Whitepaper

Using data classification

Michel Gulpen

Sr. Security Consultant at Kembit

September 2016

Data classification, aren't we supposed to do something about that? If that's a question being asked within your organization, then this whitepaper can offer you the insight you need to take action. Even if you have never even thought about data classification, this whitepaper can be a source of information. The implementation of a data classification model is the first step towards setting up a detailed information security policy within your organization.

Contents

About KEMBIT

Different times require different approaches. Information technology shouldn't just connect people, it should move with the times. KEMBIT anticipates market development and turns your IT questions into opportunities for your business. Our approach is focused on your organization and using our IT expertise for optimum results, now and in the future.

Through a strong focus on the comprehensive and personal development of our 100 or so employees, KEMBIT has been delivering high end IT professionals and solutions since 1996. Our knowledge and expertise are put to use for a broad range of organizations, from local MKB companies to large, nationally operating businesses and multinationals.

Using data classification

Introduction

Information security means securing (business-critical) data. A detailed information security policy can increase the resilience of your organization by detecting vulnerabilities and enabling you to implement the correct counter-measures.

Data classification is done based on three pillars, namely: confidentiality, integrity and availability. By establishing vulnerability and its impact within your organization (and its data) in terms of one or more of the known data-features of confidentiality, integrity and availability, the risk level can be established. Using this risk level, a set of counter-measures can be established which are in line with the wishes and requirements of the organization for a detailed information security policy.

Data classification is an important basis for a successful information security policy. Detailed analysis within your organization is required to establish business-critical processes and to protect the associated data using the necessary counter-measures.

This whitepaper offers an insight into the necessary steps for the implementation of a data classification model within your organization, as well as certain conditions that have to be met.

Management summary

Data classification is all about classifying business-critical data within your organization. This whitepaper answers questions about why data classification is important for your organization.

Data classification is one of the first steps in the implementation of a successful information security policy within your organization. In order to successfully implement data classification, the first thing that needs to be done is to map out your business-critical processes. Using a detailed risk analysis based on confidentiality, integrity and availability, suitable counter-measures can be taken for the listed data.

This whitepaper is not only a recommendation for data classification levels, it also contains a set of counter-measures that can be implemented within your organization. Examples of these are:

- Encryption
- Safe deletion of media
- Network segmentation
- Server hardening
- File server access
- User Awareness
- Data Lost Prevention solutions

As well as the security of your internal organization, the successful implementation of a data classification model also helps you meet the respective laws and legislation, provided that the correct counter-measures are implemented. How these counter-measures are structured within your organization is something to be worked out using the data classification model to be implemented

Given that a data classification model, as well as the vulnerabilities and the impact of those vulnerabilities, addresses various risks based on confidentiality, integrity and availability, it will soon become clear to you that data classification customization is required for every organization. The solutions and recommendations set out in this whitepaper will provide you with a detailed framework for successfully classifying data within your organization.

1. Information security

Information security is based on three pillars, namely: confidentiality, integrity and availability (see figure 1).

Confidentiality

Data confidentiality indicates who or what (in a process) may and can see certain data. When counter-measures based on data confidentiality fail, publication of sensitive data can be a possible outcome. One of the most commonly used practical counter-measures is the use of encryption, which is applied to business-critical data or data flows within an organization.

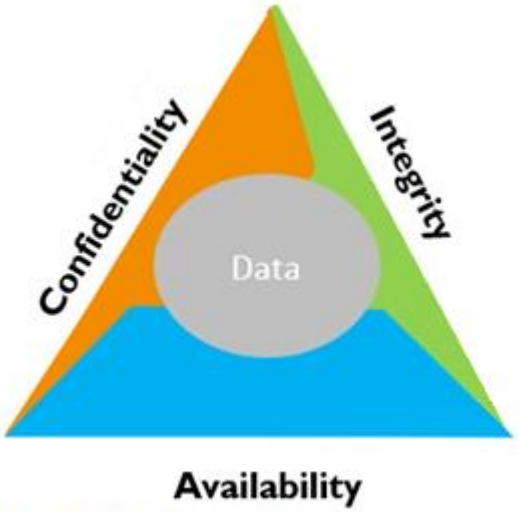

Figure 1: Data Classification

Integrity

Data-integrity is important to safeguard the accuracy of data. A common counter-measure is hashing, a technique in which the data is linked to a logical algorithm to form a given value. This value is then generated by the sender as well as the recipient with the same outcome, the data is not changed. This control factor is mainly valuable when dealing in process or finance related data.

Availability

Availability answers the question of when and to what degree information has to be available for the organization. Essentially, this phase involves mapping out business-critical processes and applying gradation in terms of the degree of availability of data. Another useful point is that a higher degree of availability results in higher costs for the counter-measures.

Using a data classification model, the impact on the three above stated pillars can be established. A risk analysis for the business-critical processes is essential for establishing a correct risk level based on confidentiality, integrity and availability

2. Three pillars of data classification

Business-critical processes

Establishing business-critical processes is essential for developing a detailed information security policy. Business-critical processes form the basis on which other processes and data in the organization, which may be of a less critical nature, can be worked out. It is important to develop a framework to establish data classification in the organization.

Risk analysis based on data confidentiality

Data confidentiality is simplified within an organization via a number of data classification levels. The higher the level, the higher the authority (i.e. options) must be for modifying, copying and deleting data. A data classification model based on confidentiality might have the following structure:

- Public: this data is not a direct risk for the confidentiality of the organization and is freely accessible.
- Internal: this data should only be available for internal staff and may not be published externally.
- Confidential: data in these classification categories may only be seen, modified or deleted by certain group(s) users.
- Secret: only senior management may see documentation in these categories.

Risk analysis based on data integrity

Data integrity means that the data is current; it determines whether or not the data is correct, complete and timely. A data classification model based on integrity might have the following structure:

- Not required: this data can be modified, there are no extra measures required to protect the data-integrity.
- Required: amending the data may cause limited direct or indirect damage to business operation. A basic set of measures is required to safeguard data-integrity.
- Obligatory: data-integrity is required for this process. Changes to the integrity level can have serious direct or indirect damage to business operation.
- Absolute: data in these categories must not contain any errors. To this end, counter-measures must be taken. Undesired changes to the data integrity level will cause significant or extreme damage to business operation.

Risk analysis based on data availability

Data availability means when and how much data has to be available for business operation.
A data classification model based on availability might have the following structure:

- Not required: not having data has no direct influence on the continuity of the process.
- Needed: data may occasionally not be available. If this is the case, then the impact on business operation and related processes is nil.
- Required: data should really be constantly available. If there is a threat to the continuity of the data, this can result in serious direct or indirect damage to business operation.
- Essential: this data has to be constantly available. A calamity based on non-availability will cause major damage, direct and indirect, to business operation.

Data classification models

After working out the data values in the categories of confidentiality, integrity and availability, it is a good idea to set up a generic classification document which recommends assigning a number between 1 and 4 to the given values, where 1 is the lowest value and 4 the highest (see figure 2).

Organizations wishing to adopt a data classification model are advised to verify a baseline based on the values attributed to integrity and confidentiality. These two values added together reflect the risk to which the organization is exposed.
As a baseline gives a minimum, it is advisable not to be too strict about the baseline. It can result in high costs for the organization.

Availability is difficult to measure. It is also advisable to document this properly and where necessary to apply suitable measures.

	Risk values			
	1	2	3	4
Integrity				
Not required	X			
Required		X		
Obligatory			X	
Absolute				X
Confidentiality				
Public	X			
Intern		X		
Confidential			X	
Secret				X
Availability				
Not required	X			
Needed		X		
Required			X	
Essential				X

Figure 2: Risk matrix

It is also advisable to use a generic standard classification model. An example of this is (from low to high):

- Unclassified
- Restricted
- Confidential
- Secret
- Top Secret

Data features

Data has three statuses, namely:

- Data in use: describes data used by an end user, a process and/or a system.
- Data not in use: describes data not in use and stored somewhere on a file server or a back-up.
- Data in transport: describes data being transported via a network.

If the requirements for each data feature might be different, then an impact analysis can be made for these data features based on confidentiality, integrity and availability.
For example, for data in transport, there might be more emphasis on confidentiality and integrity; whereas for data not is use, the emphasis might shift to availability and integrity. When it comes to data in use, all three pillars will apply (see figure 3).

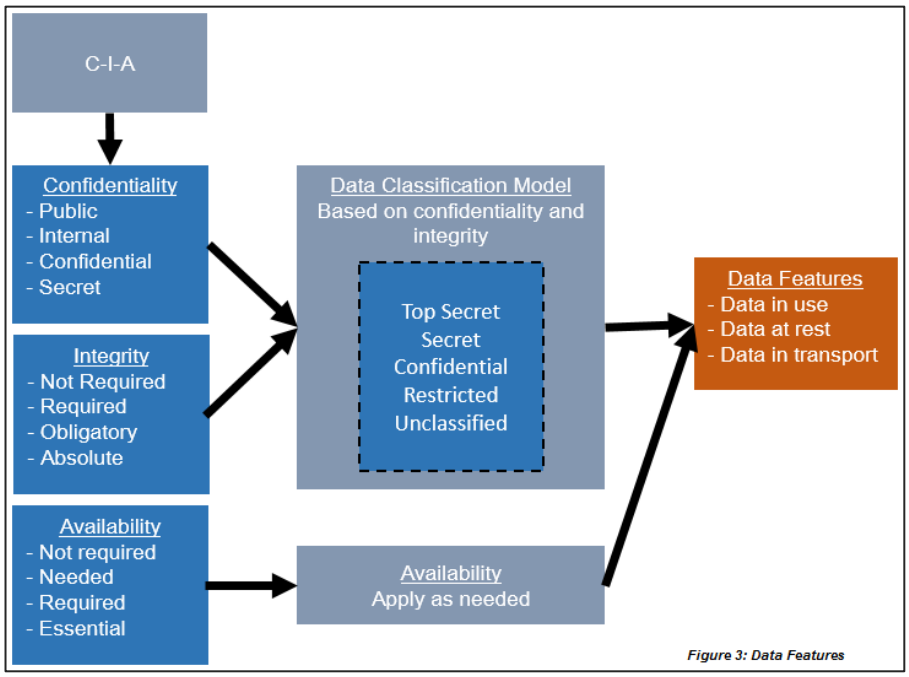

Figure 3: Data Features

Data ownership and responsibilities

The data owner is any designated person within an organization who determines the impact on data confidentiality, integrity and availability. An organization can have multiple data owners, which is why it is important from an organizational perspective to define clear roles and responsibilities. From a central perspective, guidelines should be set out in relation to the standard data classification model. It is imperative to stick to the data classification policy and the controls around this for successful data classification implementation in the organization.

3. Responsibility for reporting data leaks

In May 2016, the European Union implemented the General Data Protection Regulation (GDPR) with the intention to strengthen and unify data protection for individuals within the European Union (EU). It also addresses export of personal data outside the EU. The GDPR will lead to new Data Protection legislation in the member states of the EU. Part of this legislation is the new responsibility to report data leaks. The implementation of a data classification scheme within the organization can help to meet the required legislation.

The Data Protection Act states that 'appropriate security measures' must be taken to protect personal data. Not having this in place can result in a fine from the Data Authority. Organizations that hold patient dossiers or other personal, sensitive information must take serious steps to prevent any 'leak' in personal data.

A data classification model can help an organization to classify personal data based on confidentiality and integrity. As described in chapter 2, this can help to make sure that the correct counter-measures are implemented.
Besides, the data classification baseline gives a clear view of the standard counter-measures that need to be taken in relation to data in the organization and can also prevent or minimize data leaks.

4. Mitigating Measures

Below is a set of counter-measures. Data classification is the foundation of taking these measures when classification is also important for the information security policy within your organization.

Encryption

Data confidentiality is often ‘summarised’ as encryption. The implementation of encryption is a significant factor when it comes to requiring high levels of confidentiality.
Encryption can be implemented at file level if this is on a server. It is also possible to implement at network level, not excluding so-called VPN connections. At system level (end-points), hard drive encryption is firmly recommended.

What needs to be taken into consideration is the strength of the encryption protocol being used and if it is sufficient with regard to the security level. Also, Encryption management will have to be set up properly within the organization.

Safe deletion of media

Disclosure, publication of business-sensitive information, is a common feature in the news these days. The cause of this is not properly (or securely) clearing hard drives or other portable media. Such media storage often has a data trail and formatting this type of media is often not enough to delete all data trails. As a counter-measure the organization can agree a contract with a party that specialises in the secure removal of old data storage media.

Backup

Availability is often mentioned together with business continuity. Backup is vital when it comes to business continuity. Having a good backup policy, either on-site or off-site, can often be decisive in continuing business operation.

Using a so-called data retention policy, where a policy is defined in terms of the responsibility to store data in line with any respective legislation, is also highly recommended. In this policy, consideration also has to be given to the impact on the three pillars: confidentiality, integrity and availability.

Data Lost Prevention Solutions

These days there are all kinds of so-called data lost prevention solutions on the market, aimed at monitoring data with a particular classification. The implementation of a data classification model that suits the current organization is the basis of this. Data lost prevention solutions operate on a variety of implementation levels, namely:

- Network-based
- Cloud-based
- End-Point-based
- (File) Server-based

User Awareness

Ultimately, the end user within the organization has to work with the data classification model. That means making sure users have had the correct training to work with the data classification model to avoid any risk.