Innovative: International Multi-disciplinary Journal of Applied Technology
(ISSN 2995-486X) VOLUME 02 ISSUE 10, 2024

Big Data in Space Exploration

Matthew N. O. Sadiku
Department of Electrical & Computer 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 plays a crucial role in space exploration by enabling scientists to analyze vast amounts of data collected from telescopes, satellites, and space probes. The role of big data in space exploration is undeniable, revolutionizing the way we observe, analyze, and understand the universe. Big data analysis influences space exploration by enabling us to better understand space data, and unlock the mysteries of the universe. Big data analytics has become an indispensable tool for space agencies and researchers worldwide. Utilizing big data analytics techniques, astronomers are now able to process immense datasets and uncover patterns and trends that would otherwise remain hidden. In this paper, we will delve into the fascinating role of big data in space exploration.

Keywords: big data, big data analytics, space exploration.

INTRODUCTION

The world has seen a digital revolution where more and more work is being conducted online. Storing this activity has led to the concept of big data, i.e., large datasets that are otherwise difficult to manage. Because of these characteristics, big data requires new technologies and techniques to capture, store, and analyze. The cloud word for big data is shown in Figure 1 [1]. Typical sources of big data are shown in Figure 2 [2]. Big data plays a central role in space exploration, where satellites, telescopes, rovers, and space probes collect and analyze colossal volumes of data from space and celestial bodies. It empowers the transformation of raw data into transformative knowledge. It meticulously sifts through terabytes of information, uncovering patterns, predicting celestial events, and optimizing missions. It is about pushing the boundaries of human understanding and propelling us further into the cosmic abyss [3].

Space exploration has always been a primary goal for humanity, pushing the boundaries of our understanding and opening up new frontiers. Space exploration, once guided primarily by telescopes and human curiosity, now relies heavily on data analytics to decipher the secrets of the cosmos. Figure 3 depicts the journey through the history of space data [4]. Data analytics is important in contemporary space missions, encompassing data collection, processing, analysis, and interpretation. It plays a fundamental role in spacecraft design and trajectory optimization by employing mathematical models, simulations, and optimization techniques. It supports communication, navigation, and lunar exploration through techniques like error correction codes, encryption, and machine learning. It is essential for tasks such as filtering, noise reduction, and anomaly detection, ensuring data integrity, and instrument reliability during processing.

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 [5]. 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 [6]:

- *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 4 [7].

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 [8]. 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 [9].

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 5 [10], big data is characterized by 42 Vs. The first five Vs are volume, velocity, variety, veracity, and value [2].

- *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. weather 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 [11]:

- *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) [12].

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 [13-15].

BIG DATA SPACE EXPLORATION

Space exploration has always been a realm of discovery, pushing the boundaries of human knowledge and technology. It has always been at the forefront of human ambition, pushing the boundaries of our understanding and opening up new frontiers. The vast unknown of space has captivated humanity for millennia, but unlocking its secrets demands more than celestial gazing.

The sheer volume of data has presented a new challenge for astronomers. How to process and analyze this massive amount of information effectively is not an easy task. Processing and making sense of big data from space is a monumental task. This is where big data comes into play. Big data analysis plays a crucial role in the study of dark matter and dark energy, which together constitute around 95% of the universe. Analyzing big data from space has transformed the field of astronomy, enabling scientists to unlock the secrets of the universe.

Today, data analytics in space industry plays a pivotal role, acting as the bridge between raw information and profound knowledge. Data is the backbone of space exploration, and with advancements in technology, we are now generating an unprecedented amount of data. From analyzing planetary formation data to predicting space weather patterns, big data plays a crucial role in virtually every aspect of space exploration. Data analytics has become a pivotal tool, transforming how we explore, understand, and utilize the cosmos. It enables continuous monitoring of astronaut health, analyzing biometric data to detect potential health issues, optimize life support systems, and ensure astronaut safety. It enhances deep space communication, optimizing data transmission and reducing communication delays. Data analytics in space exploration is illustrated in Figure 6 [16].

More than 50 years ago, space exploration was a race to the moon between the US and the former Soviet Union. In 1998, the International Space Station launch marked a highlight for international cooperation in space between the U.S., Russia, Japan, Canada and Europe. Today, more countries like as India and China are pursuing their own goals in space, which could challenge the US as a world leader. Americans continue to see an essential role for the United States as a leader in space exploration. Most Americans agree that it is essential for US to be a leader in space exploration.

NASA is engaged in a wide range of activities in space, including exploration and applied and basic research. Since its creation in 1958, NASA has achieved some pretty impressive feats of science. When NASA was created, it had to invent the technology to get where we needed to go, and we will continue to push the boundaries of technology into the future. New emerging technologies that open opportunities for research and exploration with minimal investments include NASA's satellites, such as typically shown in Figure 7 [17].

There is far less public urgency for NASA to send humans to the moon or Mars and to search for other planets that could support life. Americans by and large view NASA as critical to space exploration [18]. NASA collects, stores, processes, and analyzes data. It uses a system called the Mission Data Processing and Control System (MPCS) to manage and process data. The application for the NASA's Mars Curiosity rover uses production data and creates an easy way for engineers to monitor instruments and measurements faster and more efficiently. Elasticsearch permits handling large amounts of data from Curiosity's sensors, such as the temperature on the Martian surface. Figure 8 shows the artist's concept of NASA's X-57 'Maxwell' aircraft [19], while Figure 9 displays NASA's gateway in lunar orbit [19]. The X-57 will be the first all-electric X-plane and will be flown to demonstrate the benefits that electric propulsion may yield for the future of aviation. NASA will continue to be a global leader in scientific discovery. It has continued to push the boundaries of knowledge to deliver on the promise of American ingenuity and leadership in space. Its future will continue to be a story of human exploration, technology, and science.

APPLICATIONS OF BIG DATA SPACE EXPLORATION

Big data analytics in space exploration unlocks patterns within volumes of observations to advance astrophysics, guide future missions, and reveal our universe's deepest mysteries. The cosmos may be vast, but through data, it becomes knowable. Scientists and engineers use big data for everything from predicting weather on earth to monitoring ice caps on Mars to searching for distant galaxies. Here are some ways that big data is used in space exploration [20]:

- *Astronomy:* Astronomy has always been a data-intensive science. Data science has significantly transformed the field of space exploration, particularly in the realm of astronomy. By leveraging data science techniques, researchers are able to tackle complex astrophysical problems, uncovering hidden patterns and shedding light on the universe's origin, evolution, and large-scale structure. Data science techniques can help researchers understand the universe's origin, evolution, and structure. For example, machine learning algorithms can classify galaxies and their mergers, while image recognition algorithms can identify celestial objects.
- *Predicting Space Weather*: Space weather can impact satellites, spacecraft, and even power grids on earth. Big data analysis plays a crucial role in predicting space weather patterns by

analyzing vast amounts of data from satellites, ground-based observatories, and space weather models. The predictions enable space agencies and operators to take proactive measures to protect spacecraft and astronauts and mitigate the effects of space weather on critical infrastructure. Space exploration models predict space weather events that could severely impact satellites and astronauts in space. Crucial data-driven analytics will forecast perilous space weather to ensure mission success.

- *Astronauts' Health:* Radiation exposure poses significant risks to astronauts' health during space missions, especially during long-duration flights beyond Earth's protective magnetosphere. Big data analysis of radiation monitoring data collected from spacecraft, space stations, and dosimeters worn by astronauts helps scientists assess exposure levels and mitigate health risks. This data-driven approach enables space agencies to optimize mission planning, crew scheduling, and spacecraft design to ensure the safety and well-being of astronauts.
- *Earth Observation:* Data analytics plays a crucial role in interpreting satellite data for earth observation. The Sun's activity, characterized by phenomena such as solar flares, coronal mass ejections, and sunspot cycles, influences space weather conditions in the solar system. Satellites play a crucial role in monitoring and studying earth's environment, climate, and natural resources from space. Big data analysis of satellite imagery and remote sensing data provides valuable insights into changes and trends in earth's surface, atmosphere, and oceans. By analyzing multi-spectral imagery, radar data, and other earth observation datasets, scientists can track deforestation, urbanization, pollution, and other environmental changes over time. Information derived from this analysis supports decision-making in areas such as land use planning, disaster response, and natural resource management. Big data analysis of cosmic ray data collected by ground-based detectors and space-based observatories helps scientists study the origins, propagation, and interactions of cosmic rays. Free earth-observation datasets from the European Space Agency, NASA and the US Geological Survey have helped organizations around the world develop object-detection platforms.
- *Space Telescope Data:* Modern telescopes equipped with sophisticated sensors and cameras capture detailed images of celestial objects. They generate an astronomical amount of data. Space telescopes equipped with advanced instrumentation enable astronomers to search for signs of extraterrestrial life beyond our solar system. For example, the Hubble Space Telescope has collected over 150 terabytes of data throughout its mission. Analyzing this vast amount of data manually would be an impossible task. Big data analysis of space telescope data, including spectroscopic observations and transit surveys, helps scientists identify potentially habitable exoplanets and study their atmospheres for biosignatures.

BENEFITS

Many analysis tools offer fast analysis and visualization of data, and have proven to be very valuable. They help space probes make faster and better decisions, reduce failure, and improve life on planet earth. Big data analysis of data collected from space probes, telescopes, and simulations allows researchers to identify patterns and trends that reveal the underlying mechanisms of planetary formation. Data analytics enhances deep space communication, optimizing data transmission, and reducing communication delays. As space missions grow more complex, data scientists play a vital role in ensuring success. One of the significant advantages of data science in astronomical research is the ability to unveil the secrets of the universe. Other benefits include [20]:

- *Decision Making:* From mission design to resource allocation, data science in space industry, empowers informed decision-making throughout exploration endeavors, maximizing efficiency, minimizing risks, and paving the way for future discoveries. Space agencies are using big data tools to rapidly analyze their data and make time-efficient decisions. Data science helps make informed decisions throughout space exploration, from mission design to resource allocation. The ability to tell stories with data is paramount for inspiring decision makers.

- *Farming:* The Climate Corporation uses their satellite data to enable farmers all around the globe to find more sustainable ways to grow substantially more food. Farmers can use image data to better understand what factors affect the growth of crops. There are factors that can be detected from space, such as weather patterns, exposure to sunlight, air quality or pest activity.

- *Monitoring:* Spacecraft monitor everything from our home planet to faraway galaxies, beaming back images and information to earth. Spacecraft operates in extreme conditions, making health monitoring a critical aspect of space missions. Wearable sensors and advanced analytics are transforming how astronaut health is monitored during space missions. By continually tracking vital signs, biometrics, and even microbial flora in the body, subtle shifts can trigger earlier medical alerts than relying on scheduled tests or self-reported issues alone.

- *Autonomous Systems:* Data science is instrumental in developing autonomous systems for space exploration. Autonomous robots and spacecraft rely on data analysis to make informed decisions and adapt to dynamic environments. By integrating data science techniques with robotics, scientists can develop intelligent systems capable of exploring distant planets.

- *Simulation:* Data science helps in simulating and modeling complex astronomical scenarios, allowing astronomers to test theories and hypotheses. These simulations aid in understanding celestial processes and predicting the behavior of astronomical objects under different conditions.

- *Communication*: Data analytics improves communication reliability and data quality by analyzing signal patterns, noise, and interference. Releasing public domain images of a lonely blue planet or footage of astronauts on spacewalks elegantly communicate their accomplishments to the public while inspiring future engineers and scientists.

- *Mission Planning:* Predicting celestial events becomes more accurate through the use of data analytics. Data analytics helps optimize mission trajectories and improve spacecraft design by analyzing past mission data, current space weather conditions, and predictive models.

- *Data Processing:* Data analytics transforms raw data from onboard instruments into meaningful information using signal and image processing, data mining, and pattern recognition. Data analytics in space exploration is pivotal for processing the enormous volumes of data transmitted back to earth by space probes, satellites, and telescopes. Big data analytics techniques are essential for managing, processing, and extracting meaningful information from this data deluge.

- *Data Quality:* Astronomy heavily relies on observational data to study celestial objects. By applying data analytics techniques, filtering out noise and extracting relevant information enhances the quality of data. Data science improves the precision and accuracy of observational data collection.

- *Improved Efficiency:* By automating data analysis processes, astronomers can focus more on the interpretation of results and the formulation of new theories. This leads to a significant increase in efficiency and productivity.

- *Space Photography:* Photographs are universal means of communication. Over time, photography techniques grew from pointing a camera out of a window to using powerful mirrors and telescopes to capture impossibly distant subjects. To share these images with the world, NASA needed to solve photographic and data challenges. The space photography is crucial to communicating the mission of NASA. From the beginning, NASA understood that space photography played a vital role in its missions. The data from satellites and cameras is captured and stored locally.

- *Space Tourism*: Most adults in US expect that people will routinely travel in space as tourists in the next 50 years. More than half expect space tourism to become routine. However, Americans are not enthusiastic about traveling to space themselves.

CHALLENGES

Data transmission in space is still a challenge and finding a way to ease that process would be beneficial for many space missions. Some challenges associated with data analytics in space exploration include data accuracy, privacy, ethical use of AI, and making ai algorithms compatible with harsh space environments. Space missions generate enormous volumes of data, presenting significant challenges in data storage, processing, and analysis. Other challenges include [21]:

- *Space Debris:* As Americans look to the future of space, a large share expect problems with human-made debris. There will be a major problem with debris in space from rockets, satellites, and other human-made objects. Ratings are more mixed when it comes to the job private companies are doing to limit debris from objects like rockets and satellites.
- *Safety:* Space debris poses a growing threat to spacecraft and satellites in orbit. Big data analysis of space debris tracking data helps space agencies and operators predict debris trajectories and assess collision risks.
- *Communication:* Communicating from the ground station to a spacecraft is a complex challenge, largely due to the extreme distances involved. When data are transmitted and received across thousands or millions of miles, the delay and potential for disruption or data loss is significant. Transferring data also becomes more challenging because the signal gets weaker due to distances involved. Advanced communication technologies are required to enhance deep space exploration for both robotic and human missions. Optical communications technologies can dramatically improve communications between spacecraft and earth far better than radio communications.
- *International Collaboration*: Space exploration and scientific research are inherently collaborative endeavors that require cooperation among nations and space agencies worldwide. Big data analysis of space data often involves international collaboration, with researchers sharing data, resources, and expertise to address complex scientific questions and challenges. International collaborations enable scientists to tackle fundamental questions in astronomy, planetary science, and space physics more effectively.
- *Risk:* Whenever an astronaut steps out of a space vehicle, it is an incredible, high-level-risk activity. We need to critically think about the risk to human life and the value return for assigning humans to do activities. Smart autonomous systems have the capability to improve the efficiency and speed while minimizing the risk and cost involved with planetary missions.

CONCLUSION

Space data is about to get big. Big data has revolutionized the field of space technology, enabling us to explore the universe with unprecedented precision and efficiency. The integration of big data analytics has revolutionized the way we explore and understand the cosmos. It will play an increasingly vital role in shaping the future of space exploration. As technology evolves, the future holds even greater opportunities for data analytics in revolutionizing our understanding of the cosmos. By combining powerful data analytics techniques with increasingly vast amounts of data, astronomers are revolutionizing the way they conduct research, analyze celestial objects, and make groundbreaking discoveries.

Data analytics is not just transforming space exploration; it is redefining our understanding of the universe. This versatile tool has been a cornerstone of space missions, ensuring efficiency and success in the exploration of space and celestial bodies.

It provides astronomers with the tools and techniques necessary to analyze the massive volumes of data generated by telescopes and satellites. Data-driven intelligence provides the analytical foundation needed to realize humanity's boldest visions. As we continue to push the frontiers of space exploration, there is no doubt that big data will remain a cornerstone of astronomical research. More information about big data in space can be found in the books in [22,23].

REFERENCES

1. R. Delgado, "The challenges of bringing BYOD to the military,"

   https://socpub.com/articles/the-challenges-of-bringing-byod-to-the-military-11272

2. J. Moorthy et al., "Big data: Prospects and challenges," *The Journal for Decision Makers*, vol. 40, no. 1, 2015, pp. 74–96.

   "10 Uses of data science in space exploration," March 2024,

   https://medium.com/@analyticsemergingindia/10-uses-of-data-science-in-space-exploration-d958045213c2

3. "Navigating uncharted space: Harnessing data analytics in space exploration," October 2023,
4. https://www.dasca.org/world-of-data-science/article/navigating-uncharted-space-harnessing-data-analytics-in-space-exploration
5. 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.
6. "The complete overview of big data,"

   https://intellipaat.com/blog/tutorial/hadoop-tutorial/big-data-overview/

7. 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

8. 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.
9. X. Wu et al., "Knowledge engineering with big data," *IEEE Intelligent Systems*, September/October 2015, pp.46-55.
10. "The 42 V's of big data and data science,"

    https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html

11. 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.
12. 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-iwsn-market

13. 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.

14. M. B. Hoy, “Big data: An introduction for librarians,” *Medical Reference Services Quarterly*, vol. 33, no 3. 2014, pp. 320-326.

15. 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.

16. T. Ebrahimi, “ Dispatches from deep space: Imaging data and space photography,” July 2022,

    https://blog.westerndigital.com/space-photography/

17. D. Kekare, “Data analytics in space exploration: Beyond earth's boundaries,” March 2024,

    https://www.linkedin.com/pulse/data-analytics-space-exploration-beyond-earths-durgesh-kekare-v5iwf#:~:text=Data%20analytics%20enhances%20deep%20space,mission%20control%20and%20scientific%20research.

18. “Americans’ views of space: U.S. role, NASA priorities and impact of private companies,” July 2023,

    https://www.pewresearch.org/science/2023/07/20/americans-views-of-space-u-s-role-nasa-priorities-and-impact-of-private-companies/

19. “60 Years and counting - The future,”

    https://www.nasa.gov/specials/60counting/future.html

20. “Exploring the role of big data in space exploration,” March 20 https://1ansah.in/blog/exploring-the-role-of-big-data-in-space-exploration/#:~:text=The%20role%20of%20big%20data%20in%20space%20exploration%20is%20undeniable,space%20agencies%20and%20researchers%20worldwide.24,

21. D. Werner, “Big data, advanced algorithms and new approaches for space missions,” October 2024,

    https://spacenews.com/big-data-advanced-algorithms-and-new-approaches-for-space-missions/

22. M. N. O. Sadiku, U. C. Chukwu, and P. O. Adebo, *Big Data and Its Applications*. Moldova, Europe: Lambert Academic Publishing, 2024.

23. A. Agarwal, *Interplanetary Data Governance: Navigating the New Digital Frontier in Space*. Kindle Edition, 2024.

Figure 1. The cloud word for big data [1].

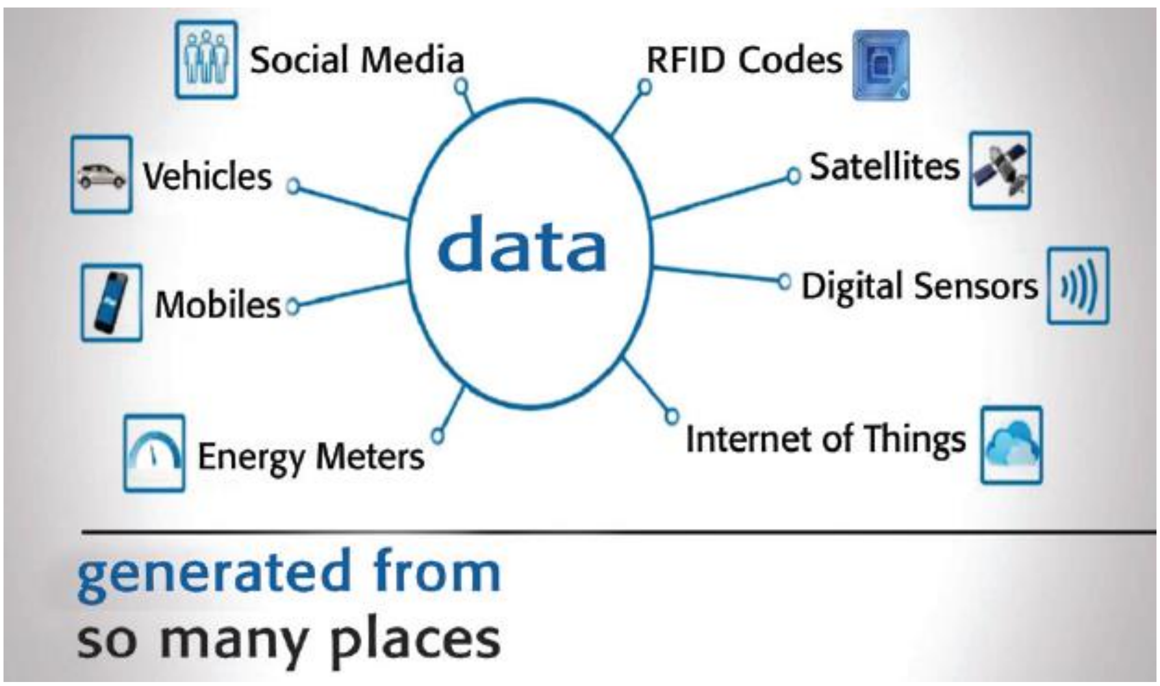

Figure 2. Typical sources of big data [2].

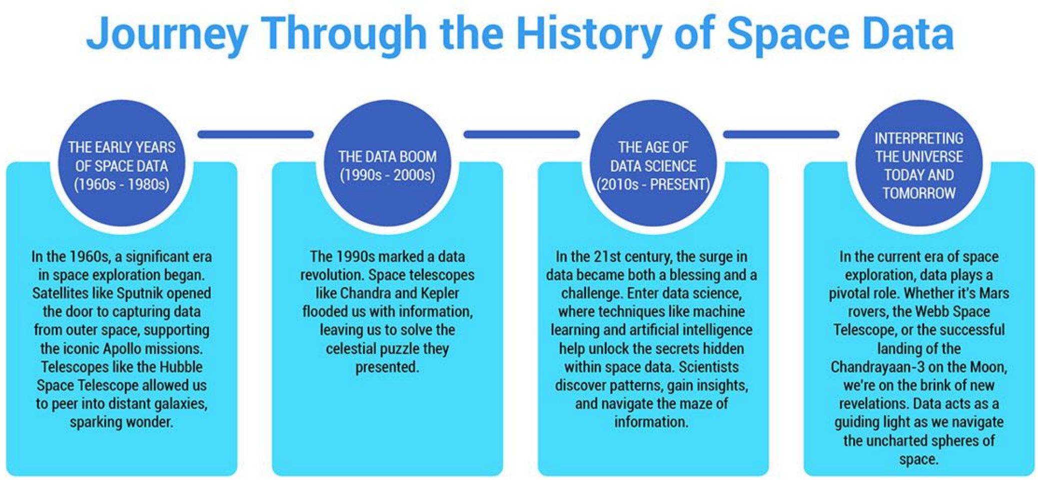

Figure 3. Journey through the history of space data [4].

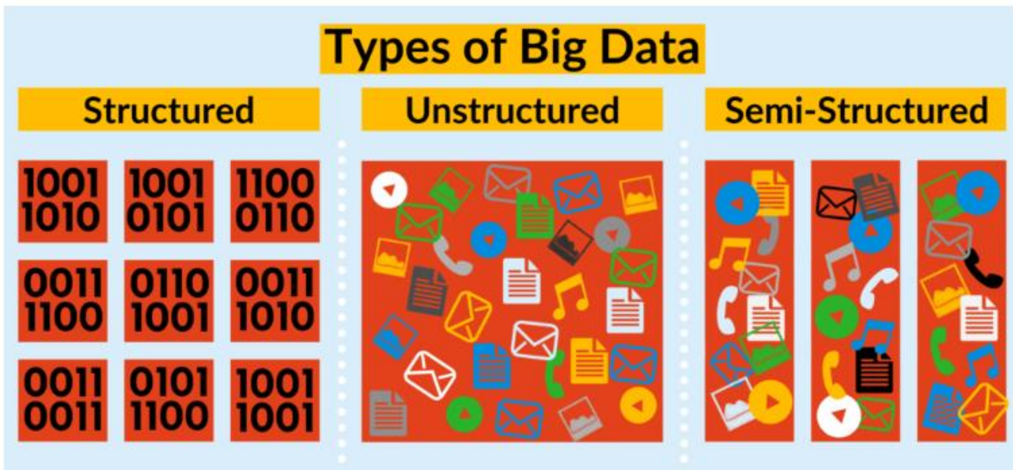

Figure 4. Types of big data [7].

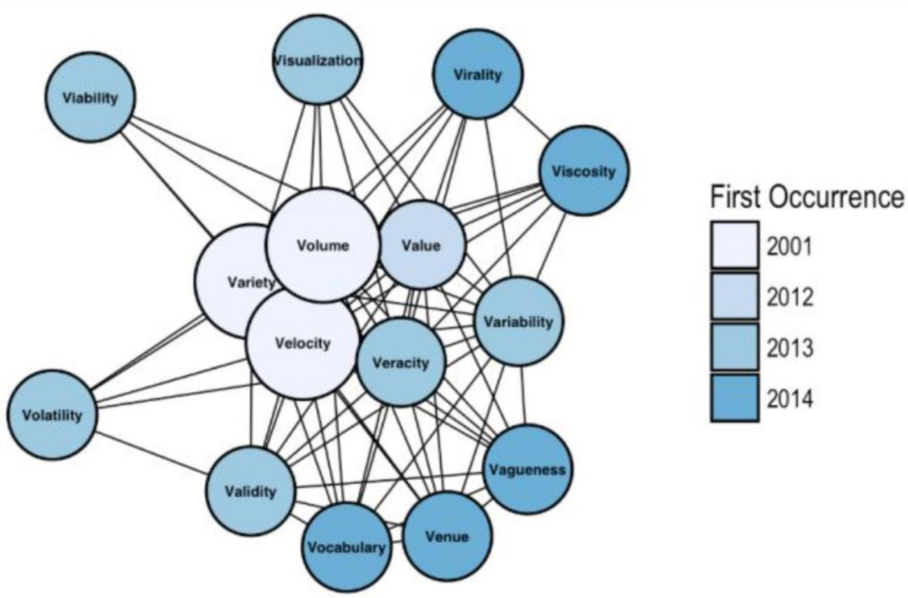

Figure 5. The 42 V's of big data [10].

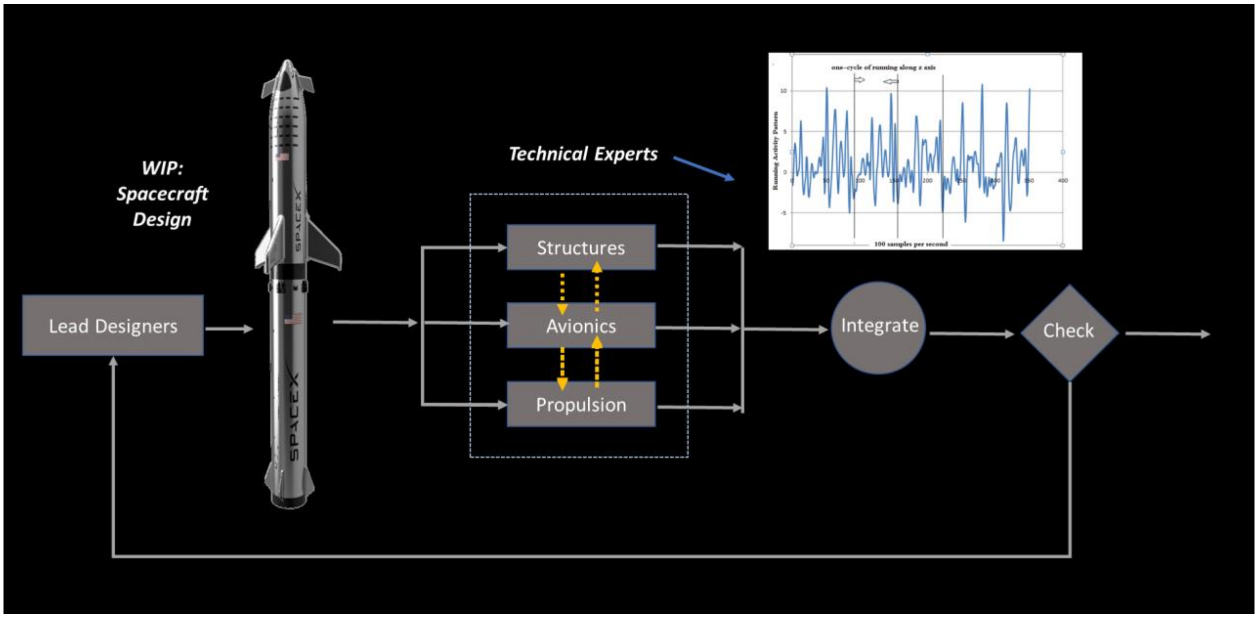

Figure 6. Data analytics in space exploration [16].

Figure 7. Satellites for exploration [17].

Figure 8. The artist's concept of NASA's X-57 'Maxwell' aircraft [19].

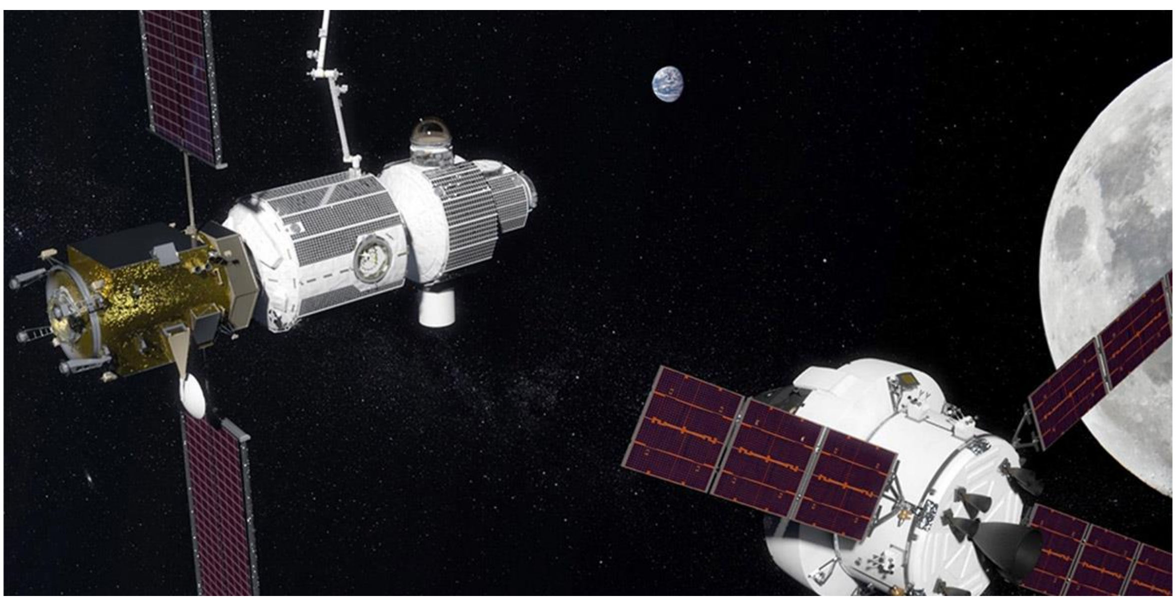

Figure 9. NASA's gateway in lunar orbit [19.