EUROPEAN PRIVACY BY DESIGN

-thesis notebook-

1. INTRODUCTION

Throughout history, privacy protection has received a deserved and rewarded attention. Yet, no attempts to conceptualize privacy managed to describe all its components[1]. This is not revolutionary, rather evolutionary and it is proven by the relative staticness[2] of data protection principles enshrined in the applicable laws and regulations. Where the basics of the legal background had not changed, the way technologies implement such requirements did, and often on such an opaque level that served the ground for privacy lawyers to dwell into everlasting litigations and debates.

Protection and *per a contrario*[3] violation of the privacy is not an exact science. It is not based on set parameters and does not allow absolute precision in its results. It is not the use-case of an expired parking ticket, where the minutes spent in delay (*i.e.* unlawful conduct of a citizen) can easily be translated into a precise administrative fine. From the perspective of law enforcement, the economic value of the said unlawful conduct aims to be dissuasive enough to suppress future attempts. From the citizen's perspective it is an exercise of risk *vs.* reward based on game-theory. Where the fine could only reach a certain level, having a limited economic value on the citizen's financial status, perhaps the benefits gained from those precious minutes of unlawful parking outweighed the financial loss suffered from paying the fine.

Why it is different when someone's privacy is violated? In case of privacy violations the citizen is subject to a tail of combined events that might trigger sever violations of the right to privacy and cause harm. As such, in case of privacy violations the damage suffered by citizens is not imminent, nor immediately detectable. Snipe eloquently argues about how citizens value privacy differently and how different members of the same network cannot actually maintain different levels of privacy practice[4]. Network refers to internet service providers and or other information technology

[1] In detail, see Acquisti et al. 2016, pp. 2-48.
[2] Staticness meaning as not changing for a long time.
[3] Known as appeal from the contrary, denotes any proposition that is argued to be correct because it is not disproven by a certain case. Arguments *per a contrario* are often used in the legal system as a way to solve problems not currently covered by a certain system of laws.
[4] Snipe 2021.

services that have a network effect (*e.g.* e-mail)[5]. These markets for privacy in the networks are absorbing the digital footprint of every user[6]. And we know, as the society digitization marches onward, the privacy markets are getting bigger, due to increasing numbers in active users. According to Jonson, as of January 2021 there were 4.66 billion active internet users worldwide, accounting for 59.5 per cent of the global population[7].

Is there an infrastructure inversion in privacy markets? The concept of infrastructure inversion was used by Andreas M. Antonopoulos, who defined it as phenomenon that is used when a new technology must first use the old infrastructure, and how that creates a conflict and pressure that can lead to an infrastructure inversion[8]. He argued that this is caused by the fact that in its first few years of its adoption it has to be carried by the existing technology that it is disrupting[9]. In privacy markets signs of infrastructure inversion are visible too. Even if the active internet users worldwide suffer from a privacy paradox, the technological advances are present. By way of a mere example, the widespread usage of AdBlock Plus[10] extensions in internet browsers may prove this assumption. The *Statista* research department provided the last quarterly results of monthly active users of mobile adblocking browsers reaching to 586 million[11]. Further, users are increasingly adopting privacy-preserving tools to protect their web usage. In this regard, the Brave browser has reached 36 million monthly active users in September 2021[12].

But what do we mean when we say privacy paradox? Privacy paradox is demonstrating the discrepancy between users' intention to protect their privacy and how these users actually behave on the privacy markets. A systematic literature review on this concept has been provided by Susanne Barth and Menno D.T. de Jong[13]. They concluded that a user's decision-making process as it pertains to the willingness to divulge privacy information is generally driven by two considerations: (1) risk-benefit evaluation and (2) risk assessment deemed be none or negligible[14].

[5] Ibid.
[6] User refer to citizens.
[7] Jonshon, 2021.
[8] Antonopoulos, 2017.
[9] Ibid.
[10] AdBlock Plus is a free extension that allows the user to customize its web experience. The user can block ads or disable tracking. More information: https://adblockplus.org/en/about
[11] Statista 2021, https://www.statista.com/statistics/606357/mobile-adblocking-browser-users-worldwide/ [09.23.2021].
[12] Brave Announcements, 2021, https://brave.com/36m-mau/ [09.24.2021].
[13] Barth - de Jong 2017, pp.1038-1058.
[14] Ibid.

Reflecting on this research, arguable the user perception against privacy risks can be illustrated in a two-to-two-dimensional matrix with privacy paradox sitting at the intersection of the four different user profiles that can be constructed from this matrix. *Figure 1* provides the overview of the privacy paradox.

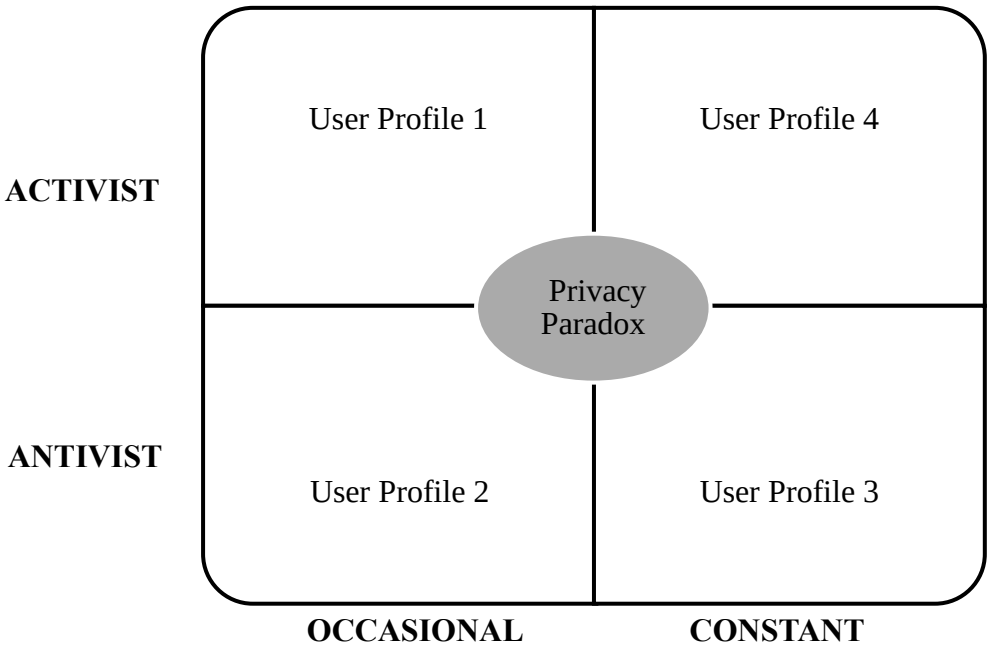

Figure 1. Overview of privacy paradox.

We can generate four different user profiles interpreting this matrix:

a. **User Profile 1** – Occasional Activists, describe users that are looking after noisy events to bring out their privacy concerns towards the public.
b. **User Profile 2** – Occasional Antivists, describer users willing to accept privacy violations based on a risk-benefit evaluation.
c. **User Profile 3** – Constant Antivists, described who's risk assessments are deemed to be negligible or none; and thus simply ignore any risks imposed on their privacy.
d. **User Profile 4** – Constant Activists, describe users never willing to accept privacy violations as their risk-benefit evaluations are always leaning towards privacy protection.

We can also expect that the digital transformation and the infrastructure inversion in the privacy markets will reshape the size of these user groups and, with that, will also alter the location of the privacy paradox. We argue that the European data protection reform and its associated strategy

will eventually lead to the complete erosion of constant antivists (*i.e.* User Profile 4). With EU data sovereignty[15] becoming more emphasized, its companies will be required to focus more on compliance with this mandate. Legislation brings privacy protection to the attention of organizations establishing the organizational privacy sphere. In this sphere, these actors need to serve their customers as well as their own personnel with efficient measures to process their data in a transparent and secure manner. This trend is translating into the updated overview of the privacy paradox, as depicted in *Figure 2*.

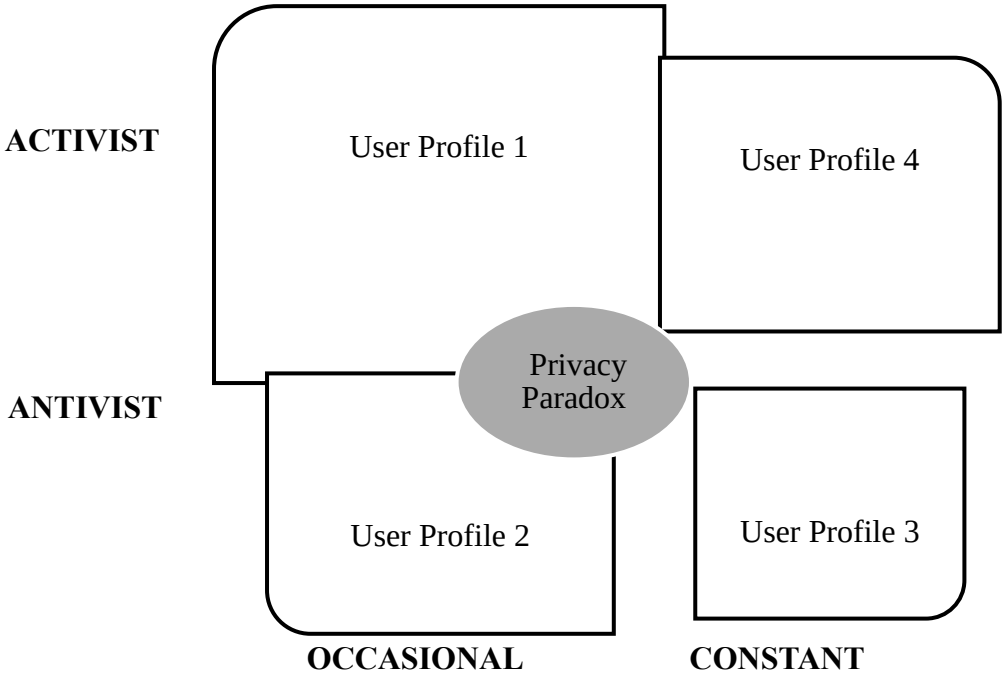

Figure 2. Updated overview of privacy paradox.

Against this background, we consider useful to understand the concept of Privacy by Design (PbD) through a European lens. We aim to do it by analyzing its nature, application and enforcement and discover new findings in our journey. The ultimate understanding the European PbD will help us navigate the infrastructure inversion in the European privacy markets.

2. RESEARCH CONTEXT

According to Peffers et al., information systems (IS) is an applied research discipline, in the sense that it frequently applies theory from other disciplines, such as economics, computer science, and

[15] Data sovereignty represents the idea that data are subject to the laws and governance structures within the nation it is collected. The EU data sovereignty refers to the data collected within the European Union.

the social sciences, to solve problems at the intersection of information technology (IT) and organizations[16]. IT is a constantly emerging industry of nowadays economy. IS developed by organizations are wide spreading even more rapidly, while the amount of data generated through these platforms are exceeding any expectations. Both personal and non-personal data are key enablers of the single digital market, a concept meant to offer business opportunities and new business models across the European Union (EU).

Recent high profile data breaches have pushed consumers to escape from service providers that did not adequately protect data. However, there are certain scenarios where no escape route is given. An example of such is the employment relationship between the employer and employee, whereas the amount of data generated and processed by the employer is dangerous towards its employee's informational privacy. Informational privacy was described by Koops et al., where the authors defined eight different types of privacy, also establishing that informational privacy is:

> *an overarching aspect of each underlying type [*of privacy*], typified by the interest in preventing information about one-self to be collected and in controlling information about one-self that others have legitimate access to*[17].

From an organizational point of view, this is a compliance and security risk, often handled by a proper data governance. An employer, who is outsourcing its services towards external service providers, escalates the compliance and security risks. Organizations using IS are directly or indirectly exposing themselves to potential data breaches. And these data breaches are treated with unmatched severity in the currently applicable legislative framework. As the constant threat is imminent, organizations need to have strong confidence in their provider's compliance levels. In case services are provided via Cloud Computing (CC), the setup can get even more complicated. In a CC environment, other entities are likely to join the infrastructure: cloud-brokers, cloud-auditors, cloud-intermediaries, and other agents. Thus, the Privacy Ecosystem (PECO) of rules relating to data processing in cloud-based IS can be defined as an interoperability zone of at least three and sometimes even more key participants[18].

[16] Peffers et al. 2007, p. 45.
[17] Koops et al. 2016, p. 568.
[18] These are the data controller as the client who is using an IS / IT service or solution, the data processor as the solution provider, the data sub-processor as entities used by the solution provides in its supply-chain, and individuals as data subjects, whose data are subject to processing. For a detailed description on these roles, see Section 4.8.6.3.

Seemingly, the current privacy and data protection requirements are pointing towards the conclusion that technological and regulatory measures failed to provide citizens with satisfactory privacy protection in Information and Communication Technologies (ICTs)[19].

This is one of the many reasons why data integrity and data security also constitutes an essential characteristic of IT itself. Notably, the EU - regime adopted the famously known set of PbD principles, curved out by Ann Cavoukian[20]. However, PbD principles, as such, were not included into the data protection principles foreseen in Article 5 of GDPR. Rather these are seen an extension of integrity and confidentiality principle, since the methodological approach of data protection by design places more accent on data security, than privacy[21]. The seven principles of PbD can be briefly described as follows:

a. Proactive not Reactive; Preventative not Remedial: anticipates and prevents privacy-invasive events before they happen. In short, comes before the fact, not after.
b. Privacy as the Default Setting: if an individual does nothing, its privacy remains intact. No action is required on the part of the individual to protect its privacy. In short, privacy is built into the system by default.
c. Privacy Embedded into Design: becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality. In short, it is not an add-on, after the fact.
d. Full Functionality – Positive-Sum, not Zero-Sum: helps avoiding the pretense of false dichotomies, such as privacy vs. security, demonstrating that it is possible to have both. In short, it is a winner for two.
e. End-to-End Security – Full Lifecycle Protection: having been embedded into the system prior to the first element of information being collected extends throughout the entire lifecycle of the data involved, from start to finish. In short, it is a cradle to grave protection.
f. Visibility and Transparency: component parts and operations remain visible and transparent, to users and providers alike. In short, it is trustworthy, but verified.

[19] van de Pas – van Bussel 2015, p. 186.
[20] Cavoukian 2013, pp. 2-3.
[21] Fabiano 2017, p. 731.

g. Respect for User Privacy: requires development teams to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options. In short, it is user centric.

3. RESEARCH OBJECTIVE AND HYPOTHESIS

The current thesis provides the results of a cross-disciplinary research plan. This plan is targeting the concept of European PbD as full functionality[22]. Its aim is to discover what the European PbD is, how it is applied and how it is enforced. Along the way, it may prove to become a future catalyst for disruptive innovation. Adding one-step to disruptive innovation, the potential for infrastructure inversion[23] is discussed. Hence, a holistic view is applied towards the applicable regulatory instruments[24] and software design[25] as a research field.

Arguably, there is an assumption of insufficient privacy and data protection strategies in IS development. Martens and Teuteberg provided that in IS literature only few explicit evaluation approaches to reference models can be found, most of which, however, do not lead to convincing results[26]. There are undoubted economic benefits in IS embedding PbD requirements. These benefits serve an important role in customer satisfaction. Implementing PbD is strategically important. Core privacy requirements have to be essential parts of the IS infrastructure.

The research objective argues that PbD strategies in software architecture are driven by the requirements stemming from privacy and data protection laws, and manipulated by business goals, that translate into infrastructure inversion as a product of disruptive innovation. To achieve this, we believe that in Europe, PbD principles are more focused on data anonymization and data

[22] Cavoukian 2006, pp. 3-4.
[23] *The concept of infrastructure inversion is used when a new technology must first use the old infrastructure, and how that creates a conflict and pressure that can lead to an infrastructure inversion. When a new technology is introduced, many are quick to say, "See it's not working, it's slow, or it doesn't work as well." This is not new. This happens every time you have a new technology that is disruptive; in its first few years of its adoption it has to be carried by the existing technology that it is disrupting. When you introduce a disruptive technology, you meet resistance. Resistance is the first reaction. The ones who succeed are the ones who continue-even though the rest of society tells them they are crazy. In the beginning, the disruptive technology has to live in a world created for the technology it is replacing. Infrastructure inversion is when you start with the new technology living on the old infrastructure and then, it flips. You build infrastructure and then the old infrastructure rides on top, on the infrastructure designed for the new technology.* (Antonopoulos 2018).
[24] Bygrave 2010, pp. 179-198.
[25] Sommerville 2008, pp. 241-266.
[26] Martens – Teuteberg 2011, p. 8.

security. In practical terms: new software solutions need to operate with the lowest amount of personal data, replacing them with equivalent non-personal data. Where replacement is not possible, their processing need to take place with the highest security standards affordable to the company. In a sense this will cause less applicability of data protection and more applicability of privacy preservation. We construct the research hypothesis around the enforcement of enforcement of PbD. We try to understand what its impacts are. We seek to learn if law enforcement agencies are pioneering a new approach to privacy preservation. This is why we want to measure their activity.

4. RESEARCH GAP AND QUESTION

Regulation of data protection and privacy is paramount[27]. The ramifications of privacy and data protection never reached this far and with such efficiency. Based on the current state of art, in light of the existing literature, it can be concluded that risks associated to insufficient privacy and data protection in IS are well founded and present a moderate level of discussion. In particular, the enforcement of PbD is under-researched.

We identify this specific sub-field that provides the research gap. Our intention is to make the field of European PbD richer, by rendering it more understandable to non-legal experts. Therefore, our research question seeks to discover if the enforcement of PbD can be measured and if yes, what are possible ways to do so?

Our research findings may support businesses to overcome difficulties in adopting a methodology that promotes data privacy in their organization. Furthermore, the findings will shed light on practices that might educate data protection authorities (DPAs).

5. RESEARCH METHODOLOGY

Mertens describes research as the systematic inquiry whereby data are collected, analyzed, and interpreted in some way in an effort to "*understand, describe, predict or control an educational or psychological phenomenon or to empower individuals in such contexts*"[28]. He further suggests that the "*exact nature of the definition of research is influenced by the researcher's theoretical framework*" with theory being used to establish relationships between or among constructs that

[27] Löhe – Blind 2015, p. 5.
[28] Mertens 2005, p. 2.

describe or explain a phenomenon by going beyond the local event and trying to connect it with similar events[29]. Mackanzie and Knipe also provides that the theoretical framework, as distinct from a theory, is sometimes referred to as the paradigm[30] and influences the way knowledge is studied and interpreted[31]. It is the choice of paradigm that sets down the intent, motivation, and expectations for the research[32]. Without nominating a paradigm there is no basis for subsequent choices regarding methodology, methods, literature, or research design[33]. Having in mind these statements, first and foremost the research paradigm has to be identified.

A vast number of research paradigms are present in the academic research such as: positivist, constructivist, interpretivist, transformative, emancipatory, critical, pragmatism and deconstructivist paradigms. However, multiple paradigms are suitable for the research problem, due to personal motivation, the pragmatic paradigm is embraced. Pragmatism researchers focus on the "what" and "how" of the research problem[34]. This is well reflected in the research question. The pragmatic paradigm also enables and encourages the use of mixed research methods, which provides the necessary flexibility to conduct comprehensive research. Creswell mentions that the pragmatic paradigm places "the research problem" as central and applies all approaches to understanding the problem[35].

This philosophical stance is recommended for the creation of methodologies that might be useful both for the academia and for business practices as well. Yet, researchers famous for delivering clear insights of the existing paradigms denote that with the research question "central", data collection and analysis methods are chosen as those most likely to provide insights into the question with no philosophical loyalty to any alternative paradigm[36]. Therefore, it can be concluded that the pragmatic paradigm is problem-centered and real-world practice oriented[37]. Nevertheless, the transformative paradigm has also key characteristics, which are embraced by the

[29] Ibid.
[30] The authors also refer to Mertens' study on paradigms.
[31] Mackenzie – Knipe 2006.
[32] Ibid.
[33] Ibid.
[34] Creswell 2003, p. 11.
[35] Ibid.
[36] Mackenzie – Knipe 2006.
[37] Ibid.

author. These are for instance the participatory and change-oriented characteristics of the transformative paradigm[38].

Throughout the research process, the research map defined by Mackanzie and Knipe, is followed[39]. The research map is used as a general guide for conducting research. It serves as a basic roadmap in conducting the research, with customized tailoring based on specific problems. Following the map also prevents problems encountered when methodological fit is low, as Edmondson and McManus have described these in their extensive guideline for finding the most adequate methodology[40]. During the research and thesis writing period, an elaborated Action Design Research methodology (ADR) was applied[41]. It required both field and desk work, with the defined unit of analysis inside an organization and sampling based on convenience. Alternatives, as action research and case-study research have been considered in a timely manner. Case-study based research is particularly appropriate for certain types of problems since it is suitable to capture the knowledge of practitioners. Baxter and Jack also concluded that qualitative case study is an approach to research that facilitates exploration of a phenomenon within its context using a variety of data sources[42]. This ensures that the issue is not explored through one lens, but rather a variety of lenses, which allows for multiple facets of the phenomenon to be revealed and understood.[43] Notably however, according to Yin, case study design should be considered when the researcher cannot manipulate the behavior of those involved in the study[44]. Yin's statement is a groundbreaker from the perspective of this research. Throughout the fieldwork, it was possible to influence the behavior of the target group that was involved in the research process. This is a reason why case-study research was not be considered the optimal solution.

On the other hand, ADR is defined as an interventionist approach to the acquisition of scientific knowledge that has sound foundations in the post-positivist tradition[45]. The interventionist approach fits the transformative paradigm as well. This as explained by researchers as a two-stage process:

[38] Ibid.
[39] Ibid.
[40] Edmondson – McManus 2007, p. 1170.
[41] Mullarkey – Hevner 2018, pp 1-16.
[42] Baxter – Jack 2008, p. 544.
[43] Ibid.
[44] Cited in Baxter – Jack, p. 555.
[45] Baskerville – Wood –Harper 1996, p. 237.

a. the *diagnostic stage* involves a collaborative analysis of the social situation by the researcher and the subjects of the research; and
b. the *therapeutic stage* involves collaborative change experiments, where such changes are introduced, and the effects are studied[46].

To achieve scientific rigor, additional structures have been imposed on ADR in this research. Thus, the action research cycle had been established with five phases that were iterated several times: (1) diagnosing, (2) action planning, (3) action taking, (4) evaluating and (5) specifying learning[47], as illustrated on *Figure 3*.

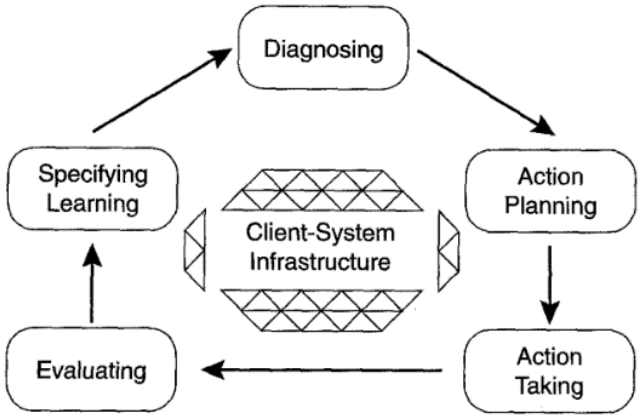

Figure 3. Action research cycle[48].

Baskerville and Wood-Harper also identified the characteristics of the method in relation to ideal domains of the action research, whereas they provide that:

a. the researcher is actively involved, with expected benefit for both researcher and organization;
b. the knowledge obtained can be immediately applied. There is not the sense of the detached observer, but that of an active participant wishing to utilize any new knowledge based on explicit, clear conceptual framework; and
c. the research is a cyclical process linking theory and practice[49].

[46] Ibid.
[47] Ibid.
[48] Ibid.
[49] Ibid, p. 239.

6. DATA COLLECTION AND ANALYSIS

First, the chosen research design required data collection in a repetitive manner systematically inquiring research groups. The research design plan implied regular monthly and weekly meetings to conduct necessary alignments with development teams working on various IS projects. The author provided advice on data privacy in all such projects. By nature of the research methods applied, here the data collection was longitudinal. We use these to navigate selected convergent technologies with PbD principles. We present our work in Chapter 6 of the thesis. This is also where, due to nature of the collaboration on different projects, we are able to witness the development of IT solutions that embrace PbD in their core, while maintaining their main functionalities.

Second, we use fully structured interviews to extract the knowledge of interview subjects that are working in a law enforcement agency (namely a national DPA). Although we contacted six interview subjects, only received responses from three of them. The selection of interview candidates is based on the highest fines issued by countries and highest number of total fines issued by countries. The interview insights are described in Chapter 7. These are analyzed with techniques described by Miles et al.[50], for potential coding of qualitative data to quantitative measures. We use anonymized interview responses.

Third, we rely on semi-structured interviews to understand what organizational privacy means for experts working with such organizations. The author uses its network to build up a pool of interview subjects covering multiple European countries. The aim here is to understand how organizations are dealing with challenges resulting from the implementation of PbD principles by ensuring continuous comprehensive and efficient privacy programs in the organizations they are working with. Practically speaking, if there is no organizational privacy program deployed, there is little chance that PbD principles find their way into such organizations business practices. We favor a variety of countries, since this way we discover multiple approaches within the EU countries. The interview notes are presented in Chapter 3. Same as above, the interview responses are anonymized.

[50] Miles et al. 2014, pp. 7-18.

The vast amount of qualitative and quantitative data that is collected also gives the possibility to carry out content analysis and pair-wise comparisons. Comparative studies of architectures for PbD are described in Chapter 5. The role of this chapter is not primarily to comment on the selected architectures. Instead, it is targeted towards displaying the richness of fields in which PbD principles can find practical application. We try to showcase how privacy protection is moving beyond pure theoretical frames.

To ultimately respond to the research question, we carry out text mining, decision tree modelling and predictive analysis of GDPR fines using machine learning techniques in Chapter 7 of the thesis. The applications deployed in this regard are Rapidminer and R. This is the most accentuated part of the thesis. We expect this to yield the most significant results and to enrich the field of legal data science.

7. TERMINOLOGY

Users nowadays pay with their personal data and privacy; they do not invariably benefit when the services are 'free'[51]. This places data on the top of the value-chain, becoming the new currency of the digital age and a concept offering business opportunities on a global level. The Big-data era is playing a pivotal role in many companies' strategic decision-making. With Big-data, new dimensions of privacy concerns are also arising[52]. More and more companies are adopting data-driven business models and strategies to obtain and sustain a competitive 'data-advantage' over rivals[53]. In order to maintain control over the excessive processing of personal data, the EU adopted a reform package[54], which aims to ensure the highest data protection standards on the globe. Legal instruments are big influencers of companies' business models by establishing the imperative norms to determine limits of legitimate activities.

[51] Stucke – Grunes 2016, p. 9.
[52] Mantelero 2017, p. 139-154.
[53] Stucke – Grunes 2016, p. 9.
[54] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, hereinafter referred to as GDPR); Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA.

Innovation is often described as the process of translating an idea or invention into a good or service that creates value or for which customers will pay. Through PbD, innovation can appear on product and process level or even at business model level. This is due to the above-mentioned aspects, by which it was shown that organizational measures are better to originate a PbD – compliant environment, as they are more executable and documentable.

Two hallmark characteristics can be concluded here regarding disruptive innovations. First, the entrants to the market can make a foothold by targeting those overlooked segments, delivering their products for a lower price. These products are representing a more-suitable functionality compared to the needs of such customers and the price accessibility is another important aspect. Besides, some disruptive innovations can originate in new-market footholds, by firms creating a market where none existed before[55]. Second, disruptive innovations do not catch on with mainstream customers until quality catches up to their standards[56].

Those responsible with drafting the GDPR either assume the existence of a healthy market for Privacy Enhancing Technologies (PETs) and products or services that may stimulate the applicability of PbD; or that this principle will help to create such a market[57]. Companies with lower resources have the opportunity to embrace the approach of building more consumer-friendly products and seduce categories of consumers, which are neglected by the incumbents. Organizations have to decide how much they dedicate themselves to the PbD principle. Certainly, this principle is a call for a race to gain consumer trust. However, it should not be forgotten that disruption is a process, not a product; far more than that, disrupters often build business models that are very different from those of incumbents and some of such innovations might succeed, while others fail[58]. Moreover, there is no actual guarantee that adopting a disrupter path will lead to a triumph as not all disruptive innovations succeed[59].

Privacy and data protection are interlinked concepts, although not synonymous. While ensuring one's privacy represents the scope, data protection provides the means to protect individual's private interests. In short, what we mean is we protect data to protect privacy. What justifies such

[55] Christensen et al. 2015, p. 47.
[56] Ibid.
[57] Bygrave 2017, p. 118.
[58] Christensen et al. 2015, pp. 48-49.
[59] Ibid, p. 50.

a legal protection? Gellert and Gutwirth believe that privacy and data protection are products of distinct practices and 'regimes of enunciation', such as politics, law, ethics, economy, and religion and so on, and that the challenge is not so much to find the foundational unity "behind" these, than it is to understand how, each being singular, they interact and articulate[60]. Privacy is enshrined in article 8.1 of European Convention for Human Rights (ECHR)[61]. The right to privacy is also consecrated in article 7 of the EU Charter for Fundamental Rights (EUCFR)[62]. Although these are regulated separately, in this chapter the reader will observe the existence of a close relationship between the cornerstone concepts of privacy and data protection. Additional instruments as "secrecy"[63], "confidentiality"[64] and "security"[65] should have their own interpretations, to the extent that a holistic and coherent approach provides a glossary of fundamental concepts throughout the thesis. We consider secrecy and security to be part of a privileged state of an individual, where confidentiality is more closely related to data protection. The concept of privacy includes all of them to a certain extent. Overlaps between privacy and data protection have been provided, where Gellert and Gutwirth argue[66]:

> *All in all, data protection and privacy overlap on a mode whereby data protection is both broader and narrower than privacy. It is narrower because it only deals with the processing personal data, whereas the scope of privacy is wider. It is broader, however, because it applies to the processing of personal data, even if the latter does not infringe upon privacy. Privacy also is broader and narrower: it might apply to a processing of data, which are not personal but nevertheless affect one's privacy, while it will not apply upon a processing of personal data which is not considered to infringe upon one's privacy. It can be said as well that a processing of personal data can have consequences not only in terms of privacy, but also in terms of other constitutional rights, and most obviously, when the processing of data relating to individuals bears risks in terms of discrimination.*

[60] Gellert – Gutwirth 2013, p. 522.
[61] European Convention of Human Rights, www.echr.coe.int
[62] EU Charter of Fundamental Rights, OJ, C 364/10, 18.12.2000
[63] Referred to as the condition of being hidden or concealed.
[64] Referred to as an ethical principle of not disclosing personal information, unless consent permitting disclosure is granted.
[65] Referred to as a set of measures safeguarding a person, building, organization, or country against threats, crimes or attacks.
[66] Gellert – Gutwrith 2013, p. 526.

The highlighted duality of privacy and data protection is a very important aspect to be (re)considered. As provided by Kokott and Sobotta, the distinction between both rights in the EUCFR is not purely symbolic[67]. The case law of the Court of Justice of the European Union (CJEU) and the European Court of Human Rights (ECtHR) is reinforcing the interferences and the differences between them. The ECtHR case law explains from the earliest stages, the concept of personal data with reference to Convention 108[68]. As indicated, the concept of personal data is defined as "any information relating to an identified or identifiable individual"[69], whereas it should cover not only information directly identifying an individual (*e.g.* surname and forename)[70], but also any element indirectly identifying a person (*e.g.* a dynamic IP address)[71].

Considerable amount of cases concerning the issue of personal data collection has been addressed. In context of covert surveillance by authorities, it was provided that the existence of adequate and sufficient guarantees against abuse is essential[72]. The position taken by ECtHR was that powers of secret surveillance of citizens are tolerable only in so far as strictly necessary for safeguarding the democratic institutions[73]. Such interference must be supported by relevant and sufficient reasons and must be proportionate to the legitimate aim or aims pursued[74]. Domestic legislation must provide safeguards that are sufficiently precise, effective, and comprehensive in respect of the ordering and execution of surveillance measures and for the securing of potential redress[75].

Modern-day challenges of data protection also resulted in the technological advances, algorithms, and growing usage of artificial intelligence. To this extent, judgments have been delivered on collection and storage of fingerprints and biological samples[76], facial recognition[77], mobile-telephone provider's practices of storing subscriber information and disclosure to authorities upon

[67] Kokott – Sobotta 2013, p. 223.
[68] Council of Europe Convention no. 108 for the protection of individuals with regard to automatic processing of personal data of 28 January 1981
[69] Amann v. Switzerland, 2000, par. 65; Haralambie v. Romania, 2009, par. 77.
[70] Guillot v. France, 1996, para. 21-22; Güzel Erdagöz v. Turkey, 2008, par. 43; Garnaga v. Ukraine, 2013, par. 36; Henry Kismoun v. France, 2013, par. 25.
[71] Benedik v. Slovenia, 2018, para. 107-108.
[72] European Court of Human Rights 2020, p. 30.
[73] Klass and Others v. Germany, 1978, par. 42; Szabó and Vissy v. Hungary, 2016, para. 72-73.
[74] Segerstedt-Wiberg and Others v. Sweden, 2000, par. 88.
[75] Szabó and Vissy v. Hungary, 2016, par. 89.
[76] S. and Marper v. the United Kingdom, 2008, par. 112.
[77] Gaughran v. the United Kingdom, 2020, para. 70-98.

request[78] or direct access by technical means of authorities to all mobile-telephone communications[79].

Further it was denoted that the risk of harm posed by content and communications on the internet to the exercise and enjoyment of human rights and freedoms, particularly the right to respect for private life, is certainly higher than that posed by the press, particularly on account of the important role of search engines[80]. The discussion on balancing the interests between freedom of expression and personal data protection is under constant evolvement. In this regard, the ECtHR provided that internet archives contribute to preserving and making available news and information[81]. The discretion afforded to member states in striking a balance between the competing rights is greater where news archives of past events, rather than news reporting of current affairs, are concerned[82]. The duty of the press to act in accordance with the principles of responsible journalism by ensuring the accuracy of historical, rather than perishable, information published is more stringent in the absence of any urgency in publishing the material[83]. By mere example, in a case concerning mass flows of personal data regarding taxation of 1.2 million individuals were published in a magazine and subsequently disseminated by means of a text messaging service the ECtHR also decided that there was no public interest of automatic dissemination[84].

In all these cases, there is an interference between privacy interests and personal data protection. They co-exist for the reason of being substantially and formally different. One of the challenges that developers and software engineers are constantly facing is the confusion around the seemingly fuzzy legal terminology that privacy and data protections laws are using. During the work on different case studies, conducted in this research, target groups endorsed this confusion. As shown, these are not interchangeable concepts, however in software engineering the concept of 'privacy engineering' captured ground. Perhaps in the same manner, 'privacy by design' and 'data protection by design' could entail different meaning, but with respect to their content, these are identical.

[78] Breyer v. Germany, 2020, par. 88.
[79] Roman Zakharov v. Russia, 2015, para. 302-305.
[80] M.L. and W.W. v. Germany, 2018, par. 91.
[81] Times Newspapers Ltd v. the United Kingdom (nos. 1 and 2), 2009, par. 45.
[82] Ibid, par. 45.
[83] Ibid.
[84] Case C-73/07, Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland [2017] ECLI:EU:C:2008:727, para. 175-197.

8. CONCLUSIONS

Three competing forces are shaping the concept of European Privacy by Design (PbD): laws and regulations, business goals and architecture designs. These forces carry their own influence in terms of ethics, economics, and technology.

In this research we undertook the journey to understand the concept of European PbD. We examined its nature, application, and enforcement. We concluded that the European PbD is under-researched in two aspects: at organizational level (compared to the individual level); and mainly in the way it is enforced by authorities. We had high hopes especially with regards to the latter, and eager to bring significant scientific contribution on this field. We were interested to learn if data protection authorities are having such impacts looking at European PbD, that can pioneer new approaches to privacy preservation. This is why we elaborated on possible ways to measure their activity, in a manner that both legal and non-legal experts can understand our work.

We promised a response to the research question can the enforcement of European PbD be measured and if yes, what are possible ways to do so? We conducted data analytics on quantitative and qualitative data to answer this question the best way possible. Our response is a moderate yes, the enforcement of PbD can be measured. Although, at this point, we need to settle with only good-enough ways of measure and not dwell into choosing the most optimal or best ways.

One reason for this is that enforcement of PbD cases are highly customized and specific to their own circumstances. We have shown this while creating models to predict the amount of administrative fines for infringement of GDPR. Clustering these cases was a daunting task.

Second reason for not delivering what could be the best way of measure is lack of data availability in Europe. This problem has its roots in the philosophical stance that the European legislator is taking on the topic of data collection within the EU. Lawmakers in Europe certainly dislike programs that collect gigantic amounts of personal data from EU citizens.

Third reason is a causal link between the inconsistent approach between the data protection authorities' practices. This is due to the different levels of competencies, reporting structures, personnel numbers, and experience in the work of data protection authorities.

Looking beyond the above limitations, there are certainly ways to measure the enforcement of European PbD. Our measurements helped us formulate the following statements:

a. **The European PbD operates in 'data saver' mode:** we argue that analogous to the data saving mode on mobile phones, where most applications and services get background data only via Wi-Fi connection, in Europe data collection and data processing is kept to minimal. Therefore, we argue that European PbD is in essence about data minimization. Our conviction that this concept is more oriented towards data security have been partially refuted.
b. **The European PbD is platform independent:** we elaborated in the thesis on various infrastructures and convergent technologies that found compatibility with the PbD principles. We consider that the indeed the concept is evolutionary and technology – neutral.
c. **The European PbD is a tool obligation:** we argue that the authorities are looking at PbD as a tool utilization obligation. In a simple language, companies should first perform a privacy impact assessment in order to find out which tools are supporting their data processing activities and then implement these, as mandated PbD.
d. **The European PbD is highly territorial:** we reached the conclusion that enforcement of PbD is highly dependent on geographical indicators (i.e. countries and counties). The different level of privacy protection cultures are still present in Europe. On a particular level, what is commonly true across all countries is that European PbD mandates strong EU data sovereignty.

It is important to note that, the efficient replacement of personal data with anonymous data results in avoiding the application of certain data protection laws and regulations. Yet, there are multiple laws to preserve privacy. Excluding one sub-set of it (*i.e.* personal data protection) shall not be interpreted as a "free-for-all" ideology, leaving the door open to massive deployment of privacy-invasive business practices. In a similar vein, multiple business models incorporate PbD as an incentive. Decision-makers in organizations with such business models are utilizing PbD as a marketing tool. They strive to extrapolate their strategies to capture and accelerate consumer loyalty. Although, organizations are not always interested in protecting privacy. Examples include

conflict between the business vision and consumer behavior, or constraints due to market conditions.

Lastly, system designers are the pivotal factor in how PbD is conceptualized in IS. Developers, on the other hand, are required to implement the ideas drawn by designers. A natural separation between their roles is a need. They have to establish and maintain a coordinated relationship on addressing different organizational aspects (*e.g.* agreed-upon share of responsibilities) tied to ICT. In searching for answers on the difficulties of developers not able to embed privacy into IS, researchers came to relevant conclusions. Senarath and Arachchilage undertook an empirical investigation that resulted in issues like contradiction between the requirements in the design and privacy requirements, lack of assurance that the implementation was undertaken in a complete and sufficient manner, lack of knowledge and confusion relating to requirements in practice[85]. Hadar et al. found another significant problem: that developers are actively discouraged from making informational privacy a priority, being expected to conform to norms and practices dictated by a negative organizational privacy climate[86].

Another finding was denoted by Bednar et al., which suggests developers are required to battle with lawyers and thus they deal with privacy related issues, mostly because they are required to do so[87]. Despite causing frustration, operationalizing informational privacy is mostly dependent on the developer's mindset. However, placing this responsibility entirely in their hands is an unnecessary burden. In exchange, if the systems designers are actively taking on fulfilling privacy related requirements, the developers feel much safer as being guided by skilled individuals. Continuous and well-designed educational programs for privacy-preserving system designs would ensure preparation of individuals with such profiles.

We also argue that the privacy system designer's role should be separated from the rest of developers. This role should focus on displaying a sketch, which considers PbD in its core. Hence, a privacy focused architecture development is realized. During the design implementation, the privacy system designers should constantly offer guidance to developers. Finally, during the verification and validation, they should provide their seal (*i.e.* approval or acceptance), which

[85] Senarath – Arachchilage 2018, p. 4.
[86] Hadar et al. 2017, p. 20.
[87] Bednar et al. 2019, pp. 137-138.

endorses conformity. A fundamental alteration to take better account from whomever is expected to implement PbD is to change the conjunction in the structure. Thus, what is needed is Privacy *from* Design, not Privacy *by* Design.

9. BIBLIOGRAPHY

1. Acquisti, A., Taylor, C. R., and Wagman, L. (2016): *The Economics of Privacy*. Journal of Economic Literature, Vol. 52, No. 2, 2016; Sloan Foundation Economics Research Paper No. 2580411.
2. Antonopoulos, A. (2018): Infrastructure Inversion by Andreas M. Antonopoulos. Steemit. Available at: https://steemit.com/bitcoin/@pbgreenpoint/infrastructure-inversion-by-andreas-m-antonopoulos [04.04.2021].
3. Barth, S. and de Jong, M. D.T. (2017): *The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review*, Telematics and Informatics 34, no. 7, pp. 1038-1058.
4. Baskerville, R. and Wood-Harper, T. (1996): *A Critical Perspective on Action Research as a Method for Information Systems Research*. Journal of Information Technology. 11. 235-246. 10.1080/026839696345289.
5. Baxter, P., and Jack, S. (2008): *Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers.* The Qualitative Report, 13(4), 544-559.
6. Bednar, K., Spiekermann, S. and Langheinrich, M. (2019): *Engineering Privacy by Design: Are engineers ready to live up to the challenge?*, The Information Society, 35:3, 122-142, DOI: 10.1080/01972243.2019.1583296
7. Bygrave, L. (2010): *Privacy and Data Protection in an International Perspective.* Scandinavian studies in law, 165-200.
8. Bygrave, L., (2017): *Data Protection by Design and by Default: Deciphering the EU's Legislative Requirements.* Oslo Law Review, Volume 4, No. 2.
9. Cavoukian, A. (2013): *Privacy by Design*. Information and Privacy Commissioner of Ontario.
10. Cavoukian, A., (2006): *Creation of a Global Privacy Standard*. Available at: www.pc.on.ca/images/Resources/gps.pdf
11. Christensen, C.M., Raynor, M. and McDonald, R. (2015): *What is disruptive innovation?* Harward Business Review, December 2015.
12. Edmondson, A. and McManus, S. (2007): *Methodological Fit in Management Field Research*. Academy of Management Review. 32. 1155-1179. 10.5465/AMR.2007.26586086.
13. Fabiano, N. (2017): *Internet of Things and Blockchain: Legal Issues and Privacy. The Challenge for a Privacy Standard*. 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Exeter, 2017, pp. 727-734.
14. Gellert, R. and Gutwirth, S. (2013): *The legal construction of privacy and data protection.* Computer Law and Security Review. 29. 522–530. 10.1016/j.clsr.2013.07.005.
15. Hadar, I., Hasson, T., Ayalon, O., Toch, E., Birnhack, M., Sherman, S. and Balissa, A. (2018): *Privacy by designers: software developers' privacy mindset*. Empirical Software Engineering. 23. 10.1007/s10664-017-9517-1.
16. Jonshon, J. (2021): Global digital population as of January 2021, Statista, Available at: https://www.statista.com/statistics/617136/digital-population-worldwide/
17. Kokott, J. and Sobotta, C. (2013): *The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR*. International Data Privacy Law. 3. 222-228. 10.1093/idpl/ipt017.

18. Koops, B.-J., Newell, B.-C., Timan, T., Škorvánek, I., Chokrevski, T., and Galič, M. (2016): *A Typology of Privacy*. University of Pennsylvania Journal of International Law 38(2): 483-575 (2017); Tilburg Law School Research Paper No. 09/2016.
19. Löhe, M. G. and Blind, K. (2015): *Regulation and standardization of data protection in cloud computing*. ITU Kaleidoscope: Trust in the Information Society (K-2015), Barcelona, 2015, pp. 1-6. doi: 10.1109/Kaleidoscope.2015.7383634.
20. Mackenzie, N. and Knipe, S. (2006): *Research dilemmas: Paradigms, methods and methodology*. Issues in Educational Research. 16. 193-205.
21. Mantelero, A. (2016): *From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era*. 10.1007/978-3-319-46608-8_8.
22. Martens, B. and Teuteberg, F. (2011): *Risk and Compliance Management for Cloud Computing Services: Designing a Reference Model*. AMCIS 2011 Proceedings - All Submissions. Paper 228. http://aisel.aisnet.org/amcis2011_submissions/228.
23. Mertens, D.M. (2005): *Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches*. (2nd ed.) Thousand Oaks: Sage.
24. Miles, M. B. (2014): *Qualitative data analysis: a methods sourcebook* / Matthew B. Miles, A. Michael Huberman, Johnny Saldaña, Arizona State University. — Third edition. Thousand Oaks: Sage.
25. Mullarkey, M. T. and Hevner, A. R. (2018): *An elaborated action design research process model*. European Journal of Information Systems, DOI:10.1080/0960085X.2018.1451811.
26. Peffers, K., Tuunanen, T., Rothenberger, M. A. and Chatterjee, S. (2007): *A Design Science Research Methodology for Information Systems Research*. Journal of Management Information Systems, 24:3, 45-77, DOI: 10.2753/MIS0742-1222240302.
27. Senarath, A., and Arachchilage, N., (2018): *Why developers cannot embed privacy into software systems? An empirical investigation*. 211-216. 10.1145/3210459.3210484.
28. Snipe, M. (2021): 'The Markets for Privacy', Yale Journal of Law & Technology. Availablet at: https://yjolt.org/blog/market-privacy
29. Sommerville. I. (2015): *Software Engineering* (8th. ed.).
30. Stucke, M. E. and Grunes, A. P. (2016): *Introduction: Big Data and Competition Policy*. Oxford University Press.
31. van de Pas J. and van Bussel G. (2015): *Privacy Lost - and Found? The information value chain as a model to meet citizens' concerns*. The Electronic Journal Information Systems Evaluation Volume 18 Issue 2, (pp185- 195).

10. CASE LAW

1. C-73/07, Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland [2007], ECLI:EU:C:2008:727
2. Amann v. Svájc, 2000
3. Benedik v. Slovenia, 2018
4. Breyer v. Germany, 2020
5. Garnaga v. Ukraine, 2013
6. Gaughran v. United Kingdom, 2020
7. Guillot v. France, 1996
8. Erdagöz v. Turkey, 2008
9. Haralambie v. Romania, 2009
10. Kismoun v. France, 2013
11. Klass and others v. Germany, 1978
12. M.L. and W.W. v. Germany, 2018

13. S. and Marper v. United Kingdom, 2008
14. Segerstedt-Wiberg and others v. Sweden, 2000
15. Szabó and Vissy v. Hungary, 2016
16. Times Newspapers Ltd. v. United Kingdom (nos 1. and 2.), 2009
17. Zakharov v. Russia, 2015