happiest minds
The Mindful IT Company
Born Digital . Born Agile

# APPLICATIONS OF BIG DATA IN CLEAN WATER AND SANITATION FOR ALL

# INTRODUCTION

Clean water and sanitation are critical aspects of human health and well-being and the functioning of ecosystems and economies. Access to safe and clean water and adequate sanitation facilities is essential for preventing water-borne diseases, reducing child mortality, and promoting economic growth. By 2030, SDG 6 wants to ensure equitable access to clean, cheap drinking water. This includes increasing the use of water-saving technologies and water-use efficiency to reduce water waste. SDG 6 also aims to eradicate open defecation and ensure that everyone has access to sufficient and equitable sanitation and hygiene facilities, focusing on the needs of women and people in vulnerable situations. This includes increasing sustainable sanitation technologies, such as composting toilets and urine-diverting dry toilets. Furthermore, SDG 6 aspires to enhance water quality globally by decreasing the share of untreated wastewater, eliminating dumping, minimizing the release of harmful chemicals and materials, and considerably increasing recycling and safe reuse. Millions of people face a significant issue: access to clean water and sanitary facilities. Lack of access to clean water and adequate sanitation facilities significantly impacts human health and well-being, leading to the spread of waterborne illnesses and environmental damage. The Sustainable Development Goals reflect the UN's recognition of the significance of access to clean water and sanitary facilities.


Big data has the potential to play a significant role in addressing the challenges of access to clean water and sanitation. By collecting, storing, and analyzing large amounts of data/ information from a variety of sources, including sensors, remote sensing, and social media, big data can provide valuable insights into the distribution and quality of water resources management, as well as the effectiveness of existing water and sanitation infrastructure. This information helps the decision-making process and drives more effective and sustainable solutions to clean water and sanitation challenges.

## PROBLEM STATEMENT

Millions of people in rural India are affected by open defecation, which harms sanitation and health. Despite government initiatives to construct toilets and increase awareness, many rural areas still need access to essential sanitation services. As a result, people start to conduct open defecation, which pollutes the environment and spreads diseases. Children's and women's safety and dignity are also impacted by poor sanitation. Developing practical ways to end open defecation and enhance the sanitation in rural India is crucial

# BACKGROUND WORK AND CASE STUDIES

Ancient civilizations created rudimentary systems for gathering and distributing water and managing waste, beginning a long and complicated water and sanitation management history. Aqueducts, sewage systems, and treatment facilities are increasingly advanced systems that have been created over time. However, as the world's population has increased in recent years, the need for water and sanitation services significantly strains the world's current infrastructure. Several issues, including restricted access to clean drinking water, inadequate sanitation facilities, and the deterioration of water resources due to pollution and misuse, describe the current status of water and sanitation management.

- **Water quality monitoring in India:** In India, the government has implemented a big data solution to monitor water quality in real time. Sensors are placed in bodies of water to collect data on pH, temperature, and other vital parameters. These are then analyzed using big data technologies to identify areas where water quality is poor and to track the spread of waterborne diseases

## Real-time water monitoring in the US:

In the US, a big data solution has been implemented to monitor real-time water consumption, enabling utilities to detect leaks and reduce water wastage. The solution collects data from water meters, analyzes it using big data technologies, and provides utilities with real-time insights into water consumption, enabling them to respond quickly to changes and reduce water wastage.



## Sanitation management in Africa:

A big data solution has been implemented to improve sanitation management. The solution collects data from sensors placed in latrines and uses machine learning algorithms to identify patterns and predict potential issues, such as clogs and malfunctions. This enables the government to respond quickly and take appropriate action to improve the efficiency and sustainability of the country's sanitation infrastructure


## Sanitation management in Africa:

A big data solution has been implemented to improve sanitation management. The solution collects data from sensors placed in latrines and uses machine learning algorithms to identify patterns and predict potential issues, such as clogs and malfunctions. This enables the government to respond quickly and take appropriate action to improve the efficiency and sustainability of the country's sanitation infrastructure

# NEED OF BIG DATA SOLUTION

Big data analytics solutions are needed in various industries, including clean water and sanitation because big data provides a means to handle the large and complex amounts of data generated and collected by multiple sources. The traditional data processing and analysis methods are often incapable of handling such large and complex data, making extracting valuable insights and making informed decisions challenging. Big data analytics solutions provide several benefits that are particularly relevant for the clean water and sanitation industry.

**Real-time monitoring and analysis:** Big data analytics solutions allow organizations to collect, process, and analyze large amounts of data from various sources, including cloud, sensors, and IoT devices, in real time. This gives organizations real-time visibility into the state of their water and sanitation systems, enabling them to promptly detect and respond to issues

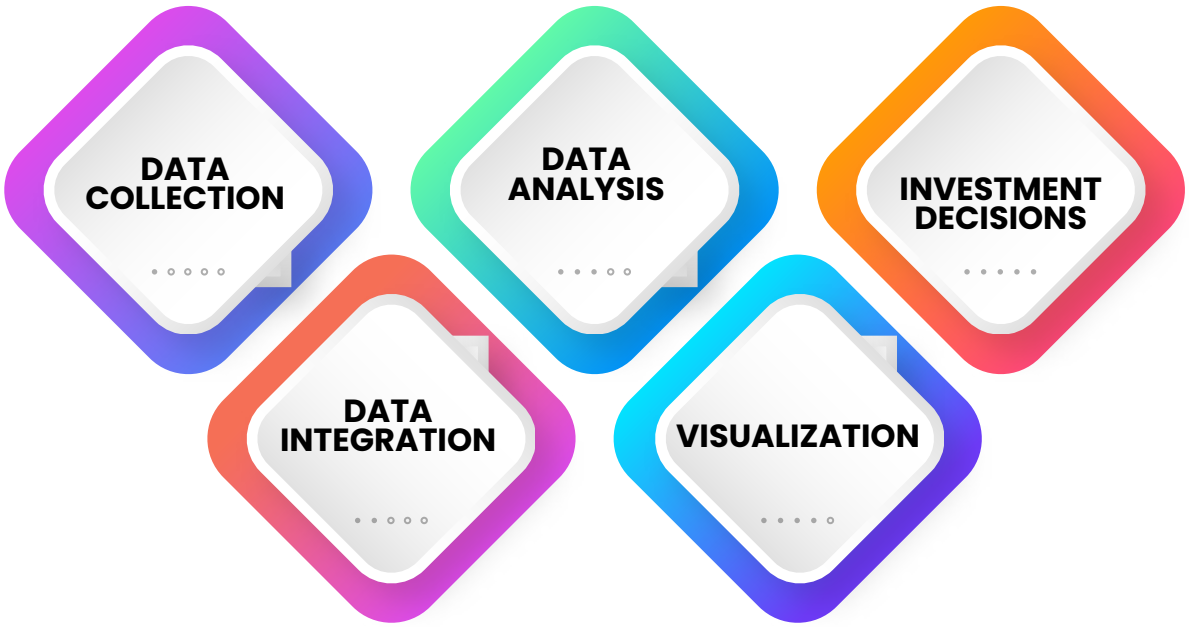
**Predictive modeling:** Machine learning algorithms can be trained on large amounts of historical data to predict future trends and patterns. This can be particularly useful in predicting water quality, water flow, and other relevant parameters, which can be used to optimize water resource management, predict and prevent water scarcity, and improve the efficiency of water treatment plants.

**Water resource management:** Big data analytics solutions can be used to store, process, and analyze large amounts of data on water resources, including satellite imagery, topography, and hydrological data. This data can then be used to optimize water resource management, predict water scarcity, and improve the efficiency of water treatment plants.

**Community engagement:** Big data analytics solutions can be used to analyze data on community engagement and public perception of water and sanitation issues. With the help of this information, communities can be better engaged and communicated with, and problems with access to and use of water and sanitation may be found and fixed.

# PROPOSED SOLUTION

**Improving sanitation infrastructure planning:** Local authorities can prioritize planned investment in new sanitation facilities by using big data analytics to identify regions with high rates of open defecation. Finding ideal places for new restrooms may entail examining data on population density, demography, and existing infrastructure




- **Data collection:** The first step is to collect data from various sources such as government databases, satellite imagery, household surveys, and other relevant sources. This data can include information on population density, demographics, income levels, existing infrastructure, and areas with high levels of open defecation.

- **Data integration:** The collected data must be integrated into a centralized data platform, such as a data lake or data warehouse, to make it accessible for analysis. This enables data from various sources to be combined, cleaned, and transformed into a usable format. Hadoop, Hive, Pig, and Spark are big data technologies that can integrate data from various sources to improve sanitation infrastructure planning. Hadoop is an open-source software framework for storing and processing large amounts of data. It provides a scalable platform for data processing and can be used to store and analyze large datasets. Hive is a data warehousing and SQL-like query language for Hadoop. It can be used to manage and analyze structured data stored in Hadoop and can also be used to process large amounts of data quickly. Pig is a high-level platform for creating MapReduce programs that can process large amounts of data. Pig scripts can extract, transform, and load data into Hadoop. Spark is an open-source, distributed computing framework that can process big data. Spark can process data in real time, providing a high-level API for writing big data applications in Scala, Java, and Python

- **Data analysis:** With integrated data, big data analytics identify patterns and trends, such as high-density areas with low levels of sanitation infrastructure, which can inform investment decisions. For example, demographic data can be analyzed to determine the age and gender of those most likely to practice open defecation. This information can target specific populations for education and awareness campaigns.

- **Visualization:** The results of the extensive data analysis can be visualized through interactive dashboards or maps, providing insights that can inform investment decisions. For example, visualizing the relationship between population density and open defecation can help identify priority areas for investment in new toilets.

- **Investment decisions:** Based on the insights from big data analytics, investment decisions can be made to address the need for new sanitation facilities in areas with high levels of open defecation. This can include the construction of toilets, upgrading existing facilities, and installing community-based solutions, such as community toilets or latrines.

## CLOUD COMPUTING

Cloud computing can store and process large amounts of data related to sanitation infrastructure and usage patterns, making it possible to analyze this data in real time. This can be particularly useful in remote or rural areas where traditional data storage and processing methods may need to be more practical.

## CLOUD SECURITY

Cloud security is critical when implementing big data analytics and cloud computing in sanitation management. Sensitive data related to sanitation facilities and usage must be protected to ensure privacy and prevent unauthorized access. To address this, it is essential to implement robust security measures, including encryption and secure access controls, to protect data stored in the cloud.

## ETHICAL PRACTICES WHILE HANDLING BIG DATA

When implementing big data analytics in sanitation management, it is essential to follow ethical practices to ensure that data is collected, stored, and analyzed in a responsible and trustworthy manner. This includes ensuring that data is collected with the informed consent of individuals, protecting sensitive information, and ensuring that data is used for the intended purposes. Additionally, it is vital to ensure that the benefits of big data analytics are shared equitably, and that the technology is not used to perpetuate existing inequalities.

# CONCLUSION AND FUTURE SCOPE OF WORK

In conclusion, implementing big data analytics and cloud computing to eliminate open defecation in India holds great potential. By analyzing population density, demographics, and existing infrastructure, local authorities can identify areas with high levels of open defecation and prioritize investment in new sanitation facilities.

This can help to improve sanitation infrastructure planning and ensure that resources are allocated effectively. However, it is essential to consider cloud security and ethical practices while handling data to ensure the privacy of individuals and communities. The proposed solution should take into account India's unique cultural and societal context and be sensitive to the diverse needs and experiences of those most affected by open defecation.

In the future, there is scope to develop further and refine the use of big data analytics and cloud computing in improving sanitation infrastructure planning. This could include incorporating real-time data from sources such as sensors and social media to provide more accurate and up-to-date insights into patterns of open defecation. Additionally, there is potential to collaborate with other countries and organizations working towards similar goals to share best practices and other advanced technology in improving access to clean water and sanitation.

# Rijan Gaha

**Business Analyst,**
**CoE-Digital Process Automation**

## ABOUT THE AUTHOR

**Rijan Gaha** is a Business Analyst, CoE-DPA at Happiest Minds with extensive experience in business analysis and identifying and implementing RPA, Intelligent Document Processing & Low-code solutions for organizations. He has a keen eye for analyzing business processes and converting them into automation opportunities. He values how Gen AI and data can transform the organization to achieve its objective. Rijan is a great team player who believes technology can transform complex business problems into simple solutions.

**Write to us at**
**Business@happiestminds.com**

happiest minds
The Mindful IT Company
Born **Digital** . Born **Agile**

**www.happiestminds.com**

Happiest Minds Technologies Limited (NSE: HAPPSTMNDS), a Mindful IT Company, enables digital transformation for enterprises and technology providers by delivering seamless customer experiences, business efficiency and actionable insights. We do this by leveraging a spectrum of disruptive technologies such as: artificial intelligence, blockchain, cloud, digital process automation, internet of things, robotics/drones, security, virtual/ augmented reality etc. Positioned as 'Born Digital . Born Agile', our capabilities span Product & Digital Engineering Services (PDES), Generative AI Business Services (GBS) and Infrastructure Management & Security Services (IMSS). We deliver these services across industry sectors such as automotive, BFSI, consumer packaged goods, e-commerce, EdTech, engineering R&D, healthcare, hi-tech, manufacturing, retail, and travel/transportation/hospitality. The company has been recognized for its excellence in Corporate Governance practices by Golden Peacock and ICSI.

A Great Place to Work Certified™ company, Happiest Minds is headquartered in Bangalore, India with operations in the U.S., UK, Canada, Australia, and Middle East.