

Big Data in Architecture

Matthew N. O. Sadiku

Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, TX, USA

Uwakwe C. Chukwu

Department of Engineering Technology, South Carolina State University, Orangeburg, SC, USA

Janet O. Sadiku

Juliana King University, Houston, TX, USA

Abstract:

Big data is a broad term for the collection of both structured and unstructured data. Data is one of the biggest byproducts of the 21st century. Almost everything we do produces data, from swiping credit cards to emailing. Our production of data is exploding. Increasingly, the value of a business is tied to its ability to mine data. In the field of architecture, data is having a similar impact. In the modern era of architecture, big data plays a crucial role in transforming design processes. Big data and architecture is one combination worth exploring in depth. This paper examines the role of big data in architecture.

Keywords: big data, big data analytics, architecture.

INTRODUCTION

Big data has been a buzzword in recent years. The increasing amount of data raises both the opportunities and the challenges of managing it. Figure 1 shows Google's data center in Mayes County, Oklahoma, USA, where data is stored [1]. When properly analyzed, big data can reveal patterns, trends, and relationships related to human behavior and interaction in the built environment. Today's organizations depend on data and insights to make most of their decisions.

Big data in architecture refers to the application of big data technologies and techniques within the field of architecture, where large volumes of data from building sensors, user interactions, and environmental factors are collected and analyzed to optimize building design, performance, and operations. It also refers to the extensive datasets collected from various technologies, such as IoT sensors, social media, and satellite imaging, which provide valuable insights into design practices and urban planning. In the architecture industry, big data can be utilized to analyze vast amounts of

information like building usage patterns, environmental data, construction costs, and design trends, allowing architects to make data-driven decisions for more efficient, sustainable, and user-centric building designs.

WHAT IS BIG DATA?

Big data applies to data sets of extreme size (e.g. exabytes, zettabytes) which are beyond the capability of the commonly used software tools. It involves situation where very large data sets are big in volume, velocity, veracity, and variability [2]. The data is too big, too fast, or does not fit the regular database architecture. It may require different strategies and tools for profiling, measurement, assessment, and processing.

Big Data is essentially classified into three types [3]:

- *Structured Data*: This is highly organized and is the easiest to work with. Any data that can be stored, accessed, and processed in the form of fixed format is known as a structured data. It may be stored in tabular format. Due to their nature, it is easy for programs to sort through and collect data. Structured data has quantitative data such as age, contact, address, billing, expenses, credit card numbers, etc. Data that is stored in a relational database management system is an example of structured data.
- *Unstructured Data*: This refers to unorganized data such as video files, log files, audio files, and image files. Any data with unknown form or the structure is classified as unstructured data. Almost everything generated by a computer is unstructured data. It takes a lot of time and effort required to make unstructured data readable. Examples of unstructured data include Metadata, Twitter tweets, and other social media posts.
- *Semi-structured Data*: This falls somewhere between structured data and unstructured data, i.e., both forms of data are present. Semi-structured data can be inherited such as location, time, email address, or device ID stamp.

The different types of big data are depicted in Figure 2 [4].

The process of examining big data is often referred to big data analytics. It is an emerging field since massive computing capabilities have been made available by e-infrastructures [5]. Big data analytics is the application of advanced analytic techniques to large, heterogeneous data sets that comprise structured, semi-structured, and unstructured data from many sources with sizes ranging from terabytes to zettabytes.

Analytics include statistical models and other methods that are aimed at creating empirical predictions. Data-driven organizations use analytics to guide decisions at all levels. Several techniques have been proposed for analyzing big data. These include the HACE theorem, cloud computing, Hadoop, and MapReduce [6].

CHARACTERISTICS OF BIG DATA

Big data is growing rapidly and expanding in all science and engineering, including physical, biological, and medical services. Different companies use different means to maintain their big data. As shown in Figure 3 [7], big data is characterized by 42 Vs. The first five Vs are volume, velocity, variety, veracity, and value .

- *Volume*: This refers to the size of the data being generated both inside and outside organizations and is increasing annually. Some regard big data as data over one petabyte in volume.
- *Velocity*: This depicts the unprecedented speed at which data are generated by Internet users, mobile users, social media, etc. Data are generated and processed in a fast way to extract useful, relevant information. Big data could be analyzed in real time, and it has movement and velocity.

- *Variety*: This refers to the data types since big data may originate from heterogeneous sources and is in different formats (e.g., videos, images, audio, text, logs). BD comprises of structured, semi-structured or unstructured data.
- *Veracity*: By this, we mean the truthfulness of data, i.e. whether the data comes from a reputable, trustworthy, authentic, and accountable source. It suggests the inconsistency in the quality of different sources of big data. The data may not be 100% correct.
- *Value*: This is the most important aspect of the big data. It is the desired outcome of big data processing. It refers to the process of discovering hidden values from large datasets. It denotes the value derived from the analysis of the existing data. If one cannot extract some business value from the data, there is no use managing and storing it.

On this basis, small data can be regarded as having low volume, low velocity, low variety, low veracity, and low value. Additional five Vs has been added [8]:

- *Validity*: This refers to the accuracy and correctness of data. It also indicates how up to date it is.
- *Viability*: This identifies the relevancy of data for each use case. Relevancy of data is required to maintain the desired and accurate outcome through analytical and predictive measures.
- *Volatility*: Since data are generated and change at a rapid rate, volatility determines how quickly data change.
- *Vulnerability*: The vulnerability of data is essential because privacy and security are of utmost importance for personal data.
- *Visualization*: Data needs to be presented unambiguously and attractively to the user. Proper visualization of large and complex clinical reports helps in finding valuable insights.

Instead of the 10V's above, some suggest the following 5V's: Venue, Variability, Vocabulary, Vagueness, and Validity) [8].

Industries that benefit from big data include the healthcare, financial, airline, travel, restaurants, automobile, sports, agriculture, and hospitality industries. Big data technologies are playing an essential role in farming: machines are equipped with sensors that measure data in their environment. Structured and unstructured data are generated in various types [10-12].

BIG DATA IN ARCHITECTURE

Big data in architecture and design leverages vast datasets to enhance and personalize user experiences, optimize product functionalities, and streamline design processes through data-driven insights. An increasing number of gadgets in the built environment, such as thermostats and refrigerators, are bolstering the Internet of things and relaying the data that they gather. Data has even manifested a physical presence. In New York, a new type of architecture is emerging in which large skyscrapers are being retrofitted into digital warehouses that accommodate computers rather than people. An example is 375 Pearl Street in New York, shown in Figure 4 [1]. Similar buildings are popping up across the United States for the purpose of storing and analyzing data.

When architects embark on a project, they gather data from a variety of sources, most often directly from the client. They determine which data points will help them create buildings that will have the greatest impact on their end users. They also tailor data analysis to every new project. Architects must gather as much data as possible in the early stages of the design process, because even small details can impact the efficiency of an entire building.

Data sources for big data in architecture include data on HVAC systems, lighting, and other building systems such as connected devices like smart meters and wearable technology. It also

includes environmental sensors that monitor air quality, temperature, humidity, and other environmental parameters. Key aspects of big data in architecture:

Data collection: Sensors embedded in buildings gather data on energy consumption, temperature, humidity, occupancy levels, lighting conditions, and more. Data collection is now part of our daily lives, from the temperature set on our thermostat to the time we walk down the sidewalk.

Data analysis: Advanced analytics tools are used to process and interpret the collected data, identifying patterns and trends to inform design decisions.

Data integration: Combining data from various sources (sensors, building management systems, weather data) can be complex.

Data is changing architecture in the following three ways [1]:

1. *Clients are demanding data from architects:* Clients are starting to ask architects to deliver more than just drawing sets. They are eyeing the data-rich BIM models that companies use to document projects as a way to supply data for downstream applications, such as facilities management. For architects, this means that their data needs to be as rigorous as their drawings.
2. *Clients are demanding data from buildings:* Clients have also become interested in the data generated by the buildings. This data enables building owners to measure and improve their facilities' performance quantitatively. Architects also need to recognize that clients are going to use this data to measure their own performance. Performance-based contracts, where a portion of an architect's fee is withheld until post-occupancy data validates its prescribed design performance, are gaining popularity.
3. *Data is changing the process as much as it changes the output:* The abundance of data may give rise to data warehouses, but the much more profound changes for architects will be procedural. If architects are to harness data from the built environment, even more significant procedural changes may be coming.

APPLICATIONS OF BIG DATA IN ARCHITECTURE

In architecture, big data applications can be used for various purposes including facility performance monitoring, predictive maintenance, energy optimization, space utilization analysis, design optimization through data-driven insights, construction project management, and even urban planning. Common applications of big data in architecture include the following [13]:

- *Building Performance Monitoring (BPM):* Sensors deployed throughout a building collect data on energy consumption, temperature, humidity, CO2 levels, and occupancy patterns, which can be analyzed to identify areas for improvement and optimize energy usage. This involves analyzing energy usage data to identify areas for improvement, optimizing HVAC systems, and adjusting lighting based on occupancy. It may monitor energy usage, indoor air quality, and occupant comfort levels in real-time to identify areas for improvement. Real-time data can be used to optimize building performance while studying occupants can be useful for user-centered design.
- *Predictive Maintenance:* By analyzing sensor data from building systems, potential failures can be predicted before they occur, allowing for proactive maintenance and preventing costly breakdowns. This may entail identifying potential issues in building systems before they occur, preventing costly breakdowns.
- *Space Utilization Analysis:* Data from occupancy sensors and building management systems can be used to understand how spaces are being used, leading to improved space planning and allocation.

- *Design Optimization:* Analyzing large datasets on climate patterns, building materials, and user behavior can inform design decisions to create more sustainable and user-friendly buildings. This may also involve analyzing user behavior to optimize building layout, lighting, and ventilation for occupant comfort. It may identify potential delays and optimize construction schedules using data from past projects.
- *City Planning:* Cities use big data to improve urban planning and urban development. Analyzing demographic data, traffic patterns, and environmental factors can help urban planners design more sustainable and resilient cities. This data-driven approach to urban planning can guide decision-making processes not only for architects and urban planners but also for city officials to create policies for more livable cities. New York City, for example, developed an initiative known as NYC Open Data, a public platform that provides over 1,300 datasets from housing and construction, education, health, and safety, to transportation. Figure 5 shows Manhattan in New York [13]. Another example is MX3D Bridge, a smart 3D-printed bridge in the City of Amsterdam to analyze pedestrian and crowd behavior. The smart bridge uses a sensor network to collect data from its users. The goal is to predict future behaviors to improve its 3D-printed manufacturing technique.
- *Smart City Design:* Heavily connected “smart cities” are becoming more prominent in society and are no longer the stuff of future-oriented dreams. Although some of them are built from scratch, other city officials bring in architects and similar professionals. Although smart cities rely on architects to design physical buildings, they also depend on data architects to develop big data-based infrastructures that coincide with the easily visible structures. Smart cities use big data for advertising purposes. They use data from traffic, pollution, and transport to optimize urban planning.
- *Synthesis Design:* Big data analytics was used to design an interior feature wall for the Watson Experience Center in San Francisco. They used data from the influence of mobile phones on monthly consumer spending to create a precise screen material that defines the wall. Figure 6 shows a synthetic wall design [14]. Without any explanation, the design gives little indication that it represents a massive data set.

BENEFITS

The application of big data in architecture brings numerous benefits that significantly influence the quality and efficiency of projects. In architecture, big data is used to improve design decisions, operational efficiency, and overall building performance. Integrating big data into architecture can lead to more informed, sustainable, and efficient design solutions. By collecting and analyzing data on past projects, construction companies can use predictive analytics to identify potential risks. Other benefits of big data in architecture include the following [13]:

- *Clash Detection:* With the help of big data, we can anticipate potential design problems before construction begins. For example, using building information modeling (BIM) it is now common practice to do collision detection to check all areas. This reduces costs for our clients and much less headaches for architects as there are fewer design changes during construction.
- *Energy Efficiency:* Another area where big data shines is improving energy efficiency. Using data sets of weather patterns, solar trajectories, wind direction, and strength, we can design better buildings that maximize natural light, ventilation, and heating for a more sustainable future.
- *Project Management:* Tracking progress on construction sites using sensors and data analytics can help identify potential delays, optimize resource allocation, and improve overall project efficiency.

- *Parametric Design:* Parametric design in architecture involves the use of parameter-based algorithms to design buildings that automatically adapt to changing variables, allowing for customizable and flexible structures.
- *Enhanced Decision-Making:* Data provides the foundation for evidence-based design decisions. Big data is transforming architectural design processes by enabling more informed decision-making through data-driven insights. It allows architects to analyze user behavior, environmental conditions, and urban trends, improving design accuracy and sustainability. The data-driven approach facilitates informed decision-making, reducing waste, and enhancing sustainability in architecture.
- *Anticipate the Future:* The architects of the past had to largely rely on guesswork when calculating things like population growth or the effects of climate change. Big data aids architects in understanding places and people, then incorporating the associated revelations into their designs. Datasets and modeling could help architects test the viability of design concepts before constructing them in real life.
- *Waste:* Waste is a major practical and political concern. It also becoming an increasingly public issue as news outlets publish pieces about pollution, plastics in the ocean, and whales killed by car parts. Figure 7 shows waste [15]. Recycling is viewed as the responsible way of dealing with waste. By embracing big data, however, businesses can streamline their disposal practices where it really matters. Data systems all play a role in moving the US towards being a zero waste system.
- *Sustainable Design:* Utilizing data to design buildings that minimize environmental impact through energy efficiency and resource optimization. Big data informs sustainable architectural design by analyzing environmental impacts, optimizing resource use, and predicting energy efficiency. It allows architects to assess site conditions, climate, and material performance to create eco-friendly buildings.

CHALLENGES

Although big data has tremendous impact on industries, it is not without challenges. The challenges of integrating big data into architectural design include managing and analyzing large volumes of data, ensuring data privacy and security, the need for specialized skills and technology, and bridging the gap between data insights and creative design processes. Opting for a big-data enabled data analytics solution is not straightforward. It requires vast technology land space for components to ingest data from numerous sources. It is also essential to have proper synchronization between these components. Keeping up with varying use cases of big data is a significant challenge for many organizations. Other challenges in using big data in architecture include the following [15]:

- *Data Security:* Although big data can provide great insight for decision-making, protecting data from theft is challenging.
- *Data Privacy:* Data collected may contain personal information. It is necessary to ensure user data is protected when collecting information about building occupants.
- *Data Visualization:* Presenting complex data insights in a clear and understandable way for architects and building owners.
- *Data Integration:* Combining data from various sources can be complex.
- *Data Quality:* Ensuring the accuracy and reliability of data is crucial.
- *Data Interpretation:* Another challenge is the interpretation and application of data. The architect are liable for making the final decision. Misinterpretation can lead to wrong decisions, leading to ineffective architectural solutions.

- *Data Storage:* Data volume remains a significant challenge because data volumes are doubling in size about every two years. Besides data size, the number of file formats used to store data is also growing. As a result, effectively storing and managing information is often a challenge for an organization.
- *Data Analytics:* Data analysis in architecture is a complicated process. Design teams must not only gather the most accurate data, but also analyze it and locate hidden patterns that predict how people will most efficiently use a building.
- *Data Quality:* Data quality aspects include accuracy, consistency, relevance, completeness, and use fitness. Data quality is a challenge anytime working with diverse data sources. Consequently, obtaining useful data requires a significant effort to clean the data before bringing it for analysis.
- *Learning New Data Tools:* Despite its transformative potential, it can be challenging for firms to incorporate big data into architecture due to the sheer volume and complexity of data. Architects often do not have the skill set required to understand complex data and algorithms. This means they have to constantly learn new tools to implement big data in our practice.
- *Scaling:* Big data solutions are used to handle large volumes of data. It can cause issues if the planned architecture is unable to scale. This may degrade the application performance and efficiency.
- *Complexity:* Big data systems can be challenging to implement since they must deal with various data types from various sources.
- *Collaboration:* This is essential during the design process, especially when architects and their clients gather data simultaneously. It is crucial to use data analysis tools that foster collaboration between all parties involved in the building's development.
- *Skillset:* Big data technologies are highly specialized, and they use frameworks and languages that are not common in more general application architectures. To operate these modern technologies and data tools, skilled data professionals are required. These will include data scientists, analysts, and engineers to operate tools and get data patterns. A shortage of data experts is one of the challenges that companies face.
- *Emerging Technologies:* Emerging technologies such as artificial intelligence, machine learning, and Internet of things are expected to play an increasingly important role in analyzing and interpreting big data. Architects need to learn about these technologies and harness the power of big data to become better designers of buildings and cities.
- *Cost:* This can also be a challenge when performing data analysis in architecture. Third-party data formatting can be a long, expensive process, depending on the scope of the project.

Some of these challenges are displayed in Figure 8 [16].

CONCLUSION

Data analytics in architecture is vital. Data analytics in architecture can be used to improve wayfinding at a facility. But in order to use data effectively, we must first understand what types of data to gather and how to analyze it properly. While integrating big data into architecture promises impressive benefits, it also brings unique challenges. Despite the challenges, big data is here to stay. Those who can harness big data will undoubtedly lead the charge toward more responsive and responsible built environments. More information about big data in the architecture industry can be found in the books in [18,19].

REFERENCES

1. D. Davis, "How big data is transforming architecture," April 2015,

- https://www.architectmagazine.com/technology/how-big-data-is-transforming-architecture_o
2. M. N. O. Sadiku, M. Tembely, and S.M. Musa, "Big data: An introduction for engineers," *Journal of Scientific and Engineering Research*, vol. 3, no. 2, 2016, pp. 106-108.
 3. "The complete overview of big data,"
<https://intellipaat.com/blog/tutorial/hadoop-tutorial/big-data-overview/>
 4. R. Allen, "Types of big data | Understanding & Interacting with key types (2024),"
<https://investguiding-com.custommapposter.com/article/types-of-big-data-understanding-amp-interacting-with-key-types>
 5. P. Baumann et al., "Big data analytics for earth sciences: The earthserver approach," *International Journal of Digital Earth*, vol. 19, no. 1, 2016, pp.3-29.
 6. X. Wu et al., "Knowledge engineering with big data," *IEEE Intelligent Systems*, September/October 2015, pp.46-55.
 7. "The 42 V's of big data and data science,"
<https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html>
 8. P. K. D. Pramanik, S. Pal, and M. Mukhopadhyay, "Healthcare big data: A comprehensive overview," in N. Bouchemal (ed.), *Intelligent Systems for Healthcare Management and Delivery*. IGI Global, chapter 4, 2019, pp. 72-100.
 9. J. Moorthy et al., "Big data: Prospects and challenges," *The Journal for Decision Makers*, vol. 40, no. 1, 2015, pp. 74–96.
<https://www.grandviewresearch.com/industry-analysis/industrial-wireless-sensor-networks-iwsmn-market>
 10. A. K. Tiwari, H. Chaudhary, and S. Yadav, "A review on big data and its security," *Proceedings of IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems*, 2015.
 11. M. B. Hoy, "Big data: An introduction for librarians," *Medical Reference Services Quarterly*, vol. 33, no 3. 2014, pp. 320-326.
 12. M. Viceconti, P. Hunter, and R. Hose, "Big data, big knowledge: Big data for personalized healthcare," *IEEE Journal of Medical and Health Informatics*, vol. 19, no. 4, July 2015, pp. 1209-1215.
 13. "The promises and pitfalls of big data in architecture," July 2024,
<https://slantis.com/blog/the-promises-and-pitfalls-of-big-data-in-architecture>
 14. "Big data becomes architecture in this CNC-milled screen wall for IBM," April 2017,
<https://www.archdaily.com/870095/big-data-becomes-architecture-in-this-cnc-milled-screen-wall-for-ibm>
 15. L. Alton, "Facing environmental crisis, data is reshaping waste management,"
https://www.smartdatacollective.com/facing-environmental-crisis-data-is-reshaping-waste-management/#google_vignette
 16. C. Gaur, "Top 9 challenges of big data architecture and its solutions | Overview," January 2023,
<https://www.xenonstack.com/insights/challenges-of-big-data-architecture>

17. M. N. O. Sadiku, U. C. Chukwu, and P. O. Adebo, *Big Data and Its Applications*. Moldova, Europe: Lambert Academic Publishing, 2024.
18. S. M. F. Akhtar, *Big Data Architect's Handbook: A Guide to Building Proficiency in Tools and Systems Used by Leading Big Data Experts*. Packt Publishing, 2018.



Figure 1. Google's data center [1].

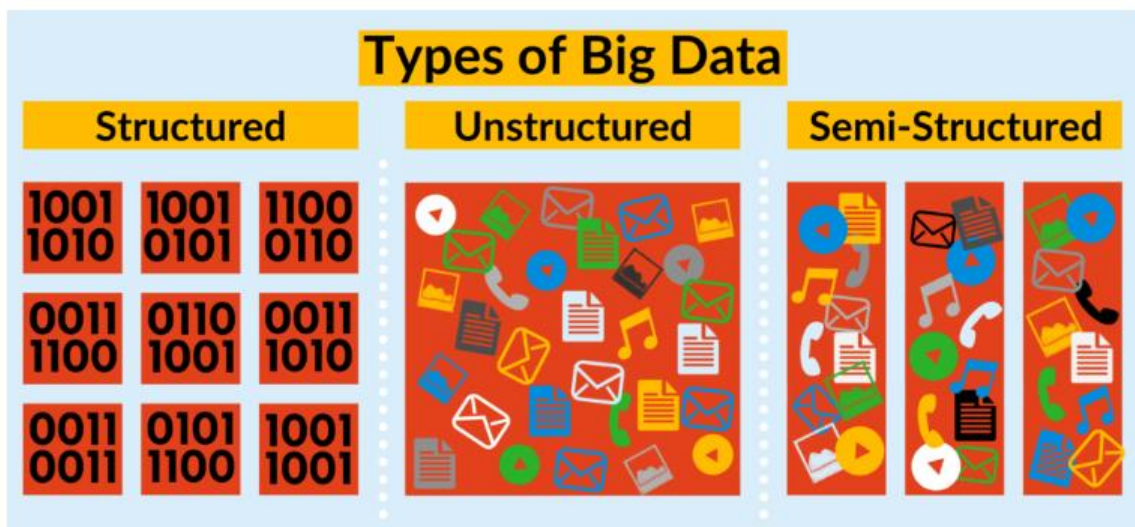


Figure 2. Types of big data [4].

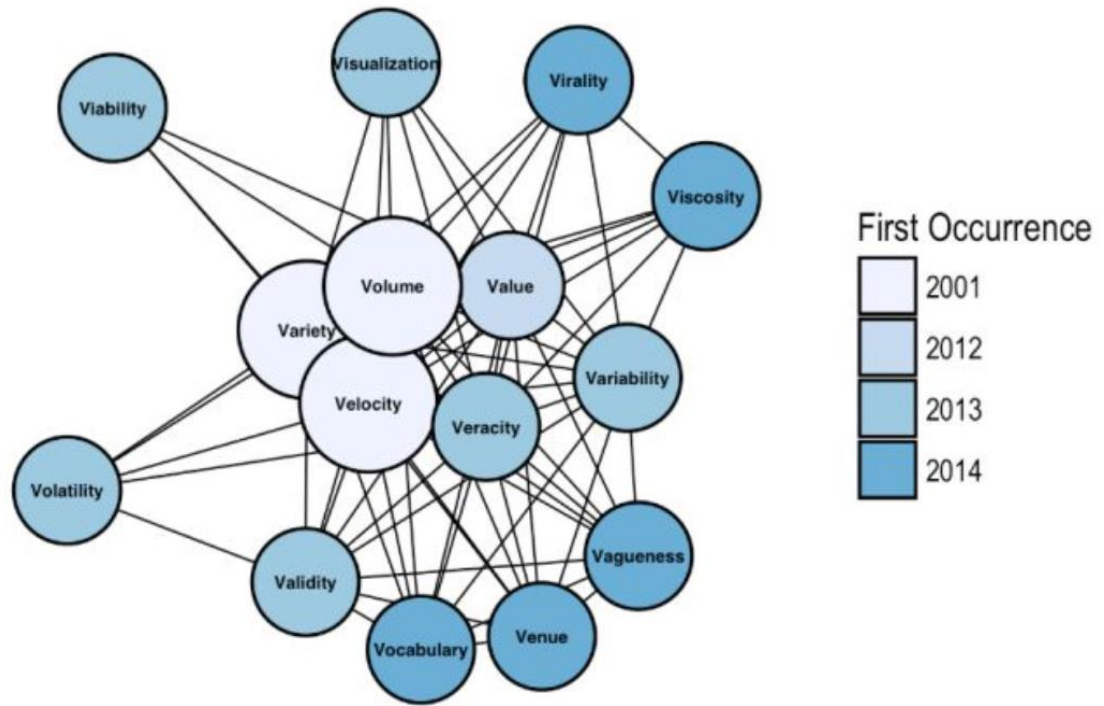


Figure 3. The 42 V's of big data [7].



Figure 4. A building in New York for storing and analyzing data [1].



Figure 5. Manhattan in New York [14].



Figure 6. A synthetic wall design [14].



Figure 7. Waste [15].

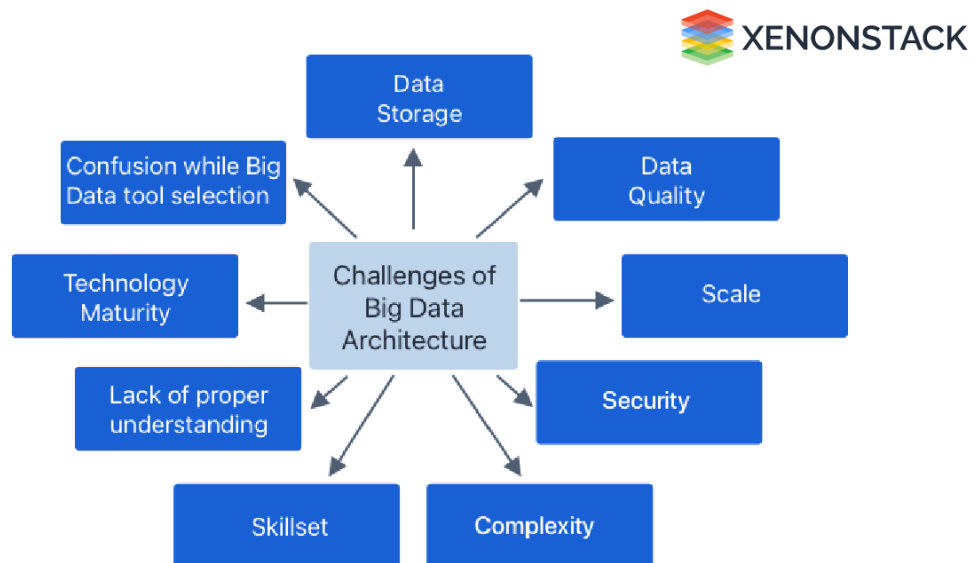


Figure 8. Some of the challenges of big data in architecture [16].