ATLANTIS PRESS

Advances in Social Science, Education and Humanities Research, volume 332

International Conference on Contemporary Education and Society Development (ICCESD 2019)

Research on Risk Prevention and Control of Educational Statistics in Big Data Era

Juan Xu
Hunan Modern Logistics Vocational and Technical College
Changsha, China

***Abstract*—Based on the analysis of the value and risk of data in the big data era, this paper studies the risks of current education statistics in big data era, including leakage, theft and misuse, tampering, and secondary utilization of data information. The credibility of the executive branch is reduced. In view of the risk of education statistics, the risk prevention and control countermeasures are proposed. Firstly, the data risk awareness should be improved. Secondly, technical and institutional levels should be enhanced. In addition, risk recovery and remediation mechanisms should be established, and equal rights and responsibilities should be established. Finally, a social integrity system should be established.**

Keywords*—Big data; Education statistics; Risk; prevention and control

I. THE VALUE AND RISK OF DATA IN THE BIG DATA ERA

Big data is popularly referred to as "huge, massive data," but it is not limited to this. The Big Data Era features big data profiles as Volume, Velocity, Variety, Value, and Veracity. What is the relationship between big data and things, migration, cloud, and intelligence? In fact, big data comes from informationization. Without the Internet, there is no big data. In the era of mobile Internet, especially the popularity of smart and mobile products, big data is also operating. Born. It can be said that there is no big data era without mobile internet.

A. The value of data in the big data era

Going out to take a taxi, there is a mobile phone; how long it takes to get to the unit, the mobile phone can tell you; Taobao, the webpage eye-catching position is the topic you care about; the goods pushed by the merchant are exactly what you are prepared to start; the friends who have been separated for many years help through the circle of friends You are looking back, etc. What is the convenience that you are used to? It is big data.

It's no exaggeration to have the power to own big data, because the value of big data is reflected in what you can't think of, no big data can't do. Speaking of the value of big data, we have to mention the business model changes brought about by big data. Big data is the foundation of the emergence of new business models. Big data is constantly subdividing and tapping market opportunities, prompting enterprises to transform new business models. Ali Finance is an example of big data applications and business model innovation. The emergence of apps, applets, mobile services, precision push, sharing economy, new retail, artificial intelligence, and micro-business are all examples of the value of big data. The value of big data is unquestionable and clearly visible. It is precisely because the value of big data has enabled many enterprises and data mining companies to fully collect, store and mine data. It can be said that no enterprise can generate immunity to data at present. However, value and risk are often concurring. Some organizations, enterprises, and individuals are driven by the value of big data, unscrupulously possessing data, illegally reselling data information, illegally using data, and monopolizing data [1] are also applied.

B. Big Data Age Risk

1) Privacy disclosure[2]

The rapid development of mobile internet and artificial intelligence brings convenience, while the personal information of the public and consumers has become the main target of data mining. With the popularity of apps and small programs, personal information is actively disclosed or passively acquired. The main reason for active disclosure is the risk awareness and risk prevention awareness of individuals lacking data. For example, most young people like to disclose personal positions, preferences, and concerns. The dynamics, etc., do not know, personal unconscious information disclosure is of great benefit to data acquirers, or commercialized, tapping useful value; or being maliciously abused, even infringing on state and trade secrets. Passive acquisition is the process owner's use of bundled information to obtain personal information[3], such as check-in and location information. According to the survey, 1/3 of consumers like to share their geographic location with retailers with smartphones, and many Both the website and the app have selected the option by default, or the "Apply for Public Location Permission", "Apply for Public Avatar Permission", etc., which appear when using the program, and the public's lack of risk prevention awareness, so it is easy to click "Agree". The phenomenon, some even more, when you refuse, certain features are not available, these are cases of passive disclosure of personal information.

2) Hacking and illegal stealing

Some specialized hacking activists, their stealing of data and information is terrible and catastrophic, because these organizations or individuals often aim at attacking or influencing the country's politics, economy, and even social disorder. Of course, there are also money and interests. Drive. These specialized organizations or individuals are often organized and led. They are unscrupulous for personal or group purposes. Once hacking attacks occur, websites, web pages are tampered with, data is interrupted, and key data is illegally stolen. The consequences are unimaginable. The security of the country will be threatened. For example, Dianchi was permanently closed due to hacking incidents. On

1.Hunan Provincial Social Science Results Review Project (Project Number:XSP19YBC289),
2.Hunan Provincial Bureau of Statistics Project(2018 A39)2: Management and Open Research on Education Statistics in the Context of Information Disclosure,
3.Hunan Modern Logistics Vocational and Technical College School-level Project (JYC201819) .

February 24, 2013, the mail system of China Tibet Net was implanted in the back door. The Chinese version of the Chinese enterprise sub-station was attacked by the US address and the page was maliciously tampered with[4], there is a hacker in the place where there is a network, there is a cyber attack.

3) Data abuse

Data abuse is mainly reflected in two aspects, the commercialization of data values and the illegal abuse of data. For most businesses, the main purpose of their data is to extract useful data from a large amount of data, and then to play the commercial value of the data, such as the birth of small programs, the promotion of personalized needs, Alipay personalized annual bills, etc. Etc. is the law found in big data and evolves into an example of stimulating business; while the illegal abuse of data is reflected in the illegal stealing of large amounts of data for illegal purposes, such as the sale of personal information, more than 5000 on typical Facebook. Ten thousand user information data is leaked; comprehensive use of multi-dimensional information such as personal information, life information, financial information, political inclinations, etc., single-dimensional personal information does not constitute a threat or risk, but the comprehensive use of multi-dimensional data information, the probability of occurrence of risks Greatly improved, big data killing and price discrimination are typical; for example, illegal misuse of data can be reflected in achieving political goals or disrupting social order, confusing the purpose of viewing, and threatening national security.

4) Reduced government credibility

From the big data killing to the information mortuary, the data risk incidents are constantly exposed, and the degree of influence of some publics on the credibility of government data management is affected to some extent. Where is the source of the problem? The development of technology is not synchronized with the pace of legislation. With the attention and fermentation of public opinion on relevant issues, individuals began to pay attention to personal information prevention to a certain extent, and began to consciously avoid information exposure and prevent information leakage. To a certain extent, it affected the open data of organizations and individuals, and reduced the credibility of government data management. .In addition, for government data managers, when the legislation, technology, and systems for data security and data risk prevention cannot keep up with the development of big data, they will consciously choose the scope and disclosure of data disclosure, thus the government's credibility. The reduction, after all, open government data is seen by many countries as a good medicine to improve the government's fairness, transparency, accountability and increase citizen participation.

5) Threatening state secrets and security

The stealing of state secrets and the threat of national security can be said to be the most serious risks, and today it is normal to use big data analysis to obtain intelligence and state secrets. Terrorists, criminals, and some hacker attackers often use revenge and resistance to use the network to retaliate, attack military and defense information to achieve their national private information, security information, defense, military information, and national critical infrastructure information. Access to threaten national security, trade secrets, infringement of national information sovereignty, influence and disrupt government business activities, disrupt social order, stir up people's hearts, and undermine national critical infrastructure. Vulnerability is the weakness of information systems. Criminals, terrorists, etc. all use cyber attacks to achieve the purpose of destroying state secrets or threatening national security. When we consciously, unconsciously, passively or actively generate data, big data breaks the tradition. The secrecy habits have changed the way information is collected, and anti-stealing and leak prevention has become a major challenge. Therefore, the government may need to re-examine and define the scope of confidentiality and confidentiality.

II. THE RISK OF EDUCATION STATISTICS IN THE BIG DATA ERA

Educational statistics are no exception, and there is also a data risk due to the enormous value of the data.

A. Data information disclosure

The disclosure of student information has become an indisputable fact. Whether it is the school management department or the enterprises and institutions related to the school, it has become a possible subject of basic information leakage. Xu Yuyu's telecom fraud case is a typical personal information leak. When this precious life is over, we can't help but sigh, who leaked the information? When did the behavior of reselling information begin? Where is the loophole? Data leaks, especially The negative impact of the disclosure of sensitive data on society is enormous, and the impact on personal safety is serious. Of course, there are also active or passive disclosures of data due to students' own risk perception and risk prevention awareness. As the saying goes, when you unconsciously click "Agree", your privacy has been collected; when using certain Software, the bundle settings you don't know have leaked your privacy; when you sigh convenient, convenient, accurate service push, you may enter the information boudoir, you may be spending "exclusive price", your Information may have been leaked.

B. Data information theft and misuse

One of the risks of big data is the secondary use of data. For data collectors, even if it signs a confidentiality and use agreement with the data subject when collecting data, it cannot control the data it collects and stores. Theft, that is, the secondary use of data, often, the secondary use of data often has a certain purpose, it is the driving of interests and purposes to have data theft, so that for data owners and data users In addition, there must be corresponding legislative norms; in addition, even if the data is collected for a fair, transparent, and open purpose, there is no shortage of other data for future use. Misuse of data, whether intentional or unintentional, deviates from the value of the data itself, especially when the data is misused to produce erroneous conclusions, and the decision-making guidance of the education management department is guided. The data itself is not right or wrong. The key is to look at how to use educational data. The key is to look at the purpose of collecting educational data. The key is to see what the channel is for accessing educational data.

C. Data information tampering

For hackers or lawless elements, illegally obtaining data, tampering with web pages, and tampering with data in the form of cyberattacks, planting viruses, etc. for economic or political interests, when data from state, regional, educational administration, schools, and other management systems are stolen, and the key data files have been tampered. When the webpage has been tampered with, it is a fatal blow to the national education cause, the education administrative department, and the education decision-making guidance to the school management. Imagine that if there is a guiding error in the education cause, the consequences will be unimaginable. It is lethal. For data theft and tampering, it is possible to obtain and tamper with the data of a country, region, and education administrative department, and apply these data to other fields to understand and control the educational development of a country or region. It is extremely easy. Therefore, this is why the National Education Statistics System must undergo strict network security monitoring when it goes online.

D. Data information secondary utilization

In the omnipotent era of big data, people have given the magic of all aspects of big data, so everyone is paying attention to and collecting big data. No matter what the collector knows and how to use it, but they always believe it. These data are always useful, or even if you don't use the data yourself, you can get the benefits by reselling the data. The key point is here that “data information is used twice”. There is no right or wrong in the data itself. When the data subject provides data, it is also informed of the data and agrees to provide it, but what you don’t know is that the data may be Sub-use, multiple use, seemingly unrelated things may use the same batch of data, for example, the admission information of students enrolled in the record may be used for lending, part-time, etc., when the student's personal privacy sensitive data is illegal When it is acquired or resold and then used twice, it will infringe on the personal safety of the students and even the family.

E. Reduce credibility of the education administration

It is the constant emergence of data risk cases in the field of education that makes the public question the credibility of the education administration. As a parent, as a social citizen, naturally, will the student's information not only be provided to the school or the education administrative department? What is the reason for the information leakage? Whether it is the reason that the students’ own risk prevention awareness is not strong, whether it is the loopholes in the management of the education administrative department or the school management department, no matter whether the network security technology can not keep up, no matter If the relevant laws and regulations are not perfect, the public will decline the credibility of the education administration department, and naturally there will be a sense of reducing the amount of data information provided. Conversely, for the education data management department, in the context of data sharing and information disclosure, considering the privacy leakage and state secret disclosure that may be involved in data open sharing, it is natural to reduce the amount of data opening and reduce the scope of data opening[7]. the data that should be shared is not shared due to lack of laws and mechanisms, which also affects the credibility of the data management of the education administration that the public expects.

III. Analysis of Risk Prevention and Control Measures for Educational Statistics

A. Improve data risk awareness

Because the public's perception of data risk is not in place, the awareness of protecting personal information is weak. Therefore, it is necessary to publicize risk awareness from macro to micro, top to bottom, raise public awareness of risks, enhance data ethics awareness and vigilance, ethical awareness, and let the public have a certain awareness of risks, so as to consider protecting personal information. Promote the relevance of big data, so that the public can prejudge the possible use of the data and the information that may be predicted to protect the privacy of the individual concerned. For students, it is mainly based on school education, actively promoting vigilance awareness and safety awareness; for data managers, enhancing moral awareness and ethical awareness, strengthening data security management skills training; for network technicians, improving themselves Quality, strictly in accordance with the network security law to regulate network management; for the management department, develop and improve data risk management regulations and systems.

B. Technically strengthen prevention

Big data comes from informationization, and data risks also come from informationization. The network plays an important role in providing convenience and promoting the development of social law, but the network has its own vulnerability. Therefore, in today’s mobile Internet era, it is increasing. There is no end to network security construction. Security is the basic guarantee for the opening of educational data. Only by doing a good job of network security, education statistics can play its due role. Network security construction requires a large number of R&D talents and funds for security. Therefore, the state, enterprises and institutions It is necessary to increase network security investment, cultivate (introducing) network technology talents, and use innovative technical means to ensure network security construction, such as establishing a data security database, and continuously improve the environmental security and data security of data systems.

C. Strengthening prevention in the system

One of the reasons for the frequent occurrence of data risk events is the absence of laws and institutions and the lack of synchronization with data development. In the areas of data openness, open standards, licensing conditions, usage patterns, early warning mechanisms, prevention mechanisms, risk recovery and accountability mechanisms, most countries are still not perfect, and comprehensive data protection legislation is still lacking. Therefore, a series of systems[5] must be established for the scope of data disclosure, data usage rights, etc.; and a protective system for the open standards of data sharing and the usage patterns of data, so as to maximize the protection of public privacy, after all, big data The subject of the source is the individual; for the administrators of the school and the education administration, the acquisition, storage and use of educational data must be in accordance with ethical and ethical norms. On the basis of publicizing the awareness of risk

prevention, the risk prevention system is formulated. Based on the awareness of the foresight, a risk early warning mechanism, a risk response mechanism, a risk remediation and risk responsibility mechanism, and a power and responsibility equalization mechanism are formulated. Identify risk management processes and build an institutional framework for risk management of educational data.

D. Establish risk accountability mechanisms and risk remedies

We do not want risks to occur, but organizations and individuals that use illegal or malicious use of data for illegitimate interests must use laws and regulations to punish them. Therefore, we must establish a risk-recovery mechanism to let risks disappear into the cage of risk-recovery mechanisms. Through the accountability of individuals and organizational units that have data risks, it can be effective, and at the same time, it can link accountability with the integrity of individuals and organizations, so that data risk generators do not hide in the supervision of social credit system. Where. When the risk occurs, the pursuit of responsibility may be second. The important thing is to take effective remedial measures to remedy the risk, reduce the risk range as much as possible, reduce the risk impact, and rationally transfer and mitigate the risk and violation of education data. Punishment and accountability may not really control the risk. On the contrary, it may proliferate. The real effective measure should be to establish and improve the legal system of education data risk management from the policy system level.

E. Establish a system of equal rights and responsibilities

Rights and responsibilities have always been contradictory bodies in pairs. They define the rights and responsibilities of managers in all aspects of the data management process. In addition to the corresponding legal requirements for the corresponding management departments and management personnel in data collection, storage, use, sharing and backup. In addition, its responsibilities, rights and responsibilities should be clarified [6]. For the data owner, the principle of clear data collection should be clearly observed. When collecting data, the data subject should sign an agreement to let the data subject understand the purpose of collecting the data and inform the data subject how to handle the third party supervision after using the data. For data users, they should bear the corresponding responsibility while excavating the value of educational data, so that data users understand that they should pay for their own behavior while acquiring value, especially for the second user of the data. Strict regulations to minimize data purchases and multiple uses of data. At the same time, there is a limit to the time users and data owners use data and the timeliness of owning data. It can also stimulate the potential of data value mining while avoiding the endless use or abuse of data.

F. Establish a social credit system

The era of big data provides an opportunity for the establishment of a social credit system. At the same time, the era of big data has forced the establishment of a social credit system. In an era when everyone is chasing big data, the risks and risks caused by big data are also drama. increase. First of all, construct a social credit system from the legal level, let the law as the fundamental guarantee for the construction of the integrity system; then, establish a credit system to focus on inter-connected, shared and open, and truly open and transparent behaviors of dishonesty; in addition, to establish and integrity In addition to the professional credit evaluation agencies, the system, the government, enterprises, institutions, industry organizations, individuals, etc., as the third party supervision body, the government, enterprises, institutions, industry organizations, personal integrity into the social integrity system Let the subject act as the supervisor, participate in the supervision of the government's credit evaluation agency's execution, supervise the social integrity behavior, and also spontaneously regulate the participants' integrity behavior. Finally, we must increase the reward and punishment of honesty and dishonesty in spirit and economy to promote the standardization of the social credit system and let the social credit system monitor data risks.

IV. CONCLUSION

Data in the era of big data is playing an increasingly important role. It can be said that having big data has the potential to master wealth, but the value of big data coexists with the risks of the big data era. Education statistics are no exception. There are also data risks associated with the great value of data, including data leakage, data information theft and misuse, data information tampering, secondary use of data information, and reduced credibility of the education administration. risk.

In view of the five risks of educational statistics, risk prevention and control countermeasures are proposed. First, we must improve data risk perception. Second, we must strengthen technical prevention. Third, we must strengthen prevention in the system. Fourth, we must establish risk recovery and risk recovery mechanisms. Fifth, we must establish a system of equal rights and responsibilities, and we must establish a social credit system.

REFERENCES

[1] Big data risk management cannot be ignored [db/ol],.http://www.ce.cn/xwzx/gnsz/gdxw/201805/17/t20180517_29163203.shtml, 2018

[2] Shijie Cai, Xia Yizhen. On the identification and prevention of government data open risk [j]. Book and Information, 2017 (4): 104-112, 121.

[3] Who is your privacy protection in the era of big data [EB/OL].http://www.sohu.com/a/217542962_785805, 2018-01-18.

[4] China's overseas hacker attacks are becoming more and more serious - more than half of the attacks originated in the United States [n]. People's Daily, 2013-3-11 (21).

[5] Dongmei Zhang. Educational Innovation Trends and Risk Avoidance in the Background of Big Data [j], China Adult Education, 2016(10): 25-27.

[6] Juan Xu. Management Reform of Big Data Driven [j]. Statistics and Decision, 2015(06).

[7] Jing Cheng. Problems and Countermeasures of Effective Utilization of Educational Statistics in the Background of Big Data——Based on the Thinking of Basic Statistics in Higher Education[J]. Value Engineering, 2017-12-28.