Proceedings of the 3rd International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/6/20230953

# The application of big data and the analysis of advantages and disadvantages in its development

**Shiyue Yang**

RDFZ Chaoyang branch school, Beijing, China, 100028

yangshiyue@rdfzcygj.cn

**Abstract.** Today, all Internet technologies are being constructed on the basis of big data, the so-called Internet of Everything, which is said to be a comprehensive service for individuals, facilitating people's life, enriching people's brain, and improving the quality of life and work efficiency. On the other hand, Restrictive laws and regulations have not been developed and improved along with big data, but rather lag behind. This has resulted in loopholes in the big data industry, giving all the bad people the opportunity to take advantage of personal data reselling, privacy disclosure, and even the use of big data to influence personal habits, such as daily work and rest, shopping and consumption, education and cultural beliefs. This paper lists the benefits of big data to various groups, as well as the hidden problems in the development of this technology. Through literature analysis, this paper finds that big data is a great threat to the privacy and security of the masses. At the end of the paper, some constructive suggestions are put forward to protect public privacy.

**Keywords:** Big Data, Privacy, Technology, Network Security.

## 1. Introduction

Big Data is a result of today's high-tech, information age. Big Data is a term used to describe data sets that, for decision-making, insight, and process optimization purposes, cannot be acquired, managed, and processed in a given amount of time using traditional software tools. Big Data calls for novel processing patterns. The storing and processing of enormous amounts of data constitute big data's primary value. Big Data, according to IBM, has the following five v's: volume, velocity, variety, value, and veracity. Vogels said, "Once you enter the world of big data, companies will have unlimited possibilities in their hands". Jack Ma, the founder of Alibaba, mentioned in his speech that the future era will not be the era of IT, but the era of DT, which is Data Technology, showing that big data is pivotal to the Alibaba Group. Data have unfathomable worth to individuals. It is possible to assert that whomever gathers more data will be able to rule the industry. The research methodology used for this work involves data induction, data analysis, and data integration. Big data has now become a part of everyday life, although not everyone is aware of its precise function. In order to lessen the likelihood of being impacted by big data hazards, this paper provides a general introduction to the role and dangers of big data. This will encourage readers to pay attention to their privacy security throughout use.

DOI: 10.54254/2755-2721/6/20230953

## 2. The benefits of big data

The development of big data can undoubtedly bring us many opportunities and benefits. First of all, big data can bring convenience to our daily life. It can bring personalized recommendations to us, which is difficult to achieve with other technologies. For example, online shopping, film and television recommendation, software recommendation and so on. This has created huge fortunes for some companies. Secondly, big data projects can accelerate the development of society and make information processing faster and more accurate. For example, weather forecasts based on supercomputers can be predicted more reliably with the help of big data technology, and outbreaks in China can be controlled more effectively through big data networks, precisely controlling the movement of each infected person. The development of big data has provided a breeding ground for most new industries, from chip development to artificial intelligence, from combustible ice harvesting to the use of hydrogen energy, all of which require data from previous research. At the same time, the rise of big data has also provided a measure of employment in society. Many industries derived from big data provide a large number of jobs and enhance people's well-being. It understands humanity even better than we do ourselves.

When the Internet was less well known, the general public had to go to physical stores to shop because of their variety and various areas. This required us to spend a lot of time looking for and trying on our preferred products, which took a lot of time, and when we were busy, we did not have a lot of time to choose, but as the Internet developed, the emergence of online shopping software allowed us to. However, as the Internet has grown, online shopping apps have appeared, enabling us to make choices while on the go and from the comfort of our homes, saving a lot of time costs and online goods compared to physical shops, there will be a certain discount, both for us to save time and save money. Big data development has increased convenience by directing us to our preferred products based on our searches and allowing us to select from our favorites.The application of big data will push us related videos based on the videos we have swiped, which really makes it possible to choose from the best, without the need for us to filter the videos. Big data is also a powerful tool for easing traffic pressure, as it can predict future traffic conditions and provide optimised solutions to improve them, which helps the transport authorities to improve their ability to control traffic on the road and prevent traffic congestion. Electronic medical record gathering is by far the most effective use of big data. Each patient gets a unique electronic record that contains information about their medical history, family history, allergies, and the outcomes of all diagnostic tests. Big Data gathers patient data to identify diseases as early as possible, which benefits patients by lowering the risk of health harm and lowering the cost of healthcare [1].

Every customer is a source of data for organizations. There are thousands of data sources when thousands of users assemble on a platform, and these data sources need to be refined in a brand-new fashion, which is cloud computing. Big data integration and analysis will result in new regulations that will help businesses better understand their customers, adopt realistic marketing plans, manage their markets, and save money. Cloud computing supports big data as the upper layer of computer resources [2]. By analysing and categorising customers' consumption behaviour through big data, customers' needs can be better understood and personalised and customised services can be provided to them. Businesses can discover new business opportunities and create new gap markets while improving the consumer experience.

The main reason why big data analysis can understand ourselves better than we do is that big data analysis is a process of analysing existing data with the help of technology, and every aspect of this process is carried out in strict accordance with the software programmes that have been written. This means that the results obtained through big data analysis are more objective and accurate.

The main difference between a computer and a human being is that it has an absolute reason and no emotions. When it performs analysis, no personal emotion can influence or interfere with it, and this is its advantage over human analysis.

## 3. The downsides of big data

As far back as ancient times, people would carve into stone to record time and events. As time went on and as times evolved, the number of things to record increased and the methods of recording improved.

DOI: 10.54254/2755-2721/6/20230953

And today's digital age requires more space to store converted digital bits. Although people have gradually entered the era of Big Data, this innovation has also created many difficulties for people. One of the typical characteristics of Big Data is that data is being generated in a constant flow, like an open water main, and data is constantly flowing out, which poses a problem for current Big Data processing systems. There were 33 zettabytes (ZB), or 33 trillion gigabytes, of data produced, copied, used, and consumed globally in 2018. This number increases to 59 ZB by the year 2020, reducing development costs, reducing storage costs and reducing operation and maintenance costs are long-term needs for enterprise big data storage, and it is important to consider these needs of enterprises while addressing the main issues of big data storage. Secondly, big data can also be maliciously used by some enterprises. Looking back at the development of big data, there have been various problems. For example, big data kill ripe, refers to in the context of big data, e-commerce platform using user information and algorithm technology on the consumer consumption of expensive data collection, analysis and mining, and then according to the consumption of expensive frequency, consumption habits, consumption of expensive geographical, consumption of expensive ability, etc. [3]. It is a form of price discrimination to carefully judge a customer's willingness to buy and pay for the same goods or services through extremely covert means. It is the result of a specific algorithmic process that is executed in batches, whereby groups of users who match the set characteristics are selected and treated in a similar way. In simple terms, when a user frequently uses software to book airline tickets, hotels, etc., the price for old customers will be more expensive than that for new customers, or in a music software frequently listen to the same song, originally not a paid song, after listening to a few times to generate a fee, this behaviour is the bad side of big data brings us. A taxi software has been exposed to analyse users' consumption level by recording their mobile phone model, frequent destinations, etc, so as to determine the user’s work and home address and distribute coupons with different prices to different customers. In this way, the company can retain the maximum number of customers for the same price. There is also takeaway software that will give different offers depending on the phone model. This phenomenon is ubiquitous in our lives.

### *3.1. Privacy and security*

Big data has ushered in a golden age of data utilization for human civilization, but it has also raised concerns about personal privacy and security that have stymied the growth of data value and government data availability. Users are tracked online and offline by a vast ecosystem of websites, applications, social media platforms, data brokers, and ad tech businesses, which collect their personal information. This information is assembled, transmitted, compiled, and used in real-time auctions, supporting a $22.7 billion business annually [4]. Without our knowledge, a vast chain of big data companies is selling our privacy for huge profits.

In order to offer convenience in this era of intelligence, data, and big data, we must also do a good job protecting information. While we enjoy the advantages, we must also know how to preserve our privacy by selectively filling out personal information in non-essential scenarios. Some people reflect on the Internet to brush up on a few English videos, then someone will call to ask if they want to buy a class or training, and there are even fraudulent calls that will precisely tell people’s personal information such as name, ID number, etc. [5]. This is actually because our information has been leaked and it is becoming increasingly difficult to keep people's personal information private, much to their chagrin. There are few people who have lost their money due to fraud. In addition to leaving people's legal rights unprotected, this raises crime rates and contributes to social unrest. It is an invasion of privacy when the big data platform and the so-called accurate pushing analyze personal data on citizens, social activities, electronic device use, etc. Individuals may quickly become misinformed about society and spend money they don't need to. The emergence of big data presents a threat to the national security of state secrets. More sophisticated technology will be needed if the state is to better defend its state secrets against leaks to other media.

DOI: 10.54254/2755-2721/6/20230953

### *3.2. The quality of data*

The need to address the data quality is the main drawback of working with big data. Before using big data for analysis, data scientists and analysts must make sure the information they use is accurate, pertinent, and in the right format for analysis. The reporting process is significantly slowed down as a result, but if businesses don't address data quality, the information they have worked so hard to gather and the data they analyze will be useless. On the other hand, information principles need to be shifted to substance. That is, people need to know exactly where the data collected from them is going and what it is being used for. Of course, this should be done with their authorisation. This is not just a technical issue, it is about how technology can serve the public more efficiently and securely. The concept of consent and the mechanism for seeking consent involve normative values, and if the technical applicability of the principle of informed consent is only mechanically emphasised, it is easy to lose sight of the ethical orientation and practical basis of privacy protection, and to deviate from a human-centred design approach and approach. Thus, it is necessary to establish the principle of "technology neutrality", to deconstruct the original principle of informed consent, and to design it in a more rational and refined manner in accordance with the new data environment. On the other hand, the core of dismantling the alienation of the informed consent principle is how to keep data privacy in the hands of users in the digital age, giving them decisive control and voice over their own data, promoting the materialisation of the informed consent principle, finding a balance between data control and data use, and avoiding the sacrifice of data privacy in exchange for convenient and efficient services. Controlling who our data is disclosed to, how it is disseminated, where it is located and how it is used requires limiting the flow of data, rather than stopping it altogether. Big data that is not handled properly can be extremely invasive of users' privacy.

### *3.3. Big data trustworthiness*

Recommending videos and ads based on our preferences is not an invasion of privacy, it records our actions and analyses them. When personal information about citizens, social activities, use of electronic devices, etc., is analysed by the platform for big data and pushed with precision, this is an invasion of privacy, as personal preferences and needs are completely autonomous and do not need to be changed by the platform to stimulate consumption and push information that is not subjective, which can easily cause unnecessary expenses and social misinformation. This can easily lead to unnecessary spending and social misinformation. Of course, big data is not absolutely correct. The results of data can be artificially altered through deliberate guidance. Some celebrities, for example, hire a water army to create public opinion for themselves. Once a topic has been raised, uninformed netizens will often follow the trend and magnify the event. The data collected by Big Data will then be deliberately directed rather than the true opinion of the people.

No one can be independent of Big Data, and most people don't want to be outside of it. In the era of big data, everything is data, everything is quantifiable, the essence of the world is data, and data, together with matter and energy, are the three elements that make up the world [6].

## 4. Conclusion

Big data brings new security problems, but it is also an important tool in its own right. Big data can do more harm than good for ordinary people, while it can do more good than harm for big companies and governments. So we need to strike a balance to make this technology work better and more securely for humanity. First of all, the amount of data is expanding quickly, current technology systems cannot support big data applications, big data theory and technology are still in their infancy, and the development of future information technology systems will necessitate disruptive innovation and change. Secondly, the masses have the right to know, and access to software should be more clearly marked with the usefulness of the intelligence collected. People should be allowed to give authorisation with full knowledge. The state should create more comprehensive laws to protect the privacy of users. Finally, fraud prevention should be made more widespread, especially among the elderly, so that losses can be minimised in the event of a breach of information. This paper only gives a general introduction to the

DOI: 10.54254/2755-2721/6/20230953

advantages and disadvantages of big data. This paper is lack detailed and enough data support. Future research could focus on how to improve the current system and legal constraints on big data to ensure users' rights and privacy.

## 5. References

[1] Zhang Haiqing. Research on Artificial Intelligence innovation and development of big data combined with cloud computing [J]. Industrial Innovation Research,2022,20:19-21.

[2] Zhang Kexin. Big data is analysed and the advantages and disadvantages of the application of [J]. Science and technology and innovation, 2021 (9):50-51. DOI: 10.15913 / j.carol carroll nki kjycx. 2021.09.022.

[3] Hu Yuanchong, Feng Yifan. Big data kill cooked [J]. Journal of consumer fair trading right protection in the inquiry of shanxi normal university (philosophy and social sciences edition), 2022 ploidy (01): 161-176. The DOI: 10.15983 / j.carol carroll nki SXSS. 2022.0112.

[4] Apple. A Day in the Life of Your Data, A Father-Daughter Day at the Playground, April 2021.:3 https://www.apple.com/uk/privacy/docs/GBEN_Privacy_ADITL.pdf

[5] Li Xi. Science and Fortune, Issue 15, chapter 7, section 1, 2018

[6] GleickJ. Theinformation: A History, A Theory, A Flood [M]. New York: Pantheon Books, 2011:7.