Universiteit Leiden

ICT in Business and the Public Sector

Behaviour models of consultants and developers about privacy-by-design and data minimisation in the ERP system

Name: Alicia Pang
Student-no: s2785412

Date: 07/04/2022

1st supervisor: O. Gadyatskaya
2nd supervisor: K. Labunets

MASTER'S THESIS

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Acknowledgements

I would like to begin this thesis by acknowledging everyone who helped and supported me during the thesis writing process. First of all, I would like to thank my first supervisor, Olga Gadyatskaya for her time, useful feedback, advice, and support during the thesis period. I would also want to thank Kate Labunets, my second supervisor, for her time and for giving me useful insights and feedback. I really appreciate your feedback and I learned a lot.

In addition, I would like to thank all my colleagues at my internship company for their time, guidance, and ideas in helping me with my thesis research. Especially my company supervisor, Kartik Joshi, for his time, help and support in these strange and difficult times.

Lastly, I would like to thank my family and friends for their motivation, support and listening to my thesis.

Abstract

Technology makes it easier for people to share their personal data. Companies are also sharing data within the company so that different departments can work faster through ERP systems. However, protecting personal data has become extremely important to protect individuals. Therefore, the General Data Protection Regulation (GDPR) has established principles to ensure that the personal data of individuals is protected. One of the biggest challenges is minimising data. Many companies find it difficult to get this right, yet it is the foundation for effective privacy measures. Integrating privacy into the structure of the organisation is one way to change employee behaviour. Privacy-by-design (PbD) is addressed as a core strategy for the company. However, there are still some challenges to overcome when implementing PbD. The aim of this study is to investigate the behaviour of developers and consultants in relation to privacy-by-design and data minimisation in ERP systems.

In addition to existing literature studies, qualitative research is conducted to investigate the behaviour of developers and consultants. Sixteen interviews were conducted using semi-structured interviews. A grounded theory approach is used for inspiration in analysing the qualitative dataset. The BJ Fogg behavioural model uses motivation, triggers and the ability to perform a behaviour to outline the behaviour.

Using the literature review and interviews, we outlined the behaviour of developers and consultants in relation to privacy-by-design and data minimisation in the ERP system. Behaviour is motivated by compliance with the GDPR and the company's privacy policies. But customer and business requirements also trigger the behaviour, and some obstacles affect the developers' and consultants' abilities.

Based on the research, we can say that some aspects influence the behaviour of developers and consultants, e.g. obstacles, GDPR and lack of experience with privacy-by-design or data minimisation in the ERP system. However, there is no difference between developers' and consultants' behaviour. On the contrary, both are triggered and motivated by the same aspects. Therefore, recommendations such as privacy awareness can improve ability and eventually behaviour.

Table of Context

List of Figures

List of Tables

1. Introduction

In the last two decades, technology has made enormous strides. Thanks to increasing technological development, companies can collect, share and shop large amounts of (personal) data. However, technological developments also provide new opportunities for cyber-attacks, which can lead to insecurities such as identity theft or data breaches. As a result, protecting the privacy of personal data has become extremely important to protect individuals (McKinsey Company, 2020).

1.1. Problem statement

The definition of privacy is dynamic, changing over time and often influenced by political and technological features of the social environment (Marc Pelteret, 2019). As newspapers posed a threat by publishing photos without people's consent, the definition of privacy used to be seen as the right to *"be left alone"*. Today, privacy is associated with personal data and information technology is seen as a risk (Samuel D. Warren, 1990).

In modern society, people want privacy, but also to share personal information to access services such as insurance and healthcare, and to make friends. Technology makes it easier for people to share personal data with each other, but it is a challenge to control personal data after it has been shared. Companies are also sharing data within the company so that different departments can work faster. An Enterprise Resource Planning (ERP) system helps companies integrate data and information across multiple departments (Hasan, 2018). ERP stores information about customers and employees in a database (Nishad Nawaz, 2013). However, the data in the ERP system needs to be protected from third-party frauds, such as phishing scams. Therefore, data protection is an essential factor in the system as it ensures that personal data is not lost and prevents privacy threats (Kelly D. Martin, 2016).

To better protect (personal) data, the EU has introduced the General Data Protection Regulation (GDPR). However, not all companies are aware of the importance of data protection and the challenges that the GDPR brings. The Association for Intelligent Information Management (AIIM), the world's largest organisation in the field of information management, conducted a survey on understanding the GDPR. More than 800 IT and business professionals responsible for data protection in companies with European clients participated in this survey. 30% of companies know little or nothing about the GDPR (AIIM, 2017).

Furthermore, most data protection officers are not prepared for the challenges of the GDPR (AIIM, 2017). One of the biggest challenges is minimising data. Today, every company needs to carefully consider how personal data is processed (e.g. collection, updating and retention) and regularly review its policies to make adjustments where necessary, just as the GDPR requires. Initially, only essential personal data should be collected and it should not be kept longer than is necessary for the reason for which it was collected. Many companies find this difficult to get right, but it is the basis for effective data protection measures (Dode, 2018).

Over the years, there have been many major data breach incidents that have cost businesses millions of dollars. A study by Intel Security found that 43% of internal employees are responsible for data leaks, and half of these leaks are unintentional. It should be the people who are the problem, not the cybercriminals (McAfee, 2017). Approximately 81 per cent of data breaches by hackers are caused by stolen or weak passwords (Verizon, 2020). In addition to these human aspects, breaches are frequently caused by unpatched software or IT managers who fail to follow best practices. In addition to these human aspects, breaches are often caused by unpatched software or IT managers who do not follow best practises.

Integrating privacy into the structure of the organisation is one way to change people's behaviour. As a result, privacy-by-design (PbD) is treated as a core strategy for the entire organisation (Cavoukian, 2011). However, there are still some challenges to overcome when implementing PbD in the business in today's world. The privacy-by-design concept has been characterised as "*vague*," leaving many unanswered concerns relating to how to apply it in system design (Jeroen van Rest, 2012). As a result, it is challenging for managers to protect the privacy-by-design concept (Spiekmann, 2012). Researchers and engineers often tend to associate the privacy-by-design concept with certain privacy-enhancing technologies (PETS). On the other hand, privacy-by-design cannot be reduced to a set of rules or the use of a specific technology. According to European Union Agency for Cybersecurity (ENISA), it is a process that involves a variety of technological and organisational measures that enforce privacy and data protection principles through the use of appropriate and adequate technical and organisational methods, including PETS (ENISA, 2022). ENISA is the European Union's management is implemented with providing a high standard of cybersecurity all over the EU (ENISA, 2022).

1.2. Research Objectives

It has emerged that the employees population does not yet have a clear understanding of what the concepts of privacy-by-design and data minimisation mean in the ERP system. On the other hand, the existing literature focuses on the risks and consequences for the company if privacy concepts such as privacy-by-design and data minimisation are not correctly integrated into the ERP system, and not on what developers and consultants understand by these privacy concepts. There is also little information in the literature about which strategies and techniques are suitable for implementing privacy-by-design and data minimisation in ERP systems. In addition, behavioural models are used in the literature to represent certain behaviours such as a person's motivation or ability to perform something. In order to understand how developers and consultants take action regarding privacy protection in ERP systems, this much-needed research aims to find out what the behaviour models are from the developers' and consultants' regarding privacy-by-design and data minimisation in ERP systems.

1.3. Research Questions

A research question has been developed in order to answer the problem specified in section 1.1, 'Problem Statement', and to achieve the research objectives.

The main question of this research is:

'What are the developers' and consultants' behaviour models for privacy-by-design and data minimisation in the ERP systems?'

The following sub-questions have been formulated in order to answer this main question:

1. What is privacy-by-design and data minimisation?
2. How do developers' and consultants' understand privacy-by-design and data minimisation in the ERP system?
3. What privacy-by-design and data minimisation strategies and techniques do developers and consultants use in the ERP systems?
4. What would help the developers and consultants to increase understanding about privacy-by-design and data minimisation techniques in the ERP system?
5. What are the differences between developers' and consultants' behaviour models of privacy-by-design and data minimisation in the ERP systems?

1.4. Research Scope

This thesis was written as part of an internship with the Microsoft team at KPMG in the Netherlands. The data protection legislation (GDPR) used in this research is only applicable in Europe. Therefore, this research focuses on European organisations working with the ERP system, where personal data of customers is collected, stored and shared. Furthermore, this research focuses on two principles of the GDPR: data minimisation and privacy-by-design. The other five principles of the GDPR are not considered in this research. In addition, only the behaviour models of developers, consultants and managers are examined and not end-users of the ERP system. Managers and consultants are often confused with each other, as they both advise clients. It is possible to evolve into the position of a manager as a consultant, which is why most managers started their work as consultants. The only difference between a manager and a consultant is that a manager can manage a project independently, whereas a consultant is supported by a (senior) manager. Managers and consultants are therefore considered the same function in this study.

1.5. Thesis Outline

This section presents the outline of the dissertation with the aim of showing the process of answering the research question.

This thesis begins with a literature review in order to answer the first sub-question and to lay the groundwork for answering the other sub-questions. After the literature review described in Chapter 2, Chapter 3 describes the research methodology. The qualitative research method is then described from the perspective of theory. Chapter 4 then presents the results of the data collection methods. Next, Chapter 5 presents the behavioural model. Finally, chapters 6 and 7 finish the thesis with a discussion and conclusion.

2. Literature review

This chapter is divided into three sections. The first section of this chapter discusses the GDPR, precisely what it means, and data minimization and privacy-by-design principles. The second section of this chapter discusses the ERP system, data types, and privacy. The last section discusses the literature studies on behavioural models.

2.1GDPR and related principles

Definition

The General Data Protection Regulation (GDPR), also known as Regulation (EU) 2016/679, is a privacy law introduced by the European Parliament. The GDPR is a set of regulations aimed to give European citizens more control over their personal information (Baxevani, 2019). It replaced the outdated Data Protection Directive (DPD) 95/46/EU- introduced in 1995 - as this regulation was no longer up-to-date with the technological changes taking place in the world. "*The development of technology today is faster than the adaptation of human thinking".* (Jan-Kyrre Berg Olsen, 2009). The new regulation was adopted in April 2016 and followed a transition period of two years, after which the regulation was applied in May 2018.

One of the most key characteristics of the GDPR is that it applies to all businesses that use personal data, regardless of where they are based. The GDPR law applies to any company that handles sensitive or personal data of European citizens (Tankard, 2016). Even if personal data is processed outside of the EU, the data controller or data processor is liable to the GDPR's regulation (Europa.eu, 2021). A data controller or data processor is someone who is responsible for the processing of personal data within an organisation. The GDPR protects the personal data of EU residents from collection, processing and use by companies (Sposit, 2018).

Both small and large businesses must comply with the GDPR or they will be subject to heavy fines. Article 83 of the GDPR provides for fines that are flexible and can be adapted to the company. Serious violations can be punished with a fine of up to 20 million euros. Less serious infringements can be punished with a fine of 2% of the company's annual worldwide turnover in the previous business year. This is a financial incentive for companies to comply with the GDPR. Many companies hire third parties to process their data in order to avoid fines and comply with the GDPR. In addition, the appointment of a Data Protection Officer (DPO) is mandatory, who is responsible for maintaining documentation and processes within an

organisation (Europa.eu, 2017). DPOs are responsible for monitoring and implementing a company's data protection strategy to ensure compliance with the GDPR regulations. The DPO's tasks are set out in Articles 33 and 34 of the GDPR (InterConsulting, 2018):

- Maintain compliance;
- Offer guidance on Data Protection Impact Assessments (DPIA)
- Collaborate with supervisory authorities
- Inform and advise on data protection responsibilities

The DPO is not personally responsible for the organisation's compliance with the GDPR (Šidlauskas, 2021). Compliance must always be demonstrated by the organisation itself. If there is a data breach affecting a user's personal data occurs, the data controller must notify the supervisory authority within 72 hours (Paul De Herta, 2012). If the breach poses a high risk to people's rights and freedoms, the company is obliged to inform people (Daniel Mikkelsen, 2019).

Principles

In order to protect personal information, especially personal information collected by companies or through the internet, the GDPR establishes seven principles that must be followed. Article 5 of the GDPR is the foundation for these principles (InterConsulting, 2018). ERP system processes personal data and therefore must comply with these principles in order to comply with the GDPR. The seven principles are showed in table 1 (InterConsulting, 2018):

Principles	Explanation
1. *Lawfulness, fairness and transparency*	Personal data must be processed lawfully, fairly, and transparently by companies.
2. *Purpose Limitation*	Companies should only collect data if they have a defined, disclosed, and valid motive for doing so. The purpose should be stated in a clear and straightforward way to the user.
3. *Data Minimisation*	Companies can collect and analyse personal data, but they must do so in a way that is sufficient, relevant, and restricted to what is required for the purpose.

4.	*Accuracy*	Personal data must be "*accurate and, when applicable, up to date"*. Companies must guarantee that old and outdated relationships are not maintained and that incorrect personal data is erased as soon as possible.
5.	*Storage Limitations*	Companies must delete users' personal information when they are no longer using it and it is not suitable for their purposes.
6.	*Integrity and confidentiality (security)*	Companies must protect personal data in a secure way, including against unauthorized processing and data leakage, destruction, or corruption. Encryption of data should be seen the core of data security.
7.	*Accountability*	This is a principle that demands companies to implement suitable technological and organisational measures and to be able to show their efficiency when asked.

Table 1 Seven GDPR Principles

These GDPR principles are mandatory for companies and employees to follow. However, The GDPR principles are difficult to implement in these circumstances because they cannot be done in a traditional, *'intuitive'* way. Processing procedures need to be reconsidered and reformed, sometimes significantly, with new actors and tasks defined and technology playing a crucial role as a guarantee aspect. Effective technological and organisational measures and controls must be established and incorporated into the processing (ENISA, 2022). Therefore, two GDPR principles are discussed more in detail in this research.

2.1.1 Data Minimisation

Large companies store a lot of personal data such as the name, address and location of the customers in their system. This personal data can be stored in the Enterprise Resource Platform (ERP) system (Subhi R. M. Zeebaree, 2020). However, personal data needs to be protected to ensure the privacy of the customers. Therefore, the GDPR has established various principles such as minimisation of data to ensure user privacy.

Definition

Data minimisation (DM) is a straightforward and easy privacy principle that advises minimising the use of personal data in software systems (Nalin Asanka Gamagedara Arachchilage, 2018). Although it appears to be straightforward, it is not in practice. For software developers who are constantly collecting data from users to provide value to the business, data minimisation is a major challenge. If a marketing company collects a lot of personal data, it can identify its customers in order to promote certain products more effectively. As a result, system developers are less likely to use DM, which could lead to privacy concerns. Cambridge Analytica (Hu, 2020) was able to obtain data from 50 million Facebook users due to Facebook's recent data breach. According to Hu (2020), software developers have tried to achieve data minimisation by focusing their system design on storage and sharing aspects.

On the other hand, it is challenging for developers to minimise the amount of data used in storage and sharing in their system designs. This is because implementing DM into a system design can be complicated and the considerations of system developers are not always aligned (Stefan Schiffner, 2018). In addition, software developers are not attuned to the privacy risks posed by the collected data and users' privacy concerns. Therefore, it is challenging for software developers to create a system that uses as little data as possible. It is important that developers understand the data involved (Colonna, 2013).

When it comes to data minimisation, the focus should be on data availability, openness, and accuracy, especially as it relates to data sharing and storage (Onno Tene, 2011). Furthermore, a better understanding of the information collected would lower the cost and problems associated with storing and protecting large amounts of data in software systems.. As a result, strict standards for the use of user data in system designs have been developed, requiring that data be used as little as possible in system designs to protect software privacy (European Union data protection, 2016).

Data Minimisation Methodology

Senerath *et al.* (2019) has developed a methodology for minimising data in software systems. The methodology consists of two principles:

1. Making data comprehensible from the user's point of view.
2. Extending data sharing and storage in a system and reducing the use of data beyond the data collecting phase.

This methodology consists of two steps. The first step is to understand the data, which entails comprehending the data's sensitivity and visibility. In order to understand the data, a decision must first be made regarding the data. Once the decision is made to collect data, the developers must first understand the data before designing the system. At this point, the developers have a good idea of what data they'll use in the system, as well as the system's context and an overview. However, adjustments or improvements can still be made during the design phase. The second step of this methodology is to make system design decisions to minimise the use of data in the system.

In order to understand data or the sensitivity of data from the user's perspective, the value of data in relation to the system, and define the visibility of data in the system design, an empirical model is developed. The methodology is used to assess the privacy risk of data from the user's perspective. Sensitivity, visibility and context are the three categories into which they can be placed. The scale is used to illustrate Figure 1. Software developers need to scale the data they use based on the parameters in order to use this model.

Scale	Sensitivity	Visibility	Relatedness
1	Category S1 : Highly sensitive data elements, loss of data would impose serious damage to the privacy of the data owner	Category V1 : Highly visible, similar to publicly posted content in Facebook, anyone can access without the knowledge of the data owner	Category R1 : Extremely related data the application cannot do without. For example, the location information for a tracking application
2	Category S2 : Sensitive elements, loss would impose considerable damage to the privacy of the data owner	Category V2 : Relatively visible, similar to *friends only* content in Faceboook, a limited set of users access the content without the knowledge of the data owner	Category R2 : Related data that provide features that add significant value to the application. For example, the location information for a restaurant finder
3	Category S3 : Low sensitive elements, loss would impose limited, calculable and bearable damage to the privacy of the data owner	Category V3 : Not visible, similar to the *only me* content in Facebook, no one can access the data without the knowledge of the data owner	Category R3 : Remotely related to the purpose and provide optional features in the application. For example, location information for a trip planner

Figure 1 Data weights in each category (Awanthika Senarath N. A., 2019)

Implementation of Data Minimisation

When implementing data minimisation, system designers use different phases of data use in their system design. According to Kneuper (2019), most implementations of data minimisation focused on data storage rather than data collection or data sharing. It was also found that most system designers did not focus on all three phases of data use in the system. However, it was noted that data minimisation is not only about the data points collected, but also about the links that organisations make between data. Consequently, comprehensive data minimisation should focus on the entire data processing chain, including data collection, capture, storage, modification, linkage and access in a system. But not all software developers adhere to this (Yang Wang, 2009).

According to the study by Senerath *et al.* (2019), just understanding data serves no purpose. Unless it has an impact on the design decisions that developers make when developing software systems. Research by Senerath *et al.* (2019) has highlighted a number of design decisions that developers should consider when implementing data minimisation in systems. Figure 2 shows the GDPR principles and indicates the phase at which data should be stored in the system.

Legal Requirement (Article 5 of Regulation (EU) 2016/679) European Union data protection (2016)	**Phase of data in the system**
It must be collected for explicit and legitimate purposes and used accordingly	Data Collection and Processing, Purpose Limitation
It must be adequate, relevant and not excessive in relation to the purposes for which it is collected and/or further processed	Data Collection and Processing, Purpose Limitation
It must be accurate, and updated where necessary; Data controllers must ensure that data subjects can rectify, remove or block incorrect data about themselves	Data Storage (accurate, access and control)
Personal Data must be processed lawfully and fairly	Transparency in data collection, storage and sharing
Data that identifies individuals (personal data) must not be kept any longer than strictly necessary	Data retention (Storage period)
Data controllers must protect personal data against accidental or unlawful destruction, loss, alteration and disclosure, particularly when processing involves data transmission over networks. They shall implement the **appropriate** security measures	Data Sharing and Storage (security of data)
These protection measures must ensure a level of protection **appropriate** to the data.	Accountability, Security

Figure 2 Data protection GDPR requirements (Awanthika Senarath N. A., 2019)

Data Minimisation Techniques

Anonymisation and pseudonymisation are two well-known strategies frequently utilized to perform data minimisation in practice, according to ENISA (2022). The GDPR defines pseudonymisation as a technology that can help with data protection by improving the design and security of personal data processing. It is a common misconception that pseudonymised data is the same as anonymised data.

However, there are other examples of methods/techniques for implementing data minimisation in software systems such as encryption, cryptography, and identity management. Only none of the offered methods describe or specify which technologies are suitable for software system design as well as how much data minimisation should be performed by developers. However, the percentage of developers who identified the implementation techniques for data minimisation in system design was not satisfactory (Nalin Asanka Gamagedara Arachchilage, 2018). Developers prefer technical instructions to concepts that guide them, according to Senarath *et al.* (2018). In addition, Oetzel *et al.* (2013) found that developers are unaware of how techniques like anonymisation may be used to implement data minimisation in a system. The reason for this is that data minimisation only asks developers to use data wisely and does not provide guidance on how to do so. Because they are not measurable, developers are unable to interpret statements like "*anonymise data if required*" and "*minimise the use of unnecessary data".* This leads to uncertainty in data protection practises, making it difficult for developers to comply. Therefore, if developers are given concrete instructions on how to use data as little as possible and still meet system requirements, they will be able to better implement data minimisation in their system designs (Nalin Asanka Gamagedara Arachchilage, 2018).

2.1.2 Privacy-by-design

Definition

In recent years, the term "*privacy-by-design (PbD)"* has gained popularity. It is similar to the term *"data protection by design,"* although the two terms are used equally in the European Commission's proposal. Data protection through technical design is another phrase for privacy-by-design. However, there is still confusion about what "*privacy-by-design*" means and how to implement this concept can be implemented in practice (InterConsulting, 2018).

According to Spiekmann (2012) PbD refers to a technical and strategic management approach that commits to choosing and implementing governance controls to reduce the privacy risks of information systems. According to Cavoukian (2011), on the other hand, the term PbD implies privacy must be considered throughout the design process, from the earliest phases until system operation. Although this is a good approach, there is a lack of ways to integrate data protection into the system development processes (George Danezis, 2014). Therefore, under Article 23 of the General Data Protection Regulation, the European Commission proposed that "*privacy-by-design*" be defined as a set of principles. These principles can be used right from the start of a system's development to avoid privacy concerns and ensure data protection compliance.

For many companies, implementing privacy-by-design is a serious issue. A major challenge for PbD is the involvement of senior management in the privacy strategy. According to research by Spiekmann (2012), the key to a successful corporate privacy policy is the active involvement of management. This is due to the fact that many business models own personal data. Furthermore, managers find the definition of privacy-by-design to be a vague concept, and they are still not clear about what it entails, which makes it difficult to protect (Jeroen van Rest, 2012). In addition, little is known about the benefits of protecting privacy in companies and the risks involved (Kathrin Bednara, 2018).

Privacy Impact Assessment (PIA)

According to Spiekmann (2012), there are no generally accepted methods for systematically building data protection into systems. Therefore, the privacy impact assessment approach was developed to overcome this (Marie Caroline Oetzel, 2014). PIA provides clear data protection objectives and sets out the means to achieve them, and it is also referred as a *"milestone towards privacy-by-design"* (The European Data Protection Supervisor, 2018). PIA provides early warning information that can be used to implement corrective actions (Amir Shayan Ahmadian, 2018). Organisations can benefit from privacy (and data protection) impact assessments as an *"early detection"*. It is a useful tool to inform management of any threats and help them make wise decisions to avoid privacy disasters (Marie Caroline Oetzel, 2014). PIA also intends to improve data quality and increase customer, employee, and consumer trust in how personal data is processed and privacy is protected (NOREA, 2015). Making PIAs mandatory for system developers could be an important step towards privacy-by-design and improve compliance with European and US data protection regulations (Spiekmann, 2012).

PIA is considered as a part of an organisation's risk management approach, according to Alshammari *et al.* (2018). The goal of a privacy impact assessment is to discover and resolve privacy issues, not just to ensure that a project complies with regulations (Wright, 2012). A PIA is also not the same as an audit. An audit is used to ensure that the PIA has been carried out correctly and that recommendations have not been followed.

If an organisation collects personal data and this processing involves a high-privacy risk, the GDPR requires it to conduct a PIA. This indicates that the processing may lead to a personal data breach. If an organisation processes sensitive personal data on a large scale or characterises individuals or monitors individuals in a public place (e.g. through cameras), it must in any case carry out a PIA (Clarke, 2009).

A PIA should be considered as a process. A process that should start in the early planning stages of a project and continue throughout its life cycle. As the project progresses, new risks may emerge. According to Wright (2012), this is the most common PIA process.

1. Identifying the PIA team and establishing a scope statement
2. Determining if a PIA is required (threshold analysis)
3. Identification of stakeholders and a description of the planning process
4. Analysis of data flows and other privacy implications
5. Interaction with stakeholders
6. Identification of risks and potential solutions
7. Formulation of suggestions
8. Report preparation and publishing
9. Suggestions implementation
10. Refreshing the PIA if the project changes
11. Third-party review and/or audit of the PIA

Design Strategies for Privacy-by-Design

As stated previously, "*Privacy-by-design*" is a vague and complicated concept that can lead to privacy-invading system design. Hoepman (2014) has developed eight privacy design strategies as a result. These are classified into two categories.

1. **Data-oriented strategies** - this category is more technical in nature and focuses on privacy-friendly data processing. This category contains four main methods.

1.1. *Minimise* - The quantity of personal data handled should be kept to a minimum level.

1.2. *Separate* - Personal data should be processed in a dispersed manner, in separate parts, wherever appropriate.

1.3. *Aggregate* - Personal data must be processed at the most accurate level possible.

1.4. *Hide* - Guarantee that personal information will be kept confidential and is not shared or revealed to the public.

2. **Process-oriented strategies** - this category is mainly concerned with the organisational factors of the processes around the responsible management of personal data.

2.1. *Inform* – When personal data is processed, data subjects must be properly informed.

2.2. *Control* - Data subjects should have control about how their personal data is processed. This strategy works in tandem with the *"inform"* strategy.

2.3. *Enforce* – A privacy policy that complies with legal requirements should be in existence, and these requirements should be followed. This strategy ensures the existence of a privacy policy.

2.4. *Demonstrate* - Demonstrate that personal data processing is carried out in a privacy-friendly manner.

	Purpose limitation	Data minimisation	Data quality	Transparency	Data subject rights	The right to be forgotten	Adequate protection	Data portability	Data breach notification	(Provable) Compliance
MINIMISE	o	+								
HIDE		+					o			
SEPARATE	o						o			
AGGREGATE	o	+								
INFORM				+	+				+	
CONTROL			o		+			+		
ENFORCE	+		+			+	+			o
DEMONSTRATE										+

Legend: +: covers principle to a large extent. o: covers principle to some extent.

Figure 3 Privacy-by-design strategies (Hoepman, 2014)

Hoepman (2014) has incorporated the aforementioned strategies into Figure 3. A privacy-by-design strategy cannot cover every legal data protection principle since it has no impact on that principle. Particular strategies only cover a part of some data protection principles, such as

purpose limitation. Organisational and procedural measures are also required to fully achieve purpose limitations.

Implementation privacy-by-design strategies

Traditionally, the development of a system is a cyclical process. The definition, design, development, deployment, operation, and evaluation phases are all part of the system life cycle. Existing techniques, such as design patterns, are mainly useful for the design and development phases, thus the privacy design strategies were developed. In addition, the privacy strategies have been developed with the aim of achieving specific (technical) goals in order to increase the privacy of the whole system - provided that the goals are achieved. It is important that all project stakeholders, including end users, are involved. The end users are the ones who will end up interacting with the data, so the design process needs to consider both the project stakeholders and the end users.

Organisations should not focus on just one strategy, according to Coleksy (2016). Although all of the strategies are obviously useful, apply them all at the same time to make the system as privacy-friendly as possible. Depending on the content of the system, certain strategies may be more efficient and appropriate than others. The processing of personal data, on the other hand, must be taken into account. Organisations need to assess whether each of the strategies (and also several techniques) is relevant to the processing of personal data.

2.2 Enterprise Resource Planning (ERP)

This section discusses the ERP system in detail. It starts with an introduction to the ERP system and the meaning of this system. Then, the data in the ERP system and privacy will be discussed.

The amount of data on the internet and in systems has increased significantly in recent years. To cope with the changing competitive environment of businesses, investments are being made in information technology (IT). Enterprise Resource Planning (ERP) systems are one of the technologies that have become indispensable for businesses. ERP systems help businesses to meet the rising expectations by providing accurate, fast and integrated information that helps businesses to make better decisions (P. Trott, 2011). Furthermore, the ERP system is a business information system meant to manage all the resources, information, and relevant works for entire business operations, according to Coe (2011). This system consists of a single database

and data software packages. The software has a feature that enables all departments to work together to manage the company's assets, see Figure 4.

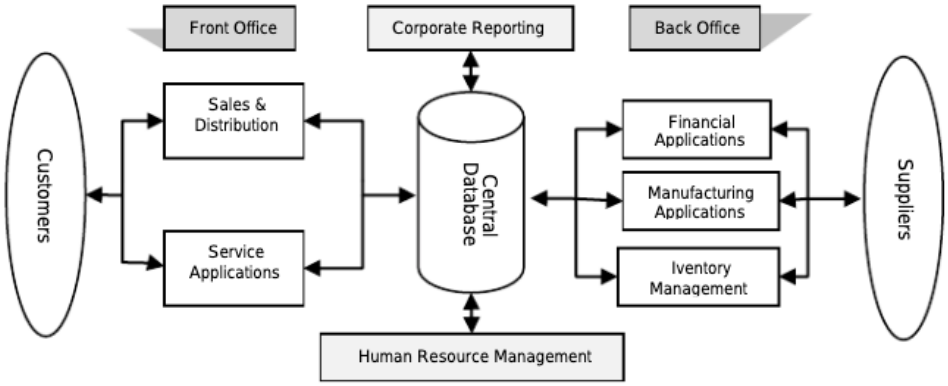

Figure 4 ERP system concept (Andrianto, 2019)

Already in 2000, more than 60% of Fortune 500 organisations have implemented an ERP system according to a study by Steward (2000). ERP is being implemented in a growing number of small and medium-sized businesses, as well as in large corporations (Hasan M. T., 2018).

ERP systems are available from different vendors, with Microsoft, SAP, and Oracle being the most popular. In 2009, these top players in this business software market accounted for more than 75% of total revenue (IMAP, 2010). Users in the United States, on the other hand, choose SAP because it enables them to make better decisions by providing company-wide information (Kakoli Bandyopadhyay, 2012). The most common ERP systems integrate modules such as production planning, purchasing, supply chain, inventory management, human resources, accounting, marketing, and finance. Integrating these modules into a single system is a requirement for an ERP system whose full potential lies in Big Data, data-driven strategic decision-making, mitigating risk, rapid reports, and performance monitoring.

On-premise ERP and hosted ERP are the two typical types of ERP systems. With an on-premise ERP, the system is completely internal. The company's own servers and computers run the software and data. Despite the advantages such as security and having your own equipment, there are also disadvantages such as investment in equipment and licencing, as well as maintenance fees (Mezghani, 2019). If it is a hosted ERP solution, the system may be hosted on a remote server outside the region. Most of the time, the services are only available through the direct network. This is also referred to as cloud-based ERP or Software as a Service (Saas) ERP and is a new model of ERP that is similar to traditional on-premise ERP in terms of systems, functions and solutions (Björn Johansson, 2013).

The SaaS ERP model is accessible through the use of an internet application, and data is structured and managed by the cloud service provider before being given access to the client for a monthly fee. A cloud-based ERP system, on the other hand, offers a more dynamic approach to hosting ERP system. Accessibility, availability, affordability and scalability are the main advantages of cloud computing, all guaranteed by Service Level Agreements (SLA) (Angela Lina, 2012).

Data in the ERP system

As mentioned earlier, there is a central database with all the company's information. However, the data in the ERP system can be interpreted from two perspectives (Sebastian Wieczorek, 2008):

1. *Business view on ERP data.* In this business view, the data is separated into master data and transactional data.
    1.1. The term "*master data*" refers to data that is static and remains valid over time. For instance, supplier information such as name and address, or product information such as product size or description. This data is rarely changed and can be automatically put into various transactions.
    1.2. Transactional data is information with a short duration. It's only used for one transaction at a time and can always be linked to master data. Transactional data includes, for example, information regarding the number of items or the delivery period for a certain order (Cong, 2010).
2. *Technical view on ERP data*. From this perspective, the difference between master data and transaction data is less important. The difference between user-generated and automatically derived transaction data is more relevant from a technological point of view. Transaction data can be derived automatically, for example, from the current date or from a previous transaction, in addition to being created by user input. The amount of a product in a sales order can, for example, be determined by the customer's request. As a result, a technical distinction is made between system data and input data, see Figure 5. Furthermore, the figure shows how system data is used as internal data. This information is stored in a database that can be accessed directly from the application. There is no external access and the system data includes both master and transaction data. All data that must be provided externally by users or external components during execution and cannot be obtained automatically is referred to as input data. Master data or transaction data can be used as input data (Sebastian Wieczorek, 2008).

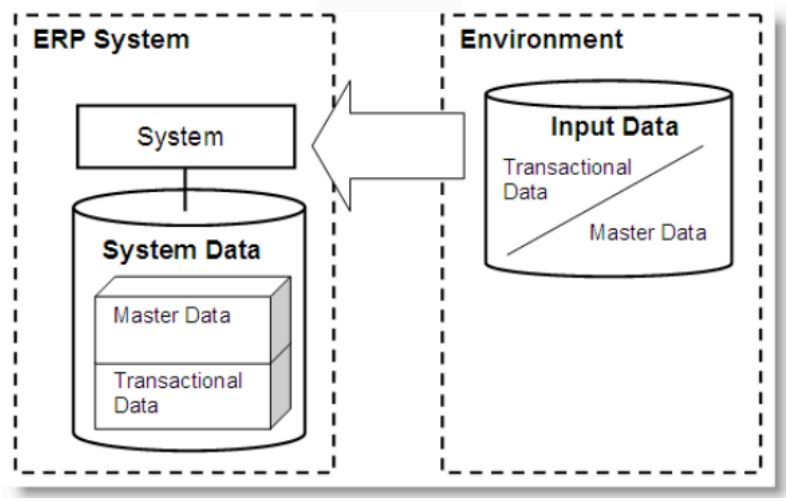

Figure 5 Overview data in the ERP system (Sebastian Wieczorek, 2008)

Privacy in the ERP system

A software solution such as an ERP system that collects, saves and analyses personal data, e.g. customer data, obviously needs to comply with data protection laws, including the GDPR *"right to be forgotten"* regulation. One of the biggest concerns facing IT managers is whether it is possible to identify and delete personal data within a certain timeframe and whether these activities could slow down or even disable some of the back-end functionality of the ERP system (Eugenia Politou, 2018).

Personal data is often not directly linked to the user ID of a database, making it difficult to find it among hundreds of tables in the ERP system. It can take a lot of time and effort to find personal data in systems of this scale and complexity (Samantha Mathara Arachchi, 2015). In addition, the GDPR must be taken into account, which imposes explicit and strict time constraints on data controllers when it comes to responding to a user's request to delete their personal data. This can create a variety of challenges in the ERP system. These challenges can be complicated by the fact that ERP backup plans can differ significantly in terms of the methodology used, such as cloud infrastructure and hard copies (Eugenia Politou, 2018).

Modern ERPs currently offer their customers features that help them comply with the GDPR in several areas. SAP, for example, offers five products that help companies comply with GDPR requirements: SAP Information Lifecycle Management, SAP Data Services, SAP Information Steward, SAP Process Control, and SAP Access Control (Eugenia Politou, 2018). There are even more tools for GDPR compliance, which can be divided into the following categories based on their main functions:

- A tool to discover where personal data is located in the ERP system. Even though these tools can find personal data on existing systems, it's unclear if they can find personal data already backed up.
- These tools should have unique logging algorithms that maintain check of data backup in their real-time databases to achieve this.
- Personal data masking tools. Masking personal data in ERP systems could be a feasible alternative to data deletion. On the other hand, many tools for masking personal data do not make such adjustments to production systems.
- Control and analysis access to the stored personal data tools. These tools only apply to run-time copies of data, so it's unclear how they'll work with backup data.

2.3 Behaviour Model

In this section we discuss the literature on behavioural models. First, an introduction to behaviour modelling based on the literature. Then we discuss the BJ Fogg behavioural model and how it is used in research.

Definition

Over the past 70 years, several theories of behaviour change have been put forward to help develop interventions to promote good habits and minimise harmful behaviours. Theories of behaviour change attempt to explain why people's behaviour changes. Environmental, personality and behavioural characteristics are identified in these theories as the most important elements in influencing behaviour. In recent years, there has been increasing interest in extending these theories to other fields, such as education, criminology and computer science, in order to improve the services offered in these areas through a better understanding of behaviour change. Many researchers have recently distinguished between behavioural models and theories of change (Andrea Carlson Gielen, 2003). Behavioural models differ from theories of change in that they are more descriptive and better understand the psychological elements that explain or predict a certain behaviour. On the other hand, theories of change are more process-oriented and typically attempt to change a specific behaviour (Darnton, 2008).

In an attempt to explain behaviour change, each behavioural change theory or model focuses on different aspects. The social cognitive theory, theories of reasoned action, the transtheoretical model of behaviour change, the health action process model and the BJ Fogg Behaviour Model (FBM) are among the most common (Fawad Taj, 2019). On the other hand,

the COM-B model is also a behavioural model. The COM-B behavioural model is used to determine what needs to change for a behaviour change intervention to be successful. The model consists of three components: Capability (C), Opportunity (O), and Motivation (M). When it comes to intervention methods, the COM -B model is particularly relevant because interveners need to ensure that the learned behaviour is maintained. It has been used primarily in the field of health care (Susan Michie, 2011). Therefore, the BJ Fogg Behaviour model is used to investigate privacy behaviour in this study.

BJ Fogg Behaviour Model

BJ Fogg developed the behavioural model for understanding human behaviour in 2009 (Fogg, 2009). Fogg claims its usefulness in the development and research of persuasive technologies. The FBM can help behaviour change experts in various areas, including education and security. It provides a framework for researchers to consider the factors influencing behaviour change. This model is both simple and powerful simultaneously, and these features make it suitable for this research.

According to the FBM (Fogg, 2009), human behaviour is the result of three factors: motivation, ability and trigger. In short, behaviour occurs when someone wants to do something (motivation), is able to do it easily (ability), and something drives the action (trigger).

Figure 6 shows Fogg's behavioural model. There is a vertical axis for motivation. A person with high motivation is on the high axis, while someone with low motivation is on the low axis. The ability axis is on the horizontal axis. A person with high ability is on the right of the horizontal axis, while a person with low ability, who has difficulty performing, for example, is on the left. In the middle of the model is "Prompts". According to Fogg (2009), prompts are known by many different names, such as trigger, call to action and request. A trigger is anything that causes the individual to act in a certain way at that moment (Fogg, 2009). The action line is either above or below the trigger. The trigger is successful if it is above the action line and it fails if it is below the action line.

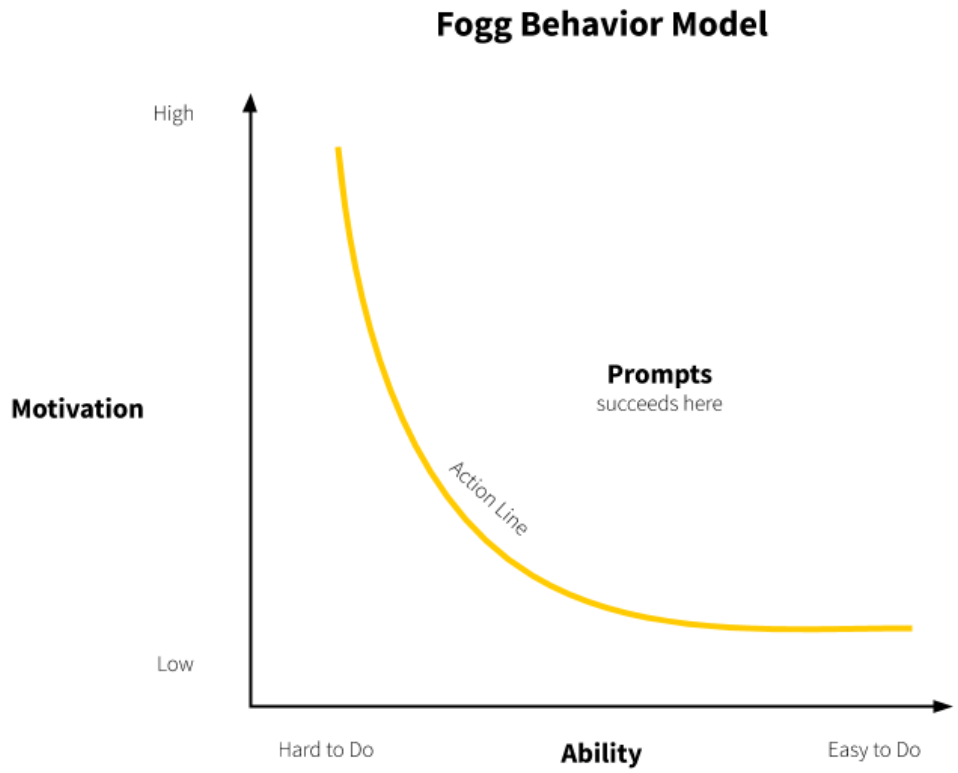

Figure 6 Behaviour model BJ Fogg (John, 2022)

- ***Motivation***

Motivation is a concept that is commonly used in a variety of fields. According to Qemajl Sejdija (2016), motivation is the process of starting, developing, and maintaining goal-oriented behaviour. Motivation encompasses the physiological, emotional, interpersonal, and cognitive factors that influence behaviour. Fogg (2009) developed a framework for motivation that includes three basic motivators, each with two sides, to clarify this concept clear in Fogg's behavioural model.

1. *Pleasure/ Pain*

What sets this motivator apart from the others is that it produces a quick, or almost immediate, outcome. There isn't a lot of planning or anticipation going on. People are reacting to what is happening right now.

2. *Hope/ Fear*

The expectation of a result defines this dimension. The expectation of something positive happening is known as hope. Fear is the expectation of something negative, most commonly failure. As proven by everyday behaviour, this dimension can be more significant than pleasure/pain at times. For example, when people update virus protection settings, they are driven by fear according to Fogg (2009).

3. *Social acceptance/ Rejection*

Much of our social behaviour is influenced by this dimension, from the way we dress to the words we use. People are clearly driven to do activities that will get them social acceptability. People are driven to avoid being socially rejected, sometimes even more profoundly.

- ***Ability***

Increasing ability in the real-world design does not mean teaching people new skills or training them to develop. As this involves work, people are often reluctant to teach and train. According to Fogg (2009), people are naturally lazy. Consequently, products that require users to develop new skills tend to fail. Therefore, designers of persuasive experiences need to make the activity simple in order to improve the user's skills. In other words, the design of persuasive experiences depends mainly on the power of simplicity. A typical example is Amazon's 1-click purchasing. People purchase more because it is simple to do so. Simplicity has an effect on behaviour. Fogg (2009) has created a framework consisting of six components and an understanding of how they interact. These five components interact with each other like links in a chain: if one component breaks, the whole chain fails, see Figure 7. Simplicity is lost in this approach.

1. *Time*

When a target behaviour demands time but the person does not have it, the behaviour isn't straightforward. For instance, if they need to fill out a 100-field form online, such behaviour might not be natural for them because they normally have other responsibilities.

2. *Money*

A target behaviour that costs money is difficult for those with low finances. That link in the chain of simplicity can probably break.

3. *Physical effort*

A behaviour that takes low physical effort and stress is more capable than one that needs a lot of physical work and stress.

4. *Mental effort (brain cycles)*

It may not be straightforward to do a target behaviour if it requires everyone to think hard.

5. *Routine*

When it comes to routine activities that people repeat again and over, they tend to find them simple.

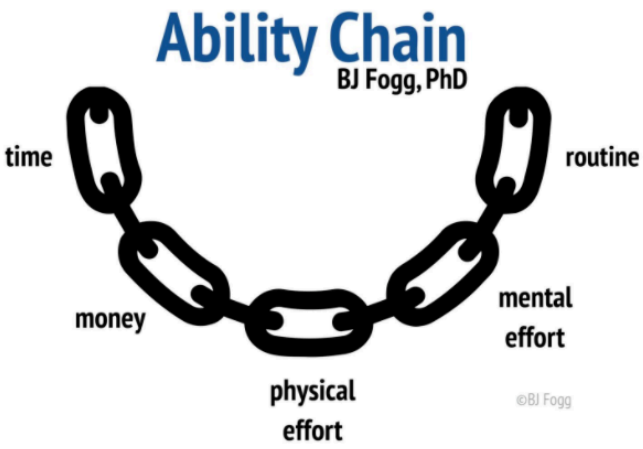

Figure 7 Ability Chain (Fogg, 2009)

- ***Triggers***

As mentioned earlier, there are various names for the concept of 'triggers', such as prompt, call to action or call. A trigger is anything that causes the individual to behave in a certain way at that moment (Fogg, 2009). There are also extrinsic and intrinsic triggers (Susanne Kießling, 2021). Intrinsic triggers are triggered by a person's own desire or motivation. Extrinsic triggers come from the outside world (signals), technology (email) or other people (meetings). Fogg (2009) describes the following three types of triggers:

1. *Spark as trigger*

When there is high ability but low motivation, this trigger is used. A motivating component should be included in the trigger's design. This is intended to emphasize the benefits of performing the behaviour. It's a collection of motivational messages. Showing the benefits of subscribing to a newsletter, for example.

2. *Facilitator as trigger*

When there is high motivation but low ability, this trigger is used. It aims to make the work easier. An good facilitator assures users that the desired behaviour is simple to do and does not necessitate the use of a resource that they do not have at the time. Software upgrades, for example, often use facilitators to get compliance by implying that the task can be completed with just one click.

3. *Signal as trigger*

When both motivation and ability are high, a trigger is used. This is only a prompt to remind you to do something. A traffic light that turns red or green is a common example of a signal. The traffic light isn't attempting to motivate people. It's just signalling when a certain behaviour is appropriate.

Scope

To outline the behaviour of the developers and consultants, the concept of FBM was used based on the literature review. We define the concepts of the FBM for this research as follows:

- **Motivation**

Motivation can be described as how much is the person driven to do something or against it. There must be a reason why a specific behaviour is being performed. A question can be asked whether it is motivation like am I willing to do this?

- For example, when an interviewee mentions GDPR compliance. GDPR compliance gives them the motivation/drives to implement privacy into the system.

'As I mentioned, every year, we sign up for GDPR acknowledgement, where we make sure that we follow all the EU protocols for that, and we don't go by it.' - Developer 4

- **Ability**

Ability refers to how easy or hard it is for someone to perform something at a particular time. There are some resources that can be used to determine whether or not someone has the ability to perform a behaviour. A question that can be asked about the ability is "Can I do this?" Fogg (2009) has defined five elements for ability, namely time, money, physical and mental effort and routine.

- For example, when an interviewee mentions that they have no knowledge about data minimisation. Because of a lack of knowledge, implementing data minimisation in the system will be difficult.

'No, I haven't seen people lecture about data minimisation much.' – Manager 4

- Another example of ability is when an interviewee mentions privacy out of the box. By privacy out of the box makes the ability of the interviewee easier because they do not have to develop anything themselves anymore.

'No, unfortunately, I have to disappoint you. I think the system does a lot for us, so it relieves us. Everything has already been described within Microsoft. It is not a complete ERP solution. That is why little attention is paid to what is necessary. But the fact no.' - Developer 3

- **Triggers**

A trigger is basically an event to start the process. What triggers the respondent to perform a behaviour? There are different triggers that have been described in the literature research such as social, forced and proactive triggers (Sauvik Das L. A., 2019).

  - For example, when an interviewee mentions business or customer requirements. They must have a requirement in order to develop something in the system; otherwise, they will be unable to go forward with their process. This can be a forced trigger because they have to implement this requirement regardless of what the interviewee thinks.

> *'Yes, business requirements. We store requested data, and we also ask whether they really can not do without it. But if a specific process is based on that, I can't get around it. There is no other choice.' - Developer 2*

2.4 Related Literature Research

This section, discusses the related literature research on behavioural models and privacy problems perception. Next, we discuss the security and privacy behaviour related to the BJ Fogg Model. Finally, we discuss the related literature research regarding perceptions of developers' and consultants' privacy problems.

Security and Privacy Behaviour

Many studies have used the FBM to explore end-users' security and privacy behaviour. However, to the best of our knowledge, no research has been done on identifying consultant and developer behavioural models related to ERP system privacy. One of the study that come close to our research is Sauvik Das *et al.* (2014). People tend not to follow the recommended security advice, according to research by Sauvik Das *et al.* (2014). There exists a variety of psychological theories that explains the lack of motivation. These includes:

- Strict security measures are sometimes hostile to the end-users' intended goal. Using additional password authentication when end-users log into the system, for example, can result in more time and effort (Serge Egelman, 2010).
- Experts also recommend against the usage of security measures, which lowers end-users' motivation to feel safe (Sauvik Das H. J., 2014).

The ability of end-users to behave in a security and privacy-compliant manner is limited by lack of awareness and knowledge (Sauvik Das H. J., 2014). Many end-users are not aware of the risks and do not know what they can do to protect themselves. According to Anne Adams *et al.* (1999), security tools are often too complicated for end-users. In addition, there is a gap between what users want and understanding how to achieve it.

In contrast to motivation and ability, security and privacy behaviour triggers are less well researched. There are social, forced and proactive triggers that reveal security and privacy behaviour, according to another study by Sauvik Das *et al.* (2019). Social triggers are direct social interactions that cause behaviour change. Examples of social triggers include observing others, sharing access with others, and seeking advice from others. Word-of-mouth conversations was also found to be a social triggers for security and privacy behaviours in this study (Sauvik Das L. A., 2019). There are also non-social and external triggers, also known as enforced triggers, which make the end-user change their behaviour against their will. An example of a forced trigger is when an end-user learns about a leak of their personal data, or an employer requires them to constantly update their passwords. Finally, there is a non-social 'proactive' trigger, which includes internal procedures such as screen locks or routine password changes that lead to a purposeful shifting behaviour. Sauvik Das *et al.* (2019) found that social triggers were the most mentioned behavioural triggers.

Finally, Sauvik Das *et al.* (2019) discovered that security and privacy triggers change according to social economic status, age and level of security behavioural intention (SBI). Security and privacy triggers are specific to each person, hence, the security and privacy behaviours keep constantly changing (Paula Braveman, 2014). In particular, people with low and medium safety behavioural intention were more likely to report changing their behaviour in response to social triggers. In contrast, people with a higher SBI were significantly more likely to report changes in their behaviour.

Sauvik Das *et al.* (2019) study focused more on the triggers of end-users in privacy and security behaviour in general. This study, however, is more focused on the behaviour of developers and consultants related to privacy in the software systems.

In addition, the research by Liljestrans et al. (2019) used the reason action approach (RAA) to help understand and adapt the user's mental model in the context of computer security. The Reasoned Action Approach (RAA) model is commonly used as a mental models of human

behaviour in general and covers the belief oriented aspect of an individual's behaviour. Liljestrans *et al.* (2019) proposed a mental model using the following components:

- *Skills and abilities*

The constraints on a person's ability to influence their behaviour are determined entirely by their skills and abilities. Even if someone has the right motivation to change their behaviour, they will be unable to do so due to a lack of skill. Someone who has the appropriate skills and abilities to behave in an ideal manner but lacks the motivation to do so will perform poorly and fail to act appropriately effectively.

- *Behavioural control*

Behavioural control is the technique through which individuals influence their own desires. There are two aspects of this behavioural control: the response efficacy and perceived self-efficacy.

- *Severity & Vulnerability, Fear, and Attitude*

The attitude is the most essential aspect of this. Attitude influences how pleasant or sour a behaviour is. Severity and Vulnerability are terms that indicate a person's view of the severity of a danger and their sense of their vulnerability to that concern. Finally, fear is a crucial part of this model since it is a constant motivator. Fear motivates through controlling a person's attitude as well as person's behaviour. Fear affects behavioural control and conversely since self-efficacy is a big part of behavioural control. If a person has a high level of self-efficacy, they are less bothered by fear, and conversely.

A user study was undertaken to assess the validity of each of the model's primary components. The findings of the user study revealed that the suggested mental model was successful, with each cognitive model's key elements changing user behaviour.

The study of Liljestrans *et al.* (2019) mainly focused on the mental models of end-users in computer security context. While this study focus on the behavioural aspects such as motivation, ability and triggers. Furthermore, this study objectives is focuses on the developers and managers behaviour related to privacy in software systems and not the end-users in computer security.

Privacy problems perceptions

Through a series of case studies, Culnan *et al.* (2009) have found the importance of an accountability infrastructure for IT organisations to successfully resolve privacy trust breaches

successfully. They emphasise that security and privacy are two different concepts and that protecting personal data stored on the Internet is not sufficient to protect users' privacy. Furthermore, Sheth *et al.* (2014) analysed and compared the views of developers and users on privacy. They showed that developers believe that anonymising data is more successful than privacy laws and practises in minimising privacy concerns. They even identified considerable gaps in user and developer privacy perceptions. For example, developers are more prepared than users to give up privacy in return for more customized or better system functionality.

According to Hadar *et al.* (2018) most developers view privacy from a data security perspective and focus on technical and security solutions (user control and access, encryption and anonymisation). Developers' understanding with security solutions rather than solutions for other privacy-related issues and developers' personal choice for privacy policies solutions rather than data protection solutions suggest that developers lack the understanding needed to develop privacy-friendly advanced technologies. Hadar *et al.* (2018) also found that the developers' work environment, namely the company's privacy culture, influences their privacy perceptions and beliefs. Despite the risk of future reputational damage, organisational culture in certain companies enables and encourages behaviour that is at odds with official, established policies or rules. However, in some companies, the organisational culture supports the privacy policy, e.g. through monitoring, communication and instruction procedures that ensure that employees are aware of and understand the policy.

To the best of our knowledge, there is limited literature research about the privacy perceptions of the managers or consultants. However, we found a study that is close to this research. Sandra Henderson *et al.* (1999) study described the implications of the information systems (IS) managers on personal information privacy. Managers need to be aware of any inherent privacy issues and be prepared to take appropriate actions and necessary steps to protect the privacy of individuals. Therefore, Sandra Henderson *et al.* (1999) developed a normative model that companies might adopt if they are concerned about privacy concerns and are willing to take action to protect individuals' privacy rights. However, the study by Sandra Henderson *et al.* (1999) focuses more on the actions the IS manager needs to take if there is a privacy concern within the system but not how the managers' perception of privacy in the system.

3. Research Approach

This chapter discusses qualitative research and the reason why we use this methodology. Furthermore, this chapter also explains the sample and methods of the data collection and analysis.

3.1 Research Design

The aim of this research is to find out the behaviour of developers and consultants about privacy-by-design and data minimisation in the ERP system. Although there is enough literature on privacy-by-design and data minimisation, there is still a lack of clarity among developers, consultants and managers about what exactly the privacy concepts mean and how to implement they can be implemented (Spiekmann, 2012). Through qualitative research, it is possible to understand what developers and consultants do in terms of privacy-by-design and data minimisation in the ERP system and how to solve the understanding.

Qualitative research involves collecting, organising and analysing textual material that comes from conversations or discussions. Qualitative research methods are designed to help researchers learn more about people and their social and cultural contexts (Cathryne Palmer, 2006). Data can be obtained through qualitative research methods in the form of written or spoken words rather than in quantitative form. This qualitative data can be collected and transformed into written text for analysis through interviews, observations, and focus groups (Juliet Corbin, 1990).

3.2 Sample

In this research two groups will be interviewed, the developers who develop the ERP system and the consultants and managers who actually support the customer and help to set up the ERP system. With this population, it will be possible to find out what the two views are on understanding privacy-by-design and data minimisation in the ERP system. The respondents are qualified for this research because of their position, experience with the ERP systems and education.

3.3 Data Collection

As mentioned earlier, interviews are conducted to collect data. According to qualitative research, it is also most common to conduct interviews to collect data (Juliet Corbin, 1990). There are several forms for conducting an interview:

- *Structured interview*

In a structured interview, the interviewer asks the respondents the same questions in the same way. This entails using a well-structured question schedule, which is comparable to a questionnaire and is frequently used in quantitative data analysis. Not only the questions, but also the probable answers, are predetermined (Nigel Mathers, 2000).

- *Unstructured interview*

Since there is no structure to the interview, all topics can be discussed in an unstructured interview. The interviewer goes through a few subjects and offers follow-up questions based on the interviewee's prior response (Nigel Mathers, 2000).

- *Semi-structured interview*

Semi-structured interviews are similar to structured interviews such as they include pre-planned topics or questions. Semi-structured interviews, on the other hand, use open-ended questions rather than closed questions. When a large amount of data is being collected or little is known about a topic, is it very useful (Nigel Mathers, 2000).

For this study, a semi-structured interview was used, as it allows us to obtain information from respondents in an open and honest way. Respondents were able to express their opinions/thoughts on certain topics through open-ended questions, which can lead to new information. To ensure that an interview goes well, an interview protocol has been prepared. See **Appendix I** for the interview protocol in English and the Dutch version.

We sent an e-mail to the interviewee asking when they would be available for an interview. An invitation to the interview and a consent form were sent to the respondent as soon as they indicated a time when they were available for an interview. The consent form can be found in the **Appendix II**. The consent form contains further information regarding the interviewer's study, the dissertation topic, and the respondent's consent to participate in the interview, including the recording of the interview. The interviewee had to send their consent by email. Nevertheless, not all respondents replied to the consent form. However, before the interview questions were asked, we requested the permission to record the interview, regardless of whether they approved the form.

The interviews were recorded and transcribed. The recorded interviews were not kept for a longer period of time for reasons of confidentiality of the information. In addition, individual interviews were conducted to gain a deeper understanding of the individual interviewee compared to a group interview and to ensure that the interviewee expressed his or her own views. Due to government regulations COVID-19 all interviews were conducted online via Microsoft Teams.

In addition, the study was approved by the Faculty of Science Ethics Committee of University Leiden. To protect the data in this study, all names and emails were password protected. The other data (interview transcripts and transcript analysis) were anonymised. For each interview, the interviewee was assigned a ID and a link "name email and ID" was placed in a secure folder. The audio files and transcripts were also stored in a password-protected folder on SURFdrive, which was anonymised by ID.

3.4Data Analysis

To analyse the data for this research, Strauss and Glasser's (1967) Grounded Theory is used as inspiration. Strauss and Glasser define grounded theory as the theory that emerges from the data that is systematically collected and analysed during the research process. Grounded theory is about the collection and analysis of data. However, Punch (2014) explained that grounded theory is not a theory, but rather a method, an approach, a strategy. It is a research strategy whose aim is to develop a theory from the data collected. The meaning of the term "*grounded*" is that the theory is developed based on the data, and the term "*theory*" means that the aim of collecting and analysing the research data is to develop a theory.

Coding is one of the most important aspects of the grounded theory method (Holton, 2010). Coding is the process of classifying and arranging qualitative data to identify different ideas and their relationships. The grounded theory approach is used in this research to develop a theory by interpreting and understanding the differences in the data collected through interviews (Ylona Chun Tie, 2019). In order to follow the grounded theory approach, Strauss and Glass proceed in several steps.

- **Open coding**

The technique of breaking down data into different units of meaning is called open coding. The basic aim of open coding is to capture and name data, such as diversity or conflict, as shown in

Figure 8. It starts with classifying a large number of different occurrences. To structure more abstract categories, separately classified concepts are collected around a common theme. Coding is *'open'* at this stage of exploring the data and searching for codes. During the open coding process, occurrences or events are labelled and grouped to create categories and attributes through ongoing comparison. However, it is important to mention the use of memos. Memos are notes that a researcher creates after data collection to elaborate on their thoughts about the data and the classified categories (Holton, 2010).

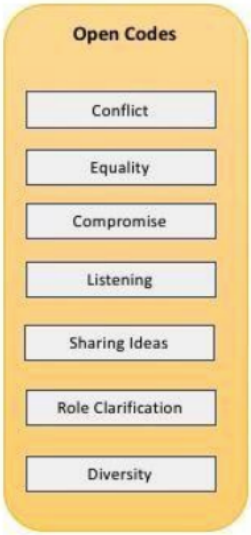

Figure 8 Open coding (Michael Williams, 2019)

- **Axial coding**

Axial coding is the second phase of the Grounded Theory method after open coding. In contrast to open coding, which divides the data into discrete sections, axial coding focuses on the connections made by the codes. It examines how categories and subcategories relate to each other by separating attributes and dimensions of categories. After axial coding, a set of codes can be used to support a set of categories, such as collaboration or communication, as shown in Figure 9. These are the categories that the codes focus on (Cliff W. Scott, 2017).

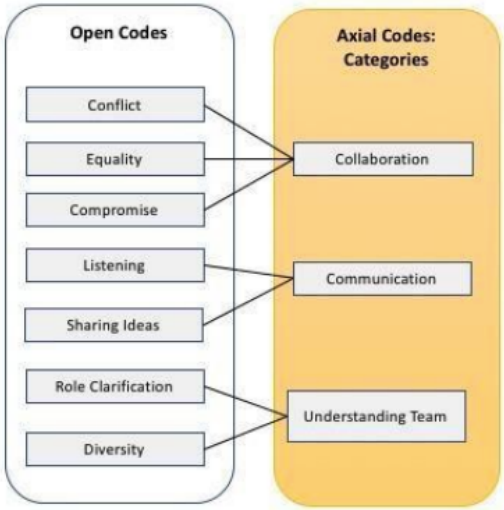

Figure 9 Axial coding (Michael Williams, 2019)

- **Selective coding**

The next stage of Grounded Theory is selective coding, where all categories are linked around a specific core theme. The core theme that is created is based on the qualitative data collected in the earlier phases of Grounded Theory, especially the axial phase. The aim of selective coding, which takes place towards the end of the Grounded Theory process, is either to develop a new theory or to modify an existing one (Maike Vollstedt, 2019). The Figure 10 shows that a core theme emerges from axial codes with categories such as collaboration, communication.

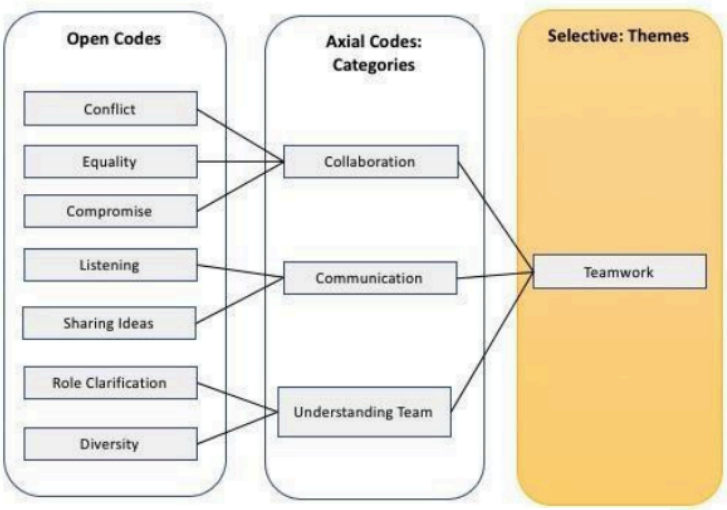

Figure 10 Selective coding (Michael Williams, 2019)

In qualitative research, coding enables the identification, structuring and development of theories. The software program Atlas.ti was used to code the transcripts for this research. Leiden University has a licence for the use of this software program. Most interviews were done in Dutch but were translated into English for analysis purposes.

3.4.1 Inter Coding Agreement (ICA)

Atlas.ti (2020) provides a coder agreement check tool that allows you to evaluate how different coders code a record. To ensure that the codebook is reliable, double coding should be performed. According to Krippendorff (2015), the reliability of the codebook indicates that different people are using the data in the same way. When a theory is based on data, it can be considered valid if it is highly reliable. The family of alpha coefficients includes a range of measurements that can be used for calculations at different levels, see Figure 11 (Krippendorff, 2015). The available coefficients measure the degree of agreement or disagreement between different coders. This metric can be used to predict reliability, but the coefficients do not measure actual reliability.

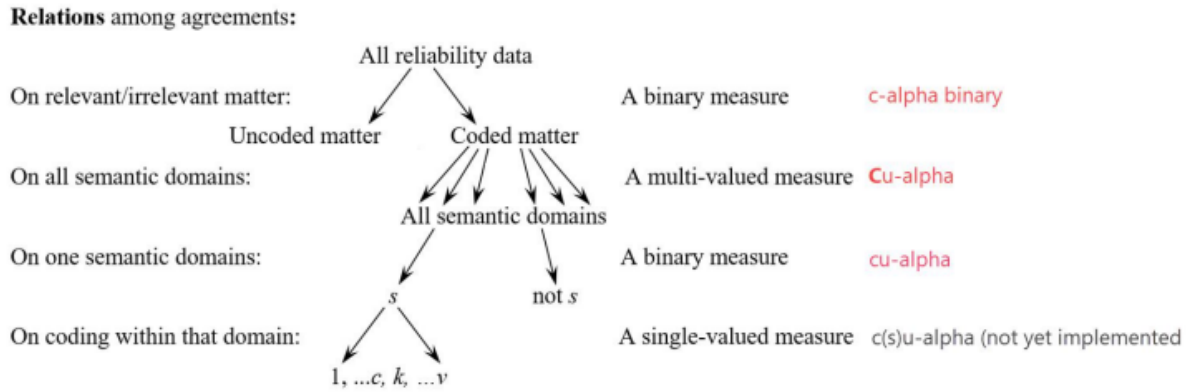

Figure 11 Inter Coding Agreement (Krippendorff, 2015)

- The most basic level is **c-alpha-binary**. Determine if diverse coders see the same areas of the data as interesting to the subjects of interest indicated by codes.
- **Cu-alpha** is the sum of all cu-alpha coefficients. It helps determine the fact that several semantic domains' codes might be applied to the same or overlapped quotations.
- The **cu-alpha** is to see if various coders could distinguish between the semantic domains codes. A semantic domain means a collection of separate ideas with similar meanings.
- **CSU-alpha** is not implemented yet. Krippendorff (2015) has not explained why CSU-alpha is not implemented. Once implemented, it allows users to dive down a step deeper and determine whether code inside each semantic domain performs well or not. It shows the agreement on coding inside a semantic domain.

When determining whether to accept or reject coded data, the ICA coefficient must be taken into account. As a result, Krippendorf (2015) (Atlas.ti, 2020) states the following:

- Data with a reliability value of less than 0.667 is unreliable.
- If the coefficient is more than 0.8, a semantic domain is considered reliable.

To ensure that the codebook was reliable and that the coefficient was higher than 0.8, double coding was performed. In the first round, the c-alpha could not be calculated because of problems with merge projects. Instead of merging the codes, Atlas.ti started duplicating the coding and the interviews, which resulted in the c-alpha value not being correct. In the second round, there was another problem with the citations. One of the researchers had a different citation, which resulted in an incorrect c-alpha. Despite the difficulties, we managed to calculate the first round of Krippendorff's alpha between two independent coders, which is 0.944 (see Figure 12). In the second round, Krippendorff's alpha was 0.971, see Figure 13. We present the complete codebook in **Appendix III.**

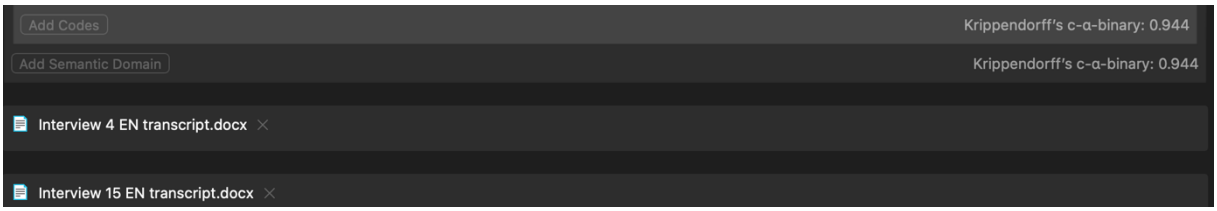

Figure 12 Krippendorff's Alpha (first round)

Figure 13 Krippendorff Alpha (second round)

4. Results

In this chapter, we discuss the results of the research. The first section of this chapter discusses the statistics of the interviews and respondents demographics. Furthermore, this chapter discusses what the developers, consultants and managers understand about GDPR and privacy concepts such as privacy-by-design and data minimisation in the ERP system. Also, the strategies and techniques developers and consultants/managers use to implement privacy concepts in the ERP system are discussed. The last section is about how to improve the understanding of developers and consultants about privacy-by-design and data minimisation in the ERP system. Quotations from the interviews are used to illustrate the results, Dutch interviews have been translated to English.

4.1 Statistics interviews

This study has been conducted from September 2022 till March 2022. The interviews were conducted in November 2021. Of the total sixteen people, nine consultants/managers and seven developers were interviewed. The roles and tasks of the consultants and managers differ from those of the developers. Consultants and managers have more interaction with the client, for example in the form of advice. In addition, managers or consultants are responsible for privacy protection and ensure that only certain people have access to specific data. The only difference between a manager and a consultant is that a manager can manage a project independently, whereas a consultant is supported by a (senior) manager. Finally, the developers are responsible for the implementation, which includes creating specific functionalities based on the customer requirements and solving technological problems.

Due to the COVID -19 circumstances, all interviews were conducted by Microsoft Teams. Seven of the sixteen interviews were conducted in English, the other nine interviews in Dutch. The interviews conducted in Dutch were transcribed and translated into English. The interviews were also coded in English codes. The following table contains the statistical data on the interviews.

	Consultants/managers (n=9)	**Developers (n=7)**	**Total (n=16)**
Average duration (min)	22	26	24
Range (min)	14 - 33	20 - 42	14 – 42

Table 2 Statistics of the interviews

Each respondent has a different background and therefore they also vary in characteristics. Table 3 shows the demographic information of each respondent.

Characteristics	Sub-characteristics	Consultants/managers (n=9)	Developers (n=7)
Educational level	Bachelor degree	3 (33%)	4 (57%)
	Master degree	6 (67%)	3 (43%)
Additional certifications	Privacy (IAPP, CIPP/E)	1 (11%)	-
Educational area	Computer Science	-	2 (29%)
	Electronics and Communication Engineering	1 (11%)	1 (14%)
	Business Information Management	3 (33%)	-
	Applied informatics	-	2 (29%)
	Law	1 (11%)	-
	Production Engineering	1 (11%)	-
	Finance and Control	1 (11%)	-
	Business Process control	1 (11%)	-
	Commercial Economics	1 (11%)	-
	Accounting and Control	-	1(14%)
	Engineering and policy analysis	-	1(14%)
Roles	ERP Developer		3 (43%)
	Technical Architecture		2 (29%)
	Data Engineer		1 (14%)
	Internal Advisor	1 (11%)	
	Legal Council	1 (11%)	
	(Senior) Consultant	2 (22%)	
	(Senior) Manager	5 (56%)	1 (14%)
Work years active	Median	10	9
	Minimum	2	2
	Maximum	31	14

Table 3 Participants' Demographics and Background

4.2 Understanding

In this section, we answer the research sub-question "*How do developers and consultants understand privacy-by-design and data minimisation in the ERP system?"* We start with how developers and consultants understand GDPR and what impact GDPR has on ERP systems. Then we discuss what the developers' and consultants' experiences with ERP systems are and what they know about privacy concepts such as privacy-by-design and data minimisation in ERP systems. In this section, the open and axial codes have emerged. Every section starts with a category, also known as axial coding, supported by open codes that is marked **blue**. In Table 4 of **Appendix IV**, we present an overview of the codebook and the code counts.

- **GDPR and other privacy regulations**

To the question *"what is GDPR?"* most of the developers were able to provide a clear **explanation** of GDPR. Three of the developers give equivalent answers. This is how the developers explained GDPR:

> *'GDPR is a ruling that came into effect in Europe that gives the right to the customers to know what information is being stored.' – Developer 1*

> *'What is GDPR? No, it's pretty tricky. It belongs to the European Union and has several rules about how data should be processed. What is allowed and what is not.' – Developer 3*

> *'GDPR is a general law of how one should handle data. And then again categorising very sensitive data versus sensitive data. And everything related to how you handle the data securely. Data protection is GDPR, in my opinion. So all principles, rules, et cetera regarding data.' – Developer 6*

However, not every consultant and manager was capable of giving a clear GDPR explanation. One of the managers struggled to give a clear explanation of what GDPR entails.

> *'In my own words. I know what it is. But it's quite difficult to explain. GDPR, I think, it's the protection of the private data from being exposed, or something, due to… The private data should not be shared until …. [not clear]. So, yeah, this much I know; I can't explain more.' – Manager 4*

Nevertheless, the majority of managers' and consultants' answers to the question of what GDPR is align. The GDPR is a European regulation that controls how personal information about individuals is processed, according to developers and consultants. GDPR has tightened the regulations and legislation around personal data, ensuring that individuals are protected and that organisations do not misuse it.

> *'Yes, this data law is very important for organisations that are actually using the consumers' or clients' data, and it ensures that people don't misuse that information. I know that the provision is all about, you know, giving, taking, preparing consent for using such personal data, which was not very mandatory before. But this law makes it mandatory.' – Manager 1*

GDPR plays a role in the work of developers, consultants and managers. Some developers stated that they have to develop extra functionalities in the ERP system to ensure the application is GDPR compliant. Another developer stated that they sign up for GDPR acknowledgement every year, following all EU protocols to ensure GDPR compliance. They explained what role the GDPR has in the work:

> *'Yes, it's a very one of the key factors. If one of our developers doesn't abide by GDPR, they might end up losing their job. Or there'll be strict actions taken against them. So the data protection is very key. And we all, as a developer, as a consultant, we are abide by it, and we have to follow it.' – Developer 4*

But another developer stated that GDPR does not directly impact their work but creates more awareness, for example, leaving all data open on the laptop.

> *'On my work? Not directly, but I try to be. I wouldn't call it a direct connection with GDPR. But I try to be careful with the data, for example, working from home so as not to leave the laptop alone are unlocked. Also, on the train, I got his privacy filter. So if I need to work or do something on the train, I can do it without anyone seeing what I have on the screen.' – Developer 7*

One of the consultants stated that GDPR partly plays a role in his work because the awareness is not quite there yet but now the data is considered to be more sensitive following the GDPR introduction they have become more aware.

'Partly yes, I will not say that there is a very strict follow-up and that the awareness is not quite there yet. When this law was not yet in place, employee data and suchlike were handled more senselessly. But now because it's more sensitive.' – Consultant 8

The GDPR has increased privacy awareness. One of the managers stated that companies have grown more awareness as a result of GDPR and that information is not shared as quickly as it once was. Before the GDPR, there was a privacy law, Wbp[1], which had less impact and visibility because of the exemption decisions (how to deal with administration, e.g.), explained the manager. However, the GDPR has created more visibility and awareness.

'Since the introduction of the GDPR, I think that more awareness has been created within companies, but companies are more involved in this. We should not just share or see something. They need to get this right, so I think it made it more accessible.' – Manager 2

'Yes, it brings more awareness to the people. Before the GDPR, the Wbp was, therefore, the Personal Data Protection Act. But it had less impact. Rules also applied there, and only that law was less visible because you all had exemption decisions under the Wbp.' – Manager 3

Despite the fact that GDPR is a mandatory law and that it has created awareness, **GDPR and privacy is not high on the priority list** as one of the developers put it.

'I have other business to do and no time for that. The GDPR will be at the lowest priority on my list. It's not the most exciting thing either, but it's helpful to know. Not an exciting workshop to follow.' – Developer 2

- **ERP systems**

GDPR also plays a role in the ERP system. One developer stated that before GDPR was introduced, there was hardly any talk about privacy or a framework built into the system. But with the introduction of GDPR, ERP systems had to comply with it.

[1] The Dutch Personal Data Protection Act, shortened to Wbp, provides rules to protect the privacy of citizens. The Act came into force on 1 September 2001. It was replaced in May 2016 by the GDPR (Autoriteitpersoonsgegevens, sd)

'It is very reactive in nature, because until GDPR came into the picture, we had nothing explicitly called privacy, specifically, or built in the framework, or people talking about it. And as GDPR came in, it was reactive because this software system had to comply with it, then Microsoft gave the framework with the classified information. That's it.' – Developer 1

Another developer stated that the role of GDPR was to ensure that in the ERP system not everyone gets access to certain data. They only need to know the information they need to know.

'Correct. We follow that standard, we make sure, especially with the ERP, you have customers coming in, and your suppliers coming in, and you have your own employees. Making sure each employee gets only the information he is required to work on. And the customers get to work on a system where they are expected to work off. And even the suppliers – the same thing. They need to know only the information they are required to. For example, in case of updating something, we don't want the information from one supplier letting go to the other supplier. You have a lot of stuff.' – Developer 4

Another finding is that with the introduction of GDPR in the ERP system, the developer also thinks more about privacy. It was always an afterthought and nobody told the developers how important privacy is in the ERP system as noted by one of the participants.

'And for the first eight years or nine years of my career, I've never bothered about privacy. Because nobody told me that it was important in ERP. I didn't know this. You sent me your thesis. I did not start thinking about privacy in the ERP, because it is always an afterthought. GDPR is a mandate.' – Developer 1

- **Privacy enablers**

Due to the role of GDPR in the ERP system, there are **privacy and data protection certifications**, as noted by one of the managers. Everything in the system is protected by the privacy certificates:

'Microsoft themselves give the system with full privacy and data protection certificate. So that comes with the tool. Within the system, for accessing data, we have the security

and role authorizations, which allow only certain people to access certain data sets within a particular form within the system. But overall, everything is protected with data privacy certificates' – Manager 4

As a result, the managers do not have to think about GDPR in the ERP system. Another finding is that they assumed the ERP system is compliant with privacy laws.

'The role of GDPR, like I said, the ERP system is already compliant. So you don't have to be mindful of the GDPR part.' – Manager 1

One of the developers stated that privacy is now becoming increasingly "**out of the box**" in the ERP system. This means that privacy is already implemented in the ERP system by the vendors. When a developer looks at the existing ERP system packages, the majority are GDPR-compliant.

'Yes, in the past, but now it is more and more added out of the box. So one way you have to add it and you go and look at the existing packages. [..] When you talk about ERP packages, it is mainly large companies that look very closely at the legislation.' - Developer 2

The vendors use privacy certifications to verify that the system is GDPR compliant, but the **ERP privacy goals are not stated explicitly**. One of the managers indicated that it is the customer's goals, not the privacy goals, which are key. It depends on the customers' requirements, from which a privacy goal is established.

'For example, suppose you have a customer who has internal audits or a control department. They may well say that they see certain risks if people can see them. It is, of course, a financial risk because the customer can receive a hefty fine. The customers can have specific goals, but from the ERP system, there are no goals.' – Manager 2

- **Challenges**

Developers, consultants and managers experience challenges in the ERP system such as **GDPR challenges and ERP privacy challenges**. Two developers stated that **classifying roles** and establishing who has **access** to which data within the system is challenging. This is necessary

to protect data, as not everyone should have access to it. According to one of the developers, GDPR is the root cause of this challenge.

> *'That is, the challenge for us is to classify those roles. So they don't see more information than they need. That's an exercise everybody does. […] GDPR is the reason that we have the challenge. Nobody would have been bothered about that. So the main thing about the roles is that nobody is able to execute an action they're not authorized to do. So that is the main reason for the roles. But that also plays into the privacy concern because it is data you're giving access to.' – Developer 1*

Another challenge is **anonymizing data**. This is how one developer explained the challenge:

> *'For example, if the dataset, such as the backups of the database, has to be transferred and we have to move deeper into the system, we have to find out why? Occasionally a copy is made of their environment, which then contains all data, including personal data. To make that set of data anonymous, update all names and addresses. That makes it all complicated. It is ultimately not done or too little.' – Developer 3*

However, one of the developers stated that they did not experience a privacy challenge in the ERP system.

> *'No challenge, I would say no.' – Developer 7*

The consultants and managers experience different GDPR and ERP privacy challenges than the developers. One of the managers stated that **testing** real-life personal information is an ERP privacy challenge because personal information cannot be copied into the system.

> *'But I can't copy our production system's personally identifiable and sensitive information to test all these things. So if I have to do that, then I have to take those details or scramble the data so that many people understand I'm going to have to put it here.' - Manager 1*

A challenge that most consultants or managers experience is the **integration of systems**. It is a challenge because all the information in one system must be transferred to the "*new*" system.

'What you also sometimes see is that the ERP system is linked to another system. Then you have to connect the data from one system to another system. How do you ensure that all applications in the ERP systems are neatly closed across all applications?' – Manager 2

'In my experience, I never had this case in terms of critical privacy challenges, except that when you want to integrate with new additional software, there could be a challenge.' – Manager 4

- **Data management**

The access control is perceived as a data management challenge but it is a task that everyone who works in the ERP system should do. One of the developers stated that it is an exercise for everyone. This involves the concept of **the principle of least privilege**, whereby the developer needs certain principles to be able to perform his tasks.

'That is, the challenge for us is to classify those roles. So they don't see more information that than they need. That's an exercise everybody does.' – Developer 1

One of the managers stated that one of their responsibilities is giving **certain users rights to access data.** In this way, the data is not visible to everyone and is protected.

'Those are the two things you have to take into account in terms of responsibility. When I design an authorization, certain people are allowed to see specific data. But also ensure that this data is not visible to everyone.' – Manager 2

The customer, not the consultant, decides which roles have access to which data. The consultant advises and supports the customer. For example, consultants ask the customer how the business process is organised and who is involved in it.

'I don't decide that myself. I'm going to talk to the customer about who can view what. [..] So during such a conversation with the customer, I will ask how do you set this up? What do your business processes look like, and who should be involved in which step in which process? What data can they view, and who carries out the work?' – Manager 2

The customer uses the system and is ultimately responsible for what they put into the system and what they do not explain to a consultant. The consultant is responsible for the implementation and database and provides technical support.

> *'The use of the system is by the client and also their responsibility with what they put in or not. Especially the technical support we perform. We have no real responsibility in this.' – Consultant 8*

However, developers and consultants find that the processing of personal data in the system does not always run smoothly. They experience **malpractices** with data. One of the consultants receives an Excel document by e-mail from a customer containing all the employees' salaries.

> *'What I did see with a customer in France, for example, when we were also doing an ERP implementation with an HR process, the customer threw her workforce in Excel over the email with all the salaries in it.' – Consultant 9*

Another developer gets a database from the customer with all the (personal) data. The developer asked the customer for a test database but this does not always happen in practice.

> *'Yes. I do try to be aware of that database that contains the data. During development, we get files, for example, the payroll interface. So we can see all the wages of companies for that month. Then I always ask for a test file, but the customer does not provide it in practice. They fail, and we get an actual file with that data. It's not OK. But no. I communicate with the customer to have a test file. But if we do get the real thing in practice, you just get to work.' – Developer 3*

Data privacy frameworks and organisational privacy policies are established to ensure that privacy is safeguarded and that there are not too many data-related malpractices, for example. It is important, according to one of the managers, to follow the guidelines and frameworks on how to handle data.

> *'So becomes very important for me to follow the guidelines and principles and work on such projects. You always suggest that whichever is the stricter one, but I think [the professional service firm] has very clear data protection, fundamental courses, and very*

rigid policies and frameworks in place on how this data should be handled. Always go back.' – Manager 1

Another manager had also experienced malpractice where personal information was exposed, but because of the organisation's privacy policy, the manager knew what actions to take.

'I did experience that and I reported that. Because the company I work for also trains you to do. That indicated that it was not good but was ultimately not processed. [..] Yes, and with the customer himself. Just follow the right processes.' – Consultant 9

- **Privacy Design Aspects**

A privacy design aspect that has emerged during the interviews is data minimisation. The concept of data minimisation is quite a familiar term for developers. Four out of seven developers could answer the question of what data minimisation means. All answers are almost the same. The developers describe data minimisation as follows:

'Data minimisation is, like, from a data perspective, having as we already said, less mandatory fields are possible to go through something. So if you don't need that, you don't need that.' – Developer 1

'Only keep data that you are only allowed to keep?' – Developer 3

'I only take data minimisation into account from the perspective of performance. So you have to see it this way. When you build a solution, you want to use as little data as possible.' – Developer 7

However, two developers do not know exactly the meaning of data minimisation as it plays a minimal role in their function. One of the developers answered with doubts, because they were not sure if it is right. Another of the developers cannot give a clear definition of data minimisation, but states that they constantly consider which data is useful and which is not.

Interviewer: And do you know what data minimisation is?
Developer 4: I'm sorry. Data minimisation? Sorry no… No. It might be done, but I don't know.

Interviewer: Data minimisation is a principle that states that data is collected and processed. Its processes should not be held or further used unless this is for a reason. So you can't have hold data too long.
Developer 4: Okay, but as I was saying, we do follow this. But I mean, I play a very minimal role in this, like you are speaking "aware".

More than half of the consultants/managers were familiar with the term *'data minimisation'*. However, not all consultants and managers deal with data minimisation. One of the managers stated that they are not involved with core development and also does not work much with developers.

'Honestly, I have not done core development or created user integration interfaces with specific data minimisation concepts. But I have worked on projects where this was a bit of a challenge of what kind of mission collect.' – Manager 1

'Data minimisation, I don't work with many developers. I saw these screens, the standard screens of SAP or by design to collect minimal information or only the required information. So I don't have any specific deal with this data minimisation or privacy by design in my day to day life.' – Manager 1

Another manager did not know what data minimisation exactly means because the manager focuses more on data visibility. One of the manager claimed that he did not know the term because people do not talk about it much.

'No. We did look at data visibility. And who can change what? And who can see what? But don't store your data for long. We hadn't included that in that project at the time.' – Manager 7

'No, I haven't seen people lecture about data minimisation much.' – Manager 4

Despite the fact that, in comparison to developers, managers and consultants are less familiar with the term "*data minimisation*" and/or are **not involved in data minimisation**. However, a few of consultants and managers were able to provide a definition.

'A lot more, a lot less. There used to be a time when you wanted the person's name, phone number, address, email, address, sex, and other essential fields in one form. But based on the nature of the purpose of the data collection, you might or might not need all this data about a person. So data minimisation is adequate collection of data for the purpose and nothing more than that.' – Manager 1

'So indeed, you have a database that you minimise that in addition to the necessary data that you have. So that you already protect something from others by the GDPR. That you reduce that so that only the essential data becomes available.' – Manager 7

There are **several factors** why data minimisation in the system should be implemented according to the interviewees. Not only does GDPR require developers and consultants to use data minimisation, but so does **customer requirement**. If a customer wants to minimise the data, the developer or consultant must minimise the data in the system.

'It all depends on the functionality that the customer needs. If I minimise the data…' – Developer 1

Another factor is the **requirement of the country**; each country has its own set of data minimisation regulations.

'For example, in Germany, it is mandatory for any expense claim for company travel to mention the address. So that is a rule by the government. We can't bypass that and say that we minimise the data. It's all driven by the regulations of the country and also the functionality that the customer needs.' – Developer 1

The **business requirement** is another factor why data should be minimised. This requirement focuses more on operational needs such as the performance of the ERP system, database storage and costs. One of the developers stated that data is always stored as little as possible because of the database size. More storage in the database may also leads to more costs for the customer.

'No, not really, we always try to store as little as possible. Not because of privacy but more because of the database size. The storage is minimised and additional storage also

costs more money for the customer. So you try to do as little as possible.' – Developer 2

According to another consultant, having too much data in the system has an impact on performance. As there is too much data in the system, it becomes slower.

'Of course. I only take data minimisation into account from the perspective of performance. So you have to see it this way. When you build a solution, you want to use as little data as possible. Because it just slows down. You delay the resolution. And the more unnecessary data you bring in, the more that system has to work to get all the data. So the longer it takes. In other words, from that perspective, I am constantly working on data minimisation. From a privacy point of view, as I said, not really. Because I assume that a requirement has been drawn up with the idea that the data is minimised. That you only use the data you need.' – Consultant 9

Furthermore, the customer can also require **holding the data for a certain time**. Manager 4 stated the following:

'We do a statement of work, agreeing with a client that they will have the legacy system available for 5 to 10 years, depending on how the security..,. how the clients, auditors are okay with. Based on that only, we decide.' – Manager 4

To minimise the amount of data, the data is removed from the system after a certain period of time. It is also possible that data is kept for a longer time because the company needs it.

'Most of the business I work with or most of the people I work with, we tend to keep the data long. I would say three to five years is the bare minimum they expect to keep, especially the supplier. The customer information will be there for long, but product information after the end of the cycle, we try to keep it at least for three to five years to make sure that if there is a service, or if there's a problem, we should be in a position to address it.' – Developer 4

The literature proposed a two-step data minimisation methodology to minimise the data in the software system (Awanthika Senarath N. A., 2019). Understanding the data and classifying

whether the data is sensitive, visible, and relevant to the system is the first step in the data minimisation methodology. The second step is to make system design decisions that will minimise the amount of data used in the system. One of the managers stated that they are looking for data visibility in the system, as well as who is allowed to view the data. However, developers and consultants do not need to understand the data because they only start minimising data in the system when they get a requirement from the business or customer. The result has shown that the developers or consultants did not follow the second step proposed in the methodology.

- **Privacy-by-design**

Privacy-by-design is not a well-known and most used privacy concept by developers and consultants. Three out of seven developers could not explain what privacy-by-design means. They have never heard of it and have no (direct) experience of it.

> *'I've never heard it before your email. But I think it's about the design and a solution that offers as much privacy as possible to the customers. It's never been called that with me, the concept of "privacy-by-design". I've never heard it before.' – Developer 2*

This applies to consultants and managers as well. Three of the nine managers/consultants are unaware of what it entails and have no prior experience with privacy-by-design. This corresponds to the literature research by Jeroen van Rest *et al.* (2012) that discusses that privacy-by-design is a vague concept for managers, and it is not clear what it exactly entails. However, the other four of the nine consultants were familiar with the term but **do not deal with** it or do not have any experience with it. Nevertheless, the developers and managers describe the concept of privacy-by-design in the same way:

> *'In my view, privacy-by-design is already taking privacy into account when implementing or configuring a system.' – Manager 2*

> *'Privacy by design is when you're building a system, or any process for that matter, something as simple as a flyer that you ask people to fill in. You have to do that with the privacy concerns of the users or the people who give their information in mind.' – Developer 1*

One of the reasons why developers were not familiar with privacy-by-design and had **no experience** in implementing this privacy concept is because everything is already done by the ERP vendors. As mentioned before, ERP vendors make sure that privacy concepts are already implemented in the system, so the developers do not have to do much. One of the developers stated the following:

> *'No, unfortunately, I have to disappoint you. I think the system does a lot for us, so it relieves us. Everything has already been described within Microsoft. It is not a complete ERP solution. That is why little attention is paid to what is necessary. But the fact no.' – Developer 3*

The reason why consultants and managers **do not have experience or do not deal** with privacy-by-design is because it is not part of their daily responsibilities/tasks. Another finding why the manager does not have experience with privacy-by-design is because they do not deal with HR data in the ERP system. So they do not have to implement additional privacy measures like privacy-by-design.

> *'So I don't have any specific deal with this data minimisation or privacy by design in my day to day life.' – Manager 1*

> *'My experience is that I don't have much experience because I don't work with HR data. This lies within the HR department or developers, and I don't have much to do with that myself. So if I look at the customer data and supplier data, no, sorry, I can't comment on that, no experience with that.' – Manager 2*

Furthermore, a developer noted that the ERP system **does not currently use privacy-by-design**. When questioned if privacy-by-design is crucial for the ERP system, the developer responded by emphasizing the importance of data security than rather emphasizing privacy solutions.

> *'Yes and no. Because without access to that… It's not like public domain data. It is a private network where only people in the organisation can see the data. And with a certain level of clearance. Yes. So that is better already.' – Developer 1*

- **Impediments**

Several developers and consultants **see privacy as a security problem.** Regarding the privacy goals of the ERP system, the developers' answers were more focused on securing information than privacy solutions such as roles system security. This is in line with what the literature has investigated. Hadar *et al.* (2018) discovered that most developers view privacy from the perspective of data security, concentrating on technical concerns and security solutions. The developer stated the following:

> *'The problem is that the ERP system is mainly used for use within companies. So from the outside, there are certain integrations and there you effectively have all the tools to close where certain subset of data is sent out. There is also an extensive roles system with security. [..] So it's all pretty watertight. Later in the stage of a project, the roles are determined, and so is the security. It probably will be.' – Developer 2*

Furthermore, there is a **conflict between what the business needs and the privacy regulations**. A customer requirement, according to one of the developers, might be a challenge since the customer wants specific data, such as birth or date, and the developer questions if this is appropriate in light of the GDPR regulations. This adds to the developers', consultants' and managers' overhead when it comes to establishing the ERP system in a certain way. One of the managers stated that the business wants data testing but that is not possible because it involved personal information. GDPR requires to protect personal information. There is a conflict between what a business wants and what privacy dictates, which can lead to a challenge.

> *'So that is the most biggest difficulty whenever you go for implementation. If you want to do some testing out of the actual real-life data, you really can't do it concerning the personally identifiable information. That's the challenge.' – Manager 1*

Another privacy impediment that makes it challenging to implement privacy techniques in the ERP system is the **mixing of different concepts about privacy**. It resulted in developers, consultants, and managers mixing up different privacy concepts. Questions were asked about privacy, but the interviewee answered back that is not entirely privacy-related. For example, one question was asked about one of the managers' privacy responsibilities and data protection. The manager answered that it was not their responsibility to talk about customer turnover.

'And don't talk much about the client's turnover or revenue, with a team. So it's very restrictive within the project team.' – Manager 4

As previously mentioned, not all developers, consultants, and managers were familiar with privacy concepts such as data minimisation and privacy-by-design. They **lack knowledge** about these privacy concepts which can make implementing these concepts almost difficult for them. One of the reasons why they do not have knowledge about these privacy concepts is that they are not talked about or lectured about it.

'No, I haven't seen people lecture about data minimisation much.' – Manager 4

4.3Strategies and Techniques

This section answers the sub-question *"what privacy-by-design and data-data minimisation strategies and techniques do developers and consultants use in the ERP systems?"* To answer this question we use the results of the previous section.

Due to the lack of knowledge, the developers and consultants/managers are not familiar with the available techniques or strategies to implement privacy-by-design and data minimisation in the ERP system. One of the developers stated that they are not aware of any guidelines for implementation of privacy-by-design in the ERP system. It is not clear how to do it, and they wish the process was more transparent. This can probably be improved by doing research to make sure that they understand the guidelines better.

No, it's never been called that with me, the concept of "privacy-by-design". I've never heard it before. […] I don't know, in terms of guidelines. There is some request or requirement for GDPR to work compliant. This means that we anonymize or delete the data as much as possible. Therefore, further research needs to be done. [..] Yes, how it can be better. – Developer 2

Privacy concepts have already been implemented in the ERP system, as mentioned earlier in the previous **section 4.2**. As a result, developers and consultants did not have to think about or implement privacy-by-design or data minimisation. One of the managers stated that there are already globally approved standard screens, therefore they do not need to do anything to accomplish privacy-by-design. This is applicable for data minimisation as well.

'But in the ERP system, as I said, these are standard screens that are globally accepted and agreed upon. I don't have to put anything specific to achieve this privacy-by-design.' – Manager 1

Another developer stated that the data is kept in the system for a certain time. After a number of years the data is automatically removed from the system. The only thing the developer has to do is to make a rule about how long the data can be kept and the system does the rest.

In the ERP, what we follow I'll tell you. As I mentioned, we tend to keep an archive for three years, five. So after that, in Oracle, you have an option to delete the data automatically on its own, after, say, ten days. So you tend to set it up that rule, and then the system itself takes care of it, irrespective of the number of years or days, whatever you set. – Developer 4

According to one of the managers, whenever a project involves a lot of personal data being processed in the ERP system, a data migration team is brought in to help. Another manager stated that implementing data minimisation in the ERP system is not their responsibility, but the customer responsibility. The following is what the manager stated:

'No, not. I think that's because it is mainly outside "my scope" or our scope of the department. That's where the customer cares. What data are we going to put in the system, and what are we going to process. This is beyond our responsibility.' – Manager 2

There are different factors to consider when implementing data minimisation and these factors also influence which strategy or technique is applied. According to one of the developers, there is no specific strategy or technique to implement data minimisation. It depends on the customer's needs and what kind of functionalities they require in the system. Other factors play a role in data minimisation, such as business processes and geography. This finding is in line with the literature research. Oetzel *et al.* (2013) found that developers are unaware of techniques to implement data minimisation. However, Arachchilage *et al.* (2018) found out that if developers get specific instructions on using as little data as possible, they can implement data minimisation in their system designs.

'It all depends on the functionality that the customer needs. If I minimise the data… If some data is mandatory for some process to be executed. So we don't have any process as such, but it's not set in stone, and there are no guidelines as such. For example, you need to process the salaries you need the bank account information. Right? [..] So the country demands that. And additionally, the ERP also has this, what is the country-specific rules setting. So we may have to collect more data than then the other countries.' – Developer 1

To the question about whether there are specific techniques or strategies for implementing privacy-by-design in the ERP system, one of the developers responded as follows:

Honestly, in current business. No, we don't tend to be entirely inclined to it. As I mentioned, we have some standards, which we tend to follow, but not precisely what the book says. – Developer 4

Common security controls the developers use to safeguard privacy are multi-factor authentication and auditing logging.

'So one essential thing is multi-factor authentication; irrespective of the application, I build it with multi-factor authentication. That's one of these, which comes as part of the privacy by design. And the second part is generally the logging and the auditing part of it, in which we make sure all the logs are captured. And auditing logs are also going to be captured irrespective of the system they work with, or we build it.' – Developer 4

To the question about whether there are specific techniques or strategies for implementing data minimisation in the ERP system, one of the consultants responded as follows:

'No, there are no guidelines for our implementations. I don't think there's a framework for that. If there is a guideline, I believe that with us, less is more—record as little data as possible.' – Consultant 8

However, there are some privacy-by-design and data minimisation strategies and techniques that developers and consultants use in the ERP system. Two managers stated that there are templates/blueprints available within the consulting's firm where the research has been done.

These templates/blueprints explain what the organisation thinks about the privacy-by-design or data minimisation concepts. Sometimes the blueprints are presented to the customer to check whether this approach is suitable for them.

> *'If we have an implementation from scratch, we have blueprints of how we think about it, especially the business process. There is also a template for how we think about the data. But in the end, it is always the customer who decides. But within the company, we have a blueprint that we can present to the customer to ask whether this would work for you.' – Manager 2*

> *There we also have those templates, those suggested templates. And if it contains a column with data that contains privacy-sensitive information, we will take a critical look at it. I think the customer too, why do you want this in your ERP? Usually, it is straightforward and understandable. – Manager 6*

Another strategy or method for implementing privacy-by-design in the ERP system is to set up controls in the design. For example, someone in the Finance department can not see the salary of someone in the HR department. By incorporating different controls into the design, the data is protected so that not everyone can access the data. This is in line with the findings of the literature review. There are eight privacy-by-design strategies, according to Hoepman (2014), including the *"control"* strategy. Another strategy mentioned by the consultant is to call in a dedicated compliance team to ensure that the system is GDPR-compliant. However, this strategy is not in line with the strategies of Hoepman (2014).

> *'So you can, for example, if you do the ERP implementation and you have someone from Finance, someone from HR and someone from Purchasing. They are given different roles in the system to force a person to see everything. So you can set up controls in the design so that privacy and data are protected that you as a Finance person can never see the salaries of someone who can see HR, for example. So then you are working on roles based on function so that only the minimal data is available to people. So that's a method.' – Consultant 9*

> *Only if you say 'certain strategies', I don't know what the strategy would be. Then I would think, make sure you are GDPR compliant. And put that upfront. Make sure*

you've thought of it. That's what the coverage and compliance unit is for. Internal controls, you name it, to check that. – Consultant 9

The literature also stated that PIA could help with the privacy-by-design approach. The PIA provides clear privacy objectives and specifies a means to achieve them. It is also known as a "*milestone towards privacy-by-design*" (The European Data Protection Supervisor, 2018). However, the PIA approach did not emerge from the findings of this study.

4.4 Improvement of Understanding

This section describes how to increase the understanding of privacy-by-design and data minimisation in the ERP system. The sub-question, *"What would help the developers and consultants to increase understanding about privacy-by-design and data minimisation techniques in the ERP system?"* is answered.

Each expressed an opinion or suggestion for improving the ERP system's understanding of privacy-by-design and data minimisation. The most recommended answer from developers and consultants is to increase privacy awareness. By raising privacy awareness, they better understand privacy concepts. This is not only a recommendation for the developers and consultants to raise their privacy awareness, but for the whole company, one of the managers said.

'What would I suggest? Yes. I think creating more awareness. First, that has helped me. So I think for all people consultants and developers, but just the whole company. And then emphasize what the risks are and the consequences of these risks, such as a fine. Of course, you also do not want a data leak that your data ends up with someone else. To create awareness within the company.' – Manager 2

Raising privacy awareness can be accomplished by providing a presentation to someone familiar with the subject. However, the developers present must take a test after the presentation. The organisation can test whether the developers actually understand it by taking a privacy exam. This can also benefit the organisation, as it can be determined whether or not the employees are aware of the importance of privacy.

'I would appoint someone who understands and has to give a presentation. The person should explain how it works, what is needed and what you should pay attention to. [..]

In short, I would give a presentation about that, and the developers who attend have to take a test. This way, you can check whether this has stuck. They are sleeping during training. They are doing something else in the meantime, but you can determine whether they have understood it with a test. This is important. Then you as a company can also demonstrate by taking tests that your people are aware of privacy and now they have to apply it.' – Manager 3

Attending regular training as part of one's work might also help the employee to raise their privacy awareness. Employees can keep up with the latest privacy developments by attending regular training. However, a manager stated that it is also critical that the training is relevant, that everyone can understand it, and it is held during working hours. If this is handled properly, people are more motivated to attend training.

'They need to follow a course. So they should not be like me, who had partial knowledge. So they should actually start a refresher course every six months. Or a training every six months, and giving them about new developments in these areas, and what has been expected of them. So without that, they can't continue, or they might lose access, something like that. And they want work on it proactively and finish the certificate. So it should be part and parcel of work.' – Manager 4

In addition, developers and consultants need general privacy guidelines in order to properly understand privacy concepts. If there are no clear procedures or standards, it is not only difficult to understand what it all means, but it can also lead to frustration. Another consultant prefers clear policies and procedures to help them decide which direction to take.

'But having your procedures in place, and simple procedures, not too much complication. When you make things more complicated, it's just becoming annoying.' – Manager 3

'Yes, I certainly prefer that we, for example, have a specific direction of the things that we should specifically pay attention to in the GDPR officer who validates the project how we should approach that way.' – Consultant 8

One of the managers recommended that developers and managers should not only be aware of privacy and establish general guidelines but also be able to sit down together to discuss and address privacy issues, for example. This raises awareness of privacy issues, but by holding a meeting, everyone is informed at the same time, rather than hearing it from other project members or colleagues. Another developer stated that when it comes to privacy, all project stakeholders should be involved to understand why these (privacy) decisions were made. By involving everyone and explaining why certain decisions are made, people get a better understanding of privacy concepts such as privacy-by-design and data minimisation. The developer also explained that everything concerning privacy is regulated in a project. Everything is watertight, so nothing can go wrong.

> *'At this point, everything is ready for you when you land on a project. You get a laptop, and you can't do anything wrong. You can't do crazy things. You can't email from that laptop. You can't just send excels from that laptop to something else. So basically, everything is already watertight.' – Developer 6*

It is also recommended to assign a data protection officer or data controller to a project so that someone with relevant knowledge and experience can guide the other members of the project team. Not everyone knows a company's privacy procedures, and if one person can guide the team members, that is sufficient. One drawback, however, is that DPOs are not project-specific. A data protection officer is often a single person within the company. Therefore, the recommended follow-up is to establish general guidelines, which managers consider sufficient.

> *'Well, you're right. They're not projects based usually, the Data Protection Officer is one person as a company. And if I have the general guidelines on how this data should be handled, I think that is sufficient for us.' – Manager 1*

5. Behaviour Model

This section looks at the behavioural models of developers, consultants, and managers. Based on the results from the previous section, we can discuss the behaviour of developers, consultants, and managers. The first section is about the behaviour of the developer, and then about the behaviour of the consultants and managers. The BJ Fogg Model is used to understand behaviour, using the formula: *motivation + ability + triggers*. We used the codes in the codebook (see **Appendix III**) to identify the aspects of behaviour in this study, these codes are marked ***blue*** in this chapter. The codes were assigned to a category such as motivation, ability, triggers and general.

5.1 Developers Behaviour

First, we start with motivation. What motivates developers to perform certain behaviours? This can be **GDPR compliance**. One of the developers states that they have to follow the GDPR requirements. If they do not follow the GDPR requirements, there are consequences for losing their job. Therefore, they are "*highly*" driven to follow GDPR requirements. This is how the developer explains why it's important to follow GDPR.

> *'Yes, it's a very one of the key factors. If one of our developers doesn't abide by GDPR, they might end up losing their job. Or there'll be strict actions taken against them. So the data protection is very key. And we all, as a developer, as a consultant, we are abide by it, and we have to follow it.' – Developer 4*

Compliance and adherence to certifications can also be a motivation. It is a high motivation because the developer states that they have to go through the mandatory stuff every year to comply.

> *'We do it once a year. We have some certifications, some mandatory stuff which we go through once a year. However, in case of any reports of loss of some certain stuff, we have ad hoc one. Once again when someone has reported, for example, a server has crashed, something went wrong… So that's when we make sure that nothing is broken and everything is as expected.' – Developer 4*

Another motivation for the developers is **the role of GDPR in ERP privacy**. For example, the developer must follow some GDPR standards in the ERP system by ensuring that each employee is given access to certain information only to do their work. Therefore, the following standards derived from the GDPR can be considered as a "*high*" motivation.

> *'Correct. We follow that standard, we make sure, especially with the ERP, you have customers coming in, and your suppliers coming in, and you have your own employees. Making sure each employee gets only the information he is required to work on. And the customers get to work on a system where they are expected to work off. And even the suppliers – the same thing. They need to know only the information they are required to. For example, in case of updating something, we don't want the information from one supplier letting go to the other supplier. You have a lot of stuff. So GDPR, we do follow it. I do take time that… We have a mandate where I have to acknowledge it every year and tell that I follow the regulations and don't let go of it.' – Developer 4*

The **GDPR affects the developers' work**. For example, one of the developers states that they cannot publish data to the outside world due to the GDPR, and data protection has become important. Therefore, the role of GDPR in work can be seen as a "*high*" motivation for the developers.

> *'After working from home, so working remotely, data protection has become the critical thing. Now, most of us work remotely but make sure that we don't tend to leave out our data to our neighbours or to or anyone for that sake, at our home.' – Developer 4*

In some ERP systems, little personal information is kept, so developers **do not deal with or work with personal information in the ERP system.** One of the developers explains that little personal information is kept in the system. They do not deal much with personal information. If they do not deal with personal information in the system, the motivation to implement privacy techniques to protect this personal information can be very low.

> *'Very little personal data is retained in the central ERP system we work with. There are systems; customers have had quite a bit of sensitive information. I haven't done much with it, to be honest' – Developer 6*

Another (de)motivation can be that the developer **does not deal with privacy;** the lack of motivation is low. This is how developer 2 expresses their thoughts on privacy.

> *'No, you don't think about privacy at that moment.' – Developer 2*

Following an **organisation's privacy policies** can also be a motivation. One of the developers states that following privacy policies is important so that everyone is aware of what the rules are. This is aligned with Hadar *et al.* (2018) findings in the literature, which indicated that companies embrace privacy policies that impact developers' privacy perceptions. In addition, the organisation's privacy policy comes from GDPR, so the developers are motivated to follow it.

> *'So what I would say, we tend to, for all of them, to make sure that everyone is aware of this privacy by design, and we stick to it, and we follow it everywhere across. Rather than few folks following here, few folks following there, and most of them aren't aware of everything. Like in my case, as I mentioned, we follow a couple of it, not everything. So if you tend to follow, if you make a policy, or if you make a rule, then if everyone follows it, it would be better for whatever we are doing in it. ' – Developer 4*

- **Ability**

It is also essential to have the ability to perform a behaviour. A **conflict between privacy and business needs** can affect the developer's ability. While the business needs something like creating a function where they can see everything in the system, this can conflict with privacy laws. This makes the developer's ability "*obstructed*".

> *'I had privacy challenges with my previous employer. Based on the roles I said, we had to create a function that could see everything. But also really everything and nothing that was allowed to change. That was a real challenge. That was intended for external audit people. They had to come to a production system, but they were not allowed to change anything, but they could see everything.' – Developer 2*

Mixing different privacy concepts can also make the developer's ability hindered. Implementing the proper privacy technique can be "*hard to do*" for the developer by mixing

different privacy concepts. The right knowledge is not applied, making the ability to implement privacy techniques in the ERP system not easy to do so.

> *'No technique, not to the best of my knowledge, but strategy, for example, we have separate environments for the role we want to do. And so, not everyone has access to that. The accesses are going to be given not by me from someone specific. So, for example, I can provide access to someone to in a particular environment, but I'm not supposed to do that. So kind of a strategy to make sure we know about who has access to this data migration environment, which has the actual data, even not all the functional consultants are allowed to see all the data. That is very necessary for having these separate environments for different purposes.' – Developer 7*

The ability to implement privacy techniques in the ERP system can also be obstructed because developers see **privacy as a security problem**. Developers do not have the proper knowledge, seeing privacy as a security problem. Not having the appropriate expertise can make implementing privacy techniques "*hard to do"*. This is how Developer 1 answered the question whether privacy-by-design is essential for the ERP system:

> *'It's not like public domain data. It is a private network where only people in the organisation can see the data. And with a certain level of clearance.' – Developer 1*

The developer's ability to implement privacy techniques in the ERP system can be hindered due to a **lack of knowledge**. Because of not having the right or sufficient knowledge, developers do not know how to implement a privacy technique. Furthermore, it affects the developer's ability if they have no experience with privacy techniques such as privacy-by-design.

> *Developer 7: From you? But before that, privacy by design? No, I don't think so.*
> *Interviewer: No. And do you have any experience with that?*
> *Developer 7 : No, I don't think so.*

However, **privacy out of the box** makes the developers' work more accessible. Through *"privacy out of the box",* where ERP providers implement privacy frameworks in the system, developers do not have to implement privacy frameworks themselves. The ERP system is already compliant, so the developer does not need to take any additional measures.

'Yes, in the past, but now it is more and more added out of the box. So one way you have to add it and you go and look at the existing packages. Or, if you look at the ISV solution, it is mentioned everywhere that it has been added. When you talk about ERP packages, it is mainly large companies that look very closely at the legislation.' – Developer 2

- **Triggers**

The behaviour can also be triggered. As mentioned earlier, triggers can be divided into social, forced or proactive (Sauvik Das L. A., 2019). A "*forced*" trigger for a developer can be a **requirement**. These can be both business and customer-related. Because the developer has no choice but to comply with this requirement, it is a forced trigger. On the other hand, the **business requirements** focus more on operational needs such as system performance, cost and storage. This is how Developer 2 explains that, based on a business requirement, they need to perform an action:

'Not because of privacy but more because of the database size. The storage is minimised and additional storage also costs more money for the customer. So you try to do as little as possible.' – Developer 2

The **customer requirement** focuses more on what the customer demands and triggers the developers' behaviour. For example, the developer explains that he cannot share code with other parties because the customer requires it.

'Yes, so right now, I'm working on one single project, and I've been working on this project for a year and a half, I think, maybe even more. But there were projects in which the requirements for privacy were much more strict. So, for example, we had a police department, which was our client. So basically, they were vetting even us to make sure that we didn't get any faster or something. And, even when we developed code, we couldn't even share a snippet of code with my friend. For example, I was working with another project who might have had the exact requirement. And so that was, that was from a coding perspective. But then, from an actual client point of view…' – Developer 5

Nevertheless, the **customer is also involved** in projects, which can trigger the developer's behaviour. Furthermore, it is a "*social*" trigger because the customer's involvement in a project

also gives the developer different insights or opinions on how, for example, data is processed and stored.

> *'You want to process that data, and also give a choice to the customers or the people who give you the data to have a say in how it is managed or how it is stored or how it is shared.' – Developer 1*

Anonymizing or deleting personal information also triggers the behaviour of the developers. One of the developers explains that it is a requirement to be GDPR compliant by deleting or anonymizing personal information. Therefore, it is a forced trigger because the developers have no choice but to follow this requirement derived from the customers.

> *'There is some request or requirement for GDPR to work compliant. This means that we anonymize or delete the data as much as possible.' – Developer 2*

However, anonymising and deleting personal information can also be an **implementation practice.** This implementation practice is then a trigger for the developer. For example, one of the developers explains that they deleted the data because GDPR requires it to be compliant.

> *'So if the customer says, I want to know who all have the information to share that. And if somebody wants to delete the data, you can contact the people who can process that in the automation.' – Developer 1*

As mentioned earlier, **GDPR compliance** can be a motivation, but it can also be a trigger for the developers. It may even be a forced trigger, as the developer has started implementing data classification because of the GDPR. This is how one of the developers explains why they have to take action because of the GDPR:

> *'There's nothing in the case of Dynamics; there is no framework to hide the data by default. So, but the only thing we do is classification for GDPR regulations.' – Developer 1*

Another forced trigger can be **the role of GDPR in work**. One of the developers states that they had to build extra things into the system to comply with the legislation. The developer had no choice but to build certain things into the system because of the GDPR.

> *'Yes, introduced, but we had to build certain things into our systems to comply with the legislation.' – Developer 2*

The role of GDPR in ERP privacy can also be a forced trigger for the developers. GDPR prompts the developers to take an action in the ERP system, such as giving access control based on the roles. This is how one of the developers explains the role of GDPR in ERP privacy, requiring the developer needs to perform an action:

> *'GDPR is the reason that we have the challenge. Nobody would have been bothered about that. So the main thing about the roles is that nobody is able to execute an action they're not authorized to do. So that is the main reason for the roles. But that also plays into the privacy concern because it is data you're giving access to.' – Developer 1*

- **Improvements of Behaviour**

Nonetheless, there are a few recommendations that can help to improve developers' behaviour. In the Fogg's model improvement of behaviour means increasing the ability or motivation, or adding a trigger, what facilitates moving across the action line. One of the recommendations is to have **general guidelines**. This may be a motivation or ability for the developer. It can be a motivation because the developers are more motivated by general guidelines to implement privacy techniques. According to one of the developers, it can be demotivating and annoying if there are no general guidelines. On the other hand, this recommendation can also improve the developer's ability and make the work "*easier*" because there are clear and general guidelines on what the developer has to do.

> *'But having your procedures in place, and simple procedures, not too much complication. When you make things more complicated, it's just becoming annoying. Like, you're going stop you from working efficiently and may cause people to try to go around these processes not so simple process processes that make sense and some central people who can manage this.' – Developer 7*

According to one of the developers, getting **involved in a project** is a recommendation. The developer explains why it is important to get involved in a project to increase the ability:

> *'That I'm getting involved in, I'm getting already, but everyone is getting a little bit involved in, 'Hey, we made this and this choice, for example. [..]' So that you get a detailed description of how it happens, take people in there and tell them why confident choices were made. Which they are not aware of. This creates more awareness within the project, so that might be a tip' – Developer 6*

Raising privacy awareness by, for instance, following training, presentations or workshops can change the developer's ability. In particular, the level of privacy knowledge can be increased, enabling the developers to implement privacy techniques in the ERP system. By raising privacy awareness, the developer's ability becomes "*easier*".

> *'By having training, having a mandatory boot camp, just to give you the policies of the company. So explain, based on the rule on what they will have, what they need to do.'- Developer 1*

Another recommendation is to **research the requirements of the GDPR**. One of the developers states that the GDPR requirements should be studied so that developers at least understand them. The developers' ability improves when the GDPR requirements are researched. However, developers are still unclear about what the GDPR requirements mean, which makes the ability to follow privacy policies or techniques very difficult.

> *'Therefore, further research needs to be done. [...] Yes, how it can be better.' – Developer 2*

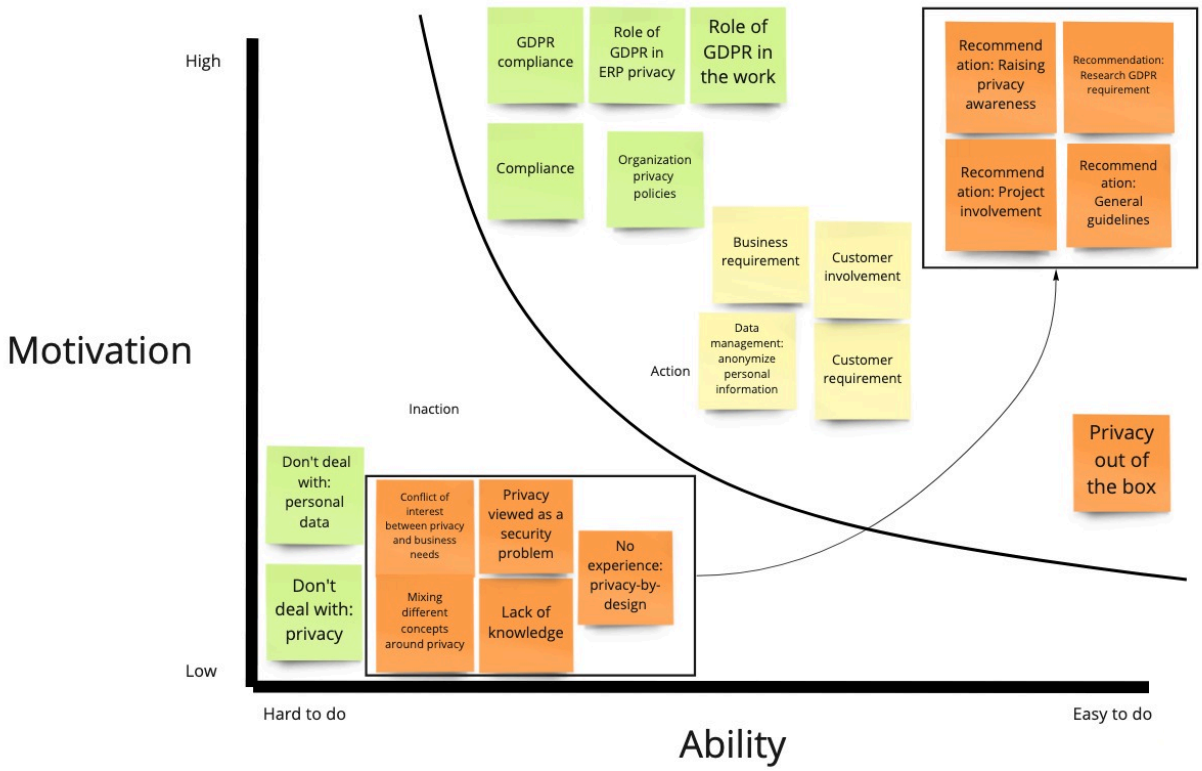

Figure 14 Developers Behaviour Model
Green stands for motivation, orange for ability and yellow for triggers

The behavioural model is illustrated in Figure 14, where green stands for motivation, orange for skills and yellow for triggers. The recommendations pointed to by the arrow can increase the ability of the developers at the bottom of the action line. The recommendations address the lack of knowledge and increase the ability to implement privacy protection strategies, thus moving the action line upwards by increasing awareness of privacy protection through training.

5.2 Consultants and Managers Behaviour

Consultants and managers are also motivated by **GDPR compliance**. One of the consultants explains that GDPR is high on the list, and therefore the motivation to be GDPR compliant is very high.

> *'Everything is legal around accounting and lawful things that need to be done, and GDPR is also high on the list. You must be compliant, and of course, Microsoft must also comply with the package.' – Consultant 9*

The role of GDPR in the work of consultants and managers can also be a motivation. According to one of the consultants, more care has been taken when handling data because of the GDPR. This is how one of the consultants explained the role of GDPR in work:

> *'Partly yes, I will not say that there is a very strict follow-up and that the awareness is not quite there yet. When this law was not yet in place, employee data and suchlike were handled more senselessly. But now because it's more sensitive.' – Consultant 8*

Another motivation for the consultants and managers can be the **organisation's privacy policy**. For example, one of the managers explained that it is essential to follow the guidelines and principles that come from GDPR. Therefore, it is a "*high*" motivation because the manager sees it necessary to follow.

> *'So becomes very important for me to follow the guidelines and principles and work on such projects.' – Manager 1*

One of the managers explains that the project's final product must comply with **privacy regulations.** And besides, it is also their responsibility to guarantee that the final product is compliant with privacy regulations. That is why the motivation is high.

> *'At that time, we were responsible for ensuring that the delivered product complies with legislation and regulations. Because, of course, you have privacy about data.' – Manager 6*

However, it can also be that the consultants and managers are not motivated. There is a lack of motivation if they **do not deal with data minimisation, privacy-by-design, privacy or personal information.** Consultants or managers do not deal with data minimisation or privacy-by-design because it is not part of their responsibility. This is how one of the managers explained why they do not deal with data minimisation or privacy-by-design:

> *'Data minimisation, I don't work with many developers. I saw these screens, the standard screens of SAP or by design to collect minimal information or only the required information. So I don't have any specific deal with this data minimisation or privacy by design in my day to day life.' – Manager 2*

Another manager explained that they **do not deal with personal information** in the system, but mainly with financial data. This creates a lack of motivation.

> *'Again, because I mainly focus on the financial part, I do not directly contact very privacy-sensitive information there.' – Manager 7*

However, it is also the case that the manager **does not deal with privacy** because it is not their responsibility. Therefore, they are not driven to implement privacy techniques in the system. One of the managers explains what their responsibility is related to privacy:

> *'None, to my knowledge. I think anyway we do the implementation and database. The use of the system is by the client and also their responsibility with what they put in or not. Especially the technical support we perform.' – Consultant 9*

- **Ability**

The ability of consultants and managers is also impeded by **conflicting interests between privacy and business needs.** For example, Manager 1 explains that the business wants to test real-life data, which is not allowed under GDPR because it contains personal information. This obstacle for the manager results in low ability ("*hard to do*").

> *'So that is the most biggest difficulty whenever you go for implementation. If you want to do some testing out of the actual real-life data, you really can't do it concerning the personally identifiable information. That's the challenge.'- Manager 1*

The lack of knowledge and the mixing of different privacy concepts also obstruct the ability of the managers and consultants to implement privacy techniques in the ERP system. Managers and consultants do not know which privacy concept applies to the ERP system because they do not have the proper knowledge. For example, when asked what the privacy goal of the system is, one of the managers answered the following:

> *And, and it's been secured and approved for, so we have Azure file storage, which is, again, a secure location. So everything what Microsoft offers is just 100% secure, but apart from that some clients need additional validations. So we have many certificates and many authentication protocols that we follow if the client is huge and if they are*

very much inclined towards data protection. But it varies from client to client. – Manager 4

The ability to implement privacy techniques in the ERP system can also be impeded if consultants and managers **see privacy as a security problem.** This is how Manager 4 explained what their responsibility is related to privacy:

'And from my side, it's about keeping the system secure in terms of proper security authorizations. And of course, on the technical side, Microsoft already gives a certificate. That's all I can say about the security part.' – Manager 4

Furthermore, the results reveal that consultants and managers have **no experience with privacy-by-design and data minimisation in the ERP system**. With the lack of experience, the ability to implement these privacy techniques can be low.

'My experience is that I don't have much experience because I don't work with HR data. This lies within the HR department or developers, and I don't have much to do with that myself. So if I look at the customer data and supplier data, no, sorry, I can't comment on that, no experience with that.' – Manager 2

However, **privacy out of the box** makes the consultants' and managers' behaviour "*easier*". This way, consultants and managers do not have to implement privacy concepts themselves and think about whether the system is GDPR compliant.

'The role of GDPR, like I said, the ERP system is already compliant. So you don't have to be mindful of the GDPR part.' – Manager 1

This also applies to **privacy certification** of the system. The managers' ability is facilitated by the fact that, as one of the managers explained, the system is already protected by a privacy certification. This is how Manager 4 explained privacy certification:

'Microsoft themselves give the system with full privacy and data protection certificate. So that comes with the tool. Within the system, for accessing data, we have the security and role authorizations, which allow only certain people to access certain data sets

within a particular form within the system. But overall, everything is protected with data privacy certificates.' – Manager 4

- **Triggers**

Consultants and managers are triggered to perform certain behaviours. **Customer involvement** can be a "*social*" trigger because the customer gives input on what data should be in the system together with the consultant. There is communication about who puts what in the system. Eventually, the consultant gives the "*power*" to the customer to enter data himself, so the consultant does not have to do anything.

'Usually, you do that with the customer. You give the customer the power to enter that kind of data so that you don't touch it yourself and don't have those files at all. At least that's how I did it on that project.' – Consultant 9

A **customer requirement** can be a "*forced*" trigger. For example, the consultants or managers must comply if the customer requires data to be kept in the system for a particular time. They have no choice but to do what the customer requires. This is how manager 4 explains the customer requirement and which action they need to take:

'So, but here, how do we imply GDPR? I don't think it could be a case because usually… This is the thumb rule that we keep: We do a statement of work, agreeing with a client that they will have the legacy system available for 5 to 10 years, depending on how the security… how the clients, auditors are okay with. Based on that only, we decide.' – Manager 4

Another "*forced*" trigger can be **the data management access control**. One of the managers explains that there is a requirement from the customer to implement access control so that not everyone can access all the data in the system.

'Requirements are collected from the customer. The customer must say that these are our requirements in the field of privacy. We do not want everyone to be able to view the data of others. If they're going to use the HR module within the ERP system, all your details are listed there, so their address and bank details. Sometimes the salary is also paid via the ERP system. Then the customer will undoubtedly have requirements regarding privacy.' – Manager 2

As mentioned earlier, an **organisation's privacy policy** motivates consultants and managers. But it can also be a trigger for consultants and managers. Consultants and managers think about how to share sensitive information in the system and hide it in files to secure everything. This action comes through the organisation's privacy policy.

> *'For example, with the design, what I just mentioned, especially with HR processes, you will put people in the system, salaries, you name it. Which is sensitive, because how will you share that? While, you, as an implementation party, have nothing to do with it. This is usually agreed upon with non-disclosure agreements and things like that. But you have to hide those files well, make sure it's only in one place, that sort of thing.' – Consultant 9*

- **Improvements of Behaviour**

A variety of recommendations given by managers and consultants can help to improve this behaviour. One of the recommendations is to appoint a **data protection officer (DPO)** for a project. This recommendation makes the consultants' and managers' behaviour "*easier*". A DPO knows how the privacy processes work, so the consultant or manager does not have to think about it himself. This is how Manager 1 explained the DPO in a project.

> *'The first thing I would say is… In my mind, do you have a data protection officer or data controller as part of your projects. That is the most important or the first question. Because when I'm doing my mergers, acquisitions and separation projects, often I have to look at the personal data that needs to be separated into two different companies that are merged between two companies. So without looking at the data, I will not be able to make a decision on that. […] So the data protection officer is the data controller where they should be aware of these processes. So every organisation has their way of doing things to be compliant with respect. That is where I started, and you have a data protection officer. Do you have a data controller for this project? ' - Manager 1*

Another recommendation is establishing a **GDPR checklist** to make the consultants' and managers' work more accessible. This checklist ensures that every sprint within a project complies with GDPR, which makes thinking about GDPR during the project much more manageable.

'We work very Agile, of course, so we all have features that are planned in sprints, etc. Such a sprint often consists of several requirements or parts to be built. You could make a kind of GDPR check with every feature. It is already suspended from the sprint whether you are already compliant. For example, does data minimisation apply to this account at all? And what have I done to minimise that? Something like that?' – Consultant 9

Having **general guidelines** can be sufficient for the managers or consultants to increase the employee's ability. This is how Manager 1 explained about having general guidelines:

'And if I have the general guidelines on how this data should be handled, I think that is sufficient for us.' – Manager 1

The ability of the consultants or managers can also be improved or made more accessible by implementing **privacy modules in the ERP system**. However, Manager 6 explained that implementing privacy modules manually takes a lot of time and effort:

'And it would be good if in the future, from Salesforce or SAP, they would already take that into account. So look at what is the law and regulations? What is privacy? And then offer modules so you don't have to create all that by hand. [..] But that has more to do with logging in and things like that. So single on and stuff like that. But I would recommend that. I think there is much demand for that. ' – Manager 6

Raising privacy awareness is also a recommendation to improve the ability of the consultants and managers. This will increase the knowledge, and as a result, the consultants and managers will be able to implement privacy techniques more easily and without problems.

'They need to follow a course. So they should not be like me, who had partial knowledge. So they should actually start a refresher course every six months. Or a training every six months, and giving them about new developments in these areas, and what has been expected of them. [...]And they want to work on it proactively and finish the certificate. So it should be part and parcel of work.' – Manager 4

The last recommendation that can improve the ability of consultants and managers is to **involve everyone** during the project. For example, Manager 2 explained that it is wise to bring

developers, consultants, and everyone involved in a project together and talk to them to solve privacy problems.

> *'That is one thing anyway, but another thing that is also important is that consultants and developers sit together. Then we can discuss what the challenges are in this area and how we can solve them, what we can and cannot do during the implementation. There is always one person in the company who is in charge of privacy in terms of data protection. That would be my advice.' –Manager*

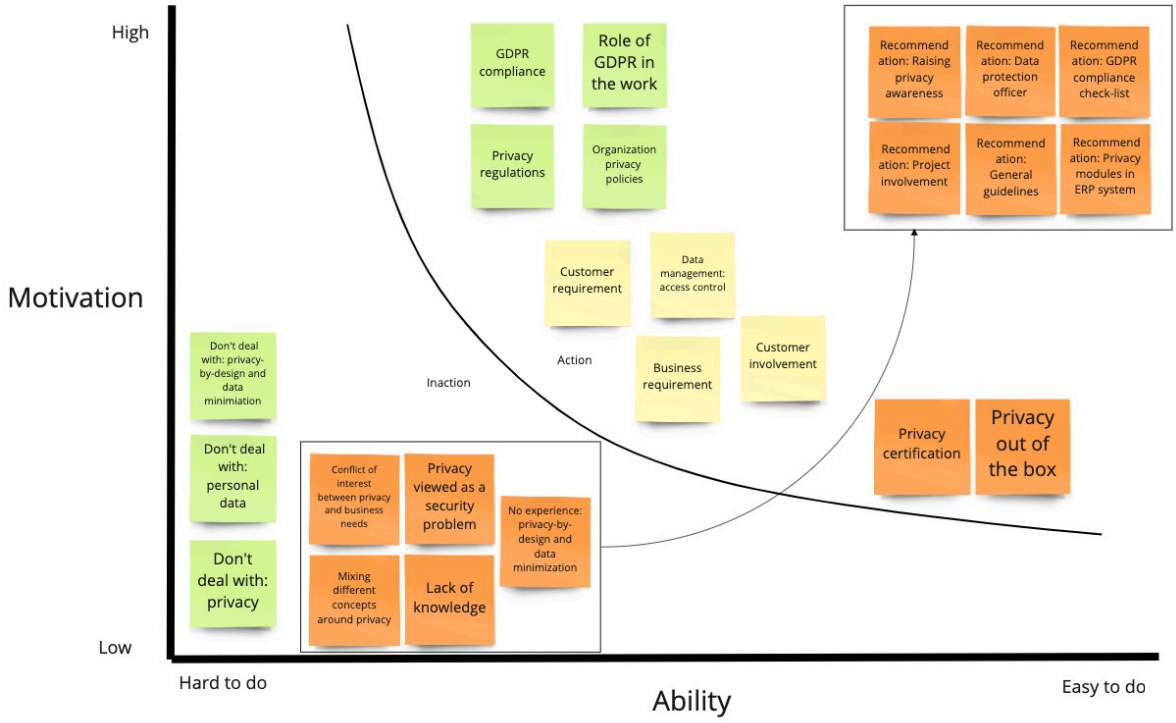

Figure 15 Consultants and Manager Behaviour Model
Green stands for motivation, orange for ability and yellow stickers for triggers

Figure 15 shows the behavioural model of the consultants and managers. Green represents motivation, orange represents ability, and yellow represents triggers, just as in the behavioural model for developers. When the recommendations are applied, the impediments such as lack of knowledge and privacy viewed as a security problem can rise above the action line. Therefore, it is recommended to create general guidelines to improve the lack of knowledge and to ensure that consultants and managers no longer see privacy as a security problem.

5.3 Behaviour Differences

This section answers the last sub-question *"What are the differences between developers' and consultants' behaviour models of privacy-by-design and data minimisation in the ERP systems?"*

Developers, consultants and managers behave almost the same way when it comes to privacy in ERP systems. They are motivated by compliance with the GDPR, the role the GDPR plays in their work and in ERP systems, and the company's privacy policy. Managers and consultants are motivated by privacy regulations, while developers are motivated by compliance in general. However, motivation is low among developers, consultants and managers when it comes to handling personal data and privacy.

Moreover, the developers, consultants and managers lack knowledge, mixing different privacy concepts and see privacy as a security problem. The perception of privacy concerns as IT security problems is consistent with the study by Hadar *et al.* (2018) that most developers view privacy from the perspective of data security, concentrating on technical concerns and security solutions. This makes their ability to implement privacy techniques such as privacy-by-design and data minimisation in the ERP system "*hindered*". On the other hand, consultants and managers do not have experience with privacy-by-design and data minimisation, which makes it very difficult to implement these privacy techniques.

Both developers, consultants and managers state that privacy frameworks or techniques have been implemented, so they do not have to think about privacy frameworks or techniques in the system. This makes it easier for both developers and consultants and managers, but the motivation to implement other privacy techniques, for example, is low. Therefore, they do not think about privacy in the ERP system.

Customer and business requirements trigger the behaviour of developers, consultants and managers. Due to these requirements, developers, consultants and managers have to take measures to minimise data due to database size. Therefore, consultants and managers are triggered to implement access control for data management, while developers are more triggered to minimise data.

Ultimately, both behaviours can be improved by the recommendations given. As mentioned earlier, developers and consultants can reduce the lack of knowledge by raising awareness about privacy. They should not mix up different privacy concepts. Consultants and managers provided more recommendations to improve behaviour compared to developers. For example, one of the managers explains that hiring a data protection officer for a project and implementing privacy modules in the ERP system is recommended. The manager does not have to do everything manually in the system.

6. Discussion

In this chapter, we discuss the outcomes of this research. It begins with a summary of the key findings, including the behaviour, and with a discussion of the key interview findings. Then the research question is answered. Finally, the limitations of this research and future work are discussed.

Summary of key findings

There are already existing literature studies on how developers and managers think about the privacy concepts derived from the GDPR, such as data minimisation and privacy-by-design. However, this research focuses more on the behaviour of the developers and consultants about privacy-by-design and data minimisation in the ERP system. The results of this research have emerged and are almost consistent with the findings from the literature research. In particular, privacy-by-design, it is a not very well known or vague concept that managers need to protect (Jeroen van Rest, 2012). However, the literature also shows that data minimisation is a challenging concept for developers because they are not attuned to the privacy risks posed by the collected data and users' privacy concerns (Stefan Schiffner, 2018). However, the existing research and literature tends to focus more on the general understanding of these privacy techniques rather than how managers and developers behave regarding privacy techniques in the ERP system.

This research shows that privacy-by-design is a concept that consultants and managers do not fully understand. Due to this lack of knowledge, the ability of the developers and consultants to implement this privacy concept in the system is hindered. On the other hand, data minimisation is a well-known concept, but not all consultants or managers have experience with it. It is not their responsibility or task and they do not deal with personal information in the system. The literature proposes a data minimisation methodology for software systems (Awanthika Senarath N. A., 2019). However, the findings of this research shows that neither developers nor consultants adhere to this methodology.

Besides the lack of knowledge, other obstacles hamper the ability of developers and managers more complicated, for example mixing different privacy concepts and considering privacy as a security issue. However, ERP vendors are making this ability more accessible. They have already implemented privacy frameworks, so that developers and consultants do not have to

think about privacy techniques themselves. Furthermore, this research shows that customer and business requirements trigger the behaviour of developers and managers. The business mainly demands operational requirements, such as minimising data due to database size. The customer, on the other hand, demands data minimisation, for example, because they want it. This can be seen as a forced trigger where the developers or consultants have to take an action based on the requirements.

Discussion of interview key findings

For this research, interviews were conducted. The following key findings emerged from the interviews. The first finding is that privacy has become more important due to the GDPR. Before the GDPR, developers and consultants did not really think about privacy in the system, but with the introduction of the GDPR which also imposes requirements, developers and consultants are thinking about privacy. However, privacy is not seen as a top priority. The second finding from the interview is that there is a contradiction between what the business needs in the ERP system and what the privacy rules impose. This leads to confusion among developers and consultants as to what they should do in the ERP system. The final finding is the perception of privacy challenges as IT security challenges such as multifactor authentication, authorisation and access control. The ENISA report (2022) does call for effective technical and organisational measures and controls. However, the perception of privacy challenges as IT security challenges is in line with the literature. Hadar *et al.* (2018) found that most developers view privacy from the perspective of data security, focusing on technical concerns and security solutions.

6.1 Research Question

In this section, we answer the main question and sub-questions of this research.

- **Sub-question 1: 'What is privacy-by-design and data minimisation?'**

The GDPR defines privacy-by-design as a set of principles that can be used directly in the early stages of software development to avoid privacy issues and ensure data protection compliance. However, there is still confusion about what "*privacy-by-design*" means and how it should be applied in practice (InterConsulting, 2018). Data minimisation is a principle that advises minimising the use of personal data in software systems. (Nalin Asanka Gamagedara Arachchilage, 2018). However, it is a difficult privacy concept for developers as they do not know exactly what data they need to protect (Sebastian Wieczorek, 2008).

In **sections 2.1, 2.2, and 2.3** of this thesis, more explanation is given of what data minimisation and privacy-by-design entail and are derived from the privacy legislation GDPR.

- **Sub-question 2: 'How do developers' and consultants' understand privacy-by-design and data minimisation in the ERP system?'**

Some developers and consultants are unfamiliar with or have no experience with "*privacy-by-design."* However, the concept is defined by the managers in such a way that privacy is already taken into account when a system is developed. On the other hand, data minimisation is a concept that both developers and consultants are familiar with, even if most consultants do not have experience with it because it is not their responsibility. Data minimisation is defined as storing as little (personal) data as possible in the system.

More explanation on the developers' and consultants' understanding of privacy-by-design and data minimisation in the ERP system is provided in **section 4.2** of this thesis.

- **Sub-question 3: 'What privacy-by-design and data minimisation strategies and techniques do developers and consultants use in the ERP systems?'**

Section 4.3 of this thesis discusses the strategies and techniques used by the developers and consultants to implement privacy-by-design and data minimisation in the ERP system. Unfortunately, there are not many strategies or techniques that developers and consultants use to implement these privacy techniques. This is because they do not know about it, or it is not their responsibility. However, some strategies or techniques are used by the consultants and developers. Control is one of the privacy-by-design strategies related to Hoepman's (2014) literature research, which identified eight privacy-by-design strategies.

- **Sub-question 4: 'What would help the developers and consultants to increase understanding about privacy-by-design and data minimisation techniques in the ERP system?'**

Consultants and developers have made several recommendations to increase the understanding of privacy-by-design and data minimisation. The most frequently suggested recommendation is to increase privacy awareness by attending training courses or presentations.

Section 4.4 of this thesis explains how to increase the understanding of privacy-by-design and data minimisation in the ERP system.

- **Sub-question 5: ‘What are the differences between developers’ and consultants’ behaviour models of privacy-by-design and data minimisation in the ERP systems?’**

It can be concluded that there is no major difference between the behaviour of developers and consultants in terms of privacy-by-design and data minimisation in the ERP system. On the contrary, both show the same behaviour and are motivated and triggered by, e.g., GDPR requirements and compliance. Therefore, the recommendation can improve the ability of developers as consultants.

The behaviour of developers and consultants, as well as the differences between the two, are discussed in **section 5.3** of this thesis.

- **Main question: ‘What are the developers' and consultants' behaviour models for privacy-by-design and data minimisation in the ERP systems?’**

The aim of this research is to find out the behaviour of developers and consultants regarding privacy-by-design and data minimisation in the ERP system. Not all developers and consultants deal with privacy-by-design or data minimisation in the ERP system, and sometimes they have no experience with it because the system already complies with privacy laws. This is reflected in their behaviour and as a result, there is little motivation to implement these privacy techniques. In addition, obstacles like mixing different privacy concepts or see privacy as a security problem have also emerged that make it difficult to implement privacy techniques in the ERP system. However, these barriers can be reduced and the ability to do so increased through recommendations such as increasing privacy awareness. Chapter 5 Behavioural model of this thesis describes the behaviour of developers and consultants in detail.

6.2 Limitations

There are limitations to this study that need to be discussed. First, the sample size limits the generalisation to a specific group of people. The behaviour of developers, consultants and managers working for a European company in a single EU country has been described. Aspects such as company culture may have an impact on the population's behavioural model, leading in different behavioural models in different companies. A future research duplicating this study across various organisations and countries could address this limitation.

The method of interviewing may have an impact on the quality of the results. The phrasing of the questions, for example, might contribute to respondent bias. This was prevented by phrasing the questions in a neutral manner and having the interview protocol reviewed by other researchers.

Respondents may not fully describe their behaviour or may have intentionally biased results. Wash *et al.* (2017) explain that behaviours are often not correlated with self-reported responses. Each interview was timed to provide respondents with as much information as possible while keeping them motivated. As the respondents had volunteered for this study, they were motivated at the start of the interview. There was no evidence that any of the respondents intentionally said less than they knew or said something different than they thought.

The interviewee has no experience with an interview or does not talk much; they want to get straight to the point. An attempt was made to avoid asking more than one question at a time by using an interview protocol and follow-up questions. Furthermore, the coding method can influence the quality of the results. As mentioned earlier, the reliability of the codebook was validated by two other researchers who did partial double coding.

Another limitation is time. Due to the workload of the researcher and the availability and time of the respondents, the transcripts could not be validated and the second round of follow-up interviews was not conducted. Therefore, it was not possible to double-check the transcripts for any misunderstandings.

7. Conclusion

The main objective of this research was to find out the behaviour of developers and consultants regarding privacy-by-design and data minimisation in the ERP system. According to existing literature studies, privacy-by-design is a vague concept that is difficult for managers to protect (Jeroen van Rest, 2012). On the other hand, data minimisation is a challenging concept for developers because they are not attuned to the privacy risks posed by the collected data and users' privacy concerns (Stefan Schiffner, 2018). This study showed how the developers and consultants behaved in relation to privacy-by-design and data minimisation in the ERP system. The focus was on what motivates developers and consultants and their ability and triggers for performing behaviour. By conducting and analysing interviews, it was possible to determine which elements caused the behaviour.

Both developers and consultants are highly motivated by compliance with GDPR and the organisation's privacy policy. Nevertheless, the ability to implement privacy protection techniques in the system is low due to obstacles such as the fact that privacy is seen as a security issue. This is in line with the literature research where Hadar *et al.* (2018) found that most developers view privacy from the perspective of a data security and focusing on technical concerns and security solutions. Moreover, we can conclude that they are triggered by requirements from both the customer and the business to perform an action, e.g., minimising data. However, developers and consultants can improve their ability by recommending to raise privacy awareness through training. Training can also improve the lack of knowledge, making it easier to implement privacy techniques such as privacy-by-design and data minimisation.

The behaviour of developers and consultants was almost identical. The ability, motivation, and triggers for the behaviour were also very similar. However, only consultants and managers had no experience with privacy-by-design and data minimisation, while developers had no experience with privacy-by-design. In conclusion, this research shows what developers and managers understand about privacy-by-design and data minimisation in ERP systems, and provides recommendations on how developers and consultants can improve their understanding.

7.1 Further Research

From this research, some points have emerged for future research.

- Given the limited sample size population, other researchers can focus on different companies or countries to generate further insights into behaviour. Perhaps consultants or developers in different European countries or companies are motivated or triggered by other factors.
- Interview more people. By interviewing more people, there will be a better understanding of people's behaviour. In this research, sixteen participants were interviewed and further research can be extended to more people, leading to more information and a better perspective of the behaviour.
- More follow-up questions for the interview. Asking more follow-up questions and in-depth questions will lead to more understanding of the behaviour.

Bibliography

A. Stanik, M. H. (2012). Hardware as a Service (HaaS): The completion of the cloud stack. *Computing Technology and Information Management (ICCM),*, 1-8.

Šidlauskas, A. (2021). The Role and Significance of the Data Protection Officer in the Organization. *Vilnius University Press Vol. 44*, 8-28.

AIIM. (2017). *Understanding GDPR Readiness 2017.* Retrieved from https://cdn2.hubspot.net/hubfs/332414/Market-Intelligence/IW.Reports/AIIM-InSight-GDPR-2017%20Final.pdf

Amir Shayan Ahmadian, D. S. (2018). Supporting privacy impact assessment by model-based privacy analysis. *Conference: the 33rd Annual ACM Symposium*, (pp. 1-8).

Andrea Carlson Gielen, D. S. (2003). Application of Behavior-Change Theories and Methods to Injury Prevention . *Epidemiologic Reviews, Volume 25, Issue 1*, 65–76.

Andrianto, A. (2019). Impact of Enterprise Resource Planning (ERP) implementation on user performance: studies at University of Jember. *Journal of Physics Conference Series 1211*, 1-9.

Angela Lina, N.-C. C. (2012). Cloud computing as an innovation: Percepetion, attitude, and adoption. *International Journal of Information Management Volume 32, Issue 6*, 533-540.

Anne Adams, M. A. (1999). USERS ARE NOT THE ENEMY . *Communications of the ACM* .

Atlas.ti. (2020). *Inter-Coder Agreement Analysis.*

Autoriteitpersoonsgegevens. (n.d.). *Wetten*. Retrieved from Over privacy wetten : https://autoriteitpersoonsgegevens.nl/nl/over-privacy/wetten

Awanthika Senarath, N. A. (2018). Why developers cannot embed privacy into software systems?: An empirical investigation. *Conference: the 22nd International Conference*, 1-6.

Awanthika Senarath, N. A. (2019). A data minimization model for embedding privacy into software systems. *Computers & Security Volume 87*, 23-30.

Barney G. Glaser and Anselm L, S. (1967). The Discovery of Grounded Theory. In B. G. Strauss, *The Discovery of Grounded Theory Strategies for Qualitative Research* (p. 284). New Brunswick (U.S.A.) and London (U.K.): Aldine Publishing Company.

Batya Friedman, H. N. (2002). Users' conceptions of risks and harms on the Web: A comparative study. *CHI: changing the world, changing ourselves*, 2.

Baxevani, T. (2019, May). GDPR Overview.

Björn Johansson, P. R. (2013). Exploring Factors for Adopting ERP as SaaS. *Procedia Technology Volume 9*, 94-99.

Brady., V. B. (2005). Success and failure factors of adopting SAP in ERP system implementation. *Business Process Management Journal, 11*, 501-516 .

Camp, L. J. (2009). Mental Models of Privacy and Security. *IEEE Technology and Society Magazine 28*, 37-46.

Cathryne Palmer, A. B. (2006). A Brief Introduction to Qualitative Research. *The Canadian journal of medical radiation technology*, 16-19.

Cavoukian, A. (2011). *The 7 Foundational Principles - Privacy by Design.* Ontario, Canada.

Christopher Kuner, F. H. (2012). The challenge of ‘big data’ for data protection . *International Data Privacy Law, Volume 2, Issue 2*, 47-49.

Christopher P. Holland, B. L. (1999). A Critical Success Factors Model for ERP Implementation. *IEEE Software 16*, 30-36.

Clarke, R. (2009). Privacy impact assessment: Its origins and development. *Computer Law & Security Review Volume 25 Issue 2*, 123-135.

Cliff W. Scott, M. M. (2017). Axial Coding. In M. M. Cliff W. Scott, *Axial Coding* (pp. 1-8). University of North Carolina at Charlotte, USA: The International Encyclopedia of Communication Research Methods.

Coe, M. J. (2011). Using Enterprise Resource Planning Systems As The Core Of An Integrated Accounting Information Systems Course. *Review of Business Information Systems (RBIS)* , 1-10.

Colonna, L. (2013). MO' DATA, MO' PROBLEMS? PERSONAL DATA MINING AND THE CHALLENGE TO THE DATA MINIMIZATION PRINCIPLE. *Stanford Law School and The Center for Internet and Society* .

Cong, L. K. (2010). Using Transactional Data from ERP Systems for Expert Finding. *DEXA 2010: Database and Expert Systems Applications* , 267–276.

Daniel Mikkelsen, H. S.-J. (2019, July 22). GDPR compliance since May 2018: A continuing challenge. *McKinsey & Company*, pp. 1-2.

Danilo Krivokapić, D. K. (2018). Impact of GDPR on Business: Focus on Data Controllers and Processors not Established within the EU . *37TH INTERNATIONAL CONFERENCE ON ORGANIZATIONAL SCIENCE DEVELOPMENT:* , 1-13.

Darnton, A. (2008). *Reference Report: An overview of behaviour change models and their uses* .

Davenport, T. H. (1998). Putting the enterprise into the enterprise system. *Havard Business Review*, 1-12.

Debin Liu, F. A. (2008). Risk Communication in Security Using Mental Models. *Usable Security*, 1-12.

Deepak Kumar Verma, R. K. (2018). Cloud computing security: A Review. 1-6.

Dennis P. Slevin, J. K. (2015). Balancing Strategy and Tactics in Project Implementation. *Sloan Management Review 29*, 33-41.

Dode, A. (2018). The challenges of implementing General Data Protection Law (GDPR). *14th International Conference*, p. 8.

E.M. Shehab, M. S. (2004). Enterprise resource planning. *Business Process Management Journal 10(4)*, 359-386.

Elahe Mahdizadeh, S. M. (2016). A Comprehensive Literature Review of Enterprise Resource Planning over Cloud Computing Infrastructure. *2nd International Conference on Management and Social Science*, 1-9.

ENISA. (2022). *DATA PROTECTION ENGINEERING from theory to practice.*

Eugenia Politou, A. M. (2018). Backups and the right to be forgotten in the GDPR: An uneasy relationship. *Computer Law & Security Review Volume 34 Issue 6* , 1247-1257.

Europa.eu. (2017). *Guidelines on Data Protection Officers ('DPOs').*

Europa.eu. (2021). *Data protection under GDPR*. Retrieved from https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/index_en.htm

European Union data protection. (2016). *Data protection.*

Farida Habib Semantha, S. A. (2020). A Systematic Literature Review on Privacy by Design in the Healthcare Sector . *Electronics 9(3):452*, 1-29.

Farzaneh Asgharpour, D. L. (2007). Mental Models of Security Risks. *Financial Cryptography and Data Security, 11th International Conference,*, 367–377.

Fawad Taj, M. C. (2019). Digital Health Behavior Change Technology: Bibliometric and Scoping Review of Two Decades of Research. *JMIR Mhealth Uhealth*, 1-20.

Fogg, B. (2009). A Behavior Model for Persuasive Design. *Persuasive Technology Lab*, 1-7.

George Danezis, J. D.-F.-H. (2014). *Privacy and Data Protection by Design – from policy to engineering.* Enisa.

Glenn Stewart, M. M. (2000). Organisational Readiness for ERP Implementation. *Information Systems Management Research Centre*, 1-6.
Hasan, M. T. (2018). Impact of ERP System in Business Management. *International Journal of Management Studies* , 24-31.
Hasan, T. (2018). Impact of ERP System in Business Management. *International Journal of Management Studies V(4(4)):24*, 24-31.
Helmut Klaus, M. R. (2000). What is ERP? *Information Systems Frontiers 2*, 141-162.
Hoepman, J.-H. (2014). Privacy Design Strategies (The Little Blue Book). *ICT Systems Security and Privacy Protection*, 445-459.
Holton, J. A. (2010). The Coding Process and Its Challenges. *Grounded Theory Review: An International Journal*, 21-40.
Hu, M. (2020). Cambridge Analytica's black box. *Big Data & Society 7(2)*, 1-6.
IBM Cloud Education. (2021, September 2). *IaaS vs. PaaS vs. SaaS*. Retrieved from https://www.ibm.com/cloud/learn/iaas-paas-saas
ico.org. (2018). *An overview of the Data Protection Act 2018.*
ico.org. (2018). *Data Protection Officer (DPO).* Retrieved from https://edps.europa.eu/data-protection/data-protection/reference-library/data-protection-officer-dpo_en
ico.org. (2018). *What are 'controllers' and 'processors'?*
Ike C.Ehie, M. M. (2005). Identifying critical issues in enterprise resource planning (ERP) implementation. *Computers in Industry, Volume 56 , Issue 6*, 545-557.
IMAP. (2010). *Computing & Internet Software Global Report.* Retrieved from https://www.yumpu.com/en/document/read/40914753/computing-internet-software-global-report-a-2010-imap
InterConsulting. (2018). *Art. 33 GDPR Notification of a personal data breach to the supervisory authority.*
InterConsulting. (2018). *Art. 5 GDPR Principles relating to processing of personal data.* InterConsulting.
InterConsulting. (2018). *GDPR Privacy by Design.*
InterConsulting. (2018). *Individual Rights.*
Irit Hadar, T. H. (2018). Privacy by designers: software developers' privacy mindset. *Empir Software Eng*, 259–289.
Isaiah Liljestrand, M. G. (2019, April). Developing a Mental Model for use in the Context of Computer Security. *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, pp. 2336–2339.
Jan-Kyrre Berg Olsen, E. S. (2009). New Waves in Philosophy of Technology. In E. S. Jan-Kyrre Berg Olsen. New York: Palgrave.
Jeroen van Rest, D. B. (2012). Designing Privacy-by-Design. *Privacy Technologies and Policy* , 55-72.
Jessica T. DeCuir-Gunby, P. L. (2011). Developing and Using a Codebook for the Analysis of Interview Data: An Example from a Professional Development Research Project. *Field Methods (23)*, 136-155.
Jim Blythe, L. J. (2012). Implementing Mental Models. *EE Symposium on Security and Privacy Wo*, 86-90.
John, R. (2022, February 4). *Improve User Engagement with the BJ Fogg Behavior Model*. Retrieved from Productled: https://productled.com/blog/the-bj-fogg-behavior-model/
Juliet Corbin, A. S. (1990). Basics of Qualitative Research. In A. S. Juliet Corbin, *Basics of Qualitative Research - Techniques and Procedures for Developing Grounded Theory* (p. 434). United States of America: SAGE Publications, Inc.

Kakoli Bandyopadhyay, C. B. (2012). An Analysis of Factors Affecting User Acceptance of ERP Systems in the United States. *International Journal of Human Capital and Information Technology Professionals, 3(1)*, 1-14.
Kathrin Bednara, S. S. (2018). Engineering Privacy by Design: Are engineers ready to live up to the challenge? *Information Society, 35(3)*, 122-142.
Kelly D. Martin, A. B. (2016). Data Privacy: Effects on Customer and Firm Performance. *Journal of Marketing 81(1)*, 36-58.
Kneuper, R. (2019). Integrating Data Protection into the Software Life Cycle. *Product-Focused Software Process Improvement, 20th International Conference*, 417-432.
Krippendorff, K. (2015). *Inter-coder agreement in ATLAS.ti.*
Lachaud, E. (2018). Certification of Data Protection Officers Should Be Regulated. *SSRN Electronic Journal*, 1-16.
Laslo Šereš, T. P. (2014). ERP & Globalization: Challenges and Responses. *Strategic Management 19(4)*, 50-54.
Linying Dong, D. N. (2009). Top Management Support of Enterprise Systems Implementations. *Journal of Information Technology 24(1)*, 55-80.
M. Granger Morgan, B. F. (2002). Risk Communication: A Mental Models Approach. *Cambridge University Press*, 1-13.
Maggie Oates, Y. A. (2018). Turtles, Locks, and Bathrooms: Understanding Mental Models of Privacy Through Illustration. *Proceedings on Privacy Enhancing Technologies*, 5-32.
Maike Vollstedt, S. R. (2019). An Introduction to Grounded Theory with a Special Focus on Axial Coding and the Coding Paradigm. *Compendium for Early Career Researchers in Mathematics Education* , 81-100.
Majed Alshammari, A. S. (2018). Towards an Effective Privacy Impact and Risk Assessment Methodology: Risk Assessment. *Trust, Privacy and Security in Digital Business* , 85-99.
Marc Pelteret, J. O. (2019). A Review of Information Privacy and Its Importance to Consumers and Organizations. *Informing Science: the International Journal of an Emerging Transdiscipline*, p. 25.
Marie Caroline Oetzel, S. S. (2014). A systematic methodology for privacy impact assessments: A design science approach. *European Journal of Information Systems 23(2)*, 1-25.
Marie Oetzel, S. S. (2013). A systematic methodology for privacy impact assessments - a design science approach. *European Journal of Information Systems (EJIS) Volume 23*, 126-150.
Mary J Culnan, C. C. (2009). How ethics can enhance organizational privacy: lessons from the ChoicePoint and TJX data breaches. *MIS Quarterly Vol. 33, No. 4*, 673-687.
McAfee. (2017). *Grand Theft Data: Data exfiltration study: Actors, tactics, and detection.*
McKinsey Company. (2020). *The Next Normal: Digitizing at speed and scale.* McKinsey Global Publishing.
Mezghani, K. (2019). From On-Premise ERP to Cloud ERP. In K. Mezghani, *Advanced Methodologies and Technologies in Business Operations and Management* (pp. 816-825). IGI GLOBAL.
Michael Colesky, C. H. (2016). A Critical Analysis of Privacy Design Strategies. *IEEE Security and Privacy Workshops*, 1-8.
Michael Williams, T. M. (2019). The Art of Coding and Thematic Exploration in Qualitative Research. *International Management Review Vol. 15 No. 1*, 1-11.

Mohsen Attaran, J. W. (2018). Cloud computing technology: improving small business performance using the Internet. *Journal of Small Business & Entrepreneurship 13*, 94-106.
Nalin Asanka Gamagedara Arachchilage, A. S. (2018). Understanding Software Developers' Approach towards Implementing Data Minimization. 1-5.
Natalie A. Jones, H. R. (2011). Mental models: an interdisciplinary synthesis of theory and methods. *Ecology and Society 16(1): 46*, 1-13.
Nigel Mathers, N. J. (2000). Using Interviews in a Research Project. In N. J. Nigel Mathers, *Research Approaches in Primary Care* (pp. 1-30). United States: Radcliffe Medical Press/Trent Focus.
Nishad Nawaz, K. C. (2013). The Impact of Enterprise Resource Planning (ERP) Systems Implementation on Business Performance . *SSRN Electronic Journal*, 30-44.
NOREA. (2015). *Privacy Impact Assessment (PIA).*
Onno Tene, J. P. (2011). Privacy in the age of big data: a time for big decisions. *Stanford Law Review Online, 64*, 57-64.
P. Trott, A. H. (2011). Enterprise Resource Planning and its Impact in the Innovative Capability of the Firm. *International Journal of Innovation Management*, 257-270.
Paul De Herta, V. P. (2012). The proposed data protection Regulation replacing Directive95/46/EC: A sound system for the protection of individuals. *Computer Law & Security Review 28(2)*, 130-142.
Paula Braveman, L. G. (2014). The Social Determinants of Health: It's Time to Consider the Causes of the Causes. *Public Health Rep*.
Pragati Chavan, G. K. (2013). PaaS Cloud. *IROCS Journals*, 1-5.
Punch, K. F. (2014). Introduction to Social Research. In K. F. Punch, *Introduction to Social Research Quantitative and Qualitative Approaches* (p. 408). University of Western Australia, Australia: SAGE PUBLICATIONS.
Renata M. de Carvalho, C. D. (2020). Protecting Citizens' Personal Data and Privacy: Joint Effort from GDPR EU Cluster Research Projects. *SN Computer Science 1(4)*, 1-16.
Rick Wash, E. R. (2015). Too Much Knowledge? Security Beliefs and Protective Behaviors Among United States Internet Users. *Proceedings of the 11th Symposium On Usable Privacy and Security,*, 309-325.
Rick Wash, E. R. (2017). Can People Self-Report Security Accurately?: Agreement Between Self-Report and Behavioral Measures. *Can People Self-Report Security Accurately?: Agreement Between Self-Report and Behavioral Measures*, 2228–2232.
Ruogu Kang, L. D. (2015). My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security. *Symposium on Usable Privacy and Security*, 39-52.
S Sheth, G. K. (2014). Us and them: a study of privacy requirements across North America, Asia, and Europe. *Proceedings of the 36th International Conference on Software Engineering*, 1-12.
Samantha Mathara Arachchi, S. C. (2015). Quality Assurance and Quality Control in ERP Systems Implementation. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS)*, 70-83.
Samuel D. Warren, L. D. (1990). The Right to Privacy. *Harvard Law Review, Vol. 4, No. 5.*, pp. 193-220.
Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing Health Volume 18, Issue 2*, 179-183.
Sandra C Henderson, C. A. (1999). Personal information privacy: implications for MIS managers. *Information & Management*, 213-220.

Sauvik Das, H. J. (2014). The Effect of Social Influence on Security Sensitivity. *Proceedings of the 10th Symposium on Usable Privacy and Security*.

Sauvik Das, L. A. (2019). A Typology of Perceived Triggers for End-User Security and Privacy Behaviors. *Proceedings of the Fifteenth Symposium on Usable Privacy and Security*, 1-20.

Sebastian Wieczorek, A. S. (2008). Test data provision for ERP systems. *1st International Conference on Software Testing*, 1-9.

Sejdija, Q. (2016). Motivation an Important Part in Management. *Academic Journal of Interdisciplinary Studies Vol 5 No 3*, 373 - 377.

Serge Egelman, D. M. (2010). Please Continue to Hold An empirical study on user tolerance of security delays. *Workshop on the Economics of Information Security (WEIS '10)* .

Shaiti, S. A. (2020). The Impact of Enterprise Resource Planning on Business Performance: With the Discussion on Its Relationship with Open Innovation. *Journal of Open Innovation: Technology, Market and Complexitity*, 1-24.

Spiekmann, S. (2012, July). Viewpoint: The challenges of Privacy-by-Design. *Communications of the ACM 55*.

Sposit, N. (2018). Adapting to Digital Marketing Regulations: An Exploratory Analysis of theGDPR and its Effects on Individualized, Behavior-Based MarketingTechniques. p. 9.

Stefan Schiffner, B. B.-F. (2018). Towards a Roadmap for Privacy Technologies and the General Data Protection Regulation: A Transatlantic Initiative. *Privacy Technologies and Policy*, 24-42.

Subhi R. M. Zeebaree, B. w. (2020). Enterprise Resource Planning Systems and Challenges. *Technology Reports of Kansai University 62(4):*, 1885-1894.

Sune Dueholm Müller, S. R. (2015). Benefits of Cloud Computing: Literature Review in a Maturity Model Perspective. *Communications of the Association for Information Systems 37*, 852-878.

Susan Michie, M. M. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 1-12.

Susanne Kießling, T. H. (2021). Salt&Pepper: Spice up Security Behavior with Cognitive Triggers. *EICC*, pp. 1-6.

Tankard, C. (2016). What the GDPR means for businesses. *Network Security*, pp. 5-8.

The European Data Protection Supervisor. (2018). *Preliminary Opinion on privacy by design.*

Vayyavur, R. (2015). ERP Implementation Challenges & Critical Organizational Success Factors. *International Journal of Current Engineering and Technology Vol. 5*, 2759-2766.

Verizon. (2020). *Data Breach Investigations Report 2020.*

Voss, W. G. (2012). Preparing for the Proposed EU General Data Protection Regulation: With or without Amendments. *Business Law Today* , 1-5.

Wash, R. (2010). Folk Models of Home Computer Security. *Proceedings of the Sixth Symposium On Usable Privacy and Security*, 14-16.

Wright, D. (2012). The state of the art in privacy impact assessment. *Computer Law & Security Review Volume 28*, 54-61.

Yang Wang, A. K. (2009). Privacy-Enhancing Technologies. 1-23.

Ylona Chun Tie, M. B. (2019). Grounded theory research: A design framework for novice researchers. *SAGE Open Med.*, 1-5.

Zhenyu Huang, P. P. (2001). ERP implementation issues in advanced and developing countries. *Business Process Management Journal;*, 276-284.

Appendix I – Interview protocol

The interview protocol is provided in both **English** and **Dutch** in this appendix. The purpose of an interview protocol is to structure the interview by preparing questions.

Context:

- Introduction about the purpose of the research;
- The goal of the interview;
- Short description of how the data will be stored and processed:
    - Participants' names and audio records will be stored in a secure location;
    - Anonymized transcripts will be analysed and stored for the research purposes and may be shared as part of the research;
- Permission to audio record the interview.

Demographics & Background:

- What is your education level?
- What is your current position?
- How many years in the current position?
- What is your total work experience?
- Which ERP system are you currently working with?
- Do you have any privacy-related certifications?

Privacy-related questions:

1. What is privacy?
2. Have you heard about the GDPR? Can you explain in your own words what this is?
3. What is privacy-by-design?
4. What is data minimisation?
5. What are your responsibilities related to privacy and data protection?

ERP-related questions:

6. What are the privacy goals of the ERP system?
    6.1. How often do you review the ERP systems compliance with the privacy goals?
7. What have been the most critical privacy challenges for the ERP systems?
    7.1. What are the consequences of these challenges for the ERP systems?
    7.2. What is the role of GDPR in these challenges?
8. What is your experience with privacy-by-design and data minimisation in the ERP system?
    8.1. Have you encountered privacy-by-design and data minimisation challenges or problems while working with the ERP system?
9. What strategies and techniques for implementing privacy-by-design in the ERP system do you use?
    9.1. How does it work in detail? Can you draw on paper how such techniques or strategies are implemented?

10. What strategies and techniques for implementing data minimisation in the ERP system do you use?
    10.1. How does it work in detail? Can you draw on paper how such techniques or strategies are implemented?
11. What would you suggest to companies to improve developers/consultants' understanding of privacy-by-design and data minimisation?

Dutch Interview Protocol

- Mijn scriptie gaat over de mentale modellen van consultants en EPR-ontwikkelaars over privacy-by-design en dataminimalisatie in het ERP-systeem.
- Een mentaal model is een uitleg van iemands denkproces over hoe iets werkt in de echte wereld, dus in dit geval voor mijn scriptie is het denkproces van ontwikkelaars en consultants over privacy-by-design en data-minimalisatie in het ERP-systeem.
- Het doel van dit interview is om uit te vinden wat u denkt over privacy-by-design en data-minimalisatie in het ERP-systeem.
- De gegevens van het interview worden opgeslagen en verwerkt. Dit betekent dat uw geluidsopnamen op een veilige plaats op mijn computer worden opgeslagen. Geanonimiseerde transcripties zullen worden geanalyseerd en opgeslagen voor onderzoeksdoeleinden en kunnen worden gedeeld als onderdeel van het onderzoek. Is het goed als ik het opneem?
- Dit interview bestaat uit drie delen, eerst zal ik vragen naar je achtergrond informatie dan zal ik privacy gerelateerde vragen stellen en als laatste zal ik ERP-systeem vragen stellen.

Demografie en achtergrond:

- Wat is uw opleidingsniveau?
- Wat is uw huidige functie? Hoe lang?
- Wat is uw totale werkervaring?
- Met welk ERP-systeem werkt u momenteel?
- Heeft u enige privacy-gerelateerde certificeringen? Heeft u enige training of workshop gevolgd die gerelateerd is aan privacy?

Privacy-related questions:

1. Wat is privacy?
2. Hebt u gehoord over de GDPR – AVG Algemene Verordening Gegevensbescherming (GDPR)? Kun je in je eigen woorden uitleggen wat dit is?
3. Wat is privacy-by-design?
4. Wat is data minimization?
5. Wat zijn uw verantwoordelijkheden met betrekking tot privacy en gegevensbescherming?

ERP-related questions:

6. Wat zijn de privacydoelen van het ERP-systeem?
a. Hoe vaak beoordeelt u of het ERP-systeem voldoet aan de privacydoelstellingen?
7. Wat zijn de meest kritieke privacy-uitdagingen voor de ERP-systemen geweest?
. Wat zijn de gevolgen van deze uitdagingen voor de ERP-systemen?
a. Wat is de rol van GDPR in deze uitdagingen?

8. Wat is uw ervaring met privacy-by-design en dataminimalisatie in het ERP-systeem?
. Bent u problemen tegengekomen met privacy-by-design en dataminimalisatie tijdens het werken met het ERP-systeem?
9. Welke strategieën voor het implementeren van privacy-by-design in het ERP-systeem gebruikt u? (Wat zijn de requirements van de klant betreft privacy?)
. Hoe werkt dat in detail? Kunt u op papier zetten hoe dergelijke technieken of strategieën worden toegepast?
10. Welke strategieën voor het implementeren van dataminimalisatie in het ERP-systeem gebruikt u?
. Hoe werkt dat in detail? Kunt u op papier zetten hoe dergelijke technieken of strategieën worden toegepast?
11. Wat zou u voorstellen aan bedrijven om ontwikkelaars/consultants meer inzicht te geven in privacy-by-design en dataminimalisatie?
12. Bent u beschikbaar voor vervolgvragen?
13. Kent u nog meer mensen die ik wellicht kan interviewen?

Appendix II – Informed consent form

This appendix contains the informed consent form for conducting interviews. Before each interview, the respondent must approve the form. There is a short informed consent form and a detailed one.

Mental models of developers and consultants of privacy-by-design and data minimisation in ERP systems

Short informed consent form

This is a short information consent form. You can find a more detailed version on the next page.

This research is conducted as a part of Master thesis research project of Alicia Pang. This research studies the differences between consultants and developers in the perception of privacy-by-design and data minimisation in ERP systems. In this interview, you will be asked several questions about your experience and knowledge on this topic. The interview does not aim to evaluate your knowledge, but rather to capture what you think about privacy-by-design and data minimisation in ERP systems.

We collect the following data (**private data**):

- Your name
- Your current employment position, number of years of professional experience, and your education background
- Your email address

Only the research team has access to your private data. It is stored securely in a password-protected folder; no online back-up exists for this data. The most sensitive private data (name and email address collected digitally) will be deleted once the research is complete and its findings are published.

If you have signed a physical informed consent form (this form), your name is also retained in the signed form. The signed consent forms will be stored in a locked cabinet at Leiden University for at least 10 years, unless a permission is received to destroy them from the Ethics Review Committee of the Faculty of Science.

The interview will be recorded and transcribed for research analysis purposes. All further processing of the interviews (media files, transcript texts, transcript analysis files) works with anonymized data. Your name and your email address do not appear in this anonymized data. Only the research team has access to the media files, transcripts, and transcript analysis. We will publicly share the **anonymized** interview transcripts at the Zenodo service (zenodo.org). They will be available there for 10 years or more.

The **fully anonymized** findings from this research will be made publicly available as a Master thesis report at the Leiden University website, and as research publications.

A selection of direct quotes from the interview can be used in research publications, but only **without attributing the source**.
Your employment history (current position and the number of years of professional experience) and education background can be shared in research publications in an **aggregated or anonymized form**.

You may withdraw yourself from the study at any time, without giving a reason and without penalty, by communicating your decision with the researcher.

Mental models of developers and consultants of privacy-by-design and data minimisation in ERP systems

Detailed informed consent form

1. **Background and aims of the study**
   The purpose of the research is to investigate the mental models of developers and consultants about privacy-by-design and data minimisation in ERP systems.
   The research aims to empirically investigate the following questions:

1. What is privacy-by-design and data minimisation?
2. How do developers' and consultants' understand privacy-by-design and data minimisation in the ERP system?
3. What are the differences between developers' and consultants' mental models of privacy-by-design and data minimisation in the ERP systems?
4. What privacy-by-design and data minimisation strategies and techniques do developers and consultants use in the ERP systems?
5. What would help the developers and consultants to increase understanding about privacy-by-design and data minimisation techniques in the ERP system?

2. **Do I have to take part?**
   You can ask questions about the study before deciding whether or not to participate. If you do agree to participate, you may withdraw yourself from the study at any time, without giving a reason and without penalty, by advising the researcher of this decision.

3. **What will happen in the study?**
   If you have agreed to take part in the study, you will be asked to participate in a 60-minute semi-structured interview in Dutch or English. The researcher will travel to a location of the participant's choice or conduct the interview by Teams. This depends on the interviewee, the researcher is available for two options, Teams or physically in the office.
   An audio recorder will be used to record the interviews and transcribe them after the interview has taken place.

4. **Are there any potential risks in taking part?**
   The study will focus on the mental models of consultants and developers on the privacy-by-design and data minimisation in the ERP systems. This may reveal potential weaknesses within the ERP system in terms of privacy and the privacy knowledge that is lacking within the consultants and developers. We aim to address these concerns by:

- Protecting the responses from being made public by keeping them securely and confidentially.
- No access for anybody outside the research team to interview answers or notes that were made during the interview.
- The private data collected and the interview media files will be deleted at the end of the study.

5. **Are there any benefits in taking part?**
   The researcher asks privacy-related questions and this may make the participant more aware of the privacy in the ERP system.

   In order to collaborate in future studies, we will seek to make our results as useful to the participants as possible. If interested, the participants can request to be provided with the final research text.

6. **What happens to the data provided?**
   The mapping table with the most sensitive private data (names, email addresses and the identifiers used for the data analysis purposes) is stored on the researcher's computer in a password-protected folder. This data is not sent over the internet and it is not stored in a cloud environment. This table will be deleted after the research is completed and published.

The identifiers are short strings (e.g., "ID1") that do not contain any elements of your name, and that are used to uniquely identify your transcript among others, but they cannot be used to link the transcript back to you without the mapping table.

The interview transcripts (anonymized text files using the identifiers) will be securely stored at a server within the Leiden University environment. The server has regular backups. Nobody except the researcher and the supervisors will have access to personal/sensitive data/research data. The anonymized interview transcripts will be publicly shared at the Zenodo service (https://about.zenodo.org/policies/; this service stores data in Switzerland and Hungary). The retention period of this anonymized data will be at least 10 years.

During the interview we will not ask for your name, except if discussing the signed consent form. You will merely be asked for your organisation, role within the organisation, and working and educational experience. As this can be used to identify certain individuals, the audio files of the interviews will be treated as personal data and be stored securely in a password-protected folder. After this research is finished, the audio files will be deleted.

We ask all participants for their permission to use direct quotes; these will be attributed to the role of the participant and a description of the company. E.g. "An ERP developer says that he finds it difficult to implement privacy-by-design in the ERP system because there are no clear rules on how we should do this as a developer".

All anonymized research data and records will be stored for a minimum retention period of 10 years[2] after publication or public release of the work of the research. Your digital private data (name, email address) will be deleted after the end of the study. The physical (paper) informed consent forms will be stored in a locked cabinet at Leiden University for at least 10 years, unless permission is granted by the Ethics Review Committee to destroy them. After 10 years, or if permitted by the Ethics Review Committee, they will be securely destroyed.

7. **Will the research be published?**
Example in case of multiple organisations participating in the study:
To protect the participants' anonymity, we will:
   a. Describe the organisations that took part in the study, rather than use individual names
   b. A selection of direct quotes from the interview will be used without attributing the source, but permission will be sought first.

If you agree to participate in this study, the research will be written up as a Master's thesis, published in a public repository.

8. **Who has reviewed this study?**
This study has been reviewed by and received ethics clearance through the Ethics Review Committee of the Faculty of Science (reference number: XXXX)

9. **Who do I contact if I have a concern about the study or I wish to complain?**
If you have a concern about any aspect of this study, please speak to the principal investigator Dr. Olga Gadyatskaya or the researcher Alicia Pang. The research team should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the relevant chair of the Ethics Review Committee of the Faculty of Science who will seek to resolve the matter in a reasonably expeditious manner:

[2] In accordance with the Leiden University Research Data Framework Policy. Available online: https://www.library.universiteitleiden.nl/binaries/content/assets/ul2ub/research--publish/research-data-management-regulations-leiden-university_def.pdf

Chair, Ethics Review Committee of the Faculty of Science, Email: ethicscommittee@science.leidenuniv.nl

10. Contact Details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Principal Investigator:
Dr. Olga Gadyatskaya (o.gadyatskaya@liacs.leidenuniv.nl)
Assistant professor, LIACS, Leiden University
Researcher:
Alicia Pang (s2785412@vuw.leidenuniv.nl)

Please tick the appropriate boxes	**Yes**	**No**
Taking part in the study		
I have read and understood the study information dated 22/10/2021, or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	☐	☐
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.	☐	☐
Use of the information in the study		
I understand that information I provide will be used for the Master thesis report and publications in academic venues (like conferences or journals).	☐	☐
I understand that personal information collected about me that can identify me, such as my name or email address, will not be shared beyond the study team.	☐	☐
I agree that my opinions can be anonymously quoted in research outputs	☐	☐
Future use and reuse of the information by others		
I give permission for the anonymized interview transcript that I provide to be archived in Zenodo (https://zenodo.org/) so it can be used for future research and learning.	☐	☐
Signatures		
_____ _____ _____ Name of participant Signature Date		
I have accurately read out the information sheet to the potential participant, or I have witnessed the accurate reading of the consent form by the participants and, to the best of my ability, ensured that the participant understands what they are freely consenting to. _____ _____ _____ Researcher name Signature Date		

Appendix III – Codebook

Challenges
8 Codes:
○ **Data analytics external**
Comment:
Code name: Data analytics external
Description: Data analytics is aimed at analysing data, with the aim of gaining insight into the processes to which that data relates.
Inclusion criteria: the statements discusses analysing data for external parties such as companies. For example, companies dump all the data to determine what sells best, who buys what.
Exclusion criteria: -

○ **Depends on the project**
Comment:
Code name: Depends on the project
Description: Every project is different and this can lead to different privacy responsibilities of consultants and developers. But also the privacy requirements/strategies for implementing in projects.
Inclusion criteria: These statements discuss it in general terms about implementing strategies/techniques and the tasks of the developers or managers are related to the project. For example, this is also very dependent on the project. If we have an implementation from scratch we have blueprints of how we think about it.
Exclusion criteria: do not refer to specific tasks/responsibilities OR the specific process for implementing privacy in the ERP system.

○ **Dynamic information**
Comment:
Code name: Dynamic information.
Description: This code means it is difficult to keep track of changing information in light of GDPR compliance.
Inclusion criteria: the statements that discuss information changes and at the same time to be GDPR compliant. For example, Information is constantly changing in the system and it is difficult to follow GDPR.
Exclusion criteria:

○ **ERP privacy challenge**
Comment:
Code name: ERP Privacy Challenge.
Description: In the ERP system there are some privacy challenges that consultants and developers experience.
Inclusion criteria: the statements that discuss ERP privacy challenges. It must be related to the ERP system. For example, that is the challenge for us to classify those roles.
Exclusion criteria: these statements discuss only ERP privacy challenges and NOT the limited frameworks.

○ **ERP systems integration**
Comment:
Code name: ERP systems integration
Description: Someone that discusses that you need to communicate with other systems in order to collect data. This can cause a privacy challenge.
Inclusion criteria: the statement discuss that ERP systems or other system need to integrate. This can be a privacy challenge. For example, most of the ERPs are going into integration mode, so they start terribly.
Exclusion: -

○ **GDPR challenge**
Comment:
Code name: GDPR Challenge
Description: With the introduction of GDPR, this can pose a challenge for consultants and developers.
Inclusion criteria: the statements that discuss the challenges posed by GDPR. For example, the GDPR has added a challenge such as complying with privacy rules.
Exclusion criteria: -

○ **No ERP privacy challenge**
Comment:
Code name: No ERP privacy challenge.
Description: There are no privacy challenges in the ERP system.
Inclusion criteria: the statements discuss that there are no privacy challenges in the ERP system. For example, it wasn't really a challenge for us.
Exclusion criteria: -

○ **Processing data outside ERP system**
Comment:
Code name: Processing data outside ERP system.
Description: Data can be processed outside the system.
Inclusion criteria: the statements that discuss processing data outside the ERP system. For example, Excel sheet with personal information is shared outside the system. This is not in the system.
Exclusion criteria: -

◈ **Data management**
11 Codes:

○ **Data classification**
Comment:
Code name: Data classification
Description: Data can be classified into different categories/levels in the ERP system. Classifying data can also ensure that not everyone has access to data, so there is privacy protection.
Inclusion criteria: the statements that discuss data classification. More an explanation of what data classification is and what the consequences are. For example, [..] so every data field captured in every table is categorised as in seven different levels.
Exclusion criteria:

○ **Data control**
Comment:
Code name: Data control
Description: There is control over the data in the system. A very general code that gives transparency to the data collection. Focus more on safeguarding and protecting data.
Inclusion criteria: the statements that discuss data control. More focus on the control in the data management. For example, it brings a lot of transparency on how the organisation's do data collection.
Exclusion criteria:

○ **Data management: access control**
Comment:
Code name: Data management: access control
Description: There is an access control to view certain data. Not everyone can see all the data in the ERP system. That is why there is an access control in the data management to ensure that not everything becomes visible.
Inclusion criteria: the statements discussing access control related to data. This is more control over who gets access to certain data and who determines that (e.g customer). For example, I have to be mindful of the securities and authorizations of who's getting what kind of authorization to view that data, but my client add that.
Exclusion criteria: these statements only discuss who gets what access to data and who determines that and NOT based on the role system they get a certain access.

○ **Data management: anonymize personal information**
Comment:
Code name: Data management: anonymize personal information
Description: Personal information sometimes needs to be made anonymous. This can be done by masking data or scrambling data.
Inclusion criteria: the statements that discuss anonymizing personal information. For example, the customer's data is anonymized. The names are there, but most critical data has been removed.
Exclusion criteria:

○ **Data management: malpractices**
Comment:
Code name: Data management: malpractices
Description: malpractices related to data management and specificly personal data handling.
Inclusion criteria: the statements that describe examples of malpractices or errors related to improper handling personal data or non-compliance with GDPR requirements. For example, "a customer in [EU country] […] the customer threw her workforce in Excel over the email with all salaries in it,".

○ **Data management: personal information withdrawal**
Comment:
Code name: Data management: personal information withdrawal
Description: Sometimes data needs to be "deleted" so that not everyone can see the data. The customer has the right to withdraw his or her consent at any time.
Inclusion criteria: the statements that discuss withdrawal data and the right they have. For example, [..] they also have a right to get it deleted.

Exclusion criteria: -

○ **Data management: storing personal information**
Comment:
Code name: Data management: storing personal information
Description: This code means that companies store personal information in their system. It is explicitly focused on storing personal information in the ERP system or by the company
Inclusion criteria: the statements that discuss storing personal information in the ERP system. For example, what personal information has been stored by which company?
Exclusion criteria: -

○ **Data privacy framework**
Comment:
Code name: Data privacy framework
Description: There is a set of privacy frameworks to protect data.
Inclusion criteria: the statements that discuss the data privacy framework that organisations, for example, follow. These frameworks are specifically intended to protect the data, and the managers/developers must also follow them. It specifically discusses the frameworks. For example, I suspect this is the framework within which to work. There are functionalities to get those GDPR compliance reports on all of someone's data and a way to keep it for a certain period to delete it afterwards.
Exclusion criteria: do not refer to the organisation privacy policy

○ **Principle of least privilege**
Comment:
Code name: Principle of least privilege
Description: It is a concept where the user is only given "privileges" that are necessary to perform their task. There are some principles that the user need to follow in order to complete his/her task.
Inclusion criteria: the statements that discuss the least principles. For example, what is the challenge for us is to classify those roles. This is more information that than they need. That's an exercise everybody does.
Exclusion criteria:

○ **Roles classification: data protection**
Comment:
Code name: Roles classification: data protection
Description: there is a division of roles to ensure that only certain people have specific access to data. This protects the data and not everyone will be able to perform an action.
Inclusion criteria: this statements only discusses that role-based system that people have a certain action such as accessing certain data. This ensures that data remains protected. For example, so the main thing that is concerned about the roles is that nobody can execute an action…. But also plays into the privacy concern because it is data you're giving access to.
Exclusion criteria: these statements only discuss a role-based system and that the data is protected and NOT about the general "data management: access control".

○ **Time for keeping data**
Comment:
Code name: Time for keeping data
Description: Data should be kept for some time.
Inclusion criteria: the statements discuss that data is kept for some duration. It can be discarded afterwards, to minimise the data, ot it can be about just the need to keep the data for some (long) time for the business need. Example “We generally have years together of data because then the business runs”
Exclusion criteria: When time of keeping the data is not important.

◈ **ERP systems**
6 Codes:

○ **ERP framework limitations**
Comment:
Code name: ERP framework limitations
Description: Not all frameworks work within the ERP system, there are some limitations associated with them. Due to the limitations in the ERP system framework, developers/managers have to perform extra activities to ensure that the ERP system functions properly.
Inclusion criteria: the statements that explicitly discuss limited ERP frameworks. For example; there is no framework to hide the data by default.
Exclusion criteria: these statements discuss only the limited frameworks in the ERP systems and NOT a challenge.

○ **ERP privacy goal**
Comment:
Code name: ERP privacy goal
Description: There are privacy goals within the ERP system.
Inclusion criteria: the statements that discuss the privacy goal of the ERP system. For example, the privacy goals of the ERP system, I can talk about production.
Exclusion criteria:

● **ERP privacy goals: not explicit**
Comment:
Code name: ERP privacy goal: not explicit
Description: There are no specific privacy goals within the ERP system.
Inclusion criteria: the statements that discuss there are no explicit or no existing privacy goals in the ERP system.
Exclusion criteria:

○ **ERP system migration**
Comment:
Code name: ERP system migration
Description: It is possible that an old ERP system needs to be migrated with a new ERP system.
Inclusion criteria: the statement discusses that ERP systems need to be migrated.
Exclusion:

○ **ERP systems**

Comment:
Code name: ERP system
Description: This is a “general” code when someone discusses the ERP system on a high level.
Inclusion criteria: the statements that discuss how the ERP systems work and what they store. For example, ERP is a case where are two types of data stored…
Exclusion criteria:

○ **Review ERP system**
Comment:
Code name: Review ERP system
Description: Sometimes a review has to be done on the ERP system whether it still complies with the law and regulations.
Inclusion criteria: this code is related to statements about reviewing the ERP system. For example, we have some certifications, some mandatory stuff which we go through once a year. In case of any reports of loss of some specific property…

◈ **GDPR/ other privacy regulations**
11 Codes:

○ **Compliance**
Comment:
Code name: Compliance
Description: The act or process of doing what developers/managers have been asked or ordered to do. This may include following the organisation's privacy policy to be compliant.
Inclusion criteria: these statements discuss high/general level compliance. Developers and managers must follow rules/requirements of the customer or organisation, for example, to be compliant. Sometimes an ERP system can be also compliant with regulations not related to GDPR. For example, ERP is software that is always in compliance with the regulations.
Exclusion criteria: do not refer to GDPR compliance OR local regulations.

○ **Exemption decisions**
Comment:
Code name: Exemption decisions
Description: Before the GDPR, there was other privacy legislation, such as Wbp, for which companies were granted an exemption.
Inclusion criteria: the statements that discuss the exemption decisions that are related to privacy rules.
Exclusion criteria:

○ **GDPR compliance**
Comment:
Code name: GDPR compliance.
Description: GDPR consists of privacy rules and laws for the processing of personal data. If you comply with the rules of GDPR, you are GDPR compliance.
Inclusion criteria: these statements discuss compliance with the GDPR. But also the reason why you comply with the GDPR. This is a high/general level. For example, the only thing we do is a classification for GDPR reasons.
Exclusion criteria: these statements discuss only GDPR compliance and NOT in detail about the implementation process to be GDPR compliant OR compliance

○ **GDPR compliance: implementation practices**
Comment:
Code name: GDPR compliance: implementation practices.
Description: There are implementation practices that ensure that the system is GDPR compliant.
Inclusion criteria: the statements discussing implementation practices to be GDPR compliant. This could be based on the customer/company requirement that they need to implement something to be GDPR compliant. For example, there is some request or requirement for GDPR to work compliant. This means we anonymize or delete this data as much as possible.
Exclusion criteria: these statements discuss only specific about the implementation practice and NOT GDPR compliance in general.

● **GDPR: Data controller**
Comment:
Code name: GDPR: data controller
Description: One of the requirements of GDPR is to assign a data controller.
Inclusion criteria: the statements that discuss the responsibility of the data controller and why it is necessary for the company. For example, a data controller should be aware of these processes.
Exclusion criteria:

○ **GDPR: Data processor**
Comment:
Code name: GDPR: Data processor
Description: A data processor transfers, organizes and processes personal data. It's more about the GDPR function/role they play within the project.
Inclusion criteria: this code is related to statements about the role of data processors within a project or organisation. For example, they have such a party that is a processor. Then you have to check whether the party fulfils its processing role properly.
Exclusion criteria: -

○ **GDPR: Data protection officer**
Comment:
Code name: GDPR: Data protection officer.
Description: A data protection officer is responsible for monitoring organisation compliance. It's more about the function/role. The DPO is involved in the project/organisation.
Inclusion criteria: this code is related to statements about the role of data protection officers within a project or organisation. For example, companies with a specific size must have a data protection officer.
Exclusion criteria: -

● **GDPR: Definition**
Comment:
Code name: GDPR: definition
Description: definition of what GDPR means.
Inclusion criteria: the statements that discuss the definition of GDPR. For example, GDPR is a ruling that came into effect in Europe….
Exclusion criteria:

○ **Privacy regulations**
Comment:

Code name: Privacy regulations
Description: Every country has different (privacy) regulations.
Inclusion criteria: the statements that discuss countries and rules. For example, a customer has 20.000 people in a system in 48 different countries. It's always a trick to make sure the system is watertight.
Exclusion criteria:

○ **Role of GDPR in ERP privacy**
Comment:
Code name: Role of GDPR in ERP privacy
Description: GDPR has an "effect" on ERP privacy. This code is explicit when they discuss ERP systems.
Inclusion criteria: the statements that discuss GDPR in the ERP privacy. For example, we had nothing explicitly called privacy or built in the framework of people talking about it. So unless GDPR came in, it was reactive because the software system had to comply with it.
Exclusion criteria: these statements discuss only the role of GDPR in the ERP system and do NOT privacy awareness.

○ **Role of GDPR in the work**
Comment:
Code name: Role of GDPR in the work
Description: GDPR has an "effect" on the work of consultants/developers. As a result, they have to build more functionalities to comply with the GDPR. This leads to more work activities.
Inclusion criteria: the statements that discuss the role of GDPR in the work. For example, we had to build certain things into our systems to comply with the legislation.
Exclusion criteria: These statements only discuss what effect it has on work and NOT what effect GDPR has on their personal knowledge.

◈ **Impediments/problems**
4 Codes:

○ **Conflict of interest between privacy and business needs**
Comment:
Code name: Conflict of interest between privacy and business needs
Description: there is a conflict of interests between what business needs and what privacy regulations or legislation require regarding privacy-related measures.
Inclusion criteria: the statements that discuss business needs, privacy measures/requirements and possible misalignment or conflict between them. For example, business needs to process credit card number and phone number for a particular process, while GDPR requires to protect this data. In this way, creating an overhead for developers to configure an ERP system in a specific way.
Exclusion criteria: the statements discuss only business needs or privacy requirements, without referring to potential conflict of interests.

○ **Lack of knowledge**
Comment:
Code name: Lack of knowledge.
Description: Know little about certain things.

Inclusion criteria: the statement that discusses when someone does not know about certain things. For example, I don't know if there are any guidelines.
Exclusion criteria:

○ **Mixing different concepts around privacy**
Comment:
Code name: Mixing different concepts around privacy
Description: Different concepts are mixed up when it comes to privacy. For example, sensitive and personal information may be mixed up.
Inclusion criteria: the statement that discusses different concepts of privacy such as sensitive information and personal information. For example, sending specific information to electric analytics is different, but it becomes challenging during integrations because the other system demands the information. It can be a registered address. It can be the company, taxation number that you share with them. What the internet can save is encrypted, but you're still sending it to another system
Exclusion criteria: -

○ **Privacy viewed as a security problem**
Comment:
Code name: Privacy viewed as a security problem
Description: The person focuses on securing information rather than on explaining privacy issues/solutions.
Inclusion criteria: this code is applied when questions about privacy are treated as questions about information security (access control, authorisations, roles, security in general) and data protection in the general sense.
Exclusion criteria: When no security solutions or challenges are discussed.

◈ **Privacy design aspects**
10 Codes:

○ **Business process**
Comment:
Code name: Business process
Description: The process whereby privacy is automated in the ERP system. This can be a process of a customer/client or a consultancy company.
Inclusion criteria: the statements that discuss the business process how, among other things, privacy is automated in the ERP system. It's more an explanation of how the process works and that they must comply with privacy aspects (such as minimising data). For example, it all depends on the functionality that the customer needs. If I minimise the data… if some data is mandatory for some process to be executed.
Exclusion criteria: These statements refer only to the business process and NOT business requirement OR customer requirement (e.g. customer requirement is adding a new field in the ERP system)

○ **Data minimisation**
Comment:
Code name: Data minimisation
Description: Limiting data collection to only that which is necessary to achieve a specific goal.

Inclusion criteria: the statements that discuss how data minimisation works within the organisation. High level of data minimisation. More focus on collecting, minimise and required data. For example, if a company pays my wags, you only need my last name and back account number, no longer like the birth date.
Exclusion criteria:

○ **Data minimisation driving factors**
Comment:
Code name: Data minimisation driving factors
Description: Some factors affect how you minimise data.
Inclusion criteria: the statements that discuss the driving factors of data minimisation such as customer, geographic and business process. For example, it all depends on the customer's functionality to minimise the data.
Exclusion criteria: don't refer to the definition of data minimisation.

○ **Data purpose**
Comment:
Code name: Data purpose
Description: There is a specific purpose when data is collected. It's a subpart of data minimisation.
Inclusion criteria: the statements that discuss the purpose of collecting data and focussing on testing data. For example, what is the purpose if you use this data, like public interest, legitimate interest?
Exclusion criteria:

○ **Design principles**
Comment:
Code name: Design principles
Description: Design principles are a set of considerations/factors that form the basis of the ERP system.
Inclusion criteria: the statements that discuss the design principles. It is very specific about the design such as implementing address tables in the ERP system. For example, there used to be a time when you wanted the person's name, phone number and other essential fields in one form.
Exclusion criteria:

○ **Mandatory fields**
Comment:
Code name: Mandatory fields Description: This is a part of the design principles. If someone discusses this explicit use this code.
Inclusion criteria: the statements that discuss the mandatory fields. It's between design principles and personal information. For example, fewer mandatory fields are possible to go through something.
Exclusion criteria:

○ **No experience: data minimisation**
Comment:
Code name: No experience: data minimisation
Description: There are developers or consultants who have no experience with the data minimisation concept.

Inclusion criteria: this code is related to statements about no experience with data minimisation and never have experience with implementing this concept in the ERP system. For example, I don't have experience with data minimisation.
Exclusion criteria: these statements only discuss that they have no experience with data minimisation and NOT that they deal with it.

○ **Organisation privacy policy**
Comment:
Code name: Organisation privacy policy
Description: It is a privacy policy that explains how an organisation handles customer/employee information. This code addresses both the consulting company's privacy policy and customer policy.
Inclusion criteria: the statements that discuss privacy policies within the organisation. For example, when you are building a system you need to think about the privacy concerns of the policies.
Exclusion criteria:

○ **Personal information**
Comment:
Code name: Personal information
Description: Personal information is data that is related to a person (sensitive/personal information). This is a higher level of personal information. Examples of personal information such as addresses.
Inclusion criteria: the statements that discuss personal information. For example, ERP is a case where two types of data are stored so employee data and suppliers data. Or another example, privacy is my personally identifiable and sensitive information that's being kept safely.
Exclusion criteria:

● **Privacy: definition**
Comment:
Code name: Privacy: definition
Description: An explanation is given by consultants and developers what privacy means according to them.
Inclusion criteria: this code is related to statements about the definition of privacy. For example, privacy is my personally identifiable and sensitive information that's being kept safely.
Exclusion criteria:

Privacy enablers
5 Codes:

○ **Privacy awareness**
Comment:
Code name: Privacy awareness

Description: The GDPR is making people increasingly aware of privacy. It is more on a personal level that people experience that GDPR more awareness is being created.
Inclusion criteria: this code is related to statements about privacy awareness. For example, every single person who's working as a part of a project should be aware of this.
Exclusion criteria: these statements only explicitly address privacy awareness they experience and NOT the role of GDPR in ERP privacy

○ **Privacy certification**
Comment:
Code name: Privacy certification
Description: Systems are certified, and this ensures privacy protection.
Inclusion criteria: the statements discuss that discuss that the system is certified and thus it is secure/offers privacy protection by default. Example: "But overall, everything is protected with data privacy certificates"
Exclusion criteria: the statements that recommend certification as a way to improve developer/consultant knowledge of privacy. OR the statements that just discuss that the system is secure/ensures privacy protection because it has been desiged in this way by e.g. Microsoft.

○ **Privacy driving factors**
Comment:
Code name: Privacy driving factors
Description: There are driving factors to implement privacy in the ERP system. This can be due to business needs or GDPR requirements that developers/managers have to implement.
Inclusion criteria: the statements that discuss the driving factors of privacy. For example, it's an overhead for a developer. So until and unless the business demands it. The people who are implementing the ERP systems demand that we don't do that.
Exclusion criteria: -

○ **Privacy driving factors: money**
Comment:
Code name: Privacy driving factors: money
Description: GDPR is the main driver that there is an (economic) conflict such as money. This is the negative side of privacy driving factors.
Inclusion criteria: the statements that discuss the (negative) driving factors of privacy in an economic sense. For example, the people who are implementing the ERP systems demand that we don't do that because everything is related to the amount of time spent or the amount of money spent on the implementation.
Exclusion criteria: -

○ **Privacy out of the box**
Comment:
Code name: Privacy out of the box
Description: Privacy is already implemented in the ERP framework. Developers/managers don't have to do anything themselves because the system already has it.
Inclusion criteria: the statements that discuss privacy is in the ERP framework. Suppliers such as Microsoft ensures GDPR. For example, […] but in the ERP system, these are standard

screens that are globally accepted and agreed upon. I don't have to put anything specific to achieve this privacy-by-design.
Exclusion criteria: do not refer to GDPR compliance OR data privacy framework

◈ Privacy-by-design
4 Codes:

○ **Privacy-by-design: definition**
Comment:
Code name: Privacy-by-design definition
Description: Explain exactly what privacy-by-design means.
Inclusion criteria: this code is related to statements that give an explanation of what people think of privacy-by-design. For example, privacy-by-design is when you design hardware that interacts with you and collects information about you that holds your data.
Exclusion criteria: this statement discusses only when people give an explanation of what they think of privacy-by-design and NOT what the goal is (such as benefits/consequences)

○ **Privacy-by-design: goal**
Comment:
Code name: Privacy-by-design: goal
Description: The goal of privacy-by-design is more of a solution if you implement this concept in the ERP system. What are the consequences if you implement this concept in the system, for example, due to privacy-by-design we are storing less data.
Inclusion criteria: this code is related to statements that give an explanation of how to achieve privacy-by-design. For example, that is to minimise the data that you take in, for whatever process. You want to process that data, and also give a choice to the customers or the people who give you the data to have a say in how it is managed or how it is stored or how it is shared.
Exclusion criteria: these statements only discuss the consequences/benefits of privacy-by-design and does NOT refer to the definition of privacy-by-design.

○ **Privacy-by-design: no experience**
Comment:
Code name: no experience: privacy-by-design
Description: There are developers or consultants who have no experience with the privacy-by-design concept.
Inclusion criteria: this code is related to statements about no experience with privacy-by-design and never have experience with implementing this concept in the ERP system.
Exclusion criteria:

● **Privacy-by-design: no usage**
Comment:
Code name: Privacy-by-design: no usage
Description: There is no "use of privacy-by-design in the ERP system.
Inclusion criteria: this code is related to statements that there is no use for privacy-by-design in the system. For example, we don't design anything by virtue, we don't use the privacy-by-design principle in ERP.
Exclusion criteria:

Recommendations
7 Codes:

○ **Recommendation: Data protection officer**
Comment:
Code name: Recommendation: Data protection officer
Description: It is recommended to have someone who is responsible for the data such as a data protection officer or data controller
Inclusion criteria: this code is related to statements about recommendations of having a person who is responsible for the data privacy in your projects. For example, do you have a data protection officer or data controller in your projects? That's the most important or the first question.
Exclusion criteria: these statements discuss only recommendations and NOT the function or the role of the data protection officer.

○ **Recommendation: GDPR compliance check-list**
Comment:
Code name: Recommendation: GDPR compliance check-list
Description: To assure that the project is developed inline with the GDPR requirements, use a check-list to control that each sprint is compliant with the GDPR.
Inclusion criteria: The statements suggesting check-lists as a way to get a transparency and better control over project compliance with the GDPR requirements. For example, if data minimisation applies to the current sprint's features.

○ **Recommendation: General guidelines**
Comment:
Code name: Recommendation: general guidelines
Description: It is recommended to have general guidelines within the organisation about data or privacy.
Inclusion criteria: this code is related to statements about recommendations of general guidelines within the organisation. It's important to have general guidelines to increase data or privacy understanding. For example, if I have the general guidelines on how this data should be handled, I think that is sufficient for us.
Exclusion criteria: -

○ **Recommendation: Privacy modules in ERP system**
Comment:
Code name: Recommendation: Privacy modules in ERP system
Description: Within the ERP system it is important to offer privacy modules so that the consultants or developers do not have to manually check whether the system or certain functionality complies with the privacy rules.
Inclusion criteria: this code is related to statements discussing that it is recommended to have privacy modules in the ERP system. For example, it would be good if in the future, from SAP,

to take that into account. So what is privacy and offer modules so you don't have to create all that by hand.

○ **Recommendation: Project involvement**
Comment:
Code name: Recommendation: Project involvement
Description: Developers or managers should be more involved with the project privacy so that they are aware of the privacy development.
Inclusion criteria: this code is related to statements discussing that it is recommended to involve developers or managers more in the project. For example, The only thing you could think of is that they participate and set up the security framework in which they work. That I'm getting involved in, I'm getting already, but everyone is getting a little bit involved in, 'Hey, we made this and this choice, for example. For working with client laptops, for these and these reasons.'
.
Exclusion criteria:-

○ **Recommendation: Raising privacy awareness**
Comment:
Code name: Recommendation: raising privacy awareness.
Description: It is recommended to create awareness about privacy so that people understand better what privacy concepts mean. This can be done through presentations, exams and trainings.
Inclusion criteria: this code is related to statements about recommendations of creating privacy awareness for developers and consultants. For example, I think creating more awareness. That has helped me. So I think for all people, consultants and developers but the whole company.
Exclusion criteria:-

○ **Recommendation: Research GDPR requirement**
Comment:
Code name: Recommendation: Research GDPR requirement
Description: It is recommended to do more research about the GDPR requirement. It is still not clear and this could probably be improved to ensure clear guidelines are in place.
Inclusion criteria: this code is related to statements about recommendations of doing research about GDPR requirements. GDPR only provides a requirement is still not clear to the developers. It is recommended to research the GDPR requirement to better understand the guidelines, for example. Another example, there is some request for GDPR to work compliant. This means that we delete data as much as possible. Therefore, research needs to be done. How it can be better.
Exclusion criteria: -

Roles/actors
3 Codes:

○ **Advisory role**
Comment:
Code name: Advisory role
Description: It gives suggestions and help people in the organisation. Give advice to consultants or other roles. It's a subset of managers/developers responsibility to give advice to the customers or projectteam.

Inclusion criteria: this code is related to statements about giving advice to the customers. For example, [..] I would be asking questions such as do you have an excel sheet […]
Exclusion criteria: do not refer to managers/developers responsibility.

○ **Engagement party**
Comment:
Code name: Engagement party
Description: A general code about engagement management. This included user engagement, manager engagement. These are all roles/actors involved in a project. Usually, the managers or the partner (higher level of the project) are responsible for the engagement of the customers.
Inclusion criteria: this code is related to statements about the engagements. For example, in most all the year engagements, you work with the user profiles and user engagement.
Exclusion criteria: these statements discuss only the role of the engagement and NOT the responsibility.

○ **Third-party**
Comment:
Code name: Third party
Description: A general code about the third party. They help organisations with privacy responsibilities or tasks.
Inclusion criteria: this code is related to statements about the third party. For example, some third-party software helps in taking a subset of data and scrambling…
Exclusion criteria: -

◈ **Tasks and responsibilities of the key factors**
22 Codes:

○ **Business requirements**
Comment:
Code name: Business requirements
Description: some design solutions and implementations are driven by the business requirements such as cutting costs or usability and other.
Includion criteria: the statements refering to the requirements from business side rather than driven by privacy goals of GDPR requirements.
Exclusion criteria: not related to the discussion of the conflicts between privacy perspective and business needs OR discussion of specific business processes (e.g., data needed for salary calulations).

○ **Customer involvement**
Comment:
Code name: Customer involvement
Description: Customers are involved with the project. They work together with the consultants/developer and can provide input in the project.
Inclusion criteria: this code is related to statements about the customer that work together with a consultant or developer. For example, the customer or the people who give you the data have a say in managing or storing it.
Exclusion criteria: not referring to user involvement.

○ **Customer requirement**
Comment:
Code name: Customer requirements
Description: Customer requirements are characteristics and specifications that a developer or management must implement for the customer in the ERP system.
Inclusion criteria: this code is related to statements that the customer has some requirements for their system. For example, we get specific requirements from the customer and you start building something based on that.
Exclusion criteria: these statements discuss the only high level the customer requirement and NOT the process (e.g. business process)

○ **Customer's responsibility**
Comment:
Code name: Customer responsibility
Description: the customer itself has some privacy-related responsibilities.
Inclusion criteria: this code is related to statements about customer responsibility.
Exclusion criteria:

○ **Customers right**
Comment:
Code name: Customers right
Description: This code means when someone discusses a customer's right related to privacy and data minimisation. This is more focused on the rights they have from GDPR.
Inclusion criteria: this code is related to statements about customers' rights. For example, GDPR is a ruling that came into effect in Europe that gives the right to the customers to what do you say to know what information is being stored.
Exclusion criteria: -

● **Develop own privacy knowledge**
Comment:
Code name: Develop own privacy knowledge.
Description: It is important to always be up to date with privacy, so it is essential to develop your own privacy knowledge.
Inclusion criteria: this code is related to statements about developing your own privacy knowledge. For example, I'm very conscious as a customer about the privacy concerns.
Exclusion criteria: -

○ **Developer's responsibility**
Comment:
Code name: Developer's responsibility
Description: A developer has different responsibilities within a project or assignment. It's more at a high/personal level. The obligation to make sure that something goes well.
Inclusion criteria: this code is related to statements about the responsibilities of a developer. For example, my responsibility is that which something has to be built and preferably before the deadline.
Exclusion criteria: the statements discuss only the developer's responsibility without referring to aspects of work (tasks)

● **Developer's task**
Comment:
Code name: Developer’s task
Description: A developer has different tasks within a project or assignment. This is more aspect of the work he/she is doing. A task must be completed according to a procedure/guidelines/instructions and it’s very specific.
Inclusion criteria: this code is related to statements about the tasks of a developer. For example, we had to build certain things into our systems DDon’t deal:to comply with the legislation.
Exclusion criteria: the statements discuss only the developer’s tasks without referring to the responsibilites.

● **Don't deal: data minimisation**
Comment:
Code name: Don’t deal: data minimisation
Description: Some consultants or developers are not involved with data minimisation. It’s outside their scope.
Inclusion criteria: this code is related to statements about the developers or consultants that are not involved with data minimisation. For example, I have not done core development or created user integration interfaces specific data minimisation concepts.
Exclusion criteria: -

○ **Don't deal: personal data**
Comment:
Code name: Don’t deal: personal data
Description: Consultants or developers don’t deal with personal data in the system.
Inclusion criteria: this code is related to statements about the developers or consultants that are not involved with processing personal data in the system. For example, we are not dealing with personal data.
Exclusion criteria: -

○ **Don't deal: privacy**
Comment:
Code name: Don’t deal: privacy
Description: Some consultants or developers are not involved with privacy because it doesn’t belong to their daily job.
Inclusion criteria: this code is related to statements about the developers or consultants that are not involved with privacy. For example, I think that’s because it is mainly outside my scope, that’s where the customer cares.
Exclusion criteria: -
○ **Don't deal: privacy-by-design**
Comment:
Code name: Don’t deal: privacy-by-design
Description: Some consultants or developers are not involved with the privacy-by-design concept. It’s outside their scope.
Inclusion criteria: this code is related to statements about the developers or consultants that are not involved with privacy-by-design. For example, I have worked with large ERP systems that are some products of [the professional service firm], where I have to be mindful of privacy by design or data minimisation. But in the ERP system, as I said, these are standard screens that

are globally accepted and agreed upon. I don't have to put anything specific to achieve this privacy-by-design.
Exclusion criteria: -

○ **Engagement management privacy responsibility**
Comment:
Code name: Engagement management privacy responsibility
Description: The engagement management has some privacy responsibility in order to protect the information in the project. They are the high level of a project.
Inclusion criteria: this code is related to statements about the engagement management is responsible for privacy within a project or organisation. For example, the engagement partner should also devise all the rules and regulations. Make sure that it gets cascaded across the team.
Exclusion criteria: these statements discuss only the responsibilities and NOT as an actor.

○ **Manager/consultant responsibility**
Comment:
Code name: Manager/consultant responsibility
Description: A manager has different responsibilities within a project or assignment. It's more at a high/personal level. The obligation to make sure that something goes well.
Inclusion criteria: this code is related to statements about the responsibilities of a manager. For example, wherever I'm working, I have to be mindful of the securities and authorizations of who's getting what kind of authorization to view what data.
Exclusion criteria: the statements discuss only the manager's responsibility without referring to aspects of work (tasks)

○ **Manager/consultant task**
Comment:
Code name: Manager/consultant task
Description: A manager/consultant has different tasks within a project or assignment. This is more aspect of the work he/she is doing.
Inclusion criteria: this code is related to statements about the tasks of a manager. For example, I have to onboard with new clients into the system in my current role.
Exclusion criteria: the statements discuss only the manager's tasks without referring to the responsibilities.

● **Privacy control**
Comment:
Code name: Privacy control
Description: A personal responsibility concerns controlling your privacy, for example, you give permission to someone who can take pictures of you. You have your own control over your own privacy.
Inclusion criteria: this code is related to statements about privacy control. For example, it's as simple as what do you say what information you wish to divulge.
Exclusion criteria: this statement discusses only the personal responsibility regarding privacy and NOT data control e.g. personal information safety.

○ **Privacy is handled outside the project's scope**
Comment:
Code name: Privacy is handled outside the project's scope
Description: privacy is considered outside the scope of a development project.

Inclusion criteria: the statements that note that the privacy is considered by the [professional service firm], but in general rather than within a scope of a specific project. For example, they mention that there is a coverage and compliance unit in the firm or that the privacy is handled upfront the project.

○ **Privacy related trainings**
Comment:
Code name: Privacy related trainings
Description: Sometimes training is provided (mandatory within the company) on privacy and this can help keep developers/administrators up to date. This way they stay informed of the latest developments regarding privacy.
Inclusion criteria: This code is related to statements that discuss privacy training that ensures developers/managers are aware of the latest development and what is happening in practice. For example, we go through all mandatory trainings, which talks about privacy and security.
Exclusion criteria: not referring to recommendations OR developing own privacy knowledge.

○ **Privacy: Low priority**
Comment:
Code name: Privacy: low priority
Description: Privacy has low priority among other developer', consultant's or manager's tasks.
Inclusion criteria: the statements discussing that privacy and GDPR have low priority/interest among other working tasks.
Exclusion criteria: -

○ **Project responsibility**
Comment:
Code name: Project responsibility
Description: The team members who participate are responsible for the entire project but also responsible for the project structure such as laptops.
Inclusion criteria: this code is related to statements about the entire project responsibilities. For example, [..] but it is the entire project responsibility
Exclusion criteria: these statements refer specifically to the entire project and not to 1 person who is responsible for a specific task.

○ **Training is not essential**
Comment:
Code name: Training is not essential.
Description: Training is no longer essential because it has become a task. It didn't give much input related to knowledge.
Inclusion criteria: this code is related to statements that training is not essential. It doesn't create value. For example, any number of training that is done did not have that much. So people need to be conscious about that and recall that.
Exclusion criteria: do not refer to recommendations OR privacy training

○ **User involvement**
Comment:
Code name: User involvement
Description: User involvement ensures that (end)users/customers voices are also "heard" and how certain functionalities in the system can be improved through their involvement. This is a

high level of (end) user involvement. An end-user is a person that actually uses the system, this can be customers or developers.
Inclusion criteria: this code is related to statements about the users that are involved in the project. They have something to say related to data privacy. For example, […] people who give you the data have a say in managing or storing it.
Exclusion criteria: do not refer to customer involvement

No code group
3 Codes:

○ **App**
Comment:
Code name: Application
Description: An ERP system is a set of applications or modules. This code is general if someone discusses the application in the ERP system.
Inclusion criteria: this code is related to statements about the application. For example, an app asking you to give your contact or location information.
Exclusion criteria: -

● **Concept: area**
Comment:
Code name: Concept: area
Description: Privacy can take place in different areas.
Inclusion criteria: this code is related to statements about the area's where privacy takes place. For example, especially in the digital space or even in the non-digital space.
Exclusion criteria: -

○ **Examples:**
Comment:
Code name: Examples
Description: This is a general code about examples. Developers or consultants can explain with examples to better explain the story.
Inclusion criteria: this code is related to statements about examples. For example, what has personal information been stored by which company?
Exclusion criteria: -

Appendix IV – Codebook used during the Qualitative Analysis

	Dev. (N = 7)		Con. (N = 9)	
	Total	Part.	Total	Part.
Code				
A. Challenges				
A.1. Data analytics external	1	1	-	-
A.2. Depends on the project	8	5	5	4
A.3. Dynamic Information	1	1	-	-
A.4. ERP privacy challenge	16	6	16	8
A.5. ERP systems integration	4	4	8	5
A.6. GDPR challenge	4	3	5	3
A.7. No ERP privacy challenge	5	5	3	2
A.8. Processing data outside ERP system	-	-	5	3
B. Data management				
B.1. Data classification	5	3	-	-
B.2. Data control	3	3	6	6
B.3. Access control	16	6	24	9
B.4. Anonymize personal information	6	4	3	2
B.5. Malpractices	1	1	4	4
B.6. Personal information withdrawal	2	2	1	1
B.7. Storing personal information	3	3	10	7
B.8. Data privacy framework	3	3	4	4
B.9. Principle of least privilege	4	2	5	4
B.10. Roles classification: data protection	5	4	4	3
B.11. Time for keeping data	3	1	1	1
C. ERP systems				
C.1. ERP framework limitations	3	2	1	1
C.2. ERP privacy goal	2	2	1	1
C.3. ERP privacy goals: not explicit	2	2	4	3
C.4. ERP system migration	1	1	2	2
C.5. ERP systems	5	4	7	6
C.6. Review ERP system	1	1	-	-
D. GDPR/ other privacy regulations				
D.1. Compliance	2	2	-	-
D.2. Exemption decisions	-	-	2	1
D.3. GDPR compliance	9	5	16	8
D.4. Implementation practices	3	2	3	3
D.5. Data controller	1	1	1	1
D.6. Data processor	1	1	1	1
D.7. Data protection officer	1	1	3	2
D.8. Definition	6	5	8	7
D.9. Privacy regulations	1	1	4	4
D.10. Role of GDPR in ERP privacy	7	4	7	5
D.11. Role of GDPR in the work	15	6	9	6
E. Impediments				
E.1. Conflict of interest between privacy and business needs	17	7	9	6
E.2. Lack of knowledge	6	4	7	5
E.3. Mixing different concepts around privacy	7	5	4	2
E.4. Privacy viewed as a security problem	9	5	3	2
F. Privacy design aspects				
F.1. Business process	8	4	3	2
F.2. Data minimisation	19	7	23	9
F.3. Data minimisation driving factors	3	3	-	-
F.4. Data purpose	7	5	8	5
F.5. Design principles	10	7	13	6
F.6. Mandatory fields	1	1	1	1
F.7. No experience: data minimisation	1	1	6	5
F.8. Organisation privacy policy	5	3	12	7
F.9. Personal information	5	3	19	9
F.10. Privacy: definition	6	6	10	9
G. Privacy enablers				
G.1. Privacy awareness	3	2	7	6
G.2. Privacy certification	1	1	2	1
G.3. Privacy driving factors	1	1	-	-
G.4. Privacy driving factors: money	1	1	-	-
G.5. Privacy out of the box	10	4	18	6
H. Privacy-by-design				
H.1. Privacy-by-design: definition	5	5	6	5
H.2. Privacy-by-design: goal	1	1	-	-
H.3. Privacy-by-design: no experience	6	6	8	6
H.4. Privacy-by-design: no usage	2	1	3	3
I. Recommendations				
I.1. Data protection officer	-	-	1	1
I.2. GDPR compliance check-list	-	-	1	1
I.3. General guidelines	3	2	3	2
I.4. Privacy modules in ERP system	-	-	1	1
I.5. Project involvement	1	1	1	1
I.6. Raising privacy awareness	8	7	11	9
I.7. Research GDPR requirement	1	1	-	-
J. Roles/Actors				
J.1. Advisory role	-	-	5	4
J.2. Engagement party	-	-	2	2
J.3. Third-party	1	1	2	2
K. Tasks and responsibilities				
K.1. Business requirements	5	4	1	1
K.2. Customer involvement	3	3	8	6
K.3. Customer requirement	6	4	6	3
K.4. Customer's responsibility	4	3	8	4
K.5. Customer's right	7	6	11	8
K.6. Develop own privacy knowledge	2	2	-	-

	Dev. (N = 7)		Con. (N = 9)	
	Total	Part.	Total	Part.
Code				
K.7. Developer's responsibility	8	5	-	-
K.8. Developer's task	3	3	1	1
K.9. Don't deal: data minimization	1	1	4	3
K.10. Don't deal: personal data	5	3	6	4
K.11. Don't deal: privacy	7	4	4	4
K.12. Don't deal: privacy-by-design	2	2	4	4
K.13. Engagement management privacy responsibility	-	-	2	2
K.14. Manager/consultant responsibility	4	2	13	8
K.15. Manager/consultant task	-	-	4	2
K.16. Privacy control	4	3	10	6
K.17. Privacy is handled outside the project's scope	-	-	5	4
K.18. Privacy related trainings	1	1	-	-
K.19. Privacy: Low priority	2	2	2	2
K.20. Project responsibility	2	2	2	2
K.21. Training is not essential	1	1	-	-
K.22. User involvement	1	1	-	-
L. No Code Group				
L.1. App	2	2	1	1
L.2. Concept: area	2	2	-	-
L.3. Examples	11	3	16	5

Table 4 Codebook used during the Qualitative Analysis