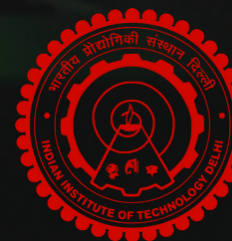


BIG DATA

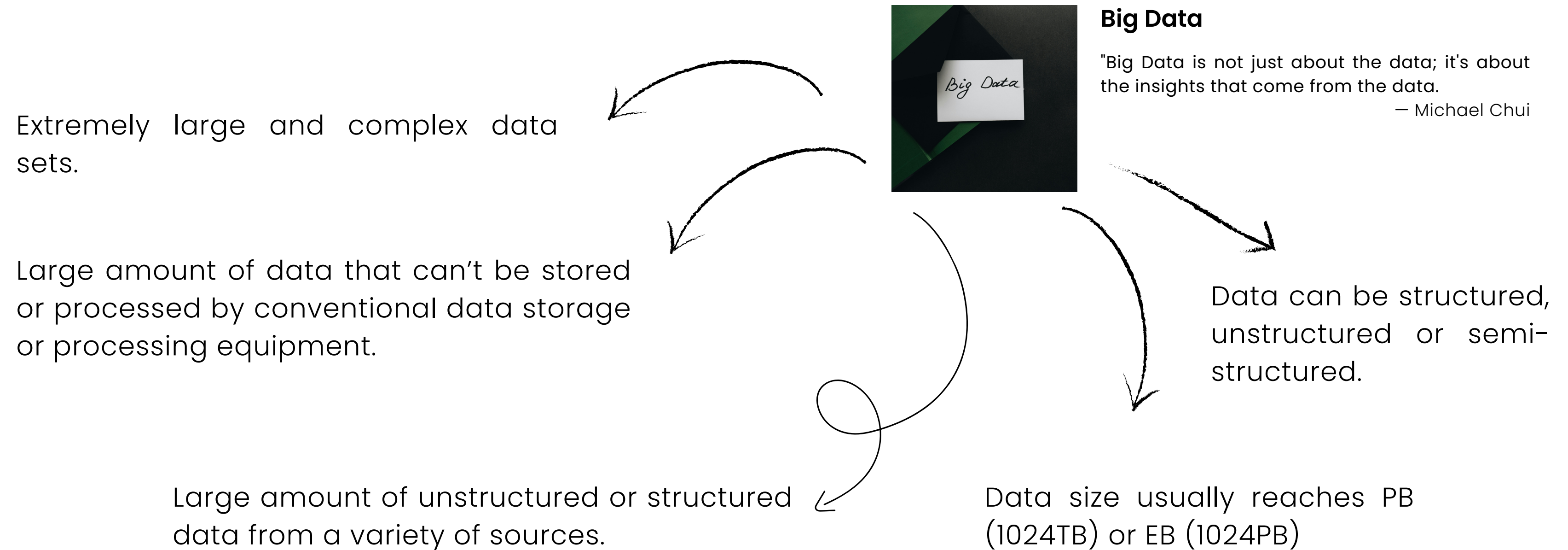
Presented by Deepali Patel
Library Intern,
Central Library,
IIT Delhi



**“Without big data, you are blind and deaf and
in the middle of a freeway.”**

– Geoffrey Moore

"Big Data in Library Science: Enhancing Operations and User Experience"



History of Big Data

1 First used by John Mashey

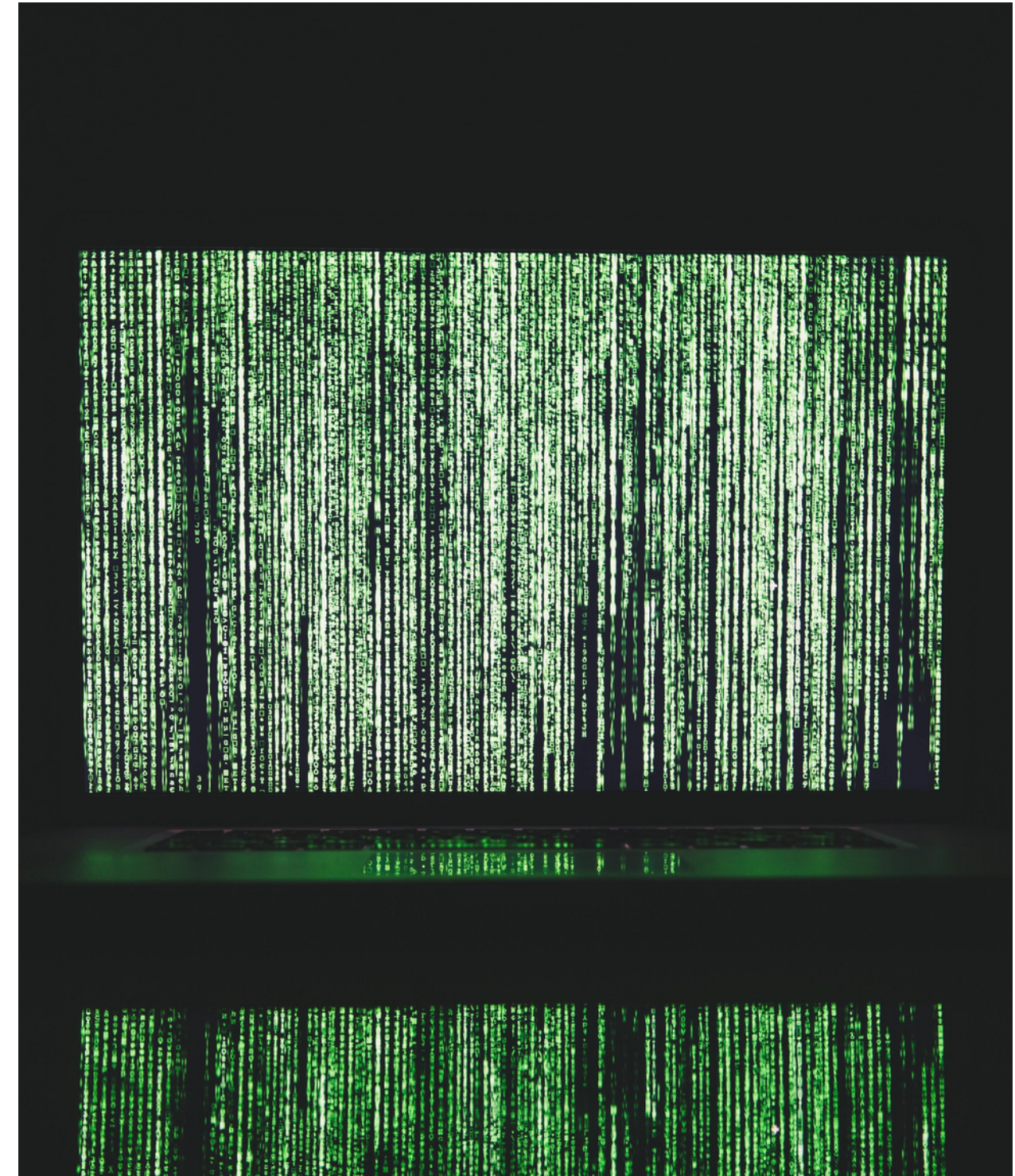
Refer to handling and analysing massive datasets.

2 "3V's": volume, velocity, and variety

big data gained strength in the early 2000s when Doug Laney gave 3V's.

3 "4V's": 3V's Veracity

The most common way of defining big data nowadays is by the "4V's" which adds veracity to the "3V's".





The evolution of Big Data



The Past

Around 2005, people began to realize just how much data users generated through Facebook, YouTube, and other online services. Apache Hadoop, an open source framework created specifically to store and analyze big data sets, was developed that same year. NoSQL also began to gain popularity during this time.

The Present

Users are still generating huge amounts of data but it's not just humans who are doing it. With the advent of the Internet of Things (IoT), more objects and devices are connected to the internet, gathering data on customer usage patterns and product performance. The emergence of machine learning has produced still more data.



The Future

While big data has come far, its value is only growing as generative AI and cloud computing use expand in enterprises.



CHARACTERISTICS OF **BIG DATA**

Two additional Vs have emerged over the past few years: value and veracity.

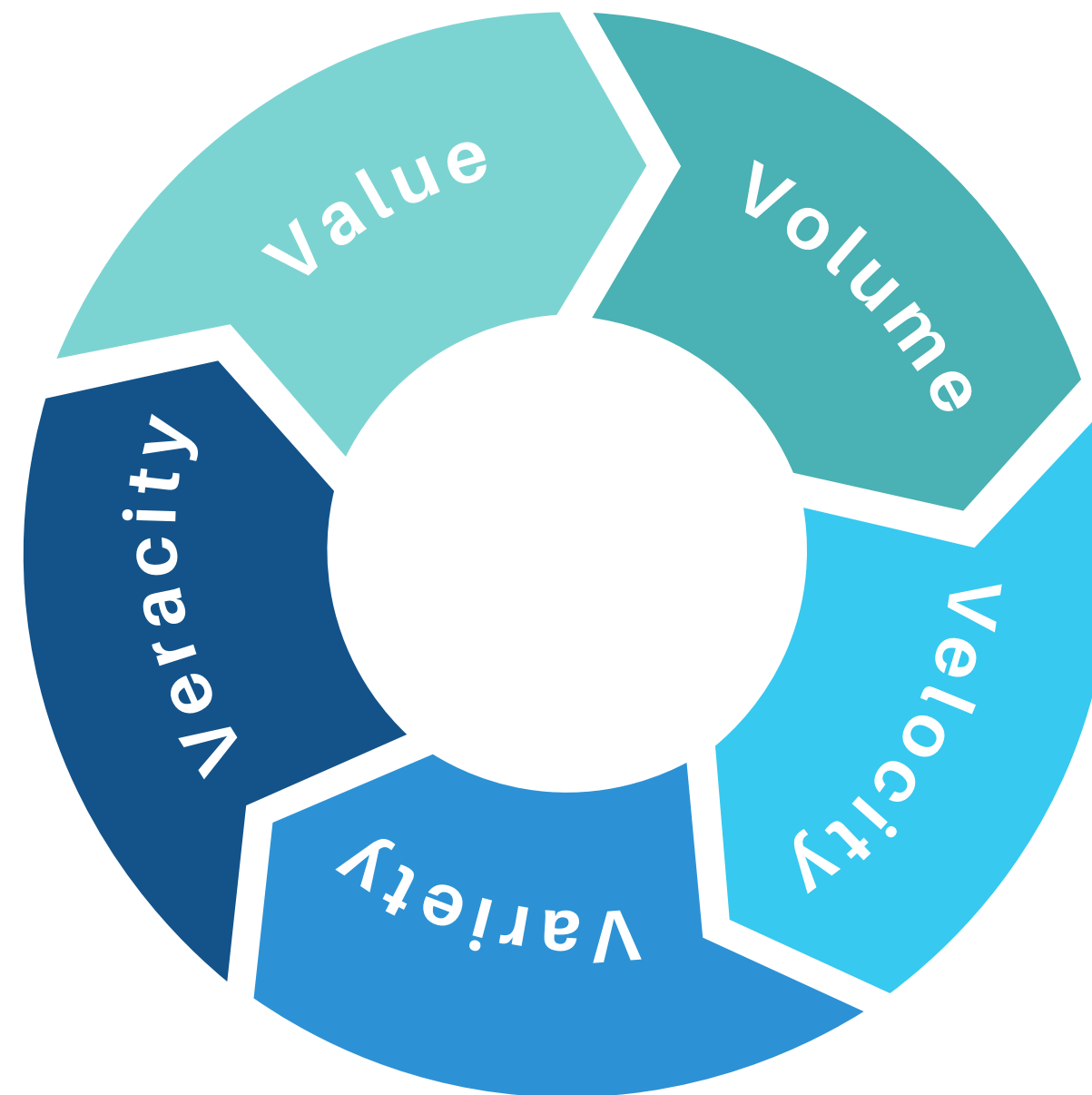


Veracity

Refers to the quality of the collected data. The data must have quality and produce credible results that enable action to produce right decision making. The veracity determines the accuracy of the data.

Value

Typically there is good information hidden amongst a larger body of non-traditional data; the challenge is identifying what is valuable and then transforming and extracting that data for analysis.



Volume

This is the primary characteristic of big data, refers to the large size (tera or petabytes) of records, transactions, tables, files, video, Web text, and sensor logs, etc.

Velocity

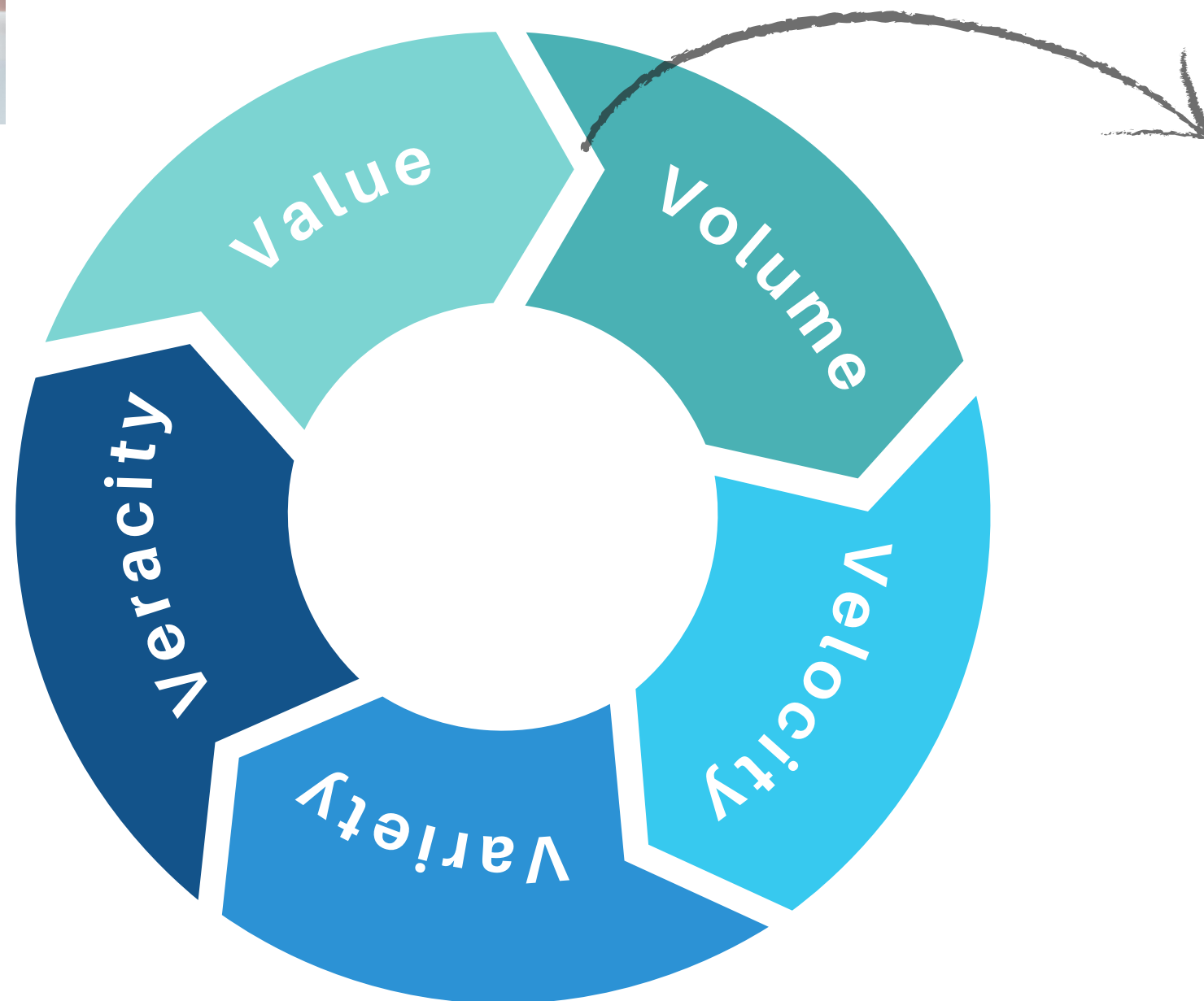
The rate at which data accumulates also influences whether the data is classified as big data or regular data.

Variety

The big data comes from a great variety of sources. Data of different formats, types, and structures need to be processed using efficient solutions.

CHARACTERISTICS OF **BIG DATA**

Two more Vs have emerged: variability and visualization.



Variability

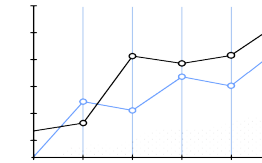
It's different from variety. With time the meaning, value, validity, visualization of same Big Data file changes, which have a huge impact on data homogenization.

Visualization

To understand enormous amount of information, proper visualization is critical. Using charts and graphs to visualize large amounts of complex data is much more effective in conveying meaning than spreadsheets and reports chock-full of numbers and formulas.

Importance of Big Data

Enables data-driven decision-making and personalized services.



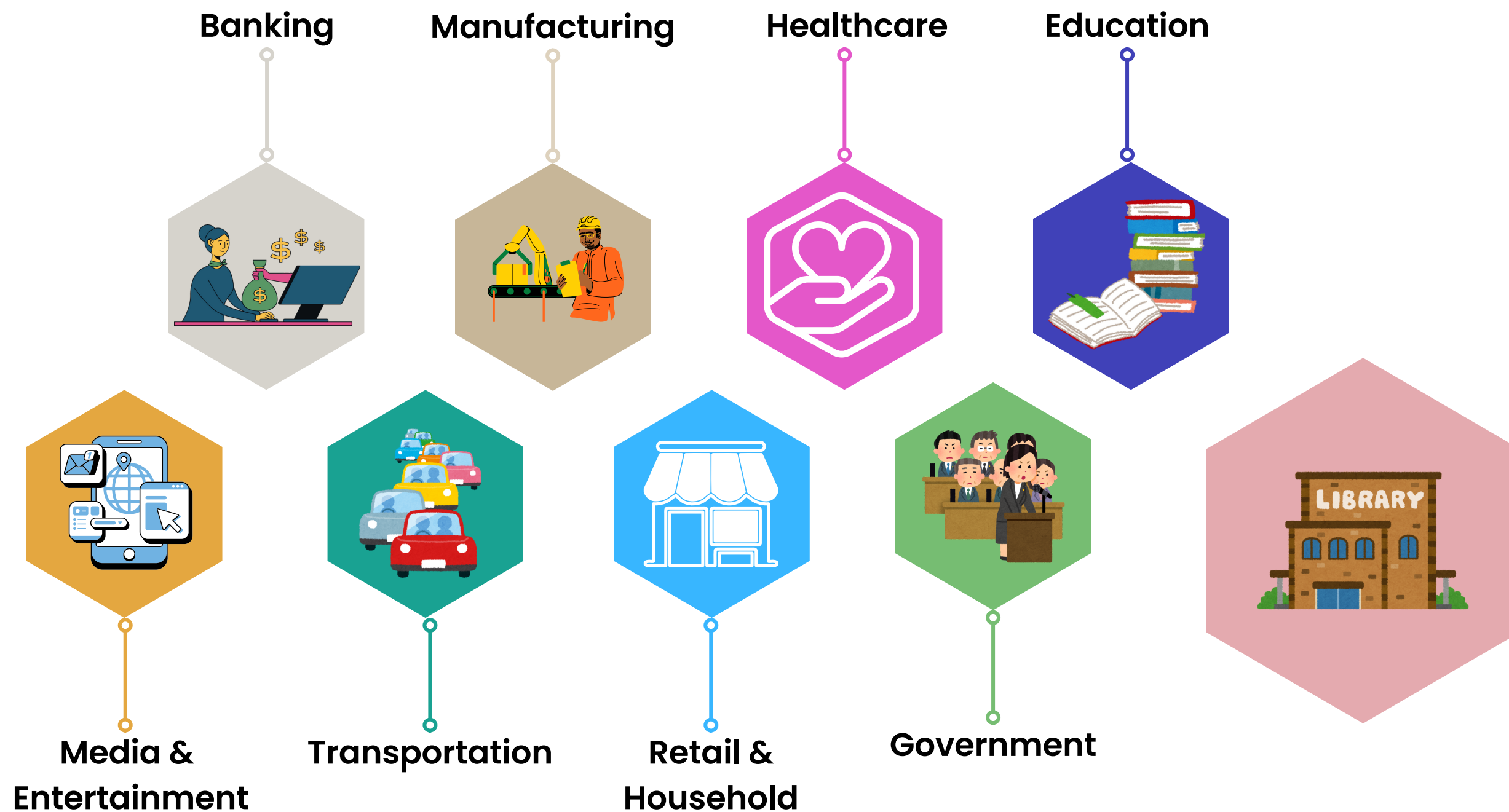
The importance of big data does not revolve around how much data you have, but what you do with it. You can take data from any source and analyze it to find answers that enable:

- Cost Reduction
- Time Reduction
- Innovation & Competitiveness
- Smart decision-making

When you combine big data with high-powered analytics, you can accomplish business-related tasks such as:

- Determining root causes of failures, issues, and defects in near-real time.
- Generating coupons at the point of sale based on the customer's buying/ habits.
- Recalculating entire risk portfolios in minutes.
- Detecting fraudulent behavior before it affects your organization.

Applications of Big Data in different **Sectors**



Big Data applications in libraries results in the creation of new knowledge and libraries that use big data analytics are more productive and efficient than others.

Applications of Big Data in Libraries

The benefits of Big Data in libraries are meant for management plans, supporting researchers, investing in the opportunities of Big Data and text-mining methodology.

According to Faisal (2015) big data and data analytics are intertwined. For good results to be obtained big data and analytics must be used together. The results also show that big data is mostly important to the researchers since they will make use of it to find problems faced by library used and suggest solutions.

There are many analytical methods that can be applied in the library and these include Visualization, Association Rule mining, Classifiers and Basic Statistical Procedures



Applications of Big Data indifferent **Academic Libraries**

As posited by Ball (2019):

1 Data sources

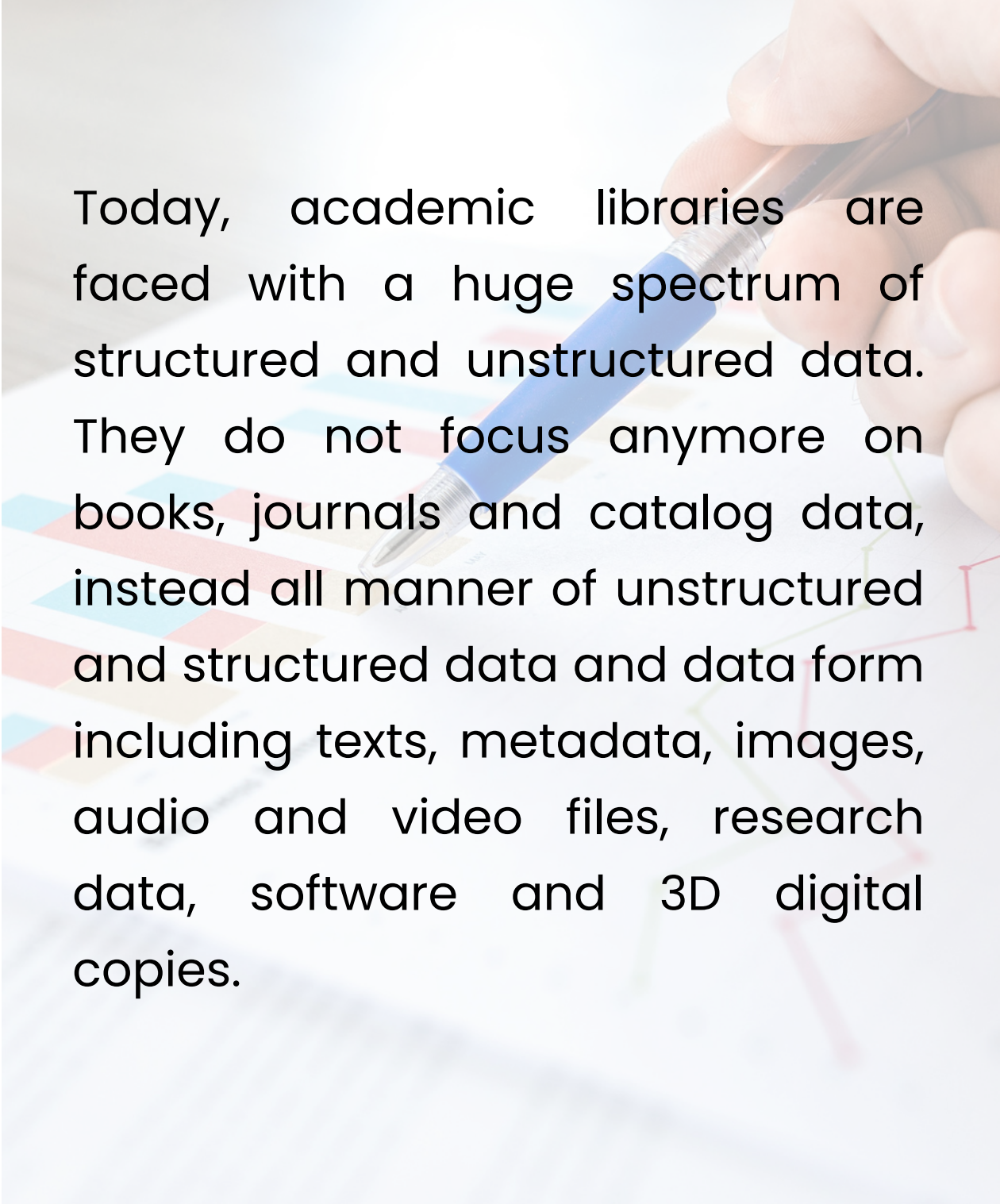
Data can now be studied either as a whole or in parts for data mining or use and other analysis methods.

2 Data analytics

Analytics involves the use of statistical techniques, information system software and operations research methodologies to explore, visualize, discover and communicate patterns or trends in data.

3 Data vizualization

There are reasons why data visualization tools should be used by academic libraries to communicate the results of data analytics.

A hand holding a blue pen is positioned over a document. The document features a line graph with a red line showing an upward trend. The background is a light blue and white gradient.

Today, academic libraries are faced with a huge spectrum of structured and unstructured data. They do not focus anymore on books, journals and catalog data, instead all manner of unstructured and structured data and data form including texts, metadata, images, audio and video files, research data, software and 3D digital copies.

Key Technologies Driving Big Data in 2025

The rise of big data in 2025 is supported by several cutting-edge technologies. These advancements have revolutionized how data is collected, stored, processed, and analyzed, driving innovation across industries.

1

Internet of Things (IoT)

By analyzing IoT data, businesses can gain insights into consumer behavior, enhance operational efficiency, and optimize systems for better performance.

2

Artificial Intelligence and Machine Learning (AI & ML)

Enable businesses to uncover hidden patterns, identify correlations, and predict future trends with remarkable accuracy.

3

Cloud Computing

Cloud platforms provide scalable storage and computing capabilities for big data.

4

Blockchain Technology

Adds value to big data by enhancing security, transparency, and data integrity.

Challenges in Adopting Big Data Technologies

1

Data Privacy Concerns

A lack of transparency in data collection and use can also lead to ethical concerns and damage customer trust.

2

Integration Complexity

Seamlessly integrating big data technologies with legacy systems is often complicated.

3

Skill Shortage

Businesses may struggle to find and retain talent capable of designing and managing complex data analytics systems.

4

Data Quality and Accuracy

Inconsistent, incomplete, or outdated data can lead to inaccurate insights and poor decision-making.

5

High Implementation Cost

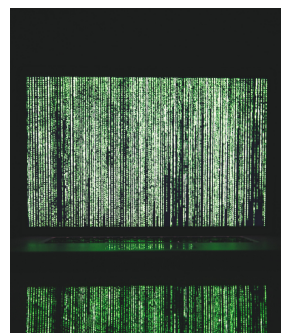
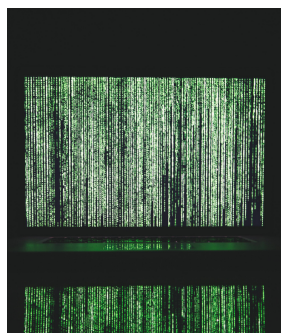
The initial investment in big data technologies, infrastructure, and skilled professionals can be prohibitively expensive for some organizations.

References :

1. Ashikuzzaman, M. (2024, June 4). The applications of big data in libraries. Library & Information Science Education Network. <https://www.lisedunetwork.com/the-applications-of-big-data-in-libraries/>
2. Bay Atlantic University. (2023, November 8). Characteristics of Big Data: Types & Examples – Bay Atlantic University – Washington, D.C. Bay Atlantic University – Washington, D.C. <https://bau.edu/blog/characteristics-of-big-data/>
3. Big data. (n.d.). Science Direct. Retrieved March 2, 2025, from <https://www.sciencedirect.com/topics/engineering/big-data>
4. BIG DATA AND ITS IMPORTANCE IN TODAY'S TECHNOLOGICAL LANDSCAPE. (2024, July 17). Association for Advancing Automation. Retrieved March 2, 2025, from <https://www.automate.org/news/big-data-and-its-importance-in-todays-technological-landscape-2>
5. Bornfight. (n.d.). 7 real-world examples of how brands are using Big Data analytics. Bornfight. <https://www.bornfight.com/blog/7-real-world-examples-of-how-brands-are-using-big-data-analytics/>
6. Chen, M. (2024, September 23). What is big data? Oracle India. <https://www.oracle.com/in/big-data/what-is-big-data/>

References :

7. Institute of Data. (2023, September 20). Why Big Data is Important: Exploring Its Benefits and Uses | Institute of Data. Institute of Data. <https://www.institutedata.com/blog/why-big-data-is-important/>
8. Kamupunga, W., Zhejiang University of Science and Technology, & Chunting, Y. (2019). Application of big data in libraries. In *International Journal of Computer Applications* (Vol. 178, Issue 16, pp. 34–35). <https://www.ijcaonline.org/archives/volume178/number16/kamupunga-2019-ijca-918955.pdf>
9. Panda, S. (2021). Usefulness and Impact of Big Data in Libraries: An Opportunity to Implement Embedded Librarianship. *Technological Innovations and Environmental Changes in Modern Libraries*, 45–60. <https://doi.org/10.31235/osf.io/fsau9>
10. Rice, M. (2024, November 6). 35 Big Data examples and applications. Built In. <https://builtin.com/articles/big-data-examples-applications>
11. Tella, A., & Kadri, K. K. (2020). Big data and academic libraries: is it big for something or big for nothing? *Library Hi Tech News*, 38(2), 15–23. <https://doi.org/10.1108/lhtn-07-2020-0069>



THANK YOU