GLAVNI KONCEPTI SIGURNOSTI PODATAKA / MAIN CONCEPTS OF DATA SECURITY

Snezhana Cherepnalkovska Dukovska[76], PhD

Atanas Kozarev[77], PhD

Violeta Rajcevska Luzevska

Prethodno priopćenje

Apstrakt: Kako do adekvatne sigurnosti podataka? Kako implementirati softverske alate koji sprečavaju curenje podataka? Kako pratiti podatke u oblaku da zaštitite podatke u mirovanju, u upotrebi i u pokretu? Kako zadovoljiti regulatorne zahtjeve da bude u skladu? Proveli smo istraživanje koje predstavlja aktivnosti i vještine za odgovor na ova pitanja. Implementacija politike koja koristi najbolje prakse je odgovarajuće rješenje za prevenciju sigurnosti podataka. Predstavljamo glavne koncepte koje ova politika treba da sadrži, a koja uključuje liderstvo, edukaciju radne snage, identifikaciju podataka koji zahtevaju zaštitu, identifikaciju svih zainteresovanih strana i korišćenje metrike za određivanje uspeha.

Ključne riječi: prevencija sigurnosti podataka, zaštita podataka, metrika za prevenciju podataka.

Abstract: How to obtain adequate data security? How to implement software tools that prevent data leakage? How to monitor data in the cloud to safeguard data at rest, in use, and in motion? How to satisfy the regulatory demands to be in compliance? We performed research that presents activities and skills to answer these questions. Implementation of a policy using best practices is the appropriate solution for data security prevention. We present the main concepts that this policy should contain, which are involving leadership, education of workforce, identification of data that requires protection, identifies all stakeholders and use metrics to determine the success.

Keywords: data security prevention, data safeguarding, metrics for data prevention.

[76] This research is a personal view of the authors

[77] MIT University, Skopje, North Macedonia

1. Introduction

Unintentional data loss occurs when confidential information leaves an organization's boundaries without explicit approval by authorized personnel. Innovations in areas such as the cloud and Internet of Things (IoT), reduces these boundaries. Entrepreneurial companies use new technologies and start up new services, which transmit and store data very frequently. The move to the cloud and the increasing usability of cloud services are accelerating the risk of potential data loss. It happens often that end users purchase applications without involvement from IT or IT security teams. Moreover, without end-user recognition, these actions create shadow IT departments. Shadow IT[78] increases the risk of data transmission and storage outside of organizational standards and controls.

The potential risk, inherited in protecting data, is higher if security measures are breached resulting in the improper use and disclosure of user data, or if services are subject to attacks that degrade or deny the ability of users to access products and services. In such cases, organizations may incur significant legal and financial exposure. Intellectual property rights are valuable, and any inability to protect them could reduce the value of an organization's assets. Privacy concerns relation to technology could damage an organization's reputation and repel current and potential users from utilizing its products and services.

Every organization has different strategies and future development directions, which influence on the process of building an appropriate orientation for data security. Therefore, every organization should pay due attention to developing a data loss prevention strategy. Moreover, if we want to apply IT tools, this strategy will have deal with details about the data, such as identification and classification, storage location and transmitting direction, level of significance and more. The main areas for implementation of prevention controls for data security are **Network** (preventing the loss of sensitive data from computer network, including email, web applications, and transfer protocols), **Cloud** (classifying and protecting sensitive data in cloud environments, including public, private, hybrid and multicloud environments) and **Endpoint management** (monitoring servers, computers, laptops, cloud repositories, mobile phones and other devices).

In addition, the most demanding part of every activity in today's working environment is compliance with regulatory requirements. Noncompliance can be costly. Europe's GDPR act, almost exactly five years since the strict rules came into force, Meta has been hit with colossal €1.2 billion fine ($1.3 billion) for sending data about hundreds of millions of Europeans to the United States, where weaker privacy rules open it up to US snooping. Meanwhile, Gartner predicts[79] that global spending on cyber security and risk management will increase by more than 11% in 2023. It estimates that organisations will spend $188 billion (€170 billion) in total on security, as threats increase and inflation drives up costs. The true cost of non-compliance for organizations due to a single event is an average of $4 million[80] in revenue.

This paper focus on data strategy and data loss prevention (DLP) framework, starting from the challenges of a need to maintain huge amount of data that is producing and using within the organization and consuming from client, as well as the cost of non-compliance with regulations.

[78] Shadow IT is any **software**, hardware or IT resource used on an enterprise network without the IT department's approval and often without IT's knowledge or oversight.

[79] Gartner Identifies Three Factors Influencing Growth in Security Spending

[80] The True Cost of Non-Compliance (saviynt.com)

In the first part of this paper, the possible states of data are explained and the main element for DLP framework are proposed. Next, the element of data security environment is discussed, which requires, as best, automation tool and (quantifiable) measures upon which the behavior of the data will be assessed. A conclusion is given at the end of the paper.

2. Data Security Framework- the main concepts

2.1. Different states of data in the system

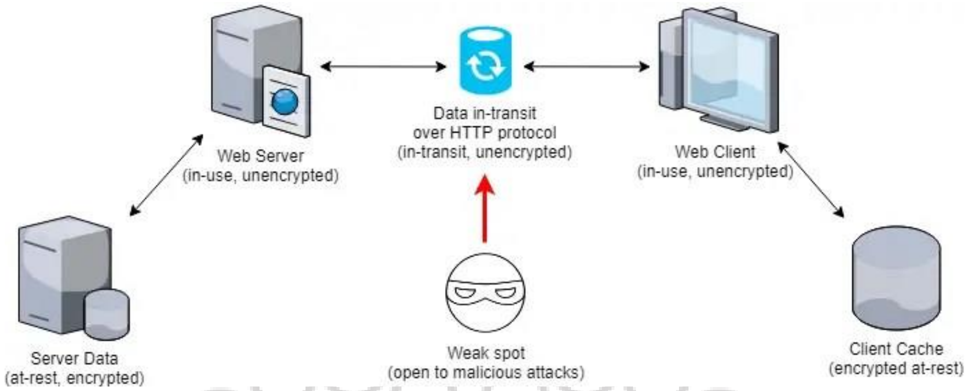
Different states[81] that data can go through and help to recognize how data is generated, stored, transferred and utilized are:

a. **Data in Creation**: This term refers to data that is being generated or produced. It encompasses the process of collecting, generating, or inputting information into a system or database. For example, when you create a new document, take a photo, record a video, or fill out a form, you are generating data in creation
b. **Data at Rest**: Data at rest refers to data that is stored or saved in a static state, typically on a storage device or within a database. In this state, the data is not actively being accessed, modified, or transmitted. It is simply stored and waiting to be used. Examples of data at rest include files stored on a hard drive, information saved in a database, or data archived in backups.
c. **Data in Transit**: Data in transit refers to data that is being transferred or transmitted from one location or system to another over a network or communication channel. During this stage, the data is in motion, moving between different devices or networks. Examples of data in transit include emails communication channel. During this stage, the data is in motion, moving between different devices or networks. Examples of data in transit include emails being sent over the internet, files being uploaded or downloaded, or any data being transmitted between servers.
d. **Data in Use**: Data in use refers to data that is actively being accessed, processed, or utilized by a system, application, or user. It is the state where the data is being acted upon or manipulated. For example, when you open a document and make changes to it, or when a program performs calculations based on input data, the data is considered to be in use.

On Figure 1 are presented IT resources where different states of data is located/transit, with appropriate controls (encryption) and weak point.

Figure 1: Different states of data where they resides in IT systems

[81] What Is Data In Motion: Encryption, States, Security And More (dataconomy.com)



Source: data at rest data in transif - Bing images

2.2 Data loss prevention framework

Focusing on the following key concepts of data security, we can establish a robust data loss prevention (DLP) framework[82] that protects sensitive data and mitigate risk of data breaches or unauthorized disclosure:

a. **Data classification**: Start by classifying your data based on its sensitivity and importance. Categorize data into different levels, such as public, internal, confidential, and highly confidential. Then, continue with categorization such as financial, non-financial, other, etc. Additionally, classify the data based on their storage location and possible transit and rest states. The classification helps determine the appropriate security measures to apply based on the data's value and potential impact if leaked.
b. **Access control**: Implement strong access control mechanisms to ensure that only authorized individuals can access sensitive data. This involves using authentication methods like passwords, multi-factor authentication, and role-based access control (RBAC) to grant permissions based on job roles and responsibilities.
c. **Encryption**: Apply encryption techniques to protect data both at rest and in transit. Use strong encryption algorithms to encode sensitive information, making it unreadable to unauthorized individuals who may gain access to the data. This includes encrypting data stored on devices, databases, or in the cloud, as well as data transmitted over networks.
d. **Data loss prevention (DLP)**: Implement DLP solutions that monitor and prevent unauthorized data transfers or leaks. DLP tools can detect and block sensitive data from being transmitted via email, web applications, removable storage devices, or other communication channels. They can also identify and prevent data breaches or policy violations in real-time.
e. **User awareness and training**: Educate employees about data security best practices and the importance of data protection. Conduct regular training sessions to raise awareness about data leakage risks, social engineering techniques, phishing attacks, and the proper handling of sensitive information. Encourage a culture of security-consciousness among all staff members. Perform penetration testing on regular basis.
f. **Monitoring and auditing**: Implement robust monitoring and auditing mechanisms to track data access and detect any suspicious or unauthorized activities. Use security

[82] What is data loss prevention (DLP)? | Microsoft Security

information and event management (SIEM) tools to monitor and analyze logs, generate alerts for potential breaches, and ensure compliance with security policies and regulations.

g. **Data disposal**: Develop proper procedures for securely disposing of data that is no longer needed. Ensure that sensitive data is permanently deleted or destroyed, following industry best practices for secure data disposal. This includes securely wiping data from storage devices or using physical destruction methods when necessary.
h. **Incident response and recovery**: Establish an incident response plan to address data leakage incidents promptly. Define roles and responsibilities, establish communication channels, and outline the steps to be taken in case of a data breach. Regularly test and update the incident response plan to ensure its effectiveness.
i. **Vendor and third-party management**: If you share data with external vendors or third parties, implement stringent security measures to protect data during transmission and storage. Conduct due diligence to evaluate the security practices of these entities and establish agreements that enforce data protection requirements.
j. **Regulatory compliance**: Stay informed about relevant data protection regulations, such as the General Data Protection Regulation (GDPR), California Consumer Privacy Act (CCPA), or industry-specific regulations. Ensure that your data leakage prevention framework aligns with these regulations and implement necessary controls to achieve compliance.

Including these several main concepts for building a data loss prevention (DLP) framework, we can revolve protecting sensitive data from unauthorized access, disclosure, or loss.

2.3. Data security environment

In data security environment, several additional important elements contribute for effective and efficient implementation of a robust data loss prevention (DLP) framework within the organization's security strategy, which follows:

a. **Leadership**: Effective data security requires strong leadership and commitment from senior management. Leadership should establish a data culture of security throughout the organization, prioritize data protection, allocate resources for security initiatives, and set clear policies and guidelines. They should also ensure that data security is integrated into the overall business strategy and promote a proactive approach to identifying and addressing potential security risks.
b. **Education of the workforce**: Workforce education is crucial for data security. Employees should receive comprehensive training on security best practices, data handling procedures, and the potential risks associated with data breaches. This includes educating employees about phishing attacks, social engineering techniques, password security, and the proper handling of sensitive information. Ongoing awareness campaigns and regular training sessions help reinforce a security-conscious culture.
c. **Identification of stakeholders**: Identifying stakeholders is crucial for effective data security. Stakeholders may include executives, data owners, data stewards, IT personnel, security teams, legal and compliance officers, customers, and third-party vendors. Each stakeholder has specific responsibilities and roles in data protection.

Understanding and involving all relevant stakeholders ensures a comprehensive approach to security, enables collaboration, and ensures compliance with regulations and internal policies.

d. **Use of metrics to determine success**: Metrics play a vital role in evaluating the effectiveness of data security initiatives and identifying areas for improvement. Key Performance Indicators (KPIs) and metrics should align with the objectives of the data security program. Examples of metrics include the number of security incidents, response and resolution time, employee training completion rates, data classification coverage, security audit findings, and regulatory compliance status. Regularly measuring and analyzing these metrics allows organizations to gauge their security posture, identify trends, and make informed decisions to enhance their data security program.

By addressing these elements in a data security environment, organizations can establish a strong foundation for protecting sensitive data, mitigating risks, and maintaining compliance with regulations.

Additionally, for more details, several controls will be explained through examples. If we are transferring data to third parties, we will need to monitor and continually improve the ways we secure it while in transit. Likewise, if we store information in cloud, we should run regular test s to make sure it is secure. Next, the biggest factor is the level of risk our organization faces. We should conduct a risk assessment as a part of our initial GDPR compliance program, but this process must be repeated annually to ensure you stay on top of threats. Risk assessments help identify the likelihood of data sets being breached and the amount of damage incidents would cause. The more substantial the threat, the more organizations must invest in defenses. This is also connected with the risk appetite that is a complex measure and hard to achieve on an adequate way. When we consider each of this issues, it is clear that there is no single answer for how much our organization should spend on compliance, one of the problems measuring leadership and strategical planning.

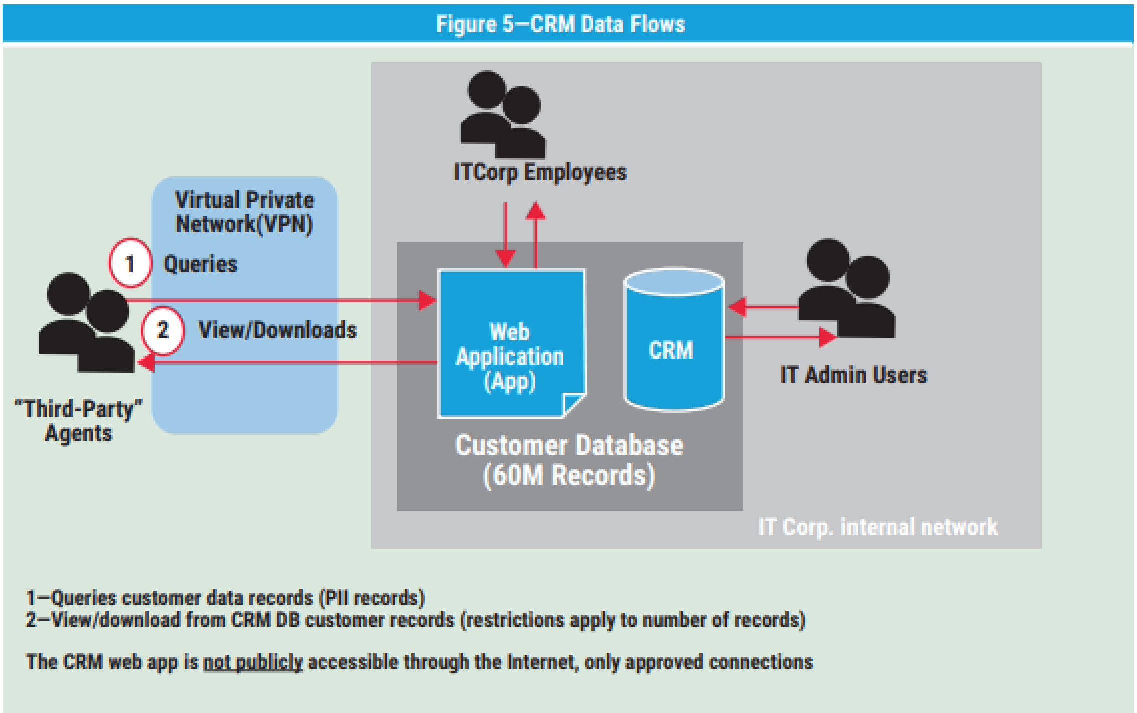
When adopting a DLP solution, we must do that justified by our needs. To avoid costly and other mistakes, we should document the deployment process, clearly define our security requirements and directly establish roles and responsibilities, with appropriate segregation of duties and assessing the misuse of policies and sensitive data.

It is very appropriate to use a tool that can block transition or encrypt data, in alignment with data classification and handling policies. It is important, this tool to protect the data at creation, in use, in transit and at rest. Further, it is important to include visibility and monitoring of data wherever they may flow. In addition, evidence available for demonstrating management and for lessons learned from incident management should be implemented. To achieve these, we must use an automation tool of the whole process to provide all steps required in our DLP framework.

To use an automation tool we, should set several measures that will indicate the system behavior as a whole. It is very important to have known indicators in advance, connected to threats, which the automation tool will act upon them based on determined rules. One example has been taken, given on Figure 2 (Layout 1 (isaca.org)), in which are demonstrated four threat scenarios with a purpose to introduce quantitative risk measures. The proposed system was not accessible to the public, and the internal threat was deemed most significant. However, in this example, external threats were not dismissed because customer data are always an attractive target to any malicious actors outside of the organization. In total, four threat scenarios were identified and documented as follows:

- Scenarios 1 and 2—Internal users access the sensitive personally identifiable information (PII) or other confidential information and extract the data for resell and misuse, applicable for privileged user (scenario 1) or general user (scenario 2),
- Scenario 3—Third-party user extracts the data via screenshots for sharing with competitors
- Scenario 4—Hackers access data for financial gain, ideology or espionage.

Figure 2: A simple data flow diagram is presenting to better determine the threat scenarios. (source: https://www.isaca.org/-/media/files/isacadp/project/isaca/articles/journal/2019/volume-4/evolving-from-qualitative-to-quantitative-risk-assessment_joa_eng_0719.pdf)



1—Queries customer data records (PII records)
2—View/download from CRM DB customer records (restrictions apply to number of records)

The CRM web app is not publicly accessible through the Internet, only approved connections

These quantitative risk assessment measures are important for DLP framework because they provide a more sound approach that is rich in meaningful data (connected to risk appetite) as opposed to the lightweight and judgmental qualitative-based method. Therefore, the example was intended to demonstrate that very simple indicators could be used for performing a quantitative risk assessment as well as more advanced simulation techniques (such as Monte Carlo simulations). Overall, the quantitative method provide a sound basis for further discussions with the stakeholders, giving them the ability to make well-informed decisions on the action plan, engaged in mitigation on risks.

3. Conclusion

The cost of non-compliance and loss of trust due to unintentional and intentional data loss are high. Leaders in the organization must focus on developing a data loss prevention (DLP) framework within data security strategy that provides the control, mitigate the risks and provide scalability, visibility and flexibility needed to meet the needs of the ever-changing threat and regulatory landscape. The proposed framework, security environment and usage of automation it toll with well-defined measures enables the organization to empower prevention of unintentional data loss.

4. Literature

[1] Benoit Heynderickx, CISA, CRISC, **Measuring Risk Quantitatively**
[2] ChatGPT, **Data Security Leadership Measures** (openai.com) (https://chat.openai.com/?model=text-davinci-002-render-sha, accessed on 30th of May 2023)
[3] **Evolving from Qualitative to Quantitative Risk Assessment. A Practitioner’s Dilemma** (https://www.isaca.org/-/media/files/isacadp/project/isaca/articles/journal/2019/volume-4/evolving-from-qualitative-to-quantitative-risk-assessment_joa_eng_0719.pdf), accessed on 30th of May 2023)
[4] **Gartner Identifies Three Factors Influencing Growth in Security Spending,** (https://www.gartner.com/en/newsroom/press-releases/2022-10-13-gartner-identifies-three-factors-influencing-growth-i), accessed on 30th of May 2023
[5] Jason Jiao, Ph.D., CPA, **Deploying a data security defense** **(**https://www.isaca.org/resources/isaca-journal/issues/2020/volume-4/deploying-a-data-security-defense, accessed on 30th of May 2023**),**
[6] Luke Irwin, **How Much Does GDPR Compliance Cost in 2023? -** IT Governance Blog En , (https://www.itgovernance.eu/blog/en/how-much-does-gdpr-compliance-cost-in-2020, accessed on 30th of May 2023)
[7] Sunil Bakshi, CISA, CRISC, CISM, CGEIT, CDPSE, AMIIB, MCA, **Performance Measurement Metrics for IT Governance** (https://www.isaca.org/resources/isaca-journal/issues/2016/volume-6/performance-measurement-metrics-for-it-governance, accessed on 30th of May 2023)
(https://www.isaca.org/resources/news-and-trends/isaca-now-blog/2019/measuring-risk-quantitatively, accessed on 30th of May 2023**)**
[8] **The True Cost of a Data Breach** (https://www.isaca.org/isaca-digital-videos/podcasts/the-true-cost-of-a-data-breach, , accessed on 30th of May 2023**)- video**