

GDPR implementation

Case: Headpower Oy

Pasi Reini

Master's thesis

March 2019

School of Technology, Communication and Transport

Master's Degree Programme in Information Technology

Cyber Security

Author(s) Reini, Pasi	Type of publication Master's thesis	Date March 2019 <hr/> Language of publication: English
	Number of pages 58	Permission for web publication: X
Title of publication GDPR Implementation Case: Headpower Oy		
Degree programme Master's degree programme in Cyber Security		
Supervisor(s) Saharinen, Karo; Kokkonen, Tero		
Assigned by Headpower Oy, Keränen, Kari		
Abstract <p>New General Data Protection Regulation (GDPR) of the European Union came into force on 25th of May 2018. Every organization processing personal data of EU citizens must comply with the aforementioned regulation. The assigner organization for the research is Headpower Oy which provides cloud services for energy sector companies. The background for the research was the Headpower's interest in complying with GDPR regulation.</p> <p>The first research problem was a gap with a knowledge whether there is a common model that explains the needed steps for how to make a software meet the GDPR requirements. The second research problem was to study the steps that need to be taken to make a software GDPR compliant. The third research problem was about ensuring that a software stays GDPR compliant in the future. The research method was systematic literature review, which means that the searches were performed in the databases containing theses of technical universities, universities and universities of applied sciences located in Finland. The literature was Master's theses and the selection of the literature was performed against the predefined search criteria. The main goal of the literature review was to find out if there is a model how to make the example software to meet the GDPR requirements.</p> <p>The previous surveys found in the literature review show that there is no common model how to implement GDPR requirements into the example software. Instead of the common model, small pieces of information can be found in the previous surveys. Headpower performed GDPR implementation by checking what personal data there is in the system and then implemented GDPR requirements into the example software.</p>		
Keywords/tags (subjects) General Data Protection Regulation, Personal Identifiable Information		
Miscellaneous GDPR (General Data Protection Regulation), PII (Personal Identifiable Information)		

Tekijä(t) Reini, Pasi	Julkaisun laji Opinnäytetyö, ylempi AMK	Päivämäärä Maaliskuu 2019 Julkaisun kieli Englanti
	Sivumäärä 58	Verkkojulkaisulupa myönnetty: X
Työn nimi GDPR Implementation Case: Headpower Oy		
Tutkinto-ohjelma Master's degree programme in Cyber Security		
Työn ohjaaja(t) Saharinen, Karo; Kokkonen, Tero		
Toimeksiantaja(t) Headpower Oy, Keränen, Kari		
Tiivistelmä <p>Euroopan unionin uusi tietosuoja-asetus (GDPR) tuli voimaan 25. toukokuuta 2018. Kaikkien organisaatioiden, jotka käsittelevät Euroopan unionin kansalaisten henkilötietoja, tulee noudattaa tätä uutta asetusta. Opinnäytetyön toimeksiantaja Headpower Oy tarjoaa pilvipalveluita energia-alan yrityksille. Tutkimuksen taustalla oli Headpowerin tarve vastata GDPR-asetuksen myötä tullessiin vaatimuksiin.</p> <p>Työn ensimmäisenä tutkimusongelma oli selvittää, onko olemassa yleistä mallia, joka sisältää tarvittavat toimenpiteet ohjelman muuttamiseksi GDPR-yhteensopivaksi. Toinen tutkimusongelma koski niitä toimenpiteitä, joilla ohjelma muutetaan GDPR-asetuksen mukaiseksi. Kolmas tutkimusongelma keskittyi kuvaamaan, miten varmistetaan ohjelman GDPR-yhteensopivuus tulevaisuudessa. Tutkimusmenetelmänä työssä käytettiin systemaattista kirjallisuuskatsausta, joka toteutettiin suorittamalla haku niihin tietokantoihin, jotka sisältävät ylempien korkeakoulututkintojen opinnäytetöitä. Kyseiset tietokannat olivat suomalaisten teknisten yliopistojen, yliopistojen ja ammattikorkeakoulujen opinnäytetyötietokantoja. Aikaisemmat opinnäytetyöt GDPR:stä haettiin ja valittiin tiettyjen ennalta määriteltyjen kriteerien mukaan. Kirjallisuuskatsauksen päätaavoite oli löytää malli, jonka avulla Headpowerin esimerkkiohjelma saadaan täyttämään GDPR:n vaatimukset.</p> <p>Kirjallisuuskatsauksen avulla löydetty aiemmat tutkimukset aiheesta osoittivat, että yleistä mallia GDPR:n toteuttamiseksi ei ole olemassa. Näin ollen aiemmista tutkimuksista saatu tieto jäi vähäiseksi ja hajanaiseksi. Headpower aloitti GDPR:n käyttöönoton selvittämällä, mitä henkilötietoja järjestelmissä on. Tämän jälkeen se toteutti tietosuoja-asetuksen vaatimukset esimerkkiohjelmaan.</p>		
Avainsanat (asiasanat) General Data Protection Regulation, Personal Identifiable Information		
Muut tiedot GDPR (General Data Protection Regulation), PII (Personal Identifiable Information)		

Contents

1	Introduction	4
2	Research implementation	5
2.1	Research questions	5
2.2	Research method	6
2.3	Conducting literature review.....	6
3	General Data Protection Regulation (GDPR)	10
3.1	Background for GDPR.....	10
3.2	Personal identifiable information definition under GDPR.....	11
3.3	Principles of GDPR	12
3.4	Natural person's rights under GDPR.....	14
3.4.1	The right to be informed	14
3.4.2	The right of access.....	14
3.4.3	The right to rectification.....	15
3.4.4	The right to be forgotten	15
3.4.5	The right to data portability.....	15
3.4.6	The right to be notified of data breaches.....	16
3.5	GDPR challenges	17
3.5.1	Notification of the data breach challenge	17
3.5.2	The right to be forgotten challenge	19
4	Example application	20
4.1	Target organization	20
4.2	History of example application.....	21
4.3	Stakeholders	21
4.4	Architecture	23
4.5	Registration into Headpower	25
5	Steps to GDPR compliance	25
5.1	Existing GDPR implementation model	25
5.1.1	Previous surveys of the GDPR implementation.....	25
5.1.2	Conclusion of the GDPR implementation model	27
5.2	Personal data mapping	28

5.3	Who has access to data and why.....	30
5.3.1	Helpdesk	30
5.3.2	The software development team.....	31
5.3.3	Corporation management, salespersons and product owners	32
5.4	Consent	32
5.5	Lawfulness, fairness and transparency	33
5.6	Data minimization.....	34
5.7	Storage limitations	35
5.8	Software’s GDPR features	35
5.8.1	Change, access and error logs.....	35
5.8.2	Do not archive unnecessary data.....	37
5.8.3	Clean databases automatically	38
5.8.4	Person search.....	43
5.8.5	Automatic file removal service	43
5.8.6	HTTPS over HTTP.....	44
5.8.7	Right to data portability	45
5.8.8	Review of existing features of the application	46
5.9	Notification of data breaches.....	47
6	How to ensure GDPR compliacity in the future.....	47
6.1	Monitor GDPR changes	48
6.2	Train employees regularly.....	48
6.3	Internal audit	49
6.4	Take GDPR into account when designing and implementing features	49
6.5	Audit software features and data.....	50
7	Conclusions	51
	References.....	55

Figures

Figure 1. Privacy principles of GDPR	13
Figure 2. Person's rights under GDPR	17
Figure 3. The example application's stakeholders.....	22
Figure 4. The software development process	23
Figure 5. The architecture of the example application.....	24
Figure 6. An example document of the personal data mapping.....	29
Figure 7: Example form to be filled in before an employee can access personal data	31
Figure 8: Part of Headpower's Data Security Report.....	33
Figure 9: Information note for users of changes that have been made to end user terms	34
Figure 10. SQL Server Agent on example server	39
Figure 11. An example procedure for deleting archived customers	40
Figure 12. An example script for an agent job that executes a procedure.....	41
Figure 13. The example schedule for the SQL Server Agent job	42
Figure 14. History of an SQL Server Agent Job	42
Figure 15. Example application's person search feature	43
Figure 16. Connections protected with HTTPS.....	45
Figure 17. Timeline of the GDPR project.....	54

Tables

Table 1. Search criteria for the literature.....	7
Table 2. Databases / Repositories for the search.....	8
Table 3. Search phrases.....	8
Table 4. Hits of the search	10

1 Introduction

The new regulation of the European Union, General Data Protection Regulation (GDPR), came into force on 25th of May 2018. Every organization operating within the European Union prepares to become GDPR compliant. GDPR will increase the individuals' (natural persons) rights to take control over their own data. The theory base introduces six privacy principles of GDPR. In addition, the rights of a natural person have been listed under the new regulation.

The research was assigned by Headpower Oy which provides cloud services. The company constantly develops its products and services to meet the needs of customers and industry. Currently, the corporation is planning to develop its software products to meet the GDPR requirements. The company has an example application with over 15-year history including many clients and users. In addition, many programmers have worked on it. Most of the application's original developers have changed their firms, which naturally causes that information has been lost with the former employees leaving the company. This is why a development team now has received the challenge to get the software to meet the GDPR requirements. The author of this research is a member of the development team. His role is a software developer whose responsibilities are designing, programming, testing and releasing new versions.

The results of this research concern the following problems: Is there an existing GDPR model for software developers, what steps need to be taken to make a software compliant with the GDPR, and how to ensure that a software stays GDPR compliant in the future. The result for each research question has been reported in the following chapters: Chapter 5, Steps to GDPR Compliance and Chapter 6, How to Ensure that Software Meets the GDPR Requirements in the Future.

2 Research implementation

2.1 Research questions

“The GDPR leaves much to interpretation. It requires that companies must provide a ‘reasonable’ level of protection for personal data, for example, but does not define what constitutes ‘reasonable’.” (General Data Protection Regulation (GDPR) requirements, deadlines and facts 2018.)

As also written in Moonsoft’s customer magazine 1/2018, many sources offer day by day more instructions about GDPR; however, it is hard to filter facts out of the information. Many consulting firms propose their instructions as facts even though there are no existing standardized GDPR practices. The problem is also that instructions and processes will be updated continuously, e.g. by authorities even with a short distance to GDPR’s deadline. The regulation does not contain a detailed requirement about the ways organizations should implement GDPR. Organizations should themselves evaluate and define how they will comply with the GDPR. (DR, GDPR vai DDR: Mitä pitää oikeasti tehdä? 2018.)

The Marketing & advertising page of Alma Talent Oy has a blog text of the GDPR, which also confirms that many consultants offer information and some of them do this without a valid competence. In addition, some consultants frighten potential customer organizations with large fines if they cannot comply with the GDPR requirements. The regulation does not seem to be ready yet and the authorities should bring more detailed information about it in the future. (Lemminki 2018.)

The need for the research is caused by a lack in knowledge on how to make software meet the GDPR requirements. When the GDPR project started in the beginning of year 2018, the assigner organization had no clear information on how to get a software to comply with GDPR requirements. There is a great amount of information, for example, on the internet, which has been confusing to the assigner organization. The main goal of the research is to find a model which helps developers to make a software GDPR compliant or create an own model for it.

The research questions are listed as follows:

1. Is there any existing model for software developers how to make a software GDPR compliant?
2. What steps need to be taken for a software to be GDPR compliant?
3. How to ensure that a software meets the GDPR requirements in the future?

2.2 Research method

A literature review is a useful method when a researcher needs to understand a subject that he is studying and discover what previous researches there have been about it. A researcher must also demonstrate that he understands the theory that has been used in the research. The literature review has four basic stages which are: getting relevant sources, studying them, extracting and taking notes and writing sections about them. The literature review is not “copy and paste” work or a search on the internet. Instead, the researcher selects the suitable content from different texts, theories and concepts relevant to the topic. Then the researcher critically evaluates the previous work of other researchers. The reason for this is to find out literature that might help to create e.g. best practices. (Hart 2018, 3–5.)

Systematic literature review minimizes doubts, because a researcher finds out as many as possible evidences of the research subject. The result of this type of literature review is e.g. a research report that will provide the best available information for decision makers how to do something better, more efficiently or how to do right things in the right way. A systematic literature review might be a useful method when an individual or an organization lacks knowledge or finite resources. (Hart 2018, 99–100.)

2.3 Conducting literature review

When preparing a literature review, it would be useful to create some criteria and think about e.g. parameters that will set a scope for a study. In other words, what kind of literature should be included or excluded. The search is mostly about locating and managing a great amount of information, and most of the findings might be irrelevant or superficial. The information selection criteria help to make the decision

between relevant and irrelevant material. For example, a researcher can use theses from universities with a good reputation or select publications from a specified time range. (Hart 2001, 23–26.)

The search criteria for this work have been defined the following way: The source material must contain researches about General Data Protection Regulation and a model of how to implement its requirements into a software. This work uses literature published in Finnish and English only, because there is no time to learn a new language. The electronic material constitutes the main form of the literature in this work; nevertheless, with access to the original documents. The detailed search criteria have been described in Table 1.

Table 1. Search criteria for the literature

Literature approval criteria	Literature rejection criteria
<ul style="list-style-type: none"> • The literature is published between 01.01.2016 and 31.01.2019. • Language of the literature is Finnish or English. • The form of the literature is the electronic material. • The literature is Master's theses • The original documents are available • Major/subject of the literature is related to the information technology or to the computer science. • The content of the literature does contain references to the model how to implement European Union's General Data Protection Regulation requirements into the software. 	<ul style="list-style-type: none"> • The literature is published before 01.01.2016 or after 31.01.2019. • Language of the literature is not English or Finnish. • The form of the literature is not the electronic material. • The original documents are not available • The literature is not Master's theses or the literature is bachelor's theses. • Major/subject of the literature is not related to the information technology or to the computer science. • The content of the literature does not contain any references to the model how to implement European Union's General Data Protection Regulation requirements into the software.

The material sources the researcher intends to search should be planned and listed (Hart 2001, 23). A quick search is a handy method for getting the overview of a topic and indicating which sources contain relevant information (Hart 2001, 8). A list of useful sources is made by a quick search, because it exposes if a source contains material regarding the topic of this work. The searches will be made into databases

containing theses of technical universities, universities and universities of applied sciences located in Finland. This search is to be conducted using keywords such as GDPR. Then the suitable source databases are selected for further searching. If a database does not contain any keyword hits, then it is to be rejected from sources. Databases and repositories are listed in Table 2.

Table 2. Databases / Repositories for the search

Database / Repository	Description
Aaltodoc	Aalto university's archive for full text materials such as theses, journal articles, conference publications and research materials.
JYX	Jyväskylä university's digital repository e.g. for Master's theses.
LUTPub	Lappeenranta university of technology's publication repository for bachelor's and Master's theses.
TamPub	Tampere university's open archive for Master's theses.
Theseus	Open Repository Theseus is Arene ry's provided service for theses and publications of the Universities of Applied Sciences.
TUT DPUB	Tampere university of technology's archive for Master's theses.

The initial search to the databases was conducted with search phrases listed in Table 3. The search was made to find out if the selected databases contain Master's theses of GDPR. The search phrases are in English and Finnish in Table 3.

Table 3. Search phrases

English search phrases	Finnish search phrases	Advanced filters
"European Union" AND "General Data Protection Regulation"	"Euroopan Union" AND "tietosuoja-asetus"	<ul style="list-style-type: none"> • Abstract contains GDPR • Level of the thesis equals Master's thesis • Publication year is 2016 - 2019
"General Data Protection Regulation"	"tietosuoja-asetus"	<ul style="list-style-type: none"> • Abstract contains GDPR • Level of the thesis equals Master's thesis

		<ul style="list-style-type: none"> • Publication year is 2016–2019
"GDPR"	-	<ul style="list-style-type: none"> • Abstract contains GDPR • Level of the thesis equals Master's thesis • Publication year is 2016 - 2019
"GDPR model"	"GDPR tietosuojamalli"	<ul style="list-style-type: none"> • Abstract contains GDPR • Level of the thesis equals Master's thesis • Publication year is 2016 - 2019
"GDPR requirements"	"Tietosuoja-asetus vaatimukset"	<ul style="list-style-type: none"> • Abstract contains GDPR • Level of the thesis equals Master's thesis • Publication year is 2016 - 2019
"GDPR software changes"	"Tietosuoja-asetus ohjelmistomuutokset"	<ul style="list-style-type: none"> • Abstract contains GDPR • Level of the thesis equals Master's thesis • Publication year is 2016 - 2019

Each search phrase was used systematically with each database / repository. The approved literature is analyzed in Chapter 5.1 Existing GDPR implementation model. The search returned many hits, which do not contain any references to the General Data Protection Regulation. The quantities of the hits are listed in Table 4.

Table 4. Hits of the search

Search phrase	AaltoDoc	JYX	LUTPub	TamPub	Theseus	TUT DPUB
"European Union" AND "General Data Protection Regulation"	2	63	85	0	4	244
"Euroopan Union" AND "tietosuoja- asetus"	0	43	0	0	0	0
"General Data Protection Regulation"	2	28	5	0	6	23
"tietosuoja-asetus"	2	1	4	1	5	299
"GDPR"	2	1	7	1	6	29
"GDPR model"	2	1	0	0	3	1739
"GDPR tietosuojamalli"	2	1	0	0	0	29
"GDPR requirements"	2	5	6	0	4	1478
"Tietosuoja-asetus vaatimukset"	2	3	3	1	5	960
"GDPR software changes"	2	15	6	0	3	1767
"Tietosuoja-asetus ohjelmistomuutokset"	2	1	0	0	0	301

3 General Data Protection Regulation (GDPR)

3.1 Background for GDPR

The main goal of the new legislation is to protect EU citizens from organizations that use personal identifiable information (PII) unlawfully. Sanctions for data breaches have also been increased, and organizations have new requirements e.g. for data breach notifications. The organizations failing to comply with the GDPR will face

penalties of €20m euro or four per cent of their global annual turnover. The new GDPR rules should also help organizations to prepare correct policies and procedures to handle cyber security incidents. In addition, GDPR will change the way organizations process and store personal identifiable information. The rights of EU citizens are to be extended and the GDPR applies to all organizations that process EU residents' PII. (EU General Data Protection Regulation (GDPR) Overview; Duncan 2018.)

The GDPR standardizes personal identifiable information (PII) protection in every European country. Organizations must consider what PII they process and how they should protect it. With GDPR, there are different roles for each organization which are data controller and processor. The controller must define how and why PII is being processed and the processor processes it. The controller might be a company, charity or government and the processor might be an information technology firm. Even organizations outside the European Union operating in European Union territory must apply the requirements of the regulation. After the GDPR legislation's due date, every organization must handle personal data lawfully and transparently. In addition, the processing of the PII must have a real purpose. When personal identifiable information is no longer required, organizations should remove it. (Curtis 2018.)

3.2 Personal identifiable information definition under GDPR

Any data related directly or indirectly to an identifiable natural person is personal identifiable information (PII). Examples of personal identifiable information are name, identification number, location data, online identifier (email or IP address), health, physical, genetic or biometric data, mental, economic, cultural or social identity of a natural person. The processing of special categories of PII is prohibited by default according to GDPR's Article 9. Racial, ethnic origin, political opinions, religious or philosophical beliefs, genetic data, biometric data, health, a natural person's sex life or sexual orientation can be counted as sensitive data. The GDPR lists some circumstances which allow organizations to process a special category of PII. (Regulation (EU) 2016/679 of the European parliament and of the council 2016.)

3.3 Principles of GDPR

The regulation's Article five describes principles to which organizations should pay attention when processing the personal identifiable information (PII). Below is a list of some principles of processing of the PII (Regulation (EU) 2016/679 of the European parliament and of the council 2016.):

- PII must be processed lawfully, fairly and in a transparent manner.
- PII must be collected for specified, explicit and legitimate purposes and not processed in a manner that is not compatible with these purposes.
- PII must be adequate, relevant and limited to the purposes for which it is processed.
- PII must be accurate, kept up to date and organizations should ensure the data is not inaccurate.
- PII must be kept in a form that permits identification of data subjects only for the time it is necessary.
- PII must be processed in a manner that is secured and protected against unauthorized or unlawful processing and not accidentally lost, destroyed or damaged.
- The controller is responsible for demonstrating that the data is compliance with the regulation.

Lawfulness, fairness and transparency can be explained in the following way: an organization must inform a person of the data processing methods. In addition, they also need to inform what kind of data is being processed. The processing methods must match up with a data security report offered by an organization. Figure 1 lists six privacy principles of the GDPR. (The Six Privacy Principles of GDPR 2017.)



Figure 1. Privacy principles of GDPR (The Six Privacy Principles of GDPR 2017)

Purpose limitations mean that the personal identifiable information (PII) can be processed for specified, explicit and legitimate purposes. A data subject is aware of the aforementioned purposes and the PII is not used for further actions without user's consent. Only data is collected that is necessary and nothing more (data minimization). Accuracy means that the PII should be kept up to date and accurate. Storage limitations mean that the data will be stored only for the time necessary and not any longer. When there is no purpose anymore for storing the PII, the data should be deleted. Integrity and confidentiality mean that the PII should be handled in a manner that it is secured against unlawful processing or accidental destruction or damage. (The Six Privacy Principles of GDPR 2017.)

GDPR's sixth Article defines lawfulness of processing, examples follow (Regulation (EU) 2016/679 of the European parliament and of the council 2016):

- Data subject has given a consent to the processing of his or her personal identifiable information for at least one or more purposes.
- Data subject is a part of a contract that requires processing of personal identifiable information.
- Data subject's or another natural person's vital interests should be protected.

- Processing is necessary from the aspect of public interest or official authority.
- When legitimate interests of the controller or third party are necessary to be protected.

3.4 Natural person's rights under GDPR

3.4.1 The right to be informed

Register owner must give the following information to a registered person (EU-tietosuoja kokonaisuudistus 2016, 14):

- Contact information of the register owner
- What is the purpose and legal base for personal identifiable information processing?
- If personal identifiable information will be given to third parties and who are the recipients of the data.
- If personal identifiable information will be transferred to third countries and how security has been taken into account.
- Storage time for each personal identifiable information and what is the legal base for storing.
- If there is automatic decision making or profiling, what is the processing logic and what are consequences for a registered person?
- What kind of data will be collected?
- What is the source of personal identifiable information?

3.4.2 The right of access

A person has the right to access to his/her personal identifiable information (PII), which means that register controller must notify the natural person if any PII will be processed and then deliver a copy of the aforementioned data (EU-tietosuoja kokonaisuudistus 2016, 14–15).

3.4.3 The right to rectification

The GDPR regulation brings for natural persons the right to demand for rectification for the wrong information in the register owner's systems (EU-tietosuojan kokonaisuudistus 2016, 15).

3.4.4 The right to be forgotten

A person has the right to ask register owners to remove his/her expired personal identifiable information (PII). In addition, a person has also the right to cancel the consent for data processing. In addition, a person has the right to demand deletion of his/her PII from register owner's systems. After this the data must be deleted if there is no legal purpose for storing it anymore. The regulation does not give any requirements from a technical aspect for data deletion. At least the data can be deleted, e.g. overwriting it so that natural persons cannot be identified from the data anymore. In addition, the data can be marked as deleted and then limitations are set for its use in information systems; however, this way the data still exists, for example in a database. Nevertheless, the destruction of physical devices that will store PII is overreacting, because it might be difficult to find locations of the data, e.g. from cloud systems. (EU-tietosuojan kokonaisuudistus 2016, 15–16.)

3.4.5 The right to data portability

The right to data portability is also a new demand of GDPR. A person has rights to get all his/her personal identifiable information (PII) in a common structured format and then transfer this data to other register controller's systems. One aspect of this data portability is that a person has rights to transfer the data directly from one register controller to another if technically possible. The right to data portability does not mean that register controllers or processors should design and implement compatible systems. When the systems are different, the PII can be transferred e.g. using external memory storage and then loading it to another register controller's system. (EU-tietosuojan kokonaisuudistus 2016, 16.)

3.4.6 The right to be notified of data breaches

One responsibility for register controllers is to notify registered persons of data breaches the data of which has been leaked. The right is in force if a breach causes great risks for an individual's rights and freedom. The aforementioned risks are, for example, identity thefts, credit card frauds or other criminal activities. The notification is not compulsory if the leaked personal identifiable information was encrypted and encryption keys were not leaked. An organization can use social media for informing of the data breach if otherwise it might cause too big a load of work. (EU-tietosuoja kokonaisuudistus 2016, 17.)

An organization must give the following details of data breaches to data subjects whose data has been leaked (EU-tietosuoja kokonaisuudistus 2016, 17):

- Clear and simple description of the data breach
- Contact information for more details
- A description what impacts a data breach might cause for a person's right and freedom
- A description of activities that the register owner has already done or will do for decreasing the impacts of a data breach.

Figure 2 visualizes natural person's rights under GDPR.

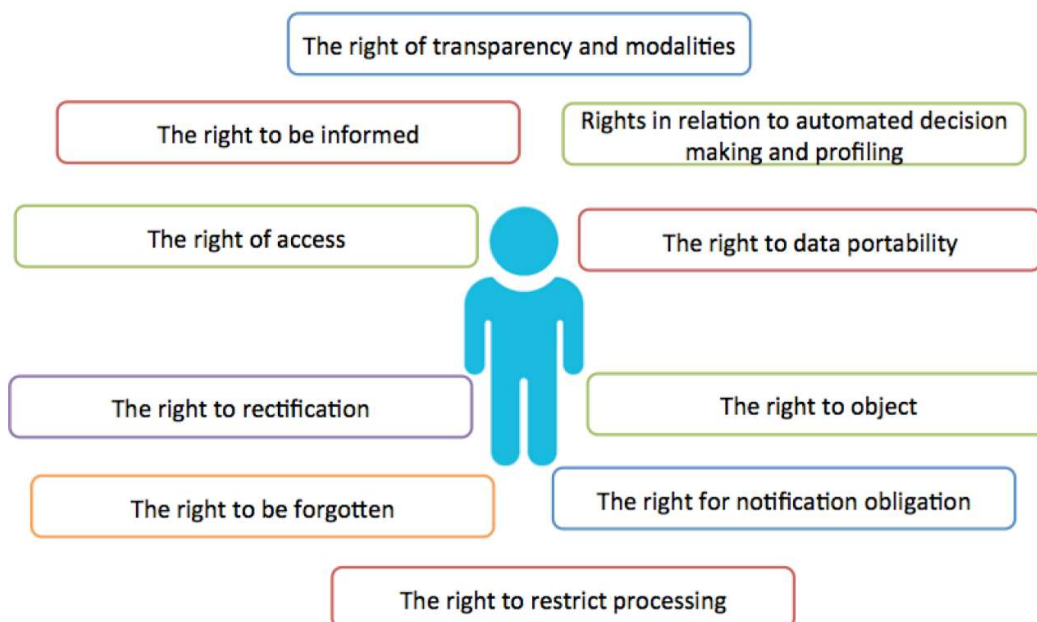


Figure 2. Person's rights under GDPR (Gunathunga 2017)

3.5 GDPR challenges

3.5.1 Notification of the data breach challenge

In November 2017, IAPP Europe Data Protection Congress 2017 took place in Brussels where hundreds of privacy professionals evaluated what risks GDPR might bring. The biggest risk with GDPR compliance was to comply with 72-hour data breach notification regulation (Survey Reveals Biggest GDPR Compliance Risks are Breach Notification, Data Mapping, Managing Consent, and Data Transfer 2017).

GDPR Article 33 states: "In the case of a personal data breach, the controller shall without undue delay and, where feasible, not later than 72 hours after having become aware of it, notify the personal data breach to the supervisory authority competent in accordance with Article 55, unless the personal data breach is unlikely to result in a risk to the rights and freedoms of natural persons. Where the notification to the supervisory authority is not made within 72 hours, it shall be accompanied by reasons for the delay." (Regulation (EU) 2016/679 of the European parliament and of the council 2016.)

In the USA there are over 48 different breach notification laws. The shortest notification timeline is 30 days in Florida, and for example, South Dakota has 60 days' timeline of the breach's discovery. GDPR's demand is 72 hours, which turns tracing a problem a race against the clock from the moment the breach has been discovered.

Organizations must be prepared to perform specified tasks when a breach has been detected. These tasks are, for example: A local supervisory authority and data subjects must be alerted, the organization must provide the nature of the breach and organize communication channels between different parties on the number of personally identified information involved and how the organization will respond to the incident. All this must be done in 72 hours or one should better have a good explanation. (Do 72 Hours Really Matter? Data Breach Notifications in EU GDPR 2018.)

Security breaches can be categorized in following ways (Article 29 data protection working party 2017):

- Confidentiality breach: an unauthorized or accidental access to data.
- Availability breach: an accidental or unauthorized loss of access to data or its destruction.
- Integrity breach: an unauthorized or accidental alteration of personal data.

A data breach can concern all abovementioned categories at the same time or any combination of these. Here are some examples of availability breaches: Data has been deleted accidentally or by an unauthorized person, data cannot be restored from a backup or data cannot be accessed because of denial of service attack or power failure. (ARTICLE 29 DATA PROTECTION WORKING PARTY 2017.)

From the perspective of software development, there is a possibility that all abovementioned data breach types might occur. A confidentiality breach might occur in a software as a result of a user's mistake. In other words, the aforementioned mistake might be social engineered credentials or sensitive information or just giving too many privileges to people; despite the fact that automated tests, regression testing and manual testing reduce software failures and bugs. There is always a chance that a bug in the software code might alter the wrong data or corrupt it

accidentally. It has not been defined clearly if e.g. the aforementioned software bug can be called as integrity breach and if an organization is obliged to report it to authorities and to persons whose data has been damaged.

3.5.2 The right to be forgotten challenge

GDPR Article 17 states: “The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay –” (Regulation (EU) 2016/679 of the European parliament and of the council 2016.)

Article 17 describes a data subject’s rights to demand data erasure from data controller and the controller must obey this request. Companies should have a mechanism and processes to make sure that the removed data will not come back in the future. (Loshin 2017.)

Integrations between systems should be implemented in a manner that the once removed data will not be restored into the system as Malste (2017, 40) wrote in his thesis about CRM system’s integration problems with another system. This is also a recognized problem with many different systems, not only in the CRM.

A very common problem appears with database backups, because if a backup is ever to be restored and once the removed data is still there, it will surely come back into the system. The aforementioned scenario might bring troubles for organizations, which is why the once removed data should not be forever stored in the backups. (Loshin 2017.)

Database backups taken from a production environment the only locations where personal data might be. There might be different test environments for many purposes with their own databases and backups. An organization should have processes and guidance, e.g. for developers how to keep test environments clean of data that containing personal identifiable information. Very often, the easiest way to hunt software bugs is that database backups have been restored in a test environment or a software developer’s personal working computer from a production environment. Hunting bugs can be challenging and requires customer’s

data that repro steps can be found. Creating complicated scenarios from a zero point is a time consuming method to fix things.

It can be read between the lines of GDPR that a data processor must provide software build-in personal data erasure mechanism for a data controller. In the future, it might be a problem that a customer or a controller might want to know why the data stored years ago does not exist anymore. This depends on each organization's policies; however, the oldest database backup might be e.g. 12 months old and this older data cannot be restored anymore because it does not exist in backups.

A processor can tackle the abovementioned problem by logging user actions in the system. The aforementioned tracing might be e.g. logging information: Who, when, what and why the data was removed. If an employee, for example intentionally or accidentally destroys data, it can be found in the logs. Anyway, there is a possibility for a software error, which might alter or destroy data accidentally. Logging the aforementioned malfunctions must be handled in the system. Organizations should also remember that they need a valid reason for storing log information. In addition, they must remove logs when there is no legal base to keep them anymore. Data in the logs should have an expiration date/time, which helps removing the expired data. There can be some circumstances when setting an expiration date/time for data afterwards can be problematic, because organizations cannot point it out clearly when data was created.

4 Example application

4.1 Target organization

Headpower Oy provides cloud services for infrastructure network and contractor companies. The products consist of applications and instructions, which will help with the daily work of network operators, designers, contractors and other suppliers in the energy industry. The corporation constantly develops its products and services to meet the needs of customers and industry better. Headpower's experts have regular dialogue with authorities and they want to have influence in the legislation in

advance so that the operational level's visions and needs are considered by legislation in the right way. (Headpower is the sum of its customers 2018.)

The corporation was established in Harjavalta in 2001, and the offices in Espoo and Jyväskylä were opened in 2005. The company acquired Cybersoft Oy in 2009 and its office is in Tampere. In 2018 Headpower consists of over 40 professionals of IT and energy industry. (Headpower is the sum of its customers 2018.)

4.2 History of example application

The example application has over 15 years of history and it has many clients and users. Many software developers have developed it during its existence. Most of the original developers have changed firms and naturally, a piece of the application's architectural information has been lost with the former employees gone. New features have been added and some old features have been removed. Databases might contain data of some former features, which do not exist anymore. The corporation's current developers might have a challenge to make the application to meet the GDPR requirements. The corporation has manufactured dozens of software products; however, it is difficult to handle all of them in one thesis, which is why only a one software has been chosen for this study. The architecture of the software has been described only at an abstract level due to product confidentiality issues.

4.3 Stakeholders

There is large number of people working around the application, and a great amount of personal identifiable information (PII) is processed in the software. It is important to understand who is using it and whose data is processed in the application. The clients are small and medium size of enterprises. The assigner corporation is a so-called data controller, and Headpower's role is to provide cloud services as a data processor. In the application, there are user accounts containing naturally personal identifiable information of the users. A typical user is an employee from the assigner corporation. Consumer customers do not use the software, however, their PII is processed inside the system, which is why Headpower must have instructions for its help desk employees for a scenario in which a consumer customer asks directly from

Headpower what kind of data the corporation has of him/her. Another group of users are Headpower's own employees working as software developers in the help desk as product owners or salesmen. These people also have user accounts and their activity is logged into the system. The conclusion is that much of the assigner's employees', Headpower employees' and consumer customers' PII's and other persons' personal data is processed in the system. Figure 3 shows few stakeholders who have a relation with the example application.

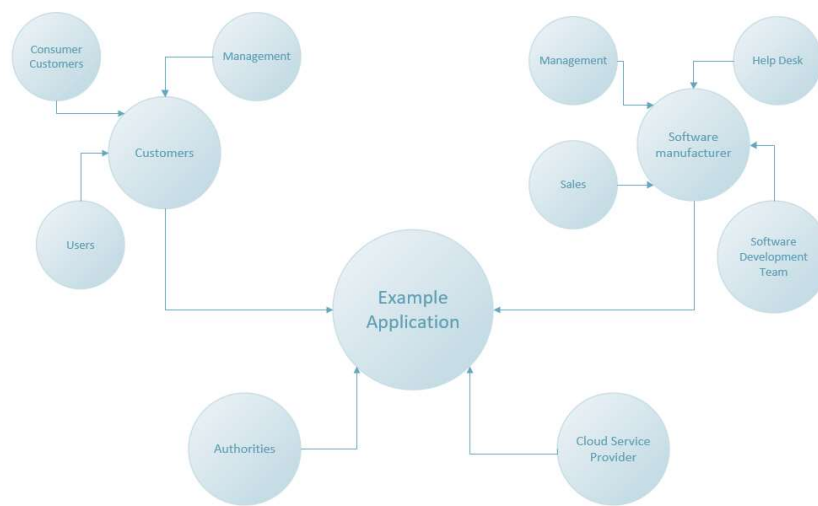


Figure 3. The example application's stakeholders

Headpower's help desk takes care of customer support tickets sent by users. The help desk has been organized into different levels of support tiers and each level has access to the system as well as to customers' data. The help desk employees communicate continuously with the software development team when they are solving customers' problems.

The software development team consists of software developers and testers. The team designs, implements, tests and releases new versions of the application. The team has a product backlog in which the team's work is listed, i.e. product backlog items (PBIs) and bugs. PBIs and bugs have been sorted out with priority from the top to the bottom, and the highest priority items come first on the list. A product owner has the largest role when the priority of PBIs is to be decided. The development has

been split into regular cycles, sprints. A sprint is usually a 10-day long period and it belongs to a release that is normally a four- sprint long period. The software development team publishes six releases during a calendar year and small hotfixes, i.e. small features or bug fixes. Every sprint begins with a sprint-planning meeting and the result of the meeting is a sprint backlog. This backlog contains only items that the team implements during next 10 days. Figure 4 presents an abstract level description of the software development process.

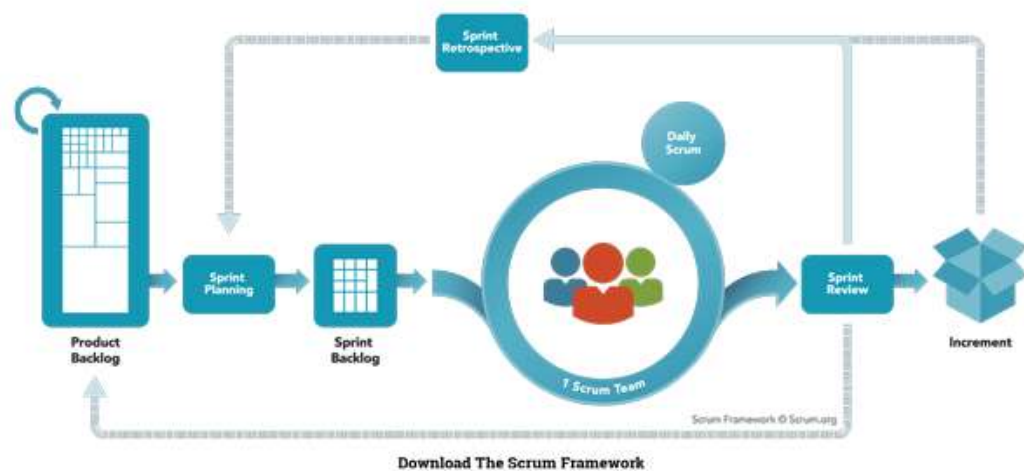


Figure 4. The software development process (What is SCRUM? 2018)

Corporation management, salespersons and product owners use the application for demonstrations when they have meetings or training sessions with their clients. In a training session, users use the application and some personal identifiable information (PII) is processed. It is important that the aforementioned data is removed after trainings and demonstrations, because if there is personal data e.g. stored in the database. All kind of PII must have a legal base for processing according to GDPR and it is better not to keep unnecessary data in databases after training sessions.

4.4 Architecture

The basic idea of the service is that a user needs only a web browser and an internet connection, after which the user is able to start using the software. The example application's architecture consists naturally of a web server, which is the base for the

application. A database server has several databases where customers' data containing personal identifiable information has been stored. Each database might contain hundreds of database tables and many million rows of data. A file server has been built for a storage for files containing several million files. The architecture of the example application can be seen in Figure 5.

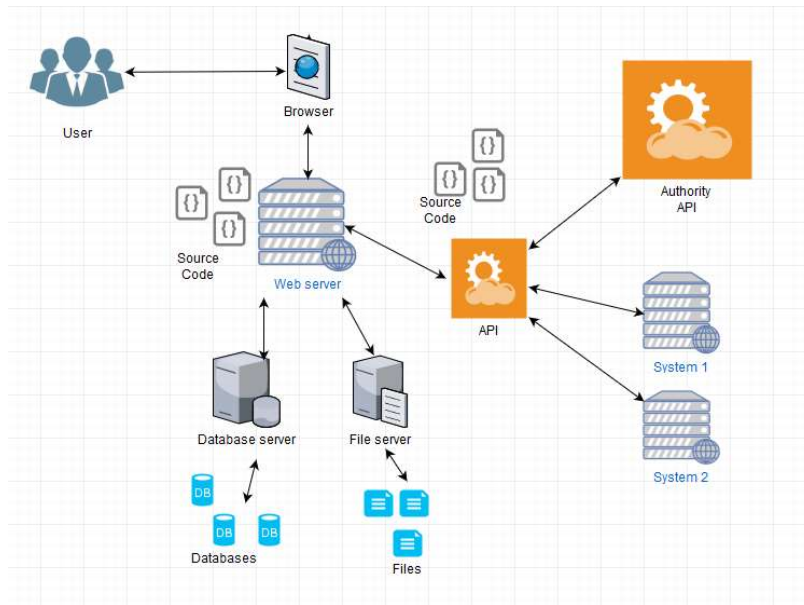


Figure 5. The architecture of the example application

The example application also has Web API used by other systems for reading data from the application and storing data into the application's database. Application Programming Interface (API) integrates two or more different applications together. An API consists of several elements; namely, functions, protocols and tools which developers use to build applications. A typical purpose for using APIs is that it accelerates development of building applications. The idea is to provide a part of the application's functionality out-of-the-box. Hence, a developer does not need to implement the same code each time they build a new application. Usually Web APIs have been served via HTTP interface, which offers a different type of media for response; however, typical response types are XML and JSON. (Pedro 2017.)

4.5 Registration into Headpower

A new client corporation must register a new account by themselves, because Headpower does not add new companies or users. After the client has registered on Headpower's web page, they must add user accounts. The client must accept the terms of service. Then, depending on the case, a service desk will open the requested product licenses and add the settings for a new client. A new client must sign a contract if they want to use Headpower applications.

5 Steps to GDPR compliance

5.1 Existing GDPR implementation model

The literature review's result, the previous surveys about GDPR, is analyzed in this chapter. The motivation for the literature review was to find a model which explains how to implement GDPR requirements into the software. The more detailed research implementation is explained in Chapter 2 Research implementation and the process of the literature review is explained in the Chapter 2.3 Conducting literature review.

5.1.1 Previous surveys of the GDPR implementation

GDPR – Six Months After the D-Day, the Master's thesis by Lehtisalo (2018) is the survey into time after the GDPR enforcement day since 25 May 2018. Lehtisalo's goal was to find out what has happened so far, whether there have been any court decisions or case studies about GDPR. In addition, he wanted to find out companies' opinions on the new regulation. (Lehtisalo 2018, 14, 19.)

GDPR-Strategy Management at a SAP Organization is the Master's thesis by Mononen (2019). The subject of the thesis is about how to lead General Data Protection Regulation implementation project in the SAP organization. Mononen describes e.g. SAP Contact Center software changes and then analyzes how the managing of the GDPR project is succeeded. Mononen writes that very often managing the GDPR project is only software or technology oriented. When leading a change project, people should be adopted into the project, because e.g. cooperation with a company's own staff increases the chance to adapt the changes more

successfully. Leadership methods have been described as a diary in Mononen's thesis. Every week during the GDPR project Mononen reported the progress of the leadership. (Mononen 2019, 6, 16–17.)

Mononen (2019) reports in his thesis that SAP Contact Center system has new features, which makes aforementioned software GDPR compatible. The Contact Center product has a new web based DPO tool, which can be used for editing and removing contact information from the database. With DPO tool contact information queries can be made e.g. with phone number, email address, network address and name. The system also has automatic methods for removing of contact information and anonymization of the data. In addition, the storage time for the data can be set up in the SAP system. (Mononen 2019, 10–11.)

Mast's (2018) Master's thesis of SAP authorization concept renewal project and GDPR in company X contains GDPR changes what is developed into SAP system from the data security aspect. Mast reports that following GDPR features have been implemented into the SAP system as follows: Read access logging, data archiving, data anonymization, data pseudonymization and data masking. Mast writes that his thesis does not contain detailed explanation of how company X has implemented GDPR requirements. (Mast 2018, 7, 36–46.)

Pulkkinen (2018) in his Master's thesis "Cloud outsourcing guidelines and data protection regulation in Europe : Context of online banking services" writes about providing a view on legal frameworks, recommendations and requirements to mobile banking channels and internet bank in the outsourced delivery environment. Pulkkinen writes that the application of his thesis might help service providers to reduce risks by providing information of security controls, processes and by reporting on a practical level. (Pulkkinen 2018, 12.)

Pulkkinen (2018, 54–56) lists a few security controls in his thesis, namely, e.g. intrusion detection, prevention and vulnerability management, ICT system hardenings and vulnerability scanning practices or implementing application firewalls, denial of service protection, key management and data encryption.

Kylmänen's (2018) Master's thesis "General Data Protection Regulation – Requirement Analysis of Customer Personal Data: Case Study" is the case study for a

company that has implemented GDPR requirements e.g. to avoid penalties and to give a customer more control over his/her personal data. One of the aforementioned thesis research questions is how GDPR requirements can be implemented into existing systems. (Kylmänen 2018, 20.)

Kylmänen (2018) writes in his thesis that first they found out which GDPR requirements the systems have already and which part of the system does not require changes. He used a table listing the aforementioned requirements and systematically checked the selected systems. This way he found out which GDPR requirements needed to be implemented into the system. Kylmänen listed GDPR system changes, which were: compliance statement and consent, cookie banner statement and access rights (e.g. mandatory password changes from generated password). (Kylmänen 2018, 42, 52–53.)

5.1.2 Conclusion of the GDPR implementation model

The previous surveys (Master's theses) about how to implement GDPR requirements e.g. into information systems show that there is not yet common guidelines how it should be done. Instead of that, small pieces information can be retrieved from the previous surveys how the GDPR requirements have been implemented in different organizations. The previous surveys also show that there is not enough technological hands-on information about how the GDPR requirements have been implemented for example into the legacy system or how to make the system meet the GDPR requirements when the system has e.g. over million files attached.

Kylmänen (2018) has the same kind of approach to the implementation of the GDPR requirements in his survey (Master's thesis), which is that the organization should get the knowledge of what personal data there is in the systems:

“Fingerprints are listed as sensitive data and are used to log in to warehouse UI. After the user first time logs in to warehouse UI using his PIN code, he can add a fingerprint to get an alternative way to log in.” (Kylmänen 2018, 39.)

Kylmänen (2018) also created the list of GDPR requirements that already exist in the systems. This way he understood which GDPR requirements still need to be implemented into the example systems. (Kylmänen 2018, 42, 52–53.) Headpower

had the same kind of approach to the GDPR compliance; however, the implemented features were different. Headpower's GDPR implementation is described in Chapter 5 Steps to GDPR compliance.

5.2 Personal data mapping

An organization must have knowledge of what personal identifiable information (PII) they process. Without this knowledge, it will be difficult to be sure that the organization processes PII in a way that GDPR demands. Article 30 of General Data Protection regulation states, for example, that the organization should name a purpose for each PII processing and set time limits how long the data will be kept in a storage. In addition, the controller and processor must be capable of making the records available to authorities and audits on request. The GDPR Article 30 does not give instructions what an organization should do to make organization's data processing activities meet the requirements. Data mapping might be a good method for this kind of purpose. (Biscoe 2017.)

Data mapping aids an organization to identify information kept in the organization and how it will be transferred from one location to another. Data mapping should contain at least the following elements (Biscoe 2017):

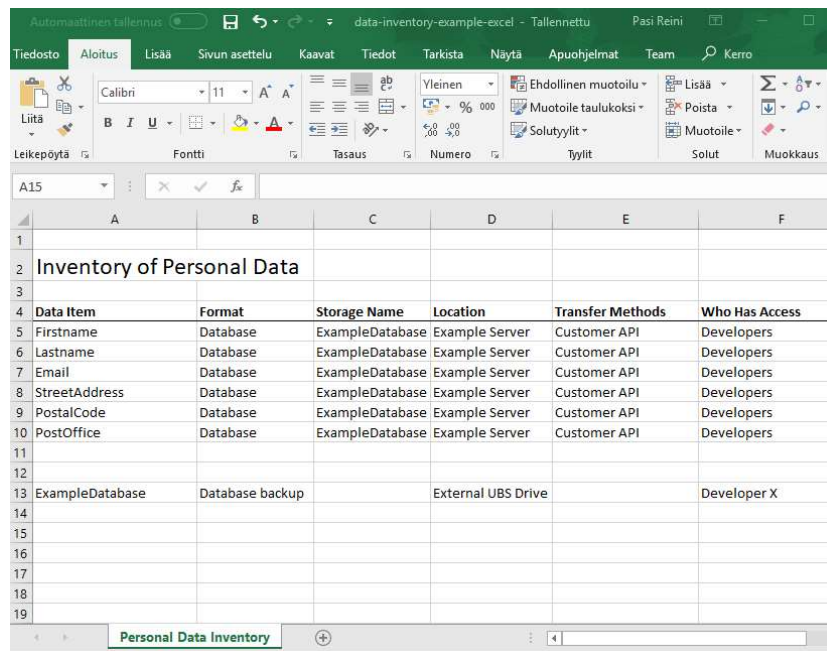
- Data items (names, addresses, emails etc.)
- Formats (e.g. databases and files)
- Transfer methods (e.g. API)
- Locations (e.g. clouds, third parties or on premises servers)
- Who has access to the data.

The application has a 15-year old history and many developers have worked on it. Old features have been removed and new ones added during the existence of the software. Some developers have also changed firms during the application's existence and technical documents probably have not been kept up to date. The developer team has a challenge with GDPR and getting the application compliant with the regulation. The development team started the work with a planning

meeting of the data mapping. The team listed possible locations where the data might have been stored. And the initial list was as follows:

- Production environment
- Test and demo environments
- Office premises servers and other devices
- Developers' workstations
- USB memory sticks, external USB drives and other similar data storages.

The abovementioned locations might contain personal identifiable information (PII) in databases, files, logs, backups and printed materials (e.g. screenshots of an application). The development team must check each location and clean them of PII which has no legal base for keeping it in storage. All possible locations for personal identifiable information were documented (for example, a data mapping document in Figure 6) and this way the development team could create a plan what tasks they need to perform for e.g. cleaning the data away from unnecessary locations. In addition, the team had an idea for evaluating how much workload it would take for performing those tasks.



Data Item	Format	Storage Name	Location	Transfer Methods	Who Has Access
Firstname	Database	ExampleDatabase	Example Server	Customer API	Developers
Lastname	Database	ExampleDatabase	Example Server	Customer API	Developers
Email	Database	ExampleDatabase	Example Server	Customer API	Developers
StreetAddress	Database	ExampleDatabase	Example Server	Customer API	Developers
PostalCode	Database	ExampleDatabase	Example Server	Customer API	Developers
PostOffice	Database	ExampleDatabase	Example Server	Customer API	Developers
ExampleDatabase	Database backup		External UBS Drive		Developer X

Figure 6. An example document of the personal data mapping

5.3 Who has access to data and why

After the personal data mapping the corporation had an idea who has access to each data location. The next phase was to define if these persons really needed to have access to the specific data. In addition, the corporation had to decide who can alter client's data and personal identifiable information. This is why policies and processes must be checked carefully to prevent easy mistakes and data breaches.

5.3.1 Helpdesk

The main role for helpdesk is to help customers with problems which might occur e.g. with applications. Sometimes it is required that an employee who is working at the helpdesk needs to access the customer's data, because otherwise it might be difficult to trace the cause of a problem. When a helpdesk employee needs access to a customer's data, he/she must fill in form where there is a field for giving the reason to get access to a customer's data. Access to customers' data has been prevented programmatically; therefore, without a valid reason, a helpdesk employee cannot access the data. A valid reason is, for example, a ticket number of the helpdesk support ticket. Headpower has programmed a form and functionality for preventing unauthorized access to customers' data as one feature of the GDPR project (Figure 7 presents the example form).

Figure 7: Example form to be filled in before an employee can access personal data

Headpower's helpdesk employees have instructions for phone and email support requests, because this way the corporation can prevent social engineering security breaches. Sometimes a person calls the helpdesk or sends an email and then e.g. asks for his/her password to the system. The password cannot be given, because the phone call might be a malicious contact. Clients will lose their trust to an organization that gives their passwords on the phone. Therefore, only simple instructions can be given in a phone call, however, no personal identifiable information or sensitive data can be exposed. Data changes can be performed e.g. when a registered user sends an email to the helpdesk.

5.3.2 The software development team

The software development team must fix bugs or solve complicated problems that might occur in the production. Sometimes it is impossible to understand the root cause of a problem without a customer's data. Building the data for a customer's incident scenario might be time consuming and difficult, because very often bugs must be fixed fast. It will be much faster, easier and cheaper for each party to use

customer's data for problem tracing, which is why the rules are the same for the development team as for employees working at the help desk. A developer must fill in a valid reason before he/she can get the access to the client's data.

There must also a policy for how developers handle database backups, because there is a large possibility for data breaches if backups are to be stored e.g. in external USB memory drives with no encryption. Somebody can steal the external memory easily or by mistake, it might be lost outside the office. Then a client's personal identifiable information and other sensitive data will be exposed to outsiders.

5.3.3 Corporation management, salespersons and product owners

Corporation management, salesmen or product owners do not need access to the example application's data or databases by default. Only the product owner has the permission to see customers' data, however, the reason for viewing the data must be logged in the same way as helpdesk employees must do. It has been prevented programmatically to gain access to the customer's data without a valid reason. Customers' contact information has been stored in CRM and salespersons can use this data.

5.4 Consent

Data controllers must keep records for how and when an individual gave consent for his/her data processing. A consent must be given by an individual and it cannot be automatically created with a system, or pre-ticket boxes cannot be used by a data controller. The controller must also understand that an individual has right to cancel a consent whenever he/she wants. (Curtis 2018.)

Headpower's solution for this is that when a user logs into the system for the first time after the GDPR project has been released, the system will prompt a confirmation box for the user and ask for a consent. The confirmation box contains links to terms of service and to a data security report which describes the needed GDPR information. In addition, there is a checkbox, which needs to be checked to give consent for personal data processing.

5.5 Lawfulness, fairness and transparency

Earlier in this document, Chapter 3.3 Principles of GDPR introduced a few principles, which means that a person's data must be processed in a way to meet lawfulness, fairness and transparency requirements. In Chapter 3.4.1 The Right to Be Informed there is a list of issues which a register owner must inform, e.g. the contact information of the register owner, the purpose for personal data processing, the source of personal data, the legal base of processing and the kind of data that is to be collected. Headpower has created the Data security report of personal data processing (Figure 8), which is available for every user.

DATA PROTECTION SUPPORT - HeadPower portal customer and user register

1. REGISTRAR

HeadPower Oy
 Harakantie 18 B, 02650 ESPOO
 tel. [+358 10 841 5400](tel:+358108415400)
helpdesk@headpower.fi

2. REPRESENTATIVE OF REGISTER

Antti Jukarainen

3. REGISTERED RIGHTS

In principle, HeadPower complies with the EU data protection regulations and current legislation in the rights of the data subject.

Correcting Personal Data: A customer with an active user ID can check the information in the registry at any time through the internet data network and modify them if necessary. In addition, the request can be made as described in Connections section.

Checking Your Personal Information: A registered user with an active user name can check their records at the HeadPower portal at any time free of charge. If this is not possible, the registrar can make a request as described in Connections section.

Deletion of Personal Information: A registered person may submit a written request for deletion of data. Data will be deleted only if it does not endanger the legal protection of HeadPower Oy or HeadPower Oy's customer companies.

Right to file a complaint with the Data Protection Authority: If the controller refuses to act, the data subject may refer the matter to the data protection authority.

4. CONNECTIONS

All questions regarding the use of your rights and the privacy statement or other contact information should be sent to helpdesk@headpower.fi

Contact must be sent from email address of registered user, so that the request can be properly targeted to correct user. If this is not possible, the data subject must be able to prove that he is registered by a written request. A written request is made at the registrar's address of section 1. A personal appointment must be agreed for the request and the registrar will present a formal identity document at the meeting.

For repeated requests, the registrar may charge a fee based on its current price list based on the employee's working time.

5. USE OF PERSONAL DATA PROCESSING

The register is used to manage the access rights of the HeadPower portal, to customer relations and to manage customer relations, to implement the rights and obligations of the customer and HeadPower Oy. The register is also used for the processing of personal data in accordance with current data protection legislation for the purposes of online services, research and marketing with the consent of the customer.

6. GROUNDS FOR HANDLING OF PERSONAL DATA

Managing customer relationships and managing and utilizing access rights for HeadPower-portal. The registered user accepts the terms and conditions of the portal and agrees to the processing of personal data on the HeadPower-portal. Without approval, the use of the services is not possible.

7. DATA CONTENT OF THE REGISTER

The register contains the following information about the registered:

- Contact information (name, position, phone number and email address)
- User ID with identifying information and password.
- The companies, languages, roles and industry profiles represented by the user
- Login times and other user-related activities
- Consent to communication and other processing operations
- Information on system management and usage rights limitations

8. REGISTERED DATA SOURCES OF THE REGISTER

Customer information is collected at the portal when registering and during the customer relationship. A customer with an active user ID can check the information stored on the registry at any time, except for usage activity data, if necessary. The register information will also be updated based on notifications made by the customer to the registrar.

Figure 8: Part of Headpower's Data Security Report

The data security report also contains descriptions for delivery of the personal data, storage, archiving and distribution of personal data and general description of technical and organic safety precautions. In addition, users have been informed of end user term changes (Figure 9).

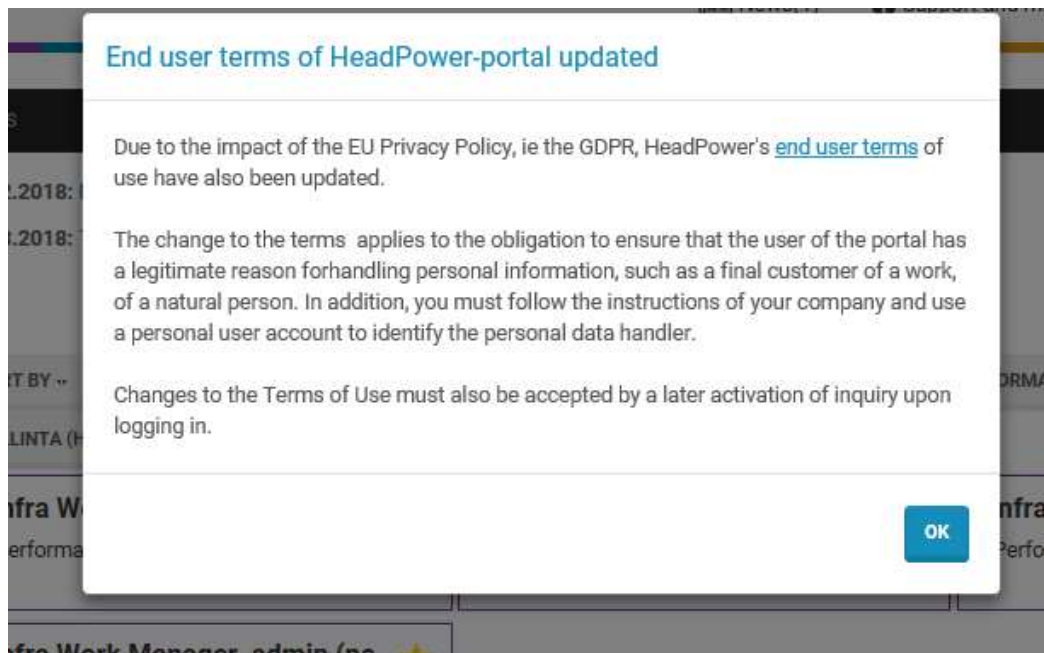


Figure 9: Information note for users of changes that have been made to end user terms

5.6 Data minimization

Earlier in this document data minimization was introduced as one of the GDPR's privacy principles. It states that only information that is needed and nothing more is to be collected. The organization should check in the personal data inventory results and e.g. the registration form and other personal data processing forms from the system that there are no special categorized data fields. Special categorized data is e.g. person's ethnical, political or health information (For more detailed information, see Chapter 3.2 Personal identifiable information definition under GDPR).

Headpower checked all software forms, web APIs and databases for the data which has a legal base and is to be processed and stored only. In addition, the corporation ensured that the special categorized data (sensitive data) is not being processed.

5.7 Storage limitations

Within this topic, storage limitation has been introduced as one of the GDPR privacy principles. The idea is to keep the data stored only the time that is necessary and not any longer. During the years, some old customers have quit using the application and terminated the license. All personal data has not been deleted at the same time and it still exists in the database. According to GDPR, there are no legal purposes to keep this kind of data in storage anymore, which is why the data processor and controller should know which data should remain in the database. The corporation should have a plan on implementation how to remove aforementioned data.

Headpower created a document (personal data mapping) of data that has been stored in databases and other locations. Then the corporation's software development team implemented automatized cleaning services for databases and file storage. The cleaning services are described in more detail in Chapters 5.8.2 Do not Archive Unnecessary Data, 5.8.3 Clean Databases Automatically and 5.8.5 Automatic File Removal Service.

5.8 Software's GDPR features

This chapter contains the example software's GDPR features, which have been implemented to help users and Headpower's employees with their daily job. The aforementioned features are also the steps taken to make the example application meet the GDPR requirements.

5.8.1 Change, access and error logs

Log information is a manually or an automatically created document of an event occurring in a network or in an information system. Logs are needed for tracing problems from a system and monitoring that a system is working as it was designed. In addition, logs offer a legal protection for users and system administrators. Best practice includes a policy for log handling, e.g. a plan for e.g. how long logs will be kept in a storage, where logs will be stored, who can access them and what the purpose and legal base is for keeping these logs. There are different types of logs that can be named e.g. maintenance log, change log, error log and access/audit logs.

A life cycle of log processing contains e.g. the following phases: collecting logs, analyzing logs, storing logs, giving logs to third parties and removing them. Handling logs is one of the most important practices for maintaining information systems and monitoring their security. (Lokiohje 2009.)

Checklist for legal log processing (Lokiohje 2009) as follows:

- Identify why and what is the purpose and legal base for each log processing.
- Evaluate the need for each log data entity that is being stored.
- Identify what personal data might be stored into logs.
- Find out how each log should be protected.
- Pay attention to legal aspects of log monitoring and when necessary, perform e.g. co-operation negotiations with employees.
- Notify users and other stakeholders of log processing.
- Pay attention to requirements on the person register given by authorities if the log is to be a person register.
- Design and document a purpose for log storing and ensure its implementation.

Headpower has listed possible places where log data has been stored and then identified a storage purpose for each log data. In addition, a time was evaluated for each type of log that how long records will be kept e.g. in storage. An expiration date flag was generated for each log row where it was technically possible. Then an automated service was created for removing the expired log records automatically.

The employees of the organization were educated to meet the new requirements of personal identifiable information handling. Especially, accessing customers' data must have a valid purpose. An employee must write manually a valid reason why he/she wants to gain access to a customer's data. An aforementioned reason is the compulsory log information and without it, the employee cannot access the data. It has also been prevented programmatically in the application that without a valid reason the access to the data is impossible. In addition, employees have signed a contract for personal identifiable information handling, which has been instructed by the corporation management.

5.8.2 Do not archive unnecessary data

There are two common strategies for deleting data from a database: soft and hard deleting. Soft deleting means that in a database table there is a column named for example as `isDeleted`. When using this type of delete, a software will set a flag into the aforementioned column stating that the data is deleted (archived). The data will not be removed from the database and it stays there until hard delete has been performed. The soft delete might cause complex database queries in the future, because developers must every time programmatically filter the archived data out of the query results. In addition, database queries might become slower in the future, because archived data still lays in a database and it will be handled in the same way as not archived data. When using a hard delete, the data will be permanently removed from a database. The hard delete brings, for example, the following advantages: A smaller table is easier from the aspect of maintaining. Rebuilding indexes is faster and their size will be smaller. In addition, a smaller table has better performance. (Kloeten 2009; Pinal 2010.)

Storage limitation is one of the GDPR's six privacy principles (3.3 Principles of GDPR) which stand for that the personal identifiable information should be removed after it has no purpose or legal base for keeping a in storage anymore. The example application had been implemented in the way that it used soft deleting for data. Now the former soft delete features (data archiving functionality) have been changed to hard delete, because the target corporation does not want to keep useless data in their databases. Perhaps in the future, the useless data will be corrupted and even cause problems for the software or when developing new features. Some advantages of hard deleting are listed above. It might also make software code easier to understand and there will be fewer possibilities for incidents that old and corrupted data might cause when data is being transported, for example via Web APIs into other system.

Sometimes there is a technical or juridical reason why, for example, user accounts cannot be removed from the system. Deletion of a person might cause that the data will lose its referential integrity. This can easily cause data corruption. A way to handle this problem is to create e.g. a database view that returns users who cannot

be removed from a system. A juridical reason for keeping archived users in a database is e.g. when the data has been retrieved from an API provided by authorities who demand that a log information of queries shall be stored for a defined time.

5.8.3 Clean databases automatically

“SQL Server Agent is a Microsoft Windows service that executes scheduled administrative tasks, which are called jobs in SQL Server 2017” (SQL Server Agent 2017).

SQL Server Agent runs scheduled jobs, for example a scheduled procedure for taking database backups. These tasks can be set to run automatically e.g. every Friday at 23:00 o'clock. If the task encounters a problem, the agent records the event and it can be configured, for example, to send notifications to a person responsible for databases and aforementioned jobs. A job is a series of steps (actions) that the SQL Server Agent is running. Jobs can be executed one or many times depending on the nature of a task and a schedule. A job's success or failure can be monitored easily. A job can run on a local or on many remote servers. (SQL Server Agent 2017.)

Stored procedure is a Transact-SQL script containing one or more statements and is usually run on SQL Server. If a developer needs the same database operation in many different scenarios, it might be useful to write a stored procedure. This eliminates the need of writing the same code multiple times. When a database operation is on the data tier only, it reduces network traffic between a client and a server, which also brings stronger security, because the procedure controls what activities can be performed e.g. to databases instead of giving this permission for multiple users or programs. In addition, when a procedure is used over the network, an execute statement is the only visible part of it, and malicious users cannot see an underlying database or table object names. SQL injections can also be prevented, because a malicious user's own SQL statements cannot be embedded if a procedure's parameters are treated as a literal value and not as executable code. Maintenance is also one of the benefits, because one needs to change only a structure of underlying databases and the client program does not need to know anything of these changes. Stored procedures have also better performance time, because they will be compiled

at the first time when they are executed. Therefore, later the server does not have to create an execution plan every time, when a procedure is to be used. (Stored Procedures (Database Engine) 2017.)

The software development team decided to implement database procedures and jobs into SQL Server Agent (Example of SQL Server Agent in Figure 10) for taking care of GDPR requirements (the benefits of the database procedures were explained above). The team designed different procedures for cleaning databases of the data that is useless.

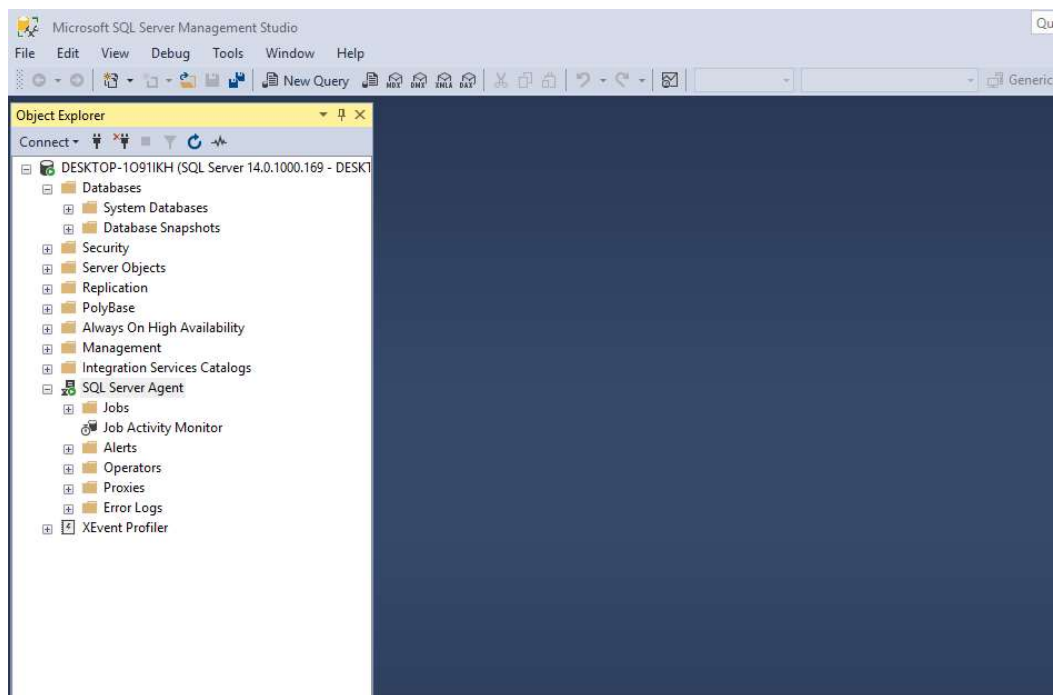


Figure 10. SQL Server Agent on example server

In addition, the software development team created procedures for auditing databases so that there is no GDPR incompliant data. All data removes will be also logged so that later the team can check why e.g. a user does not exist anymore and who removed it. Audit logs of data deletions have been stored also in another location, because if a source database is corrupted, then it is important to check logs and trace which data has been removed earlier. This way the corporation can

prevent a problem that the removed data might be restored back to a database accidentally.

Procedures for cleaning and auditing tasks were scheduled into SQL Server Agent as jobs (dummy example of database procedure in Figure 11). This automation was made, because nobody remembers to run procedures manually e.g. every Friday and it is a waste of time. The software development team has more important tasks to do than manually execute stored procedures every week. The SQL Server Agent also writes an event of errors, which a procedure might encounter during its execution. The agent was also configured to send a notification of events to the software development team. Should a problem occur, somebody in the team can start tracing the problem thrown by the agent. Hence, the team must monitor the agent and history of its jobs regularly, because one cannot fully trust that the automation always does the job without interrupts.

```

USE [ExampleDatabase]
GO
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
--CREATE PROCEDURE [dbo].[RemoveArchivedCustomers]
AS
SET NOCOUNT ON;
DECLARE @customerId INT;
DECLARE @dbCursor AS CURSOR;
-- All customers which needs to be deleted
SET @dbCursor = CURSOR FOR SELECT c.Id FROM ExampleDatabase.dbo.Customers c WHERE c.Archived = 1;
-- Open cursor
OPEN @dbCursor;
-- Get customer's id for delete
FETCH NEXT FROM @dbCursor INTO @customerId;
-- Loop through all customers which were found
WHILE @@FETCH_STATUS = 0
BEGIN
-- Delete customer
DELETE FROM ExampleDatabase.dbo.Customers WHERE Id = @customerId;
-- Set log text
DECLARE @logText VARCHAR(1000);
SET @logText = 'Deleted customer, id = ' + CAST(@customerId AS VARCHAR(1000));
-- Write audit log that we know who/what deleted the customer
INSERT INTO [AuditLogs].[dbo].[AuditLog] (
    [Operation] -- Name of the operation (Insert, Update, Delete)
    , [DatabaseName] -- Name of the database
    , [TableName] -- Name of the database table
    , [TargetId] -- Target Id (Id of a customer), if we must restore a database, we know which records we should remove automatically
    , [LogText] -- Audit log text
    , [Logger] -- Script that was executed
    , [UserName] -- Name of the user who executed
    , [PersonName] -- First- and lastname of the user
    , [Company] -- Company of the user
    , [LogDatetime] -- Timestamp
    , [ExpirationDatetime] -- Expiration datetime for a log record
) VALUES (
    'Delete'
    , 'ExampleDatabase'
    , 'Customers'
    , @customerId
    , @logText
    , 'PROCEDURE dbo.RemoveDeletedCustomers'
    , 'sa'
    , ''
    , ''
    , GETDATE()
    , DATEADD(MM, 6, GETDATE()) -- Keep log records for six months as long as database backups will be stored
);
-- Get next customer's id for delete
FETCH NEXT FROM @dbCursor INTO @customerId
END
-- Close cursor
CLOSE @dbCursor;
DEALLOCATE @dbCursor;
GO

```

Figure 11. An example procedure for deleting archived customers

The team must create a plan and schedule dates when to check the history of agent jobs. Otherwise, nobody will monitor them until a problem has been detected. In the worst scenario, users will notice something strange in a software and report it to the help desk. An example of a database job can be seen in Figure 12.

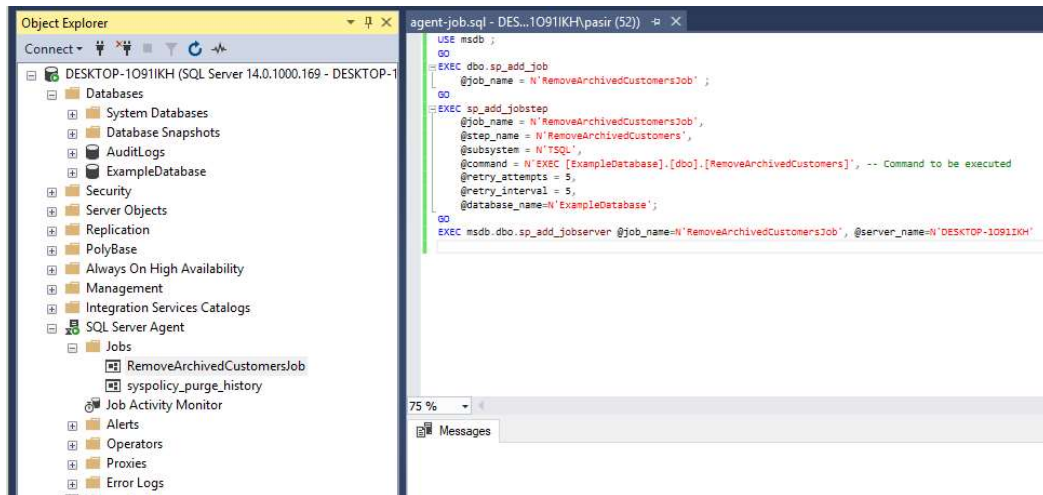


Figure 12. An example script for an agent job that executes a procedure

A job can be scheduled by using Transact-SQL or doing it from properties of a job (Figure 13).

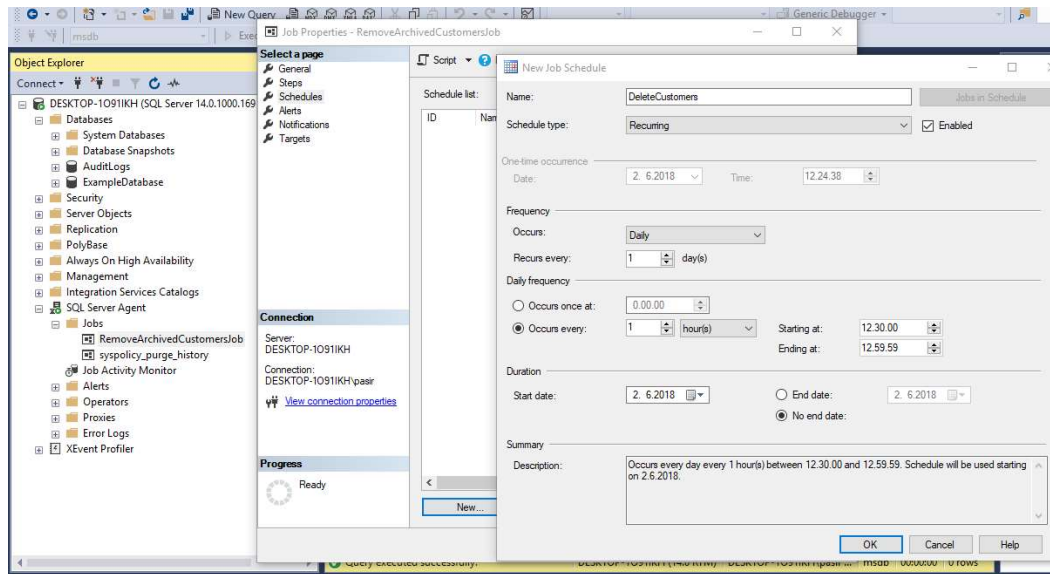


Figure 13. The example schedule for the SQL Server Agent job

The execution history of a job can be viewed by right clicking the job with a mouse and selecting “View history” (Figure 14).

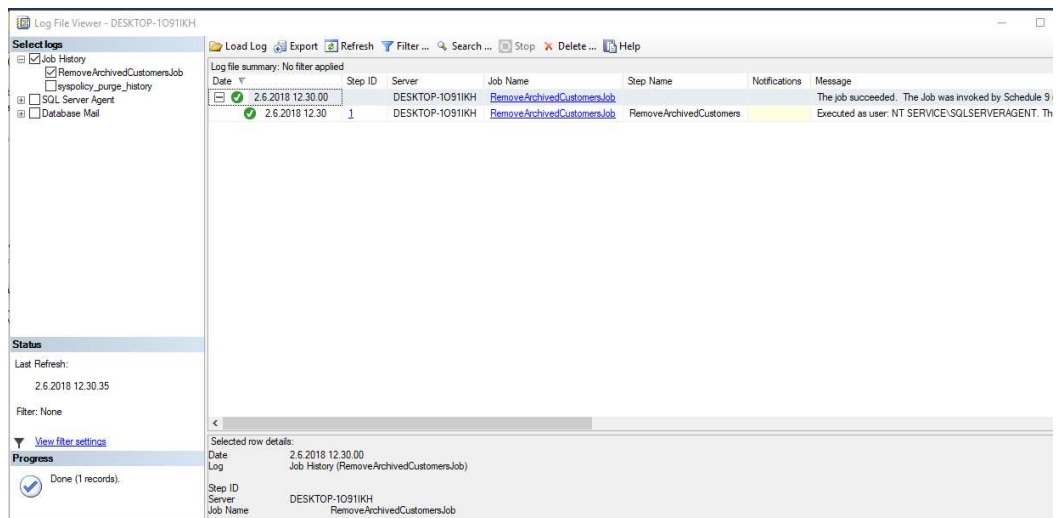


Figure 14. History of an SQL Server Agent Job

5.8.4 Person search

Headpower implemented for its clients a feature, which allows customers to search e.g. a consumer's personal data from a system. This way users can make searches themselves and will not send this kind of person search support requests to helpdesk all the time. In the system there is data of thousands of people and this will save the time of helpdesk employees when they can focus on solving a customer's other problems. By default, the feature will be open for customer's admin user. The example application's person search demo can be seen in Figure 15.

Example Application

Person search

Firstname
Example

Lastname
Customer

Search 3 items found

#	Item	Created	Modified	
1	Contract 3	1.10.2017	1.10.2017	View
2	Contract 2	1.12.2015	1.12.2015	View
3	Contract 1	1.11.2014	1.11.2014	View

Previous 1 Next

Figure 15. Example application's person search feature

5.8.5 Automatic file removal service

The system has thousands of customers' files, and it might be impossible to check which one of them contains personal data, which is why the deletion of files has been automatized by creating a service to take care of cleaning them out of the system. All unnecessary files will be removed automatically when there is no purpose for storing them anymore. A user marks a file to be removed and the service will read from a database which files it should remove. The benefits of getting rid of

unnecessary files are following: There are no useless files for taking e.g. a disk space; the customer does not need to check from the user interface file by file for files that should be removed. This feature will save time of both organizations, the data controller and processor. The customer is not interested in removed files and neither is the data processor. In addition, the storage limitation of the GDPR's six privacy principles (3.3 Principles of GDPR) has been implemented.

5.8.6 HTTPS over HTTP

The website transfers content from a server to a user's browser through Hypertext Transfer Protocol (HTTP) and it is a text file with clear text data. Anyone might see the content of this file, because it was not encrypted. Nowadays, most of the web sites use HTTPS instead of HTTP to ensure privacy and security of data. Therefore, the main benefit of using HTTPS with an SSL certificate is security, because all content will be encrypted and secured, e.g. username, password and customer's data. (Hopping & Millman 2018.)

Network traffic between a web server and user's browser and traffic between APIs and other systems should be protected by using HTTPS and SSL certificate. This prevents data breaches that might occur when e.g. API sends data to another system over the internet. HTTPS and SSL certificates are not a new thing for the research target organization; however, the connections between different locations must be checked and secured. This way the part of GDPR's principles has been implemented, which demands that personal data must be processed in a manner that it is secured and protected against unauthorized or unlawful processing (3.3 Principles of GDPR). The corporation's software development team ensured that all network traffic between each server and clients has been protected with HTTPS or encrypted with other methods. Some of the example application's network traffic can be seen in Figure 16.

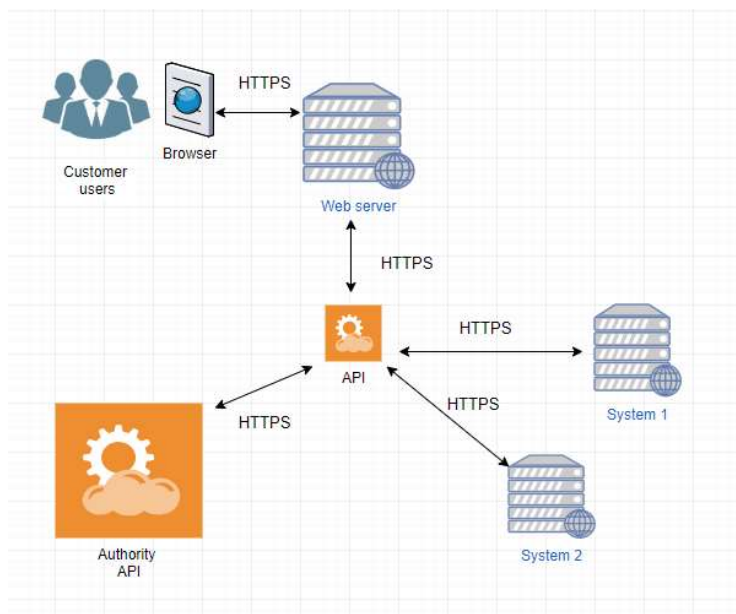


Figure 16. Connections protected with HTTPS

5.8.7 Right to data portability

GDPR's Article 20 stands for the right to data portability, which means a person has the right to get the personal data that is concerning him/her. The data must be e.g. in a structured, commonly structured and machine-readable format. In addition, a person shall have the right to transmit the data to another controller's system. (Regulation (EU) 2016/679 of the European parliament and of the council 2016.)

This is a problematic requirement of the GDPR, because not all data can be transported to another controller's system. The example application contains e.g. contracts between data controllers and consumer customers. The data is important from business point of view to a data controller and it has a so called business purpose for processing and storing it. The example application has been built to provide e.g. contract data as PDF files, which can be given to the customers when needed.

Users might have qualifications, e.g. certifications for dangerous work. The earlier mentioned certifications are sometimes person's own data and can be considered to be transported between different companies. A typical situation is e.g. when a contractor rents its employees to another company's projects. Naturally, there

should be then a feature for this kind of data portability, because it will simplify trading employees between projects.

5.8.8 Review of existing features of the application

When an application has e.g. over 15 years of history behind it and many developers have worked on it, it might be a good idea to review which features are still valid and which should be removed. Some features have been added and removed during the existence of the application. For sure there will be GDPR incompatible data located without any purpose e.g. in databases, because probably nobody cleaned it from the database when old features were deactivated.

A very common technique to prevent abovementioned situations is refactoring, which means that the internal structure of an application will be changed without changing its external behavior. There are at least three different refactoring areas: Refactoring source code, database and user interface. (Veerraju & Srinivasa & Murali 2010.)

Refactoring code brings following benefits (Veerraju & Srinivasa & Murali 2010):

- Makes the code readable for other developers.
- Makes easier to maintain and upgrade code.
- Increases the quality of application design and implementation.
- Can be understood as an investment for the future.

Removing unnecessary data and tables from database not only increases the quality of the application but also helps an organization to get rid of GDPR incompatible data. Nowadays, some cloud provider companies will charge for used disk storage. It is probably not a problem with small databases; however, when the database size is e.g. 1000 Gigabytes, cleaning the database of unnecessary data might bring some performance advantages or cheaper bills. Code refactoring might also reduce possible attack surface from malicious users. Thus, making the application GDPR compatible is not only a heavy load for organizations but it is also a chance to improve matters. Companies should refactor their applications even if there was no

GDPR. The corporation's software development team refactored and removed few unnecessary features out of the example application during the GDPR project.

5.9 Notification of data breaches

One of GDPR's requirements is that the local supervisory authority must be notified of data breaches as soon as possible. The target organization must create a policy for this kind of action. There must be a defined group of people who will take over when there is a suspicion or an information security incident has already happened.

In Finland, the notifications of data breaches must be sent to Office of The Data Protection Ombudsman. The general instruction is that should a data breach cause risk to the rights and freedom of a natural person, the notification must be sent without undue delay (within 72 hours) when the controller has become aware of the breach. Typical data breaches are stolen computers, lost memory sticks, hackings, malware infections, fire in the sensitive data center, cyber-attack and mailing an official statement to wrong person. (Personal data breaches.)

The controller shall provide e.g. the following information (Personal data breaches):

- A description of the personal data breach (nature)
- Contact details of data protection officer or other contact who will offer more information
- Consequences of the breach
- Actions that the controller has taken to mitigate the effects of the data breach.

Headpower has an incident management team which takes over when big problems occur, which might cause a business risk and bad reputation to the company.

6 How to ensure GDPR compliacity in the future

The third research question answers the problem of how to ensure that the software meets the GDPR requirements in the future. The company must understand that it is not possible that once implemented processes, policies and software features keep

the organization or software GDPR compliant forever. There will be changes to the legislation and new software features will be designed and implemented.

6.1 Monitor GDPR changes

The organization needs to act fast when the government changes the legislation and rather before than when they take effect. Authorities might order fines, lawsuits or closure to non-compliant companies. New legislation might cause changes to existing procedures and policies and creating new ones always takes some effort. In addition, the employees must be trained for their new roles and responsibilities. A business cannot get rid of laws and it is important to guarantee business continuity. (Masson 2017.)

The organization must monitor GDPR changes to ensure that the newest regulation updates are taken into account, e.g. in the company's policies and processes. As it was stated earlier in this document in Chapter 6 (Is There Existing GDPR model), the regulation needs some generally defined instructions what firms need to do with the regulation, because the regulation leaves much to interpretation, which is why it is important to monitor the legislation changes and react to them.

6.2 Train employees regularly

The finest security controls and firewall will not help one's organization if e.g. a criminal gets a password into the network or sensitive information systems by using social engineering. A social engineering actor might make a phone call to a company's service desk and pretend to be a fellow employee or a partner in order to get sensitive information. In addition, an employee might be tricked to click a malicious link that came from an email or a fake source of social media. A person with bad intentions might just walk into the organization's facilities and install malicious USB sticks, steal assets or even harm people. (Goodchild & Hulme 2017.)

The employees should be trained regularly to prevent e.g. social engineering breaches, which is why the company has at least two different courses in security and GDPR for its employees. In addition, help desk instructions will be kept up to date all the time and when something new occurs, these changes will be trained to

employees. It is also important to train development team members in security policies, new technology, security controls and tools.

6.3 Internal audit

It is important to ensure that information security and regulatory compliance have been implemented correctly. An internal audit is a good tool for helping organizations to define which areas might need improving. It also exposes possible risks. Normally the scope of an internal audit consists of monitoring, analyzing and assessing e.g. risks of an organization. In addition, compliance with law might be reviewed. Recommendations are also a result of the internal audits. (How Do Internal Audits Work 2018.)

Laws will change and it is important to monitor these changes so that the organization stays compliant with them. A company must alter its processes and policies as a result of law changes, which is why the organization must conduct internal audits regularly to stay e.g. GDPR compliant. Security policies and software development processes must be audited so that the company can reduce security breaches and the performance stays high.

6.4 Take GDPR into account when designing and implementing features

Software development processes should contain working phases for analyzing information security requirements for personal data. Security requirements vary depending on the sector in which the corporation is specialized. Technical implementations must be designed in such a way that they go hand in hand with the risk level of the data. It is important to include impact assessment in the software development process from the very beginning to get e.g. data security requirements implemented correctly. Wrongly designed systems might be difficult to transform to be compatible with data security requirements afterwards. It is also important to ensure that security controls are implemented correctly during the development phase and afterwards. Technical security methods might be for example, Access control and data encryption / anonymization. (EU-tietosuojan kokonaisuudistus 2016, 22–23.)

When features are designed and implemented into the example application, GDPR and information security controls should be taken into account. This way the organization automatically creates GDPR compatible applications. The software development team is the Scrum team, which develops application features in sprints. GDPR requirements can be taken into account before, during and after the sprints. When a sprint begins, there is a sprint planning meeting and the team designs the tasks of given user stories for the upcoming sprint. A task contains a description how a feature should be implemented. Hence, a sprint planning meeting could be one event where to check if the tasks require GDPR work. During a sprint when a task has been done, the team members will review and test implementations. GDPR requirements can be reviewed and tested at the same time. After the sprint, there will be a retrospective meeting and the team has a chance to evaluate processes, tools, implementations and create recommendations for the future work. This way the GDPR requirements can be taken into account during the whole development lifecycle of each software feature.

6.5 Audit software features and data

GDPR and other features implemented into the example software and automatic cleaning processes need to be tested and audited regularly to check that they are working properly. In addition, unnecessary data and files have to be removed. The job history of database cleaning can be viewed in SQL Server's agent. Anyway, it is worth having e.g. an SQL script for checking all necessary database tables that might contain personal data so that there is no unnecessary data. The file cleaning service also needs automation for testing that all files marked removable have been deleted. Without automation it might be a frustrating task to find files to be removed. This way it can be tested that e.g. database job is fully functional. When a new feature is implemented or an existing one altered, the developers must alter and test automatic cleaning services to keep them doing what they were designed to do. The software should also be audited from the GDPR point of view so that all personal data has a legitimate purpose for processing and keeping in a storage.

7 Conclusions

As written in Chapter 5.1 (Existing GDPR implementation model), the result of the literature review is that, unfortunately, there is no such model which might make an organization's processes, policies or software GDPR compliant just like that. GDPR leaves much for interpretation and does not define what the correct level for the protection of personal data is, which the reason is why an organization itself must evaluate and define what a reasonable implementation is to meet the GDPR requirements. There are many different organizations and e.g. software products and cloud solutions that probably one standard or model is not enough to cover all industries. Maybe in the future there will be common industry specific GDPR instructions.

An organization should find information from documents, which authorities have released; online sources might also be an option and there should also be further training on the subject. Material offered by different consults should be analyzed carefully, because there is not just one way to implement GDPR. Then one could get a partner (lawyer) who is specialized in GDPR and other data security legislation. In addition, contracts should be made with each client and partner companies to guarantee that the responsibilities between each organization have been understood and agreed.

An organization must know what kind of data is being processed in its systems, where the data is located and where it will be transferred. It is a big challenge e.g. when a software has a long history. Data inventory and mapping will light up the organization's knowledge of the processed data items, the formats of data, transfer methods, locations and who has access to data. In addition, GDPR privacy principles must be noticed when personal data is being processed.

It is important to understand that only changes in a software's GDPR will not guarantee that everything is well done and the regulation requirements have then been met. As written in chapter 7.2 (Who Has Access to Data and Why), it is important to keep record who is browsing customers' data and why. In addition, the organization must prevent data access without a valid reason. In addition, for those

working at a help desk need instructions in order to prevent e.g. social engineering. The staff must also be trained for new data handling requirements and instructions.

Automatization helps to save time and the organization can focus on its main processes, e.g. software development. Reading or removing data manually from multiple databases and files from thousands of folders might be otherwise a frustrating task. Automatic database and file cleaning services still require monitoring that they are working as was planned. Scheduling the cleaning services also need strict planning; e.g. a database might contain several automated processes which might set, for example, locks to database tables. This might lead to a problem when reading or writing data from a database might fail and a cleaning or other automated process cannot complete its task. A database can also be in use all the time when data is written or read, e.g. via Web APIs or the software has many users online.

Monitoring GDPR and other legislation changes is important so that the organization is able to meet the requirements caused by the aforementioned changes. In addition, this way the organization can avoid expensive fines that might be set by authorities if matters have not been dealt with properly. It is also a good sign to customers and partners if the organization adapts legislation changes well and makes an effort for cyber security.

Training employees regularly from the aspect of security is important, because it helps to avoid easy mistakes e.g. at the service desk. A password or other sensitive information leaked through social engineering might cause a great financial loss for the organization. For this reason, it is important to train and create instructions for employees so that the staff can recognize potential social engineering threats.

Internal security and GDPR courses are one solution for making people aware of personal data processing principles and security threats. Internal audits and possibly external help assists to check if e.g. an organization's processes and policies have been done properly and employees follow the instructions. In addition, should the legislation change, it might cause changes to organization's processes and policies.

When an idea of new software feature arises, it is important to pay attention to what kind of data is to be processed. In addition, attention should be paid to what kind of security controls must be implemented to ensure that legislation requirements have

been taken into account properly. It is important to ensure that a software does not process or store information which does not have a legal base (e.g. special categories of personal information). In a sprint planning meeting when a software development team starts planning tasks of a feature, there must be definitions of business logic, what kind of data and what security controls must be implemented. Then a team will design its tasks from a technical point of view as for how the feature will be implemented, tested and released during a sprint. The team cannot commit to the feature in a sprint if there are ambiguities with the definitions. The sprint planning meeting is approximately four percent of a whole workload for a sprint. In a two-week sprint, the planning meeting is about four hours long. The team cannot design badly described features in the sprint planning meeting. If they do so, they probably cannot get anything designed in four hours.

General Data Protection Regulation is very a topical subject in many organizations even if it still leaves much to interpretation. During the GDPR project, many important GDPR features have been implemented and the organization has a better understanding now of how the GDPR affects the daily work. The most important result of this thesis are the steps that have been taken to get the example application GDPR compliant. In addition, a common GDPR model of what kind of features software developers should design and implement e.g. for applications to make them to meet GDPR requirements. It requires also further studying how to keep matters GDPR compliant in the future. In this thesis, a few ideas have been opened as for which matters still need a deeper analysis. The first phase of the GDPR project has ended; yet, Headpower continuously develops its software products and instructions to better meet the requirements of GDPR.

The corporation's GDPR project began in the end of year 2017. An initial list of GDPR features for the example application were planned in the beginning of 2018. As can be seen in Figure 17, the project's phases took about less than six months. The reporting of the GDPR project from the thesis perspective started after the project results had been released into production. Personally, it felt easier to do things in this order, because after the project as a writer I had a deeper perspective to research problems.

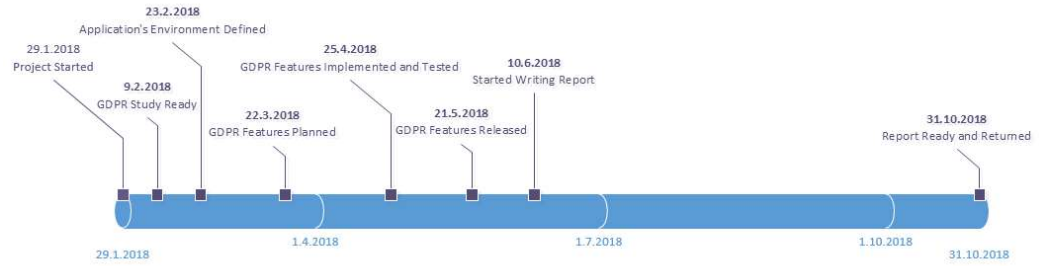


Figure 17. Timeline of the GDPR project

References

- ARTICLE 29 DATA PROTECTION WORKING PARTY. Directorate General Justice. 2017. PDF document published by Directorate General Justice 3 October 2017. Accessed 17 March 2018. Retrieved from http://ec.europa.eu/newsroom/document.cfm?doc_id=47741
- Biscoe, C. 2017. Data mapping: Where to start for GDPR compliance. Accessed 6 June 2018. Retrieved from <https://www.itgovernance.co.uk/blog/data-mapping-where-to-start-for-gdpr-compliance/>
- Curtis, J. 2018. What is GDPR? Everything you need to know post-compliance deadline. Accessed 11 May 2018. Retrieved from <http://www.itpro.co.uk/it-legislation/27814/what-is-gdpr-everything-you-need-to-know>
- DeRose, J. 2018. How Do Internal Audits Work? 27 April 2018. Accessed 1 July 2018. Retrieved from <https://www.ispartnersllc.com/blog/how-do-internal-audits-work/>
- Do 72 Hours Really Matter? Data Breach Notifications in EU GDPR. 2018. Trend Micro. Accessed 13 May 2018. Retrieved from <https://www.trendmicro.com/vinfo/us/security/news/online-privacy/do-72-hours-really-matter-data-breach-notifications-in-eu-gdpr>
- DR, GDPR vai DDR: Mitä pitää oikeasti tehdä [DR, GDPR or DDR: What you really need to do]? 2018. Moonsoft Oy. PDF document published by Moonsoft Oy 1/2018. Accessed 26 June 2018. Retrieved from https://www.moonsoft.fi/materials/asiakaslehti_01_2018.pdf
- Duncan, E. 2018. What is GDPR in a nutshell? Accessed on 4 June 2018. Retrieved from <https://www.ftadviser.com/regulation/2018/04/12/what-is-gdpr-in-a-nutshell/>
- EU General Data Protection Regulation (GDPR) Overview. Vigilant software. Accessed 4 June 2018. Retrieved from <https://www.vigilantsoftware.co.uk/topic/eu-gdpr>
- EU-tietosuojan kokonaisuudistus [A reform of EU data protection]. Valtiovarainministeriö [The Ministry of Finance]. 2016. PDF document published by The Ministry of Finance 1/2016. Accessed 7 July 2018. Retrieved from <http://urn.fi/URN:ISBN:978-952-251-778-4>
- General Data Protection Regulation (GDPR) requirements, deadlines and facts. Nadeau, M. 2018. Accessed 26 June 2018. Retrieved from <https://www.csoonline.com/article/3202771/data-protection/general-data-protection-regulation-gdpr-requirements-deadlines-and-facts.html>
- Goodchild, J. Hulme, G. 2017. What is social engineering? How criminals take advantage of human behavior. Accessed 30 June 2018. Retrieved from <https://www.csoonline.com/article/2124681/social-engineering/what-is-social-engineering.html>
- Gunathunga, S. 2017. Individual's rights under GDPR. Accessed 6 July 2018. Retrieved from <https://medium.com/@sagarag/individuals-rights-under-gdpr-3256fb3f356c>
- Hart, C. 2001. Doing a Literature Search. First edition. SAGE Publications Ltd.

Hart, C. 2018. Doing a Literature Review. Releasing the Research Imagination. 2nd Edition. SAGE Publications Ltd.

Headpower is the sum of its customers. Headpower Oy. Accessed 7 May 2018. Retrieved from <https://headpower.com/>

Hopping, C. Millman, R. 2018. HTTP vs HTTPS: what difference does it make to security? Accessed 20 June 2018. Retrieved from <http://www.itpro.co.uk/network-internet/30416/http-vs-https-what-difference-does-it-make-to-security>

Kloeten, O. 2009. Soft-deletes are bad, m'kay? Accessed 18 June 2018. Retrieved from <https://weblogs.asp.net/fbouma/soft-deletes-are-bad-m-kay>

Kylmänen, A. 2018. General Data Protection Regulation - Requirement Analysis of Customer Personal Data: Case Study. Master's thesis, university. Tampere University of Technology. Degree programme in industrial and information management. Accessed 25 January 2019. Retrieved from <http://URN.fi/URN:NBN:fi:tty-201808292218>

Lehtisalo, I. 2018. GDPR-Six Months After the D-Day. Master's thesis, polytechnic. Haaga-Helia University of applied sciences. Degree programme in information systems management. Accessed 28 January 2019. Retrieved from <http://urn.fi/URN:NBN:fi:amk-2018112718374>

Lemminki, R. 2018. Sahaammeko omaa oksaamme liian tiukalla GDPR:n ja ePrivacyn tulkinnalla [Do we cut our own branches with too strict an interpretation of GDPR and ePrivacy]? Accessed 26 June 2018. Retrieved from https://www.marmai.fi/blogit/mainostajien_blogi/sahaammeko-omaa-oksaamme-liian-tiukalla-gdpr-n-ja-eprivacyn-tulkinnalla-6704012

Lokiohje [Log guidelines]. Valtiovarainministeriö [The Ministry of Finance]. 2009. PDF document published by The Ministry of Finance. Accessed 24 June 2018. Retrieved from https://www.vahtiohje.fi/c/document_library/get_file?uuid=775179cb-6c54-4dfb-b65d-e925d47c61d2&groupId=10229

Loshin, P. 2017. The GDPR right to be forgotten: Don't forget it. Accessed 13 May 2018. Retrieved from <http://searchsecurity.techtarget.com/feature/The-GDPR-right-to-be-forgotten-Dont-forget-it>

Malste, M. 2017. Muutossuunnitelman laadinta EU:n tietosuoja-asetuksen pohjalta : Case: Kotimainen monialakonserni [Drafting change plan based on EU Data Protection Regulation. Case: Domestic multi-industry corporation]. Thesis, polytechnic. Jyväskylä University of applied sciences, Degree programme of information technology. Accessed 13 May 2018. Retrieved from <http://urn.fi/URN:NBN:fi:amk-201802162519>

Masson, D. 2017. Adapting to regulation: How to cope when government changes the rules. Accessed 29 June 2018. Retrieved from <https://www.theglobeandmail.com/report-on-business/careers/leadership-lab/adapting-to-regulation-how-to-cope-when-government-changes-the-rules/article34862926/>

Mast, J. 2018. SAP authorization concept renewal project and GDPR in company X. Master's thesis, polytechnic. Turku University of applied sciences. Degree

- programme in international business. Accessed 28 January 2019. Retrieved from <http://urn.fi/URN:NBN:fi:amk-2018060713056>
- Mononen, M. 2019. GDPR-Strategy Management at a SAP Organization. Master's thesis, polytechnic. Karelia University of applied sciences. Degree programme in technology competence management. Accessed 29 January 2019. Retrieved from <http://urn.fi/URN:NBN:fi:amk-201901281616>
- Pedro, B. 2017. What are Web APIs. Accessed 10 June 2018. Retrieved from <https://hackernoon.com/what-are-web-apis-c74053fa4072>
- Personal data breaches. OFFICE OF THE DATA PROTECTION OMBUDSMAN. Accessed 24 June 2018. Retrieved from <https://tietosuoja.fi/en/personal-data-breaches>
- Pinal, D. 2010. SQL SERVER – Soft Delete – IsDelete Column – Your Opinion. Accessed 18 June 2018. Retrieved from <https://blog.sqlauthority.com/2010/09/03/sql-server-soft-delete-isdelete-column-your-opinion/>
- Pulkkinen, T. 2018. Cloud outsourcing guidelines and data protection regulation in Europe : Context of online banking self-service channels. Master's thesis, polytechnic. Jyväskylä University of applied sciences. Degree programme in cyber security. Accessed 26 January 2019. Retrieved from <http://urn.fi/URN:NBN:fi:amk-201805107482>
- Regulation (EU) 2016/679 of the European parliament and of the council. 2016. Official Journal of the European Union 4 May 2016. Accessed 25 February 2018. Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=FI>
- SQL Server Agent. Microsoft. 2017. Technical document in docs.microsoft.com page 19 January 2017. Accessed 7 June 2018. Retrieved from <https://docs.microsoft.com/en-us/sql/ssms/agent/sql-server-agent?view=sql-server-2017>
- Stored Procedures (Database Engine). Microsoft. 2017. Technical document in docs.microsoft.com page 14 March 2017. Accessed 4 June 2018. Retrieved from <https://docs.microsoft.com/en-us/sql/relational-databases/stored-procedures/stored-procedures-database-engine?view=sql-server-2017>
- Survey Reveals Biggest GDPR Compliance Risks are Breach Notification, Data Mapping, Managing Consent, and Data Transfer. 2017. TrustArc. Accessed 13 May 2018. Retrieved from <https://www.prnewswire.com/news-releases/survey-reveals-biggest-gdpr-compliance-risks-are-breach-notification-data-mapping-managing-consent-and-data-transfer-300551549.html>
- The Six Privacy Principles of GDPR. 2017. MTHREE Consulting. Accessed 12 May 2018. Retrieved from <https://www.mthreeconsulting.com/blog/2017/04/the-6-privacy-principles-of-gdpr>
- Veerraju, A. Srinivasa, R. Murali, G. 2010. Refactoring and Its Benefits. Accessed 5 June 2018. Retrieved from <https://aip.scitation.org/doi/abs/10.1063/1.3516393?journalCode=apc>

What is SCRUM? Scrum.org. 2018. Accessed 25 February 2018. Retrieved from <https://www.scrum.org/resources/what-is-scrum>