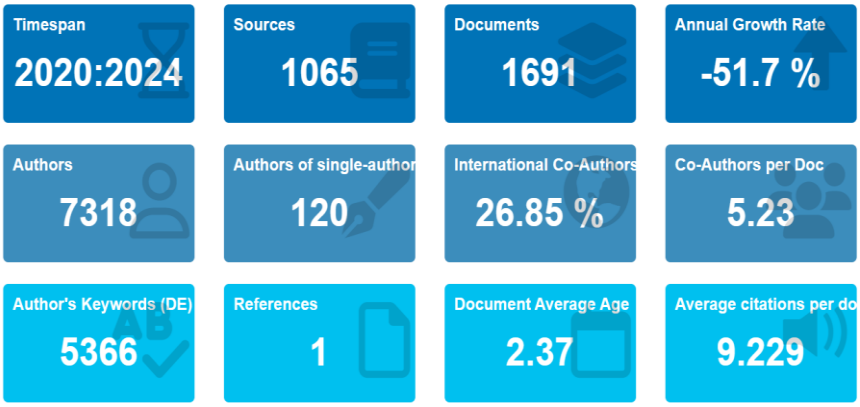

Figure 2. Summary of analysis

The analysis's summary is shown in Figure 2. Based on the downloaded documents, which spanned the years 2020 to 2024, 1691 papers with international co-authors (26.75%) and 7318 authors were cited.

2. Research Material and justification

The process involves data preparation, data analysis, and output, renaming the dataset to BibTeX for bibliographic organization, extracting valuable information, and downloading graphs for decision-making.

3. Recommendations

From the results of the bibliometric analysis, this section identifies the techniques and discusses their limitations.

C. Result and Discussion

Figure 3 indicates data science for energy applications with an energy efficiency of 7%, that the energy utilization in Africa is 6%, and that 5% of digital signage is also highlighted as a need for data science for energy applications.

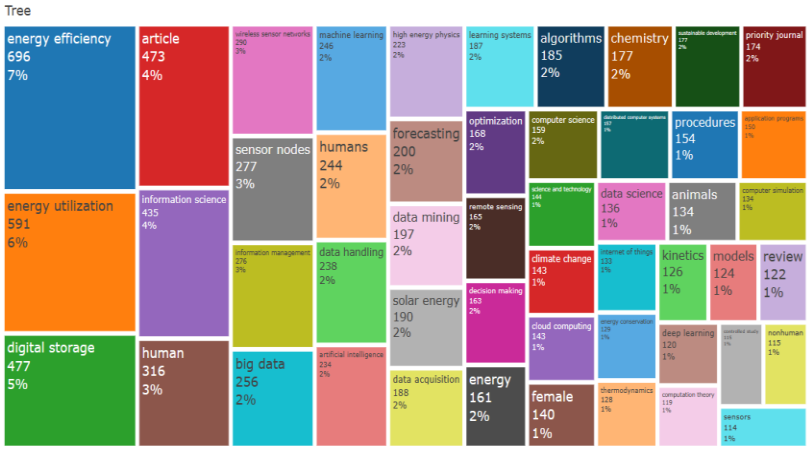

Figure 3. Analysis on Data science for energy applications

This figure (4) shows the most common research words from bibliometrics analysis, and the purpose of this section is to identify the most popular tools or techniques to support data science energy applications.

Figure 4. Frequently search words

As part of the bibliometric analysis, tools and techniques highlighted for energy efficiency, energy sustainability, and data science applications are machine

learning, digital storage, artificial intelligence, remote sensing, data mining, wireless sensor networks, and solar energy algorithms(see Figure 4).

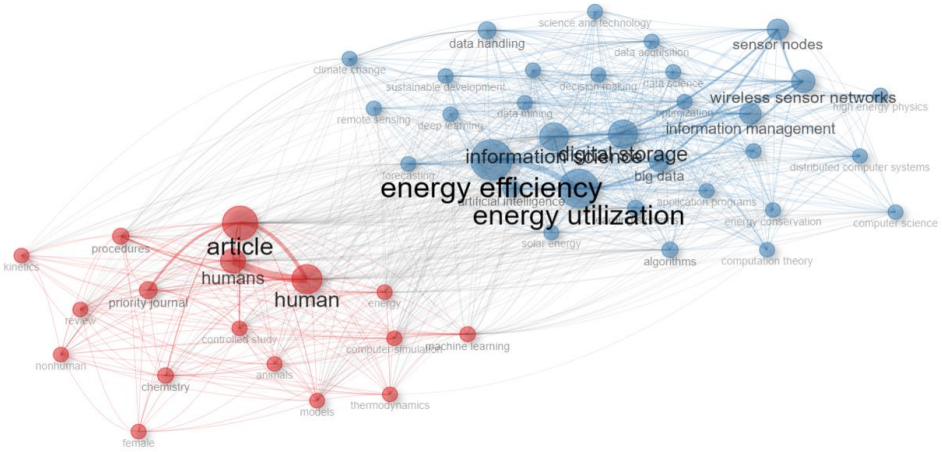

Figure 5. Data science for energy applications

According to Figure 5, there is a relationship between energy efficiency and energy integration with the support of the tools and techniques that are indicated in Figure 4.

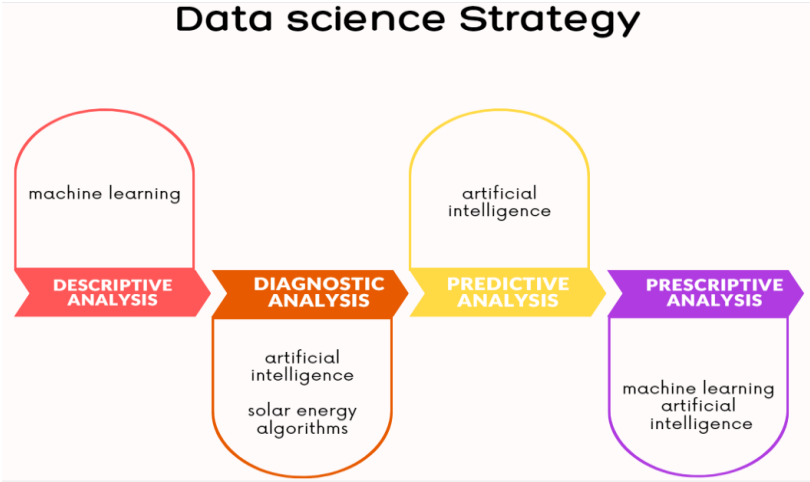

Figure 6. Data science strategy

Data science can be used to conduct: (i) descriptive analysis (manipulation of data to get insights into the past and present happening) machine learning; (ii) diagnostic analysis (examining data to understand the reasons why something happened) artificial intelligence and solar energy algorithms; (iii) predictive analysis (which uses historical data to make forecasts about the future) artificial intelligence; and (iv) predictive analysis (using predictively analyzed data to suggest the best solution) machine learning and artificial intelligence(see figure 6).

Afrrica faces energy crises, requiring diverse applications. Data science offers potential for energy solutions. Understanding African researchers' use of this technology is crucial, this study analysed data from databases like the Institute of Electrical and Electronics Engineering (IEEE), Web of Service, and Scopus to explore data science for energy applications.The findings reveal that tools and techniques highlighted for energy efficiency, energy sustainability in data science

applications are machine learning, digital storage, artificial intelligence, remote sensing, data mining, wireless sensor networks, and solar energy algorithms.

As a result, it is necessary to take into account the limitations of bibliometric analysis when interpreting the findings. The Institute of Electrical and Electronics Engineering (IEEE), Web of Service, and Scopus was used as a database in the study, but PubMed, and Google Scholar were not included,

It is likely that influential publications were missed, so future researchers need to combine more databases. They can also analyze data with bibliometric software such as Bibexcel and VOSviewer. In future studies, a comprehensive search that includes all articles is needed to increase the number of studies reviewed. The results may not reflect all studies on the adoption of data science for energy applications.

As African countries face energy crises in different parts of the country, different applications are needed to provide solutions. There is great potential for data science to provide good information and knowledge that can contribute to energy solutions in African countries, so they can provide energy efficiency and energy sustainability through data science applications. Machine learning, digital storage, artificial intelligence, remote sensing, data mining, wireless sensor networks, and solar energy algorithms are recommended as techniques to enhance the energy strategy.

D. Conclusion

Data science could be applied in areas that could enhance the lives of human beings and can be used to conduct descriptive analysis and manipulation of data to get insights into past and present happenings and Diagnostic analysis in order to examine data to understand the reasons why something happened and also Predictive analysis, which uses historical data to make forecasts about the future, and finally, predictive analysis, which uses predictively analyzed data to suggest the best solution, In this regard, this paper proposes machine learning, digital storage, artificial intelligence, remote sensing, data mining, wireless sensor networks, and solar energy algorithms as techniques to enhance the energy strategy.

Since in this paper only IEEE Xplore and Scopus database was used, in the future this study will be expanded by including another research database. Future researchers can also analyses the data using bibliometric software like Bibexcel and VOSviewer. The results may not reflect all studies on adoption of ICT. Future studies will need a more comprehensive search that includes all articles in order to increase the number of studies reviewed.

Since this paper uses only The Institute of Electrical and Electronics Engineering (IEEE), Web of Service, and Scopus as databases, but PubMed and Google Scholar were not included, it is likely that influential publications were missed, so future researchers need to combine more databases. They can also analyze data with bibliometric software such as Bibexcel and VOSviewer. In future studies, a comprehensive search that includes all articles is needed to increase the

number of studies reviewed. The results may not reflect all studies on the adoption of data science for energy applications.

E. Acknowledgment

The paper is part of the North-West University PhD program in Computer and Information Sciences with Information Technology.

F. References

[1] Mattia Casini, Paolo De Angelis, Eliodoro Chiavazzo and Luca Bergamasco. Current trends on the use of deep learning methods for image analysis inenergy applications, journal of Energy and AI,15 (2024) [Online].Available:. https://doi.org/10.1016/j.egyai.2023.100330

[2] Ephraim Daka ,Adopting Clean Technologies to Climate Change Adaptation Strategies in Af-rica: a Systematic Literature Review, Journal of Environmental Management,vol.4,no.71 pp 87–98,2023, https://doi.org/10.1007/s00267-022-01704-w

[3] Masud Kabir and Sami Ekici ,Energy-agriculture nexus: Exploring the future of artificial Intelligence applications journal of Energy Nexus, vol, no.13.[Online].Available:. https://doi.org/10.1016/j.nexus.2023.100263

[4] Désiré.Avom et al.,ICT and environmental quality in Sub-Saharan Africa: Effects and trans-mission channels", journal of Technological Forecasting & Social Change,vol,no.155,2020, https://doi.org/10.1016/j.techfore.2020.120028

[5] Raheemat Yussuf and Omar Asfour,Applications of artificial intelligence for energy efficiency throughout the building lifecycle: An overview ",journal of Energy & Buildings. vol,no.305,2024 [Online].Available:.https://doi.org/10.1016/j.enbuild.2024.113903

[6] Xue et al., "Energy Internet: A Novel Green Roadmap for Meeting the Global Energy Demand," 2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2), Taiyuan, China, 2021, pp. 3855-3860, doi: 10.1109/EI252483.2021.9713467.

[7] Shahzad Hussain, Raazia Gul and Sabeeh Ullah ,Role of financial inclusion and ICT for sustainable economic development in developing countries "journal of Technological Forecasting & Social Change, vol,no.194,2023. .[Online].Available:. https://doi.org/10.1016/j.techfore.2023.122725

[8] Qin Jia; Jing-Lei Yu and Zan Tao,Airports Integrated Multi-energy Complementary Energy System Design and Multi-scenario Application Analysis," 2023 International Conference on Power Energy Systems and Applications (ICoPESA), Nanjing, China, 2023, pp. 514-519, doi: 10.1109/ICoPESA56898.2023.10141512.

[9] Ralf Klinkenberg,Real-time big data stream analytics and complex event detection: modular visual framework, data science platform, and industry applications"DEBS '21: Proceedings of the 15th ACM International Conference on Distributed and Event-based SystemsJune 2021, pp 102 [Online].Available:.https://doi.org/10.1145/3465480.3468676

[10] Pratik Kumar Singh and Shiau Wei Chan,The Impact of Electronic Procurement Adoption on Green Procurement towards Sustainable Supply

Chain Performance-Evidence from Malaysian ISO Organizations",Journal of Open Innovation: Technology, Market, and Complexity,vol,no.61,2022. https://doi.org/10.3390/joitmc8020061

[11] Hongquan.Sun.et al., Application of green energy and environment-friendly decorative composite materials by using kinetic model with Sodium Alanate Nano carbon particle for building technologies, Journal of Sustainable Energy Technologies and Assessments, vol,no.60 ,2023 [Online].Available:.https://doi.org/10.1016/j.seta.2023.103502

[12] Mohsin.Shahzad et al.,Adoption of green innovation technology to accelerate sustainable development among manufacturing industry, Journal of Innovation & Knowledge,vol,no.7,2022. https://doi.org/10.1016/j.jik.2022.100231

[13] Pawel .Bratek , Lukasz Szustak, Roman Wyrzykowski and Tomasz Olas,Reducing energy consumption using heterogeneous voltage frequency scaling of data-parallel applications for multicore systems" Journal of Parallel and Distributed Computing,vol,no.175 ,2023 [Online].Available:. https://doi.org/10.1016/j.jpdc.2023.01.005

[14] Chaowen Zhong, et al., Energy Efficiency Solutions for Buildings: Automated Fault Diagnosis of Air Handling Units Using Generative Adversarial Networks",Journal of Energies Vol, no. 12, [Online].Available:. https://doi.org/10.3390/en12030527

[15] Zulfiqar Ahmad Khan et al., Towards efficient and effective renewable energy prediction via deep learning" journal of Energy Reports,Vol,no.8,2022, pp.10230-10243 [Online].Available:.https://doi.org/10.1016/j.egyr.2022.08.009

[16] Devinder Kaur, Shama Naz Islam, Md and Apel Mahmud,A Bayesian Deep Learning Technique for Multi-Step Ahead Solar Generation Forecasting",journal of Computer Science and Machine Learning,21 Mar 2022 [Online].Available:.https://doi.org/10.48550/arXiv.2203.11379.

[17] Yi Zhou, Nanrun Zhou, Lihua Gong and Minlin Jiang,Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine",journal of Energy ,Vol,no.204,2020, 117894,[Online].Available:. https://doi.org/10.1016/j.energy.2020.117894

[18] Hu Yue, Zizhuo Zhou and Hanwen Liu, How does green finance influence industrial green total factor productivity? Empirical research from China "journal of Energy Reports,vol,no.11,2024 [Online].Available:. https://doi.org/10.1016/j.egyr.2023.12.056