Designing Data Classification and Secure Store Policy According to SOC 2 Type II

Oleh Deineka[1], Oleh Harasymchuk[1], Andrii Partyka[1], Anatoliy Obshta[1], and Nataliia Korshun[2]

[1] *Lviv Polytechnic National University, 12 Stepana Bandery str., Lviv, 79000, Ukraine*
[2] *Borys Grinchenko Kyiv Metropolitan University, 18/2, Bulvarno-Kudriavska str., Kyiv, 04053, Ukraine*

Abstract

This paper discusses the design of a data classification policy for SOC 2 Type II compliance. SOC 2 Type II is a significant certification that attests to a service organization's ability to meet the Trust Services Criteria, which encompass security, availability, processing integrity, confidentiality, and privacy. Data classification is a critical first step in establishing a robust data security strategy, as it helps organizations understand what data they have and assigns a level of sensitivity to that data, which informs the security controls that should be applied. The main objectives of data classification are to organize and manage data in a way that enhances its protection and aligns with the overall data security strategy of an organization. Data security plays a pivotal role in the data classification process, as it directly influences how classified data is protected and managed. Designing a data classification policy for SOC 2 Type II compliance involves several challenges and considerations that organizations must navigate to effectively protect sensitive information and maintain the integrity of their service delivery. These challenges and considerations include understanding the scope of data, aligning with the Trust Services Criteria, balancing security with usability, training, and awareness, regular updates, and reviews, defining classification levels, ensuring consistency, automating classification, integration with other policies and controls, dealing with third-party vendors, monitoring and enforcement, and legal and regulatory compliance.

Keywords

SOC 2 Type II, data classification, data security, access management, storage.

1. Introduction

The modern world is characterized by a rapid growth of information assets, which contain a rather high percentage of critical information. Large volumes of such information primarily require classification by various parameters and features, their reliable storage and transmission, as well as protection from unauthorized access. Recently, the number of possible attacks on information resources has been constantly increasing [1–3]. Cybersecurity specialists are constantly developing new standards, approaches, and methods to counteract such malicious acts, as well as the development of infrastructure in this direction [4–9]. An important direction is the development of standards for safe data storage [10, 11]. Security standards allow a better understanding of how exactly an institution controls access to data and ensures their security and confidentiality [12].

The standards and requirements for data storage for organizations can vary depending on the country, the organization's industry, the sensitivity level of the information, and other factors. For a specific organization, there may

CPITS-2024: Cybersecurity Providing in Information and Telecommunication Systems, February 28, 2024, Kyiv, Ukraine
EMAIL oleh.r.deineka@lpnu.ua (O. Deineka); garasymchuk@ukr.net (O. Harasymchuk); andrijp14@gmail.com (A. Partyka); anatolii.f.obshta@lpnu.ua (A. Obshta); n.korshun@kubg.edu.ua (N. Korshun)
ORCID: 0009-0005-9156-3339 (O. Deineka); 0000-0002-8742-8872 (O. Harasymchuk); 0000-0003-3037-8373 (A. Partyka); 0000-0001-5151-312X (A. Obshta); 0000-0003-2908-970X (N. Korshun)

be specific standards and requirements dictated by its needs and legal requirements.
Most organizations or institutions form their security policy based on international standards, which are mostly carried out with the participation of external auditing companies that certify compliance with the standard [13, 14].

However, there are still many problems that professionals who deal with secure storage of large volumes of data encounter. For instance, they have to grapple with issues of data integrity, confidentiality, and accessibility. Ensuring that the information remains unaltered from creation through storage and retrieval can be a daunting task. Moreover, professionals have to guarantee confidentiality, so that only authorized individuals can access the data. They also need to ensure that the data is readily accessible when needed, which can be challenging in an era of rapidly increasing data volumes.

While there are a variety of effective approaches, methods, and ways to organize big data storage, there are still certain problems in this area. The issue of searching for the necessary information in unstructured data can be identified as a significant drawback.

ISO 27001 is a standard designed to ensure proper management of a company's digital assets, including financial information, intellectual property, employee data, and trusted third-party information.

In turn, SOC 2 certification is more recognized and is usually preferred by American and Canadian companies.

Another important point: SOC is divided into SOC 1, SOC 2, and SOC 3. The first is exclusively about financial control, and the third is mostly used for marketing purposes, so SaaS providers can focus solely on SOC 2.

The Service and Organization Controls 2 standard was developed by the American Institute of Certified Public Accountants using the Trust Services Criteria reliability criteria. SOC 2 provides an independent assessment of risk management control procedures in IT companies that provide services to users.

The standard pays special attention to data privacy and confidentiality, so it is turned to by such giants as Google and Amazon—for them, a high level of security and transparent data processing processes are especially important. External auditors are invited for certification. Their task is to study the implemented practices, check how the company follows its procedures, and how it registers changes in processes.

SOC 2 Type II is a significant certification within the landscape of data security and compliance. It serves as an attestation by an independent auditor that a service organization has not only designed its systems to meet the Trust Services Criteria but also that it operates effectively over time. The Trust Services Criteria encompass several critical areas: security, availability, processing integrity, confidentiality, and privacy.

The importance of SOC 2 Type II lies in its ability to build trust with clients and stakeholders. By demonstrating a commitment to stringent data management practices, companies can assure clients that their sensitive data is handled responsibly. This is especially crucial in sectors where data privacy and security are paramount, such as financial services, healthcare, and cloud computing.

Moreover, the audit process SOC 2 Type II helps organizations identify and mitigate potential security risks, ensuring that they maintain a strong security posture. This proactive approach to risk management is critical in an era where cyber threats are constantly evolving, and data breaches can have catastrophic consequences. Therefore, there is a constant search for new approaches and methods to ensure reliable data storage and user and device authentication where this data is stored [15–17].

In an increasingly regulated environment, SOC 2 Type II compliance can also support adherence to legal and regulatory requirements. This can help organizations avoid costly penalties and legal issues associated with non-compliance.

From a business perspective, SOC 2 Type II compliance can serve as a competitive differentiator. It signals to the market that an organization is a reliable and secure partner, which can be instrumental in winning new business and retaining existing customers [18].

The result of implementing SOC 2 is a report based on the AICPA Attestation Standards, section 101, Attest Engagement.

Types of SOC 2 reports:

Type I report contains information about the design of control procedures and the result of an assessment of the internal control system as of the date of the check. This type of report

is a starting point for further building SOC 2 Type II compliance.

Type II report proves compliance with requirements over a certain period. The organization must demonstrate adherence to control measures and policies during this period, which usually requires a certain degree of automation and long-term commitments.

Goal of the work: Development of a solution for optimizing the classification of organizational data and its appropriate storage by the SOC 2 Type II standard.

Task: Analyze the main requirements for data classification and their storage organization, identify shortcomings, and search for the optimal solution in terms of speed and economic efficiency to ensure compliance with the SOC 2 Type II standard.

2. Overview of Data Classification and its Role in Data Security

Data classification is the process of organizing data into categories that make it easier to manage and protect based on its level of sensitivity and the impact on the organization should that data be disclosed, altered, or destroyed without authorization. It is a critical first step in establishing a robust data security strategy because it helps organizations understand what data they have and assigns a level of sensitivity to that data, which informs the security controls that should be applied. Recognizing this is important as big data plays a key role in data analytics. It's the analytics that allows us to correctly understand and interpret this data so that it can be used for making correct and justified decisions, predicting trends, etc. It's important to understand that big data repositories are not just a "large database". The main difference lies in the fact that databases typically store structured data and have a fixed schema, while repositories of unstructured data can also store unstructured data and process large volumes of information. The main objectives of data classification are to organize and manage data in a way that enhances its protection and aligns with the overall data security strategy of an organization. The process involves assigning categories to data based on its level of sensitivity and the potential impact on the organization if that data were to be improperly accessed, modified, or destroyed. The objectives are as follows:

Identify Sensitive Data: Data classification enables organizations to determine which data is sensitive and requires more robust protection measures. This includes data such as Personal Identifiable Information (PII), financial details, health records, and intellectual property.

Facilitate Risk Management: By classifying data, organizations can better understand the risks associated with each type of data. Higher classification levels typically indicate a higher need for protection due to increased risk.

Enhance Regulatory Compliance: Many industries are governed by regulations that mandate the protection of certain types of data or dictate specific rules for their storage and access.

Data classification is critical for compliance with regulations such as the General Data Protection Regulation (GDPR), the Health Insurance Portability and Accountability Act (HIPAA), and others.

Enable Focused Security Measures: Through data classification, organizations can apply appropriate security controls where they are most needed. This targeted approach ensures that the most sensitive data receives the highest level of security, optimizing the use of security resources.

Support Access Controls: Proper data classification assists in the implementation of effective access controls. It ensures that access to sensitive data is restricted to authorized individuals based on their roles and the need-to-know principle.

Inform Data Lifecycle Management: Classification helps determine how data should be handled throughout its lifecycle, including retention, storage, archiving, organizing access to it, and secure destruction policies.

Prioritize Security Efforts: In the event of a security incident, understanding the classification of affected data can help prioritize response and recovery efforts, thereby minimizing the potential impact on the organization.

Raise Awareness and Accountability: It promotes awareness among employees about the types of data they handle and their responsibilities in safeguarding it, thereby fostering a culture of security and accountability within the organization [19, 20].

Data security plays a pivotal role in the data classification process, as it directly influences how classified data is protected and managed. The role of data security in data classification can be described through several key functions:

Defining Protection Measures: Data security is the driving force behind the selection of protection measures for each classification level. Once data is categorized, data security principles guide the application of appropriate security controls, such as encryption, access controls, and monitoring systems.

Risk Mitigation: Data security practices are essential in mitigating the risks associated with the handling of data. By understanding the classification of data, organizations can implement security measures that are commensurate with the level of risk, ensuring that sensitive data is afforded stronger safeguards.

Regulatory Compliance: Many data security frameworks and regulations require the classification of data as part of their compliance standards. Data security ensures that classified data is handled in a manner that complies with legal and industry-specific requirements, thus avoiding potential fines and legal action.

Access Control: Data security policies determine who has access to various classes of data, based on their need to know and authorization levels. By enforcing strict access control measures, data security helps prevent unauthorized access to sensitive information.

Data Lifecycle Management: The role of data security extends throughout the entire lifecycle of the data, from creation to disposal. Security measures are applied differently at each stage of the lifecycle, depending on the classification of the data.

Incident Response and Recovery: In case of a data breach or other security incident, the classification of the compromised data guides the incident response and recovery efforts. Data security teams can prioritize their actions based on the sensitivity of the data involved, ensuring that the most critical data is addressed first.

Awareness and Training: Data security involves educating and training employees on the importance of data classification and the correct handling of data according to its classification. This increases awareness and reduces the likelihood of accidental data breaches or leaks, enhancing the reliability of their storage and control of access to them.

Auditing and Compliance Monitoring: Data security involves regular auditing and monitoring to ensure that classified data is being managed by established security policies and procedures. This helps to identify and rectify any deviations or weaknesses in the protection of classified data [21, 22].

3. Challenges and Considerations in Designing a Data Classification Policy for SOC 2 Type II

Designing a Data Classification policy for SOC 2 Type II compliance involves several challenges and considerations that organizations must navigate to effectively protect sensitive information and maintain the integrity of their service delivery. Here are some of the key challenges and considerations:

Understanding the Scope of Data: Organizations must first identify and understand the types of data they handle, which can be a complex task, especially for large or data-intensive businesses. This involves mapping out where data resides, how it flows through the organization, and what data is critical for the operation or sensitive by nature.

Aligning with Trust Services Criteria: SOC 2 Type II revolves around the Trust Services Criteria set by the AICPA, which include security, availability, processing integrity, confidentiality, and privacy. A data classification policy must ensure that controls are in place to address these criteria appropriately for different categories of data.

Balancing Security with Usability: Implementing too stringent controls can hinder business operations, while too lenient controls can expose the organization to risk. Organizations must find the right balance to ensure data is both secure and accessible to authorized users as needed and with appropriate rights and privileges.

Training and Awareness: Employees must be aware of the data classification policy and understand their roles in maintaining

compliance. Training programs are essential to ensure that all personnel can correctly handle data according to its classification.

Regular Updates and Reviews: Data classification policies must be dynamic, reflecting changes in the business environment, emerging threats, new data types, and regulatory requirements. Regular reviews and updates to the policy are necessary to maintain SOC 2 Type II compliance.

Defining Classification Levels: Organizations need to define clear and practical classification levels that reflect the sensitivity and value of the data. These levels will determine the corresponding controls and handling procedures.

Ensuring Consistency: Consistency in how data is classified across different departments and systems is crucial. Inconsistencies can lead to gaps in protection and potential compliance issues, which can result in possible data loss or unauthorized access.

Automating Classification: Manual data classification can be error-prone and inefficient and can be quite time-consuming. Implementing automated classification solutions can help, but it is essential to choose tools that align well with the organization's specific needs and compliance requirements.

Integration with Other Policies and Controls: The data classification policy must integrate seamlessly with other organizational policies, such as access control, incident response, and data retention policies, and not slow down their operation.

Dealing with Third-Party Vendors: If third-party vendors manage or have access to the organization's data, they must also adhere to the data classification policy. This requires careful vendor management and sometimes additional contractual agreements or audits, regarding their rights and privileges.

Monitoring and Enforcement: Ongoing monitoring is needed to ensure that the data classification policy is being followed and that controls are effective. This includes regular audits and reviews, which are part of SOC 2 Type II requirements.

Legal and Regulatory Compliance: Organizations must consider various legal and regulatory frameworks that apply to their data and ensure that the classification policy helps them meet these obligations and does not contradict current legislation.

Addressing these challenges and considerations requires a strategic approach and ongoing commitment to maintaining a robust data classification policy. Organizations may seek guidance from compliance experts, legal counsel, and SOC 2 audit professionals to design and implement a policy that not only meets SOC 2 Type II requirements but also supports the organization's overall data governance strategy [23, 24].

4. Data Classification Policy Design

4.1. Requirements

While SOC 2 Type II itself does not prescribe specific data classification policies, it does require organizations to effectively manage and protect the confidentiality, privacy, and security of information, by the Trust Services Criteria (TSC). A Data Classification Policy is a critical component of meeting these criteria, particularly the Security criterion, which is common to all SOC 2 audits.

A SOC 2 audit measures the effectiveness of your processes and systems based on the Trust Service Criteria and checks compliance with information security standards and rules, including Common Criteria standards. Here are some general requirements that a Data Classification Policy should address to support SOC 2 Type II compliance:

Identification of Data Type: The policy should define the types of data handled by the organization, including sensitive data subject to SOC 2 considerations, such as PII, business confidential data, and intellectual property.

Classification Levels: The policy must establish clear classification levels that reflect the sensitivity of the data. Common levels include public, internal use only, confidential, and highly confidential.

Ownership and Responsibilities: The policy should define roles and responsibilities for data classification, including data owners, custodians, and users, and outline their responsibilities in maintaining data classification.

Handling Requirements: For each classification level, the policy should specify handling requirements, including storage, transmission, access controls, encryption standards, and end-of-life procedures.

Labeling and Marking: The policy should provide guidelines on how data should be

labeled or marked according to its classification to ensure that it is easily identifiable and handled appropriately.

Access Controls: The policy must address access controls, ensuring that access to data is based on the principle of least privilege and that only authorized individuals can access sensitive data.

Retention and Disposal: The policy should outline data retention periods and secure disposal methods for each classification level, ensuring data is not kept longer than necessary and is disposed of securely.

Training and Awareness: The policy should mandate regular training and awareness programs for employees to understand the importance of data classification and their role in it.

Auditing and Monitoring: The policy should include provisions for regular auditing and monitoring to ensure that classification controls are effective and being followed.

Incident Response: The policy should be linked to an incident response plan that addresses potential data breaches or loss, with procedures tailored to the classification level of the data involved.

Review and Update: The policy should specify intervals for reviewing and updating data classification procedures to ensure they remain relevant and effective as the organization evolves, data volumes increase, and new threats emerge.

Third-Party Vendors: If data is shared with or handled by third-party vendors, the policy must extend to these vendors, often requiring them to adhere to similar or compatible classification and handling standards.

To ensure alignment with SOC 2 Type II requirements, developing a Data Classification Policy usually demands a comprehensive understanding of the AICPA's TSC and the unique data protection requirements of the organization. Engaging with seasoned compliance experts or auditors who can give tailored advice and oversee compliance with the standard's stipulations is highly recommended. The AICPA's guidance and frameworks such as ISO 27001, when consulted and utilized, can offer invaluable inputs for the creation and sustenance of a strong data classification policy. It is crucial to identify and categorize data based on its sensitivity, importance, and regulatory mandates. Moreover, regular reviews and updates of the policy should be conducted to ensure its efficiency and continued compliance with SOC 2 Type II requirements [25–29].

4.2. Representation

A high-level overview of the interaction between a system and its users, outlining the different functions (use cases) the system is expected to perform and the roles that interact with these functions.

Considering the aforementioned requirements, we have developed the following structure (Fig. 1), which fully allows for data classification, ensures their storage, and authorizes access to them by SOC 2 Type II. The diagram mentioned in the document illustrates and provides detailed information about the various actions, processes, and roles that are necessary to fulfill the requirements for coverage.

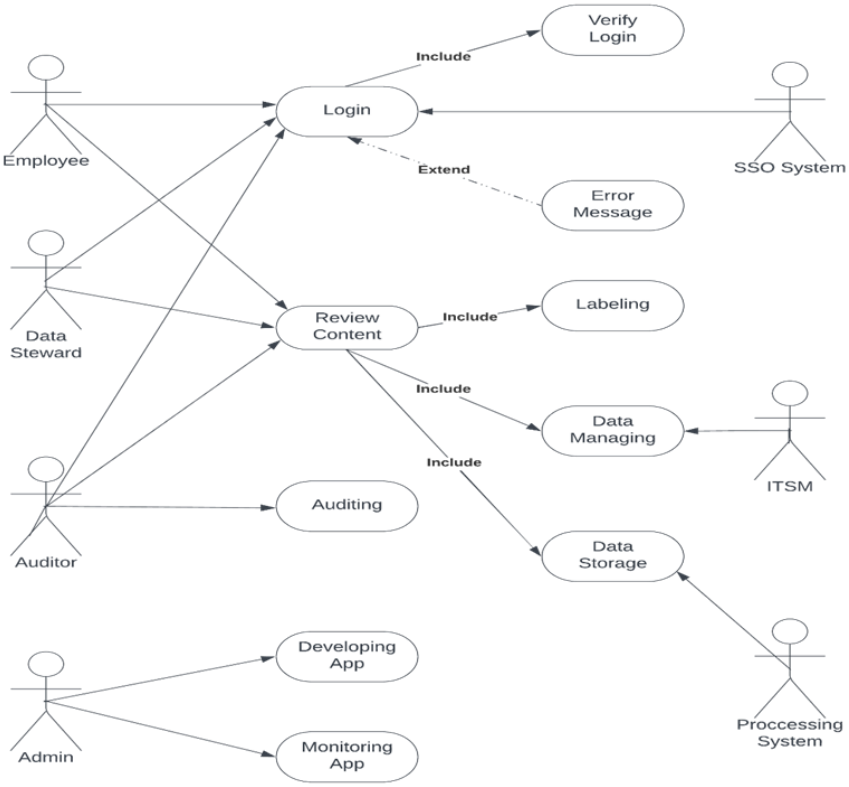

Figure 1: Use case diagram

4.3. Roles

We suggest applying for the following roles:

Employee: An employee is responsible for adhering to the data classification policy, correctly handling data according to its classification, and reporting any incidents or violations. As the primary users of data within an organization, employees are responsible for correctly handling data according to its classification level. This means that employees must understand the different classification levels and the corresponding handling requirements, such as storage, transmission, access controls, and end-of-life procedures. In addition to correctly handling data, employees are also responsible for adhering to the data classification policy and reporting any incidents or violations. This includes reporting any suspected data breaches, loss, or unauthorized access to data. By promptly reporting incidents, employees can help the organization to quickly respond and mitigate any potential damage or block unauthorized access to data. To fulfill these responsibilities, employees must receive regular training and awareness programs to understand the importance of data classification, the rules, and methods of such classification, and their role in it. This training should cover the data classification policy, the different classification levels, and the handling requirements for each level.

Data Steward: A data steward is responsible for the management and governance of data within the organization. They ensure that data is classified correctly, that the classification policy is being followed, and that data is being used in compliance with legal and regulatory requirements. Data stewards play a crucial role in maintaining the integrity of the data classification policy and ensuring that it is effectively implemented throughout the organization. They work closely with data owners, custodians, and users to ensure that data is correctly classified and that the appropriate controls and handling procedures are in place. Data stewards also monitor compliance with the data classification policy and report any incidents or violations to the appropriate authorities.

Auditor: An auditor plays a crucial role in assessing an organization's compliance with SOC 2 requirements, including the data classification policy. They are responsible for independently reviewing the policy, processes, and controls to ensure that they meet the Trust Services Criteria. The Trust Services Criteria

encompass several critical areas: security, availability, processing integrity, confidentiality, and privacy. The auditor's role is to provide an objective evaluation of the organization's compliance with these criteria and to identify any areas where improvements may be needed. This helps the organization to maintain a strong security posture, and reputation among its clients and partners, and to demonstrate its commitment to protecting sensitive data [19].

Admin: An admin plays a crucial role in maintaining the smooth operation of an organization's IT systems. They are responsible for deploying new app versions, monitoring system performance, and patching all operational staff. Here are some of the key responsibilities of an admin in this context:

Deploying new app versions: Admins are responsible for rolling out new versions of applications to ensure that users have access to the latest features and security updates. This involves testing the new version, preparing the deployment plan, and coordinating with other teams to ensure a smooth rollout. Monitoring: Admins are responsible for monitoring the performance and availability of IT systems. This involves tracking key metrics, identifying and resolving issues, and ensuring that systems are operating at optimal levels. Patching all operational staff: Admins are responsible for ensuring that all operational staff have the latest security patches and updates installed on their systems. This involves identifying and deploying patches, testing their effectiveness, and ensuring that all systems are up-to-date and secure.

SSO system: A Single Sign-On (SSO) system is responsible for managing user authentication and access control, the rights, and privileges of authorized individuals. It ensures that users are correctly authenticated and that they have access only to the data that they are authorized to access based on their roles and the data classification [30].

Processing System: A processing system plays a crucial role in managing data in line with the data classification policy. It ensures that data is processed, stored, and transmitted securely, adhering to the policies set forth by the organization. The processing system also involves data indexing, which is a method of organizing data to optimize its retrieval. This function is crucial as it makes the data search process more efficient, enabling users to locate and retrieve the necessary data quickly. Machine learning classification is an integral part of a processing system [31–34].

ITSM: IT Service Management (ITSM) is responsible for the delivery of IT services that support the data classification policy. This includes the provision of systems, tools, and processes that enable the organization to effectively classify, manage, and protect its data. ITSM also plays a key role in access management, request fulfillment, and incident management [32].

4.4. Actions and Processes for Review Customer Data

New customer information or category appears: If new customer information or category appears it should be added to Customer Data Catalog. Adding new customer information could include collecting additional details with any relevant information.

Adding a new category could involve creating a new grouping or segmenting the existing customer data into different categories. Adding new customer information or categories is usually done to improve the effectiveness of the Customer Data Catalog and enable to make more informed business decisions. However, it's important to ensure that the new information or category is collected and stored in compliance with data protection regulations and customer privacy laws.

The sensitivity level has changed: The sensitivity level of the data category refers to how valuable or confidential the information is, and how much damage or harm could be caused if it were to be disclosed or accessed by unauthorized individuals or entities.

When the sensitivity level of the data category has changed, it means that the level of importance or confidentiality of the data has increased or decreased. If a previously non-sensitive data category has now become sensitive due to changes in regulations, business practices, or legal requirements, the sensitivity level of that data category has increased. Conversely, if the sensitive data category has become less important or valuable due to changes in business practices or legal requirements, the sensitivity level of that data category has decreased. It is

important to review and assess the sensitivity level of the data category to ensure that it is being protected adequately and to make any necessary adjustments to security measures and access controls.

The description of an existing category has changed: If the data being collected for a specific category is changing or expanding, the description of that category may need to be edited to reflect the new data category being collected. Editing the description of an existing customer data category is usually done to ensure that the information being collected is accurately and completely described [22].

4.5. Data Flow Design

This diagram provides all data flow steps:
Step 1: Understanding the Types of Data Your Company Owns

The first step in creating a Data Flow Diagram is to understand the types of data your company owns. Data can be broadly classified into three categories: structured, semi-structured, and unstructured.

Structured data refers to data that is organized in a predefined manner, such as data stored in a relational database. Structured data is easy to search, analyze, and manipulate, as it follows a consistent format. Semi-structured data refers to data that has some level of organization but does not follow a strict format. Examples of semi-structured data include XML and JSON files, which contain data in a hierarchical format, but do not have a fixed schema.

Unstructured data refers to data that has no inherent structure or organization. Examples of unstructured data include text documents, images, and videos. Unstructured data can be difficult to search, analyze, and manipulate, as it does not follow a consistent format [35, 36].

Step 2: Understanding the Metadata Associated with Your Data Once you have identified the types of data your company owns, the next step is to understand the metadata associated with that data. Metadata refers to data that provides information about other data. For example, the metadata associated with a text document might include the author, date of creation, and file size. Understanding the metadata associated with your data can help you to better organize, manage, and analyze your data [22].

Step 3: Using Integration Tools to Manage and Store Your Data

After you have identified the types of data your company owns and the metadata associated with that data, the next step is to use integration tools to manage and store your data. Integration tools allow you to extract data from various sources, transform it into a common format, and load it into a data store. This process, known as Extract, Transform, Load (ETL), allows you to consolidate your data into a single location, making it easier to manage and analyze [31–34].

Step 4: Creating a Data Model

Once your data has been extracted, transformed, and loaded into a data store, the next step is to create a data model. A data model is a visual representation of the relationships between different data elements. It provides a framework for organizing and structuring your data and can help you to identify patterns and trends within your data [37].

Step 5: Classifying and Linking Your Data to Metadata

After you have created a data model, the next step is to classify your data and link it to the metadata associated with it. This involves assigning a level of sensitivity to your data, based on its importance and the potential impact if it were to be lost or stolen. Once your data has been classified, you can link it to the metadata associated with it, providing additional context and information about the data [38].

Step 6: Visualizing and Managing Your Data

The final step in creating a Data Flow Diagram is to create an application that allows you to visualize and manage your data. This application should provide a user-friendly interface for accessing, analyzing, and manipulating your data. It should also include logic for managing access, requests, and incidents, and should be integrated with your ITSM system to ensure that data is handled according to your company's policies and procedures [39, 40].

This solution presents a host of advantages over traditional product-based offerings from various companies. One of the key benefits is the flexibility to choose the hosting environment that best fits your needs, be it on-premise or cloud-based. This allows you to

align the solution with your operational requirements and infrastructure capabilities.

Furthermore, you have the freedom to select the technology stack that best suits your project. This means that you're not limited to a predetermined set of technologies, but can tailor the solution to leverage the most relevant and efficient tools for your specific needs.

In terms of team composition, you can assemble a team that is uniquely suited to the project at hand. This flexibility ensures that the right expertise and skills are applied to deliver the best possible outcomes.

Another advantage is the budgeting flexibility. Unlike vendor-specific solutions that may come with fixed licensing costs, the budget for this solution can be adjusted according to your financial capacity and project requirements. This can result in significant cost savings without compromising on quality or performance.

Lastly, this solution offers robust change and feature management capabilities. This means that it can easily adapt to evolving business needs, with the ability to incorporate new features and make necessary changes in a timely and efficient manner. This flexibility ensures the solution remains relevant and continues to deliver value over time.

4.6. Value of SOC 2 Type II Compliance

The SOC 2 Type II report has become a standard requirement for businesses looking to assure clients, partners, and stakeholders about the security of their data and systems. This report, issued by an independent auditor, offers an in-depth review and attestation of the effectiveness of a company's information security controls over some time.

The main reason why the SOC 2 Type II report is valuable to a company is that it provides clear evidence that the company has robust and effective controls in place to protect customer data. In today's digital age, data security is a top priority for businesses and customers alike. A data breach not only leads to financial loss but also damages a company's reputation.

The SOC 2 Type II report helps build trust with customers by demonstrating that a company has taken necessary measures to protect its data. It's a clear signal to clients that their data is safe, secure, and handled in a manner that meets or exceeds industry standards.

Another benefit of SOC 2 Type II is that it can provide a competitive edge. Companies that have achieved SOC 2 Type II compliance can differentiate themselves from competitors that haven't. This can be a decisive factor for potential customers when choosing between different service providers.

Furthermore, the SOC 2 Type II report can help companies avoid penalties related to non-compliance. Various laws and regulations require businesses to take certain steps to protect customer data. By achieving SOC 2 Type II compliance, companies can demonstrate that they are meeting these requirements, thus avoiding potential fines and legal complications.

The SOC 2 Type II report can also help companies identify and address vulnerabilities in their information security controls. The process of achieving compliance requires a comprehensive review of a company's information security policies and procedures. This can help identify any weaknesses or gaps that need to be addressed, thereby strengthening the company's overall security posture.

Lastly, the SOC 2 Type II report can help improve a company's internal processes. The process of achieving compliance requires a company to document and formalize its information security policies and procedures. This can lead to more efficient and effective processes, as well as a greater understanding of the company's information security risks and controls among employees [41, 42].

5. Conclusions

According to the document: In conclusion, designing a data classification policy for SOC 2 Type II compliance is a complex but crucial task for organizations. SOC 2 Type II is a significant certification that attests to a service organization's ability to meet the Trust Services Criteria, which encompass security, availability, processing integrity, confidentiality, and privacy. Data classification is a critical first step in establishing a robust data security strategy, as it helps organizations

understand what data they have and assigns a level of sensitivity to that data, which informs the security controls that should be applied. The main objectives of data classification are to organize and manage data in a way that enhances its protection and aligns with the overall data security strategy of an organization. Designing a data classification policy for SOC 2 Type II compliance involves several challenges and considerations that organizations must navigate to effectively protect sensitive information and maintain the integrity of their service delivery. These challenges and considerations include understanding the scope of data, aligning with the Trust Services Criteria, balancing security with usability, training, and awareness, regular updates, and reviews, defining classification levels, ensuring consistency, automating classification, integration with other policies and controls, dealing with third-party vendors, monitoring and enforcement, and legal and regulatory compliance. Addressing these challenges and considerations requires a strategic approach and ongoing commitment to maintaining a robust data classification policy. Organizations may seek guidance from compliance experts, legal counsel, and SOC 2 audit professionals to design and implement a policy that not only meets SOC 2 Type II requirements but also supports the organization's overall data governance strategy. The proposed solution aims to demonstrate the simplicity of the process that can be developed using the technologies and resources that are acceptable to the company within an affordable budget.

References

[1] B. Matturdi, et al., Big Data Security and Privacy: A review, China Communications, 11(14) (2014) 135–145. doi: 10.1109/CC.2014.7085614.

[2] V. Susukailo, I. Opirskyy, S. Vasylyshyn, Analysis of the Attack Vectors Used by Threat Actors During the Pandemic, IEEE 15th International Scientific and Technical Conference on Computer Sciences and Information Technologies (2020) 261–264.

[3] M. Islam, et al., Security Threats for Big Data: An Empirical Study, Int. J. Inf. Commun. Technol. Human. Dev. 10(4) (2018)1–18.

[4] A. Singh, A. Kumar, S. Namasudra, DNACDS: Cloud IoE Big Data Security and Accessing Scheme Based on DNA Cryptography, Frontiers Comput. Sci. 18(1) (2024) 181801. doi: 10.1007/s11704-022-2193-3.

[5] O. Harasymchuk, et al., Generator of Pseudorandom Bit Sequence with Increased Cryptographic Security, Metallurgical and Mining Industry Sci. Tech. J. 5 (2014) 25–29.

[6] V. Lakhno, et al., Management of Information Protection Based on the Integrated Implementation of Decision Support Systems, Eastern-European J. Enterprise Technol. Inf. and Controlling Syst. 5(9(89)) (2017) 36–41. doi: 10.15587/1729-4061.2017.111081.

[7] H. Hulak, et al., Formation of Requirements for the Electronic Record-Book in Guaranteed Information Systems of Distance Learning, Cybersecurity Providing in Information and Telecommunication Systems Vol. 2923 (2021) 137–142.

[8] V. Maksymovych, et al., Development of Additive Fibonacci Generators with Improved Characteristics for Cybersecurity Needs, Appl. Sci. 12(3) (2022) 1519. doi: 10.3390/app12031519.

[9] V. Maksymovych, et al., Combined Pseudo-Random Sequence Generator for Cybersecurity, Sensors 22(24) (2022) 9700. doi: 10.3390/s22249700.

[10] URL: https://secureframe.com/hub/soc-2/compliance-documentation

[11] URL: https://www.iso.org/standard/27001

[12] V. Buriachok, et al., Invasion Detection Model using Two-Stage Criterion of Detection of Network Anomalies, in: Workshop on Cybersecurity Providing in Information and Telecommunication Systems, vol. 2746 (2020) 23–32.

[13] P. Anakhov, et al., Evaluation Method of the Physical Compatibility of Equipment in a Hybrid Information Transmission Network, J. Theor. Appl. Inf. Technol. 100(22) (2022) 6635–6644.

[14] P. Skladannyi, et al., Improving the Security Policy of the Distance Learning

System based on the Zero Trust Concept, in: Cybersecurity Providing in Information and Telecommunication Systems, vol. 3421 (2023) 97–106.
[15] V. Maksymovych, et al., Simulation of Authentication in Information-Processing Electronic Devices Based on Poisson Pulse Sequence Generators. Electronics 11(13) (2022) 2039. doi: 10.3390/electronics11132039.
[16] J. Yi, Y. Wen, An Improved Data Backup Scheme Based on Multi-Factor Authentication, IEEE 9th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing (HPSC), IEEE Intl Conference on Intelligent Data and Security (IDS) (2023). doi: 10.1109/BigDataSecurity-HPSC-IDS58521.2023.00041.
[17] D. Shevchuk, et al., Designing Secured Services for Authentication, Authorization, and Accounting of Users, in: Cybersecurity Providing in Information and Telecommunication Systems II Vol. 3550 (2023) 217–225.
[18] A. Calder, S. Watkins, IT Governance: An International Guide to Data Security and ISO27001/ISO27002, Kogan Page (2019).
[19] ARMA International, "Information Classification: Getting It Right". URL: https://www.arma.org/
[20] Vic (J.R.) Winkler, Securing the Cloud: Cloud Computer Security Techniques and Tactics (2011). doi: 10.1016/C2009-0-30544-9.
[21] D. Alexander, et al., Information Security Management Principles, BCS, The Chartered Institute for IT, Updated edition (2013).
[22] M. Rhodes-Ousley, Information Security: The Complete Reference, Second Edition (2012).
[23] M. Harkins, Managing Risk and Information Security: Protect to Enable (2016).
[24] T. Peltier, Information Security Policies, Procedures, and Standards: Guidelines for Effective Information Security Management (2016).
[25] AICPA "SOC 2®—SOC for Service Organizations: Trust Services Criteria". URL: https://us.aicpa.org/ interestareas /frc/assuranceadvisoryservices/soc-for-service-organizations
[26] E. Gelbstein, IS Audit Basics: The Domains of Data and Information Audits, ISACA J. 6 (2016).
[27] U. Mattsson, Practical Data Security and Privacy for GDPR and CCPA, ISACA J. 3 (3) (2020).
[28] G. Pearce, Boosting Cyber Security With Data Governance and Enterprise Data Management, ISACA J. 3 (2017).
[29] D. Cannon, IT Service Management: A Guide for ITIL Foundation Exam Candidates, BCS (2012).
[30] A. Harper, et al., Gray Hat Hacking: The Ethical Hacker's Handbook, McGraw Hill (2015).
[31] C. Cote, M. Lah, Professional Microsoft SQL Server 2014 Integration Services (SSIS), Wrox (2014).
[32] S. Chauhan, Mastering Apache Airflow (2020).
[33] A. Gaikwad, Learning AWS Glue (2021).
[34] D. Anoshin, R. Avdeev, R. van Vliet, Azure Data Factory Cookbook (2020).
[35] N. Karumanchi, Data Structures and Algorithms Made Easy: Data Structures and Algorithmic Puzzles (2011).
[36] R. Watson, Data Management: Databases and Organizations (2017).
[37] S. Hoberman, Data Modeling Made Simple: A Practical Guide for Business and IT Professionals (2005).
[38] C. Aggarwa, Data Classification: Algorithms and Applications (2014).
[39] Y. Duhamel, Microsoft Power Platform Enterprise Architecture (2020).
[40] R. Collie, A. Singh, Power BI: Moving Beyond Power Pivot and Excel (2020).
[41] AICPA, Understanding SOC 2 Reports. URL: https://www.aicpa.org/ interestareas/frc/assuranceadvisoryservices/aicpasoc2report.html
[42] Why SOC 2 Type II Certification Matters. URL: https://www.alertlogic.com/blog/why-soc-2-type-ii-certification-matters/