Recommendations for a methodology of the assessment of severity of personal data breaches

Working Document, v1.0, December 2013

European Union Agency for Network and Information Security

www.enisa.europa.eu

About ENISA

The European Union Agency for Network and Information Security (ENISA) is a centre of network and information security expertise for the EU, its member states, the private sector and Europe's citizens. ENISA works with these groups to develop advice and recommendations on good practice in information security. It assists EU member states in implementing relevant EU legislation and works to improve the resilience of Europe's critical information infrastructure and networks. ENISA seeks to enhance existing expertise in EU member states by supporting the development of cross-border communities committed to improving network and information security throughout the EU. More information about ENISA and its work can be found at www.enisa.europa.eu.

Authors

This paper has been produced by experts of the Data Protection Authorities of Greece and Germany in collaboration with ENISA.

Editors

Clara Galan Manso (Seconded National Expert)

Sławomir Górniak

Contact

For contacting the authors please use sta@enisa.europa.eu

For media enquires about this paper, please use press@enisa.europa.eu.

Acknowledgements

We would like to thank the representatives of the Data Protection Authorities participating in the Article 29 Working Party Technology Subgroup for their valuable feedback which will help us further improve the methodology. We would also like to thank them for the submission of test cases which helped us evaluate the effectiveness of the methodology.

ISBN: 978-92-9204-078-9 DOI: 10.2824/27590

Acronyms

Art.29 WP – Working Party on the protection of individuals with regard to the processing of personal data

CB – Circumstances of the breach

DC – Data Controller

DPA – Data Protection Authority

DPO – Data Protection Officer

DPC – Data processing context

EI – Ease of identification

ENISA – European Union Agency for Network and Information Security

TS – Technology Subgroup (of the Art.29 WP)

Executive summary

The European Union Agency for Network and Information Security (ENISA) reviewed the existing measures and the procedures in EU Member States with regard to personal data breaches and published in 2011 a study on the technical implementation of the Art. 4 of the ePrivacy Directive, which included recommendations on how to plan and prepare for data breaches, how to detect and assess them, how to notify individuals and competent authorities and how to respond to data breaches. A proposal of a methodology for personal data breach severity assessment was also included as an annex to the above-mentioned recommendations, which was, however, not considered mature enough to be used at national level by the different Data Protection Authorities.

Against this background, the Data Protection Authorities of Greece and Germany in collaboration with ENISA developed, based on the above mentioned work, an updated methodology for data breach severity assessment that could be used both by DPAs as well as data controllers. This working document is a first result of the co-operation between experts of the two DPAs and ENISA. It is planned to further develop the methodology with the aim to generate a final practical tool for a data breach severity assessment.

An overview of the proposed methodology is presented in section 2 of this paper, and further elaborated in subsequent sections. Severity of a breach is defined as the estimation of the magnitude of potential impact on the individuals derived from the data breach. The core elements that have to be taken into account when assessing this severity are:

- Data processing context – type of breached data adjusted to the context in which they are used
- Ease of identification of the individual based on the data breached
- Circumstances of the breach, having additional influence on the severity of a breach

The methodology presented in this study is based on an as objective approach as possible whilst still being flexible enough to be adopted by various Data Protection Authorities by adjusting it to their size, national legal system and other factors. According to different requirements, the scoring of some categories can be adjusted to produce the most appropriate results.

Table of Contents

Acronyms iii

Executive summary iv

1 Introduction 1

1.1 Background information 1

1.2 Objectives 1

1.3 Severity of data breaches 2

2 Overview of the methodology 3

2.1 Criteria 3

2.2 Calculation of the severity 3

3 Detailed description of scoring and severity levels 4

3.1 Scoring of the criteria 4
3.1.1 Data Processing Context (DPC) 4
3.1.2 Ease of identification (EI) 4
3.1.3 Circumstances of the breach (CB) 5

3.2 Definition of severity level 6

3.3 Flags 6

4 Use of the methodology 7
4.1.1 Notification to competent authorities 7
4.1.2 Notification to the individuals 7

5 Closing remarks 8

Annex 1 – Data processing context 9

A1 Assessment tables 9

A2 Description of contextual factors to be considered in DPC scoring 11

A3 Examples of DPC scoring/adjustment per category of data 12

Annex 2 – Ease of identification (EI) scoring 17

Annex 3 – Examples of the circumstances of the breach (CB) scoring 19

A1 Loss of confidentiality 19

A2 Loss of integrity 19

A3 Loss of availability 19

A4 Malicious intent 20

1 Introduction

1.1 Background information

With the amendment of Directive 2002/58/EC[1] (ePrivacy Directive) an obligation for the notification of personal data breaches by the providers of publicly available electronic communication services to competent authorities and affected individuals was introduced (art 4). The European Commission, as also stipulated in the Directive, has published implementing measures mainly on the format and circumstances of the personal data breach notification – Commission Regulation (EU) No 611/2013[2].

Following the provision of the ePrivacy Directive, a proposal for a general obligation of the data controllers for the notification of personal data breaches under certain conditions has been introduced in the draft Regulation on the Protection of Personal Data (art. 31) in the context of the overall Data Protection Reform Package[3].

Both these legislative provisions constitute important developments with the potential to increase the level of data security in Europe and foster reassurance amongst citizens on how their personal data are being secured and protected by data controllers. As part of the effective implementation of the data breach notification obligation, the Article 29 Working Party concluded, through its ePrivacy and technology subgroups, that there is a serious need for the development of a methodology for the assessment of the severity of personal data breaches. The discussions within the Article 29 Working Party are currently still ongoing.

The European Union Agency for Network and Information Security (ENISA) reviewed the already existing measures and the procedures in EU Member States with regard to personal data breaches and published in 2011 a report on the technical implementation of the Art. 4 of the ePrivacy Directive[4]. A proposal of a methodology for personal data breach severity assessment was included as an annex to that report, which was, however, not considered as mature enough to be used at national level by the different Data Protection Authorities.

Against this background, the Data Protection Authorities of Greece and Germany in collaboration with ENISA developed, based on the above mentioned work, an updated methodology for data breach severity assessment that could be used both by DPAs as well as data controllers. This working document is a first result of the co-operation between experts of the two DPAs and ENISA. It is planed to further develop the methodology with the aim to generate a final practical tool for a data breach severity assessment.

1.2 Objectives

The proposed methodology presented in this document has been designed with the following objectives:

- To provide data controllers with a quantitative tool (to the extent that this is possible) to assess the severity of data breaches and accordingly notify the competent authorities as well as the

[1] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:337:0011:01:EN:HTML
[2] http://eur-lex.europa.eu/JOHtml.do?uri=OJ:L:2013:173:SOM:EN:HTML
[3] http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf
[4] Recommendations on technical implementation guidelines of Article 4, http://www.enisa.europa.eu/activities/identity-and-trust/risks-and-data-breaches/dbn/art4_tech

affected individuals. The tool could also serve as a means for data controllers to quickly determine the necessary mitigation measures.

- To provide the national competent authorities with a tool to assess the severity of the breaches notified by the data controllers.
- To support the national competent authorities in the process of performing detailed analyses and statistics concerning the reported personal data breaches.
- To contribute to the harmonization of the severity assessment of personal data breaches in the European Union, by proposing a common methodology and severity scoring. This would be especially important in the case of cross-border breaches.

1.3 Severity of data breaches

The severity of a personal data breach is defined, in the context of this methodology, as the ***"estimation of the magnitude of potential impact on the individuals derived from the data breach"***.

As stipulated in Directive 2009/136/EC[5] (recital (61)), the impact of a personal data breach may include "for example, identity theft or fraud, physical harm, significant humiliation or damage to reputation in connection with the provision of publicly available communications services in the Community".

With the use of this methodology the data controller is guided through the process by specific quantitative criteria in order to make the overall assessment. The same criteria can be used by the competent authorities, together with the information provided by the controller in the notification form[6], in order to make its own assessment of the breach.

It should be noted that the controller is applying the methodology using the information that is in his/her possession at the time of the breach. In that sense, the methodology cannot always cover all the possible casuistic, including likely impacts on specific groups of individuals or very special cases that cannot be fully addressed under a general methodology. Therefore, it has to be reminded that both, data controllers and competent authorities should put particular care when dealing with cases that, because of their specificities, could not be rightly assessed using this methodology.

[5] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:337:0011:0036:en:PDF

[6] An example of a template of a data breach notification form to the competent authorities is available in the Appendix A of http://www.enisa.europa.eu/activities/identity-and-trust/risks-and-data-breaches/dbn/art4_tech

2 Overview of the methodology

2.1 Criteria

The main criteria taken into account while assessing the severity of a personal data breach are:

- **Data Processing Context (DPC):** Addresses the type of the breached data, together with a number of factors linked to the overall context of processing.
- **Ease of Identification (EI):** Determines how easily the identity of the individuals can be deduced from the data involved in the breach.
- **Circumstances of breach (CB):** Addresses the specific circumstances of the breach, which are related to the type of the breach, including mainly the loss of security of the breached data, as well as any involved malicious intent.

2.2 Calculation of the severity

Based on the above criteria, the approach of this methodology is the following:

- DPC is at the core of the methodology and evaluates the criticality of a given data set in a specific processing context.
- EI is a correcting factor of the DPC. The overall criticality of a data processing can be reduced depending on the value of EI. In other words, the lower the ease of identification is, the lower gets the overall score. Therefore, the combination of the EI and DPC (multiplication) gives the initial score of the severity (SE) of the data breach.
- CB quantifies specific circumstances of the breach that may be present or not in a particular situation. So, when present, CB can only add to the severity of a specific breach. For this reason the initial score can be further adjusted by the CB.

Thus, the final score of the severity assessment can be extracted using the following formula:

$$SE = DPC \times EI + CB$$

In this way, in order for the controller to get the severity result, all three criteria should be scored.

The result belongs to a certain range of values which corresponds to one of the four severity levels: low, medium, high and very high[7]. At the end of the assessment, other possibly relevant criteria (number of individuals and unintelligibility of data) that have not been considered in the methodology are evaluated and flagged to the competent authority when applicable.

It is essential to bear in mind that all scores and/or rankings used in this methodology were solely set for the use within the severity formula. They are not meant to bear any significance to a conclusion about the weighting or ranking of certain types of data in general, let alone an indication to any legal consequences or precedents as to the use of this data for other purposes.

[7] Severity levels will be explained in details in section 3.2.

3 Detailed description of scoring and severity levels

3.1 Scoring of the criteria

3.1.1 Data Processing Context (DPC)

In order to define the score for DPC, the data controller should follow the next steps:

- Step 1: Definition and classification of the types of personal data

  a) Define the types of the personal data involved in the breach.

  b) Classify the data in at least one of the four categories: simple, behavioural, financial, and sensitive data (these categories are explained in details in Annex 1). In this way a preliminary basic DPC score is obtained.

  The list of data types described under the four categories is not exhaustive; however most data involved in real cases can be matched to at least one of the categories. Credentials are not considered as a specific data category and should be handled based on the type of data processed by the systems where they provide access to.

- Step 2: Adjustment by contextual factors related to the data processing

  c) Assess the occurrence of certain factors that could increase or decrease the basic score (data volume, special characteristics of the controllers or the individuals, invalidity/inaccuracy of data, public availability (before the breach), nature of data).

  d) In case such factors exist, accordingly increase/decrease the basic score. Assessment Table 1 provides the adjustment scales per category of data, together with example cases that could lead to lower/higher scores.

Please refer to the Annex 1 for a list of contextual factors and specific examples of DPC scoring.

Note: Even though, for the purpose of the methodology, four data categories are ranked, the categorization itself is not to be seen as a general ranking of the types of data at hand. Much more, additional contextual factors related to the data always need to be taken into account when regarding the processing of a certain type of data. Therefore, the basic score is to be seen just as a general indication of the criticality connected to a certain category of data and the DPC scoring of any data type can always vary from 1 to 4.

If the data matches more than one category the above mentioned steps have to be followed for each category applicable. In these cases the value to use for the overall calculation of the severity will be the highest score reached.

If the controller chooses to alter the DPC basic score (within the range of Assessment Table 1), the new score has to be supported by an explanation describing the particular contextual factors of the breach and their influence.

3.1.2 Ease of identification (EI)

Ease of identification (EI) evaluates how easy it will be for a party who has access to the set of data to univocally match them to a certain person.

For the purpose of this methodology we have defined four levels of EI (negligible, limited, significant and maximum) with a linear increase in score. The lowest score is given when the possibility to identify the individual is negligible, meaning that it is extremely difficult to match the data to a particular person, but still it could be possible under certain conditions. The highest score is selected when identification is possible directly from the data breached with no special research needed to discover the individual's identity. Annex 2 describes these levels in details.

When defining EI, it should be taken into account that identification may be directly (e.g. on the basis of a given name) or indirectly (eg. on the basis of ID number) possible from the breached data, but may also depend on the specific context of the breach. Therefore, certain identifiers may lead to different EI scores according to the specific case of the breach.

Please refer to the Annex 2 for specific examples of EI scoring using common identifiers.

In addition, when defining EI the controller should take into account all the means likely reasonably to be used by any person to identify the individuals. This includes information that is public, held or obtained otherwise, including over the Internet, as well as possible cross-matching with other sources than can be accessed by the data controller or a third party.

3.1.3 Circumstances of the breach (CB)

The elements that are considered under CB are the loss of security (confidentiality, integrity, availability) and malicious intent and are complementary to DPC and EI, as follows:

Loss of confidentiality: Loss of confidentiality occurs when the information is accessed by parties who are not authorized or don't have a legitimate purpose to access it. The extent of loss of confidentiality varies by the scope of disclosure, i.e. the potential number and type of parties that may have unlawfully access to the information.

Loss of integrity: Loss of integrity occurs when the original information is altered and substitution of data can be prejudicial for the individual. The most severe situation occurs when there are serious possibilities that the altered data have been used in a way that could harm the individual.

Loss of availability: Loss of availability occurs when the original data cannot be accessed when there is a need for it. It can be either temporal (data are recoverable but it will take a period of time and this can be detrimental for the individual), or permanent (data cannot be recovered).

Malicious intent: This element examines whether the breach was due to an error or mistake, either human or technical, or it was caused by an intentional action of malicious intent. Non malicious breaches include cases of accidental loss, inadequate disposal, human error and software bug or misconfiguration. Malicious breaches include cases of theft and hacking aiming to harm the individuals (e.g. by exposing their personal data to unauthorised third parties). In other cases malicious intent might include transfer of personal data to third parties for profit (e.g. selling of lists of personal data). In some cases malicious intent could also be inferred from actions aiming to harm the data controller (e.g. through stealing and exposing the personal data to unauthorized parties). Malicious intent is a factor that increases the likelihood that the data is used in harmful way, since this was the initial purpose of the breach.

With regard to CB scoring, contrary to DPC and EI where the maximum score reached is chosen, the points obtained for each CB element are added to obtain the final score, as different circumstances can occur in the same breach. Assessment Table 3 provides different scores per CB element and for different types of circumstances.

Please refer to the Annex 3 for specific examples of CB scoring.

3.2 Definition of severity level

As introduced in the Section 2.2, the overall severity (SE) is calculated by the following formula:

$$SE = DPC \times EI + CB$$

The final score shows the level of severity of a certain breach, taking into account the impact to the individuals[8].

Severity of a data breach		
SE < 2	Low	Individuals either will not be affected or may encounter a few inconveniences, which they will overcome without any problem (time spent re-entering information, annoyances, irritations, etc.).
2 ≤ SE < 3	Medium	Individuals may encounter significant inconveniences, which they will be able to overcome despite a few difficulties (extra costs, denial of access to business services, fear, lack of understanding, stress, minor physical ailments, etc.).
3 ≤ SE< 4	High	Individuals may encounter significant consequences, which they should be able to overcome albeit with serious difficulties (misappropriation of funds, blacklisting by banks, property damage, loss of employment, subpoena, worsening of health, etc.).
4 ≤ SE	Very High	Individuals may encounter significant, or even irreversible, consequences, which they may not overcome (financial distress such as substantial debt or inability to work, long-term psychological or physical ailments, death, etc.).

3.3 Flags

Once the severity level has been defined, it can be accompanied by flags indicating certain elements of the breach that, although they do not affect a priori the scoring, they are important for the final assessment. For the purpose of the methodology, two flags have been considered:

Number of individuals breached exceeds 100. Data of an individual, breached in the context of a bigger incident, can potentially be more easily disclosed, whereas at the same time a high number of affected individuals influences the overall scale of the breach.

Data unintelligible. Unintelligibility (e.g. in the form of strong encryption and without key compromise) can substantially decrease the impact to individuals, since it highly decreases the possibility of unauthorized parties accessing the data.

[8] Table setting the levels of severity of a data breach was first introduced in the "Recommendations on technical implementation guidelines of Article 4", page 24, but has been made more precise in this document.

4 Use of the methodology

4.1.1 Notification to competent authorities

The severity level of the breach (together with the flags), as calculated through this methodology, could be integrated in the notification sent by the controller to the competent authorities. This can be done either automatically in the notification template or through the use of a standalone tool. The competent authorities will be free to evaluate the result (using the same template/tool and the information provided by the controller) and accept it or reject it, following their own assessment.

If for some reason the final severity level is deemed to be incorrect by the data controller, it could be possible for the controller to state the "correct" level, including his/her arguments for the different result, in a free text box (ex. integrated in the notification form). Moreover, any change in the default score for the DPC criterion should also be explained by the controller using a free text box. These boxes could also be used to leave comments without changing the severity level.

4.1.2 Notification to the individuals

The severity level could be used by the controller and the competent authority to determine if there is a need to notify the individuals. The level upon which notification should be deemed necessary could be mutually accepted by all competent authorities or vary depending national-wide criteria.

5 Closing remarks

On the previous pages we have presented a working document on our proposed methodology for assessing data breaches. This proposal ultimately aims towards being integrated into a notification template (such as in ENISA's Recommendations for the technical implementation of the Art.4 of the ePrivacy Directive[9]) to achieve an as far as possible automated severity assessment.

The methodology presented in this study is based on an as objective approach as possible whilst still being flexible enough to be adopted by various Data Protection Authorities by adjusting it to their size, national legal system and other factors. According to different requirements, the scoring of some categories can be adjusted to produce the most appropriate results:

- Data processing context can be adjusted according to the importance of data (simple, behavioural, financial etc.) assigned by the specific DPA
- Ease of identification can take into account the reality in a given country and its legal system (public availability of personal numbers, names, addresses etc.)
- Circumstances of a breach offer the highest flexibility to adjust the final result to the needs of a DPA
- Flags can either be adjusted (for example in function of records breached), complemented by new ones or moved to the CB (circumstances of a breach) table.

ENISA and the DPAs of Greece and Germany aim at further developing the work presented in this document with the final scope of publishing a practical data breach severity tool that can be useful both for the DPAs and the data controllers across the EU.

[9] http://www.enisa.europa.eu/activities/identity-and-trust/risks-and-data-breaches/dbn/art4_tech

Annex 1 – Data processing context

A1 Assessment tables

Table 1: Data Processing Context (DPC)		Score
Simple data	Eg. biographical data, contact details, full name, data on education, family life, professional experience, etc.	
	Preliminary basic score: when the breach involves ''simple data'' and the controller is not aware of any aggravating factors.	**1**
	The DPC score could be increased by 1, e.g. when the volume of "simple data" and/or the characteristics of the controller are such that certain profiling of the individual can be enabled or assumptions about the individual's social/financial status can be made.	2
	The DPC score could be by 2, e.g. when the "simple data" and/or the characteristics of the controller can lead to assumptions about the individual's health status, sexual preferences, political or religious beliefs.	3
	The DPC score could be increased by 3, e.g. when due to certain characteristics of the individual (e.g. vulnerable groups, minors), the information can be critical for their personal safety or physical/psychological conditions.	4
Behavioural data	Eg. location, traffic data, data on personal preferences and habits, etc.	
	Preliminary basic score: when the breach involves ''behavioural data'' and the controller is not aware of any aggravating or lessening factors.	**2**
	The DPC score could be decreased by 1, e.g. when the nature of the data set does not provide any substantial insight to the individual's behavioural information or the data can be collected easily (independently from the breach) through publicly available sources (e.g. combination of information from web searches).	1
	The DPC score can be increased by 1, e.g. when the volume of "behavioural data" and/or the characteristics of the controller are such that a profile of the individual can be created, exposing detailed information about his/her everyday life and habits.	3
	The DPC score can be increased by 2, e.g. if a profile based on individual's sensitive data can be created.	4

Financial data	Any type of financial data (e.g. income, financial transactions, bank statements, investments, credit cards, invoices, etc.). Includes social welfare data related to financial information.	
	Preliminary basic score: when the breach involves ''financial data'' and the controller is not aware of any aggravating or lessening factors.	**3**
	The DPC score could be decreased by 2, e.g. when the nature of the data set does not provide any substantial insight to the individual's financial information (e.g. the fact that a person is the customer of a certain bank without further details).	1
	The DPC score could be decreased by 1, e.g. when the specific data set includes some financial information but still does not provide any significant insight to the individual's financial status/situation (e.g. simple bank account numbers without further details).	2
	The DPC score could be increased by 1, e.g. when due to the nature and/or volume of the specific data set, full financial (e.g. credit card) information is disclosed that could enable fraud or an detailed social/financial profile is created.	4
Sensitive data	Any type of sensitive data (e.g. health, political affiliation, sexual life)	
	Preliminary basic score: when the breach involves 'sensitive data'' and the controller is not aware of any lessening factors.	**4**
	The DPC score could be decreased by 1, e.g. when the nature of the data set does not provide any substantial insight to the individual's behavioural information or the data can be collected easily (independently from the breach) through publicly available sources (e.g. combination of information from web searches).	1
	The DPC score could be decreased by 2, e.g. when nature of data can lead to general assumptions.	2
	The DPC score could be decreased by 1, e.g. when nature of data can lead to assumptions about sensitive information.	3

A2 Description of contextual factors to be considered in DPC scoring

Increasing factors:

- ✓ **The volume of the breached data (for the same individual)**: this factor can increase the basic DPC score, due to the increment of the quantity of the breached information (i.e. acting as aggravating factor). The volume should be considered both in terms of time (e.g. same type of data over a certain period of time) and content (complementing data of the same type). For example, in case of a breach of traffic data at an ISP, the DPC score would be higher (for the same individual) if the data cover a period of one year than if they are limited to one week (time). As another example, in case of a breach at a bank, the DPC score of the complete file of an individual would be higher than that of a single document from the same file (content).
- ✓ **Special characteristics of the data controller**: this factor relates to the field of operation and the activities of the data controller, which could increase the basic DPC score of the data, revealing additional information for a certain data set. For example, the DPC score of a customers' list would be higher if it comes from an online pharmacy than from a stationery shop.
- ✓ **Special characteristics of the individuals**: the basic DPC score of a certain data set could also be increased in case that the individuals belong to a social group with particular needs (e.g. minors, individuals of a particular group with special characteristics). For example, the DPC score of a list of telephone numbers would increase if it concerns known members of the national parliament.

Decreasing factors:

- ✓ **Invalidity/inaccuracy of the data**: the basic DPC score of a certain data set can be decreased if the invalidity or inaccuracy of the data is known to the controller (e.g. due to their age or content) and, thus, their significance is reduced. The controller needs to be certain of this circumstance to include it in the assessment. For example a postal service's list of addresses where letters could not be delivered would be considered as inaccurate (i.e. most probably the individuals have moved to another address).
- ✓ **Public availability**: the basic DPC score of a data set can also be decreased in case the breached data were already publicly available before the breach or can be easily collected and/or accessed through publicly available sources.
- ✓ **Nature of data**: another decreasing factor could in some cases be the very nature of a particular data set that, despite its initial DPC scoring, is of lower significance, in terms of the information that it can reveal about the individual. This is, for example, the case of a medical certificate that is just certifying that the individual is in a good state of health without disclosing any other information. In this case, although the basic score would be 4 due to health data being sensitive data, the specific data set's final DPC score would be 1, as it cannot per se affect the individual's personal life. This factor, however, should be considered with great care and clear explanation of the reason why a particular data processing is by nature lower than its basic DPC score.

A3 Examples of DPC scoring/adjustment per category of data

Simple data

Example 1: List of names and telephone numbers

- **Case 1: the list is the customer list of a supermarket/restaurant.**

  Score = 1 (no alteration due to contextual factors)

- Case 2: the list comes from a company selling luxury cars/homes.

  Score = 2 (by characteristics of controller leading to assumptions about financial/social status)

- Case 3: the list comes from a specialized electronic pharmacy selling products for patients with diabetes.

  Score = 3 (by characteristics of controller leading to assumptions about the individual's health status)

- Case 4: the list includes the names of people working undercover for secret police.

  Score = 4 (by characteristics of individuals that could be critical for their personal safety)

Example 2: A professional CV database

- **Case 1: the data come from an online career site where access to CVs is available for registered users.**

  Score = 1 (no alteration due to contextual factors)

- Case 2: the data come from an organization helping unemployed people to find employment.

  Score = 2 (by characteristics of controller leading to assumptions about financial/social status)

- Case 3: the data come from an institution supporting gay people rights.

  Score = 3 (by characteristics of controller leading to assumptions about the individual's sexual life).

- Case 4: the data come from an organization helping recovering drug addicts to find employment.

  Score = 4 (by characteristics of individuals that could cause them serious damage).

Example 3: List of names and postal addresses

- **Case 1: the list is the customer list of a flower shop.**

  Score = 1 (no alteration due to contextual factors)

- Case 2: the list comes from an investment bank.

  Score = 2 (by characteristics of controller leading to assumptions about financial/social status)

- Case 3: the list is the delivery addresses of an adult books store.

  Score = 3 (by characteristics of controller leading to assumptions about the individual's sexual preferences)

- Case 4: the list concerns persons that have been accused for child abuse.

  Score = 4 (by characteristics of individuals that could cause them serious damage)

Behavioral data

Example 1: Telephone call history (traffic data – no content)

- Case 1: the data come from the helpdesk of an ISP and include incoming calls from subscribers to the helpdesk regarding technical problems.

  Score = 1 (by nature of the data set).

- **Case 2: the data come from an ISP and include subscribers' call history of the last week.**

  Score = 2 (no alteration due to contextual factors)

- Case 3: the data come from an ISP and include subscribers' call history of the last year.

  Score = 3 (a detailed profile of the individual can be created).

- Case 4: the data come from a psychological support centre for people suffering from a serious illness and includes incoming calls.

  Score = 4 (by characteristics of controller disclosing health status).

Example 2: Data in a fidelity card

- Case 1: the card is from a supermarket and includes only the number of points gained from purchases.

  Score = 1 (by nature of the data set).

- **Case 2: the card is from a supermarket and includes information on purchase history for the last month.**

  Score = 2 (no alteration due to contextual factors).

- Case 3: the data come from a transportation card and includes information about location/movements over the last year.

  Score = 3 (a detailed profile of the individual can be created).

- Case 4: the card is from a pharmacy and includes information on recent purchases of medical products.

  Score = 4 (by nature of data set disclosing sensitive data).

Example 3: Data from a social network

- Case 1: the data are publically available on the internet (e.g. pictures that the user has published online).

  Score = 1 (by public availability).

- **Case 2: the data include information about the user's preferences and everyday life of the last month that the user has shared with his/her friends (e.g. information published on user's wall).**

  Score = 2 (no alteration due to contextual factors).

- Case 3: the data include information about the user's preferences and everyday life of the last year that the user has shared with his/her friends (e.g. information published on user's wall).

Score = 3 (a detailed profile of the individual can be created).

- Case 4: the data include personal communication (e.g. personal messages) that may expose information about users' sexual life or health status.

  Score= 4 (leads to detailed profiling related to sensitive data).

Financial data

Example 1: Bank statements

- Case 1: the data come from a bank and include only a letter, through which the individual is identified as a client without providing any information about the specific relations between the client and the bank (e.g. only his/her name and address but no account number or information about transactions).

  Score = 1 (by nature of the data set).

- Case 2: the data come from a bank and include only a transactions history of one day without further details (e.g. account number, name and transaction).

  Score = 2 (by nature of data, info that can lead to limited information about financial behaviour).

- **Case 3: the data come from a bank and concern account balances of clients of the last month.**

  Score = 3 (no alteration due to contextual factors).

- Case 4: the data come from a bank and include account balances of clients of the last year, showing all transactions and related with them details.

  Score = 4 (by volume and nature of data leading to profiling)

Example 2: Income declaration

- Case 1: the data contains a statement to confirm that the individual has submitted his/her income declaration.

  Score = 1 (by nature of the data set).

- Case 2: the data contains the percentage of the taxation which the individual must pay.

  Score = 2 (by nature of data, info that can lead to limited information about financial status).

- **Case 3: the data contains all the fields of one year income declaration of the individual.**

  Score = 3 (no alteration due to contextual factors).

- Case 4: the data contains all the fields of the income declaration of the individual for the last 10 years.

  Score = 4 (by volume and nature of data leading to detailed profiling).

Example 3: Credit card information

- Case 1: the data come from a bank and contain credit card details of individuals but these data are more than ten years old and therefore these cards are not valid.

  Score = 1 (by age of the data set).

- Case 2: the data come from an online shop and contain some credit card information of individuals, but not the necessary set of details to perform financial transactions.

  Score = 2 (by nature of data, info that can lead to limited information about financial status).

- **Case 3: the data come from a bank and contain some credit card information of individuals, but not the necessary set of details to perform financial transactions. However, they contain information about certain purchases of the individuals for the period of one year.**

  Score = 3 (no alteration due to contextual factors).

- Case 4: the data come from an online shop and contain full credit card details that can be used for financial transactions.

  Score = 4 (the data set could be used for fraud).

Sensitive data

Example 1: Data on blood analysis

- Case 1: the data come from a laboratory and include only an indication that the data individuals have performed general blood tests.

  Score = 1 (by nature of data).

- Case 2: the data come from a laboratory of an emergency room of a hospital and only include information about the fact that the individuals performed blood tests (without further details).

  Score = 2 (by nature of data leading to general assumptions).

- Case 3: the data come from a laboratory and includes an indication that the individuals have performed test for a certain disease, without indication of results. Score = 3 (by nature of data leading to assumptions that could cause damage to the individual).

- **Case 4: the data come from a laboratory and includes results of the tests.**

  Score = 4 (no alteration due to contextual factors).

Example 2: Data on political affiliation

- Case 1: the data come from a major political party and include names of prominent members who hold public positions and their affiliation with the party is publically known.

  Score = 1 (by nature of data).

- Case 2: the data come from a company organizing events and include the names of individuals that attended a charity event sponsored by a specific political party.

  Score = 2 (by nature of data leading to general assumptions).

- Case 3: the data come from a political party and include the names of individuals that attended a specific conference organised by the party.

  Score = 3 (by nature of data leading to assumptions that could cause damage to the individual).

- **Case 4: the data come from a closed internet forum and include the political opinions expressed by the members of the forum.**

  Score = 4 (no alteration due to contextual factors).

Example 3: Data on sexual life

- Case 1: the data come from an online discussion forum about relationships and include only the name of registered users without any further information.

  Score = 1 (by nature of data).

- Case 2: the data come from a dating site and include only the name of customers without any further information.

  Score = 2 (by nature of data leading to general assumptions).

- Case 3: the data come from a dating site specialized, but not exclusive, in heterosexual or gay dating and include the name of customers.
- Score = 3 (by nature of data leading to assumptions about sensitive information).
- **Case 4: the data come from a dating site and include the declared sexual orientation of customers.**

  Score = 4 (no alteration due to contextual factors).

Credentials

Example: Username and password of registered users in an online service

- Case 1: the credentials are used for accessing users' accounts in an electronic music store.

  Score = 1 / Simple data (no alteration due to contextual factors)

- Case 2.1: the credentials are used for access to users' accounts in a supermarket's web site, including information about previous shopping lists.

  Score = 2 / Behavioural data (no alteration due to contextual factors)

- Case 2.2: the credentials are used for access to user's accounts in a social media site.

  Score = 3 / Behavioural data with detailed profiling

- Case 3.1: the credentials can be used for access to user's account in the national tax system, proving information about users' income.

  Score = 3 / Financial data (no alteration due to contextual factors)

- Case 3.2: the credentials can be used for online banking with the possibility to perform financial transactions (e.g. transfer of money).

  Score = 4 / Financial data with full financial information and possibility for fraud.

- Case 4: the credentials are used for access to users' accounts in an online community related to sexual preferences.

  Score = 4 / Sensitive data (no alteration due to contextual factors)

Annex 2 – Ease of identification (EI) scoring

This annex will present the examples of EI scoring for common identifiers.

Identification can be direct or indirect and is performed with the use of certain identifiers, taking also into account the overall context of the processing of personal data. The next examples show a (non-exhaustive) list of common identifiers and different cases of their possible use for EI scoring.

It should be noted that in many cases the breach will include a combination of different identifiers, which automatically increases the ease of identification. This is a very important element that should be taken into account by the controller and is reflected in the examples below.

Full name (first name, surname)

It is considered as the most common direct identifier but EI score may vary depending on the case, since the full name does not always in itself uniquely single out the individual. For example, when identification is performed using only the individual's full name:

- EI = 0,25 (Negligible) throughout a country's population where many people share that same full name
- EI = 0,5 (Limited) throughout a country's population where few people share that same full name.
- EI = 0,75 (Significant) throughout a small city's population where few or no people share that same full name.
- EI=1 (Maximum) throughout a country's population using also date of birth and email address.

ID card/passport/social security number

They are all considered as unique identifiers and they can be used to single out the individual, as long as it is possible to link them to a reference database (e.g. linking an ID card to a particular person). For example, when identification is performed using only one of these numbers:

- EI=0,25 (Negligible) when no other information is provided about the individual or it is not possible to find additional information unless access to the reference database is obtained
- EI=0,75 (Significant) when the identifier reveals additional identification information about the individual (e.g. social security number revealing date of birth) and is linked to other data (e.g. postal address or email).
- EI=1 (Maximum) when information from the reference database is also available (e.g. ID card and full name and/or picture).

Telephone number/Home address

They are both indirect identifiers, which can also be used to communicate with or access the individual. When identification is based only on one of these two identifiers:

- EI=0,25 (Negligible) throughout a country's population when the number/address is not registered in a publicly available register.
- EI=0,5 (Limited) throughout a small city's population and the number/address is not registered in a publicly available register (identification possible through communication).
- EI=1 (Maximum) throughout a country's population and the number/address is included in publicly available register.

Email address

It is also an indirect identifier, which can be used to communicate with the individual and in some cases it may include information about his/her name (first name and/or surname). When identification is based on email:

- EI=0,25 (Negligible) when the email address does not reveal any other identification information (e.g. name) and is not used as a primary address of the individual in internet sites, forums or social networks.
- EI=0,75 (Significant) when the email address does not reveal any other identification information (e.g. name) but is used as a primary address of the individual in internet sites, forums or social networks (searchable on the web).
- EI=1 (Maximum) when the email address reveals the individual's name and is used as his/her primary address in internet sites, forums or social networks (searchable on the web).

Picture

It may be a direct or indirect identifier, depending on the case. For example, when identification is based only on a picture:

- EI=0.25 (Negligible) when the picture is unclear or vague (e.g. CCTV footage from a long distance).
- EI=05 (Limited) when the picture is unclear or vague but it includes additional information (e.g. surroundings that show a specific location) that could lead to the identification of the individual.
- EI=0,75 (Significant) when the picture is clear but no other identification information is linked to it.
- EI=1 (Maximum) when the picture is clear and linked to some additional information (e.g. information about membership to a specific group, home address, etc).

Coding/Aliases/Initials

Coding refers to the assignment of a unique ID number to each individual, e.g. in the context of a specific database. The use of aliases is a form of pseudonymisation, in the sense that a specific identifier (usually the individual's full name) is substituted by an alias (pseudonym). The initials are a type of alias that is extracted from the full name of the individual. Like in the case of unique identifiers, codes and aliases can be used to identify the individual as long as it is possible to link them to a reference database (e.g. linking the code/alias to the full name of a particular person) When identification is based on coding or use of aliases:

- EI=0,25 (Negligible) when the code/alias does not reveal and cannot be linked to any other personal data about the individual unless access to the reference database is obtained.
- EI=0,75 (Significant) when the alias reveals some data about the individual (e.g. first name) and is linked to other personal data (e.g. the individual's email address).
- EI=1 (Maximum) when the alias reveals the individual's full name or data from the reference database are also available.

Annex 3 – Examples of the circumstances of the breach (CB) scoring

A1 Loss of confidentiality

0 – Examples of data exposed to confidentiality risks without evidence that illegal processing has occurred.

- ✓ A paper file or laptop is lost during transit.
- ✓ Equipment has been disposed without destruction of the personal data

+0.25 – Examples of data disposed to a number of known recipients:

- ✓ An email with personal data has been wrongly sent to a number of known recipients.
- ✓ Some customers could access other customers' accounts in an online service.

+0.5 – Examples of data disposed to an unknown number of recipients:

- ✓ Data are published on an internet message board.
- ✓ Data are uploaded to a P2P site.
- ✓ An employee sells a CD ROM with customer data.
- ✓ A wrongly configured website makes publically accessible on internet data from internal users.

A2 Loss of integrity

0 - Examples of data altered but without any identified incorrect or illegal use:

- ✓ The records of a database with personal data have been wrongly updated but the original has been obtained before any use of the altered data occurred.

+0.25 – Examples of data altered and possibly used in an incorrect or illegal way but with possibility to recover:

- ✓ A record that is necessary for the provision of an online social service has been changed and the individual needs to ask for the service in an offline way.
- ✓ A record that is important for the accuracy of an individual's file in an online medical service has been changed.

+0.5 – Examples of data altered and possibly used in an incorrect or illegal way without possibility to recover:

- ✓ The previous examples + the original cannot be recovered.

A3 Loss of availability

0 – Examples of data being recoverable without any difficulty:

- ✓ A copy of file is lost but other copies are available.
- ✓ A database is corrupted but can be easily reconstructed from other databases.

+0.25 – Examples of temporal unavailability:

- ✓ A database is corrupted but can be reconstructed from other databases, although some processing is required.
- ✓ A file is lost but the information can be provided again by the individual.

+0.5 – Examples of full unavailability (data cannot be recovered from the controller or the individuals):

- ✓ A file is lost/database corrupted, there is not back up of this information, and it cannot be provided by the individual.

A4 Malicious intent

+0.5 – The breach was due to an intentional action, e.g. in order to cause problem to the data controller (e.g. demonstrate loss of security) and/or in order to harm the individuals.

- ✓ An employee of a company intentionally shares private data from customers in a social media public site.
- ✓ An employee of a company sells private data from customers to another company.
- ✓ A member of a social network intentionally sends information about other members to their family members in order to harm them.

TP-04-13-118-EN-N

ENISA
European Union Agency for Network and Information Security
Science and Technology Park of Crete (ITE)
Vassilika Vouton, 700 13, Heraklion, Greece

Athens Office
1 Vass. Sofias & Meg. Alexandrou
Marousi 151 24, Athens, Greece

ISBN 978-92-9204-078-9
9 789292 040789
doi: 10.2824/27590

PO Box 1309, 710 01 Heraklion, Greece
Tel: +30 28 14 40 9710
info@enisa.europa.eu
www.enisa.europa.eu