Exploring the ethical implications of Big data and Data mining

Husnara Khan Daljeet Kaur

Department of Computer Science Department of Computer Science

Gyan Ganga College of technology Gyan Ganga College of technology

Jabalpur M.P India-482002 Jabalpur M.P India-482002

Email:husnarakhn@gmail.com Email:daljeetkaurkalsi83@gmail.com

**ABSTRACT**

Big Data and Data Mining analysts are a growing methodology for analysing data, extracting information, and relating it to other information in a variety of application domains. However, dealing with the vast volumes of data that are already available presents numerous difficulties in the context of public policy. To assess this heterogeneous and multi-sourced data, new techniques and technologies must be developed. Therefore, it is crucial for both individuals who utilise big data and data mining, as well as those who are the target of its usage, to identify and investigate its ethical implications. Our study methodology is founded on practitioner specialists who shed light on the moral dilemmas raised by big data analytics. Thus, in addition to making a significant theoretical contribution to the literature on information systems, our growing empirical insights also feed management and policy guidelines for putting big data and data mining analytics into practise and reaping their ethical benefits.

**Keywords-** Big data ethics , Data mining ethics, Awareness, Privacy, Security

**I. INTRODUCTION**

Every industry in the world, from business to emergency services to health care to science and space, has been dramatically altered by the use of big data analytics in how they manage and carry out their daily operations. All of these areas are now able to forecast future outcomes, make better key decisions, and take more efficient actions thanks to big data analytics. Businesses now have the ability to gather, access, and use customer data, and even influence their behaviour. Big data is being used extensively across many industries, particularly in business. To collect consumer data and access their personal information, including interests, behavioural patterns, thoughts, security which they retain and use to their advantage. The majority of the time, this information is acquired and accessed without the consent or awareness of their customers information, interpersonal connections, etc., they are combining big data analytical techniques. [1].

Big data has great potential for science and society, large amounts of diverse data can be produced and made available to researchers much faster than "traditional" data. Although this potential is acknowledged, there are ethical issues that users of big data must consider. Due to the amount and variety of information contained in big data, the risk of disclosure is greater. Researchers and data access services working with highly detailed and sensitive protected data have been struggling with this problem for years. The industry has developed both ethical guidelines and statistical disclosure control techniques that those working with big data can use. We discuss the challenges, present some references and techniques, and conclude with recommendations for secure access to big data.[10]

There is no consensus on what makes data "big," but a common way of thinking about big data is that it consists of multiple data sources that are combined or linked to create a specifically meaningful data source. Big data can be considered to have several main characteristics: volume, variety and speed, it refers to the size of a dataset from multiple data sources combined by some link variable. They are much larger than social research datasets and require significantly more computing power . Combining data sources leads to another feature - diversity. A large dataset contains information about many different areas of people's lives. For example, digital tracking data may combine information about retail stores, location history collected by mobile phone GPS, websites visited, etc. The last characteristic, speed, refers to the increased speed of data collection and processing.

## That big data offers big opportunities is certain, but it is not a unanimously positive picture. As with all data, big data has challenges and issues that have been widely discussed1. Many ethical issues related to consent and privacy have been raised in relation to big data. These ethical issues are complex and no attempt is made here to provide a complete overview. Nor is it intended to provide a complete guide to the ethical use of big data. This article focuses on disclosure risk (the risk that individual data subjects are identified) as a central big data ethical issue and explores what lessons can be learned from the experiences of secure data access services. There are many protected information access services in different parts of the world.

## These services specialize in producing highly detailed and sensitive information from administrative, official sources and large social surveys. Due to the detail and sensitivity of the data, these data sets may be revealing. In other words, there is a risk that individual data subjects will be re-identified based on the use of the data. Secure data access services have spent years developing the infrastructure and technologies to ensure secure access to this data, so it's a natural place to turn when thinking about how to protect yourself from exposure to big data.

## This article focuses on access to protected data in the UK, although the approaches presented here reflect those across Europe and the US. Ethical use of big data In this article, ethical use generally means that the use of big data does not harm scientific research. Although many participants in the research process, such as the data source and the wider research community, may be harmed, the focus here is on the potential harm to the data subject. Many organizations and authors have prepared research ethical criteria7. Although they vary according to the academic discipline and the type of information, they share some common features:

1. The project or research must have some generally useful or useful part

2. Security and confidentiality of data must be guaranteed

3. Registered persons must not be identified or harmed as a result of the research . The scientific community must demonstrate reliability

4. The research methodology must be reliable and statistically valid. results Current ethical frameworks can indicate consideration of these five areas.

For example, researchers define the usefulness or benefits of their research and the data must be protected and the results reliable argues that vulnerable people should not be harmed by a research project and that researchers should demonstrate their credibility so that they do not need strict supervision. The world of big data is not inactive in this area, and there is a lot of discussion about ethics and the ethical use of data. In the search for ethical frameworks, big data could benefit from learning from the world of social and administrative data, which has several well-established ethical frameworks. The UK Statistics

Authority has recently invested heavily in the ethical use of secure or legally controlled data. To achieve consistent and ethical research practice, the UK Statistics Authority has developed a comprehensive framework: a self-assessment tool that enables researchers to carry out a comprehensive ethical assessment of their research projects. This is a mandatory part of the ONS Protected Data application process, but recommended for all projects using secondary data sources.

# II .Big Data Is Ethically Neutral

While big data technology offers the opportunity to combine information and innovate new products and services for profit and greater social benefit, it is, like all technologies, ethically neutral. That is, it has no built-in perspective on what is right or wrong, or what is good or bad about using it. Big data technology has no value framework. However, individuals and companies have value systems, and only by asking and seeking answers to ethical questions can we ensure that big data is used in a way that is consistent with those values. Such discussions require an explicit examination of these values ​​and the development of ethical perspectives, which can be difficult. Ethics is a very personal subject and has a lot of polarizing vocabulary like good, bad, right and wrong. We all have personal moral codes, which of course differ from person to person. The lack of common vocabulary between what we personally believe and what we, as members of a joint enterprise, plan to do with big data can limit productive discussion and prevent consensus. However, this is not a book about dictating operational principles or judicial or legal changes. Business leaders, managers, judges and elected officials need to be concerned about this. Nor is it a book about business ethics - at least not in the traditional sense. Business focuses primarily on profit and innovation. Ethical inquiry as a formal practice is of interest only to the extent that it affects profitable operations and the continuous development of products and services that meet the needs of a dynamic market.

However, there is an inherently social component to business, and in fact, big data and social media have only exaggerated this reality in recent years. Business itself, the exchange of goods and services for assets (often in the form of currency), is an activity that usually involves people. And people have values. The purpose of this book is to provide a framework to facilitate ethical conversations in business environments designed to uncover these values ​​and help organizations act accordingly. Big data enforcement brings business functions and individual values ​​closer together. Big data is pushing business activities more and more fully into the lives of individual people due to the sheer volume, variety and speed of the data being produced.

The design, development, sale and management of big data is increasing its influence and impact on people's lives in ways that can change the common meaning of words like privacy, reputation, ownership and identity. Its sheer size and pervasiveness fundamentally forces new questions about our identity, the evolution of personal privacy, data ownership, and how our online information pathways affect our reputations—both online and offline. From enterprise to education and research to manufacturing to professional services, organizations have access to vast amounts of information about their customers, their activities and almost every other measurable aspect of their existence.

Before the rapid growth of big data technology over the past five years, changes in organizational processes or practices had a delayed impact, if any, on customers' lives. Whether or not a customer's personal data was available usually depended on how many people or organizations had access to the customer data. Big data is now operating at such a scale and such speed that such changes in policies and practices will reach ever faster and affect more people. Thus, changes in business operations have a much greater impact on people's lives.

The expansion of traditional activities touches our lives every day in ways we can hardly observe, let alone control. The reality is that the ways in which legislation, social norms, economics or reasonable expectations of normal interaction will change due to the increasing presence of big data are simply unknown. And precisely because these things are unknown, ethical dialogue should be encouraged. An open and clear dialogue about aligning values ​​and practices to balance the risks and benefits of big data innovation is one way to ensure that the business negotiations come out well – and in your favor. Identifying the moments when decisions become actions or ethical decision points is the first step in developing the ability to have these conversations both "in space" and more formally in forming transparent perspectives and perspectives. politics

**III. COMPARITIVE STUDY OF BID DATA AND DATA MINING**

| **Data Mining** | **Big Data** |
| --- | --- |
| It is one of the method in the pipeline of Big Data. | Big Data is a technique to collect, maintain and process the huge information. It explains the data relationship. |
| Data mining is a part of Knowledge Discovery of the Data. It is close view of the data. | It is about extracting the vital and valuable information from huge amount of the data.It is a technique of tracking and discovering of trends of complex data sets. It is a large or overall view of the data. |
| The goal is same as Big Data as it is one of the tool of Big Data. | The goal is to make data more vital and usable i.e. by extracting only important information from the huge data within existing traditional aspects. |
| It is manual as well as automated in nature. | It is only automated as computing huge data is difficult. |
| It only focuses on only one form of data. i.e. structured. | It focuses and works with all form of data i.e. structured, unstructured or semi-structured. |
| It is used to create certain business insights. Data mining is a manager of the mine. | It is mainly used for business purposes and customer satisfaction. Big Data is a mine. |
| It is a sub set of Big Data. i.e. one of the tools. | It is a super set of Data Mining. |
| It is a tool to dig up the vital information from the large data. Data can be large as well as small. | It is more involved with the processes of handling voluminous data. Data can only be large. |

## IV. LITERATURE REVIEW

Numerous studies demonstrate that, despite the fact that data mining has numerous benefits, it can be potentially dangerous when used to survey a person without their consent. Several publications cover the value of data mining for both corporate and public organisations. Data mining is advancing quickly, as a result of advancements in computing and information technology. Data gathering techniques have advanced recently, with digital technologies replacing manual techniques. These digital techniques include radio frequency identification (RFID), biometric tagging, cell phones, bar code readers, smart cards, and GPS location. The technique of gathering data has been improved by these devices.

However, the development of computing and information technology exposes a great deal of personal data that can be exploited without the owner's consent. new ethical and privacy concerns are emerging from data consumption, storage, and mining as a result of the present technological breakthroughs. Exploring people's data leads to ethical problems with data mining,. Government authorities have the right to access personal data from both public and private databases, despite the fact that the majority of public and private businesses have stringent privacy regulations. As an illustration, Austrian taxation

In order to obtain personal data about a person's property, employment, and earnings from multiple databases in order to conduct an investigation into tax fraud, the office (ATO) requested a waiver of privacy policies . While identifying tax evaders is the primary goal of this ATO programme, using personal financial information for other purposes is unethical and should never be done without authorization. Even while government entities may have a legitimate justification for collecting personal information, doing so is often regarded as unethical because it violates a person's right to privacy. For the purpose of enhancing security, governance, and other social services, numerous governments around the world engage in extensive data mining operations.

Government data mining initiatives in the USA should abide by rules and guidelines that respect the law. The US government use cutting-edge data mining methods to stop incidents like 9/11 from happening. The paper also claims that this procedure may violate people's constitutional and civil rights, such as their right to privacy and their freedom of speech.

According to the research, innocent people have been mistaken for terrorists and subjected to travel restrictions, while others have used this information to damage the reputations of well-known persons. For instance, in the 2008 US presidential election, candidates' reputations were ruined by the exploitation of private information Controversies have always occurred from private and public institutions mining the personal medical records of numerous people .Even though many researchers in the medical sector use this information without the patient's previous consent, getting medical information from a patient without his consent is unethical. The government organisations and medical researchers should tell people about the information they are seeking before receiving it. [5]

**V. ETHICAL IMPLICATIONS OF BIG DATA**

Information technology (IT) is one of the most growing areas in the last ten years of the world. Millions of people do IT-related jobs as professionals. Recently, advances in processing a large amount of data have created new opportunities to improve individual lives and the welfare of our societies.The opportunity for profit comes through data mining, forecasting analytics tools, and other methodologies. Because of that big companies are finding paths to comprehensive more and more information about consumers. It shows how companies might realize the value of big data among data mining. It depicts big data and data mining as the most specific area of IT.

Big data analysis raises a number of ethical questions, especially when companies start trading their data externally for purposes other than those for which the data was originally collected. The scale and simplicity of analysis completely changes the ethical framework today. We can now do things that were impossible just a few years ago, and the current ethical and legal framework cannot dictate what we should do. Although there is still no black or white, experts agree on some principles: [2]

1. Private customer information and identity must remain private: Privacy does not mean secrecy, as personal data may need to be reviewed based on legal requirements. but personal information obtained from a person with his consent must not be disclosed to other companies or individuals, so that they have traces of their identity.

2. Private Information Shared Must Be Treated Confidentially: Third-party companies share sensitive information—medical, financial, or location—and need restrictions on whether and how that information can be shared.

3. Customers should have a transparent view of how our data is used or sold and the ability to control the flow of their private data between large, third-party analytics systems.

4. Big data should not interfere with human will: big data analysis can moderate and even define who we are before we decide. Companies need to think about what predictions and conclusions to allow and what not.

5. Big data must not institutionalize unfair biases, just as machine learning algorithms can absorb unconscious biases in a population and amplify them using training samples.

We will certainly need to develop more principles as more effective technology becomes available. Data scientists, information designers, database managers and all participants in big data processing should make their voices heard in the ethical debate about data usage. Companies should discuss these issues openly in formal and informal forums. If people don't see ethics in their organization, they will disappear in the long run.

**A. Personal Data Collection**

Data thieves or data robbers may sell the data to foreign companies, government research institutes or agencies and use the data for other purposes, such as data mining. Academic research centers, private organizations and research institutes may also use this information for analytical purposes. Thus, potential harm caused by wrong information or wrong conclusions creates problems for society and users. Data owners must therefore protect their privacy and confidentiality, but they must also be informed about how and why their data is being used. Big data and data mining projects involve personal data processing methods. The term "personal information" refers to information that can directly or indirectly identify individuals by collecting information such as an identifiable name, ID card number or other personal characteristics.

**B. Ethical Issues on Selling Personal Information**

Improper use of data is called data abuse. This is called an "offense or activity" that violates certain codes of conduct and can be intentional or accidental. Data loss and misuse is another issue related to data misuse. When valuable information is disclosed, it is harmful to individuals. There will always be criminals with ulterior motives looking for ways to profit from using this information. Big data is bigger and there are no special machines to actively monitor and manage sensitive human information. Sometimes big data analysis is useful for people in healthcare, education etc. Data mining with big data is a good activity, but the privacy of individuals' data must be protected. We use technology to make our lives run smoothly, but sometimes it causes us problems. Millions of people use email as a means of communication. But sometimes they get mostly spam. Anyway, problems arise because of technology, we have to move forward with new technology. But we should manage and reduce potential problems related to technology and society. Through the process of data mining, it brings more benefits to the greatest number of people. Individuals must protect their personal data against unethical purposes to reduce professional problems related to big data and data mining. Therefore, when they fill in personal data in the form, individuals must ask how this data is used, why it is used, which part of the data is processed and how long they keep this data, etc. But individually and socially, one of the biggest concerns of big data and data mining is the security and privacy of big data. Because they are widely accepted in our daily life. A vast amount of information has been produced based on various aspects of individuals. Without adequate security and privacy protections, information would be exposed intentionally or unintentionally. It can also endanger individuals. We must also recognize that we cannot expect simple answers to the complex moral issues surrounding data. This is often impossible, especially since the information environment changes so rapidly.

**C . Privacy Concerns**

For companies, managers, marketers and analysts, big data can mean the key to a new era of business intelligence and personal business services. Big data analytics enables companies to extract and visualize hidden patterns and meaningful insights from a variety of internal and external data sources. Using this information, they can add intelligence to their processes, simplify operations, gain competitive advantages and increase sales. Importantly, by accurately analyzing complex, heterogeneous and large data sets (i.e. data from internal sources and the increasing rapid flow of heterogeneous data available externally), business analysts can, among other things, optimize their marketing and advertising strategies, gain real-time insights about customer needs, usage and purchasing habits, and potentially identify early emerging (product/market) trends. Many companies, especially online platforms (Internet and social media), retailers and marketplaces (Amazon, etc.), are already applying big data techniques to their massive databases of consumer purchase history, transaction data and inventory information;

I understand their customers better. ii. Provide personalized products, services and recommendations to current and prospective customers and iii. Anticipate changes and changes in demand. With the help of big data analytics, sales and marketing specialists can also take advantage of the mountain of data that customers use online services with smartphones, make online purchases with electronic cards or share their location and personal thoughts on social networks. media to target the right consumer at the right time with the right message. But while the growth of big data creates enormous opportunities for business leaders, organizations and society as a whole, it also creates serious privacy concerns. Big data has allowed companies and analysts to create and use consumer data without the knowledge and direct permission of consumers, and is often used unexpectedly after acquisition.

This not only affects privacy, but also leads to secondary harms such as profiling, surveillance, discrimination, exclusion, government surveillance and loss of control. In addition, some data telecom giants and social platforms have been observed to begin selling consumer data to buyers who want to use that data to make critical business decisions and drive sales. Although this data is said to be used only to drive targeted marketing strategies and make critical business decisions, there is also the possibility of data misuse. This not only prevents consumers from providing their personal information, it also compromises their security and thus exposes them to potential harm. These questions not only cast doubt on whether the right of consumers to control the data they directly and indirectly disclose can be fully realized, but also raise the question of whether individual control of personal data is an achievable goal of big data.

Given the complex life cycle of personal data in big data, the consumer's right to personal access raises an important question: how can the safety of consumers and the right to access and manage all their data be ensured. level of infrastructure? Similar concerns have been raised in the medical field, where the electronic storage and distribution of people's health information can potentially reveal information not only about individuals but also about people related to them. This was seen in a case in the United States where the genome data of a sick woman was publicly available, which raised objections from family members.

Although his data was later deleted, it raised concerns among open-access advocates that privacy concerns could significantly slow the progress of big data research. However, health data collected without personal data would still provide accurate research results to understand the exact time of occurrence, temperature, age and gender that affect our lives in the situation of out-of-hospital cardiac arrest. consumer privacy and security of their information and communication with others is unsecured/unencrypted storage of information leaving it unprotected, insecure and vulnerable to attack. Finally, unauthorized personnel can access and hijack individuals' data for nefarious purposes, posing a huge threat to an individual's safety, security and overall well-being. A typical example of the access and misuse of private information is the case of Hilary Clinton, who ran for the US presidency in 2016 and whose email was hacked and used against her, according to several political analysts. and individuals (Davis, 2018), this act lost him the presidential election agencies monitor the email and mobile phone communications of millions of residents, perhaps the pinnacle of big data analytics challenges.

Therefore, some mainstream commentators argue that the approach violates privacy. Although some citizens seem worried about privacy violations, the government's official defense tends to emphasize that data collection did not examine the content of emails, private conversations, and phone calls, but focused on more general metadata. Despite existing data protection frameworks that guarantee the right not to be subjected to such surveillance, big data analytics can increase the ability of governments and corporations to monitor private citizens, ultimately undermining individuals' rights to privacy, anonymity and freedom of expression. . Without consensus in society about the scope of data retention and control laws and how they interact with privacy laws, big data programs risk being negatively impacted, derailed, or even halted by public backlash and legal challenges.

**D. Awareness Concern of Big data**

Awareness addresses what people know about big data analytics practices, such as how organizations analyze and use their data to make critical decisions. An ethical dilemma arises when people lack information about why organizations use big data analytics and related processes .Areas of awareness that highlight ethical issues in big data analysis include understanding the concept of big data analysis, understanding the rights associated with big data analysis and knowing who holds the data and what it is used for. Unfortunately, most of the world is not aware of what big data analysis is, how it works, and that their data and information are used by various institutions and organizations.

Organizations typically collect big data directly and indirectly without the explicit informed consent of subjects and often hide secondary uses of data from individuals .Lack of awareness is an underlining factor that has created the biggest challenges to people's privacy, security and freedom. Many managers and experts argue that individuals need to learn about big data analysis, how it works, and how it affects their choices and behavior .They also believe that individuals should participate in public information programs and understand the appropriate uses and implications of big data analysis .

In this way, people can learn to better balance the personal costs and benefits of big data analysis.. Second, individuals must be aware of the laws and practices that protect them from the potential negative consequences of big data analysis [e.g. European General Data Protection Regulation (GDPR)]. Third, individuals need to know what information organizations collect about them, who owns and controls it, and which third parties may have access to it. Individuals need to be aware of , of these practices because the analysis of their data will ultimately affect their lives.

**E. Intellectual Property Concern of Big data**

Just like creativity and innovation, data is invaluable. Creativity and innovation are called intellectual property and are protected by certain laws (copyright/patent/trademark) against theft, misrepresentation and infringement. Laws protecting intellectual property rights give their owners full rights to their creations and prevent them from stealing, claiming or compromising. Unfortunately, conventional data is not considered intellectual property and is not legally searchable. The reason is that, according to law firms, it does not meet the requirements for patent or copyright protection. This limitation affects the concept of ownership of data.

This means that individual data obtained, stored and used by a certain entity can also be freely used without restrictions or charges. In this case, it is easy to damage and even misuse the data of individuals. This significant challenge concerns data collectors, managers and agencies regarding data ownership and responsibility and the security of individuals' data. It therefore questions the responsibility of data users and recommends significant changes to existing legal constitutions and practices to support data ownership and protection.

**VI. Ethical implications of Data Mining**

Data mining, especially when using data about people, has important ethical implications. Companies face an ethical dilemma even when deciding whether a company should inform an individual that their data is being stored for future data mining. By giving an individual the option to opt out of data collection, a company can harm its competitive advantage in the marketplace. The company must decide whether the lack of ethical concern will result in a loss of consumer goodwill and whether the company will experience consumer backlash. Companies using data mining techniques must act responsibly and be aware of the ethical issues associated with their implementation. they must also consider the wisdom of their actions.

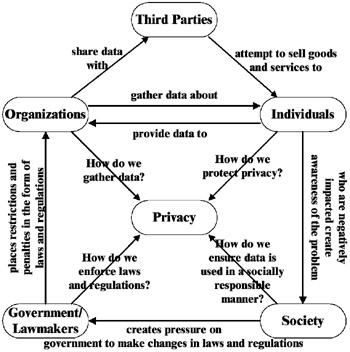
For example, data mining can sometimes be used to discriminate against people, especially on the basis of race, gender and religious orientation. This type of data mining is considered not only unethical but also illegal. Individuals must be protected against unethical use of their personal data, and before making a decision to disclose their data, they must know how this data is used, why it is used, which parts of the data are used. and what are the consequences of this action. By doing this, individuals are informed and told directly about the reasons and consequences of using their data. Ethical issues in data mining can be seen in two main ethical issues related to privacy and individuality. As mentioned earlier, misuse of data can lead people to do unethical things that are also considered illegal. The importance of privacy and individuality must be valued and believed to be protected to ensure that people are treated fairly. People should be aware of the significance of hazards and risks and constantly discuss these ethical issues. Experts consider data mining to be morally neutral, however, the way this data is used can raise questions and concerns about ethics. The information must be used for the right purpose to ensure the safety of people.[3]

## A. Security Concerns of Data Mining

Data mining is the process of creating a series of valid and meaningful queries to extract information from large volumes of data in a database. As we know, data mining techniques can be useful to recover database security problems. However, with the growth of development, there has been a serious concern that data mining techniques can create security problems. Many security experts consider data mining to be one of the most important challenges facing consumers in the next decade. The clear complexity of data mining lies in the construction of accurate models to analyze the data without giving the right to use the information of certain customer records, which protects the database against abuse. The development of such models can reduce the security problems faced by users.

Data security issues in data mining are one of the most popular issues because when using data mining, individuals usually work with and easily access large amounts of data. It is dangerous if this information is not used in a secure manner. Since data mining provides many new areas to extract information from old databases as well as future databases that can be developed with data mining as a supporting objective, a session of data mining in some large enterprises may suggest that data security problems may be important. data mining. Having said that, it is not recommended to shed light on data mining, but one important aspect and thing to mention is security that needs to be assessed and addressed.

Data warehouse companies must control who has access to data and what parts of the data service have access. An example of a company that allows limited access to its data warehouse for data mining purposes is Wal-Mart. Wal-Mart has a very extensive database of all its warehouses, stores and collected data. Companies that have Wal-Mart products can access Wal-Mart's database. This allows these companies to mine this data to obtain information about the sales of their products. By limiting these companies' access to only company-supplied products, Wal-Mart demonstrates that it is aware of the security and privacy concerns associated with data mining.

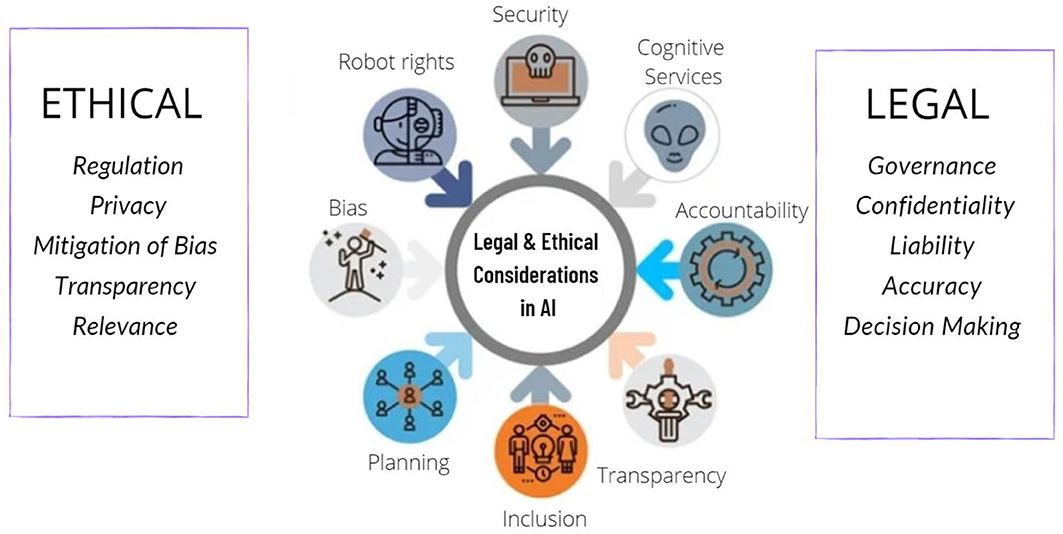


## B. Legal and Privacy Concerns  of Data Mining

In data mining, potential privacy and legal issues are key drivers of increased conflict. The way data mining is used raises questions about privacy. Every year, the state and companies collect a huge amount of information about customers and store it in data warehouses. Part of the concern is that once data is collected and stored in a data warehouse, who has access to that data? Often, the consumer may not be aware that the information collected about them is not only shared with the data collector. With the technologies available today, data mining can be used to extract information from data warehouses, discover various customer information and relationships and draw connections based on this extraction, which can compromise customer information and privacy.

Data mining requires data management, which may include consumer data that may compromise confidentiality and privacy. One way to do this is through data aggregation, where data is collected from different sources and put together for analysis. Companies like IBM are working on data mining methods that allow complete privacy by creating accurate data models. IBM has developed a method called Privacy Preserving Data Mining. By randomizing a consumer's personal data before sending it using IBM's privacy-preserving data mining method, the company can still collect the data it wants without compromising its customer's right to privacy.

It is logical that many companies and government agencies must use data mining as part of their work, but the doubt is whether this data is used correctly. For example, data mining can be useful for some companies in targeting the right market. In the age of technology and information, it seems that obtaining information about customers and employees has become much easier than before. The rapid transfer of personal data has led to the risk of identity theft. Data protection issues become an important issue in data mining due to the associated risks, especially the fact that many consumers buying products or services are not aware of data mining technology.



## C . Discussion

The ethical status of data mining by government agencies for security reasons remains a controversial issue. A conflict arises between public interest and individual rights and freedom. Although governments claim to find relevant information in personal records, mining this information is legally wrong, unethical and also a violation of individual privacy .Although complete privacy is not possible due to the need to communicate in society, personal data can only be accessed by the owner, who must decide what to share with other people. Therefore, it is not morally right for the government to access this information and use it to improve security.

If a person feels that some of his information is confidential, accessing that information is inappropriate because he may face many unexpected risks. To avoid privacy concerns, government agencies would be wise to seek consent from individuals before using such information. Politicians and legislators around the world are giving individuals the right to control the movement of their personal data. These laws state that an individual has a right to privacy and that information obtained from him or her may not be transmitted or used for any purpose other than that for which it was obtained

For example, a banker may not share its customers' credit card details with any other legal entity without the customer's consent. Thus, data mining is illegal from a legal point of view. Therefore, it is ethically wrong for public authorities to use personal data for secondary purposes. Data mining methods require searching multiple records from different sources. This means that the quality and authenticity of this information cannot be guaranteed. Even with data cleaning and analysis, most data sources can give the wrong impression of an individual.

Data mining is prone to mistakes, errors and poor quality results. Such poor quality information can have a significant impact on an individual or society. Relying on such information can lead to false accusations that affect an individual's career, family and social life. When a person is labeled as a suspect, there are negative consequences such as discrimination, injury or death during shootings, loss of reputation, and lawsuits. Trust is another ethical issue arising from data mining. As data mining by government agencies and private companies increases, most people gradually lose trust in these agencies and are unwilling to share information.

In summary, data mining by government agencies compromises individual privacy, which is unethical and illegal. However, the government's actions seem justified in the fight against corruption, crime and terrorism. Based on the discussion, there are three possible solutions to the current problem. First, the government may be empowered to invade privacy in hopes of preventing crime and terrorism. This would lead to abuse of ethical and privacy rules to improve security and social governance. While this solution is good, there is no measurable evidence that data mining positively identifies criminals and improves security.

Another solution involves a compromise between data mining, ethics and privacy. In other words, allow the government to use unethical practices in special cases. This settlement sets limits or situations in which the government should have access to personal information. Such rules determine the type and amount of data collected, as well as the methods of their processing. Although this solution is good, it does not prevent the government from collecting data through its secret agencies if their demands are justified. A third solution would be to ban personal data mining. This would mean that the government would have to rely on other sources of information to detect and prevent crime, corruption and terrorism. This solution would force government agencies to follow good ethical and moral behavior and comply with privacy laws. Deal with security and social arrangements

**VII. conclusion**

In summary, the ethical issue of data mining by government agencies remains highly controversial. Laws and decision-making bodies support and consider privacy as one of the most important rights of an individual. Therefore, it is ethically wrong for government agencies to violate this fundamental right. Although data mining can have potentially beneficial results in curbing crime and terrorist activity, the breach of individual private data can have harmful effects and is unethical. The government should therefore find other ways to improve management and security. The government should also enforce strict privacy laws that prevent any organization from collecting personal information during data mining.[5]

Everywhere we use IT technology to achieve a better life, but sometimes it causes us difficulties, such as abuse of IT professionals. Discover a better world for us in the process of big data and data mining. It is important to control crime and anti-social behavior under the control of abuse. When it comes to big data and data mining, customer privacy is essential. As professionals, the IT industry also needs to strengthen professionalism. The government should strengthen this data protection. The Data Protection Act should also include some big data and data mining laws, and customers have the right to opt out of global sharing of their data, while online sales policies should also cover this. In addition, big data and data mining professionals should include some rules and regulations. IT professionals must take full responsibility for their work. Business data must be managed in corporate law. Finally, the Privacy and Data Protection Act should also include some privacy protection laws.

**References**

[1].Vol. 2, Issue 5 pp.43-48 (2020) ISSN: 2668-778X www. techniumscience.com

* [2] Ida asadi someh European Conference on Information Systems (ECIS) At: Istanbul, Turkey

[3] [www.mbaknol.com/information-systems-management/ethical-security-legal-and-privacy-concerns-of-data mining](http://www.mbaknol.com/information-systems-management/ethical-security-legal-and-privacy-concerns-of-data%20mining)

[4] Published by  Doctor Erick at  March 4, 2017

[5] ivypanda.com/essays/ethical-implications-of-data-mining-by-government-institutions

[6] Social, Ethical and Legal Issues of data Mining by john Wang idea group publishing 2003

[7] Marina Da Bormida978-1-80262-414-4, eISBN: 978-1-80262-411-3 ISSN: 2398-6018 Publication date: 9

December 2021

[8] Amin choudhary digital development advisor at USAID published oct 10 ,2015

[9] Innov clin neurosci 2020 Oct-Dec; 17(10-12): 24–30. Published online 2020 Oct 1.

[10] Volume 8 Issue 2 february 2022 e08981Deborah Wiltshire