2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)

2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE) | 979-8-3503-4745-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICDCECE57866.2023.10151409

Research of Cybersecurity Measures for Data Governance

1st Weili Jiang
School of Cyberspace Security, Hainan University
Key Laboratory of Internet Information Retrieval of Hainan Province
Haikou, China
jiangweili@hainanu.edu.cn

2nd Jun Ye
School of Cyberspace Security, Hainan University
Key Laboratory of Internet Information Retrieval of Hainan Province
Haikou, China
yejun@hainanu.edu.cn

3rd Yuyin Tan
School of Cyberspace Security, Hainan University
Key Laboratory of Internet Information Retrieval of Hainan Province
Haikou, China
tanyuyin2000@hainanu.edu.cn

***Abstract*—With the popularity and development of network technology, China has entered the digital era, data has become an important factor of production, people can more easily access, apply and process data during the development of life and work, and improve the utilization of data resources, in the digitalization of industry, especially the industrial Internet, CNC machine tools, industrial software and other relatively mature system is also actively exploring data-based digital transformation program. With the increasing importance of data in the digital economy infrastructure, the data security governance issues have become the focus of data governance as well as a difficult point.**

Keywords—Data Governance, Cybersecurity, Data Security

I. Introduction

With the development of the Internet, data in various fields are showing a surge, and these data are increasingly used to provide insights through analysis to inform critical business decisions. Several sensitive and private component contents are also covered in these data, and the cyberspace security [1] issues derived from them have become important issues in the current development of computer network technology. Among them, the illegal flow of public security data [2] has a high degree of threat to the overall security of the country, and such data once leaked will pose a substantial threat to national public security and political security. Therefore, data security governance [3] has also become an important part of the overall national security

In the context that China has entered the era of a digital economy, data resources have become an important strategic resource for the country and contain the significant strategic value, while the security of data is also closely related to the overall security of the country. The promulgation and implementation of "The Data Security Law Of The People's Republic Of China " in September 2021 marks that data security in China is formally protected by relevant laws, which also means that data security governance has raised higher requirements.

However, China's data security industry started late, and there are still many unresolved issues.

II. Taxonomy

This Research analyzes the existing cybersecurity problems of data governance from both management and technical aspects, and makes corresponding recommendations on the problems of data security governance, and forms a solution. According to this solution, a data security governance platform system is proposed.

III. Related Work

Since the era of the digital economy, data transmission has been increasing, data types have become more complex and data security risks have been increasing. For example, XinRui Wang et al. [4] analyze the existing problems of data security and introduce the important role of international standards in the field of big data security and privacy governance and the necessity of current domestic legislation in the field of data security. However, there is still a lack of data security governance-specific architecture analysis. Xianghui Tao et al. [5] introduce the risks and problems faced in data governance in the AI domain and propose a data governance framework for the AI domain. However, no specific measures are proposed for the overall data security situation and security measures. M. Kantarcioglu and F. Shaon [6] analyze the problems faced by artificial intelligence in the context of big data and propose a data management system for the field of artificial intelligence. Xiaolan Yu [7] analyzed data security issues in computer networks and proposed countermeasures to these issues but did not develop a framework for data security governance. K. A. Saed and N. Aziz et al. [8] explore data governance security in cloud data centers and propose a data governance assessment methodology that focuses on data security in cloud data centers.

IV. Security issues in data governance

With the development and application of artificial intelligence [9], Internet of Things [10], blockchain [11], cloud computing [12], and other technologies, global data are multiplying at a high speed, bringing convenience to people while also bringing challenges to data security governance issues. Although various international actors have realized the importance of data security governance, no unified governance framework has been formed for global data security governance. China needs to comprehensively and systematically analyze the various major risk factors affecting data security, accurately grasp global data security trends, and further optimize China's strategic choices in global data security governance.

A. Data risks arising from the diversification of new technology applications

With the popularity and development of the Internet, the continuous integration of applications and innovations among various Internet technologies, a large and complex amount of

979-8-3503-4745-6/23/$31.00 ©2023 IEEE

data is constantly generated, which brings new challenges to data security governance because some technical theories are not yet perfect or there are security flaws in the technology itself.

1) Large volume and many types of data

In the context of the digital era, a large number of information tools are used in various industries, resulting in a large amount and complex type of data every day, and a large number of data resources are in dynamic change. Traditional data governance cannot adapt to the current needs of data, and it is necessary to improve and innovate data governance to achieve the relative security issues of both security and availability.

2) Fragmentation of data governance rules

The fragmentation of rules for data governance is manifested in two main aspects.

First, different standards exist for different technologies to regulate related data, for example, there are huge differences in data security rules and data security research priorities for data sets related to the artificial intelligence field and cloud computing.

Second, from the industry's perspective, the types of data generated by each enterprise have large differences, the definition and use of data in different departments of each enterprise has a large difference, making some of the same data given different meanings, which invariably increases the difficulty of data governance. The differences in the strategies adopted by different departments in data governance make it impossible for different departments to achieve uniformity in data collection, storage, and processing, which affects the sharing and security of data within the enterprise, and the lack of uniformity in data processing within the enterprise also affects the authenticity, availability, and accuracy of data.

3) Increased Data Governance Requirements

The digital economy has given more characteristics to data, such as a large amount of data, many types, low-value density, as well as fast dissemination and high timeliness, thus putting forward higher requirements for data governance and making data governance more complex.

4) Technology Risk

Due to the characteristics of the technology itself or the technology is not perfect, resulting in the risk of data security issues, such as cloud computing and other technologies that rely heavily on data, there is a risk of sensitive data being accessed and leaked by the computing party when the user is using.

B. Data risk due to human factors

1) Hacker Attacks

Hackers usually have a high level of network security technology and are the main initiators of network attacks, which are a great threat to the security of social data governance. They use computer operating system vulnerabilities or network security system vulnerabilities to launch attacks, steal, tamper with, and delete data, resulting in data loss and economic loss to the target of the attack. Currently, common hacking techniques include backdoor attacks [13], information bombs, denial of service, network monitoring, DDoS [14], etc.

In addition, hackers also use computer viruses to steal, tamper, delete and other attacks on data. If it is a virus that specifically destroys the computer, it will change the system or software program, damage the computer function. Eventually, the computer server will crash due to the virus, and even spread to infringe the whole computer network.

2) Human operation problem

The computer is the platform for application software operation, and in the process of human operation of application software, the privacy data is leaked due to insufficient awareness of network security. Usually, shopping, email transmission, and information exchange in the network are all data circulation behaviors that leave traces in the network and do not pay much attention to the surrounding environment, which may lead to the leakage of users' secret information. In the case of enterprise, computers may lead to theft and tampering of enterprise data. The cybersecurity issue of data governance is triggered.

C. Lack of adequate understanding of data governance leads to data risks

According to a survey, no respondent was able to define data governance exactly [15]. Meanwhile, the results of another study showed that the lack of adequate knowledge of data governance in SMEs is one of the reasons for its failure [16]. The lack of knowledge of data governance in SMEs and the lack of clarity of responsibility in the implementation of data governance as well as the confusion of data access rights lead to data management chaos and damage to data.

D. Data security governance technology application level issues

Data security governance and privacy protection are inevitable requirements for the development of the digital age. Data security governance cannot be carried out by management and regulation alone. Also, technologies such as privacy protection and data encryption algorithms are very important. During the actual network security technology application, the lack of network security technology application and limited protection will inhibit the enhancement of the effectiveness of network security protection and the enhancement and innovation of technical advancement. Data governance technology is the basic guarantee to achieve data security governance and an important way to achieve data security governance.

E. Risks arising from data exchange

The value of data lies in the exchange, and with the development of the digital age data storage and network security issues arising from the exchange process are becoming increasingly important issues in data governance. The IP protocols used in the transmission of data do not take into account data transmission security issues and are easily intercepted and tampered with by unlawful elements during transmission. In addition, data storage due to software, hardware or system defects hackers can easily perform illegal operations on sensitive data.

V. Data Security Governance Solutions

A. Improve the legal system and raise awareness of data security protection

The implementation of data governance cybersecurity measures needs to be supported by a sound and powerful system to standardize and institutionalize cybersecurity management, to form a deterrent and deterrent effect on cyber-attacks. On the one hand, the data governance network security is elevated to the legal level, and the specification requirements for the use, sharing, and storage of big data are clarified to determine the operational responsibility of the data: on the other hand, helping network users to establish awareness of data security protection is an important way to improve the level of data governance network security, and users' security protection needs to be carried out from various aspects to enhance the computer network security mechanism from within, which can better resist network attack infringement. Conclusions

Data is valuable, which is an inevitable trend for the future transformation of society. At the same time, data security faces high cyber security risks, and the cyber security of data can be achieved through data security governance, which integrates the latest security technology and security management concepts, and has key positive significance in ensuring data quality and security. Based on the need for data governance, we will improve the construction of a network security protection system to comprehensively protect data security and maximize the value and benefits of data governance.

B. Scenario-based data security governance

In today's data volume proliferation, different data can have different security needs and business characteristics in different scenarios, so enterprises as security management policymakers need to start from a scenario-based perspective, to develop corresponding security management policies for different scenarios. In addition, it is necessary to use the principle of minimization to control the scope of user access, and to control the time and dimension of user access under the condition of ensuring data security and not affecting usage.

In addition, the government should improve the relevant mechanisms, promote the construction of scenario-based data security standards, and develop standards and regulations for various application scenarios.

C. Building a Data Security Governance Framework

An effective organizational structure is a strong guarantee for the success of data governance, with clear responsibilities and objectives that facilitate the achievement of strategic data governance goals. Data governance requires not only each part to be managed, but also an overall structure and idea to lead the construction of the program. The overall goal of the governance framework is data protection and security governance. This paper proposes an effective data governance framework. As shown in Fig.1

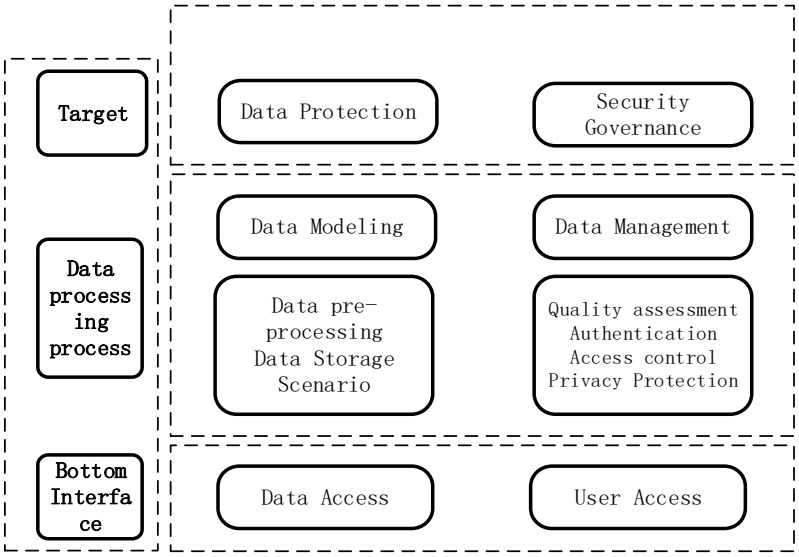

Fig. 1. Data Governance Framework

1) Determine the design principles of the data security governance framework

First of all, the data is standardized and classified, and different user rights are given to different data, with data as the core, to ensure that no matter in what scenario can get enough security protection, and can be used in a reasonable and legitimate place.

a) Data standardization modeling

The data in the new era is widely distributed and requires high privacy, which prevents efficient data sharing and the formation of data silos, and prevents these data from being fully utilized to bring out the value of the data. Realizing data standardization, data specification, data modeling, and other operations to provide support for data management and intelligence. Different data sources and data application scenarios need to be standardized to make the data easy to manage and meet the requirements. The specific scheme is shown in Fig.2.

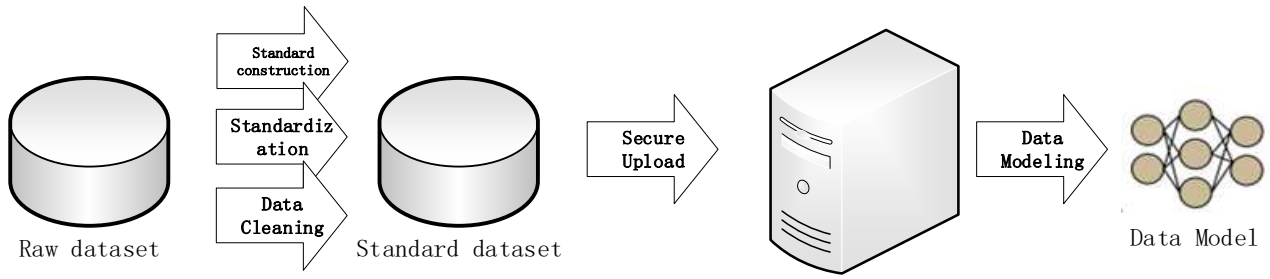

Fig. 2. Data Modeling Process

b) Data classification

The core of data protection is data classification and rating systems. Only by classifying data scientifically can we avoid data security problems with core data, perform detailed security management tasks, and strike a balance between data sharing and security.

c) User rights management

User rights are managed at a finer granularity, including different roles giving different corresponding rights. Through rights management, hierarchical control of data can be realized, and based on effective control of user behavior, malicious damage to data sources by illegal users can be effectively prevented, while the reasonable access needs of legitimate users can be ensured.

2) Effective application of data security governance key technologies

In data governance, network security protection requires the application of technical means to implement, which is an important guarantee of data security governance. The full application of network security technology and the design of a scientific and reasonable network security defense system can provide effective protection for data.

a) Data security risk-aware technology

Data security risk-aware technology is the basis of data security governance. During the construction of the network security protection structure, data security risk-aware technology is applied to analyze and monitor computer data hierarchically, and to conduct a preliminary analysis to identify potential security risks. The technology consists of two basic layers, namely the algorithm analysis layer and the data layer, where the data layer serves as the basis for classifying data assets, traffic, users, and security data, and then processing dynamic policies, features, statistics, and physical security in the algorithm analysis layer.

b) Authentication and authorization mechanism

The effective use of authentication and authorization mechanisms allows strict real-name authentication of user access to network resources and reduces the threat of illegal users to data security. It is usually divided into 3 parts - information authentication, identity authentication and authentication protocol.

c) Effective application of intelligent firewall technology

The application of network firewall technology is the main structure of modern network security protection, establishing a reliable line of protection for network data security. It can prevent network attacks and computer viruses from invading the network, and avoid network security problems caused by human operations. First of all, on the premise of clarifying the situation of the computer equipment system, we should choose the intelligent firewall technology with applicability, build an effective firewall system, and ensure that the application parameters of firewall technology can match the parameters of computer software. Secondly, the application of firewall technology should be reasonably planned according to the actual situation to ensure the perfection of functions and the full play of technical advantages, so that the firewall system can monitor computer vulnerabilities and risks in real-time and enhance security protection.

d) Data encryption technology

Data encryption technology is one of the effective means to ensure network information security. By encrypting computer network data, the information can be effectively concealed and prevented from data theft. Different encryption algorithms are adopted for different data, and plaintext data is converted into ciphertext data, so that even if the encrypted data is stolen during transmission, the stealer cannot recover the information content, thus ensuring the security of information data in the transmission process.

e) Data Security Risk Assessment and Policy Alignment

Perceived information about data risks is the basis for conducting data security risk assessments. Based on the collection of such information, a comprehensive risk assessment engine can be used to generate preliminary assessment results. The actual assessment process, is often not limited to five aspects, such as environment, risk, subject, object, and action, and can be maintained, deleted, or expanded according to actual needs.

f) Data Quality Assessment

Define high-quality data, identify data quality rules, perform data quality assessment, and improve data quality. Test all aspects of data to ensure compliance, availability, validity, completeness, accuracy, etc. Improve data quality, and data quality improvement can better reflect the value of data.

3) Build data governance platform

The unprecedented level of informatization in the era of big data, and the research of data generation and processing methods

are advancing by leaps and bounds, paving the foundation for intelligent management. Data is defined as an important factor of production in the new era and is a fundamental strategic resource for the country. Therefore, building a data governance platform, improving data governance capabilities, and better-supporting information governance is necessary to ensure data security governance. The framework of the built data governance platform is shown in Fig.3.

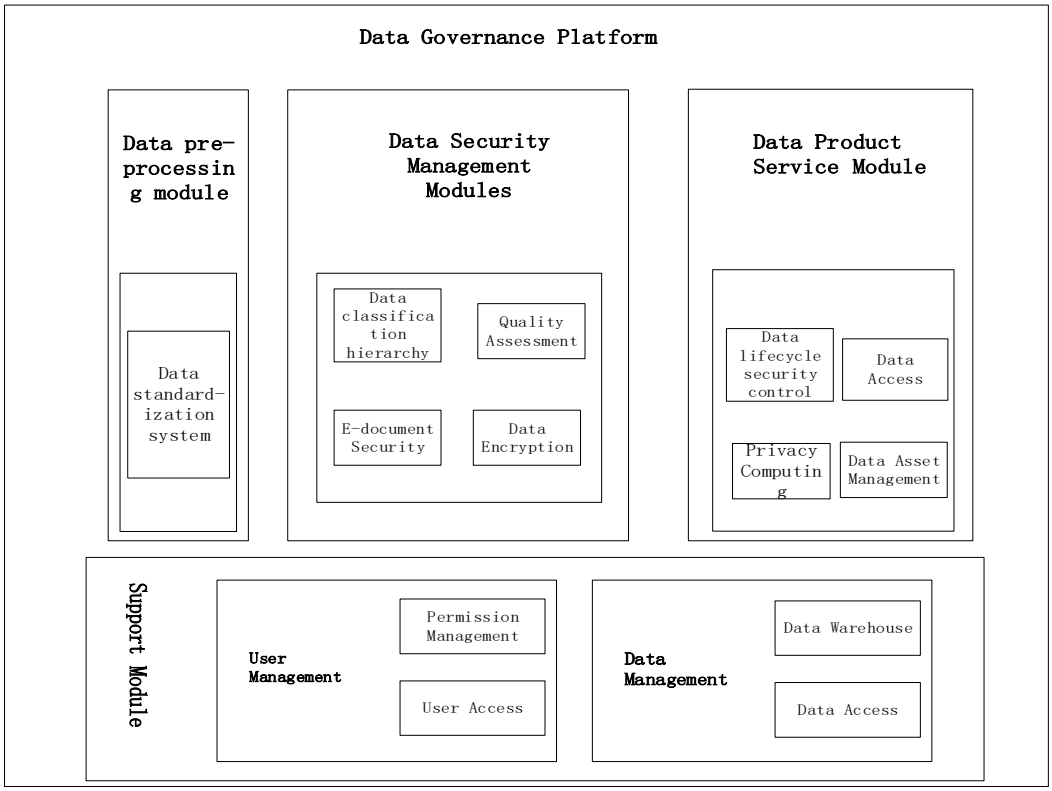

Fig. 3. Data Governance Platform

a) Data preprocessing module

Before the data circulation, the data is stored locally by the users themselves, and the form of the data is also developed by the users themselves. Therefore, we need to perform unified modeling and pre-processing operations before adding users' data to the data circulation. Data standardization operations include data standard construction, data mapping, standard maintenance, etc. Data standard development requires different standards for different application scenarios, and data mapping refers to the existence of different mapping relationships in the data standardization process, and the subsequent continuous maintenance of standards and mapping as well as the storage and unified management of these uploaded data to prevent management chaos.

b) Data Security Management Module

The data security governance platform will store the data after receiving the user data, but the data stored in the database will be exposed to the risk of data loss, tampering, and user data forgery, so the quality of the data needs to be assessed and monitored. The user sends the data to the data security governance platform, which verifies the availability of the data and ensures that the data has the appropriate value. To prove the compliance of their data, users first generate a commitment of their data compliance based on the data and then send the commitment to the data governance system, and then the data governance system generates a validation message and sends the validation message to the user, and the user generates a validation result based on the commitment generated by the user and the validation message from the data governance system, and the user sends this validation result to the data integrated governance system. Then the data governance system checks this validation result to determine whether the user's data is finally compliant. Similarly, for other assessments such as data quality, the data governance system follows the same process and finally integrates all the assessment results to get a comprehensive assessment result that can finally measure the data. After the quality of the data has been verified and stored in the database, the platform will continue to monitor the quality of the data.

c) Data Product Service Module

Let the whole process of data circulation is in a state of supervision, effective protection of data, and efficient management of data. The key lies in how to clarify the distribution location of the data, how the sensitive data is accessed, and in what way the data exists and is managed by those users. The platform needs to be clear and explicit about the

authorization of data on the platform, as well as the authorization of users and needs to have a perfect and strict access control system. Secondly, to manage the data comprehensively, it is necessary to manage the data scheduling. In the whole life cycle of the data, from uploading to the cloud to analyze the results, it is necessary to schedule the tasks to meet the accurate analysis. A scheduling system balances the work of the entire system and allows the platform to perform complex tasks in an organized manner. Finally, the platform uses privacy computing technology to achieve available invisibility of data, which guarantees the security and availability of data. Without fully trusting the user, only the calculation results are exposed to the user, and the data privacy is protected in the case that the user only gets the results, which also meet the user's needs.

VI. CONCLUSIONS

Data is valuable, which is an inevitable trend for the future transformation of society. At the same time, data security faces high cyber security risks, and the cyber security of data can be achieved through data security governance, which integrates the latest security technology and security management concepts, and has key positive significance in ensuring data quality and security. Based on the need for data governance, we will improve the construction of network security protection system to comprehensively protect data security and maximize the value and benefits of data governance.

ACKNOWLEDGMENT

This work is partially supported by the National Natural Science Foundation of China (No. 62162020), the Science Project of Hainan University (KYQD(ZR)20021).

REFERENCES

[1] C. Paulsen, E. McDuffie, W. Newhouse and P. Toth, "NICE: Creating a Cybersecurity Workforce and Aware Public," in IEEE Security & Privacy, vol. 10, no. 3, pp. 76-79, May-June 2012.

[2] Ivanc B, Blazic B J. Information Security Aspects of the Public Safety Data Interoperability Network[C]// Intelligence &Security Informatics Conference. IEEE, 2017.

[3] Liyuan Sun, Hongyun Zhang, Chao Fang,Data security governance in the era of big data: status, challenges, and prospects,Data Science and Management,Volume 2,2021,Pages 41-44,ISSN 2666-7649.

[4] X. Wang, W. Luo, X. Bai and Y. Wang, "Research on Big Data Security and Privacy Risk Governance," 2021 International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR), Shanghai, China, 2021, pp. 15-18.

[5] X. Tao and H. Zhang, "Research on data security governance based on artificial intelligence technology," 2021 International Conference on Big Data, Artificial Intelligence and Risk Management (ICBAR), Shanghai, China, 2021, pp. 102-105.

[6] M. Kantarcioglu and F. Shaon, "Securing Big Data in the Age of AI," 2019 First IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA), Los Angeles, CA, USA, 2019, pp. 218-220.

[7] X. Yu, "Analysis of the Security Strategy of Computer Network Data under the Background of Big Data," 2021 4th International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, 2021, pp. 13-16.

[8] K. A. Saed, N. Aziz, A. W. Ramadhani and N. Hafizah Hassan, "Data Governance Cloud Security Assessment at Data Center," 2018 4th International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 2018, pp. 1-4.

[9] Srinivas, M. and Dr. C. Krishna Mohan. “Medical Image Indexing and Retrieval using Multiple Features.” (2013).

[10] A. Samuel and C. Sipes, "Making Internet of Things Real," in IEEE Internet of Things Magazine, vol. 2, no. 1, pp. 10-12, March 2019.

[11] P. Frauenthaler, M. Sigwart, C. Spanring, M. Sober and S. Schulte, "ETH Relay: A Cost-efficient Relay for Ethereum-based Blockchains," 2020 IEEE International Conference on Blockchain (Blockchain), Rhodes, Greece, 2020, pp. 204-213.

[12] Guangyao Zhou, Wenhong Tian, Rajkumar Buyya,Multi-search-routes-based methods for minimizing makespan of homogeneous and heterogeneous resources in Cloud computing,Future Generation Computer Systems,Volume 141,2023,Pages 414-432,ISSN 0167-739X.

[13] Y. Ren, L. Li and J. Zhou, "Simtrojan: Stealthy Backdoor Attack," 2021 IEEE International Conference on Image Processing (ICIP), Anchorage, AK, USA, 2021, pp. 819-823.

[14] Ali Mustapha, Rida Khatoun, Sherali Zeadally, Fadlallah Chbib, Ahmad Fadlallah, Walid Fahs, Ali El Attar,Detecting DDoS attacks using adversarial n neural etwork,Computers & Security,Volume 127,2023,103117,ISSN 0167-4048.

[15] UBM, “The state of data quality,” 2018.

[16] C. Begg and T. Caira, “Exploring the SME Quandary : DataGovernance in Practise in the Small to Medium-Sized Enterprise Sector,” Electron. J. Inf. Syst. Eval., vol. 15, no. 1, pp. 3–13.