sustainability

Concept Paper

The Potential of Big Data Application in Mathematics Education in Malaysia

Sim Jia Jin [1], Abdul Halim Abdullah [1,*], Mahani Mokhtar [1] and Umar Haiyat Abdul Kohar [2]

[1] School of Education, Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, Johor Bahru 81310, Johor, Malaysia
[2] Faculty of Management, Universiti Teknologi Malaysia, Johor Bahru 81310, Johor, Malaysia
* Correspondence: p-halim@utm.my

Abstract: The world is facing rapid changes after the emergence of innovative technologies. These changes aim to ensure that a country keeps track of current world developments, strengthens its economy, and reduces its dependence on imports. Hence, every country is now amid technological transformation in the industrial sector by replacing manpower with machines to increase production and efficiency, allowing for mass production. Technology advancements in control, information technology, and automation that are applied to business and industry production processes are referred to as 'Industry 4.0'. The objective is to increase the autonomy, adaptability, and effectiveness of decision-making and production processes utilizing cyber-physical systems (CPS), Big Data (BD), artificial intelligence (AI), and the industrial Internet of Things (IoT). Specifically, this article first introduces Industry Revolution (IR) 4.0, followed by a delineation of the concept of BD. Correspondingly, we discuss BD in education and relate mathematics education with BD. The article concludes with the implications of BD for Malaysian teaching and learning practices.

Keywords: Big Data; Malaysian education system; mathematics education

Citation: Jin, S.J.; Abdullah, A.H.; Mokhtar, M.; Abdul Kohar, U.H. The Potential of Big Data Application in Mathematics Education in Malaysia. *Sustainability* **2022**, *14*, 13725. https://doi.org/10.3390/su142113725

Academic Editors: Angelo Natalicchio, Firoz Alam and Umberto Panniello

Received: 25 August 2022
Accepted: 20 October 2022
Published: 23 October 2022

1. Introduction and Motivation

Innovative technologies have brought about rapid global transformation (Chae, 2019). The purpose of these modifications is to ensure that a country stays abreast of global developments, strengthens its economy, and reduces its reliance on imports. Consequently, every nation is currently undergoing a technological transformation in its industrial sector by replacing manpower with machines to increase production and efficiency in cases permitting mass production. People believe that technological transformation in the industrial sector can fundamentally alter how we live and work, thereby enhancing the economic development of a nation's business and manufacturing sectors. However, it is not easy to foresee exactly how this transformation may affect different industries and countries (Industry4WRD(Ministry Of International Trade And Industry, Kuala Lumpur, Malaysia), 2018).

2. Industry Revolution (IR) 4.0: Big Data and Education

Industry Revolution (IR) 4.0 is a process of revolution that builds on the third revolution, where the combination of technology with industry heads toward automation and data exchange between the physical, digital, and biological spheres. It has brought significant changes and challenges in various sectors, especially in economic, social, political, and organizational fields. It is noteworthy that the industrial Internet of Things (IoT), Big Data (BD), cloud computing, cyber-physical systems (CPS), algorithms, data mining, analytics, machine learning, and artificial intelligence (AI) have become some of the most significant technical developments and concepts in recent years [1]. These are accomplished by assisting manufacturing in moving toward the newest technical frontier and strengthening economic conditions (Industry4WRD, 2018). By investing in cutting-edge infrastructure and high-quality education, countries increasingly emphasize creating advanced manufacturing capabilities. This is because school conditions directly impact student performance in learning. This situation was proved by Ramli and Zain [2], where insufficient attention to infrastructural factors influenced student academic achievement, especially in higher education. Hence, it is crucial to ensure that education facilities are equipped with high-tech infrastructure to prepare the future generation well and to cope with the challenges of IR 4.0. According to Cherri [3], technology advancements in automation, control, and information technology are referred to as IR 4.0 and are applied to the manufacturing procedures of businesses and sectors. Attempting to make decision-making and production processes more effective, autonomous, and adaptable is the aim of using the IoT, CPS, BD, and AI. Based on this statement, many decisions are made based on the transformation of collected data to maximize profitability and efficiency, as large-scale data are closer to nature populations, and models do not converge asymptotically [4].

The world community has had multiple revolutions during the past 200 years that have had a significant impact on and have affected many elements of human life. In order to increase the productivity of industrial operations, humans developed mechanical, electrical, and information technology during the first three IRs. The first IR supported the development of machine tools, steam power, and hydropower. The second industrial revolution gave rise to the areas of large-scale manufacturing and electricity production. Additionally, the third industrial revolution, which was brought about by the development of computers and the internet, increased the industrial efficiency of the manufacturing and production sectors. The fourth industrial revolution, led by CPS technology, incorporates the real world with the information age for prospective industrial development [5]. The term 'Industry 4.0' was first coined in Germany in 2011 and described a methodology to promote the shift from manufacturing dominated by machines to digital manufacturing. An essential objective of the nation was to encourage industrial reform and take the lead in global manufacturing [6]. India, Thailand, Indonesia, Malaysia, Brazil, Nigeria, and Pakistan are developing and emerging nations that have had trouble adjusting to Industry 4.0 [7,8]. Thus, this revolution demonstrates how the development of CPS has encompassed whole new capacities for technological approaches, computers, and humans. However, automation technology is viewed as a technical advancement that does not always directly involve labor resources but that can fundamentally alter the nature of work [9].

In education, this fourth IR includes the use of automation, digitization and AI, which has led to new challenges with the discovery of information from large amounts of BD. Analytics and simulation data (BD), the IoT, and modeling are also crucial elements in the IR 4.0 era [10]. Based on the Malaysian National Fourth Industrial Revolution (4IR) Policy [11], IR 4.0 refers to industries' disruptive transformation through the use of emerging technologies. It features new technologies that combine the physical, digital, and biological realms, affecting all fields and industries, as well as the economy. For example, bioprinting uses digital (digital) files to print objects, such as organs (physical), using cells and bio (biological) materials. In addition, one of the cores of their policy is focused on talent and education management components. It includes the potential workforce and the whole present and future labor force. Both the public and private sectors have talented workers today. These are now enrolled in the educational system at the tertiary, secondary, and primary levels. It shows that Malaysia has started to focus on cultivating new talent and is trying to educate them to fulfil the workforce in the future when IR 4.0 arrives.

The digital revolution known as 'IR 4.0' belongs to the process of technology advancement in a particular industry that involves the automation and production sectors at more sophisticated and systematic levels. This process of changing an industry's operation is also known as 'Smart Factory' where each system put in place is more systematic and flexible, in addition to adapting to more difficult, new scenarios [12]. The National Automotive Policy (NAP) 2020 launched earlier this year showed that Malaysia is ready to move in the same direction as other developed countries in the challenges of Industry 4.0. Based on a report from "Positioning Malaysia as a Regional Leader in The Digital Economy: The Economic Opportunities of Digital Transformation and Google's Contribution" [13], if digital technologies are fully utilized, Malaysia's economy could reach MYR 257.2 billion (USD 61.3 billion) annually by 2030. Digital technologies can also significantly boost Malaysia's economy. Notably, eight essential technologies can revolutionize the nation. These include the IoT, remote sensing, advanced robotics, additive manufacturing, cloud computing, BD, AI, and financial technology (Fintech). According to "Positioning Malaysia as a Regional Leader in the Digital Economy: The Economic Opportunities of Digital Transformation and Google's Contribution" [13], these technologies can increase productivity, save time, and cut costs, as well as boost tax collection, wages, and revenue for Malaysia's businesses, workers, and government, and increase the gross domestic product (GDP).

The World Economic Forum (WEF) outlined the significance of 21st-century skills, including mastery of 4C aspects, which comprise problem solving, critical thinking, communication, creativity, and collaboration at all levels of education, to meet the issues of IR 4.0. If higher-order thinking skills (HOTS) become the main focus, students should have to master 4C skills in their life. In addition, every student should have problem-solving skills, be an expert in written and verbal communication skills, possess leadership skills, have high self-confidence, and always be creative and wise to take advantage of all available opportunities [9]. In the meantime, the Malaysia Education Blueprint 2013–2025 states that science and mathematics should be emphasized because these subjects can stimulate creativity and innovation. By 2020, the initiative should expanded to 10,000 schools, serving 450,000 teachers and 5.2 million students [14].

In addition to industry manufacturing fields, Industry 4.0, more or less, has impacted the academic field in Malaysia, mainly focusing on the innovation element. The fourth IR has forced the academic field to struggle with innovative, open, and flexible learning environments to equip students with the mindset and skills of Industry 4.0 to prepare them for the work ecosystem of the future [15]. According to Li [16], the transformation and creation of CPS are so compelling that changes in higher education are needed, which include training talents and educating the future workforce. These changes in the education system are expected to enhance students' innovation and productivity, as well as society; the future workforce can guide the country to compete globally. These changes can be made in the subject of mathematics, as it is a subject that aims to form students who can think logically and systematically in decision making and deal with daily life challenges [17]. Therefore, in the latest revised curriculum, the contents of mathematics textbooks implemented many elements where students are expected to adapt knowledge and mathematics skills to various strategies and apply them in solving daily life problems in Chinese national primary schools. All the efforts have shown that the Malaysian government wants to cultivate a young generation that can lead Malaysia to a higher phase and be able to cope with the challenges met in IR 4.0.

In conclusion, Industry 4.0 refers to an industry's disruptive transformation through emerging technologies. The industrial IoT, Big Data, cloud computing, algorithms, data mining, analytics, machine learning, and AI are recent technological advances. The fourth industrial revolution merges the real world and the information age. Malaysia's National Automotive Policy (NAP) 2020 shows that the country is ready for Industry 4.0.

3. Big Data

Researchers have differentiated and characterized Big Data (sometimes referred to as BD) in a variety of ways. Therefore, it is crucial for us to be able to draw the best conclusions from vast amounts of data and apply them to our judgments [4]. According to International Business Machines, the concept of BD is simply the management of enormous amounts of data and the ability to process them swiftly [18]. Many experts have attempted to categorize and define BD in various ways. Numerous authors have discussed BD from various perspectives, emphasizing various features. The research by Sabaityt et al. [19] demonstrated how V-models have changed over time and interacted with BD models.

Velocity, variety, and volume are the three critical characteristics of BD that make up the 3V paradigm [20]. Apart, velocity, variety, and volume alone are insufficient to adequately characterize BD, as they only take into account the specifics of the data and ignore the larger BD context [21]. In more recent works, authors have expanded the idea of BD by using a 5V model, which is supported by a 3V model that includes two additional elements of value and veracity [22,23], as well as 6V [24] and 7V [25] models containing visibility and validity features presented to define BD more precisely. V-models have evolved from the past literature about BD by complementing new elements with time-based BD properties. Furthermore, Emmanuel and Stanier [21] described BD in the context of the 'why' of BD, the 'what' of BD, the data characteristics of BD, the 'how' of BD, the supporting and processing architectures utilized with BD, and BD applications.

Torrecilla and Romo [4] showed that BD made data learning a core scientific discipline. They noted that the concept of BD was an important aspect of innovation that has recently attracted much attention from academics and practitioners [26]. BD is the term used to describe vast amounts of data that cannot be handled by conventional methods, technologies, or ways of processing data. Numerous industries depend heavily on BD, including banking, agriculture, education, data mining, finance, chemistry, cloud computing, healthcare, and marketing [27]. These days, many scholars are interested in studying BD [28]. In our daily lives with digital hardware and software, we are constantly generating information that can be used to identify where we go, what we like, whom we know, how we feel, what we do, what we consume, and so on [1]. The technological revolution and digitization of most fields have caused information and data to rise exponentially. The implementation of recommendations that can boost reality and decision-making processes can be facilitated by analyzing and interpreting large-scale data collected in the context of various phenomena and the factors that affect their evolution and growth. Thus, BD or mass data, can be said to be born for this purpose [29]. Based on the data collected and analyzed, artificial intelligence (AI) increasingly understands individuals and even significant populations or societies to serve everyone better based on their needs. This can be generalized to the fact that most low-cost activities generate data with valuable information, and better decisions are made based on evidence rather than intuition [4]. The ever-changing modern society calls for a transformation in education [30].

Other than that, the concept of BD refers to the phenomenon of the creation, collection, communication and use of digital data that has grown rapidly and gradually replaced traditional datasets based on databases and data analysis. This clearly shows that BD encompass all data or information and are more complex than traditional datasets [31]. In the traditional perspective, structured data contain quantitative and qualitative data that are mainly generated and that may be displayed using standard statistical charts or indicators. Moreover, they follow established structures and guidelines [32]. Nonetheless, BD also comprise semi-structured, unstructured, and heterogeneous data that contain signals that can be captured and stored, in addition to structured data [33]. Gandomi and Haider [25] stated that unstructured data are represented by text, image, audio, and video files, whereas structured data are tabular data found in databases and spreadsheets. Therefore, it is a prime example of semi-structured data for Extensible Markup Language. Additionally, according to Naveen [34], there are five key sources of BD: the cloud, media, data, the web, and the Internet of Things (IoT).

In conlcusion, BD include the management and quick processing of large amounts of data. BD promote data science. Banking, agriculture, data mining, finance, chemistry, cloud computing, healthcare, and marketing rely heavily on BD. Mass data (BD) are digital data creation, collection, communication, and use. Digital data replace databases and data analysis. BD are more complex than structured and unstructured data.

3.1. Background of Big Data

The development of Big Data (BD) occurred with the rise of the IoT in the present educational system. IoT technology can collect information from various smart devices or sensors following an agreed protocol and then transfer the information to an application platform to process the information, known as BD [10,35]. Many researchers have tried to describe and define BD in various ways and from many perspectives. According to past BD studies, different researchers have described BD from many points of view. In addition, they have highlighted BD from specific definitions and aspects. For example, BD where data are distinct in many respects, such as volume (too big), velocity (faster arrival), variability (rapid changes), veracity (commotion), and variety (diversity), are described by Memon et al. [27]. Other experts have characterized BD as a sizable, rapidly expanding, and diverse information asset that needs a new processing model to enhance capacities for vision, process optimization, and decision making [36–39].

According to Ma et al. [38], BD focus on a variety of data generated by Internet behavior, including producer and user preferences and intentions, as well as information about nontraditional structures. The emergence of BD represents a new way of thinking to help people explore new knowledge and information for a large amount of data information, create new value, enhance new abilities, form new understandings of the world, and transform the world [40]. In addition, BD are also called large amounts of data, which comes from the article “Big Data Age”. People generally agree that the purpose of BD is to develop new processing models to promote decision making and insight. Meanwhile, at the same time, it can optimize a large number of processes, increase information asset diversity, and promote high growth rates from large amounts of data. It mainly includes the following four basic characteristics: low-value density, which refers to the increasing number and types of information under large data. This leads to a decline in information identification and affects value density. Multi-type refers to users not only using information, but also publishing information under large data. Hence, the number and types of information are constantly increasing. Fast change refers to large data [41]. From the above definitions, we can conclude that the collection of huge volumes of data, including semi-structured, unstructured data, and structured, is known as BD, which are collected through technology by an organization to analyze and obtain useful information for use in the improvement or the development of a particular project. The lifespan of BD is depicted in Figure 1 and includes data generation, acquisition, storage, analytics, and visualization [42].

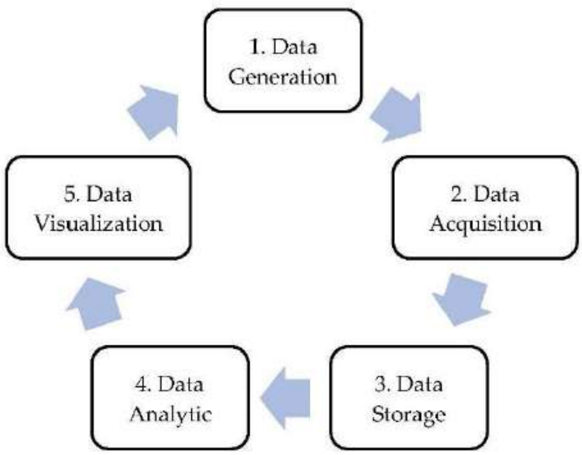

Figure 1. Big Data life cycle [42].

The BD life cycle is divided into five parts: data generation, data acquisition, data storage, data analytics, and data visualization. Data generation gathers information from a variety of sources (video, sensors, click streams, etc.) Meanwhile, data acquisition is the process of gathering information from data by choosing relevant information that is used for the analysis and preprocessing of information for identifying, removing, and filtering irrelevant and inconsistent information. Data analytics, often known as analytics, is the process of obtaining useful insights through quantitative and qualitative methods. After selecting, gathering, and preprocessing data, it is necessary to transform the preprocessed data into a format that can be used for data mining. Data analysis refers to analyzing data after they are transformed. This can be achieved using a variety of statistical techniques and data-mining algorithms, including classification, clustering, and regression. Finally, data insights are represented by evaluation, which quantifies the data analysis outcomes, and the interpretation is displayed in an interactive manner via data visualization [42].

The BD life cycle is modified per the fundamental knowledge of BD in the BD module so that mathematics teachers can easily explain the basic knowledge of BD in their mathematics lessons based on the module provided. However, it is necessary to draw attention to the issue that results from teachers' lack of expertise in managing and analyzing BD. Most teachers do not possess this practical knowledge. Hence, it would be significant to include specialized staff to perform this from a pedagogical perspective and to build training programs in this domain [1]. Given this, it is necessary to reformat the initial and ongoing training of future teachers if BD is to be used in education.

After obtaining a standard definition and concept of BD, experts have begun categorizing and defining BD's characteristics. At first, there were three main characteristics of volume, velocity, and variety, which were introduced by Gartner and are now expanded to be more refined, following in the footsteps of technological developments [39]. According to Laney [20], the three main characteristics were modeled into 3V, which explained that BD had three features: volume, velocity, and variety. Volume is a characteristic that indicates the amount of data collected or generated by an individual or organization [43]. The definition of data volume ranges depending on elements, such as time and data type, as data warehouse capacity continually increases and enables the collection of increasingly huge datasets [25]. This implies that what is currently seen as BD today may not be the same in the future.

Velocity is a term that describes how quickly data are generated and processed [22]. Data processing was initially slow and expensive since firms employed cluster-processing systems (whenever a document was accessed, a program was called to process it) [43]. However, data speeds and the prevalence of digital gadgets have sped up this process over time. This has made it simple for a variety of businesses to carry out real-time BD analysis. Lastly, variety is the last attribute of BD. The term 'variety' refers to the variety of data types that are gathered in numerous datasets. Data are gathered from a variety of sources, including text files, databases, digital data streams, and spreadsheets [44].

Some scholars have summarized that BD have the 4V characteristics of volume, variety, velocity, and veracity [45–47]. Uprichard [48] noted that few other V-words could be adopted to define BD: vampire-like, volatility, venomous, versatility, very violent, vulgar, vitality, virtuosity, vigor, visionary, vibrancy, virility, valueless, viability, and violation. Meanwhile, Lupton [49] recommended abandoning V-words and switching to P-words to characterize BD's traits. BD could be characterized by 13 P-words, including predictive, perverse, portentous, productive, political, personal, practices, partial, privacy, polyvalent, polymorphous, provocative, andpPlayful. The research of Sabaitytė et al. [19] showed that the V-models evolved, and there was interaction between BD models. The V-models became more precise and accurate as the study findings went on. According to Murthy et al. [50], a six-fold taxonomy that emphasized how it was handled and processed rather than its core properties was used to categorize BD. BD had a six-fold taxonomy that included data (a) temporal latency for analysis: near real-time, batch, and real-time; (b) structure: semi-structured, unstructured, and structured; (c) computed infrastructure (streaming or batch); (d) storage infrastructure (NoSQL (Apache Cassandra 4.0.6, Creator: Avinash Lakshman and Prashant Malik; Wakefield, Massachusetts, U.S.A.), Volt Active Data (11.3, Creator: Michael Stonebraker, Sam Madden, and Daniel Abadi; Bedford, Massachusetts, U.S.A.), and SQL 2016(Creator: Donald D. Chamberlin and Raymond F. Boyce; U.S.A.; (e) analysis (reinforcement machine learning, data mining, and statistical techniques or semi-supervised, supervised, and unsupervised); and (f) visualization (abstract, maps, real-(security, data privacy, and management), and interactive). Recently, authors have used a 5V model to expand the BD idea. Veracity and value, two new elements added to the 3V model, make up the 5V model [22,23]. Since most information obtained from numerous sensors or the internet is inaccurate, veracity is crucial [51].

1. The requirements for data collection should outline what constitutes accurate and valid data. Erroneous data should, therefore, be eliminated before the analysis starts. The organization should specify what information is required and for what purposes to minimize inaccurate or unneeded information [23]. Data flows are accelerating, and processing big volumes of data is increasingly challenging, demanding more resources and creating errors.
2. Analyzing the gathered data offers insightful recommendations on how firms can modify their strategies, boost output, and acquire a competitive advantage. This leads to value, which is an additional trait. It is also critical to realize that the value of data comes from analysis, not from the data themselves [51]. The data value creation chain presented by Cavanillas et al. [52] included steps for data collection, analysis, mentorship, storage, and usage (refer to Figure 2). In addition, data analysts may filter the data and track business trends using a range of queries, which is a key element of BD [53,54].

Figure 2. The chain of BD value creation (established by Cavanillas et al. [52]).

In addition, Rabhi et al. [42] and Jiang and Wang [55] extended 3V to 5V, which included volume, velocity, variety, value, and veracity. Furthermore, some of the literature talks about 7V traits to describe BD, including velocity, volume, veracity, variety, visibility, validity, and valuable [56]. From the study by Leung et al. [57], they characterized BD's 7Vs as velocity, volume, veracity, variety, volatility, validity, and value. Nevertheless, Saggi and Jain (2018) defined BD with a 7V model: velocity (the speed of the data), volume (the amount of data), veracity (the accuracy of the data), variety (the structure of the data), volatility (the validity of the data), variability (the meaning of the data), and value (the use of the data). 6V [24] and 7V [25] models featuring visibility and validity features have been proposed for the analysis of the BD idea.

3. Validity: The data logic is described by this trait [58]. This aspect might initially be thought to be related to veracity. However, when discussing validity, it is more crucial to consider if the data match the facts or if they have a meaning that is attributed to them. Furthermore, since it is important to ensure the source of the data is accessible and suitable for storing such data, the need for explanation can be related to personal data protection.
4. Visibility is essential for proving the authenticity and accessibility of the data. With the aid of this function, the user is able to access the data and view them [25].

V-models are changing as new components based on BD attributes are added. Therefore, V-models can be regarded as models of BD characteristics because they are based on the characteristics of BD (refer to Figure 3).

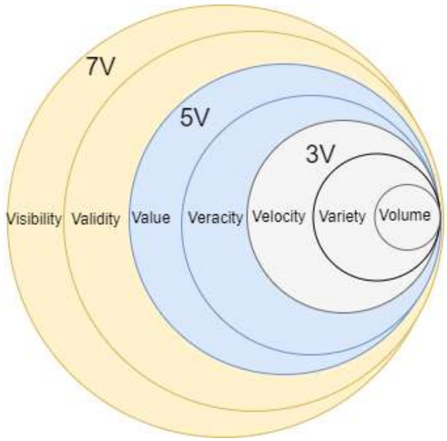

Figure 3. Interaction between Big Data models [22,23,25,58–62].

Recently, the model of 7V has extended its definition to 10V, where the latest models have almost covered all the characteristics of BD [63,64]. The following are the 10Vs:

1. Volume: a vast volume of developed data that conventional processing and storage methods can no longer handle;
2. Velocity: the high speed at which data are processed after being gathered;
3. Variety: the nature of the data, which can take on several forms, such as semi-structured, unstructured, or structured data;
4. Variability: constantly changing data meaning, which may affect homogeneity and consistency;
5. Veracity: characteristics of BD quality. Before continuing, collected data must be preprocessed because they are frequently ambiguous and erroneous;
6. Viscosity: the interval of time between the occurrence and its description, which mostly occurs when working with large datasets
7. Virality: network data distribution and insight-sharing rates;
8. Validity: the quality and suitability of the data for the intended use;
9. Value: the knowledge and insights derived from complex, concealed, and diverse BD;
10. Visualization: A thorough visual data representation that makes it possible to spot hidden patterns and correlations.

In addition to the characteristics of BD in a 10V model, Sandhu [65] also summarized BD classification, as shown in Figure 4.

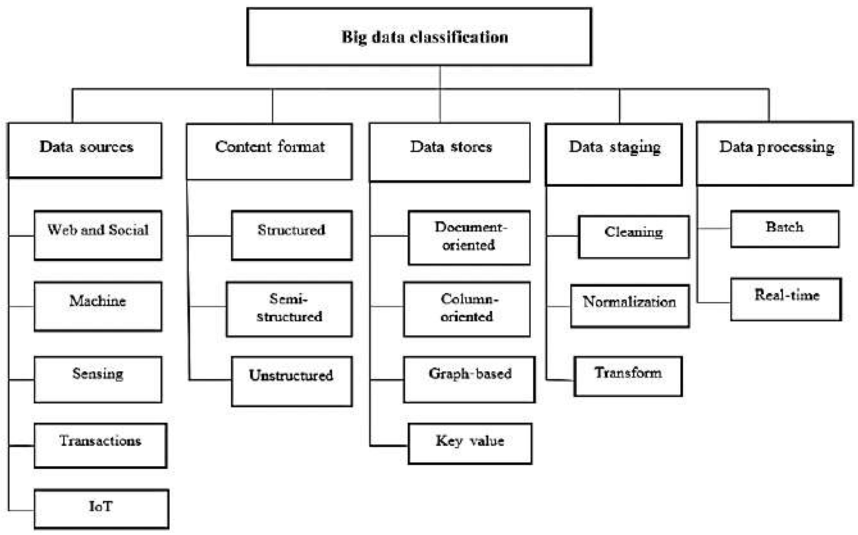

Figure 4. Types of Big Data [65].

Furthermore, BD processing and architectures were defined by Emmanuel and Stanier [21] as scalable architectures that could handle the processing needs of data that were high in volume, had a range of data formats, and included high-velocity data acquisition and processing. Additionally, they defined BD applications as a family of programs that used BD, including but not restricted to BD analytics. In short, BD are complex and have a variety of features that are taken into account to maximize the information that can be obtained.

3.2. Current State of Big Data

Big Data (BD) are closely related to Industry 4.0 technology, which fundamentally changes how students learn and educators teach. Apart from that, the concept of BD is now also an important part of the education sector. Recently, the same learning system in the form of an online platform that shares our lives with us on social media, smartphones, and the web has become more important as it takes over the educational landscape [1]. Therefore, BD have transformed traditional education by facilitating personalized lessons, detailed student modeling, and ongoing longitudinal research that evaluates changes across regions, years, and classes (Lynch, 2017). In addition, this shows that BD can help a learner learn more, learn faster, and learn better based on their education. In addition, BD can help a learner learn better and help teachers and school leaders simplify their jobs. They help teachers review and evaluate courses by tracking student engagement and grades and even allow school and university leaders to simultaneously review and evaluate the performances of institutions and staff [1].

This shows that BD could overcome problems or issues faced in the education sector and realize a progressive vision. For example, BD and dynamic technological platforms are thought to be the beginning of a revolution in education that can replace the old-style classroom model and reorganize learning around technological platforms that let students choose their own pace, teaching methods, and subjects, realizing the progressive goal of self-initiated and interest-driven learning [66]. Two brand-new educational subfields have evolved in the last several years to tackle particular issues brought on by the BD phenomenon. These two approaches to BD in education are educational data mining and learning analytics, each taking a distinct stance. The goal of educational data mining is to examine and understand learning evidence from massive, noisy datasets. Keystrokes, log files, clickstream data, and discussion threads in natural language are a few examples. At the same time, structured data, especially with 'built-in' data models, tend to focus on learning analysis [67]. For example, Clayton and Halliday (2017) demonstrated that, in education, Big Data could predict a student's performance on upcoming exams based on their use (or non-use) of an online platform through which they accessed student learning.

If the knowledge of BD is to be integrated into education, then stakeholders need to develop a training program for teachers in the education sphere. However, analytics based on BD need machine-learning algorithms to maximiae data potential value. Based on Oussous et al. [68], the goal of machine learning is to discover knowledge and execute wise conclusions. Additionally, granular computing, deep learning, incremental and ensemble learning, and data-stream learning were pointed out in their study. However, deep learning is an important reference for improving modules with a BD knowledge module. This is because deep learning is more effective at solving data analyses and learning issues in complex, large-scale datasets. For example, the analysis of large data volumes, data tagging, semantic indexing, information retrieval, and discriminative tasks such as classification and prediction can all be made simpler using deep learning. It is also based on hierarchical learning and the extraction of various levels of complex data abstractions [68].

In summary, the concept of BD is linked to Industry 4.0 technology, which changes how students learn and teachers teach because online learning systems that share our lives on social media, smartphones, and the web are taking over the educational landscape. Education data mining and learning analytics are BD approaches. Big Data can predict student exam performances based on their use of online learning platforms. Deep learning is better at handling complex, large-scale datasets.

4. Big Data in Education

In the present, Big Data (BD) technology is practical. It has stretched into different areas, including social security, finance, smart medical treatment, e-commerce, transportation, and education, providing services and assistance in people's daily lives, enterprise operations, and management [38,69]. BD technology has accelerated in recent years in the business world, but it has also entered the educational system [70,71]. Alternatively, Ruiz-Palmero et al. [72] claimed that BD technology transforms decision making across a variety of industries, including education. Since it offers information for improving and regulating the Malaysian educational system, studying BD in education is valuable. According to Talib and Khalid [10], BD in education benefits the educational system. This point of view is consistent with their claim.

Education Big Data, a subset of BD, is a collection of data produced as educational activities advance and is generated in the field of education. Education Big Data primarily comprises teaching by teachers and learning by students [73]. Educational data can be obtained from various sources in many educational settings, including student records, distinct learning management systems or traditional classrooms, student learning behaviors, teacher classroom management, exam performances, social forums, teaching effectiveness data, and others [47,74]. Schools, colleges, and universities store a large amount of information about faculty, teachers, and students that can be analyzed to gain information and improve organizational operations in education. For example, BD in education gather and assess student information, including attendance, grades, test scores, and disciplinary issues [10].

According to Cui et al. [74], BD's primary uses in teaching and education are for educational purposes. BD allow teachers to instantaneously identify the learning needs of all students, change their role from knowledge providers to learning and teaching designers, assess the effectiveness of their instruction from a variety of viewpoints, and implement personalized learning. Moreover, BD allow teachers to identify future issues and forecast learning and teaching success. Secondly, education Big Data can dismantle conventional, experience-based teaching management strategies and carry out quantitative management decision making, which is crucial for school management and teaching decision making, as well as provide data assistance.

Figure 5 illustrates how Zheng et al. [75] viewed learner data, learning resource data, and learning behavior data as important types of education Big Data. The term 'learner data' refers to a learner's personal data, for instance, their educational history, age, gender, and place of origin. The term 'learning resource data' often entails data about the learning platform and the learning content, including the strengths of the faculty, information on a platform's features, and the rate at which resources are updated. Finally, data on learning behavior are information that students create as they go through the learning process. Examples include learning time, the number of clicks, activities, learning progress, etc.

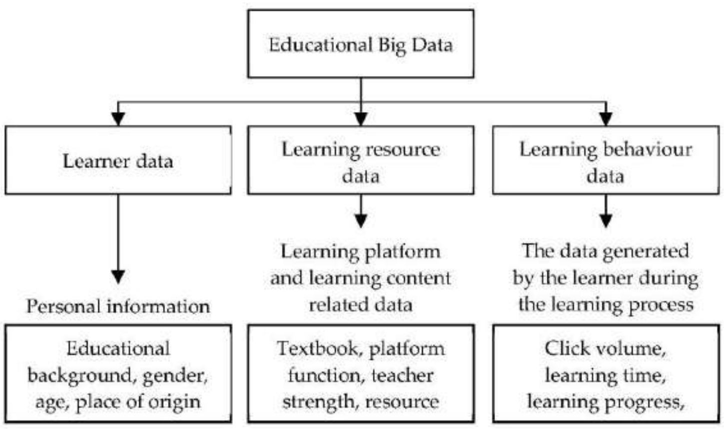

Figure 5. Research on learning analysis technology in education Big Data [75].

Since education Big Data play an important role in increasing the quality of teaching management, knowledge about BD and skills are needed to cultivate and train students' BD thinking and application capabilities. The transmission, acquisition, storage, visualization, comprehension, analysis, and application of data are fundamental requirements for BD application capability [35]. Among these, data analysis skills are the most important [76].

1. A basic data capability includes an understanding of BD. It covers the capacity to extract data under organizational requirements, analyze data dialectically, and use data as the foundation for decision making. It also includes knowing specific concepts, such as data life cycle, data format, and data type. This is a crucial component of information or data literacy.
2. BD acquisition refers to identifying a data source's location, nature, and format in various situations. These scenarios may entail a variety of topics, including physics, materials, electronics, etc.
3. Raw data must be processed using techniques including data auditing, conversion, reduction, cleaning, desensitization, integration, and labeling to better represent the BD value, verify the data accuracy, and increase the efficiency of data processing. These data-processing activities introduce a processor's design and thought, concentrate the value of BD's sparse distribution, and realize the data value-adding process [77].
4. The outcomes of the data analysis need to be controlled before being used for long-term storage and further data analysis applications.
5. BD analysis differs greatly from conventional data analysis and is based on open-source software. Finding useful information from enormous datasets requires conducting accurate statistical analyses of the data, which is one of the key characteristics of BD [78].

Data visualization is a significant way to show the value of BD. Its main goals are to describe BD's general attributes effectively and to clarify in detail why they are valuable. In addition, visualization can offer visualized expression tools for extended BD applications and assist people in understanding data and making decisions. The BD revolution's main advantages are frequently seen to be the extraction of useful knowledge and practical structure from data [79,80].

In conclusion, BD capabilities span a variety of sophisticated content and have varying layers of capability goals [35]. Without analysis, information is merely text, but analysis ablility can help us. Therefore, primary students can only learn to transfer knowledge from BD to real-life applications and apply basic BD knowledge to solve some practical problems in daily life, which is the primary objective of Big Data education. To initiate education Big Data, relevant technical schools and universities recommend creating a new curriculum focusing on four essential elements of the BD sector, which inlcude volume, velocity, variety, and veracity. These elements also serve as the basis for choosing all core courses. The relationships between data capabilities and relevant courses are displayed in Table 1 [35].

Table 1. The correspondence between data capabilities and related courses [35].

Components	Capabilities	Related Courses
Data resource layer	Acquisition and exchange of raw data	Information resource management, software architecture, computer interface technology
Basic capability layer	Data storage, data processing, and databases	Programming language, program design, data structure, database principles
Analysis and display layer	AI, statistical applications, DM, ML, visualization, and data analysis	Distributed systems, AI, applied statistics, multivariate statistical machine learning and data mining, data visualization, cloud computing, and BDA
Application layer	Data application combined with major characteristics and university background	Non-computer core major courses

Gafarov and Galimyanov [81] described the features of BD in education with the 5V rule as follows. 1V (volume) represents a substantial amount of physical data. For instance, all exam results might be found in a single database. 2V (velocity) measures how quickly data are gathered and results are processed. For instance, grades are stored in a database regularly. 3V (variety) stands for a variety of processing methods for various types of findings. For instance, grades for pupils in a school may be shown in a variety of ways (by age, teacher, gender, etc.) 4V (veracity) is a highly reliable data collection that makes it possible to create representational outcomes, for instance, exam evaluations and test outcomes. The 5V (value) principle states that the collected facts are valuable if they may be used to develop practical, varied parts of the educational system, for instance, there is correlation between classes.

A successful school is where students realize their distinctiveness and reach their greatest potential. As Khan Academy has demonstrated to educators over the past 15 years, integrating BD into education is becoming increasingly important for upcoming schools [71]. Nevertheless, it is crucial to draw attention to the issues that result from teachers' lack of expertise in handling and interpreting BD. Most teachers do not possess this operational knowledge. As a result, it influences student performance on international mathematics tests, including the Program for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) [1].

According to Wulandari et al. [82], the teaching of mathematics teachers plays an important role in ensuring that a pupil masters mathematics skills. Following a successful career in education, advisors are teachers who share their expertise and knowledge with their peers by creating permanent training programs that address the various needs of teaching staff [72]. If knowledge of BD is to be incorporated into education, then stakeholders need to develop a training program for teachers in the field of education.

To maximize the effectiveness of a program or lesson, much data must be collected and analyzed carefully. This is because the analyzed data can guide certain parties to choose a program with the highest effectiveness. According to Pratsri and Nilsook [83], there are three general levels of data that are relevant to the educational context and can be used, namely micro-level (click-stream data), meso-level (text data) and macro-level (institutional data).

1. Micro-level BD are fine-grained interaction data that collect individual data from possibly millions of students. When students interact with their learning environments, which can include massive open online courses (MOOCs), intelligent tutoring systems, games, and simulations, the majority of microlevel data are automatically collected.
2. In addition to course assignments, online discussion forum participation, intelligent tutoring systems, and social media interactions, meso-level BD also include computerized student writing artefacts that are routinely gathered during writing activities in a variety of learning environments. Notably, meso-level data offer chances to inadvertently record unprocessed information about learner advancements in social and cognitive skills, as well as affective states.
3. Institutional-level data collection constitutes macro-level BD. Macro-level data examples include information on campus services, student demographics and admittance, college major requirements, course enrollment and class schedules, and degree completion. Macro-level statistics are typically gathered over long periods of time. Still, they are rarely updated, frequently only once or twice every term (for instance, grade records and course schedule information) [83].

Apart from these, BD are now considered a 'human behaviour dashboard' as per Rick Smolan and Jennifer Erwitt, the primary authors of the illustrated book named *The Human Face of Big Data*. BD methods and analysis contribute to seeing human behavior to measure and analyze data flow constantly, which is captured by sensors, satellites, and GPS-enabled devices. Currently, there are many innovations in technology. BD implementation was performed for the first time with Google's web search algorithm to show user search behavior on Google [84]. The company Netflix has changed how people vote and consume movies and television through a recommendation engine implemented by BD. In the education and training domains, BD can be deemed a novel opportunity to adapt education to our needs and the learning processes of participants. Using BD in education is no longer fiction and has been carried out in several institutions. With the help of stress-analyzing software, students know the progress of their learning. This software collects data from students' devices and then gathers grades, weak and strong points, study skills, and hesitation patterns when using a computer mouse. BD implementation can see the potential for customized learning. It also affects teaching, giving teaching staff more time to support individual students and being able to see the needs of students by creating massive actual data.

There is a need to consider ethical and moral dimensions, especially developing anonymization of solid data. There are several issues of concern in BD implementation in education and training:

1. Data privacy: Privacy is how one has control over the extent, time, and circumstances of self-sharing (physical, behavioral, or intellectual) with other people. Privacy is also called information privacy, namely individual or organizational abilities to evaluate shareable data in a computer system.
2. Confidentiality: This is related to the treatment of information expressed by someone in a relationship of trust with the hope that it is not inconsistently disclosed without an understanding of the original disclosure without permission.
3. Data security: This is closely related to how securely the data is stored and used so that it cannot be accessed illegally. It includes standards that can be followed to obtain proper access to the appropriate data.
4. Security breach: This is related to follow-up if theft or other unauthorized access has occurred for data with sensitive personal information, leading to a possible breach in data confidentiality.

According to Siyamta et al. [85], for supporting the performance of BD in education, one program that can be used Hadoop. Hadoop is a Java-based, open-source framework licensed by Apache for supporting applications running on BD. Hadoop was created to cover the gaps in a database using a traditional approach. In addition, Hadoop supports the distributed processing of data to clusters on a computer. To support such a performance, Hadoop is supported by MapReduce and Hadoop distributed file system (HDFS) components. The HDFS can be defined as a distributed file storage system that splits large files into smaller files and then distributes them to clusters on computers. In conclusion, BD application in education has led to various positive results, such as tracking the rate of students who drop out of learning, identifying student performance, helping and evaluating the current teaching and learning system, and so on. Hence, it is necessary to expose the knowledge of BD to teachers by providing a teaching module to increase knowledge about BD among instructors and increase student achievement in mathematics by exposing them to BD knowledge.

5. Big Data in Malaysian Education

Big Data (BD) is the eighth out of nine main pillars of the Ministry of Education's (MOE's, Cyberjaya, Selangor, Malaysia) Malaysia in Digital Education Strategy 2020–2025 [86]. Therefore, this shows that BD is now beginning to be discussed and is now among the National Education System's principal pillars. In addition, Ayob et al. [87] stated that the Malaysia Digital Economy Corporation (MDEC, Cyberjaya, Selangor, Malaysia), one of the government agencies under the Malaysian Communications and Multimedia Commission (MCMC, Cyberjaya, Selangor, Malaysia), has implemented several initiatives to develop the digital economy in the country. Among them is digital implementation in digital free-trade zone strategy initiative, terms of 4WRD industry basis, a BD analytics framework, the direction of national e-commerce, a national artificial intelligence (AI) framework, and a national object internet framework.

In addition, the Ministry of Higher Education Malaysia (MOHE, Cyberjaya, Selangor, Malaysia) has prepared a Strategic Plan for Higher Education to increase Malaysia's competitiveness to a more global level in facing Industry Revolution (IR, Cyberjaya, Selangor, Malaysia) 4.0 through high-quality, effective, relevant education and training systems, such as lifelong learning, blended learning, and flipped classroom [9]. This is in line with the view of Wahab, Muhammad, and Ismail (2020) in their report, which stated that the emergence of IR 4.0 had a big impact on education following the tendency of students today, who are more interested in learning using paragogic and cybergogy methods, the functional diversity of modern and application gadgets and tools, blended learning, WhatsApp, learning via skype, gamification, FaceTime, and Google Hangouts, as well as heutagogy processes.

The traditional teaching and learning processes only take place in a classroom and use books. Now, with the convenience of information technology and the internet, the lesson process can take place anywhere, without being limited by time and place, by just using a device. Furthermore, the development of digital education nowadays is not seen to be the main focus that needs to be emphasized, as the present generation is a digitally native generation exposed to digital devices and technology earlier in life. The experience of using digital devices since childhood allows this generation to familiarize themselves and master digital skills naturally, in line with the development of digital technology in countries [87].

Based on the information and communications technology (ICT) Transformation Plan for the MOE 2019–2023, the aspirations summarized in the Malaysia Education Blueprint (PPPM) set the target for the MOE by the end of 2025. The aspirations determined and identified with the stakeholders were as follows:

1. An accessible, inclusive workforce capable of supporting the strategic goals of the Malaysian government and the MOE;
2. A dynamic, high-performing, and innovative workforce to advance 21st-century education;
3. Leading the adoption of appropriate digital tools and technology to advance the educational mission;
4. A data-driven ministry with a strong culture of knowledge sharing;
5. A platform that connects the education community and allows collaboration and idea sharing;
6. Educational institutions that have high-performance, high-quality Internet connectivity;
7. Transitions in student educational planning that are efficient and effective.

According to the Ministry of Science, Technology and Innovation Strategic Plan (MOSTI) 2016–2020 [14], their views align with the circulation of IR 4.0. The plan stated that MOSTI was the driver and leader of the Science, Technology and Innovation (STI) agenda at the national level. Additionally, it implemented new programs, such as Big Data Analysis (BDA), a digital economy, the Internet of Things (IoT), biotechnology, and nanotechnology. In addition, it also stated that school or polytechnic-level education in Malaysia needs to lead to the formation of minds of innovation. In Syaidatus et al.'s (2019) past research about BD in education, BDA was utilized to surpass the typical BD method for collecting data and processing it in meaningful ways using cutting-edge analytical techniques. Advanced analytics methods are necessary for real-time data that are too complex and need to be constantly updated,. Data science, business intelligence, and business analytics are all used in this analysis. The extraction of knowledge and understanding from BD has given rise to a name for BDA. Rather than asking 'what,' we use the question 'why' to determine the definition of BDA. The ability to gather and process different types of structured, semi-structured, or unstructured data in real-time for analytics tasks is the commonality between BDA and BD

BDA can rate or categorize data based on how significant the information is and can be used to reveal underlying facts. This analysis is crucial for society and industry to solve erroneous data analyses, lower risks, and provide better outcomes through wise decision making. When deploying BDA for reusability, business, manageability, interoperability, security, and maturity are the primary issues that must be taken into account. BDA has the potential to foster innovation in education, particularly in higher education, by assisting an institution in improving upon its current administrative, academic, teaching, and learning processes. Descriptive, predictive, and prescriptive analytics are three types of analytics that BDA recognizes. Each of these analytics has its benefits and is connected to the others.

According to Talib and Kahlid [10], BD in education has further improved the education system [10,88]. According to Bhat and Ahmed [89], BD can provide comprehensive tools for researchers to manipulate and predict future learning and teaching data. In addition, the use of BD can birth new technologies in the education system. In the education of the revolution, these fourth IR industries include the use of automation, digitization, and AI, which can lead to new challenges with the discovery of information from Big Data (BD) by the field of data science. The IoT and BDA, including modeling and simulation (Simulation and Modeling), are also key elements in the IR 4.0 era.

In today's education, it can be seen that the analysis of student data has become an important phenomenon [10,90]. Data are used to aid student learning. They can be tracked through research on artificial intelligence in education (AIED) and intelligent tutoring systems (ITS) [70]. An increase in the use of this dataset can assist teachers in making decisions about student learning [10]. For example, cohorts can be used to predict student achievement and identify specific gaps, which can be addressed to accelerate the learning process further as it changes patterns and trends in the behavior of individual learning of students [10]. Furthermore, the usage of public data in the education system can identify new strategies to further improve the learning system in a better direction [91].

The digital revolution leading to the advancement of innovative technology is also a factor affecting educational change in the educational environment by mixing with smart technology, including learning management systems (LMS) such as Moodle, online chat, and many more. LMS are vital in a teaching model to further improve learning [92]. Several researchers have investigated the implementation of BD in education. BD in education is closely associated with educational data mining (EDM) and learning analytics (LAs). It is notable that methods, instruments, and research are used in EDM to obtain new aspects and patterns of BD sets. Statistical machine methods and database systems [10] are also employed to obtain meaningful information from a large repository of data related to one's learning activities in an educational setting [93]. EDM is utilized to gather data from academic records, online logs, and exam results. Following that, all the data are examined to reach conclusions. A deep knowledge discovery database (KDD) is known as a type of EDM and contains a different and useful information discovery area from a number of datasets collected.

In conclusion, all these efforts show that Malaysia is actively developing digital facilities and capabilities. All this effort also clearly shows that the results achieved by Malaysian students have improved from over time in TIMSS and PISA assessments. All these results prove that the Malaysian government is on the right track in developing the country through education.

6. Influence of Mathematics in Big Data Research

The study of mathematics teaches the brain to approach problems and make decisions logically and systematically. The very essence of mathematics encourages meaningful education and mental challenge. Therefore, mathematics is one of the crucial subjects in the development of humans. In line with the National Philosophy of Education, the standard school mathematics curriculum has been streamlined and restructured to ensure the relevance of the curriculum to global developments. This curriculum's restructuring allows for continuity to the next level. The restructuring this mathematics curriculum was based on the need to provide intellectual and mathematical skills to students from different backgrounds and abilities. With mathematical knowledge and skills, they can explore knowledge, adapt, revise, and innovate when managing changes and dealing with future challenges [94]. A person's level of intellectual development in logical reasoning, abstract thinking, spatial visualization, and analytical skills can be greatly enhanced by studying mathematics. Students can advance their reasoning, thinking, numeracy, and problem-solving abilities by understanding mathematical ideas and applications throughout mathematics lessons. Students now have the chance to express their creativity and enjoy the thrill of learning new things thanks to mathematics. Due to these opportunities, students are more motivated to learn and are more likely to pursue higher degrees of education in mathematics [95]. To face the changes in the world of technology, the content of mathematics learning in schools has experienced some changes based on the Malaysia Education Blueprint 2013–2025. According to the blueprint, the government has emphasized science, technology, engineering, and mathematics (STEM) learning in schools so that students are trained to identify and understand problems, apply the concepts learned, and transform them into solutions. STEM is expected to produce creative, innovative, and inventive students in line with the skills required in Industry Revolution (IR) 4.0 and the 21st century. In conlusion, the study of mathematics teaches the brain to approach problems and make decisions logically and systematically. Students can advance their reasoning, thinking, numeracy, and problem-solving abilities. To face the changes in the world of technology, the content of mathematics learning in schools has made some changes.

7. Training on Big Data

The mathematics curriculum was revised in 2017 and aimed to produce students with 21st-century skills through a new curriculum. The skills prioritized in the revised curriculum were study and innovation skills, life and career skills, and information and communications technology (ICT) skills based on the practice of noble values [96]. Between learning skills and innovation, problem-solving skills were an element emphasized in mathematics aimed at producing students who could compete globally. Accordingly, teachers need to plan lessons that allow students to master these skills by selecting an appropriate approach to imparting Big Data (BD) thinking skills on data-handling topics. A variety of teaching approaches can be used by mathematics teachers for imparting BD thinking skills to students. Teaching approaches such as simulation, contextual learning, problem-based learning, mastery learning, project-based learning, discovery inquiry, modular approach, and STEM approach bring many benefits to a teacher's teaching in the classroom [97]. Through these approaches, students can learn data-handling skills in a more fun, meaningful, and challenging environment and have the opportunity to build self-generalizations on mathematical concepts. Thus, a strong mastery of skills and a deep understanding of concepts can be formed among students.

Huang et al. [98] stated that utilizing BD technologies could help students develop their skills by implementing the principle of 'learning by doing.' However, for students to conform to applications and satisfy social needs, they must receive training in all areas, such as data processing, analysis, and handling, as well as BD-handling skills. Furthermore, they contended that teachers should assign students real-world knowledge problems to answer so that students could see the value of using BD to address issues. According to their investigation of the instructors and students in a program, the most significant issues with BD were:

1. Because of the course's extensive material, solid theoretical foundation, and comprehensiveness, students found it challenging to learn, and teachers found it challenging to instruct.
2. Without experimental material, students could not fully comprehend how BD processing worked.
3. With fewer class hours, it was challenging to maximize the educational impact in the allotted time. The majority of students believed that courses should include focused experiments, in addition to providing theory introduction and practices of BD.

Thus, the 'learning by doing' approach could motivate students to take the initiative to develop their communication and cooperation abilities. In order to encourage students to participate actively in this study, an experimental group teacher was assigned to create a teaching lesson on the topic of data handling using the 'student-centered' philosophy. The experimental class adopted a grouping method to enhance the efficacy of studying through teamwork ability [98]. Since BD involve a large dataset, they bring diversity to the online teaching environment in education. Consequently, the development of a school or university curriculum should fully utilize the benefits of the multimedia teaching mode and motivate teachers to innovate with curriculum teaching mode to improve the fundamentals of instruction [99]. For example, learning and teaching could happen on a learning platform. Through a learning platform, teachers can provide students with personalized instruction and provide the most recent research findings and knowledge about the curriculum to students to ensure optimal learning. Additionally, the BD environment put forward particular demands for the innovation of instructional approaches [99]. Contrarily, this led to the teaching of unrealistic material to students and interrupted the learning progress of students who lacked self-control.

For primary and secondary education in a foreign country, Park and Shin [100] proposed using block-based programming languages, such as Scratch, to teach K-12 children the fundamentals of BD and artificial intelligence (AI). Using a drag-and-drop method, young students could quickly encounter BD and AI. In addition, there are numerous programming environments that K-12 students can use to learn the fundamentals of BD and AI, as per Park and Shin [100] and other authors [101–104]. These settings enable basic BD and AI projects to be implemented by young students. For instance, students can:

1. Visualize national cabbage production data [100];
2. Perform sentiment analysis using Twitter data [105];
3. Develop a machine-learning model that can tell a dog from a cat (Carney et al., 2020);
4. Develop a game application deploying a model that is already trained to recognize the position of a face [104].

Students frequently utilize block-based programming languages, such as Scratch, to establish these types of programs. The Tooele extension was used in eight examples of applications in Park and Shin's paper from 2021. The first four (saving quiz results to a CSV file, top five movies in theaters, weather forecast, and COVID-19 sashboard) were associated with K-12 BD education. Meanwhile, the final four were associated with K-12 AI education (handwriting recognition, color classification, facial image classification, and object classification). The results of the study demonstrated the efficacy of the drag-and-drop method for teaching K-12 students about BD and AI. Additionally, Rao et al. [101] introduced 'Milo,' a block-based programming model that was practical for beginners learning data science and accessible to students lacking prior programming knowledge. Conversely, DeepScratch, a novel programming language extension of Scratch, was presented by Alturayeif et al. [106]. Concerning serving various age groups and educational levels, it offered two choices for implementing deep-learning models. They claimed that deep learning aimed to discover sophisticated relationships between data that produced useful outcomes and predicted across various applications.

Instead of using the traditional and formal teaching approach, Morrar et al. [12] provide an example of students being able to choose their educational approach and specify exactly what they needed from courses and knowledge. This allowed for incremental and radical innovation in the educational system, with a primary focus on project-based, research-based, and interactive learning. A few solutions have been developed to give children or high school students a clear and basic understanding of data science and machine learning. For instance, the Scratch Community Blocks system teaches kids how to reason through complicated information representations by allowing them to programmatically analyze and visualize data about their engagement in Scratch [107]. The block-based language Snap has an extension called DataSnap that allows users to retrieve and analyze data from web sources [108]. With the help of Scratch Nodes Machine Learning (ML), which Agassi et al. [109] introduced, students can sample gestures and validate classifiers by learning about data gathering, data labeling, and classifier testing ML processes. Then, using these classifiers as Scratch blocks, they can make their ML applications. Other children can utilize these gesture recognition Scratch blocks to make their own interesting experiences. They can work together and assume different roles, such as Scratch application developer, data collector, and classifier trainer.

Other than these, Lazić et al. [110] mentioned the impact of project-based learning on elementary school student performances in mathematics. In order to teach students how to solve issues, think creatively while working, organize their work using scientific principles, effectively collaborate with others, and evaluate their performances, project-based learning places a strong emphasis on practical experiences for students. Additionally, in project-based learning, the tasks are related to real-world situations. Under the guidance and supervision of teachers, students are given a chance to learn and act independently in terms of project ideas, action plans, designs, implementation, and presentation [111]. Additionally, students collaborate through project-based learning to work through difficult and real-world problems [112]. As a result, project-based learning can help students' problem-solving skills, creativity, internal drive and interest, responsibility, interpersonal communication, and social skills flourish [112]. The 4Cs—critical thinking, collaboration, communication, and creativity—should be integrated into the requirements for project-based learning in the twenty-first century. This means that the benefits of project-based learning should include the development of creativity, teamwork, and critical thinking [113].

Many 21st-century competencies, such as creativity, critical thinking, and problem solving, are aligned with computational thinking [114] (lines of problem-based learning), which enables students to learn about a subject. This can be achieved by presenting them with a variety of problems and asking them to build their understanding of a subject's object using these problems. According to the constructivist theory of knowledge, all information is modified as it is digested and integrated because it is filtered via the recipent's prior knowledge. In other words, knowledge is not simply received by students, but is created. Constructivism contends that learning occurs whenever events take place that are unexpected. The disequilibrium brought on by these events encourages adaptability. As a result, constructivist-based pedagogy must always consider prior knowledge and think about how to produce disequilibrium that results in adjustments to students' present knowledge [115]. Nine fundamental computational-thinking concepts and abilities were listed by Yadav et al. [116]. The authors also noted that, when students practiced process skills, such as collecting, presenting, and interpreting data from scientific experiments, they engaged in some of the computational-thinking processes associated with computational thinking. This was in line with the survey results of [117], which indicated that teachers and, by extension, students were already engaging in activities that promoted or had the potential to support the development of some aspects of ICT in the classroom. Unsurprisingly, data collection, analysis, and representation were placed with the most emphasis. The use of computational thinking in science and mathematics, according to Weintrop et al. [118], could be divided into four main categories: data practices, modeling and simulation practices, computational problem-solving practices, and system-thinking practices.

Project-based learning is one of the instructional strategies that might encourage computational-thinking abilities [119]. For teaching and learning processes to be effective, project-based learning teaches students how to master skills and their applications in real-world situations [120]. In order to extend computational-thinking skills to digital production, design, and personal expression, Ornelas [121] emphasized that combining computational thinking with project-based learning was crucial. Programming-oriented thinking also emphasises computational thinking, recognizing the relationship between logical and programming thinking while specifying what that link entails. In essence, the term above was developed by programming-oriented thinking, a branch of computational thinking that offers programming-related curriculum in ICT education [122].

In summary, BD's large datasets have diversified online education teaching. Computational thinking aligns with creativity, critical thinking, and problem-solving. Project-based learning emphasizes hands-on learning, and students learn to plan, design, implement, and present projects independently.

8. Conclusions

The period of Big Data (BD) has begun due to advancements in communication and information technology, carrying with it new possibilities and obstacles for the world. BD and the development of information technology, such as artificial intelligence (AI), tend to modify our world and how we can understand it. Meanwhile, with the popularization of online activities, such as increasingly frequent social-networking activities, the widespread use of smart devices, e-commerce, the rise of cloud computing, and Internet of Things (IoT) technology, data have started to evolve geometrically. Consequently, the world has entered an era that cannot be separated from BD. This situation clearly shows how important it is to integrate BD into education to produce a new generation that can develop Malaysia to compete globally with a mastery of BD. Hence, this article discussed Industry Revolution (IR) 4.0, BD, and BD in education, as well as the relationship between mathematics education, BD, and teaching approaches in Malaysia. Knowledge of BD can guide teachers and students in comprehending the usage of BD and the information they want to express through BD. Apart from that, BD in education can improve education management and provide an opportunity for individuals to know themselves more deeply and improve their abilities. In addition, BD can guide researchers to explore deeper into Malaysia's education sector.

Author Contributions: Conceptualization, A.H.A. and S.J.J.; formal analysis, A.H.A. and S.J.J.; writing—original draft preparation, S.J.J.; writing—review and editing, M.M.; visualization, U.H.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the UTM Fundamental Research Grant (Q.J130000.3813.21H85) and, in part, by the Ministry of Higher Education Malaysia and Universiti Teknologi Malaysia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Williamson, B. Who owns educational theory? Big data, algorithms and the expert power of education data science. *E-Learn. Digit. Media* **2017**, *14*, 105–122. [CrossRef]
2. Ramli, A.; Zain, R.M. The impact of facilities on students' academic achievement. *Sci. Int.* **2018**, *30*, 299–311.
3. Cherri, L.H. The Mathematics of Industry 4.0. Medium. Available online: https://medium.com/@luizcherri/the-mathematics-of-industry-4-0-81904d7bd3cb (accessed on 5 February 2022).
4. Torrecilla, J.L.; Romo, J. Data learning from big data. *Stat. Probab. Lett.* **2018**, *136*, 15–19. [CrossRef]
5. Zhou, K.; Liu, T.; Zhou, L. Industry 4.0: Towards Future Industrial Opportunities and Challenges. In Proceedings of the 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Zhangjiajie, China, 15–17 August 2015; pp. 2147–2153.
6. Bartodziej, C.J. *The Concept Industry 4.0*; Springer Gabler: Wiesbaden, Germany, 2017.
7. Bahrin, M.A.K.; Othman, M.F.; Azli, N.H.N.; Talib, M.F. Industry 4.0: A Review on Industrial Automation and Robotic. *J. Teknol.* **2016**, *78*, 137–143.
8. Berawi, M.A. Utilizing big data in industry 4.0: Managing competitive advantages and business ethics. *Int. J. Technol.* **2018**, *9*, 430–433. [CrossRef]
9. Noorashid, N.A. Revolusi industri 4.0: Impak terhadap perkembangan pendidikan tinggi di Malaysia. *J. Sci. Manag. Res.* **2019**, *5*, 55–67.
10. Talib, J.; Khalid, F. Peranan Data Raya Dalam Sistem Pendidikan: Satu Tinjauan [The Role of Big Data in the Education System: A Review]. *AL-QIYAM Int. Soc. Sci. Humanit. J.* **2020**, *3*, 23–35. Available online: http://al-qiyam.net/index.php/isshj/article/view/78 (accessed on 7 February 2022).
11. Economic Planning Unit. *Malaysian National Fourth Industrial Revolution (4IR) Policy*; Prime Ministrer's Department: Putrajaya, Malaysia, 2021.

12. Morrar, R.; Arman, H.; Mousa, S. The fourth industrial revolution (Industry 4.0): A social innovation perspective. *Technol. Innov. Manag. Rev.* **2017**, *7*, 12–20. [CrossRef]
13. AlphaBeta. *Positioning Malaysia as a Regional Leader in the Digital Economy: The Economic Opportunities of Digital Transformation and Google's Contribution*; AlphaBeta Strategy × Economics: Singapore, 2021. Available online: https://alphabeta.com/wp-content/uploads/2021/08/Malaysia-Digital-Transformation.pdf (accessed on 13 February 2022).
14. MOSTI. *Ministry of Science, Technology and Innovation Strategic Plan 2016–2020*; Bahagian Dasar dan Penyelarasan Strategik: Putrajaya, Malaysia, 2018. Available online: https://www.mosti.gov.my/wp-content/uploads/repository/penerbitan/2018/Pelan-Strategik-Kementerian-Sains-Teknologi-Dan-Inovasi-MOSTI-2016--2020.pdf (accessed on 22 February 2022).
15. Kowang, T.O.; Bakry, M.F.; Hee, O.C.; Fei, G.C.; Yew, L.K.; Saadon, M.S.I.; Long, C.S. Industry 4.0 Competencies among Lecturers of Higher Learning Institution in Malaysia. *Int. J. Eval. Res. Educ.* **2020**, *9*, 303–310. [CrossRef]
16. Li, T. The Use of Cognitive Diagnostic Modeling in the Assessment of Computational Thinking. *AERA Open* **2020**, *8*, 1–17. [CrossRef]
17. Ministry of Education Malaysia. *Ringkasan Eksekutif: Pelan Transformasi ICT Kementerian Pendidikan Malaysia 2019–2023*; Ministry of Education Malaysia: Putrajaya, Malaysia, 2019.
18. Lo, F. What Is Big Data. 2018. Available online: https://datajobs.com/what-is-big-data (accessed on 31 January 2022).
19. Sabaitytė, J.; Davidavicienė, V.; Karpoviciutė, R. Learning skills for enhancing the use of Big Data. *World J. Educ. Technol. Curr. Issues* **2020**, *12*, 23–36. [CrossRef]
20. Laney, D. 3D Data Management: Controlling DataVolume, Velocity, and Variety. Available online: http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf (accessed on 17 January 2022).
21. Emmanuel, I.; Stanier, C. Defining big data. In Proceedings of the International Conference on Big Data and Advanced Wireless Technologies—BDAW 2016, Blagoevgrad, Bulgaria, 10–11 November 2016. [CrossRef]
22. Sharma, S. Rise of Big Data and related issues. In Proceedings of the 2015 Annual IEEE India Conference (INDICON), New Delhi, India, 17–20 December 2015. [CrossRef]
23. Yin, S.; Kaynak, O. Big data for modern industry: Challenges and trends [point of view]. *Proc. IEEE* **2015**, *103*, 143–146. [CrossRef]
24. Groves, P.; Kayyali, B.; Knott, D.; Kuiken, S. *Van The 'Big Data' Revolution in Healthcare: Accelerating Value and Innovation*; McKinsey & Company and Center for US Health System Reform Business Technology Office: New York, NY, USA, 2013. Available online: https://www.mckinsey.com/~/media/mckinsey/industries/healthcare%20systems%20and%20services/our%20insights/the%20big%20data%20revolution%20in%20us%20health%20care/the_big_data_revolution_in_healthcare.pdf (accessed on 20 January 2022).
25. Gandomi, A.; Haider, M. Beyond the hype: Big data concepts, methods, and analytics. *Int. J. Inf. Manag.* **2015**, *35*, 137–144. [CrossRef]
26. Baig, M.I.; Shuib, L.; Yadegaridehkordi, E. Big data in education: A state of the art, limitations, and future research directions. *Int. J. Educ. Technol. High. Educ.* **2020**, *17*, 44. [CrossRef]
27. Memon, M.A.; Soomro, S.; Jumani, A.K.; Kartio, M.A. Big Data Analytics and Its Applications. *Ann. Emerg. Technol. Comput.* **2017**, *1*, 45–54. Available online: https://scholar.archive.org/work/dgdxq26lvrgflk7ylkgpnpb33m/access/wayback/http://aetic.theiaer.org:80/archive/v1n1/P6.pdf (accessed on 13 February 2022). [CrossRef]
28. Anshari, M.; Alas, Y.; Sabtu, N.I.; Yunus, N. A survey study of smartphones behavior in Brunei: A proposal of modelling big data strategies. *Int. J. Cyber Behav. Psychol. Learn.* **2016**, *6*, 60–72. [CrossRef]
29. Rivas-Rebaque, B.; Gértrudix-Barrio, F.; de Britto, J.C. La percepción del docente universitario ante el uso y valor de los datos abiertos. *Educación* **2019**, *22*, 141–163. [CrossRef]
30. Paek, S.; Um, T.; Kim, N. Exploring latent topics and international research trends in competency-based education using topic modeling. *Educ. Sci.* **2021**, *11*, 303. [CrossRef]
31. Nobanee, H.; Dilshad, M.N.; Al Dhanhani, M.; Al Neyadi, M.; Al Qubaisi, S.; Al Shamsi, S. Big data applications the banking sector: A bibliometric analysis approach. *SAGE Open* **2021**, *11*, 21582440211067230. [CrossRef]
32. Yang, J.; Chen, T.; Zheng, L.S. *The Influence of the Big Data Age on the Teaching of Statistics*; Education Teaching Forum: Shijiazhuang, China, 2017.
33. Li, Y.; Huang, C.; Zhou, L. Impacts on statistics education in big data era. *Educ. Sci. Theory Pract.* **2018**, *18*, 1236–1245. [CrossRef]
34. Soh, H.M.; Abdullah, A.H.; Mokhtar, M.; Ali, D.F.; Jumaat, N.F.; Ashari, Z.M.; Samah, N.A.; Rahman, K.A.A. Active learning using digital smart board to enhance primary school students' learning. *Int. J. Interact. Mobile Technol.* **2019**, *13*, 4–16. [CrossRef]
35. Li, X.; Fan, X.; Qu, X.; Sun, G.; Yang, C.; Zuo, B.; Liao, Z. Curriculum reform in big data education at applied technical colleges and universities in China. *IEEE Acces* **2019**, *7*, 125511–125521. [CrossRef]
36. Su, G. Study on the effect of big data on ideological and political education and its application strategy. In Proceedings of the 2018 International Conference on Education, Psychology, and Management Science (ICEPMS 2018), Shanghai, China, 13–14 October 2018. [CrossRef]
37. Liang, Q.; Hainan, N.C. Adaptive learning model and implementation based on big data. In Proceedings of the 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, 25–28 May 2019. [CrossRef]
38. Ma, H.; Lang, C.; Liu, Y.; Gao, Y. Constructing a hierarchical framework for assessing the application of big data technology in entrepreneurship education. *Front. Psychol.* **2020**, *11*, 551389. [CrossRef] [PubMed]

39. Faroukhi, A.Z.; El Alaoui, I.; Gahi, Y.; Amine, A. An adaptable big data value chain framework for end-to-end big data monetization. *Big Data Cogn. Comput.* **2020**, *4*, 34. [CrossRef]
40. Shan, Z. Hold on to the Opening and Sharing. People's Daily. Available online: http://theory.people.com.cn/n/2015/1120/c40531-27835936.html (accessed on 20 January 2022).
41. Xiaohui, Z. Construction of personalized English teaching model driven by big data. In Proceedings of the 2019 International Conference on Arts, Management, Education and Innovation (ICAMEI 2019), Seoul, Korea, 12–14 January 2019. [CrossRef]
42. Rabhi, L.; Falih, N.; Afraites, A.; Bouikhalene, B. Big data approach and its applications in various fields. *Procedia Comput. Sci.* **2019**, *155*, 599–605. [CrossRef]
43. Lee, I. Big data: Dimensions, evolution, impacts, and challenges. *Bus. Horiz.* **2017**, *60*, 293–303. [CrossRef]
44. Goasduff, L. Gartner Says Business Intelligence and Analytics Leaders Must Focus on Mindsets and Culture to Kick Start Advanced Analytics. Gartner. Available online: https://www.gartner.com/en/newsroom/press-releases/2015-09-15-gartner-says-business-intelligence-and-analytics-leaders-must-focus-on-mindsets-and-culture-to-kick-start-advanced-analytics (accessed on 24 August 2022).
45. Barwick, H. The 'Four Vs' of Big Data. Implementing Information Infrastructure Symposium. 2012. Available online: http://www.computerworld.com.au/article/396198/iiis_four_vs_big_data/ (accessed on 25 January 2022).
46. Hilbert, M. The Theory, Practice and Limits of Big Data for the Social Scciences. 2015. Available online: http://www.martinhilbert.net/the-theory-practice-and-limits-of-big-data-for-the-social-sciences/ (accessed on 2 February 2022).
47. Khan, S.; Alqahtani, S. Big data application and its impact on education. *Int. J. Emerg. Technol. Learn.* **2020**, *15*, 36–46. [CrossRef]
48. Uprichard, E. Focus: Big data, Little Questions? *Discover Society.* Available online: https://archive.discoversociety.org/2013/10/01/focus-big-data-little-questions/ (accessed on 20 February 2022).
49. Lupton, D. The Thirteen Ps of Big Data. This Sociological Life. Available online: https://www.researchgate.net/profile/Deborah-Lupton/publication/276207564_The_Thirteen_Ps_of_Big_Data/links/5552c2d808ae6fd2d81d5f20/The-Thirteen-Ps-of-Big-Data.pdf (accessed on 2 March 2022).
50. Murthy, P.; Bharadwaj, A.; Subrahmanyam, P.A.; Roy, A.; Rajan, S. Big Data Taxonomy. Big Data Working Group, Cloud Security Alliance. 2014. Available online: https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Big_Data_Taxonomy.pdf (accessed on 23 January 2022).
51. Rajaraman, V. Big Data Analytics. *Resonance* **2016**, *21*, 695–716. Available online: https://www.ias.ac.in/public/Volumes/reso/021/08/0695-0716.pdf (accessed on 27 January 2022). [CrossRef]
52. Cavanillas, J.M.; Curry, E.; Wahlster, W. (Eds.) *New Horizons for a Data-Driven Economy: A Roadmap for Usage and Exploitation of Big Data in Europe*; Springer Nature: Saarbrücken, Germany, 2016. [CrossRef]
53. Sivarajah, U.; Kamal, M.M.; Irani, Z.; Weerakkody, V. Critical analysis of Big Data challenges and analytical methods. *J. Bus. Res.* **2017**, *70*, 263–286. [CrossRef]
54. Ibrahim, A.; Ibrar, Y.; Nor, B.A.; Salimah, M.; Abdullah, G.; Samee, U.K. The rise of "big data" on cloud computing: Review and open research issues. *Inf. Syst.* **2015**, *47*, 98–115. [CrossRef]
55. Jiang, F.; Wang, S. Research on the new changes and countermeasures of human resources management under big data and artificial intelligence. In Proceedings of the 2018 International Conference on Education, Psychology, and Management Science (ICEPMS 2018), Shanghai, China, 13–14 October 2018. [CrossRef]
56. Mayer-Schönberger, V.; Cukier, K. *Aprender Con Big Data [Learn with Big Data]*; Turner: Madrid, Spain, 2018.
57. Leung, C.K.; Chen, Y.; Shang, S.; Deng, D. Big data science on COVID-19 data. In Proceedings of the 2020 IEEE 14th International Conference on Big Data Science and Engineering (BigDataSE), Guangzhou, China, 31 December 2020–1 Janary 2021. [CrossRef]
58. Khan, N.; Yaqoob, I.; Hashem, I.A.T.; Inayat, Z.; Ali, W.K.M.; Alam, M.; Shiraz, M.; Gani, A. Big data: Survey, technologies, opportunities, and challenges. *Sci. World J.* **2014**, *2014*, 712826. [CrossRef] [PubMed]
59. Alharthi, A.; Krotov, V.; Bowman, M. Addressing barriers to big data. *Bus. Horiz.* **2017**, *60*, 285–292. [CrossRef]
60. Rai, P.K.; Lee, J.; Kailasa, S.K.; Kwon, E.E.; Tsang, Y.F.; Ok, Y.S.; Kim, K.H. A Critical Review of Ferrate (VI)-Based Remediation of Soil and Groundwater. *Environ. Res.* **2018**, *160*, 420–428. [CrossRef]
61. Russom, P. *Big Data Analytics*; TDWI: Renton, WA, USA, 2011.
62. Tanwar, M.; Duggal, R.; Khatri, S.K. Unravelling unstructured data: A wealth of information in big data. In Proceedings of the 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO), Noida, India, 2–4 September 2015. [CrossRef]
63. El Alaoui, I.; Gahi, Y.; Messoussi, R. Full consideration of Big Data characteristics in sentiment analysis context. In Proceedings of the 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, China, 12–15 April 2019. [CrossRef]
64. Moro Visconti, R.; Morea, D. Big data for the sustainability of healthcare project financing. *Sustainability* **2019**, *11*, 3748. [CrossRef]
65. Sandhu, A.K. Big data with cloud computing: Discussions and challenges. *Big Data Min. Anal.* **2021**, *5*, 32–40. [CrossRef]
66. Dishon, G. New data, old tensions: Big data, personalized learning, and the challenges of progressive education. *Theory Res. Educ.* **2017**, *15*, 272–289. [CrossRef]
67. Cope, B.; Kalantzis, M. Big data comes to school: Implications for learning, assessment, and research. *AERA Open* **2016**, *2*, 2332858416641907. [CrossRef]

68. Oussous, A.; Benjelloun, F.Z.; Lahcen, A.A.; Belfkih, S. Big data technologies: A survey. *J. King Saud Univ. Inf. Sci.* **2017**, *30*, 431–448. [CrossRef]
69. Moreno-Carriles, R.M. Big data; But what is it? *Angiologia* **2018**, *70*, 191–194. [CrossRef]
70. Daniel, B.K. Big Data and data science: A critical review of issues for educational research. *Br. J. Educ. Technol.* **2019**, *50*, 101–113. [CrossRef]
71. Matas-Terrón, A.; Leiva-Olivencia, J.J.; Negro-Martínez, C. Tendency to use big data in education based on its opportunities according to andalusian education students. *Soc. Sci.* **2020**, *9*, 164. [CrossRef]
72. Ruiz-Palmero, J.; Colomo-Magaña, E.; Ríos-Ariza, J.M.; Gómez-García, M. Big data in education: Perception of training advisors on its use in the educational system. *Soc. Sci.* **2020**, *9*, 53. [CrossRef]
73. Branch, R.M.; Kopcha, T.J. Instructional design models. In *Handbook of Research on Educational Communications and Technology*; Springer: New York, NY, USA, 2014; pp. 77–87.
74. Cui, L.; Liao, Y.; Wang, Y.; Dong, X. College Mathematics Teaching Method Based on Big Data. *Int. J. Emerg. Technol. Learn.* **2019**, *14*, 47–58. Available online: https://www.learntechlib.org/p/210555/ (accessed on 31 January 2022). [CrossRef]
75. Zheng, S.Y.; Jiang, S.P.; Yue, X.G.; Pu, R.; Li, B.Q. Application Research of an Innovative Online Education Model in Big Data Environment. *Int. J. Emerg. Technol. Learn.* **2019**, *14*, 125–138. Available online: https://www.learntechlib.org/p/209718/ (accessed on 22 February 2022). [CrossRef]
76. Zhenchong, C.; Tiantian, H. Data science:The demand and development of talents. *Big Data Res.* **2016**, *2*, 95–106.
77. Hao, X.; Yue, Q.; Lan, H. Liberal-education-oriented DataScience curriculum construction. *Comput. Educ.* **2016**, *14*, 158–162.
78. Maurer, K.T. Applications of Technology and Large Data in Statistics Education and Statistical Graphics. Ph.D. Dissertation, Iowa State University, Ames, IA, USA, 2015. [CrossRef]
79. Mayer-Schönberger, V.; Cukier, K. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*; Houghton Mifflin Harcourt: Boston, MA, USA, 2013.
80. Jagadish, H.V.; Gehrke, J.; Labrinidis, A.; Papakonstantinou, Y.; Patel, J.M.; Ramakrishnan, R.; Shahabi, C. Big data and its technical challenges. *Commun. ACM* **2014**, *57*, 86–94. [CrossRef]
81. Gafarov, F.M.; Galimyanov, A.F. Identification of key factors in the formation of an individual trajectory of teacher professional development in digital environment based on big data. *ARPHA Proc.* **2020**, *3*, 625.
82. Wulandari, M.A.; Arga, P.H.S.; Kelana, J.B.; Altaftazani, D.H.; Ruqoyyah, S. Analisis pembelajaran "daring" pada guru sekolah dasar di era Covid-19. *Jurnal Ilmiah* **2020**, *7*, 164–168. [CrossRef]
83. Pratsri, S.; Nilsook, P. Design on big data platform-based in higher education institute. *High. Educ. Stud.* **2020**, *10*, 36–43. [CrossRef]
84. Abdullah, A.H.; Zakaria, E. The activities based on Van Hiele's phase-based learning: Experts' and pre-service teachers' views. *J. Math. Statist.* **2012**, *8*, 385–395. [CrossRef]
85. Tahir, L.M.; Musaha, M.B.; Abdullah, A.H.; Musta'amal, A.H.; Abdullah, M.H.A. Technical college teachers sharing their knowledge: Does leadership, institutional factors or barriers predict their practices? *Educ. Stud.* **2016**, *42*, 465–492. [CrossRef]
86. Ministry of Education Malaysia. *Strategi Pendidikan Digital 2020–2025*; Bahagian Sumber dan Teknologi Pendidikan (BSTP): Putrajaya, Malaysia, 2020.
87. Ayob, N.H.; Hamzah, I.S.; Aziz, M.A. Merapatkan jurang digital dalam pendidikan: Dasar dan strategi di Malaysia. *J. Tour. Hosp. Environ. Manag.* **2021**, *6*, 157–170. [CrossRef]
88. Meghji, J.; Lesosky, M.; Joekes, E.; Banda, P.; Rylance, J.; Gordon, S.; Jacob, J.; Zonderland, H.; MacPherson, P.; Corbett, E.L.; et al. Patient outcomes associated with post-tuberculosis lung damage in Malawi: A prospective cohort study. *Thorax* **2020**, *75*, 269–278. [CrossRef] [PubMed]
89. Bhat, A.Z.; Ahmed, I. Big data for institutional planning, decision support and academic excellence. In Proceedings of the 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC), Muscat, Oman, 15–16 March 2016. [CrossRef]
90. Lodge, J.M.; Corrin, L. *The Connected Community of Practice in Educational Technology: A Model for Future Networked Professional Development*; Implementing Communities of Practice in Higher Education Singapore; Springer: Singapore, 2016; pp. 573–595. [CrossRef]
91. Mor, Y.; Ferguson, R.; Wasson, B. Editorial: Learning design, teacher inquiry into student learning and learning analytics: A call for action: Learning design, TISL and learning analytics. *Br. J. Educ. Technol.* **2015**, *46*, 221–229. [CrossRef]
92. Cantabella, M.; López, B.; Caballero, A.; Muñoz, A. Analysis and evaluation of lecturers' activity in Learning Management Systems: Subjective and objective perceptions. *Interact. Learn. Environ.* **2018**, *26*, 911–923. [CrossRef]
93. Aldowah, H.; Al-Samarraie, H.; Fauzy, W.M. Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telemat. Inform.* **2019**, *37*, 13–49. [CrossRef]
94. Ministry of Education Malaysia. *Pelan Pembangunan Pendidikan Malaysai 2013–2025: Laporan Tahunan 2017*; Kementerian Pendidikan Malaysia: Putrajaya, Malaysia, 2018.
95. Ministry of Education Malaysia. *Analisis Prestasi Murid Matematik TIMSS 2015*; Educational Policy Planning and Research Division (EPRD): Putrajaya, Malaysia, 2017.
96. Ministry of Education Malaysia. *Standard Document of Curriculum and Assessment for Form Two Mathematics Curriculum*; Bahagian Pembanguna Kurikulum: Putrajaya, Malaysia, 2016.

97. Misrom, N.S.; Abdurrahman, M.S.; Abdullah, A.H.; Osman, S.; Hamzah, M.H.; Fauzan, A. Enhancing students' higher-order thinking skills (HOTS) through an inductive reasoning strategy using geogebra. *Int. J. Emerg. Technol. Learn.* **2020**, *15*, 156–179. [CrossRef]
98. Huang, X.; Qin, N.; Zhang, X.; Wang, F. Experimental teaching design and practice on big data course. In Proceedings of the 2017 12th International Conference on Computer Science and Education (ICCSE), Houston, TX, USA, 22–25 August 2017. [CrossRef]
99. Zhang, D. Analysis on the application of big data in classroom teaching. In Proceedings of the 2018 International Conference on Education, Psychology, and Management Science (ICEPMS 2018), Shanghai, China, 13–14 October 2018. [CrossRef]
100. Park, Y.; Shin, Y. Tooee: A novel scratch extension for K-12 big data and artificial intelligence education using text-based visual blocks. *IEEE Access* **2021**, *9*, 149630–149646. [CrossRef]
101. Rao, A.; Bihani, A.; Nair, M. Milo: A visual programming environment for data science education. In Proceedings of the 2018 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), Lisbon, Portugal, 1–4 October 2018. [CrossRef]
102. Williams, R.; Park, H.W.; Oh, L.; Breazeal, C. Popbots: Designing an artificial intelligence curriculum for early childhood education. In Proceedings of the AAAI Conference on Artificial Intelligence, Atlanta, GA, USA, 8–12 October 2019. [CrossRef]
103. Lee, J.; Lee, S. A study on experts' perception survey on elementary AI education platform. *J. Korean Assoc. Inf. Educ.* **2020**, *24*, 483–494. [CrossRef]
104. Lane, D. *Machine Learning for Kids: A Project-Based Introduction to Artificial Intelligence*; Starch Press: San Francisco, CA, USA, 2021.
105. Druga, S. Growing up with AI: Cognimates: From Coding to Teaching Machines. Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 2018.
106. Alturayeif, N.; Alturaief, N.; Alhathloul, Z. DeepScratch: Scratch programming language extension for deep learning education. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 642–650. [CrossRef]
107. Dasgupta, S.; Hill, B.M. Scratch community blocks: Supporting children as data scientists. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017. [CrossRef]
108. Hellmann, J.D. DataSnap: Enabling Domain Experts and Introductory Programmers to Process Big Data in a Block-Based Programming Language. Ph.D. Dissertation, Virginia Tech, Blacksburg, VA, USA, 2015.
109. Agassi, A.; Erel, H.; Wald, I.Y.; Zuckerman, O. Scratch nodes ML: A playful system for children to create gesture recognition classifiers. In Proceedings of the Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow Scotland, UK, 4–9 May 2019. [CrossRef]
110. Lazic, B.; Knežević, J.; Maričić, S. The influence of project-based learning on student achievement in elementary mathematics education. *S. Afr. J. Educ.* **2021**, *41*, 1–10. [CrossRef]
111. Nilsook, P.; Chatwattana, P.; Seechaliao, T. The project-based learning management process for vocational and technical education. *High. Educ. Stud.* **2021**, *11*, 20–29. [CrossRef]
112. Shin, M.H. Effects of project-based learning on students' motivation and self-efficacy. *Eng. Teach.* **2018**, *73*, 95–114. [CrossRef]
113. Dag, F.; Durdu, L. Pre-service teachers' experiences and views on project-based learning processes. *Int. Educ. Stud.* **2017**, *10*, 18–39. [CrossRef]
114. Lye, S.Y.; Koh, J.H.L. Review on teaching and learning of computational thinking through programming: What is next for K-12? *Comput. Human Behav.* **2014**, *41*, 51–61. [CrossRef]
115. Rich, K.M.; Binkowski, T.A.; Strickland, C.; Franklin, D. Decomposition: A K-8 computational thinking learning trajectory. In Proceedings of the 2018 ACM Conference on International Computing Education Research, Espoo, Finland, 13–15 August 2018. [CrossRef]
116. Yadav, A.; Gretter, S.; Good, J.; McLean, T. Computational thinking in teacher education. In *Emerging Research, Practice, and Policy on Computational Thinking*; Rich, P.J., Hodges, C.B., Eds.; Springer: Cham, Switzerland, 2017; pp. 205–220.
117. Mannila, L.; Dagiene, V.; Demo, B.; Grgurina, N.; Mirolo, C.; Rolandsson, L.; Settle, A. Computational thinking in K-9 education. In Proceedings of the ITICSE'14: Innovation and Technology in Computer Science Education Conference 2014, Uppsala, Sweden, 23–25 June 2014; pp. 1–24.
118. Weintrop, D.; Beheshti, E.; Horn, M.; Orton, K.; Jona, K.; Trouille, L.; Wilensky, U. Defining computational thinking for mathematics and science classrooms. *J. Sci. Educ. Technol.* **2016**, *25*, 127–147. [CrossRef]
119. Hadi, M.E.; Atiqoh, K.S.N. Improving students' mathematical computational thinking using scratch program through based learning: A development research during pandemic COVID-19. In Proceedings of the 2021 9th International Conference on Cyber and IT Service Management (CITSM), Bengkulu, Indonesia, 22–23 September 2021. [CrossRef]
120. Tinenti, Y.R. *Model Pembelajaran Berbasis Proyek (PBP) Dan Penerapannya Dalam Proses Pemelajaran Di Kelas*; Deepublish Publisher: Yogyakarta, Indonesia, 2018.
121. Ornelas, N.O. *An Ecosystem: Computational Thinking, Project-Based Learning & Logo*; Portfolio School: New York, NY, USA, 2017.
122. Kusaka, S. Systematizing ICT education curriculum for developing computational thinking: Case studies of curricula in the United States, Australia, and the United Kingdom. *J. Educ. Learn.* **2021**, *10*, 76–83. [CrossRef]