International Journal of Environmental
Sciences ISSN: 2229-7359
Vol. 11 No. 15s,2025
https://theaspd.com/index.php

A Review of the Big Data Application for Achieving Sustainable Development Goals (SDGs)

Rajib Shaw[1], Ozaki Yasutaka[1], Cao Yuqiu[1], Yinjia Zhang[1], Yanwu Zhang,[1] Sakiko Kanbara[2], Takahiro Ando[3]

1: Graduate School of Media and Governance, Keio University, JAPAN
2: Kobe City College of Nursing, JAPAN
3: Data Cradle, Kurashiki, JAPAN

E-mail: Rajib Shaw: shaw@sfc.keio.ac.jp

Paper type: Review article

Abstract: To achieve Sustainable Development Goals by 2030, big data has emerged as a critical tool in realizing these objectives. This paper provides an analytical review of how big data application contributes to the United Nations' Sustainable Development Goals. It synthesizes insights from 40 scholarly articles, categorized into six groups, to examine big data's role in various sectors such as smart cities, health, education, and e-commerce. The review highlights the transformative impact and challenges of big data in advancing the SDGs, offering a critical perspective on its potential and limitations. By exploring big data's utility in deciphering complex patterns and informing strategic decisions, this study underscores the importance of integrating advanced data analytics in sustainable development efforts. This work serves as a valuable resource for future research in big data's application towards achieving the ambitious objectives of the UN Agenda 2030.

Keywords: Sustainable Development Goal (SDGs); big data analysis; big data; application.

1. Introduction

In March 2017 the United Nations (UN) Statistical Commission adopted a measurement framework for the UN Agenda 2030 for Sustainable Development, this framework consisted of 232 indicators designed to measure the 17 Sustainable Development Goals (SDGs) and their respective 169 targets. (UNITED NATIONS, 2015) These universal goals cover all three key development pillars: economic, social, and environment, and include critical enablers such as institutional coherence, policy coherence, and accountability. (MacFeely 2019) Mogens Lykketoft, President of the 70th session of the UN General Assembly, aptly described the ambition of this challenge as an 'unprecedented statistical challenge' (Lebada, 2016).

Meanwhile we are in an era witnessing an era marked by an explosion of data generation. Big data has become an essential tool in deciphering complex patterns and informing strategic decision-making. Big data is a term that describes large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information.(Allen et al. 2021). As nations and organizations globally endeavor to meet the ambitious SDGs, the role and potential of big data are increasingly recognized as crucial.

Urban big data results from the increasing availability of the daily data we generate in the urban environment. These include, for example, data associated with urban sensors, e.g., Internet of Things (IoTs), administrative records, individual- or household-level survey data, geospatial imagery, commercial information, citizen science, and social media. Such data is critical in filling the gap for existing sustainability assessment tools and indicators. (United Nations Economic and Social Council (UNESC), 2015)

Despite the wealth of data being amassed, the interpretative frameworks to leverage this data in tackling complex urban sustainability challenges remain underdeveloped. The SDGs, in particular, face a pronounced data deficit that hampers effective monitoring and implementation. The challenges and opportunities of urban big data in public policy have also not been fully understood. The Sustainable Development Goals (SDGs) suffer from the problem a lack of data needed for effective monitoring and implementation. However, with 169 targets and 230 + indicators, the monitoring challenge is even more monumental. The official framework of indicators for monitoring the SDGs has undergone several revisions since its adoption in 2017. (United Nations General Assembly, 2017)

The applications of big data are diverse, impacting fields ranging from business intelligence to public policy, and significantly altering traditional problem-solving approaches. Key sectors contributing to big data include smart cities, healthcare, education, transportation, social media, business operations, promotions, event organizations, and e-commerce. These sectors generate substantial data, providing rich sources for big data processing and storage (Rahul, Banyal, and Arora, 2023).

This review aims to critically analyze the current applications of big data in advancing the Sustainable Development Goals. It highlights the transformative potential and the limitations of these technologies in driving sustainable development. For this analysis, we have systematically reviewed 40 papers that discuss big data applications across the aforementioned domains, categorizing them and examining their connection with the SDGs. This paper also delves into the challenges and opportunities presented by big data in this context. Extracted from various databases and analyzed using common keywords, this paper serves as a foundational study for future action research projects on the theme of big data application in SDGs.

2. Methods

To identify relevant literature on sustainable development goals and big data, this review utilized the Scopus database (accessed on October 17, 2023) following the recommendations from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. We ensured coverage of sustainable development goals and big data by using relevant search terms. A search string was applied to the article title, abstract, and keywords fields of Scopus, limited to “all open access”, “article,” and “published in English” with publication dates confined to the years 2018 to 2023: TITLE-ABS-KEY ((“sdgs” OR “sustainable development goals”) AND (“big data” OR “open data” OR “data analytics”)) AND PUBYEAR > 2017 AND PUBYEAR < 2024 AND (LIMIT-TO (OA, “all”)) AND (LIMIT-TO (DOCTYPE, “ar”)) AND (LIMIT-TO (LANGUAGE, “English”)). A total of 165 publications were identified. Since a UNGA resolution on July 6, 2017, identified specific targets for each SDG and provided indicators to measure progress, we selected publications published after 2018.

A multi-step screening process was employed to select the most relevant studies, based on the following criteria:

1. Relevance to SDGs: Direct addressing of one or more of the 17 SDGs.
2. Practical Application: Demonstrated practical applications of big data in achieving SDGs.
3. Innovative Use of Big Data: Novel approaches in applying big data to SDG challenges.
4. Geographical and Thematic Diversity: Representing a range of contexts and SDG themes.
5. Publication Date and Accessibility: Focus on papers published between 2018 and 2023 and prioritization of open-access papers.
6. Impact and Citation Metrics: Consideration of citation counts and journal impact factors.

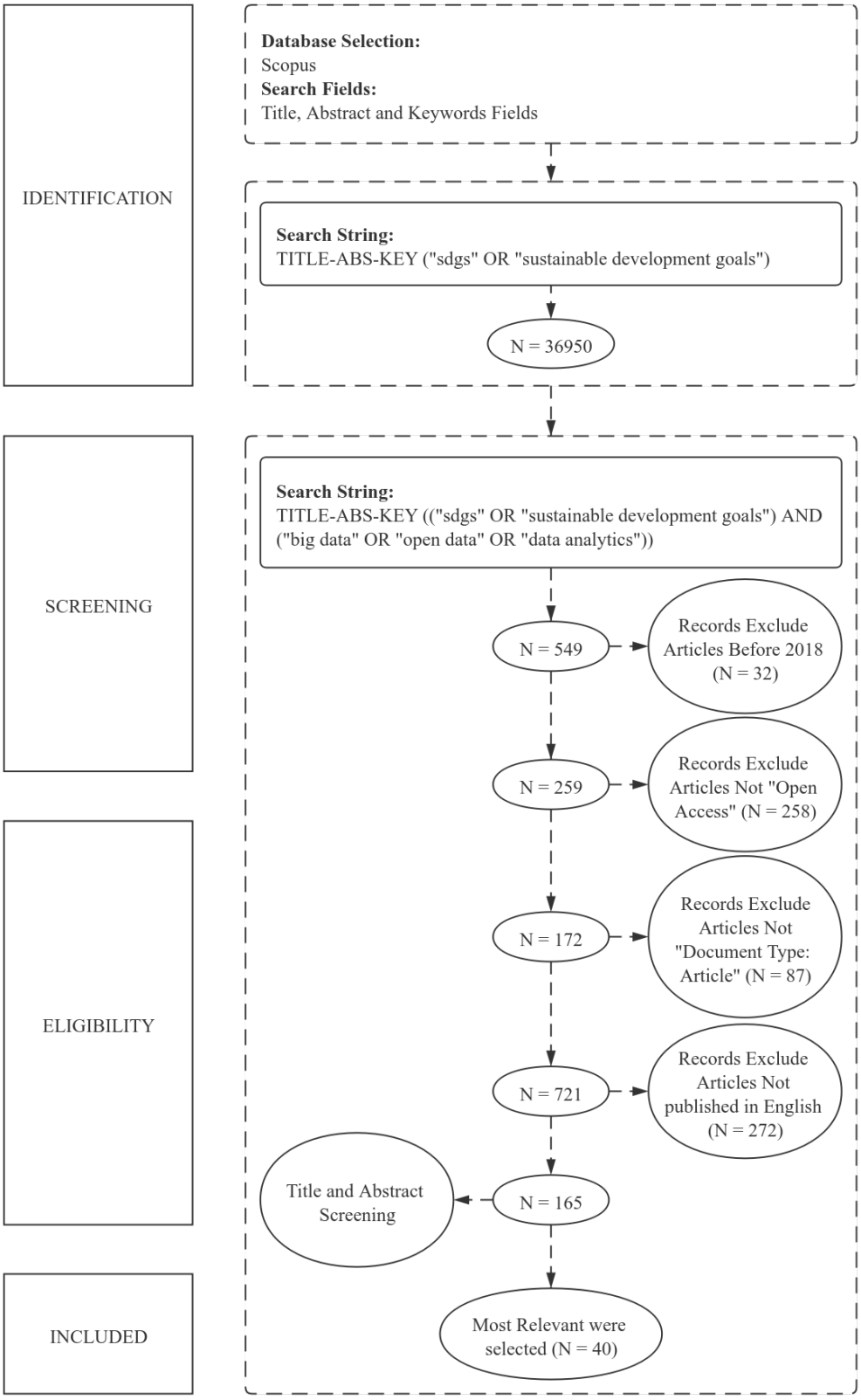

Figure.1: Steps of the Literature Search Process.

Table.1 List of the papers (Category, SDGs, Citations, Authors, Journals, Published year).

Category	SDGs	Ref. No.	Citations	Author	Journal	Year
Education	4	8	104	Buenaño-Fernández et al.	Sustainability	2019
	4	9	6	Matas-Terrón et al.	Sustainability	2020
	4	10	22	Jokhan et al.	Sustainability	2022
Health	3	11	47	Lawal et al.	Geo-Spatial Information Science	2019
	3, 1, 2	12	0	Liu et al.	Sustainability	2022
	3	13	1	Bello et al.	Texila International Journal of Public Health	2022
	3, 4	14	1	Goel et al.	Journal of Engineering Science & Technology Review	2023
	3, 5	15	0	Hammond et al.	Sexualities	2023
	3, 6	16	1	Lee et al.	Journal of Water, Sanitation and Hygiene for Development	2021
	3	17	3	Jakubik et al.	Production and Operations Management	2022
Media	3, 9, 16, 17	18	13	Li J et al.	Social Sciences & Humanities Open	2020
	9, 12, 16	19	7	Patuelli et al.	Scientific Reports	2023
Government Policy	8, 11, 13	20	1	Mihály, Szabolcs et al.	Climate	2022
	8, 9, 16	21	2	Liu, Yu et al.	Big Earth Data	2021
	all	22	7	Chen, I-Chun et al.	Communication Earth and Environment	2023
	8, 9, 11	23	2	Arfanuzzaman et al.	Sustainable Environment Research	2022
	8, 11, 13	24	10	Han, Liying et al.	Environment and Sustainable Indicators	2021
	11	25	4	Xie, Zunyi et al.	Remote Sensing	2022
	2, 8, 15, 17	26	64	Xie, Z. et al.	Remote Sensing of Environment	2019
	17	27	38	Wu, B. et al.	Geography and Sustainability	2020
	8, 9, 17	28	28	Guo, H. et al.	Big Earth Data	2021
	6	29	0	Poleshchuk, E. et al.	Statistical Journal of the IAOS	2022
	8, 9, 17	30	8	Owers, C.J. et al.	Big Earth Data	2021

	15, 17	31	61	Giuliani, G. et al.	Big Earth Data	2020
	1, 6, 9, 15, 17	32	19	Van Den Homberg, M. et al.	ISPRS International Journal of Geo-Information	2018
	4, 11, 15, 17	33	5	Assarkhaniki, Z. et al.	Big Earth Data	2021
	8, 9, 11, 17	34	4	Han, L. et al.	Remote Sensing	2022
	9, 15, 17	35	7	Borkowska, S. et al.	Sustainability (Switzerland)	2022
	11, 17	36	78	Ravanelli, R. et al.	Remote Sensing	2018
	8, 11, 17	37	6	Kulmala, M. et al.	Big Earth Data	2021
	11, 17	38	37	Cuca, B. et al.	Sustainability (Switzerland)	2021
Enterprise/ Retail/ E-commerce	9, 17	39	3	Syed Nazim Ali et al.	Journal of King Abdulaziz University, Islamic Economics	2020
	9, 17	40	0	Gholami et al.	Sustainability (Switzerland)	2023
	8, 9, 17	41	56	Gössling et al.	Journal of Sustainable Tourism	2020
	9, 17	42	102	Kumar et al.	Annals of Operations Research	2022
Smart City	11, 16, 17	43	2	Sucupira Furtado et al.	Journal of Urban Management	2023
	9, 11	44	22	Allam et al.	Smart City	2019
	9, 11	45	1	Abdalla et al.	Sustainability (Switzerland)	2022

Table.2 Application used and Evaluation.

Ref. No.	Application used	Evaluation
[8]	The utilization of educational data mining techniques to forecast a student's academic performance based on their historical grades.	Opportunity: Conduct predictive analysis on the final grades of computer engineering students to improve academic quality and reduce dropout rates. Challenge: Future work involves creating a big data architecture to handle the university's extensive academic data. This database should be well-documented, scalable, and flexible, enabling efficient indexing and access by students, teachers, and administrators.
[9]	The Big Data Applied to Education Scale (VABIDAE) is a 31-item questionnaire designed to collect opinions and perceptions about the use of Big Data technologies in education and classrooms.	VABIDAE offers a way to assess future educators' views on Big Data in education. It's valuable for higher education institutions to understand the perspectives of students and faculty on this technology, aiding in decision-making for its implementation. Additionally, VABIDAE could inspire a broader tool to evaluate educators' attitudes towards new

		technologies in a rapidly evolving socioeducational environment.
[10]	An AI-based analytics tool created to predict student performance.	The method effectively predicts student performance early in courses, enabling timely intervention. During COVID-19, the model showed high accuracy, precision, and recall in forecasting teaching behaviors and student e-learning in virtual education systems.
[11]	The study uses open data and geospatial analysis to evaluate the accessibility of Primary Health Care Facilities in Nigeria, incorporating data on location, elevation, population, and networks.	Opportunity: This approach can inform policy decisions by providing insights into healthcare accessibility, facilitating efficient planning and resource allocation. Regular accessibility assessments can identify underserved areas, guiding the placement of new facilities. Challenge: Which is integrating additional data like poverty levels, population structure, and facility capacity to create a more comprehensive measure of healthcare access.
[12]	The study in India leverages satellite data on a cloud platform to analyze village-level environmental factors related to child malnutrition indicators (stunting, underweight, wasting) for Sustainable Development Goals research.	Analyzing spatial variation patterns of these environmental variables at the village level and their correlation with child malnutrition indicators aids in understanding malnutrition causes and enhances SDG 3 implementation planning.
[13]	The study investigated the geographic and spatial accessibility of the primary health care network in the Bauchi Local Government area of Bauchi State – Nigeria, utilizing open data and geospatial analysis methods.	It identifies factors affecting healthcare accessibility, aiming to guide equitable distribution of facilities and reduce accessibility inequality, supporting efficient healthcare delivery and Universal Health Coverage goals in similar resource settings.
[14]	Big data in healthcare enhances patient care and service delivery through electronic health records management, remote monitoring, disease prediction, personalized medicine, and integration with AI and wearables for real-time monitoring, aiding telemedicine especially in remote areas.	Opportunity: The study highlights the importance of aligning AI-enhanced health products and services with the SDGs, emphasizing ethical AI for human rights protection and advocating for collaboration across various fields for technology development. Challenge: It notes the potential of medical AI to bridge health service gaps between urban and rural areas in developing countries but recognizes obstacles like poor data collection and systemic inefficiencies impacting research and advancement.
[15]	Using Big Data to address sexual and reproductive health concerns	The challenge is handling the sensitive and stigmatized nature of sexual and reproductive health data in Big Data contexts, particularly in Low and Middle-Income Countries.
[16]	A pilot Rainwater for Drinking (RFD) system was set up at a local hospital in Vietnam, where water quality and quantity were monitored regularly by trained local operators. The collected data was uploaded to an open data access platform for analysis and technical support, ensuring the provision of safe drinking water.	It enhanced the resilience of the RFD system and contribute to resolving drinking water challenges in rural areas of developing countries, aligning with Sustainable Development Goal 6.

[17]	A data-driven decision model that determines an optimal allocation of development aid across countries so that the expected rate of new HIV infections is minimized.	The decision model meets the need for improved tools in development aid allocation, aiding managers in funding bodies to finance development activities more effectively and cost-efficiently, particularly towards ending the HIV epidemic.
[18]	An unsupervised machine learning methodology using natural language processing to collect and analyze data from the popular social media platform Twitter with the aim of detecting self-reported experiences with corruption, including in the health sector.	The study showcases the potential of technology as a tool to detect and analyze corruption across sectors, aiding in future anti-corruption initiatives. It also highlights the role of the Sustainable Development Goals in promoting technology-driven anti-corruption strategies through governance and multistakeholder partnerships.
[19]	An interdisciplinary approach based on analyzing big data from an online social network (Twitter) with complex network methods from statistical physics.	This study offers a detailed, near real-time overview of UK firms' involvement with the SDGs, revealing key findings such as the prominence of social themes, varying attention to SDGs by sector, and heightened stakeholder engagement in global challenge discussions.
[20]	The study utilizes accessible and reliable big Earth observation (EO) data or geodata to launch 'DREAMS' project: aims to forge sustainable pathways for African cities using participatory scenario modeling, impact assessment, and integrated strategic planning to address essential urban SDGs.	Geospatial data comes from diverse sources and modern infrastructures offer prompt data access, aiding in analysis and promoting evidence-based decision-making in SSA. The challenges are Earth Observation (EO) data limitations, including low resolution that fails to capture rapid environmental changes and insufficient temporal resolution for monitoring sudden events. Along with the lack of accessible accuracy documentation and standardized metadata.
[21]	"Development of the Earth Observation Information System data infrastructure and services" project focuses on collecting, processing, storing, and archiving Earth observation data through resources provided by the ESA(European Space Agency).	These services are designed to support public administration tasks, lessen administrative work for businesses, supply data to technical systems, and offer enhanced data services to profit-driven entrepreneurs. However, it required a flexible and scalable hardware and software environment for potential expansions, adherence to industry standards with a focus on open-source tools, and the development of specialized online interfaces and data channels for various stakeholders.
[22]	This study developed an Artificial Neural Network (ANN) model using multi-source, low-cost big data to assess the Sustainable Development Goals (SDG) performance of cities of various sizes in China by quantifying the performance of 17 individual SDGs for 254 cities with statistical data and assesses the SDG Index for all 337 different-sized cities with open-source big data.	The assessment results provide a scientific reference for Chinese cities to achieve sustainability which can help governments at all levels monitor and evaluate SDG progress, especially cities that have limited statistical data availability and weak big data processing capabilities. However, due to data limitations, not all 17 SDGs were measured using appropriate big data monitoring indicators and its calculation methods lack the quantification of interactions among the SDGs.
[23]	This study uses open datasets of regional socioeconomic and environmental geo-information to develop a regional SDG	The BRQ model offers land stakeholders (government and investors) a flexible decision-making tool that integrates open, real-time social, economic, and environmental datasets at the

	map and screen suitable brownfield locations for redevelopment, also it has designed a Brownfield Redevelopment Query (BRQ) model with sustainable potentiality for evaluating benefit-cost ratios (BCRs).	township level which reduces investigation costs and information uncertainty, aiding in evaluating brownfield redevelopment scenarios. Additionally, governments can use this system as a platform for public communication and to assess the benefits of different brownfield plans.
[24]	Using Earth Observation (EO) data to assess and address issues such as urban heat intensities, flash floods, and other sudden environmental events.	It helps enhance resilience and promote evidence-based decision-making in South Asia, contributing to the achievement of specific SDGs by providing timely and accurate information for systemic responses to environmental challenges. However, challenges like data security, data integrity, data protection and privacy, dataveillance, data quality and ethics, social and digital wellbeing still remained.
[25]	This study developed an assessment framework for SDG indicators 11.2.1, 11.3.1, and 11.7.1 at the neighborhood level using high-resolution (HR) satellite images, gridded population data, and other geospatial big data (e.g., road network and point of interest data).	The evaluation results confirmed the applicability of SDG11 indicators to neighborhood-level assessment and local urban governance and planning practices and the evaluation framework of the SDG11 indicators based on HR satellite images and geospatial big data showed great promise to apply to other cities for targeted planning and assessment. However, geospatial data with enhanced spatial and temporal resolution should be produced and applied to improve the accuracy and reliability of SDG evaluation results in future studies.
[26]	The study demonstrated a novel approach for mapping vegetation cover changes using big data and remote sensing technologies. The main application used in this study was Google Earth Engine (GEE). The method was able to categorize vegetation cover changes into five classes: significant decrease, potential decrease, stable, potential increase, and significant increase.	The application faces challenges in analyzing large datasets across extensive areas and distinguishing human-induced changes from natural variations. Applying these insights to identify lands needing conservation across diverse geographical and climatic conditions is also complex. However, there are opportunities as the developed methods can be tailored for global application, significantly aiding land conservation and achieving SDGs. The use of cutting-edge remote sensing technologies, such as Google Earth Engine (GEE), offers a valuable resource for environmental monitoring and informed decision-making in conservation efforts.
[27]	Applications of the methodologies and data sources mentioned in the paper span various fields such as agricultural monitoring, water resource management, urban development, climate resilience, disaster risk reduction, and marine resources. Specific platforms like CropWatch Cloud and Copernicus services are highlighted for their contributions to global crop condition monitoring and providing essential information across multiple domains.	The application grapples with challenges such as high costs and complexities in obtaining quality data, redundancy issues, slow traditional data collection methods, and the need for validating public Earth Observation data. However, it also uncovers opportunities such as cost savings from new technologies, accessible cloud data, smartphone-based data collection advancements, and high-resolution satellite data products. Further, cloud computing and machine learning offer promising prospects for producing data products and facilitating more inclusive data disaggregation.
[28]	The applications cited in the paper are diverse, spanning across several SDGs.	The challenges identified in the application include integrating diverse data, managing large data volumes, and

	Examples include using Big Earth Data (BED) for monitoring crop production, urban expansion, water quality, and environmental impacts on food production.	ensuring interoperability among various data systems. Conversely, the opportunities are significant. Big Earth Data (BED) holds potential for advancing sustainable development policies, resource management, disaster preparedness, and assessing climate change impacts. To capitalize on these opportunities, the paper emphasizes the need for better communication methods, efficient data management technologies, and advanced data visualization techniques.
[29]	The paper focused on using Big Data, specifically geospatial and Earth Observation (EO) data, combined with advanced technologies to create new methods for producing environment statistics. This approach is exemplified through the development of SDG indicator 6.6.1, which monitors the change in the extent of water-related ecosystems over time.	Integrating Earth Observation (EO) data into official statistics presents challenges such as the requirement for pre-processing satellite imagery, a skills gap in national statistical offices for such pre-processing, and the need for technical, human, and financial resources for data storage, processing, and analysis. The article also points out limitations due to certain countries' Statistics Laws on using Big Data for official statistics. Despite these hurdles, there are opportunities for collaboration among national statistical offices, mapping agencies, space agencies, and specialized environment agencies to address these challenges.
[30]	The paper discussed a global land cover classification framework to support SDGs using Living Earth, an open-source Python library. This high-performance tool utilizes libraries like xarray, NumPy, GDAL, Rasterio, and Open Data Cube for data handling. Living Earth effectively applies classification rules for producing detailed level 3 and 4 outputs. Its flexible design accommodates diverse data inputs and is compatible with national super-computing resources and cloud services, including Amazon Web Services (AWS).	The paper addressed challenges in using Earth Observation data for land cover classification, like difficulties in deriving specific environmental descriptors, necessitating adjustments to the FAO LCCS-2 framework. However, it also identifies opportunities with Living Earth's flexible software, enabling detailed classifications with limited data. This software adapts to various environmental descriptors, suitable for different landscapes globally. It also integrates recent technologies, supporting consistent land cover analysis and helping nations achieve SDGs.
[31]	The study utilized various applications and tools, including the Trends.Earth model, the Virtual Laboratory (VLab), and Python scripts for processing and analysis. Trends.Earth, a system compatible with both desktop and cloud, generates sub-indicators and calculates the SDG15.3.1 indicator, integrating with QGIS and Google Earth Engine. Additionally, the Swiss Data Cube, based on the Open Data Cube software, aids in analysis, with support from a Python API for creating tailored tools and applications.	The main challenges include limited data availability, especially the need for longer time-series data for dependable analysis, and addressing technical issues related to data gaps and integrating various satellite data sources. On the other hand, there are significant opportunities for enhancing the methodology, specifically in generating sub-indicators more effectively and broadening its application scope from national to regional and global levels.
[32]	The paper focused on SDG 6.1.1 in rural Malawi, highlighting the involvement of various government and NGO actors in water supply projects, each with distinct	The paper highlights challenges like an underdeveloped data ecosystem, fragmented data management, complex funding, and overall data deficiency. However, it sees opportunities in enhancing data ecosystems to utilize non-official datasets,

	funding sources and little overarching governance. It identified several geospatial data sharing platforms and online information systems. However, these systems face challenges like limited internet connectivity and low data literacy among users, resulting in low adoption rates.	potentially improving SDG reporting. The proposed framework aims to identify geographical data gaps, dimensional deficiencies, and areas lacking collaboration. Addressing these can evolve the data ecosystem, crucial for informed decisions and achieving SDGs.
[33]	The study focused on using open data for detecting informal settlements, especially beneficial when costlier high-resolution data sources are unavailable, the target area is city-sized or smaller, and the informal settlements have distinct physical features compared to nearby formal areas. This approach demonstrates how globally accessible data can be applied to enhance resilience and aid in achieving SDGs in informal settlements.	The study notes a challenge in using only OpenStreetMap (OSM) structural features for training data, risking inaccurate classification due to ignored factors like housing quality and rooftop materials in satellite images. It suggests expert visual interpretation and comprehensive sampling of all physical and structural features to mitigate this. The study also sees potential in using OSM for better sampling to distinguish between formal and informal settlements.
[34]	The study employed a combination of geospatial big data techniques and high-resolution satellite imagery analysis. These applications were pivotal in classifying land cover, analyzing urban functional boundaries, and calculating various SDG indicators. The study also used network analysis to assess public transport accessibility and object-based image analysis to determine open public spaces.	The study encountered data quality issues, especially with historical road network information, and grappled with the complexities of urban growth requiring detailed, localized data for effective SDG monitoring. However, it also showcases the potential of geospatial big data and high-resolution satellite imagery for extensive urban monitoring and planning, illustrating how technology can facilitate reaching SDG targets in urban settings.
[35]	The applications used in the article included GIS software—ArcGIS Pro and Statistica. These tools were employed to perform spatial analyses and assess the quality of OSM data. The choice of these applications was integral to the methodology, allowing for a detailed and accurate analysis of the data quality in OSM compared to the BDOT10k database.	A significant challenge identified in the study is the inconsistent quality of OpenStreetMap (OSM) data, notably in less urbanized and economically disadvantaged areas. However, there is an opportunity to use OSM data to supplement other spatial data sources for monitoring sustainable development, particularly in well-developed and urban areas where OSM data quality is generally higher. The findings suggest potential uses for crowdsourced geographic information in aiding the monitoring and achievement of SDGs, especially in urban planning and infrastructure development.
[36]	The study used Google Earth Engine and the Climate Engine to analyze the link between Land Surface Temperature and Land Use/Land Cover changes in six U.S. cities from 1992 to 2011. It involves processing over 6000 Landsat images and employing the National Land Cover Database. A key aspect is developing a Detrended Rate	The main challenge is calibrating Land Surface Temperature from Landsat images. Another significant issue is handling and processing the large volume of satellite data. However, it also presents opportunities for enhancing urban sustainability through informed planning. It demonstrates the potential of using satellite data and platforms like Google Earth Engine for environmental monitoring, offering a scalable approach that can be adapted for various urban environmental aspects.

	Matrix to show how land cover changes affect Urban Heat Islands.	
[37]	This research has varied applications including environmental monitoring, policy-making, business guidance, and education. A key example is the VÄRI project in Eastern Lapland, leveraging SMEAR I station data to support local economies and raise environmental awareness. This project made research results accessible to local businesses and residents for product development and marketing. Additionally, SMEAR data inform urban and regional environmental strategies, aiding in air pollution monitoring and understanding its sources and impacts in large cities like Beijing and Nanjing.	The challenge lies in integrating diverse environmental data and insights across scales and disciplines. However, this complexity offers opportunities for multidisciplinary collaboration and innovation. SMEAR data inform climate change efforts, while the PEEX program fosters international collaboration in environmental research, benefiting regions like Northern Eurasia and China and informing practical solutions and policies.
[38]	The study primarily uses optical multi-spectral satellite datasets, specifically Landsat and Sentinel-2 products. The Landsat science program offers the longest continuous global record of Earth's surface, providing multi-spectral images at a medium resolution. The Sentinel-2 is part of the Copernicus program, offering recent and high-resolution data. These datasets were pivotal in detecting, identifying, and characterizing significant morphological transformations within and around Nicosia's walled city.	The primary challenge lies in accurately tracking urban changes in a complex area like Nicosia, while integrating EO data into existing urban planning poses another challenge. However, the research opens significant opportunities in sustainable urban planning. It demonstrates how EO and geospatial information can play a crucial role in understanding and managing urban changes, particularly in historic city centers. This approach offers a new perspective on urban development, with a focus on sustainability and cultural heritage preservation.
[39]	The paper discusses various applications of big data in enhancing financial inclusion, particularly in low-income and developing countries.	By leveraging data-driven insights, Islamic finance can play a pivotal role in achieving a more equitable and sustainable future and it advances several SDGs related to poverty reduction, economic growth, gender equality, and sustainable practices. However, The difficulty in regulating large digital platforms due to their size and lobbying power, and the need for better utilization of big data in the Islamic finance industry to develop personalized products and services remains a challenge
[40]	In this study, big data is used to analyze and improve various aspects of product sustainability which includes enhancing product end-of-life management efficiency, product quality and durability, and functional performance. Also, it contributes to advancing SDGs related to clean energy, economic growth, innovation, responsible consumption, and partnership.	Stakeholders can gain insights into product sustainability indicators across economic, environmental, and social dimensions, leading to more informed decision-making and strategies that align with these SDGs. However, challenges such as understanding the interdependencies of identified indicators, developing effective measurement frameworks, and optimizing their contributions still remain.

[41]	Big data and AI are increasingly used for COVID-19 management, security , and service automation in the tourism industry. AI also powers digital assistants, chatbots, and voice-controlled systems, enhancing personalization and efficiency.	Opportunities for personalization, market insights, virtual experiences, education, and strategic planning is mentioned. However, challenges include insufficient sustainable and ethical practices, complex AI technologies creating regulatory gaps, labor displacement from automation, and concerns about societal and economic impacts. also, the need for better governance and critical academic assessment is mentioned.
[42]	This paper uses big data and machine learning techniques to analyze a large volume of scholarly research.	By leveraging big data analytics, the study provides insights into the evolution of topics and identifies major themes that form the intellectual structure of sustainable finance research. This method offers a novel way to comprehensively understand and map the landscape of sustainable finance research, highlighting its development and key contributions.
[43]	This study presents how digital tools and policies can touch upon several sustainability goals when focused on benefitting vulnerable citizens, then build a framework that further links those examples with the SDGs and the Brazilian Letter for Smart Cities, which is a national set of guidelines towards equitable smart transitions.	The Letter is directly connected to SDG 16 and SDG 17, both centered around governance. The challenge related to effectively integrating Digital Transformation projects with the aim of targeting SDGs and addressing governance challenges. The challenge is to innovatively use this collaboration to solve real governance issues, particularly in the context of the global demand for IT workers and the rapid growth of the industry.
[44]	This paper uses the concept of Big Data beyond infrastructure to include that of urban health through human anatomy.	Coupling both data forms will be key in supplementing the contemporary notion of Big Data for the pursuit of more contextualized, resilient, and sustainable Smart Cities, rendering more liveable fabrics. One notable challenge is the need to define strict protocols, as health data may not be available or accessible due to its complex, confidential, or sensitive nature.
[45]	This study explores the impact of Big Data and smart city initiatives on supply chain management by creating a comprehensive model for supply chains, Big Data, and smart cities.The effects of smart cities and SCM integration on sustainable development are also examined.	Big Data, smart cities, and supply chains have more than merely causal interactions, and Big Data and smart cities will hugely impact sustainable development and SCM operations. Challenges associated with supplier network design and governance processes, where organizations must comprehend big data challenges and recognize how the overall environment may alter when new actors join.

3. Review Results

3.1. Education

Of the 40 papers reviewed, 3 published in the Sustainability journal, focused on education. They all concentrate on the integration and impact of data-driven technologies in the educational field, emphasizing that quality education is closely tied to sustainable development goals.

Papers [8] and [10] explore the use of data analytics and artificial intelligence (AI) in predicting student performance. They both highlight the importance of predictive analytics in improving student outcomes and facilitating timely interventions, crucial for maintaining academic quality. Paper [10] additionally notes the adaptability of these methods to virtual learning environments. These papers primarily address student-focused applications of data technology.

In contrast, Paper [9] targets educational institutions and educators. It introduces the Big Data Applied to Education Scale (VABIDAE), a questionnaire designed to gauge opinions on Big Data in education. VABIDAE is essential for guiding decisions on technology implementation in higher education and is instrumental in understanding educators'

attitudes toward emerging technologies, thus contributing to the broader goal of integrating sustainable practices in educational settings.

3.2. Health

7 papers under review present a diverse yet interconnected exploration of health-related issues, each contributing to the Sustainable Development Goals, with a particular emphasis on SDG 3, aimed at ensuring healthy lives and promoting well-being for all ages.

A common theme across these studies is the emphasis on data-driven approaches, highlighting the critical role of data in understanding and dealing with health challenges. From open data to big data, each study leverages different types of data to inform health-related decisions and policies.

Geospatial analysis emerges as a key tool in several papers, especially in understanding the accessibility of healthcare facilities and its impact on public health. Papers [13], for instance, delve into the geographic accessibility of health care facilities in Nigeria, using open data and geospatial techniques to identify areas that are underserved. This focus is crucial for achieving equitable health care distribution.

In contrast, Paper [12] pivots to examine the environmental determinants of child malnutrition in India, employing satellite big data to understand the interplay between environment and health at a granular village level. This study underscores the environmental dimensions of public health challenges. Digital Health Interventions are central to Paper [14], which diverges from traditional health research by highlighting the role of AI, wearable devices, and telemedicine in improving patient care. This reflects a growing trend towards integrating technology into healthcare solutions. Paper [15] addresses a unique challenge within the realm of big data - the handling of sensitive sexual and reproductive health data, especially pertinent in low and middle-income countries. This paper brings to the fore the complexities and sensitivities involved in managing health data. The case study in Vietnam presented in Paper [16] is distinct in its localized focus. It explores the implementation and monitoring of a Rainwater for Drinking system at a rural hospital, demonstrating how systematic data collection and analysis can ensure the provision of safe drinking water. Finally, Paper [17] presents an innovative approach to public health funding. It introduces a data-driven decision model for allocating development aid in combating HIV/AIDS, showcasing how data can be used for effective and efficient resource allocation in global health initiatives. Overall, these papers collectively emphasize the increasing reliance on data and technology in the health sector. They illustrate varied applications of data-driven approaches, from local-level interventions to global health challenges, each contributing uniquely to the overarching goal of improving health outcomes and advancing Sustainable Development Goals.

3.3. Media

The two papers under analysis provide intriguing insights into the use of social media, specifically Twitter, as a rich data source for understanding complex societal issues in the context of Sustainable Development Goals. The first paper employs unsupervised machine learning to detect self-reported experiences of corruption. The second paper takes a different route, analyzing how UK firms engage with SDGs on Twitter. Together, 2 papers underscore the versatility of social media data in researching societal issues. One focuses on uncovering grassroots-level problems, while the other provides a macro view of corporate involvement in SDGs, demonstrating the diverse applications of Twitter data in understanding and addressing global challenges.

3.4. Government Policy

Of the 40 papers reviewed, 16 were in the government policy category. Most of them were published in climate and environment-related journals, which demonstrates the trend of today's policy-makers to emphasize the environment and the coherence of climate with urban planning and human well-being. It is important to note that 14 of these 16 papers emphasized more than 1 SDGs, and each of them belonged to the Environment, Society and Economy category, highlighting the growing importance of data-driven approaches in sustainable development policy.

In each of the 16 papers, the research aims to provide tools or frameworks that can assist in decision-making and strategic planning. Whether it is assessing the performance of the SDGs, evaluating brownfield redevelopment, or analyzing urban resilience, the goal is to inform and enhance decision-making. A common theme across these papers is the utilization of big data to address specific SDG targets, such as urban heat island effects, vegetation cover change, water ecosystem monitoring, and land degradation. In addition, it is important to note that in the data sources section, most of the studies have utilized Earth Observation (EO) and geospatial open data, which are essential for monitoring and assessing various aspects of urban development and environmental change. As for the

challenges section, a number of studies identified challenges related to the resolution, accuracy and integration of big data, which affect the reliability and validity of data-driven analysis and decision-making.

In contrast, these 16 papers all have specific goals and applications. Paper [20] and [24] shows the DREAMS project and the research conducted in South Asia focusing on climate resilience and urban informal settlements. Papers [21] and [25], the Hungarian project and the research conducted in Guilin, China, emphasize stakeholder management and participation and neighborhood-level SDG assessment, respectively. Paper [22] and [24], the ANN model and the BRQ model in Taiwan, focuses on evaluating urban SD, SDG performance and brownfield redevelopment programs. Papers [26] and [36] demonstrate the use of Google Earth Engine (GEE) and Climate Engine (CE) tools, respectively, for detecting vegetation cover changes in rangelands and monitoring the impact of land cover change on surface urban heat islands. Papers [27], [28], and [31] discuss the use of crowdsourced geographic information for acquiring geographically located data, which, when combined with public Earth Observation data, is critical for several SDG indicators such as forest cover, crop area, and land degradation. Paper [29] focuses on using geospatial and Earth Observation data to develop new methods for producing environment statistics, exemplified by the development of SDG indicator 6.6.1, which monitors changes in the extent of water-related ecosystems. Paper [30] demonstrates how the FAO Land Cover Classification System (LCCS-2) within the Living Earth software package supports the SDGs by providing a platform for standardized, reliable mapping of land cover changes. Paper [32] presents a study conducted in rural Malawi, focusing on characterizing data ecosystems to support reporting on SDG indicators and highlighting the complexities and potential for improvement in data ecosystems for effective decision-making. Paper [33] examines the use of open data to detect the structure and pattern of informal settlements in Jakarta, Indonesia, contributing to the achievement of SDGs by enabling better planning and upgrading of these areas. Paper [34] shows how a combination of high-resolution satellite imagery and geospatial big data was used to assess SDG11 indicators in China, emphasizing improvements in public transportation accessibility. Paper [35] reveals that the quality of OpenStreetMap spatial data is highest in urbanized areas, indicating its potential to support other spatial data for sustainable development monitoring. Paper [37] discusses the SMEAR network's comprehensive long-term in-situ observations, integrated with remote sensing and modeling. Lastly, Paper [38] highlights the contribution of earth observation and geospatial information to the urban planning of Nicosia, Cyprus. In addition, some studies are region-specific (e.g., Sub-Saharan Africa, Hungary, Poland, Cyprus, China, Indonesia, South Asia), while others have broader applications or focus on specific local challenges. It is also worth noting that stakeholder engagement varied across the 16 papers, with some studies (e.g. the Hungarian project) emphasizing the development of interfaces and channels for different stakeholders, while others focused more on the technical aspects of data analysis.

Overall, these papers collectively illustrate the different ways in which big data can be used in government policy, particularly in urban and environmental planning. The studies highlight the potential of big data to inform and improve sustainable development efforts, while also recognizing the challenges that need to be addressed to make full use of this potential.

3.5. Enterprise/Retail/E-commerce

Of the 40 papers reviewed, four were in the corporate/retail/e-commerce category. Most of these were focusing on SDGs 8, 9 and 17, and it should be pointed out that SDG17 call for a reinvigorated global partnership, demonstrating that economic development has become a global issue with globalization, whether through financial inclusion, product sustainability, tourism or sustainable finance, and each of these studies aligned their objectives with the broader SDGs.

Each 4 studies utilized big data to gain insights into their respective fields. This includes analyzing product sustainability, financial research trends, tourism dynamics, and the role of Islamic finance in economic development. Meanwhile, Each study acknowledges the challenges in applying big data, such as regulatory issues, understanding complex artificial intelligence techniques, and developing effective measurement frameworks. These challenges suggest that the field of big data analytics needs further development.

In comparison, each of the four papers focuses on a specific application area: paper [39] focuses on inclusive finance in low-income and developing countries; paper [40] deals with product lifecycle, quality, and performance; paper [41] explores the application of AI and big data in COVID-19 management and service automation; and paper [42] analyzes academic research that paints a picture of the future of sustainable finance. In addition, each paper highlights unique challenges related to its fields, such as ethical practices in tourism, product sustainability in manufacturing, financial regulation in Islamic finance, and research analysis in sustainable finance. Overall, the papers collectively illustrate the multiple applications of big data in the corporate/retail/e-commerce space, while also emphasizing the need to address domain-specific challenges in order to make the most of big data.

3.6. Smart City

Of the 40 papers reviewed, three fell under the category of smart cities. All three papers emphasize the use of big data to support the development of smart cities and contribute to the achievement of the SDG11. At the same time, they recognize the role of big data in enhancing the quality of life and governance in urban settings. This demonstrates a holistic approach to smart city development, whereby big data is not isolated but part of a larger urban ecosystem. At the same time, all papers recognize challenges in applying big data, such as governance issues, data accessibility and sensitivity (especially health data), and the complexity of integrating new systems or players into existing frameworks.

To make comparison, each paper has its own specific area of focus: paper[43] focuses on the digital transformation of governance, aligning with SDG 16 (Peace, Justice and Strong Institutions) and SDG 17 (Partnerships for Purpose); paper [44] examines big data from a health perspective, aiming to improve urban livability; and paper [45] explores the impact of big data and smart city initiatives on supply chain management. In addition, each paper emphasized unique challenges related to its field, such as governance integration in digital transformation and confidentiality in holistic health. Overall, They underscore the need to address specific challenges in each area to fully leverage the potential of big data in creating more efficient, healthy, and sustainable urban environments.

4. Conclusion

The comprehensive review of 40 papers on the application of big data for achieving Sustainable Development Goals (SDGs) across various sectors– education, health, media, government policy, enterprise/retail/e-commerce, and smart cities – presents a multifaceted view of the current landscape. Since this is a paper targeting the SDGs and big data application, the research sectors are wide-ranging, so we categorize each sector and summarize the common issues regarding “big data application” and “SDGs” in each sector and these are extracted below.

1. In the realm of education, the integration of data-driven technologies is pivotal for enhancing academic outcomes and adapting to evolving learning environments.
2. The health sector’s reliance on big data, from geospatial analysis to digital health interventions, illustrates its critical role in understanding and addressing public health challenges.
3. The use of social media data in media studies reveals the power of big data in uncovering societal issues and corporate engagement with SDGs.This sector faces the challenge of analyzing vast amounts of unstructured data and deriving meaningful insights that can inform policy and practice.
4. The government policy sector, with the highest number of papers reviewed, shows a strong inclination towards environmental and urban planning issues. The application of Earth Observation (EO) and geospatial data is prominent, aiding in decision-making and strategic planning. Challenges in this sector include the resolution, accuracy, and integration of big data, which are crucial for reliable policy-making.
5. In the enterprise, retail, and e-commerce sector, big data applications are aligned with economic development and global partnerships (SDGs 8, 9, and 17). Challenges such as regulatory issues, understanding AI, and developing effective measurement frameworks are prevalent, suggesting the need for further development in big data analytics.
6. Lastly, in the smart city category, big data is seen as a key enabler for urban development, aligning with SDG 11. Challenges here include governance integration, data accessibility, and the complexity of integrating new systems.

Across all sectors, a common theme is the need for a nuanced understanding of big data's capabilities and limitations. The challenges identified – ranging from data resolution and accuracy to ethical considerations and regulatory complexities – suggest that while big data offers significant opportunities for advancing SDGs, its effective utilization requires careful consideration of these constraints.

This review not only sheds light on the current state of big data applications in various domains but also sets the stage for future research. It underscores the necessity for continued exploration into innovative big data solutions, while also emphasizing the importance of addressing the challenges to fully harness its potential for sustainable development.

In conclusion, big data stands as a powerful tool in the quest to achieve the SDGs, offering insights and solutions across multiple domains. However, its full potential can only be realized through a balanced approach that considers both its transformative capabilities and the challenges it presents. This review provides a foundation for understanding the current landscape of big data applications in SDGs and paves the way for future research and development in this critical area.

Acknowledgements
This research is part of the CBAS Keio University partnership titled: “Effectiveness of “Big Data” analysis for small and medium sized cities in Japan to achieve their SDGs.

Reference

[1] IAEG-SDGs. (2022). Tier Classification for Global SDG Indicators.

[2] UNITED NATIONS. (2015). TRANSFORMING OUR WORLD: THE 2030 AGENDA FOR SUSTAINABLE DEVELOPMENT. UNITED NATIONS.

[3] United Nations General Assembly. (2017). Work of the statistical commission pertaining to the 2030 agenda for sustainable development. United Nations General Assembly. New York.

[4] Lebada, A. M. (2016, march 8). Member States, Statisticians Address SDG Monitoring Requirements. Retrieved from SDG knowlegde Hub: https://sdg.iisd.org/news/member-states-statisticians-address-sdg-monitoring-requirements/

[5] Allen, Cameron, Maggie Smith, Maryam Rabiee, and Hayden Dahmm. 2021. “A Review of Scientific Advancements in Datasets Derived from Big Data for Monitoring the Sustainable Development Goals.” Sustainability Science 16(5): 1701–16.

[6] MacFeely, Steve. 2019. “The Big (Data) Bang: Opportunities and Challenges for Compiling SDG Indicators.” Global Policy 10(S1): 121–33.

[7] Rahul, Kumar, Rohitash Kumar Banyal, and Neeraj Arora. 2023. “A Systematic Review on Big Data Applications and Scope for Industrial Processing and Healthcare Sectors.” Journal of Big Data 10(1): 133.

[8] Buenaño-Fernández, Diego, David Gil, and Sergio Luján-Mora. "Application of machine learning in predicting performance for computer engineering students: A case study." Sustainability 11.10 (2019): 2833.

[9] Matas-Terrón, Antonio, et al. "Validity of the “Big data tendency in education” scale as a tool helping to reach inclusive social development." *Sustainability* 12.13 (2020): 5470.

[10] Jokhan, Anjeela, et al. "Increased digital resource consumption in higher educational institutions and the artificial intelligence role in informing decisions related to student performance." *Sustainability* 14.4 (2022): 2377.

[11] Lawal, Olanrewaju, and Felix E. Anyiam. "Modelling geographic accessibility to primary health care facilities: Combining open data and geospatial analysis." *Geo-Spatial Information Science* 22.3 (2019): 174-184.

[12] Liu, Xue, et al. "Spatial Variations of Village-Level Environmental Variables from Satellite Big Data and Implications for Public Health–Related Sustainable Development Goals." *Sustainability* 14.16 (2022): 10450.

[13] Bello, Isah Mohammed, Kabiru Ibrahim Musa, and Atagbaza Ajiri Okpure. "A Robust Approach to Determining Under-served Settlements for Health Using Geographic and Spatial Coverage Modelling in Bauchi Local Government Area." *Texila International Journal of Public Health* 10.2 (2022): 196-207.

[14] Goel, Raj Kumar, and Shweta Vishnoi. "Strengthening and Sustaining Health-Related Outcomes Through Digital Health Interventions." *Journal of Engineering Science & Technology Review* 16.2 (2023).

[15] Hammond, Natalie, and Angelo Moretti. "Data like any other? Sexual and reproductive health, Big Data and the Sustainable Development Goals." *Sexualities* (2023): 13634607231152599.

[16] Lee, Hakyung, et al. "Enhancing the resilience of the rainwater for drinking (RFD) system through systematic monitoring: a case study at the Ly Nhan rural hospital in Vietnam." *Journal of Water, Sanitation and Hygiene for Development* 11.6 (2021): 1048-1061.

[17] Jakubik, Johannes, and Stefan Feuerriegel. "Data-driven allocation of development aid toward sustainable development goals: Evidence from HIV/AIDS." *Produc*tion and Operations Management 31.6 (2022): 2739-2756.

[18]Li J, Chen W H, Xu Q, et al. Detection of self-reported experiences with corruption on twitter using unsupervised machine learning[J]. Social Sciences & Humanities Open, 2020, 2(1): 100060.

[19] Patuelli, Alessia, and Fabio Saracco. "Sustainable development goals as unifying narratives in large UK firms' Twitter discussions." Scientific Reports 13.1 (2023): 7017.

[20] Banzhaf, Ellen, Henry N. Bulley, Justice Nana Inkoom, and Sebastian Elze. 2022. "Mapping Open Data and Big Data to Address Climate Resilience of Urban Informal Settlements in Sub-Saharan Africa." Climate 10(12): 186.

[21] Mihály, Szabolcs et al. 2021. "Earth Observation and Geospatial Big Data Management and Engagement of Stakeholders in Hungary to Support the SDGs." Big Earth Data 5(3): 306–51.

[22] Liu, Yu, Bo Huang, Huadong Guo, and Jianguo Liu. 2023. "A Big Data Approach to Assess Progress towards Sustainable Development Goals for Cities of Varying Sizes." Communications Earth & Environment 4(1): 66.

[23] Chen, I-Chun, and Bo-Chieh Yang. 2022. "Developing Decision Model and Sustainable Mapping to Screen the Efficiency of Brownfield Redevelopment Based on Socioeconomic Open Data." Sustainable Environment Research 32(1): 29.

[24] Arfanuzzaman, Md. 2021. "Harnessing Artificial Intelligence and Big Data for SDGs and Prosperous Urban Future in South Asia." Environmental and Sustainability Indicators 11: 100127.

[25] Han, Liying et al. 2022. "Assessing Spatiotemporal Changes of SDG Indicators at the Neighborhood Level in Guilin, China: A Geospatial Big Data Approach." Remote Sensing 14(19): 4985.

[26] Xie, Zunyi & Phinn, Stuart & Game, Edward & Pannell, David & Hobbs, Richard & Briggs, Peter & McDonald-Madden, Eve. (2019). Using Landsat observations (1988-2017) and Google Earth Engine to Detect Vegetation Cover Changes in Rangelands - A First Step Towards Identifying Degraded Lands for Conservation. Remote Sensing of Environment. 232. 10.1016/j.rse.2019.111317.

[27] Wu, Bingfang & Tian, Fuyou & Zhang, Miao & Zeng, Hongwei & Zeng, Yuan. (2020). Cloud services with big data provide a solution for monitoring and tracking sustainable development goals. Geography and Sustainability. 1. 10.1016/j.geosus.2020.03.006.

[28] Guo, Huadong & Liang, Dong & Chen, Fang & Shirazi, Zeeshan. (2021). Innovative approaches to the Sustainable Development Goals using Big Earth Data. Big Earth Data. 5. 263-276. 10.1080/20964471.2021.1939989.

[29] Poleshchuk, Ekaterina & Marshall, Ayodele & Tottrup, Christian & Ghafari, Dany & Coppens, Ludgarde & Crane, Stuart. (2022). Use of big data for official environment statistics: The measurement of extent and quality of freshwater ecosystems. Statistical Journal of the IAOS. 38. 1-16. 10.3233/SJI-220041.

[30] Owers, Christopher & Lucas, Richard & Clewley, Daniel & Planque, Carole & Punalekar, Suvarna & Tissott, Belle & Chua, Sean & Bunting, Pete & Mueller, Norman & Metternicht, Graciela. (2021). Living Earth : Implementing national standardised land cover classification systems for Earth Observation in support of sustainable development. Big Earth Data. 5. 1-23. 10.1080/20964471.2021.1948179.

[31] Giuliani, Gregory & Chatenoux, Bruno & Benvenuti, Antonio & Lacroix, Pierre & Santoro, Mattia & Mazzetti, Paolo. (2020). Monitoring land degradation at national level using satellite Earth Observation time-series data to support SDG15 -exploring the potential of data cube. 10.1080/20964471.2020.1711633.

[32] van den Homberg, Marc & Susha, Iryna. (2018). Characterizing Data Ecosystems to Support Official Statistics with Open Mapping Data for Reporting on Sustainable Development Goals. International Journal of Geo-Information. 7. 456. 10.3390/ijgi7120456.

[33] Assarkhaniki, Zahra & Sabri, Soheil & Rajabifard, Abbas. (2021). Using open data to detect the structure and pattern of informal settlements: an outset to support inclusive SDGs' achievement. Big Earth Data. 5. 1-30. 10.1080/20964471.2021.1948178.

[34] Han, Liying & Lu, Linlin & Lu, Junyu & Liu, Xintong & Zhang, Shuangcheng & Luo, Ke & He, Dan & Wang, Penglong & Guo, Huadong & Li, Qingting. (2022). Assessing Spatiotemporal Changes of SDG Indicators at the Neighborhood Level in Guilin, China: A Geospatial Big Data Approach. Remote Sensing. 14. 4985. 10.3390/rs14194985.

[35] Borkowska, Sylwia & Pokonieczny, Krzysztof. (2022). Analysis of OpenStreetMap Data Quality for Selected Counties in Poland in Terms of Sustainable Development. Sustainability. 14. 3728. 10.3390/su14073728.

[36] Ravanelli, Roberta & Nascetti, Andrea & Cirigliano, Raffaella & Rico, Clarissa & Leuzzi, Giovanni & Monti, Paolo & Crespi, Mattia. (2018). Monitoring the Impact of Land Cover Change on Surface Urban Heat Island through Google Earth Engine: Proposal of a Global Methodology, First Applications and Problems. Remote Sensing. 10. 1488. 10.3390/rs10091488.

[37] Kulmala, Markku & Lintunen, Anna & Ylivinkka, Ilona & Mukkala, Janne & Rantanen, Rosa & Kujansuu, Joni & Petäjä, Tuukka & Lappalainen, Hanna. (2021). Atmospheric and ecosystem big data providing key contributions in reaching United Nations' Sustainable Development Goals. Big Earth Data. 5. 1-29. 10.1080/20964471.2021.1936943.

[38] Cuca, Branka & Agapiou, Athos. (2021). Contribution of Earth Observation and Geospatial Information for Urban Planning of Historic Cities' Centres: The Case Study of Nicosia, Cyprus. Sustainability. 13. 7023. 10.3390/su13137023.

[39] Syed Nazim Ali, Syed Nazim Ali. 2020. "Big Data, Islamic Finance, and Sustainable Development Goals." journal of king Abdulaziz University Islamic Economics 33(1): 83–90.

[40] Gholami, Hamed, Jocelyn Ke Yin Lee, and Ahad Ali. 2023. "Big Data Analytics for Sustainable Products: A State-of-the-Art Review and Analysis." Sustainability 15(17): 12758.

[41] Gössling, Stefan. 2021. "Technology, ICT and Tourism: From Big Data to the Big Picture." Journal of Sustainable Tourism 29(5): 849–58.

[42] Kumar, Satish et al. 2022. "Past, Present, and Future of Sustainable Finance: Insights from Big Data Analytics through Machine Learning of Scholarly Research." Annals of Operations Research. https://link.springer.com/10.1007/s10479-021-04410-8 (November 29, 2023).

[43] Sucupira Furtado, Lara et al. 2023. "A Framework for Digital Transformation towards Smart Governance: Using Big Data Tools to Target SDGs in Ceará, Brazil." Journal of Urban Management 12(1): 74–87.

[44] Allam, Zaheer, Houriiyah Tegally, and Meelan Thondoo. 2019. "Redefining the Use of Big Data in Urban Health for Increased Liveability in Smart Cities." Smart Cities 2(2): 259–68.

[45] Abdalla, Alaa Amin et al. 2022. "Connections between Big Data and Smart Cities from the Supply Chain Perspective: Understanding the Impact of Big Data." Sustainability 14(23): 16161.