

Received 20 August 2022, accepted 4 September 2022, date of publication 15 September 2022,
date of current version 29 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3206805

TOPICAL REVIEW

Domain and Challenges of Big Data and Archaeological Photogrammetry With Blockchain

OMER AZIZ¹, MUHAMMAD SHOAIB FAROOQ¹, (Member, IEEE), AND ADEL KHELIFI²

¹School of Systems and Technology, University of Management and Technology, Lahore, Punjab 54770, Pakistan

²Department of Computer Science and Information Technology, Abu Dhabi University, Abu Dhabi, United Arab Emirates

Corresponding author: Omer Aziz (f2018288006@umt.edu.pk)

ABSTRACT With gigantic growth of data volume that is moved across the web links today, there has been a gigantic measure of perplexing information produced. Extremely huge sets of data including universities, organizations framework, institution gas, petroleum sector, photogrammetry, healthcare, and archaeology, that have so enormous thus complex information with more differed structure. The major challenge is how to handle this significant volume of data, also in archaeological photogrammetry which alluded to as Big Data. Although big data has to be securely flying and conveyed through the internet. It cannot be controlled with regular conventional methods that fail to handle it, so there is a need for more up-to-date developed tools. The big data have frequently divided into V's characteristics beginning from three V's: volume, velocity and variety. The initial three V's have been stretched out during time through researches to arrive 56 V's till now. Among them are three newfound by the author that implies it multiplied near twenty times. Researcher had to dive to search for all of these characteristics in many researches to detect and build comparisons to answer the old, current, and restored essential inquiry, "how many V's aspects (characteristics) in big data with archaeological photogrammetry and blockchain." This paper provides a comprehensive overview of all secured big data V's (characteristics) as well as their strength and limitations with archaeological photogrammetry and blockchain.

INDEX TERMS Big data (BD), archaeology, photogrammetry, open-source software (OSS), blockchain.

I. INTRODUCTION

Modern era moved into the data-driven age, the volume of data being processed on a day-to-day basis is growing at a breakneck speed. The reason for increased data is the abundance of data-generating devices like wearable electronics, smartphones, tablets, and a whole host of other sensory gadgets. An estimated 59 zettabytes of data will be generated and processed in the 2021 alone. By 2025, the International Data Corporation (IDC) predicts that the amount of data saved would have increased to 163 zettabytes and shown in fig1.

As a result, data storage capacity has risen from megabytes to exabytes, with zettabytes per year predicted in the coming years [1]. The amount of data that will be generated during the next three years will be more than that generated during the last thirty years. The amount of data produced in the next

five years will be three times that produced in the previous five years. The task of handling and managing continuously increasing data is becoming a problem. Another issue with data is that it is being generated in new formats and in unstructured forms, such as photographs, audio, tweets, text messages, server logs, and so on. The petabyte era is coming to an end, leaving us at the threshold of the exabyte era. The technological revolution has aided billions of people by generating massive amounts of data, which has been dubbed "Big Data." [2].

According to research [3], Big Data (BD) basically meant the amount of data that could not be processed in an efficient manner by the traditional database tools and methods.

Every time a new medium of storage was devised, the amount of data that could be retrieved became larger because it was now easier to do so. The first concept of BD centred on organised data, but many academics and practitioners noticed that the vast majority of information on the planet

The associate editor coordinating the review of this manuscript and approving it for publication was Mehdi Sookhak¹.

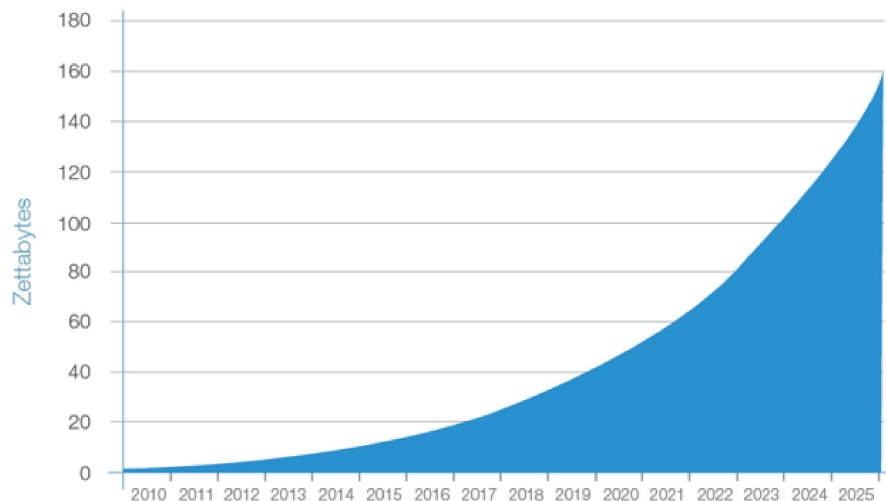


FIGURE 1. Evolution of data.

is in the form of unstructured data, with text and picture being the most common. The profusion of complex and heterogeneous information being generated from all over the world may undoubtedly be classified as a BD age; in fact, this phenomena has been dubbed the Data Deluge by some researchers [4].

The concept of BD has been defined in a variety of ways by IBM:

- “Data, coming from all kinds of sources; posts and stories on social media sites, transaction record of purchases, sensors gathering climate information, GPS signals, text, pictures, and videos”
- “Big Data is considered to be gigantic data sets that are disorganized and unstructured,” and
- “Big Data is a volume of data that has surpassed the processing abilities of traditional database engines and infrastructures” [5].

Study goal only five years ago was to have 10 to 100s of GB of data storage on systems. So, it is being aimed for data storage capacities of tens to hundreds of terabytes. As a result, large data is always evolving. To put it another way, it's the amount of data that's just out of reach right now. Author realise that the target has changed as soon as we have a solution for what our target was in terms of processing and storing the data. The current rate of rise in the volume of data collected is astounding.

BD Executive Survey of 2013 has projected a rise of BD greater than 10MM from 19% to 50% during 2013-2016 [6]. This increased to over 40 trillion GB in 2020 and is expected to double every 2 years and have dominion over information technology's industries up to year 2030. Only 25% of the world's data is being saved digitally before to BD. The rest of the data was saved on films, papers, and analogue media. However, digital data storage has grown so swiftly that just about 2% of the world's data is not digital [7].

The application of BD will be a critical component of individual company growth and rivalry. Every organisation should take BD seriously from the standpoint of competition and potential value extraction. Established organisations and new entrants in every field will use the most up-to-date data gathering methods to innovate, compete and capture value gathered and also the real-time data. Every field we looked at has examples of this type of data utilisation.

It want to execute an operation with BD, its vast volume offers a challenge. However, how will we know if the operation was successful? And do you know if it's correct or not? The truth or validity of Big Data is a major issue because it is nearly hard to check spelling, slang, and vocabulary with such a large amount of data. If the information isn't accurate, it's useless.

Modern data-driven technologies, as well as an increase in processing and data storage capacities, have greatly aided the growth of the BD industry. Companies like Google, Microsoft, Amazon, and Yahoo are collecting and maintaining data that can be measured in proportional greater than exabytes. Furthermore, social media sites such as YouTube, Facebook, Twitter, and Instagram have billions of users who produce massive amounts of data every second of the day. Various organisations have invested in the product's development and research. BD Analytics is a prominent topic in data science research since several firms have invested in building products to handle their monitoring, testing, data analysis, simulations, and other knowledge and business demands.

The core of the Big Data analytics is the processing and generation of meaningful patterns for making inferences, predictions and decision. There are also other challenges that BD analytics need to overcome for data analysis and machine learning. Variation in raw data format, speed of streaming data, data analysis reliability, vast and distributed input sources, noisy and low quality data, scalability of algorithms, increased dimensionality of data, uncategorized data,

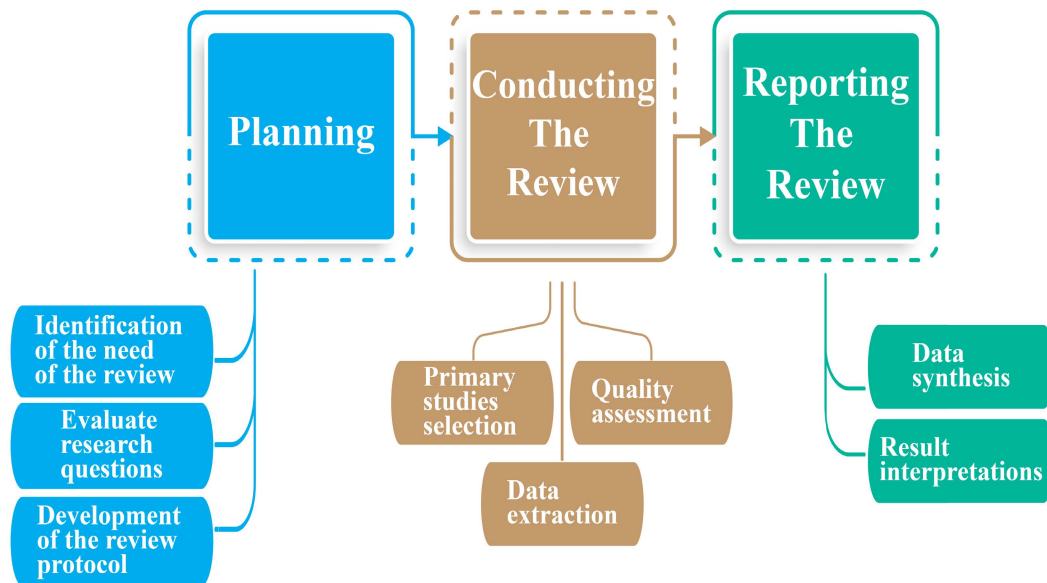


FIGURE 2. Mapping Study.

unsupervised data, limited amount of labelled data, imbalance input information, and so on are some of the challenges [8].

Data acquired from sensors, social media, and financial records is inherently noisy, incompleteness and inconsistent. For assessment of huge data, complex tools for rapidly analysing or anticipating future action with lot of precision, as well as advanced decision-making techniques, are necessary. Variety and speed of data grows, it brings uncertainty, a loss of trust in analytical process and judgment is also taken as a result. In BD analytics, traditional data methodologies and platforms are more accurate, faster, and scalable than artificial intelligence technology (such as machine learning, natural language processing, and computational intelligence) [9].

In [10] six major issues in BD analytics were discussed, including uncertainty. They are primarily concerned with how uncertainty affects the performance of BD learning, but managing uncertainty inherent within a vast data set is a different topic. The importance of data-driven decision-making is becoming more generally recognised, and the notion of Big Data is gaining traction. At every stage of the data-value-creation pipeline, heterogeneity, scale, timeliness, complexity, and privacy concerns hinder development. The problems begin during data collection, when the onslaught of data drives us to make arbitrary decisions about what data to keep and what to delete, as well as how to store what we keep in a consistent manner with the required information.

The fundamental transporter of the advanced change of social legacy the executives, and computerized innovation devices, 3D innovation, laser innovation, geographic data innovation, data set demonstrating and other new advances have given significant certifications to the gamble the board,

observing, arranging and representation of social legacies, advanced innovations like information base administration framework [11].

A massive benefit of latest technology is the pretty growing skills and user friendliness over price ratio, which inspires archaeologists to go into the rising realm of Digital Archaeology. For any metric to be widely accepted in the archaeological community as a benchmark evaluation tool for contrasting various archaeological item detection procedures, this is a crucial need. The required archaeological data for additional (field) investigation is provided by the centroid-based and pixel-based measurements. We anticipate that from now on, the community will view these two metrics as a common performance evaluation tool [12]. Over time, archaeological photography has undergone intense scrutiny and been improved. Methodological and technical advancements in the form of equipment development and digital control of photographic products and environments are significant advancements in archaeological photography [13].

With the emergence of “big” data projects, it is important to think about how these new data scales and perspectives on historic sites and landscapes might complement or conflict with local residents’ modes of knowing. Big data has a lot to offer the archaeological discipline, allowing for the use of never-before-seen scales of data to ask questions and observe sites from novel perspectives, as this issue of JFA demonstrates [14]. Heritage sites now face both new potential and difficulties as the big data era begins. Big data has enormous commercial value, particularly in the application area. However, the market demand cannot be satisfied by the current domestic cultural site development. It is challenging to implement innovative cultural tour service models because the majority of historic site tourism service modes

are still the traditional presentation of ruins and print media commentary [15].

Data's worth skyrockets could be linked to other data, making integration of data a big value provider. Mostly data is now being created directly in digital form, there is the possibility and the task of influencing the generation and automatically linking previously created data. Other core difficulties include data analysis, organisation, retrieval, and modelling. Data analysis appears to be a hurdle in many systems, owing to the less scalability of given methodologies and complexity of data to be observed [16].

Over time, archaeological photography has undergone intense scrutiny and been improved. Methodological and technical advancements in the form of equipment development and digital control of photographic products and environments are significant advancements in archaeological photography [13]. Security, healthcare, education, industry, and government agencies all have expanding Big Data processing requirements. The Big Data discipline is emerging around Vs. We will strive to uncover additional Vs in the future, and there is a possibility that this list will grow to more than 100 Vs.

Many benefits and applications, there are many provokes in enormous information to be handled for better nature of administration that is huge information examination, large information the executives, and huge information protection and security. Blockchain with its decentralization and security nature has the incredible potential to further develop enormous information administrations and applications [17]. Such a methodology is vital on the grounds that neither innovation nor society perspectives can be dealt with independently to get deployable arrangements of a more extensive social, and, surprisingly, public significance. Obviously, while the conventional ways to deal with social legacy safeguarding will stay a highest quality level, they will be progressively supplemented by computerized conservation strategies. Accordingly, in light of functional executions and illustrations learnt in different regions, this multidisciplinary structure paper examinations existing problematic data advances organizations. In accordance with the discoveries it presents a clever mechanical engineering customized to the necessities of social legacy safeguarding that sends an open blockchain design. The design protects the upsides of customary blockchains, which made this innovation so significant, while empowering energy proficient executions that can be sent in versatile applications [18].

BD and the IoT generally give a progressive answer for guaranteeing that electrical energy connected smart matrix, otherwise called the energy Internet. The blockchain has a few huge elements, making it a material innovation for savvy network guidelines to settle the security issues and trust difficulties. This study will give a thorough survey of blockchain executions the network safety discernment and energy information securities in brilliant frameworks [19].

Research project is organised into six major sections: Section II gives an overview of relevant search investigations

that have been done in a similar area in the past in Figure 4 display the article's structure; Sections III and IV discuss the step by step way for conducting a Literature Review, including the some RQs, strings to search, IE(Inclusion/exclusion) criteria, QA, and conclusion. Section V discusses the proposed taxonomy, including main findings and open challenges; and Section V discusses the obtained results. Finally, Section VI brings the article to a close.

Background:

VOLUME: Vast Amount, Scalability, and Size

The term "data volume" alludes to the massive amount of data derived from science and technology, as well as organizations, innovation, and people collaboration records. Volume alludes to the amount of data extracted from various sources such as sound, video, text, research work, long-range interpersonal communication, space images, clinical data, climate forecasting, wrongdoing reports, and catastrophic events, among others.

Regardless, data volume takes up a significant amount of time and effort to manage [20]. Although, because of the speed with which capacity innovations are created on the one hand and the capacity cost is reduced on the other, the capacity limit poses less of a challenge in terms of handling.

As a result, cost-effective data storage arrangements, Cloud advancements, and now Edge developments provide organisations with more options for data storage. In any event, data volume has an impact on executives' data handling and dynamic data [21], [22].

VELOCITY: G(Generation) Speed

It controls the rate where data flows in diverse sources such as corporations, machinery, human communication, and online media destinations. The growth of data might be enormous or nonstop. Importing data can be done in one of two techniques: 1st is batch data and 2nd is streaming data. It is critical when selecting a BD examination stage since constant cycle frequently is time-delicate and requests quicker and close moment investigation results.

The speed of Hadoop is ideal for batch processing of archive data, on the other hand the performance of Apache Spark is excellent for interactive task and real time analysis [23].

In some cases, 5 seconds is past the point of no return. For time-touchy cycles/processes such as detecting fraud, BD should be used as it flows into the attempt to increase its value. 5,000,000 exchange occasions and activities are investigated to discover potential extortion every day [24], [25].

VARIETY: Multiple Data Sets, Heterogeneity

The degree of data arrangement is referred to as data variety. Unstructured data lacks sufficient organisation, whereas structured data has a high degree of organisation [21]. The diversity and fruitfulness of data representations in text, audio, video, pictures, and other formats are measured by data variety.

From an analytic standpoint, it is most likely the most significant impediment to properly utilising large amounts of data. The fact that Data appears in a variety of shapes adds

to the overall complexity. Unstructured and semi-structured data, on the other hand, are more difficult to analyse and make judgments with.

It (Traditional data analysis systems) have incorporated RDBMS. These are only capable of handling structured data and require expensive hardware [26].

VERACITY: Truthfulness, Quality, Meaningfulness.

It alludes the consistency and accuracy of the data. Veracity and Variability was presented by IBM and Microsoft [27] and [22] in which values are added as another extent.

It is almost chatting and extracting useful data information for a certain problems that we are investigating. Data veracity is concerned with data efficiency and accuracy, and secure data when making significant decisions based on the information gathered.

Due to data inconsistency, incompleteness, ambiguity, delay, deception, and approximations, data is graded as good, horrible, or undefined [28].

VALUE: Usefulness, Mining

The importance of data analytics in influencing decision-making [29], [30] Anyone is always eager to mine and maximise the useful data. Because it's has a direct impact on company earnings, it is one of the most crucial variables in BD.

In case, McKinsey [31] The 53 plausible advantages may represent 300 billion to 450 billion in decreased medical care, or 12 to 17% of the 2.6 trillion baselines in US health-care expenses, according to the report, assuming the early achievements were increased to deliver framework wide effect.

Furthermore, Steve identify in [32] that high and low performers in the sector are distinguished by the use of business data and analytics. As a result, the value is found in the rigorous study of precise data.

BD is a massive information asset that necessitates cost-effective and innovative data processing in order to improve decision-making insight [33]. Although this definition isn't perfect, it does provide us with a clear differentiation. We cannot retrieve the data of a dataset using this definition.

As a result, for capturing the value of data, an evaluation-based definition is required. Obviously, acquiring and keeping massive volumes of data is not the purpose of every company or organisation. However, they are all interested in analysing data in order to extract and generate genuine business value [34].

Davenport [35] shared real world and anecdotal sample of how companies establish plan for utilising and extraction from acquired dataset. Furthermore, comprehensive study from [36] in terms of profitability and productivity, data-driven decision-making has been shown to outperform other decision-making strategies.

A number of researchers [37] have underlined the challenges in extracting and obtaining business value from BD analytics. Some firms may afford to pay a higher price for storage associated with higher tiers since the security is better at those levels, resulting in a better value and cost ratio [38].

VALIDITY: Governance, Understandability, Excellency

Ideas for data validity and data truthfulness may be comparable. However, they do not share the same ideas and theories. Data should be legitimate when it transitions from exploratory to actionable stage. To put it another way, a data collection may not have problems with veracity, yet it may not be legitimate and is not properly accepted or understood. Validity of BD is necessitated by occurrence of some hidden connections among pieces within large number of BD generating sources.

As [30] the terms "validity of data" and "veracity of data" are often used interchangeably. They are not the same notion, yet they are similar. Validity refers to the data's correctness and accuracy in relation to its intended use. To put it another way, data may not have any concerns with truthfulness, but it may not be legitimate if it is not correctly understood.

Importantly, the same collection of data may be appropriate for one application or even use but not really for another. Despite the fact that we are working with the information where connections may not be distinct or in beginning phases, it is basic to confirm connections between parts of information to some even out to validate it against utilization.

VOLATILITY: Lifetime, Availability, Durability

In BD it defines as the length of time in which data is valid [24]. We need to figure out when real-time data is no longer effective and applicable for present research in this field. The data should always be present in some sources, but this may not be the case in others. As a result, it is necessary to comprehend the data's requirements, availability, and longevity.

Data is retained for decades in a data standard context to develop a knowledge of the value of data [30]. We can readily recall the structured data retention policy that we employ every day in our organisations when it comes to the volatility of large data. We may easily destroy it once the retention term has expired.

This guideline and policy in real-world data storage apply equally to BD. Such a problem is amplified in the BD world, and it's not as simple to solve as it is in the traditional data world. The retention time for BD may be exceeded, and storage and security may become prohibitively expensive to execute. Because of the variety, volume, and velocity of data, volatility becomes significant.

VIABILITY: Activeness

BD ought to be able to stay alive and active indefinitely, as well as evolve and produce additional data as needed. However, researcher must do more to examine large data sets instantaneously, which necessitates thorough evaluation of the traits and aspects most likely to predict critical business effects [39]. It collect multidimensional data using Big Data, which encompasses a growing number of factors rather than just a big number of records. What impact does the time of day or week have on purchasing decisions?

VISUALIZATION: Attractiveness, Intelligence

It is mentioned that Data Visualization and Interpretation has a lot of appeal. Visualization is important in remotely to

get data characteristics that are related to present targets or objectives, not merely to help consumers or decision-makers understand BD.

To effectively employ Visualization, in BD remote sensing can be combined from different resources in a large Volume and converted into a model which help them to make decision immediately. It's a serious challenge for PB level or bigger inputs, such as in applications related to monitoring the danger.

A Complex with multiple variable's data while staying intelligible/legible are referred to as visualisation. Making that massive amount of data intelligible, easy to grasp, and readable is the difficult part of BD. Raw data can be put to use with the correct analysis and visualisations; otherwise, it is virtually useless [40].

VERIFICATION: Desired Outcome, Authenticity

The process of determining if data is true, precise, or valid [41]. The authenticity and expected outcomes of data processing are referred to as BD verification. Everyone wants to get the most value out of every dataset, to have the data's actual and real value. Furthermore, the value of data must be greater than the cost, maintenance, or ownership of the data.

Stakeholders should pay careful attention to data storage investments. While storage may appear to be less costly and substantially less expensive at purchase time, such a shortfall could jeopardise incredibly sensitive data. For example, storing clinical data for a novel treatment on low-cost, unstable storage save extra funds today and jeopardise data later [42].

VERBOSITY: Loquacity, Volubility, Garrulity,

BD is a significant volume of information that comes from a range of sources, including organised and unstructured data, as well as good and poor information. Information that is wrong, inaccurate, or out of date is referred to as bad data. The dangers of storing this kind of data and information can arise at times.

As a result, double-check that the data you've saved is secure, relevant, complete and reliable. Verbosity knew to quickly differentiate meaning of your recall from repetition is crucial to process speed [43].

VERSATILITY: Alterable, Adaptable

BD is changing to meet the demands of a wide range of businesses, researchers, and government agencies.

It assists businesses with urban planning, visualisation, computational analysis, quality classification, environmental security, and manufacturing through cost models/frameworks and advanced research outcomes. The flexibility of data refers to how useful it is in a variety of situations [44].

VARIABILITY Dynamic, Sources' Changing Behavior Inconsistent data flow is known to as "Data variability." This potential has been put to the test due to the rising need for digital media, which is the primary cause of data load peaks [45].

VISCOSITY: Complexity Because connecting, transformation and matching are all critical activities in BD administration, sophisticated in data management for massive data

sets, particularly when they come from diverse sources, can be extremely difficult [46].

Big Data Complexity refers to the degree of correlation and inter dependencies in large data structures such that tiny changes can have a substantial impact on system behaviour or may not require any changes at all [3].

VOCABULARY: Models, Semantics, Structures

BD is prevalent in academic study, spanning the full spectrum. We will almost likely come across a vast amount of data; this is due to current technology, which permit us gather, analyze, and sample massive amounts of data.

The challenge is converting BD into useful, meaningful and actionable information. This demands a wide range of mathematical, statistical, and computer science tools, as well as approaches that can be intimidating to the uninitiated.

All metadata shapes that explain the data's structure, syntax, content, and origin, such as data models, schema, semantics, ontologies, taxonomies, and other contents [47].

VENUE: Distributed, Heterogeneous, Multiple

Geo-tag real-time location data will soon be included in Online Social Networks (OSN) data, in addition to OSN interaction [48]. Data based on location will soon extend beyond landscape.

The gauntlet of prime types of technology for 3D interaction and also volume rendering technology based on GPU technology is addressed in one study.

This project investigates data-oriented and visual s/ware for the hydrological environment. It also generates surface contour mapping, dynamic simulations and element field mapping of existing fields [49], [50].

VIOLATION: Terrorist Activities, Crimes

BD have been utilised by businesses and governments to examine and address a variety of data science problems. Governments have employed BD apps to detect terrorist, track down criminals, activity, and improve services.

Smart city, for example, detectors/sensors can be used to track movement of vehicles in order to determine traffic volumes and trends [51]. This data can then be combined with information from vehicle owners to identify the correlations between trip times, age groups, and places. This data can be used to improve planning [52].

VERSIONING: Version Control System

Writing codes is a part of both data science and software development. Data science is more iterative and cyclical, with each cycle beginning with a basic comprehension of the data.

The data is collected, explored, cleaned, and transformed, and then machine-learning models are built, validated, and deployed. Researchers and "data science" teams aim to gather, analyse, and cooperate on large datasets in order to extract meaningful insights or condense scientific knowledge.

This type of collaborative data analysis is frequently ad hoc, including a lot of back-and-forth among team members as well as trial-and-error to find the correct analytic tools, programmes, and parameters. The ability to keep track of and

reason about the datasets that are being used is required for this form of collaborative study. In essence, a system to track dataset versions throughout time is required [53].

VAGUENESS: Confusion over the meaning of words and the tools utilised. Regardless of how much data is accessible, the significance of found data is frequently obscure. It was all about the truth in information, with no or little regard for what each would say [54].

VITALITY: Another important idea is data criticality, which is incorporated into the theory of Value. Prioritize data that is more important or relevant to the fundamental business goal [55].

VIRALITY: Speed of Spreading

Virality is the rate of data spreads/broadcasts by user and received by people [54].

VALOR Dealing with Serious Issues

Valor refers to the techniques that we should use to tackle the massive problems in BD. Take on the BD's major challenge [56].

VANE: In the right direction

In Data science and BD, indicate the ability to move in the right path for the right decision-making at the right moment [20].

VANILLA: Immediate Worth

Vanilla refers to the simplest Big Data models that have been well developed and may provide immediate value. When used with care, simple procedures can be beneficial [57]

VANTAGE: Provision of composite systems

Vantage refers to Big Data's ability to provide us with a unique perspective on composite/complex systems [56].

VARIFOCAL: Exceptional comprehension Varifocal refers to BD's talents, which let us to recognize and interpret both the forest and the trees together [56].

VARMINT: Generation of bugs

Varmint is defined as the rate at which bugs age in software when the BD grows massively at a rapid rate [43].

VARNISH: Polish Stain denotes the end-methodology user's for connecting with our work, and shine is important 43.

VASTNESS: Bigness acceleration

It refers to the increased speed that has come with the advent of IOT; BD's 'bigness' is also speeding up [43].

VATICINATION: Prescient analytic Analytical foresight Vaticination is a type of foresight analysis that allows you to judge the future. Depending on the meticulousness and complexity of the situation, these conjectures can be fairly accurate [47], [57].

VAULT: Enlargement in Data Security Nowadays Security increase has become a very important aspect of data storage, by the increase software and information science applications dependent on large and delicate information.

VEER: Prerequisite's route Veer empowers us to have capacity to explore the client's prerequisites and shift bearings rapidly when needed with the progress of agile Data science.

VEIL: Inactive variables reflection

The Veil enables us to see beyond the current barriers and observe the latent characteristics in data.

VERDICT: Fondness

An ever-increasing number of individuals are influenced with replicas' choices, Validity and Veracity have become really significant and are receiving greater development.

VERSED: Required more familiarity

Versed talks about another Big Data opening through which data analysts must have a basic understanding of a variety of topics like as programming, science, measuring, and so on.

VET: Possibility's complication Vet of BD alludes to the methodology wherein Data science allows us to verify our suppositions by proof.

VEXED: revealing Complex issues

The capability of Big Data to reveal insight into con-founded and immense issues in the Data science.

VIBRANT: amorousness, Vitality, Active

The amorousness, dynamic, strong, active, and sparkling practices of BD come through loud and clear. These features provide us with experiences, thoughts, and provision in many features of our data science endeavors.

VICTUAL: Fuels, Nutrition, Nourishment Victual denotes supplies of information to data science shape of BD.

VIRTUOSITY: Required data Data researchers consistently require to realize a little with regards to numerous things in regards to data science, yet normal client likewise ought to develop to realize a ton about this quick bobbing surge of data.

VISIBILITY: Completely Observable BD ought to be apparent to all unreservedly. The Visibility gives a similar significance as Voluntariness. BD provides perceivability into complex BD issues in Data science [56].

VIVIFICATION: All strategies activity Enormous Data in large expanse of Data science has the capacities to energize all method for dynamic and business measures, from showcasing to extortion discovery.

VOGUE: Influence, fashion

The presence of commoners is affected by BD. Nonetheless, trends in business development are always shifting, such as from ML to AI, and from network computing to edge, and from (IoT) to (IoE).

VOICE: Conveys Information, Speaks Noisily

Data science gives the capacity to talk noisily and current information (however only specific information) as BD, on an assorted and different scope of points.

VOODOOISM: Not Juju, Not Wizard, Real

It must persuade as a user to potential clients of Data science's worth in delivering outcomes that can be verified. Because BD and DS have no voodoo talents. To sum up, BD will continue to be authentic.

VOYAGE: Explorer, getting knowledge

It take on a variety of perplexing problems that BD and data science provide. The nature of this attempt to be a Voyager

continuously creates an environment in which data science can be learned.

VULPINE: Every on Crafty

Each BD user utilizes different informal communities/social network, attempting to act shrewdly. Be that as it may, BD and DS ought to have the capacities to deal with these kinds of clients.

VENDIBLE: The actual presence of customers for BD shows urgently that it is obvious—this is clear from the correspondence of some known method for exchanging with supporter's data.

VORACITY:

BD is conceivably excessively voracious such that it might accomplish the impact, oversee, and the likelihood to burn-through itself.

VANITY:

Ineffective of data suggests that it is happy with the effect it produces on others.

VULNERABILITY:

This implies that no system is great and error-free, which implies it's plausible there is a way to understand its hardware and software, progressively implying that any related information can be attached or controlled.

VISUAL: Researcher in a universe of keen-sighted, observing, and trading photographs and recordings, regardless of this if they are close to home or item images or climate photographs through the Internet.

VINCULARITY:

It shows the true meaning of linkage or connectivity. This thought is extremely relevant in the present interconnected world through the web.

VALENCE:

It is an activity demonstrating how thick the data is.

VERITABLE:

Data being truth be told the thing named and not unreal, false, or fanciful

VIRILITY:

With BD it implies that it makes itself. The more BD you have, the more BD gets strength and strong.

II. RESEARCH METHODOLOGY

This SLR (Systematic Literature Review) research expects to assess the current exploration distributed on Big Data with utilizing a set up outlining method and to observe and break down various BD encounters, techniques, technologies, procedures and also tactics. To check this, SMS taken as an examination philosophy.

Deliberate Charting Study intricate the planning phases go before to accomplish this artefact, that is possessing 3 dissimilar strides as 1. Arranging, 2. directing the planning study, and 3. ends and surveys.

A SMS is not like as SLR [58], [59], It content with respect to the development of the association of the show, stressing the acknowledgment of the maximum comparative example, and any place specific sign no-show or enough unequivocal in natural writing. In the interim, it isn't the chalice of

the review, as the explained investigation of articles/papers isn't here. The principle objective is having grouping, content investigation and acknowledgment of distributing gatherings [59], [58].

A. RESEARCH OBJECTIVES

- The goal of this study is to review previous studies and their conclusions, as well as to summarize Big Data and archaeological photogrammetry with blockchain research activities in the ground of DS (Data Science).
- For V's of Big Data, a taxonomy has been proposed.
- To recognize future research opportunities, identify the primary difficulties and outstanding challenges.
- To offer a map of ongoing research in order to obtain results that may be applied in actual and to find new trends in research.
- Another aim is to determine the search domains of the article.

B. RESEARCH QUESTIONS

Six (06) questions are being developed to conclude the chosen research shown in Table 1. A total of 06 RQs have been delegated for specific criteria with relevant causes. Extant studies, challenges, and more commandments relating to Big Domain and Challenges will be classified using the provided answer.

C. SELECTION PROCESS

TABLE 1. R. questions.

No.	Question	Main motivation
Q1	How can be popularity of technology connected to Big Data issues and fields varying with the passage of time?	To classify the publication propensity of Big Data Domains in research throughout time.
Q2	Which Venus publications are the most sought-after domains in Big Data research?	To classify where Domains of Big Data research should be published as a leading publication Venus for upcoming study.
Q3	What are the distinct kinds of BD (Big Data) research?	To find out more about the many forms of study that have been defined in Big Data areas.
Q4	What types of studies have been submitted in a concurrent study on Big Data V's?	To conduct a search for recent studies that will aid in future research on Big Data domains
Q5	What are the current issues and breaches in the keen BD field?	Recognize and determine unreciprocated RQs in contemporary Big Data domains.
Q6	What kind of new domains and aspects of BD and archaeological photogrammetry with blockchain have been described?	To establish the current new areas reported in the current arena of Big Data and archaeological photogrammetry with blockchain.

TABLE 2. Strings used in databases.

Venue	String
Google Scholar	1,2,3,4,5
Research Gate	2,3,4
Springer	2,5
ACM	3,4
Elsevier	2
MDPI	1

TABLE 3. Standards to include and exclude for the selected study.

IC	EC
IC-1- Papers that have been accepted for publication in conferences and journals.	EC-1- Duplicated articles.
IC-2- Papers published in 2011 to 2022	EC-2- The Domains and V's of BD are not addressed in the articles.
IC-3- Articles represents Domains and areas of BD	EC-3- Thesis, books and abstracts.
IC-4- Articles are from any location in the world.	EC-4- Articles are pre 2011 published.

D. SEARCH SCHEME

The subsequent systematic databases/sources were employed to obtain relatively relevant papers for the selected achievement: Research Gate, Springer, Association for Computing Machinery (ACM), Elsevier, GS (Google Scholar), and Multidisciplinary Digital Publishing Institute (MDPI). Google scholar has been functionally employed to advance bibliometric investigations. Three distinct strings are worn to direct spontaneous study for a certain repository see in Fig 3.

Examination system pursued in different data sets to discover significant studies in table 2. A few research strings are usually utilized in various data sets to track down the expected result.

E. PROCESS FOR SELECTION

Selected criteria are intended to identify the areas of investigation that are broadly applicable for it. Same articles by various sources is likewise prohibited. All articles are carefully scrutinized for keywords, abstracts, and titles in order to determine if they should be added or not. Exclude the similar titles, they are not for the review. The next step was to choose articles according to given Table 3 rules. Figure 10 demonstrate of a selection process. A total of 29 publications have been chosen from a total of 340 identified studies.

F. QUALITY ASSESSMENT

The type of the contained article was another critical stage in the evaluation process. Moreover questions were designed [60], [58] to access articles.

1. Answer can be found in the articles. No +0, Moderate +0.5, and Yes +1 are the most likely outcomes.

(2) The publications' contribution to how Big Data and BD topics are discussed. No (+0), Moderate (+0.5), and Yes (+1) are the three options.

(3) Future research goals and gaps are clearly indicated. No +0; Modest +0.5, and Yes +1.

(4) Articles are from known databases. Conference and journal rankings are listed below (CORE): This query was ranked using (Q#1, Q#2, Q#3, and Q#4), as well as JCR reports. Answers to these questions that are feasible:

Session position for changed core (C):

- C(A) have ranking(1.5)
- C(B) have ranking(1)
- C(C) have ranking(0.5)
- If there is no ranking in C(0)

Ranking in Journals:

- If Q.1 is rated (2)
- If Q.2 is rated (1.5)
- If Q.3 and Q.4 is rated (1)
- When it is not rated in JCR list (0)

By adding evaluation for each question, we were able to provide an overall score for each article (ranging from 0 to 5).

G. METHOD TO EXTRACT DATA

The goal is to get favorable perceptions to the presented questions.

Q1. To avoid publication drift, articles must be categorized according to the year they were published.

Q2. It is essential to determine the printing media and basis for these questions (RQ).

Q3. The following aggregation can be used to define research genre [61]:

- Solution proposal: It is proposed that fields and new features of Big Data be discussed. A new resolve or vital refinement of a previous method can be demonstrated. Some examples of argumentation, prospective performance, and the resolution's link.
- Conceptual Proposals: Studies processed by closely observing and analysing the already exists in the Big Data domain. This does not include any practical inquiries.
- Evaluation Research: Big Data domains that have already been defined are assessed and analysed. It refers to recognising difficulties when exploring new BD's area.
- Others: As reviews comparative analysis, Experimental, analytical surveys, performance analysis, development, Investigation, and case study.

Q4. The main RQ of research is apprehension incumbent study in the direction of big data and Vs. We are sure in giving a generic understanding of big data that is also tract the current study trends after compiling all relevant investigations from scientific sources.

This research will enhance current studies and practical information on existing research challenges, assisting in the process of increasing the number of Vs in big data. In the

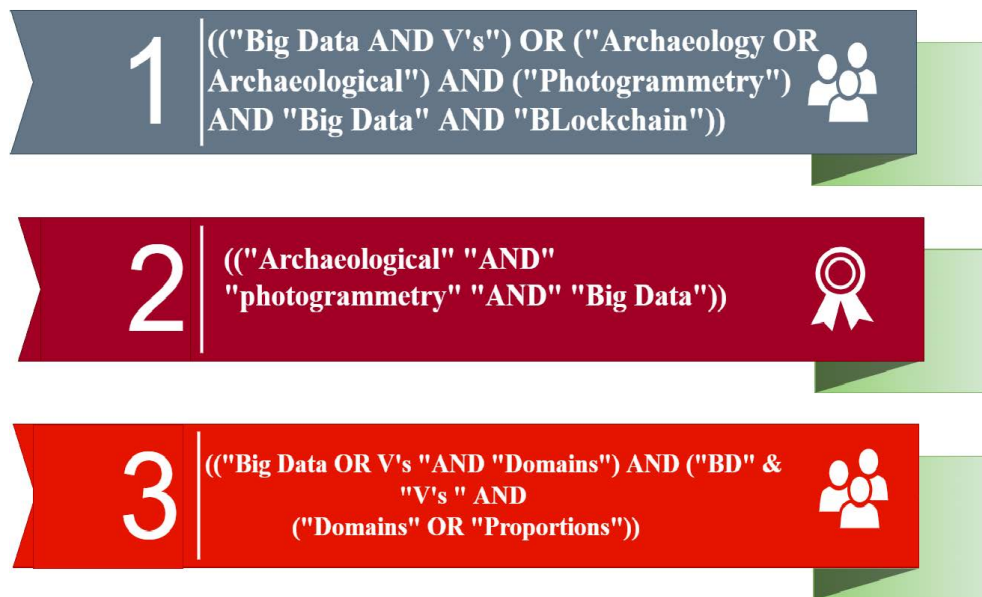


FIGURE 3. Strings.

classification table, several tactics are listed so that we can deliberate the amount of Vs rendering to the years.

Q5. This SLR aids in identifying current study breaches about those research issues that will allow researchers to upgrade and embodiment on the ground wherever the substitute enquiry was needed. The current amount of Vs in BD will help in the realization of unresolved research questions.

Q6. In [61] the following aggregate, the given approach can be categorised as a recommendation:

- Method: A process incidental the phases are accepted to expand amount of Vs in BD.
- Model: The gratified of the system in BD and archaeology with blockchain that alters the reasoning of Vs evolution.
- Infrastructure: Administrative frameworks that are required for commercial function in BD with blockchain.
- Architecture: A preparation, designing, and process buildings for Big Data.
- Framework: A hypothetical outline for influencing or straightening the artefact of somewhat that enlarges the object into something useful in the search for Vs in Big Data.
- Guideline: In Vs., an example of a shape or design that can be used to create a sequence of condition.
- Tool: Whatever used to implement Big Data with archaeological photogrammetry and blockchain.
- Other: outline, stage.

The amalgamation thought was adapted to take into account the research keys that are classified as a result of all RQ, prevailing researches that stand on the base of Questions and Answers, and offering graphical demonstration for the resolution of categorized outcomes.

TABLE 4. QA score.

Reference.	T.S	Total No.
[27] [22]	4	2
[26]	3.5	1
[20] [32] [34] [38] [46] [47]	3	6
[25] [30] [36] [48] [51] [43]	2.5	6
[33] [24] [21]	2	3
[28] [35] [45] [57] [56]	1.5	5

III. RESULTS

This section specifies the results relating to the RQs defined in the specified Table 1. For each RQ's results, a number of publications are picked to pretence the model. We predicted that they are critical and represent a significant undertaking for BD domains.

A. RESULTS OF SELECTION

340 research studies were carefully scrutinised based on their keywords, abstracts, and titles, with 317 papers being rejected and 29 publications being carefully crafted. The acknowledgement of 29 publications was detected in order to state the RQs stated. The breakdown of selected articles is provided in the 5 table, along by a summary of classification consequences and their QA.

1) Q1.HOW CAN BE THE POPULARITY OF TECHNOLOGY CONNECTED TO BIG DATA ISSUES AND DOMAINS VARYING WITH THE PASSAGE OF TIME?

In Fig 5(a) Pretend overall chosen duration of original study, fig5 shows the variation of Vs in BD by years. So that, fig 5

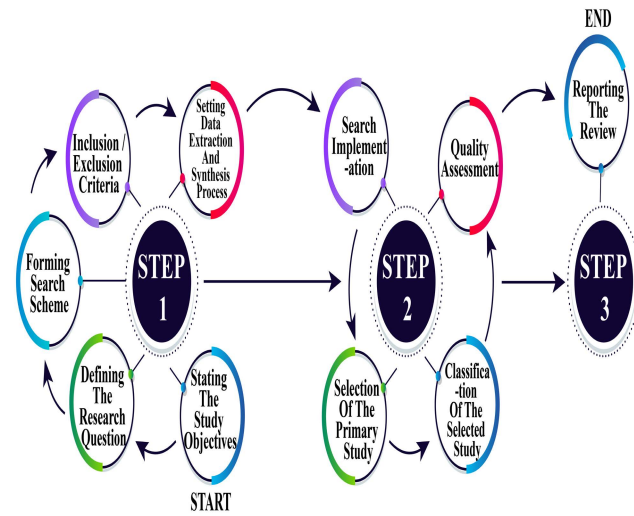


FIGURE 4. Procedure.



FIGURE 5. Paper Evaluation.

shows publication databases the distribution. Figure5 depicts an annual summary of published research from 2011 through 2022. Over the years of evolution and adaptation, very few research publications have been published. The Vs vantages adherent detriment BD architects focused could be described in this way. There is no study it until 2022 but may be administered by the time, and is fair to fetch out the exceptional studies in 2022. Figure 5 stated that there were a lot of Vs in journals. As seen in figure 5, very few studies were published in journals each year.

2) Q2.WHICH VENUS PUBLICATIONS ARE THE MOST SOUGHT-AFTER DOMAINS IN BIG DATA RESEARCH?

Figure 6 depicts the publication databases of the selected papers. The published sources covered in this mapping research were journals, Books and conferences. The majority of publications were published in journals (24) (83%), Books (01)(03%) and conferences (4)(13%). Furthermore, the table 6 lists all of the channels through which the chosen



FIGURE 6. R.Databases.

articles were published. As seen in the table 6, all of the selected articles were used in numerous publications.

3) Q3. WHAT ARE THE DISTINCT KINDS OF BIG DATA RESEARCH?

As shown in figure 8, four (4) categories of research were discovered in this study: survey (11 articles) (37%), solution proposal (10 studies) (35%), and investigative study (5 articles) (18%) and experimental study (3 articles) (10%). The maximum articles (Solution proposal) are experimental in DB, although few Vs in BD shown in figure 9. Several sorts of research models are listed in the next paragraph: Multifaceted information organization can be very composite for massive information sets specifically when they come from unlike foundations, because relating, corresponding and alteration are momentous actions in Big Data administration [46]. Complication of BD contracts by grade of association or mutuality in BD constructions in small variations and have noteworthy consequence in system, conduct and may not put on any variations [3]. The problem of assisting information in DB quickly and securely for the next Vs era has been overcome. Supported approaches vs. The growing use of social media, which is the key origin of the rise in information loads, has put this property to the test [45].

4) Q4. WHAT TYPES OF STUDIES HAVE BEEN SUBMITTED IN A CONCURRENT STUDY ON BIG DATA V'S?

We have evaluated 29 research papers after an extensive analysis of 340 papers. We found only 6 conference papers on the Big data domain and characteristics, none of them is published in any well-reputed conferences. On the other hand, we have found 23 journals,5 conferences and only 1 book is published with good ranking and we found 11 journals that are published without ranking.

5) Q5. WHAT ARE THE GAPS AND CURRENT ISSUES IN THE SMART BD FIELD?

Nonetheless, several research breaches have been recognized. The crucial research breach is that present studies do not address the role of Vs in BD-based technique, which requires

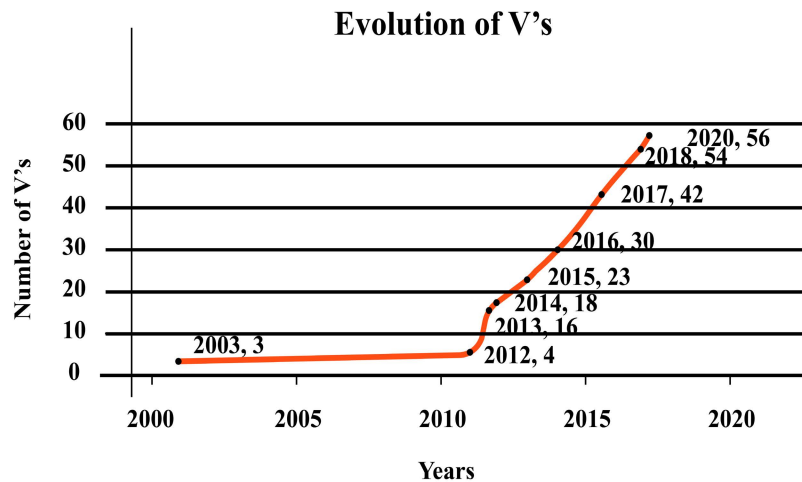


FIGURE 7. Evolution of V's Over Year.



FIGURE 8. Study Genre.

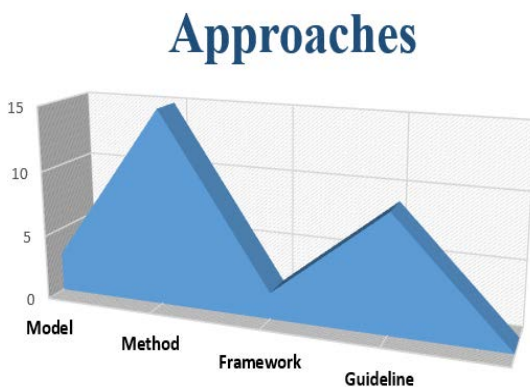


FIGURE 9. Approaches.

further investigation. There are extremely few papers that have been published in known conferences. Publications should be in known conferences.

6) Q6. WHAT KIND OF NEW DOMAINS AND ASPECTS OF BD AND ARCHAEOLOGICAL PHOTOGRAMMETRY WITH BLOCKCHAIN HAVE BEEN DESCRIBED?

The data explosion is arrived, and with more data being produced every day, there is no sign of it slowing down or stopping. These dimensions are expanding as time passes;

this list will grow from 3 to 56V's in 2022, with several further to derived. So many comes in the future years, it will reach 100 V's. Different open source software can be used in archaeological photogrammetry as it will also be helpful to create a big data. This created data can also be secured by integrating block-chain.

B. QUALITY ASSESSMENT

The QA scores are given in the table 4. Almost 25% articles are having an average, 59% have standard, and 16% papers are without any score. QA can helps to choose sited articles with defined asserted.

IV. DISCUSSION

In this SLR Vs in Big Data are hypothetically answer the given RQ's proffered by investigation. It display that Vs have been examined over years.

A. TAXONOMY

The term "Big Data" was originally used in a paper by Diebold in the year 2000. The Big Data age has brought with it a slew of new potential for promoting economic growth, improving education, advancing science, strengthening health care, and mounting social collaboration and entertainment options. However, while big data has its assistances, it also has its drawbacks in form of challenges and issues.

Since then, the number of Vs has steadily increased. In 2001, it is accredited with devising the three big Vs of BD: variety, volume, and velocity. Many individuals began to add up more number of V's to the characterization of BD when it gained a lot of attention. Other authors referred to the characteristics of big data as pillars.

Varacity is included to the V family as the 4th V in 2012. So many studies classified the evolution of V's over time for different ways, but we sought to find a pattern of V's growth through time in this work. In 2013, Big Data becomes increasingly popular, and individuals begin to discuss it. Different researchers added 12 new pillars of big data to the list in 2013.

TABLE 5. Classification of articles.

Ref.	Classification			QA							
	Years	P.channels	Articles Types	R. Topics	Approaches	No. of Vs	1.	2.	3.	4.	TS
[20]	2019	Conference	Solution Proposal	The 51 v's of big data: survey, technologies, characteristics, opportunities, issues and challenges	Guideline	51	1	1	1	0	3
[21]	2014	Journal	Investigation Study	Big data characteristics	Method	8	0.5	0.5	1	0	2
[22]	2014	Journal	Survey	Data-intensive applications, challenges, techniques and technologies: A survey on Big Data	Guideline	1	0.5	0.5	1	2	4
[24]	2014	Journal	Survey	Big data: survey, tech- nologies, opportunities, and challenges	Guideline	6	0.5	0.5	1	0	2
[25]	2018	Conference	Solution Proposal	The 10vs, issues and challenges of big data	Method	10	1	0.5	1	0	2.5
[26]	2014	Journal	Survey	Big data: A survey	Method	5	0.5	0.5	1	1.5	3.5
[27]	2016	Journal	Solution Pa- per	Big data: From begin- ning to future	Method	5	0.5	0.5	1	2	4
[28]	2020	Journal	Experimental Research	Demystifying Big Data: A Practical Guide To Transforming The Busi- ness of Government	Model	5	0	0.5	1	0	1.5
[30]	2014	Conference	Survey Study	Seven v's of big data un- derstanding big data to extract value	Guideline	7	1	0.5	1	0	2.5
[32]	2011	Journal	Experimental Research	Big Data, Analytics and the Path From Insights to Value	Method	1	0.5	0.5	0.5	1.5	3
[33]	2013	Journal	Survey	Undefined by data: a sur- vey of big data defini- tions	Guideline	5	0.5	0.5	1	0	2
[34]	2015	Journal	Solution Proposal	How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study	Framework	5	0.5	0.5	1	1	3

TABLE 5. (Continued.) Classification of articles.

[35]	2016	Journal	Solution Proposal	Competing on Analytics	Method	1	0.5	0.5	0.5	0	1.5
[36]	2012	Journal	Solution Proposal	Big data: The Management Revolution	Guideline	4	0	0.5	1	1	2.5
[38]	2013	Journal	Experimental Research	The information artifact in it governance: toward a theory of information governance	Model	1	0	0.5	0.5	2	3
[45]	2013	Conference	Solution Proposal	Big data: issues, challenges, tools and good practices	Method	3	1	0.5	0	0	1.5
[46]	2015	Journal	Survey Paper	A survey on indexing techniques for big data: taxonomy and performance evaluation	Method	6	0	0.5	1	1.5	3
[48]	2014	Journal	Solution Proposal	The real-time city? Big data and smart urbanism	Method	1	1	0.5	1	0	2.5
[51]	2013	Journal	Investigative study	Who's afraid of big data?	Method	1	0.5	0.5	1.5	0	2.5
[43]	2020	Journal	Survey Paper	How Many Old and New Big Data V's Characteristics, Processing Technology, And Applications (BD1)	Method	56	1	1	0.5	0	2.5
[47]	2020	Journal	Survey Paper	Fifty-six big data v's characteristics and proposed strategies to overcome security and privacy challenges	Method	56	1	1	1	0	3
[57]	2020	Journal	Survey Paper	Using hadoop technology to overcome big data problems by choosing proposed cost-efficient scheduler algorithm	Method	56	0.5	0.5	0.5	0	1.5
[56]	2019	Book	Survey Paper	Big Data in Healthcare: A Survey	Method	42	0.5	0.5	0.5	0	1.5
[13]	2022	Journal	Investigative study	Do-It-Yourself digital archaeology: Introduction and practical applications of photography and	Method	0	0	0	0.5	0.5	1

TABLE 5. (Continued.) Classification of articles.

				photogrammetry for the 2D and 3D representation of small objects and artefacts							
[12]	2022	Journal	Investigation study	Deep Learning for Ar- chaeological Object De- tection on LiDAR: New Evaluation Measures and Insights	Framework	0	0.5	0	0	1	1.5
[17]	2022	Journal	Solution Proposal	A survey on blockchain for big data: Approaches, opportunities, and future directions	Model	1	1	0	0	0.5	1.5
[19]	2022	Journal	Survey	Blockchain Technology on Smart Grid, Energy Trading, and Big Data: Security Issues, Chal- lenges, and Recommen- dations	Guideline	0	0	1	1	0	2
[11]	2020	Conference	Investigation	Tourism Industry is Leading the Digital Transformation of Cultural Heritage Management : Bibliometric Analysis Based on Web of Science Database	Guideline	0	0.5	0	0.5	0	1
[18]	2022	Journal	Solution	Cultural heritage preservation by using blockchain technologies	Guideline	0	0	0	0	0.5	0.5

Scholars found Visibility and Vocabulary the following year, and it was encompassed to the BD family. In 2015, the number of Vs increased to 23, and many scholars began working in the Big Data sector, discovering new domains and testing these features in a variety of applications. The new V-features of BD have arisen to aid in the efficient acquisition of information from BD. The accumulation of additional V-features increases the dimensionality of large data, resulting in a reduction in data complexity.

In 2017, 42 V's were identified, opening up new avenues of knowledge and solutions to the Big Data's complexity. Big Data has become a communal norm in recent years, and it is also the most essential topic in the IT area as well as other fields such as business.

These dimensions are growing in size as time goes on; in 2017, there are 42 V's. so in this list it will grow from three

Vs to 56 Vs by 2020 or more. Many researchers believe that in the future years, it will reach 100 V's. Figure 7 depicts the evolution of V's over time.

Also the Taxonomy has been proposed in figure 11. These features deliver explore prospect to the scholar and practitioners in command to efficiently accomplish BD. The complete study in BD circles around these features. Additionally it can resolve alot of problems related to BD. It also helps to differentiate the BD nature.

B. MAJOR FINDINGS

A total of 340 articles were carefully scrutinised based on their titles, keywords, and abstracts; 317 papers were being rejected, and 29 articles were chosen. A total of 29 papers were chosen to provide solutions to the specified

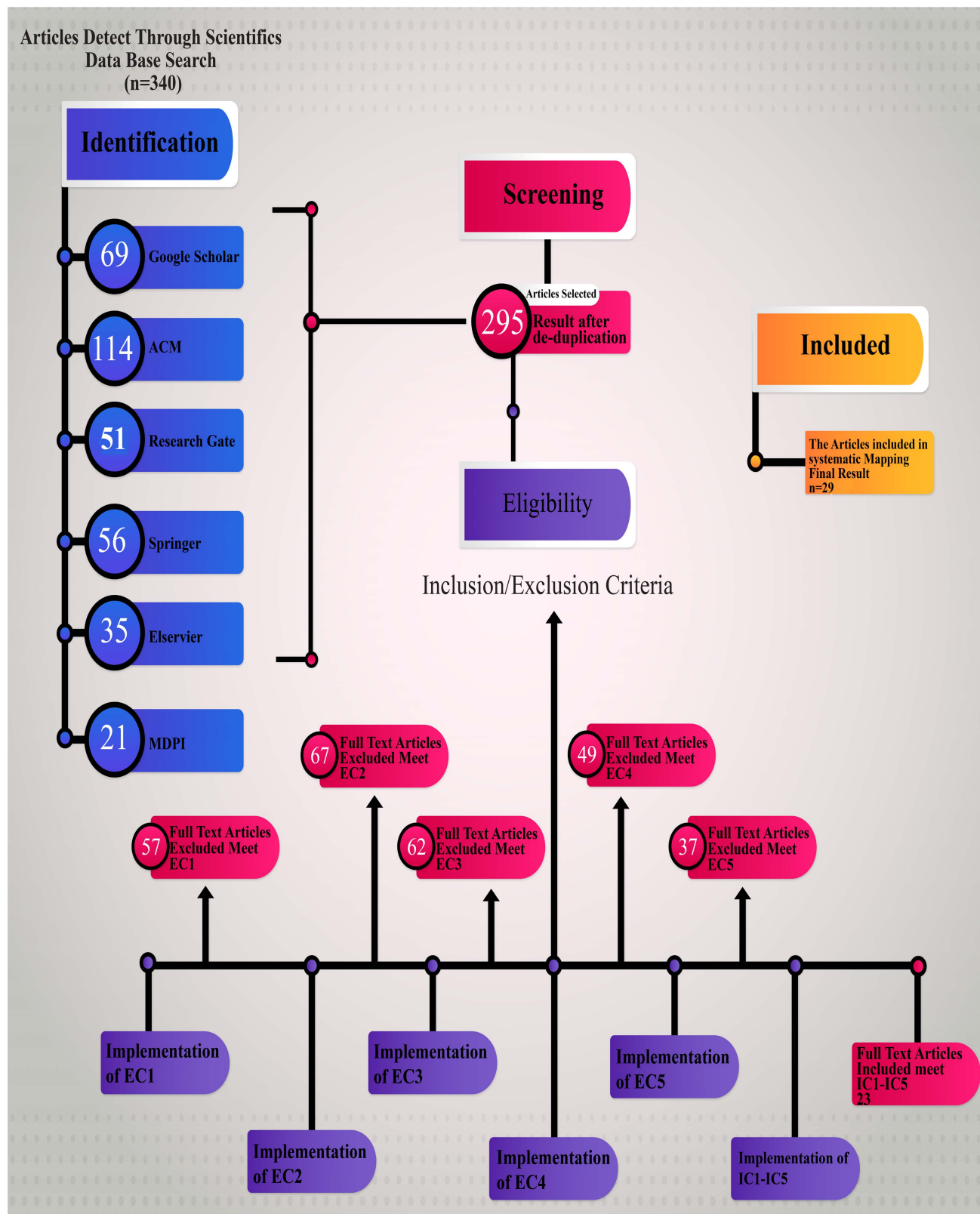


FIGURE 10. Systematic Mapping Review.

questions. The explanation of categorization conclusions and their quires is given to the chosen investigations. Our main findings are listed below:

- The majority of the investigations were conducted between 2011 and 2022. The advancement of technology continues to be a challenge.

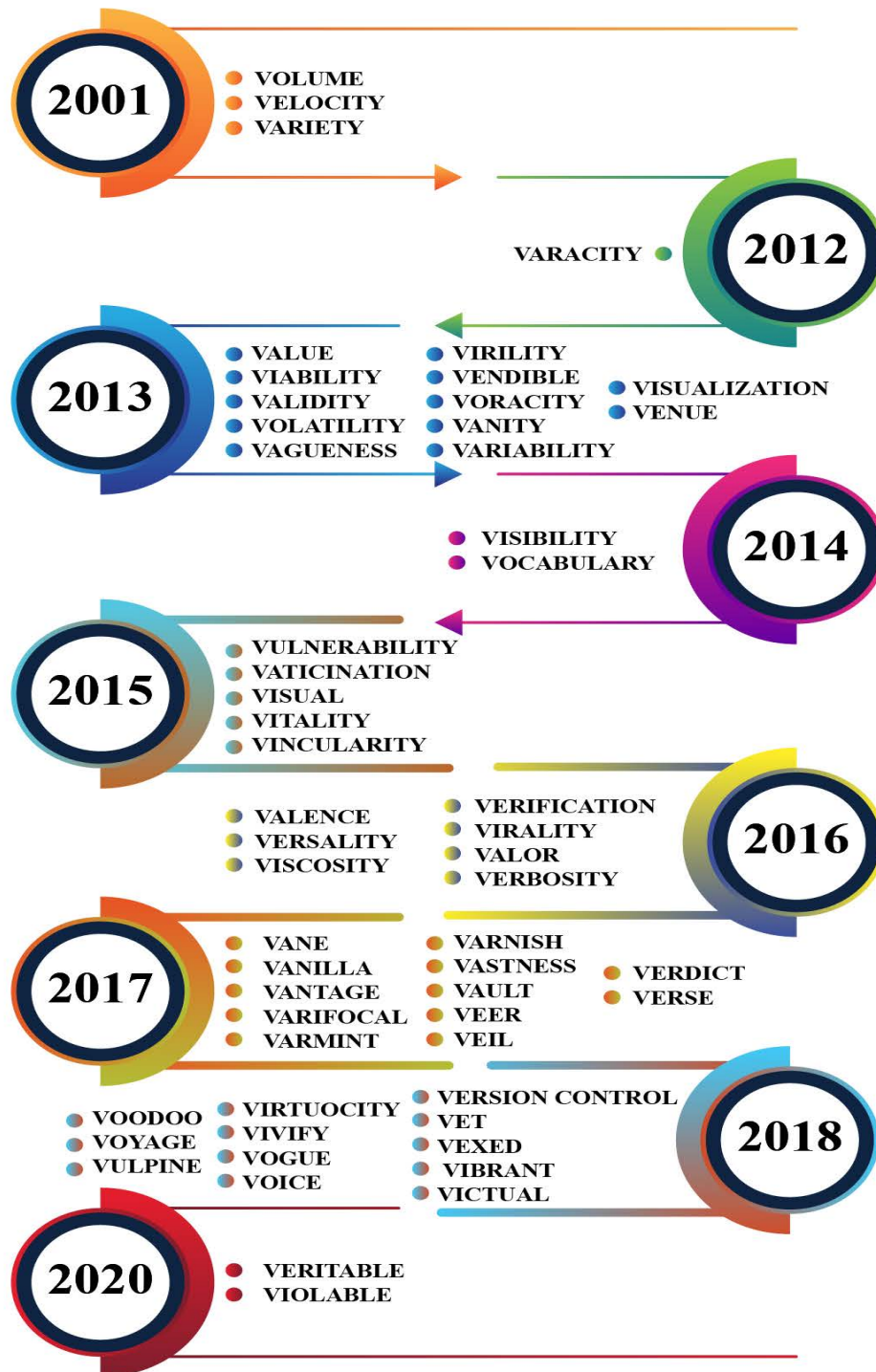


FIGURE 11. Taxonomy.

- Three separate publishing routes were used, with 83 percent of articles appearing in different journals, 3 percent books and 13 percent appearing at conferences.
- There were four (4) categories of research discovered, this included 11 survey articles(37%), 10 studies on solution proposal(35%), 4 articles on experimental studies(18%), and investigative studies(18%).
- The majority of studies and research are covering not more than four to five V's. We haven't been able to locate a huge quantity of V's.

TABLE 6. Databases of selected articles.

PS(Publication source)	References
Proceedings of the international conference on omni-layer intelligent systems	[20]
Journal of The Macrotheme Review	[21]
Elsevier Journal of Information sciences	[22]
Hindawi The scientific world journal	[24]
Proceedings of the 2018 international conference on big data and education	[25]
Springer Journal of Mobile networks and applications	[26]
Elsevier International Journal of Information Management	[27]
TechAmerica Foundation, Washington	[28]
IEEE Proceedings of the 2014 zone 1 conference of the American Society for Engineering Education	[30]
Journal MIT sloan management review (ICISCT) IEEE	[32]
arXiv preprint arXiv:1309.5821	[33]
Elsevier International Journal of Production Economics	[34]
Cambridge Harvard business review	[36]
Journal of Management Information Systems	[38]
IEEE 2013 Sixth international conference on contemporary computing (IC3)	[45]
Springer Knowledge and information systems	[46]
Springer GeoJournal	[48]
IEEE IT Professional	[51]

- The current issues are that, according to this study, there is no recent literature on Vs in BD and archaeological photogrammetry with blockchain.
- The majority of research methodologies use the method 52% of the time and the model 10% of the time. The remaining approaches are divided into two categories: framework (6%) and guideline (31%).
- Three New domains and V's are introduced my author as discussed below:
- **VENALITY**: The state or nature of being venal. Willing to sell one's administrations or force/power

- **VISCIDITY**: being firm and sticky.
- **VICISSITUDE**: property alludes to the test of scaling Big Data complex work processes

C. OPEN CHALLENGES

In this section, open challenges have been discussed. Following the amount of Vs, certain difficulties have been noted, such as the fact that several of Vs have not been deliberated in this study. Another, uncertain encounters are that the acceptance of Vs in Big Data and archaeological photogrammetry with blockchain is quite slow. It has been observed that every latest V is being explored in a new publication. Researcher has also observed about the open source software can be used in archaeological photogrammetry. Though, it is strive to publish a paper that mentions the greatest number of Vs. In order to achieve this, research and studies on Vs in BD have to be built up on by establishing a certain standard.

V. CONCLUSION

Big data of the Vs with archaeological photogrammetry and blockchain is discussed in-depth by analyzing 29 different articles. After a thorough examination of past research, it is determined that the majority of Vs are not covered. The major goal of research was to search the already available data and condense into as many Vs as possible. From 2011 to 2022, 340 publications were selected from an initial list of 80699 studies, and 29 were characterized as intent criteria: research and contribution kind, number of Vs, issues investigated articles, and techniques. Vs in BD are thought to have received little attention until 2022. The majority of the chosen studies were published in various journals, although some mature publications came from conferences as well. There are two types of esquires: experimental solution and suggestion of solution. The articles in this mapping study did not contain the design and implementations of Vs. In this study, three more Vs have been presented. In addition, this study includes a taxonomy that can assist other specialists in identifying numerous approaches that can improve the study's performance. On the other hand, evolution research must be regulated in order to assess existing strategies. As a result, the current study has deeply explored Big Data with archaeological photogrammetry and blockchain, along with the associated Vs and their brief explanations. Open source software can also be used in archaeological photogrammetry where data can be securely saved as big data. The Big Data and archaeological photogrammetry with blockchain discipline is emerging around Vs.

REFERENCES

- [1] S. Bahri, N. Zoghiami, M. Abed, and J. M. R. Tavares, "Big data for healthcare: A survey," *IEEE Access*, vol. 7, pp. 7397–7408, 2018.
- [2] B. Hong, X. Meng, L. Chen, W. Winiwarter, and W. Song, *Database Systems for Advanced Applications: 18th International Conference, DASFAA 2013, International Workshops: BDMA, SNSM, SeCoP, Wuhan, China, April 22–25, 2013. Proceedings*, vol. 7827. Springer, 2013.
- [3] S. Kaisler, F. Armour, J. A. Espinosa, and W. Money, "Big data: Issues and challenges moving forward," in *Proc. 46th Hawaii Int. Conf. Syst. Sci.*, Jan. 2013, pp. 995–1004.

- [4] U. Sivarajah, M. M. Kamal, Z. Irani, and V. Weerakkody, "Critical analysis of big data challenges and analytical methods," *J. Bus. Res.*, vol. 70, pp. 263–286, Jan. 2017.
- [5] P. Michalik, J. Štöfa, and I. Zolotova, "Concept definition for big data architecture in the education system," in *Proc. IEEE 12th Int. Symp. Appl. Mach. Intell. Informat. (SAMi)*, Jan. 2014, pp. 331–334.
- [6] T. Davenport, "Big data executive survey 2013: The state of big data in the large corporate world," NewVantage Partnes LLC, Boston, MA, USA, Tech. Rep., 2013.
- [7] K. Cukier and V. Mayer-Schoenberger, "The rise of big data: How it's changing the way we think about the world," *Foreign Aff.*, vol. 92, p. 28, 2013.
- [8] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, no. 1, Dec. 2015.
- [9] R. H. Hariri, E. M. Fredericks, and K. M. Bowers, "Uncertainty in big data analytics: Survey, opportunities, and challenges," *J. Big Data*, vol. 6, no. 1, p. 44, Dec. 2019.
- [10] X. Wang and Y. He, "Learning from uncertainty for big data: Future analytical challenges and strategies," *IEEE Syst., Man, Cybern. Mag.*, vol. 2, no. 2, pp. 26–31, Apr. 2016.
- [11] X. Zhou, C. Zhang, Z. Pan, and P. Wei, "Tourism industry is leading the digital transformation of cultural heritage management: Bibliometric analysis based on web of science database," in *Proc. Manage. Sci. Informationization Econ. Innov. Develop. Conf. (MSIED)*, Dec. 2020, pp. 273–276.
- [12] M. Fiorucci, W. B. Verschoof-van der Vaart, P. Soleni, B. Le Saux, and A. Traviglia, "Deep learning for archaeological object detection on LiDAR: New evaluation measures and insights," *Remote Sens.*, vol. 14, no. 7, p. 1694, Mar. 2022.
- [13] J. N. Cerasoni, F. do Nascimento Rodrigues, Y. Tang, and E. Y. Hallett, "Do-it-yourself digital archaeology: Introduction and practical applications of photography and photogrammetry for the 2D and 3D representation of small objects and artefacts," *PLoS ONE*, vol. 17, no. 4, Apr. 2022, Art. no. e0267168.
- [14] A. Mickel, "The proximity of communities to the expanse of big data," *J. Field Archaeology*, vol. 45, no. 1, pp. S51–S60, Feb. 2020.
- [15] J. Qiu, J. Li, and H. Sun, "Innovative and applied research on big data platforms of smart heritage," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vols. 2, pp. 257–261, Aug. 2015.
- [16] D. Agrawal et al., "Challenges and opportunities with big data 2011–1," Tech. Rep., 2011.
- [17] N. Deepa, Q.-V. Pham, D. C. Nguyen, S. Bhattacharya, B. Prabadevi, T. R. Gadekallu, P. K. R. Maddikunta, F. Fang, and P. N. Pathirana, "A survey on blockchain for big data: Approaches, opportunities, and future directions," *Future Gener. Comput. Syst.*, vol. 131, pp. 209–226, Jun. 2022.
- [18] D. Trček, "Cultural heritage preservation by using blockchain technologies," *Heritage Sci.*, vol. 10, no. 1, pp. 1–11, Dec. 2022.
- [19] M. K. Hasan, A. Alkhalifah, S. Islam, N. B. M. Babiker, A. K. M. A. Habib, A. H. M. Aman, and M. A. Hossain, "Blockchain technology on smart grid, energy trading, and big data: Security issues, challenges, and recommendations," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–26, Jan. 2022.
- [20] N. Khan, A. Naim, M. R. Hussain, Q. N. Naveed, N. Ahmad, and S. Qamar, "The 51 V's of big data: Survey, technologies, characteristics, opportunities, issues and challenges," in *Proc. Int. Conf. Omni-Layer Intell. Syst.*, May 2019, pp. 19–24.
- [21] P. Géczy, "Big data characteristics," *Macrothème Rev.*, vol. 3, no. 6, pp. 94–104, 2014.
- [22] C. L. P. Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on big data," *Inf. Sci.*, vol. 275, pp. 314–347, Aug. 2014.
- [23] S. Salehian and Y. Yan, "Comparison of spark resource managers and distributed file systems," in *Proc. IEEE Int. Conf. Big Data Cloud Comput. (BDCloud), Social Comput. Netw. (SocialCom), Sustain. Comput. Commun. (SustainCom) (BDCloud-SocialCom-SustainCom)*, Oct. 2016, pp. 567–572.
- [24] N. Khan, I. Yaqoob, I. A. T. Hashem, Z. Inayat, W. K. M. Ali, M. Alam, M. Shiraz, and A. Gani, "Big data: Survey, technologies, opportunities, and challenges," *Sci. World J.*, vol. 2014, Jul. 2014, Art. no. 712826.
- [25] N. Khan, M. Alsaqer, H. Shah, G. Badsha, A. A. Abbasi, and S. Salehian, "The 10 VS, issues and challenges of big data," in *Proc. Int. Conf. Big Data Educ.*, Mar. 2018, pp. 52–56.
- [26] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Mobile Netw. Appl.*, vol. 19, no. 2, pp. 171–209, Apr. 2014.
- [27] I. Yaqoob, I. Hashem, A. Gani, S. Mokhtar, E. Ahmed, N. Anuar, and A. Vasilakos, "Big data: From beginning to future," *Int. J. Inf. Manage.*, vol. 36, no. 6, pp. 1231–1247, Dec. 2016.
- [28] S. Mills, S. Lucas, L. Irakliotis, M. Rappa, T. Carlson, and B. Perlowitz, "Demystifying big data: A practical guide to transforming the business of government," TechAmerica Found., Washington, DC, USA, Tech. Rep., 2012.
- [29] M. D. O. Igarashi, P. E. Sartorelli, and M. Z. de Lima, "Twenty years survey of big data: Definition, concepts, and applications in engineering," in *Proc. Brazilian Technol. Symp.*, Springer, 2019, pp. 509–517.
- [30] M. Ali-ud-din Khan, M. F. Uddin, and N. Gupta, "Seven V's of big data understanding big data to extract value," in *Proc. Zone 1 Conf. Amer. Soc. Eng. Educ.*, 2014, pp. 1–5.
- [31] B. Kayyali, D. Knott, and S. Van Kuiken, "The big-data revolution in us health care: Accelerating value and innovation," *Mc Kinsey Company*, vol. 2, no. 8, pp. 1–13, 2013.
- [32] S. LaValle, E. Lesser, R. Shockley, M. S. Hopkins, and N. Kruschwitz, "Big data, analytics and the path from insights to value," *MIT Sloan Manage. Rev.*, vol. 52, no. 2, pp. 21–32, Dec. 2011.
- [33] J. Stuart Ward and A. Barker, "Undefined by data: A survey of big data definitions," 2013, *arXiv:1309.5821*.
- [34] S. F. Wamba, S. Akter, A. Edwards, G. Chopin, and D. Gnanzou, "How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study," *Int. J. Prod. Econ.*, vol. 165, pp. 234–246, Jul. 2015.
- [35] T. Davenport et al., "Competing on analytics. Harvard business review," vol. 84, no. 1, p. 14, 2006.
- [36] A. McAfee and E. Brynjolfsson, "Big data: The management revolution," *Harvard Bus. Rev.*, vol. 90, no. 10, pp. 60–68, 2012.
- [37] P. M. Hartmann, M. Zaki, N. Feldmann, and A. Neely, *Big Data for Big Business? A Taxonomy of Data-Driven Business Models Used by Start-Up Firms*. Cambridge Service Alliance, 2014, pp. 1–29.
- [38] P. P. Tallon, R. V. Ramirez, and J. E. Short, "The information artifact in IT governance: Toward a theory of information governance," *J. Manage. Inf. Syst.*, vol. 30, no. 3, pp. 141–178, Dec. 2013.
- [39] S. Dhamodharavadhani and G. Rajasekaran, "Unlock different V's of big data for analytics," *Int. J. Comput. Sci. Eng.*, vol. 6, no. 4, pp. 183–190, 2018.
- [40] M. P. M. Sahoo and S. S. Rautaray, "The emergence of big data: A survey," *Int. J. Comput. Sci. Mobile Appl.*, vol. 6, pp. 23–32, 2018.
- [41] K. Sravanthi and T. S. Reddy, "Applications of big data in various fields," *Int. J. Comput. Sci. Inf. Technol.*, vol. 6, no. 5, pp. 4629–4632, 2015.
- [42] J. S. Hurwitz, A. Nugent, F. Halper, and M. Kaufman, *Big Data for Dummies*. Hoboken, NJ, USA: Wiley, 2013.
- [43] A. A. Hussien, "How many old and new big data V's characteristics, processing technology, and applications (BD1)," *Int. J. Appl. Innov. Eng. Manage.*, vol. 9, no. 9, pp. 15–27, 2020.
- [44] S. Priyadarshy, "The 7 pillars of big data," *Petroleum Rev.*, vol. 14, 2015.
- [45] A. Katal, M. Wazid, and R. H. Goudar, "Big data: Issues, challenges, tools and good practices," in *Proc. 6th Int. Conf. Contemp. Comput. (IC)*, Aug. 2013, pp. 404–409.
- [46] A. Gani, A. Siddiqi, S. Shamshirband, and F. Hanum, "A survey on indexing techniques for big data: Taxonomy and performance evaluation," *Knowl. Inf. Syst.*, vol. 46, no. 2, pp. 241–284, 2016.
- [47] A. A. Hussein, "Fifty-six big data V's characteristics and proposed strategies to overcome security and privacy challenges (BD2)," *J. Inf. Secur.*, vol. 11, no. 4, p. 304, 2020.
- [48] R. Kitchin, "The real-time city? Big data and smart urbanism," *Geo J.*, vol. 79, no. 1, pp. 1–14, 2014.
- [49] T. H. Silva, P. O. S. Vaz de Melo, J. M. Almeida, and A. A. F. Loureiro, "Large-scale study of city dynamics and urban social behavior using participatory sensing," *IEEE Wireless Commun.*, vol. 21, no. 1, pp. 42–51, Feb. 2014.
- [50] H. Guo, X. Li, W. Wang, Z. Lv, C. Wu, and W. Xu, "An event-driven dynamic updating method for 3D geo-databases," *Geo-Spatial Inf. Sci.*, vol. 19, no. 2, pp. 140–147, Apr. 2016.
- [51] P. A. Laplante, "Who's afraid of big data?" *IT Prof.*, vol. 15, no. 5, pp. 6–7, Sep. 2013.
- [52] J. A. Shamsi and M. A. Khojaye, "Understanding privacy violations in big data systems," *IT Prof.*, vol. 20, no. 3, pp. 73–81, 2018.
- [53] A. Bhardwaj, S. Bhattacharjee, A. Chavan, A. Deshpande, A. J. Elmore, S. Madden, and A. G. Parameswaran, "DataHub: Collaborative data science & dataset version management at scale," 2014, *arXiv:1409.0798*.

- [54] A. Panimalar, V. Shree, and V. Kathrine, "The 17 V's of big data," *Int. Res. J. Eng. Technol.*, vol. 4, no. 9, 2017.
- [55] C. Cartledge, "How many VS are there in big data," Tech. Rep., 2016.
- [56] M. M. Farooqi, M. A. Shah, A. Wahid, A. Akhunzada, F. Khan, N. U. Amin, and I. Ali, "Big data in healthcare: A survey," in *Applications of Intelligent Technologies in Healthcare*. Springer, 2019, pp. 143–152.
- [57] A. A. Hussein, "Using Hadoop technology to overcome big data problems by choosing proposed cost-efficient scheduler algorithm for heterogeneous Hadoop system (BD3)," *J. Sci. Res. Rep.*, pp. 58–84, Nov. 2020.
- [58] O. Aziz, M. S. Farooq, A. Abid, R. Saher, and N. Aslam, "Research trends in enterprise service bus (ESB) applications: A systematic mapping study," *IEEE Access*, vol. 8, pp. 31180–31197, 2020.
- [59] A. Khelifi, O. Aziz, M. S. Farooq, A. Abid, and F. Bukhari, "Social and economic contribution of 5G and blockchain with green computing: Taxonomy, challenges, and opportunities," *IEEE Access*, vol. 9, pp. 69082–69099, 2021.
- [60] A. Fernandez, E. Insfran, and S. Abrahão, "Usability evaluation methods for the web: A systematic mapping study," *Inf. Softw. Technol.*, vol. 53, no. 8, pp. 789–817, 2011.
- [61] Z. A. Barmi, A. H. Ebrahimi, and R. Feldt, "Alignment of requirements specification and testing: A systematic mapping study," in *Proc. IEEE 4th Int. Conf. Softw. Test., Verification Validation Workshops*, Mar. 2011, pp. 476–485.



OMER AZIZ received the M.S. degree in computer science from the National College of Business Administration and Economics, Lahore, Pakistan. He is currently pursuing the Ph.D. degree in computer science with the University of Management and Technology, Lahore. He is also working as a Lecturer at the Department of Computer Science, NFC Institute of Engineering and Technology, Multan. He has 14 years of professional experience in education and industry. He has developed software applications, websites, and mobile application for different companies around the globe. He has strong analysis and software architecture design skills according to emerging software market demand of data science, machine learning, cross platform, and artificial intelligence.



MUHAMMAD SHOAIB FAROOQ (Member, IEEE) is currently working as a Professor of computer science with the University of Management and Technology, Lahore. He is also an Affiliate Member at George Mason University, Fairfax, VA, USA. He possesses more than 24 years of teaching experience in the field of computer science. He has published many peer-reviewed international journal and conference papers. His research interests include theory of programming languages, big data, the IoT, the Internet of Vehicles, machine learning, and distributed systems and education.



ADEL KHELIFI received the Ph.D. degree from the Engineering School of High Technology, Canada, in 2005. He was the Dean of computer information technology, American University in the Emirates, Dubai, United Arab Emirates. He has held impressive past careers. He was a Lecturer with the Engineering School of Technology, Canada. He was also a United Nations MSF, Canada; the Ministry of Relations with the Citizen and Immigration, Canada; and the Ministry of Finances, Tunisia. He is currently a Faculty Member with Abu Dhabi University. He holds a high-level of knowledge and expertise. He is involved in prompting the open source software paradigm in the region. As a Canadian ISO Member in software engineering, he is contributing in developing software measurement standards. His research interest includes archeology.

• • •