Review of Big Data Integration in Construction Industry Digitalization

Omar Sedeeq Yousif1*, Rozana Binti Zakaria1, Eeydzah Aminudin1, Khairulzan Yahya1, Abdul Rahman Mohd Sam1, Loganathan Singaram1, Vikneswaran Munikanan2, Muhamad Azani Yahya2, Noraziah Wahi3 and Siti Mazzuana Shamsuddin3

1Construction Research Center, School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia, 2Department of Civil Engineering, Faculty of Engineering, Universiti Pertahanan Nasional Malaysia, Kuala Lumpur, Malaysia, 3Faculty of Architecture Planning and Surveying, Universiti Teknologi MARA, Shah Alam, Malaysia

The 2030 agenda for sustainable development has embraced the importance of sustainable practices in the construction industry. Parallel to the Industry revolution 4.0, the construction industry needs to keep pace with technological advances in data management to keep pace with the revolution through the ability to process and extract value from data. This phenomenon attracts the requirement of Big Data (BD). The construction industry deals with large volumes of heterogeneous data, which is expected to increase exponentially following an intense use of modern technologies. This research presents a comprehensive study of the literature, investigating the potential application of BD integration in the construction industry. The adoption of such technologies in this industry remains at a nascent stage and lags broad uptake of these technologies in other fields. The Construction Industry is driving to boost its productivity through the implementation of data technologies; hence, significant research is needed in this area. Currently, there is a lack of deep comprehensive research on BD integration applications that provide insight for the construction industry. This research closes the gap and gives an overview of the literature. The discussion presented the current utilization, the issues, and ways for potential works along with the challenges companion with the implementation.

Keywords: big data, construction digitalization, data integration, project management, technological innovation

INTRODUCTION

Data integration is a series of processes applied to retrieve and mix data from several sources into meaningful and useful information. An entire data integration solution delivers trusted data from a range of sources (IBM, 2020). Traditional data integration techniques mainly depended on the ETL (Extract, Transform, and Load) process to ingest and clean data, thereafter loading it into a data repository. Presently, a massive volume of data is collected from numerous heterogeneous data sources which are generating data in real time with different qualities which are considered Big Data (BD) (Konikov and Konikov, 2017). BD integration is extremely challenging especially after the normal data integration techniques are let down to handle it. BD integration varies from traditional data integration in many parameters—Volume, Velocity, Validity, Visualization, Variety, Value, and Veracity—which are the BD’s main characteristics. Construction project success is directly related to continuous access to accurate project data. A construction project generates an enormous amount of complex, specific, and professional data, although much effort has been applied to overcome these challenges (Kagan, 2019). This research is aiming to provide an overview of the big data integration
technologies, opportunities, and challenges and to show some of the latest studies made in this domain.

The construction industry is dealing with significant data arising from diverse disciplines throughout the life cycle of a project. The main problem and questions are what is the definition of these data, from where to extract, how to transform, then how to visualize to integrate them to one platform. How can this data integration be measured? The capability to analyze big amounts of data and to extract beneficial insights from it has revolutionized the community and the innovation; hence, it has led to tremendous sustainable development around the globe. The construction industry is not new to data analytics application; the utilization of such application in the industry is left over at a budding stage and slows the wide usage of these applications from other industries (Bilal et al., 2016a).

Therefore, significant research is needed in this area. Consequently, this paper intends to establish a current overview of data technologies from the viewpoint of the construction sector with the aim of providing some possible directions for accelerating future works, presenting opportunities and barriers, and demonstrating some of the recent research in this area. These were accomplished by defining and evaluating the implementation of data technologies for integrating construction data to get the benefit from it, from research papers within the construction industry according to the context and the collection of keywords used for this analysis from 2015 till 2021. Based on the in-depth critical literature review of journal articles, thesis', books, reports published, and conference proceedings and on content analysis, the main keywords on which the research was built were big data, construction digitalization, data integration and visualization, and data-driven construction project management. The flow map of the research are shown in Figure 1 alongside the objectives to determine the relative studies in English language that can be accessed from several databases such as Scopus, Web of Science, IEEE Xplore, ScienceDirect, and Google Scholar and analyze the potential applications of big data integration in the construction industry.

The research has analyzed the literature in depth to uncover many aspects, starting from the data revolution in the construction industry, and explore ways to digitize the construction industry. This is followed by an overview of data generation and how it is utilized with a summarization of the data collections, approaches, and systems. To finalize the research and form an integrated and comprehensive picture, all techniques, stages, preprocessing techniques, parameters, benefits of adoption, and recent challenges for adopting the integration of the analyzed data technologies were covered. In conclusion, a construction research context and detail of the research area has been analyzed and presented with highlighting the important keywords of the big data integration research from various literatures.

REVOLUTION IN CONSTRUCTION INDUSTRY DATA

Data Revolution is improving the working dynamics of projects by expediting innovations in its aspects and refining decision making, boost productivity, and organizational abilities. The construction industry is currently suffering from a variety of problems, the most important of which is the severe inefficiency that leads to lower productivity, which in turn risks costing the global economy an estimated $1.6 trillion a year (Barbosa et al., 2017). Adopting technologies like BD integration sounds inevitable. Moreover, loss of human life due to poor safety and hazardous nature of work are due to the above problems. The construction industry is experiencing outstanding challenges
like the shortage of digitalization, poor project management, ineffective design, poor worker safety, increased greenhouse gas emissions, and a volatile construction economy (Li et al., 2017c; Li et al., 2017d).

The incorporation or implementation of innovation and other mechanisms of change is one of the fundamentals of development, but the construction industry was lagging back in comparison to the manufacturing industry. Although the breadth of data may not be as “large” as the volume of data in retail or financial sectors, the construction sector deals with very heterogeneous data from various sources such as plans, bill of quantities, building specifications, rate of price, and day-to-day data project progress. Given this significant demand, there is an urgent need to analyze and integrate the above multivariate data (Moeni et al., 2017).

Peiffer (2016) declared that BD integration is on top of the considerable leading factors in configuring the way which would lead toward improving the efficiency of the industry. Also in the architecture, engineering, and construction industries, the adoption of data and information technologies enhances a deep transformation of the industry for modeling, designing, and managing intelligent construction (Parisi et al., 2021). However, construction digitalization in data science management is incompetent, which according to (Castagnino et al., 2016) was the outcome of the late and slow locomotion made by the industry to take over new technologies. This is confirmed by the MGI’s indicator that places the construction industry on the bottom of the list of the digitized industries in the world. Renz et al. (2016) added up to the deliberate modification made by the industry that is caused by inadequate data-driven decision-making.

**DIGITALIZING THE CONSTRUCTION INDUSTRY**

Digitalizing the construction industry and integrating the data with suitable methods of visualization over the sector, deployment of such technologies and processes is essential to the required advancement of the construction industry. This advancement gives new opportunities over the entire value chain, through the project life-cycle phase. While having different characteristics to data analysis, the increasing volume of asset data and rising use of Building Information Modeling (BIM) is likely to create further demand for data analysts with roads agencies and local governments (Hart et al., 2018). Within 10 years, according to (Gerbert et al., 2016) estimation, full-scale digitalization in non-residential construction will lead to annual global cost savings of $1.2 trillion (21%) in the engineering and construction phases and $0.5 trillion (17%) in the operation phases. However, governmental projects like transportation and power plants will achieve a 15%–25% decrease in construction and engineering expenses. The operation of construction will face an 8%–13% possible reduction for all infrastructure projects (Gerbert et al., 2016).

Data analysis and integration make new contemplation from the big data collections generated from the project life cycle. New virtual reality (VR) and augmented reality (AR) methods help to distinguish interconnections and collisions during the design and construction phases. Through the use of mobile communication and AR, project management can engage and communicate in real time and provide engineers with additional information on site. The digital technologies currently available on the market are easily usable throughout the project (Gerbert et al., 2016; Renz et al., 2016). **Figure 2** shows the application areas of digital technologies with the construction industry.

Data integration technologies and their usefulness in reducing costs and optimizing asset use and value through more effective asset and demand management will drive demand for technologist skills (to design, implement, manage, and secure data systems) as well as informers (data analysts, economists, planners) who can crunch the data, develop policies, and communicate strategies within agencies and across their stakeholders. Designers themselves (engineers, spatial scientists), via their skills incorrectly interpreting data, are likely to provide a crucial link between technologists and informers. Consequently, the engineer of the future will likely have a different range of skills beyond traditional civil functions in a more technological, data-infused future. The digitalized uprising has affected the construction industry highly to some extent as the industry is working with a diverse and big amount of data (Bilal et al., 2016b).

3D printing, BIM, real-time data, cloud computing, AR, VR, drone scanning, BD, and The Internet of Things (IoT) are not all the tools used in today’s world; the integration and deployment of these technologies is the future (You and Feng, 2020). Utilization of these technologies is the solution to breaking the challenges faced by the construction industry; also, it requires stakeholders who think fast and take smart decisions for the gross profits of their organizations and customers (Chaurasia and Verma, 2020). This adoption lets the construction sector players from the project manager to the workers practice the right and quick decisions, optimize design and sector automation, and decrease the risk of construction. Finally, the construction duration can be shortened by boosting productivity (Hafifi Che Wahid et al., 2019).

In North America and Europe, 40% or more of the companies analyze big data, and these companies currently benefit from these analyses (Bange et al., 2015). One of its strongest benefits is strategic decisions with an increase (69%), better control over the operational process (54%), greater customer absorption (52%), and lower costs (47%). Moreover, these companies increased their revenue from data analysis by 8% and a decrease in spending by 10% (Bange et al., 2015). Analyzing and combining big data has many benefits, and many challenges have to be faced, for example, data privacy and data security, which are the most important factors to be wary of in those companies that initiated the use of big data in place (Bange et al., 2015).

In the construction industry, the current operations in the construction data platform depend on the exchange of data between the structural, architectural, electrical, and mechanical divisions; however, they involve challenges of interconnection. Thus, to achieve various usage cases such as the comprehensive integration of data from BIM with material data, geographical
information, machine sensors, and much more, this large-scale operation will allow the construction industry to realize the automation of its operations, which will improve the overall efficiency of the project. Integrated data that are available to users on the Internet have proven effective (Pauwels and Terkaj, 2016).

Technologies such as Social–BIM and BIMcloud are used in many industries; this would improve the adoption of BD (Bilal et al., 2016b). Plageras et al. (2018) proved the usability of IoT based on BD for smart buildings. For infrastructure and smart city planning, the Master Data Management has been used as a data-based application (Ng et al., 2017). Various studies have demonstrated the big opportunities engaging the BD-related technology for construction (Barima, 2017).

The construction sector has some barriers to adopt innovation at a larger scale because of the risks that can be associated with anything new that has not been well tested (Yousif et al., 2018; Qubole, 2019). Therefore, there is still more room to improve in developing an effective data integration platform that helps in storing and processing interconnected and diversified project data. This platform led to the foundation for creating efficient applications for smart buildings (Bilal et al., 2016b).

DATA AND INFORMATION GENERATION AND UTILIZATION

Since the project data are collected by different units, methods, and times and stored in various formats, it is necessary to integrate the data (Foote, 2019; Ramlia et al., 2019). The integration process converts data into useful information and therefore can improve the decision-making process for all levels of management (Ribeiro et al., 2015). The BD consolidation project is a difficult challenge as contractors must provide the required data in appropriate forms and in real time to benefit from them in decision-making (Vinitha and Ravichandran, 2018). Three categories describe the evolution of an effective data management process in terms of collecting project data, storing data in the system, and using approaches to produce information and improving business decisions (Woldesenbet et al., 2016). Figure 3 shows two generations of data and information development involving the three categories to be used in the construction industry.

BIG DATA INTEGRATION TECHNOLOGIES

Data integration over many application domains is required to fetch data from one environment to another, in another meaning from the source to the destination. Thus, therefore, to grasp the current status around big data integration, the key literature in this area has to be analyzed. The ETL (Extract, Transform, and Load) process has been utilized to achieve this in traditional data repository domains (Alshiekh, 2021). The ETL process combines three essential procedures required to bring the data from the source to the desired destination. The functions are first to extract and read data from the source, second to transform the form and data integration of the extracted data so that it fits the requirements of the targeted destination, and last to load the data to the final destination repository (Wang et al., 2019).

Conventionally, batch processing and ETL have been utilized together in data warehouse environments (Hohensinn, 2021). Data warehouses provide users with a route to integrate data and information over various sources to analyze and visualize relevant data to their specified work focus. The ETL process is used to modify the data into the layout desired by the data warehouse. The modification happens in an intermediate spot before the data are loaded to the data warehouse. Many application vendors, for
example Talend, IBM, Informatica, Oracle, Pentaho, and Microsoft, provided traditional ETL software applications tools (Ahmed et al., 2017).

After the appearance of BD, cloud computing, the internet of thing, BIM, and artificial intelligence the conventional data repository systems insufficiency to, which raise the necessity for ameliorating and utilizing more efficient and effective technologies (Motawa, 2017; Lu et al., 2019; Marzouk and Enaba, 2019). As the number of construction projects that were built over time increased, an enormous amount of data had to be handled (Martínez-Rojas et al., 2015). One of the fastest-growing technologies in several industries is data integration and its applications. Data integration can be defined as a series of computer technologies that can store, process, combine, and manage much more data than were handled before. From this next-generation technology, data can be easily gathered and analyzed. The utilization of data integration in an optimum way could unquestionably be the new limit of innovation to the construction industry. Therefore, data integration technology needs to be applied (Foote, 2019). Thus, therefore it is a solution to handle effectively such data with the existing methodology for data storage and processing.

The components of the BD integration application administer and organize data in new ways in contrast to the classical relational database, consequent to the necessity for scalability and the high ability for administering all types of structured, semi-structured, and unstructured data. Moreover, conventional ETL tools are improving to leverage the new BD features; however, conventional types of integration carry on new signification in the era of BD, and the data integration technologies lack a mutual platform that supports data fineness and tuning. As abovementioned, conventional data integration was done by utilizing batch processing, while BD integration can be done in real time. That could make the ETL process be reordered to turn into ELT in some conditions, so the data are extracted then loaded into distributed file systems and then transformed before being used (Bansal and Kagemann, 2015).

Johnson and Sargunam (2020) defined BD as a massive scale of data that is complex in structure and has difficulties in storing and analyzing the data. Also, when there is a huge amount of data, it will be hard to handle and perform an effective analysis by using traditional techniques or methods. A data preprocessing technique can overcome the problem of data inconsistency, incompleteness, scalability, timeliness, and data security, as shown in Figure 4. Ridzuan et al. (2021) analyzed the issues and challenges of big data in three main areas which are storage issues, management issues, and processing issues. Data management insights require getting all the data collected and linking it in a way to extract important information. In response to that, Rawat and Yadav (2021) demonstrated a theoretical review on big data including its application, opportunities, challenges, techniques, and problems. However, the technology needs to be continually progressing to address the ever-changing issues and challenges in BD.

In the construction industry in parallel with all other sectors, BD refers to the huge quantities of information that have been stored in the past and that continue to be acquired today. The construction industry is familiar to work with a tremendous amount of data; the exploitation of these data would be the new

![Diagram of Two generations of data and information management for construction projects.](https://www.frontiersin.org)
limits for construction industry growth (Ismail et al., 2018). The
determination of BD integration could vary from diverse
literatures. The VS family is typically used to characterize BD
as volume, velocity, variety, veracity, etc., as shown in Figure 5
(Adluru et al., 2015; Hashem et al., 2015; Patgiri and Ahmed,
2016).

Figure 5 shows the data integration parameters and their
meanings, as well as their ability to add a more detailed meaning
to the main elements. First is Velocity; it refers to the amount
of data sources growing and the ratio of data produced over time
extremely increasing, particularly after the outcrop of social
media and the Internet of things. Second is Variety, which
refers to extra data sources that imply that we have a wide
variety in how data are stored. There is a wide range of forms
like photos, texts, tracks, documents, and spatial data. Third is
Veracity, which refers to the features listed above, giving rise to
that we own various data qualities, so we can find fuzzy or
anonymous data, especially from social media, and authorize
users to publish these types of data (Halaweh and Massry, 2015).
Last is Volume; it is the premier parameter of big data (Gandomi
and Haider, 2015; Alavi and Gandomi, 2017). Currently, the
number of connected devices, machinery, and people are higher
than before. In contrast, this connection highly affected the
number of data providers and the size of data widespread
(Shaw et al., 2021).

It was found that construction projects are directly
proportional to the development of digital data and their
analysis to improve and increase profitability. A lot of BD
applications are under construction to do this. Thus, it has
become easier to use these tools, especially by increasing
previous practical experiences that can be followed. At the
same time, customers are hoping that construction
companies will turn more toward big data, to follow the lead
of other industries (Veras et al., 2021). Extracting facility
management information digitally from BIM rather than
going deep into large amounts of documents and paper
drawings and handing it over to a new customer as fast to
use digital files is a distinct application. The way to continuously
profitably use BD depends on the use of such applications
(Zhang et al., 2017).
Achieving a more accurate vision in the future relies on a solid database to monitor changes, mainly dependent on the development of BD (Li L. et al., 2017; Li T. et al., 2017). Recently, studies have shown how big data integration contributes as a technological compliance for massive information analysis on the unsuspected risk to occur. The future will largely depend on the concept of "datafication" and technological progress in terms of its creation and participation, as machines will communicate with each other via data networks, thereby reducing human participation in the process (Hasan et al., 2019).

The increasing adoption of digital technology and the rapid proliferation of data have spurred the application of data analytics and BD to drive smart project and asset management (Aibinu et al., 2019). Data analytics is an enabler to risk prevention and immediate mitigation. Knowing its great benefits, the construction industry is left behind in integrating big data and adopting digital technologies (Bilal et al., 2016b). Cabrera-Sánchez and Villarejo-Ramos (2020) found that companies’ ability to collect and store valuable data greatly affects the adoption of BD integration in various industries, including construction (Raguseo, 2018). The adoption of BD integration enhances productivity and organizational capabilities and also enhances the skills of collecting and analyzing huge amounts of data. Examples of digital technology adoption benefits in the Architecture, Engineering and Construction (AEC) industry are shown in Figure 6.

Changes in administrative processes or methods lead to risks related to the process of data integration, and those risks affect many industries, including the construction industry. The different skills represent an important factor for the emergence of a strong enthusiasm to adopt this modern technology, and successful case studies are a supporter of adding value to it (Madanayake and Egbu, 2019). The reluctance of the risk that the sponsors take to the big changes in their companies constitutes an obstacle to the spread of this technology, which in turn may lead to losing it. The risks involved should be mentioned, such as data security and privacy, costs involved in the process, and finally weak analyses (Oudjehane and Moeini, 2017). Among the challenges in this area are as follows (Shamsuddin and Hasan, 2015):

1. Inefficient data scientists who are responsible for the collection, analysis, and presentation of corporate data. Experts extract knowledge of the deep analyses of data found in dynamic databases.
2. The high cost will be an obstacle for companies to adopt this technology, as an investment in these operations requires a large and sophisticated integration platform.
3. Fear of sharing data, there is a fear of governments and companies in sharing information that could be useful to the world. Therefore, policies relating to the protection of data and information collected about them must be formulated.

Currently is the perfect time to digitize data, as many different use cases have benefited from this technology. A kickstart should be carried out with a pilot project that includes different processes and data types. It is also necessary to give workers in the field to contribute to this by giving suggestions to help improve the process. Furthermore, senior management is the main driver for the success of the process (Bange et al., 2015). It should be borne in mind that change must not be from within, meaning that...
<table>
<thead>
<tr>
<th>Research context</th>
<th>Important keywords</th>
<th>Detail of research area</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Productivity</strong></td>
<td>Output, production, rate</td>
<td>Data from technologies and auto-machines can boost construction performance and quality. Real-time data analysis is beneficial to work productivity.</td>
<td>Skibniewski and Golparvar-Fard (2016)</td>
</tr>
<tr>
<td><strong>Financial management</strong></td>
<td>Cost and benefits</td>
<td>A cost–benefit analysis of indexing big data with map reduction, project expense management using tender price assessment and BD, real-time data for cost–effective design</td>
<td>Chau et al. (2018)</td>
</tr>
<tr>
<td><strong>Complex construction</strong></td>
<td>Complexity, complication, sophisticated</td>
<td>Association analysis of piping materials of an offshore structure using big data technology, using BD and visual analytics for measuring building performance, geospatial/geo-location data for construction, BD in the development of waste simulation application</td>
<td>Zhang et al. (2015)</td>
</tr>
<tr>
<td><strong>Data-driven design</strong></td>
<td>Models, planning, and concepts</td>
<td>Efficient design model comprising BD, driven design decision from past projects BD, development of BD-integrated models for designing green buildings, integration of VR with BD for best building designs and algorithms, urban planning, and smart city construction</td>
<td>Bilal et al. (2016a)</td>
</tr>
<tr>
<td><strong>Asset management</strong></td>
<td>Assets and resources</td>
<td>Tracking, management, governance, and realization of project resources</td>
<td>Akhavan and Behzadan (2015)</td>
</tr>
<tr>
<td><strong>Digital technologies</strong></td>
<td>Digital</td>
<td>Effective predicting, forecasting, and monitoring through the use of digital technologies and automated systems, digital technologies to facilitate a circular economy in the built environment</td>
<td>Hampton (2015)</td>
</tr>
<tr>
<td><strong>Integration systems</strong></td>
<td>Integration, combining</td>
<td>Integrating BIM, IoT, sensors, scheduling data with external data, solve housing needs by leveraging big data and combined with 3D building printers, adoption of LM, VE, BIM, and SDA to address productivity and digitalization issues in the AEC</td>
<td>Martinez-Rojas et al. (2015)</td>
</tr>
<tr>
<td><strong>Prognosis</strong></td>
<td>Detection</td>
<td>Technical assessment of the probable course and outcome of a building defect, structure integrity, generate prediction system for construction businesses bankruptcy</td>
<td>Moustafa (2020)</td>
</tr>
<tr>
<td><strong>Knowledge management</strong></td>
<td>Knowledge</td>
<td>Efficient handling of information and resources within organizations and construction projects</td>
<td>Madanayake and Egbu (2019)</td>
</tr>
<tr>
<td><strong>Construction Project Management</strong></td>
<td>Monitoring</td>
<td>Progress/performance monitoring through IoT devices, evaluate the performance of an intelligent building infrastructure</td>
<td>Janda et al. (2015)</td>
</tr>
<tr>
<td><strong>Project risk</strong></td>
<td>Risk and opportunity management using data simulation for reduction actions, reducing the risk of construction using IoT using IoT sensors</td>
<td></td>
<td>Akbarnia (2018)</td>
</tr>
<tr>
<td><strong>Time, cost</strong></td>
<td>Better time and cost management</td>
<td></td>
<td>Augur (2016)</td>
</tr>
<tr>
<td><strong>Quantifying</strong></td>
<td></td>
<td>Big data for quantity survey, potential opportunities, and data value from the quantity-surveying context</td>
<td>Caron (2015)</td>
</tr>
<tr>
<td><strong>Safety</strong></td>
<td>Site safety, workers’ safety behavior</td>
<td>Envisage and predict workers’ behavior in real time using BD, manage site safety with BD and drones, increase project safety with the integration of BD and BIM, fire detection, and evacuation process for various projects and forests to assist authorities.</td>
<td>Oudjehane and Moeini (2017)</td>
</tr>
</tbody>
</table>

(Continued on following page)
current employees are trained to counter this wave of technological advances in data management. Data privacy and security are a challenge by classifying data and developing appropriate guidelines and responsibilities to improve its security, give it realism, and reduce uncertainty (Bange et al., 2015).

### RESEARCH IN BIG DATA INTEGRATION FROM THE CONSTRUCTION INDUSTRY PERSPECTIVE

The data revolution has started putting its feet in the construction industry in synchronizing with the other industries that have hugely profited from BD integration. In this matter, the construction industry would take advantage of BD integration similar to anticipation by the other industries and services. As mentioned before, this will include improving efficiency and decision-making. Bilal et al. (2016a) substantiated that the expectations on the usability of big data integration in construction would be aggrandized as the exciting constituents. Consequently, the impulse of these constituents and elements would be the indicators to push the industry to the new limits of data-driven actions.

Based on Ismail et al. (2018), the impractical direction gain from the processes of analysis would be summarized as risk and security, venture management, energy utilization, decision-making, design structure, and resource arrangement. Of the above mentioned, BD integration for construction management was specified as the zone in which the intensification of researchers around is observed. Undoubtedly, the construction industry is a data-subordinate industry; therefore, data should be administered efficiently with the proper application and system to assure the success of any project. Current data revolution studies and implementation extracted from diverse studies are abstracted in Table 1 highlighting the context of research, significant keywords specified, and detail of the research area from the overview which stand out.

A web-based application for the BD integration is one of the best solutions to provide a platform for monitoring from manual to automatic, to obtain the information required to evaluate the construction, and to help in speeding and smarter important decisions, improving the construction performance in terms of the sustainable and green approaches. Also, Sun and Zhang (2020) concluded that the integration of smart cities and big data technologies leads to the growth of smart city constructions and enhances green city developments. Above all, it provides an effective platform for competitors to be players in sustainable construction industry development. This innovation makes data analysis and management more user-friendly and efficient, reduce data complexity, have ease in data collaboration, and have data integrity.

Although there are few data integration technologies on construction, not all the tools are fully utilized due to certain issues (Chonpitakwong et al., 2021; Srinavin et al., 2021), such as lack of awareness and responsibility; the apparent lack of use of modern technologies to integrate data to obtain useful information in the fastest time, with minimal effort and cost; lack of proper training; and lack of advertising the tools and monetary allocation to manage the projects introduced.

Encouragement from top management is very important to enhance the application of integration tools in the industry. Besides, stakeholders should also be responsible for their projects and maximize the sustainability of their projects for the good of the people in the future. The use of such technologies and making the best use of information and BD integration technologies are two of the ways to achieve a sustainable project.

### CONCLUSION

Although the construction industry produces a lot of data during the life cycle of the projects, the utilization of such technology in this industry lags the advancement made in other industries. The importance of construction data integration and up-to-date

---

**Table 1** | (Continued) Detailed context of the big data integration research area from various literatures.

<table>
<thead>
<tr>
<th>Research context</th>
<th>Important keywords</th>
<th>Detail of research area</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy management</td>
<td>Consumption, building performance</td>
<td>BD, data visualization, prototype software, analysis of power utilization and social network, real-time monitoring, and computational models; all these technologies are used for establishing integrated projects to improve energy efficiency for better energy management and lower energy-consuming. Virtual sensor modeling method to determine the actual energy efficiency</td>
<td>Lin et al. (2018); Hasan et al. (2019); Janda et al. (2015); Redmond et al. (2015); Moreno et al. (2016); Koseleva and Ropate, (2017); Qian et al. (2020); Martinez-Rojas et al. (2015); Bernstein (2017); Akhavian and Behzadan (2015); Augur (2016); Sadhu (2016); Oudjehane and Moeini (2017)</td>
</tr>
<tr>
<td>Decision-making design framework</td>
<td>Decision-making</td>
<td>Fast and knowledgeable decisions based on BD systems</td>
<td></td>
</tr>
<tr>
<td>Resource management</td>
<td>Resource tracking</td>
<td>IoT with BD system for resource tracking and usage efficiency</td>
<td></td>
</tr>
</tbody>
</table>

---

---
monitoring of the well-integrated data in the construction process of projects is essential.

To improve the adequacy of the construction, the construction players need to leverage these technologies for processing, analyzing, and storing the data. This research has analyzed the range to which the industry has utilized this integration technology. To this end, we reviewed the latest research that has been published in which the new techniques have been established in many fields of application. The fundamentals of BD integration technologies are demonstrated to assist the readers in understanding the complex theme. Significant, previous applications and the adoption of such technologies by many divisions of the industry are discussed.

This technology has been set to leverage the industry; this research outcome is the catalyst in making the required knowledge to back up the implementation of data integration in the industry. This can drive the industry to prepare in improving the abilities in employing this integration technology and to support the organizational development to catch the arriving surge of the data revolution. Collecting and analyzing the data and information to present in one phase or from a project can be utilized as data and/or information in another phase, or information generated from a project can be utilized by another project to manage active projects and/or plan future projects. This phenomenon creates data and information to be used by the construction industry at all times. Furthermore, future research is needed for examining the integrated data to be used for global commercialization. Lastly, continuous research on the applicability of development of a mobile application to be connected to the big data integration systems to present real-time data with low cost and efforts on the data collection process in the construction industry is needed.

**AUTHOR CONTRIBUTIONS**

All authors listed have made a substantial and intellectual contribution to the work and approved it for publication.

**ACKNOWLEDGMENTS**

The authors would like to thank and show their appreciation for UTM and the research team for their inspirational support and guidance from the beginning to the end.

**REFERENCES**


**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Yousif, Zakaria, Aminudin, Yahya, Mohd Sam, Singaram, Munikanan, Yahya, Wahi and Shamsuddin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.