UNIVERSITY OF TARTU
Institute of Computer Science
Cyber Security Curriculum

Ilhan Çelebi

Privacy Enhanced Secure Tropos: A Privacy Modeling Language for GDPR Compliance

Master's Thesis (30 ECTS)

Supervisor: Raimundas Matulevičius, PhD

Tartu 2018

Privacy Enhanced Secure Tropos: A Privacy Modeling Language for GDPR Compliance

Abstract:
The European Union General Data Protection Regulation (GDPR) compliance is becoming a legal necessity for software systems that process and manage personal data. As a result of that fact, GDPR compliance and privacy components need to be considered from the early stages of the development process and software engineers should analyze not only the system but also its environment. Hereby with this study, Privacy Enhanced Secure Tropos (PESTOS) is emerging as a privacy modeling language based on Tropos methodology, which covers the goal and rule perspective, for helping software engineers by assessing candidate PETs, while designing privacy-aware systems, in order to make them compatible with GDPR. Although in Article 5(2) of the GDPR, the accountability principle requires organizations to show compliance with the principles of the GDPR, (To the best of our knowledge, currently there is no other privacy modeling language especially focuses on the GDPR compliance and enhanced based on Security Risk-Aware Secure Tropos methodology) there were not any practical social modeling languages supply the demand driven by industrial and commercial needs. This is a serious issue for public institutions and private sector in EU-zone because GDPR brings very serious charges for data controllers and data processors, therefore organizations do not feel themselves ready to face with those regulations and software engineers have a lack of methods for capturing change requests of the information systems. This paper applies a structured privacy modeling language that is called as PESTOS which has a goal-oriented solution domain that aims to bring a high compatibility with GDPR by covering Privacy by Design strategies for assessing proper privacy-enhancing technologies(PETs) in a respect of the goal-actor-rule perspective. Among 99 articles of GDPR, 21 articles can be identified as technical level of requirements that PESTOS is able to transform them into GDPR goals needs to be fulfilled in order to support business assets. A survey conducted by identity & security experts validates that proposed model has a sufficient level of correctness, completeness, productivity and ease of use.

Keywords: Privacy, Data Protection, Privacy by Design, Regulation, GDPR, PETs, Privacy-enhancing Technologies, Tropos, Secure Tropos, Compliance, GDPR Compliance, Requirement Analysis, Privacy Modeling, European Union

CERCS: T120 - Systems engineering, computer technology

Privacy Enhanced Secure Tropos: Modelleerimiskeel Euroopa Liidu isikuandmete kaitse üldmääruse (GDPR) vastavuse jaoks

Lühikokkuvõte:
Euroopa Liidu isikuandmete kaitse üldmäärusele (GDPR) vastavuse tagamine saab õiguslikult hädavajalikuks kõigis tarkvarasüsteemides, mis töötlevad ja haldavad isikuandmeid.

Sellest tulenevalt tuleb GDPR-i vastavuse ja privaatsuse komponentidega arvestada arendusprotsessi varajastes etappides ning tarkvara insenerid peaksid analüüsima mitte ainult süsteemi, vaid ka selle keskkonda. Käesolev uuring keskendub viimasel ajal tähepepanu pälvinud modelleerimiskeelele Privacy Enhanced Secure Tropos (PESTOS), mis põhineb Tropos metoodikal hõlmates eesmärkide ja reeglite vaatenurka, mis aitab tarkvarainseneridel hinnata erinevaid Privacy-enhancing Technologies (PET-e) kandidaate, arendades samas privaatsustundlikke süsteeme, et need oleksid GDPR-iga kooskõlas. Kuigi GDPR artikli 5 lõikes 2 sätestatakse, et vastutuse põhimõtte kohaselt peavad organisatsioonid suutma näidata vastavust GDPR põhimõtetele (meie teadmiste kohaselt ei ole praegu veel ühtegi teist privaatsuse modelleerimise keelt, mis keskendub eelkõige GDPR nõuetele ja mis põhineb Security Risk-Aware Secure Tropos metoodikal) ei olnud saadaval ühtegi praktilist sotsiaalset modelleerimise keelt, mis rahuldaks tööstus- ja ärivajadusi. See on Euroopa Liidu piirkonna avalikele asutustele ja erasektorile tõsine probleem, kuna GDPR toob vastutavatele töötlejatele ja volitatud töötlejatele kaasa väga tõsiseid trahve. Organisatsioonid ei oma piisavat kindlustunnet regulatsioonide täitmise osas ja tarkvara inseneridel puuduvad meetodid saamaks ülevaadet infosüsteemide muutmistaotlustest. Käesolevas lõputöös rakendatakse struktureeritud privaatsuse modelleerimise keelt, mida kutsutakse PESTOS-iks. Selle eesmärk on tagada kõrgetasemeline vastavus GDPR-i nõuetele kattes PET-e eesmärk-tegija-reegel perspektiivis hindamiseks ka lõimitud andmekaitse põhimõtted. 99-st GDPR artiklist 21 artiklit saab identifitseerida tehniliste nõudmistena, mile osas PESTOS suudab ettvõttetel aidata GDPR-ist tulenevaid kohustusi täita. Identiteedi- ja turvaekspertide seas läbiviidud uuring kinnitab, et kavandatud mudelil on piisav õigsus, täielikkus, tootlikkus ja kasutusmugavus.

Võtmesõnad: privaatsus, andmekaitse, privaatsuslõime, määrus, isikuandmete kaitse üldmäärus, PETs, Privacy-enhancing Technologies, Tropos, Secure Tropos, vastavus, vastavus isikuandmete kaitse üldmäärusega, nõuete analüüs, privaatsus modelleerimine, Euroopa Liit

CERCS: T120 - Süsteemitehnoloogia, arvutitehnoloogia

Contents

1 Introduction

Privacy fills our lives and thoughts. It showed up first as a social need and, later, also as an essential human right, protected in international laws. Finally, by emergence of Internet era and development of computerization, privacy in cyberspace becomes a key issue for our society. Numerous information technologies affect and change the way of communication, education, business, government, healthcare and many other public services. Additionally, Internet provides a medium for those technologies, as well as it creates a networked, globalized society connected by them. On the other hand, this huge social progress for the society is accompanied by a lot of privacy and security concerns such as data breach, identity theft, surveillance in terms of interception of electronically transmitted information, video surveillance, cyber fraud, and so on.

Since software systems handle more and more sensitive information about users, another key thing to remember is there are social concerns fed by privacy and security issues mentioned previously. As ex-MP of Finland, Esko Seppänen, once said during a debate on information society in European Parliament: "The information society is the chaos society" [EuP]. Indeed, the fact is that he has a point that every member of our generation, potentially can be targeted by cyber attackers, monitored by authorities, and subjected to incorporeal suffering because of loss of privacy. Regarding the so-called "chaos society", it is natural to mention about a discrepancy called "privacy paradox". A research made by Susan Barnes shows that users of social network sites provide a large amount of personal information on public profiles, meanwhile being worried about their privacy [Bar06]. Thus, building privacy-sensitive technologies becomes gradually more important and that is exactly where this thesis idea also comes from.

The 21st century has become the century of Big Data and highly advanced Information and Communication Technologies (ICT) dealing with storing and processing exabytes of data. But most of those advances in ICT threaten privacy increasingly, and have reduced the level of control over personal data. As a result, possibility of a range of adverse consequences as a result of access to personal data has been intensified [SFEb]. Both private companies and public authorities make use of personal data on an unprecedented scale in order to pursue their activities [GDP]. Additionally, members of the society increasingly make personal information available publicly and globally.

Despite all, in the areas of legal practice, privacy comes into prominence as a principal value, and GDPR appears as the most distinct example of that. GDPR brings rights for users and aims to give them control on their personal information and enforce institutions to minimize personally identifiable information on their datasets. However, GDPR costs a lot for the industry. The impact of GDPR noncompliance will be massive for organizations that employ information systems to support their critical business process, as well as manage privacy-sensitive information. It is estimated that in the US, one in four companies with more than 5,000 employees will spend over $1M on GDPR compliance [Bab17].

In spite of this, providers of software development tools and the research community are still far-away from offering tools that enable the intuitive implementation of privacy properties and facilitate interpretational difficulties of software engineers in eliciting of privacy & security requirements from GDPR text. Additionally, privacy modeling is a fresh concept, because for many years privacy was overlooked in IT by being considered just a part of the security, rather than a relevant but different notion.

This master's thesis work aims to propose a sustainable, valid and easy-to-use modeling language, which is called Privacy Enhanced Secure Tropos (PESTOS). It adopts a socio-technical approach based on Secure Tropos language for designing privacy-sensitive software solutions to ensure GDPR compliance. Early stages of the system development life cycle are crucial to the successful development, subsequent deployment, and ongoing evolution of the software system, and thus PESTOS especially aims to capture early privacy requirements for satisfying GDPR [Yu97]. Therefore, It is an honor to emphasize that this dissertation brings novelty and practical approach into the privacy modeling by paying regards to needs of society.

1.1 Research Questions

The main research question (MRQ) of the thesis :
MRQ - How to extend Tropos methodology for building the GDPR compliant software systems?

MRQ can be disintegrated into some minor research questions(RQ).
RQ1 - How to make software systems GDPR-compliant?
RQ2 - How can Tropos methodology implement GDPR-compliance principles?
RQ3 - How to apply enhanced Tropos methodology to the GDPR compliance?

1.2 Solution Domain of PESTOS

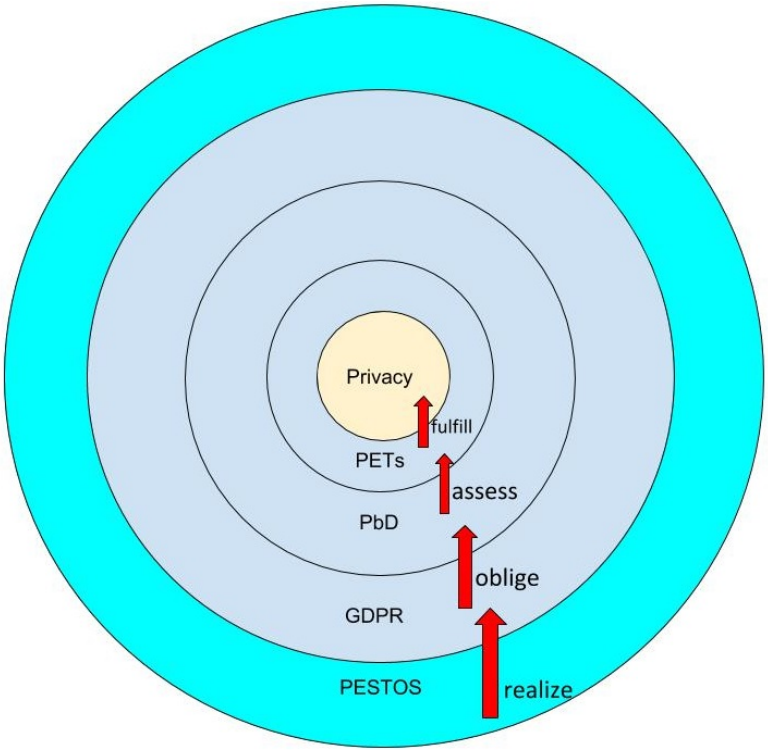

Figure 1. Solution Domain of PESTOS

Figure 1 represents the approach, used to find a solution for the main research question. According to this figure, PESTOS aims to realize GDPR compliance, which obligates system engineers to take into account PbD (Privacy by Design) strategies, while software systems are being built. Those strategies are designed to support the assessment of appropriate PETs (Privacy-Enhancing Technologies), which help to fulfill privacy requirements of the system.

2 Privacy & Privacy Components

In this chapter, many aspects of the notion of privacy are mooted. This chapter aims to provide readers with the necessary background about privacy concepts and to answer several sub-research questions of RQ1. Those questions can be listed as : SRQ1.1 - What is privacy? SRQ1.2 - What are PbD strategies? SRQ1.3 - What are PETs?

2.1 Definition and Ontological Problems of Privacy

A well-defined privacy ontology, that embodies privacy related concept along with their interrelations and a deeper appreciation of its status, would form a great step forward in designing, developing and deploying privacy-sensitive systems by helping software engineers in capturing a clear and robust set of privacy requirements upon them [GGM16] [CW10].

Approaches to addressing privacy issues tend to assume privacy is well understood [CW10], but nothing could be further from the truth. Privacy has serious ontological problems as a concept and those problems became more apparent especially since the information is being processed in digital form by information systems. Furthermore, today there is no doubt of that, typical approach to the privacy-related problems from a security perspective is not anymore working since security is more concerned with safety than with privacy.

Whereas now privacy is re-framed as a many-faceted concept by researchers [HZNF15] [Sol06], before privacy debate has co-evolved with the development of information technology, it was claimed and believed by the legal authorities that privacy could be defined as a unitary concept which is widely cited as "the right to be free of unnecessary public scrutiny or to be let alone" [GH09] [CW10].

According to Solove, "Privacy is a concept in disarray and nobody can articulate what it means" [Sol06]. Heurix et al. claim that "Privacy is a notion known to virtually everybody, yet it is surprisingly difficult to define" [HZNF15]. Furthermore, Gharib et al. make a similar inference, as mentioned in their literature review "Privacy is an elusive and vague concept". They also emphasize in the same paper that although several efforts have been made to clarify this concept by linking it to more refined concepts such as secrecy, control of personal information, person-hood, etc, still there is no consensus, neither on the definition of privacy nor on the analyze methods of privacy [GGM16]. And finally, BeVier identifies the privacy as vague, evanescent and protean [BeV95].

Beyond all of those problems about the definition of privacy, there is also value of privacy is brought up by Chen et al. in their article for proposing a new ontological approach. Chen et al. announce that in order to understand the importance and the role of privacy we should know to what extent is it of value to us. Here, they mention how philosophers justify an object's value as a compound property of intrinsic value and extrinsic value (a.k.a instrumental value). The intrinsic value of something is said to

be the value that that thing has “in itself” or “for its own sake” or “as such” or “in its own right”, while extrinsic value is value that can be generated from intrinsic values. Intrinsic value is absolute but can be contextual, whereas extrinsic one can be subjective and contextual [CW10] [SFEa].

According to Chen et al., privacy has a connotation of "rights" and because of its association to those rights which are allegedly essential to human dignity, they corroborate that privacy has intrinsic values. In human society, individuals develop relationships with dignity and mutual respect. The rights, which are associated with the privacy, aim to prevent the lost of dignity. In that sense, [CW10] states consequently privacy has also extrinsic values.

Justification of privacy is not that easy. Moor, whose alleges in his paper *"Towards a Theory of Privacy in the Information Age"* that dignity differentiates into two types as natural dignity and normative dignity. And he mentions about natural privacy and normative privacy, accordingly. In this regard, he proposed a Core Value Framework (CVF) to uncover common existences in all human cultures as a means to justify the importance of privacy. He calls those set of values, which are shared and fundamental to human evaluation, as core values. When he asks if privacy is a core value, he assumes the concept of privacy has a distinctly cultural aspect that goes beyond the core values. Moor accepts that some cultures may value privacy and some may not, like small tribes who live in Amazon rainforest and have no contact with the outside world [CW10] [Moo97].

2.2 Privacy by Design (PbD)

Privacy by Design (PbD) connotes a development approach for building privacy-sensitive systems and services. Those systems and services could be realized as specific technologies, business operations, physical architectures and networked infrastructure, and even to entire information ecosystems and governance models. The term, PbD, is originally introduced by Cavoukian [Cav11], almost three decades ago and since that time, it is an ever-evolving concept. In fact, PbD does not merely rely on technical solutions but it also involves organisational procedures and business models which implement privacy and data protection principles in order to render them an organization’s default mode of operations. Thus, PbD can be understood as PETs plus privacy enhancing processes [DDH$^+$15] [RB11].

2.2.1 Privacy by Design Principles

Cavoukian claims that, PbD introduces a proactive approach to avoid data breaches and their accompanying harm rather than simply offering mechanism for redness [Cav11]. It is based on practicing 7 principles which are highly inspired by the principles of the Fair Information Practices(FIPs) but aiming also to extend beyond of them and to be operationalized by organizations. Those are respectively: 1) Proactive not Reactive,

Preventative not Remedial 2) Privacy as the Default Setting 3) Privacy Embedded into Design 4) Full Functionality (Positive Sum not Zero-Sum) 5) End-to-End Security (Full Life-cycle Protection) 6) Visibility and Transparency 7) Respect for User Privacy (Keep it User-Centric). These seven foundational principles are characterizing properties rather than instructions for specific measures to be taken [DDH+15] [Cav11]. For detailed explanation about those principles please refer to Appendix III.

Leading agencies and regulators such as the OECD (Organisation for Economic Co-operation and Development), UK Information Commissioner's Office (ICO) and the National Institute of Standards and Technology (NIST) recognized the importance of PbD. Moreover, in October 2010, 32nd International Conference of Data Protection and Privacy Commissioners (ICDPPC) unanimously accepted to concept of PbD as an essential component of fundamental privacy protection and encouraging its widespread adoption. Finally, we would like to mention that PbD takes an important role within the scope of new GDPR, Article 25 obliges to implement appropriate technical and organizational measures by considering the concept of data protection by design [Cav11] [RB11] [KW14].

On the other hand, PbD has also been subject to criticism. Davies assumes that PbD is more a mutual consent concerning the challenges of data protection rather than presenting the targeted solutions. He argues that PbD offers a significant overlap between two domains, which are, the regulative and the engineering. And he notes that the principles of PbD could be motivating, however, they are offering too less technical substances and not enough connection points for economical interests. Rubinstein et al. state that PbD have not yet widely adopted by private sector, but a few firms, and no one among them has conducted before and after studies to determine if they achieved better privacy results. Also, currently, it wouldn't be a wrong accusation if someone said that we lack of tools to realize privacy by design and there are limitations of the approach that induced by its state-of-the-art beside inherent constraints [DDH+15] [Cav11] [RB11] [KW14].

2.2.2 Privacy by Design Strategies

Hoepman [Hoe14] defines eight different privacy design strategies in order to support privacy by design throughout the full software development life cycle. Hereby, this paper is going to call them as Privacy by Design Strategies. Hoepman lists those strategies as *minimize, hide, separate, aggregate, inform, control, enforce and demonstrate* based on both the legal and the technical perspective on privacy protection. In the paper of ENISA [DDH+15], *minimize, hide, separate, aggregate* are categorized under "data oriented strategies" while the rest are considered as "process oriented strategies".

Hoepman outlines a design strategy as a fundamental approach to achieve a certain design goal, that has certain properties that allow it to be distinguished from other approaches which, achieve the same goal [Hoe14]. In a similar manner, he describes a

privacy design strategy as a design strategy that achieves some level of privacy protection as its goal [Hoe14].

Many design patterns are very specific, and therefore cannot be applied directly in the concept development phase [DDH+15]. Hoepman proposes to use privacy strategies to express higher level abstractions than privacy patterns. Thus, a privacy design pattern may sometimes implement several PbD strategies for instance *Attribute Based Credentials* which is a design pattern brings out *minimize* and *hide* as design strategies [Hoe14] [oIb].

Minimize: *Minimize* is counted as the most basic PbD strategy that represents data minimization which expresses that the amount of personal data that is processed should be as minimal as possible. By ensuring that no, or no unnecessary, data is collected, the possible privacy impact of a system is limited. Employing that kind of strategy means one has to answer whether the processing of personal data is proportional (with respect to the purpose) and whether no other, less invasive, way exist to achieve the same purpose [DDH+15].

Common design patterns that implements this strategy are "select before you collect", "anonymisation" and "use pseudonyms" [DDH+15].

Hide: The second PbD strategy is *hide*, states that any personal data, and their interrelationships, should be hidden from plain view, so it cannot be abused easily by an adversary. The strategy does not dictates from whom the data should be hidden but the intent is to hide the information from any untrusted party [Hoe14] [DDH+15].

Common design patterns are the use of encryption (locally, or on the network using SSL), the use of mix networks to hide traffic patterns, or techniques to unlink certain related events (e.g., anonymous cash or attribute based credentials). In essence, the *hide* strategy aims to achieve unlinkability and even unobservability [Hoe14].

Separate: The third PbD strategy is *seperate*, reflects data or process separation. The strategy states that: "The processing of personal information should be done in a distributed fashion whenever possible". By decentralizing the processing or storage of various personal information sources that belong to the same data subject, whole profile of one person cannot be seized. In particular, data from separate sources should be stored in separate databases, and these databases should not be linked. Data should be processed locally whenever possible, and stored locally if feasible as well. Database tables should be split when possible. Rows in these tables should be hard to link to each other, for example by removing any identifiers, or using table specific pseudonyms [DDH+15] [Hoe14].

Abstract: The forth PbD strategy is *abstract* (former it was called as aggregate [Hoe14]), states that: details should be limited as much as possible by summarizing or grouping any storage, collection or operation on personal data, within the constraints of the agreed upon purposes [CHH16].

Aggregation of information over groups of attributes or groups of individuals, restricts the amount of detail in the personal data that remains. This data therefore becomes less

sensitive if the information is sufficiently coarse grained, and the size of the group over which it is aggregated is sufficiently large. Here coarse grained data means that the data items are general enough that the information stored is valid for many individuals hence little information can be matched with a single person, thus protecting privacy of data subjects [DDH+15].

Common design patterns are aggregation over time (for example used to provide some level of privacy protection in smart metering and smart grid systems), dynamic location granularity (used in location based services where the accuracy of the reported location of a user is adapted dynamically to ensure that a reasonable number of other users are at the same location) and Sweeney's k-anonymity concept [Hoe14].

Inform: The fifth PbD strategy is *inform* and it underlines a notification event which also obliged by GDPR (please refer section 3.2, Data Controller) for the sake of transparency. Therefore, data subjects should be adequately informed about which information is processed, for what purpose, and by which means whenever personal information is processed. This also includes information about the measures that are applied in order to protect personal data in the system. Moreover, data subjects should be informed about third parties with which information is shared [Hoe14].

Data breach notifications and Platform for Privacy Preferences (P3P) are possible design patterns for this PbD strategy [Hoe14].

Control: This PbD strategy states that data subjects should be capable to hold the control over the processing of their personal information. In that sense, GDPR brings Data Subject Rights to enable the data subjects to view, update and even ask the deletion of personal data collected about their-selves. Design patterns who implement this strategy are supposed to give users the tools to exert their data protection rights (Data Subjects Rights in the GDPR context) [Hoe14].

Enforce: The seventh PbD strategy is *enforce*, states that: a privacy policy compatible with legal requirements should be in place and should be enforced [Hoe14]. The strategy ensures that the system is compatible with data protection legislation (in our case GDPR), both at the time when the system is developed, as well as when the system is in operation.

Demonstrate: The final PbD strategy is *demonstrate*, that represent to be able to demonstrate compliance with the privacy policy and any applicable legal requirements [Hoe14]. According to GDPR context, this responsibility belongs to data controller and data processor (in case of any).

2.3 Privacy Enhancing Technologies (PETs)

Starting in the 70s, the research community and especially Chaum explored the field of privacy technologies. In 1995, the idea of structuring technology according to privacy principles was discussed among Privacy and Data Protection Commissioners. At the time the main principles were data minimization and identity protection by anonymization and pseudonymization. Finally, this discussion lead to the term Privacy Enhancing

Technologies (PETs) [DDH+15]. Initially, most of these technologies were designed by the individuals for the individual end-user, rather than incorporated into an institutional setting and they were proposing a technological fix for a technological problem which is called electronic communication surveillance [SJBK03].Among the earliest PETs were cryptographic systems that protected communications between individuals from undesired eavesdropping [Phi04].

PETs are technical mechanisms, which eliminate or reduce personal data thereby prevent unnecessary or unwanted processing of personal data, without loss of the functionality of the information system and enable individuals to take control over their personal data [GCW18]. They aim to preserve the privacy of individuals or groups and help achieve compliance with legal frameworks such as data protection legislation and data protection regulation. In this thesis, PESTOS use PbD design strategies in order to elaborate the information system by assessing appropriate PET/PETs while considering its compliance with GDPR, by doing that PESTOS makes law and technology complement each other and form an alliance to protect personal rights by steering the best practices in respect of privacy by design concept.

Nowadays, we do not talk only about one type of privacy violation but many different. On the other hand, there are thousands of PETs existing, most of which are available online. A list of PETs at Stanfords' CyberWiki [Pet11] shows the free technologies aimed at empowering Internet users to gain better control over their data. Today, an IT product is made of pre-existing building blocks and technology. Thus, to select the right PET block as a privacy solution from the list for a given purpose and to utilize it in the most effective way is undeniably important. In order to achieve that, we believe, first, privacy threats must be well identified for each concerning process, that software deals with. Secondly, as ENISA report suggests, we should take account of implementation, applicability, quality and maturity of the anticipated PET [fNE16].

Thus, classification of PETs emerges as a topic at issue. Pullonen et al. propose a classification (classification table can be found in Appendix V) which group PETs according to their application goals and the controls which they aim to bringwith examples of technologies [PMB17]. As Matulevičius et al. mentioned, the same PET should fall into more than one category sometimes, for example encryption can be used for data protection and also for secure communication [PMB17].

Shen and Pearson [SP11] offer another categorization for PETs such linking them to Solove's [Sol06] privacy taxonomy. Solove's taxonomy categorize privacy crimes under four harmful activities such as: information collection, information processing, information dissemination and invasion. Solove's taxonomy endeavors to guide the law toward a more coherent understanding of privacy and it emerges as a widely-accepted framework in the field of privacy law, already.

GDPR, in Article 25, it is written that "the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself,

implement appropriate technical and organizational measures, such as pseudonymization, which are designed to implement data-protection principles, such as data minimization, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation" [GDP], in that way GDPR promotes privacy by design principles and PETs which play a major role in implementing such Privacy by Design approaches into real-world systems [fNE16].

However, except a numerous number of technologies, PETs have not become a standard and widely used component in system design [DDH+15]. According to Borking [Bor11], the main reason that lying behind of that fact is existence of obstacles in adopting PETs. Borking mentions three different obstacles; first one is lack of availability of PETs and lack of user friendliness and the second is lack of support by current regulations, finally the last one is subsistence of drawbacks in infrastructure deployment.

Table 2 in Appendix VI provides a list of PETs that each one shows which kind of PbD strategies it could be associated with and also which category it can fall into according to its characteristic privacy solution. Those strategies represent the ways to achieve a certain level of privacy protection within PbD and PETs are the technologies which turn PbD strategies into the reality by implementing them to the system. In this dissertation, classification of the PETs is predicated upon Table 2.

2.4 Summary

The purpose of that chapter was to find answer for those sub-research questions:
SRQ1.1 What is privacy? : Privacy is a many-faceted concept and has no one-size-fits-all definition, however it is a value which is associated with rights that are allegedly essential to human decency. It is a right of an individual to keep their personal matters and relationships secret, for the sake of dignity and information security.
SRQ1.2 What are PbD strategies? : Those are eight different design strategies in order to support Privacy by Design (PbD), a development approach for building privacy-sensitive systems and services, throughout the full software development life cycle.
SRQ1.3 What are PETs? : PETs are technical solutions, which could eliminate, segregate or minimize identifying particulars thereby prevent unnecessary or unauthorized processing of personal data, without loss of the functionality of the information system and enable individuals to take control over their personal data.

3 The General Data Protection Regulation

In this chapter, GDPR and GDPR-related concepts are covered. This chapter aims to provide readers with the necessary background about EU data protection regime and to answer several sub-research questions of RQ1. Those questions are : SRQ1.4 - What is GDPR? SRQ1.5 - What are GDPR actors? SRQ1.6 - How PETs and PbD strategies can guarantee GDPR compliance?

3.1 Definition and Motivation

General data protection regulation (GDPR) is a data protection regime that brought into force by European Union. According to EU, it is the most important change in data privacy regulations in 20 years [EuG17]. Since May 25th, 2018 the GDPR came into effect as directly applicable in all Member States, that means it does not need to be transposed into any national law in contrast to the Directive 95/46/EC a.k.a Data Protection Directive which is the unsuccessful predecessor of the GDPR in a way [VB17] [Wik].

GDPR aims :

1. to give control back to citizens and residents over their personal data [Wik]
2. to prohibit the personal data citizens provide (being) used for a purpose other than that for which it was collected.
3. to simplify the regulatory environment for international business by unifying the regulation within the EU. (Because GDPR is directly applicable in all Member States) [Wik]
4. to bring legal certainty and remove potential obstacles to the free flow of personal data. (by equalizing the law within the EU) [VB17]

New data regulation brings better incentives for compliance with the regulatory privacy and data protection framework as well as serious sanctions and fines for non-compliant organizations. Thus, PESTOS aims to help system designers and developers in order to plan, implement, maintain and demonstrate GDPR compliance(also known as GDPR accountability) and finally it aims to protect also business sectors from impending serious charges.

Main objectives of this modeling language: to ensure compliant processing of personal data by pointing proper PETs as technical measures should be assigned and build a privacy-sensitive system which is capable according to PbD strategies. Nevertheless, the GDPR does not limit the scope of appropriate measures as it states by Article 32 [GDP].

GDPR applies to any kind of processing of personal data which means any operation or set of operations whether manual or automated that is performed on personal data [VB17].

3.2 GDPR Actors

In GDPR [GDP], we encounter with five different main actors which each of them represents either a natural person or a legal person according to social modeling concept. Section 4.1.1 presents the concept of actors , for more detailed information please refer to that section. The main actors are *supervisory authority, data subject, data controller* (or simply *controller*), *data processor* (or simply *processor*) and *data protection officer,* respectively. A UML class diagram just on below displays GDPR actor generalization.

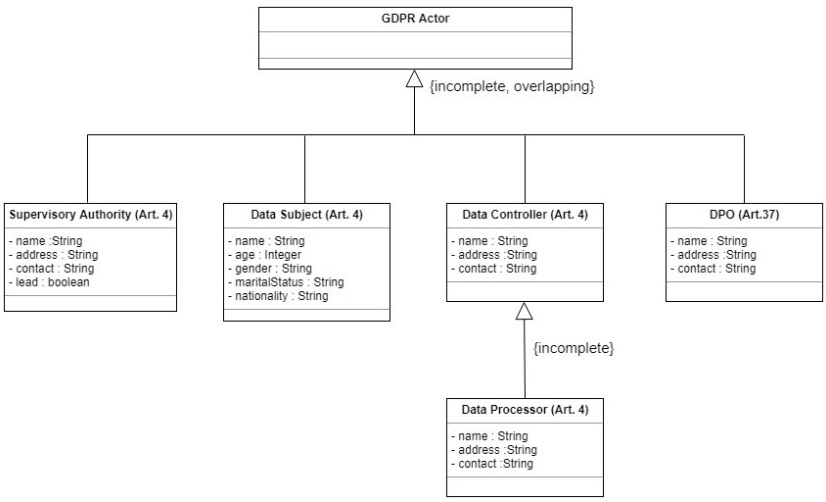

Figure 2. GDPR Actor Generalization

Supervisory Authority: Article 51 of the GDPR outlines the Supervisory Authorities. According to the Article 51 : "Supervisory Authority should be an independent authority that is responsible for monitoring the application of the GDPR, in order to protect the privacy rights of natural persons within the union" [GDP]. Also, we believe it is useful to mention here that supervisory authorities are able to perform their duties or carry out regulatory tasks on behalf of an authority in another EU State, according to Article 6 of GDPR recitals [GDP].

Each Member State shall provide for one or more Supervisory Authorities in order to perform the tasks which set out in Article 57 of the GDPR [GDP]. The tasks listed below:

- Enforcing the GDPR;
- Monitoring the GDPR compliance
- Promoting privacy awareness
- Giving advice to data controllers about their obligations
- Giving advice to data processors about their obligations, in case of any data processors exist;

- Dealing with complaints;
- Conducting investigations;
- Conducting audits;
- Monitoring the impact of technologies on data protection;
- Monitoring the impact of commercial practices on data protection;
- Adopting standard contractual clauses
- Approving binding corporate rules (BCR);
- Maintaining a list in relation to the requirement for privacy impact analyses (PIA);
- Encouraging codes of conduct.
- Encouraging certifications and seals.

Data Subject: GDPR defines the data subject in Article 4 as an identifiable natural person, that can be identified by references such as a name, an identification number, location data and so on [GDP]. By GDPR, Individuals have rights against data processing entities, those rights can be listed as [VB17]:

- to request for information that relating to them.
- to obtain confirmation from the controller as to whether or not their personal data is being processed. (Right to Access)
- to lodge a complaint with the Supervisory Authority. (pursuant to Article 77)
- to require to be clearly informed by the existence of their rights.
- to obtain and, importantly, reuse their personal data.
- to object to the data processing for specified purposes. (Right to Object)
- to obtain from the controller without undue delay the rectification of inaccurate personal data concerning them. (Right to Rectification)
- to have incomplete personal data completed.(Right to Rectification)
- to have their data erased upon a request. (Right to Erasure)
- to transmit their personal data from one controller to another. (Right to Data Portability)
- to receive compensation from the controller or processor(in case any) for the damage they have suffered due to infringement of the GDPR. (Right to Claim Compensation)

Data Controller: Article 4 of the GDPR [GDP] describes the data controller as the natural or legal person, that determines the purposes and means of the processing of personal data, according the union law. Controller can commit his/her duty alone or jointly with other controllers.

Responsibilities of the controller have been stated in Article 24 of the GDPR.

- The controller should take into account the scope and purposes of processing
- The controller shall inform the data subject upon a request
- The controller should take into account the risks.
- The controller shall implement appropriate technical and organizational measures to ensure that processing is performed in accordance with this Regulation.
- The controller shall review those measures.
- The controller should update the measures if the Supervisory Authority thinks updating is necessary.
- The controller shall notify the Supervisory Authority within 72 hours after becoming aware of the data breach.
- The controller should be able to demonstrate compliance with the GDPR.
- The controller shall cooperate with the Supervisory Authority upon the request.

Data Processor: Article 4 of the GDPR describes the data processor as a natural or legal person, who processes personal data on behalf of the controller. The article also clarifies what processing means, it connotes any operation which is performed on personal data [GDP].

Article 28 of the GDPR remarks the responsibilities of the processor such as on below:

- to process the personal data on behalf of the controller.
- to inform the controller of any intended changes on data processing.
- to inform the controller without undue delay after becoming aware of the data breach.
- to cooperate with the Supervisory Authorities

Data Protection Officer (DPO): Article 37, 38 and 39 covers DPO-related matters in the GDPR text [GDP]. According to the articles, the data protection officer is someone who is responsible to involve in issues, which relate to the protection of personal data. In this way, the DPO performs a group of tasks, which are sorted below. Finally, the DPO should also report the highest management level of person, whom he is supported by in fulfilling the tasks. That person is either data controller or data processor, in case there is a third party responsible in processing of data.

- to notify the data controller.
- to notify the data processor in case of any.
- to notify the employees who carry out processing of their compliance obligations.
- to advise the data controller.
- to advise the data processor in case of any.
- to advise the employees who carry out processing of their compliance obligations.
- to monitor compliance with the GDPR
- to monitor compliance with data protection policies of non-EU countries if data process to those countries.
- to audit processing of data through other member state data data protection provisions.
- to audit processing of data with the policies of the controller.
- to audit processing of data with the policies of the processor in case of any.
- to audit assignment of responsibilities.
- to audit awareness-raising of staff involved in data processing operations.
- to provide advice where requested as regards the data protection impact assessment (pursuant to Article 35).
- to monitor performance of data protection impact assessment (pursuant to Article 35).
- to cooperate with the supervisory authority.
- to act as the contact point for the supervisory authority on issues relating to data processing.

3.3 GDPR Compliance Meta-model

In this meta-model, we display the GDPR actors and other GDPR concepts by mapping them regarding associated rights and responsibilities to each actor in order to reach the GDPR compliance.

In section 3.2, we have described the GDPR actors and have listed their rights and responsibilities. Here we are going to introduce the rest of the concepts that take place in the meta-model.

Personal Data: Personal data means any stored information, signs or indications that make possible the identification of a person, directly or indirectly, based on it. For example : social insurance number, location data, phone number, blood type, IP address [VB17].

The following categories of personal data are considered *sensitive*, as set out in Article 9 : *racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, data concerning health or sex life and sexual orientation, genetic data, biometric data.*

Data Processing: It means any kind of treatments of personal data such as [GDP]: collecting, recording, organizing, structuring, storing, altering, restricting, erasing.

Data controller is responsible for lawfulness of those treatments. In Article 6, GDPR states that personal data may be processed only if has at least one lawful basis. Lawful basis could be [GDP]:

- a given consent by a data subject for a specific purpose of processing his/her personal data.
- a legal justification by law for the processing activity
- necessary processing for the performance of a task carried out in the public interest.
- necessary processing for protecting the vital interests of the data subject.

Cross-border Processing: We can mention about cross-border processing if at least one of those two situations emerges [GDP]:

- if a controller/processor is established in more than one Member State, so processing of personal data takes in more than one place.
- if a controller/processor has data subjects in more than one Member State, thus processing of personal data takes in more than one place.

Filing System: It is described in the Article 4, as any structured set of personal data which are accessible according to specific criteria, whether centralized, decentralized or dispersed on a functional or geographical basis [GDP].

Measure: Data protection measures could be technical or organizational. In this study, we propose PETs and PbD patterns as technical measures should be applied to personal data and information system designed to process personal data.

Consent: In Article 4 of the GDPR, consent is identified as any freely given, specific, informed and unambiguous indication of the data subject's wishes by which it, by a statement or by a clear affirmative action, signifies agreement to the processing of its personal data [VB17]. Consent must be clear and distinguishable from other matters and provided in an intelligible and easily accessible form, using clear and plain language. It must be as easy to withdraw consent as it is to give it [GDP].

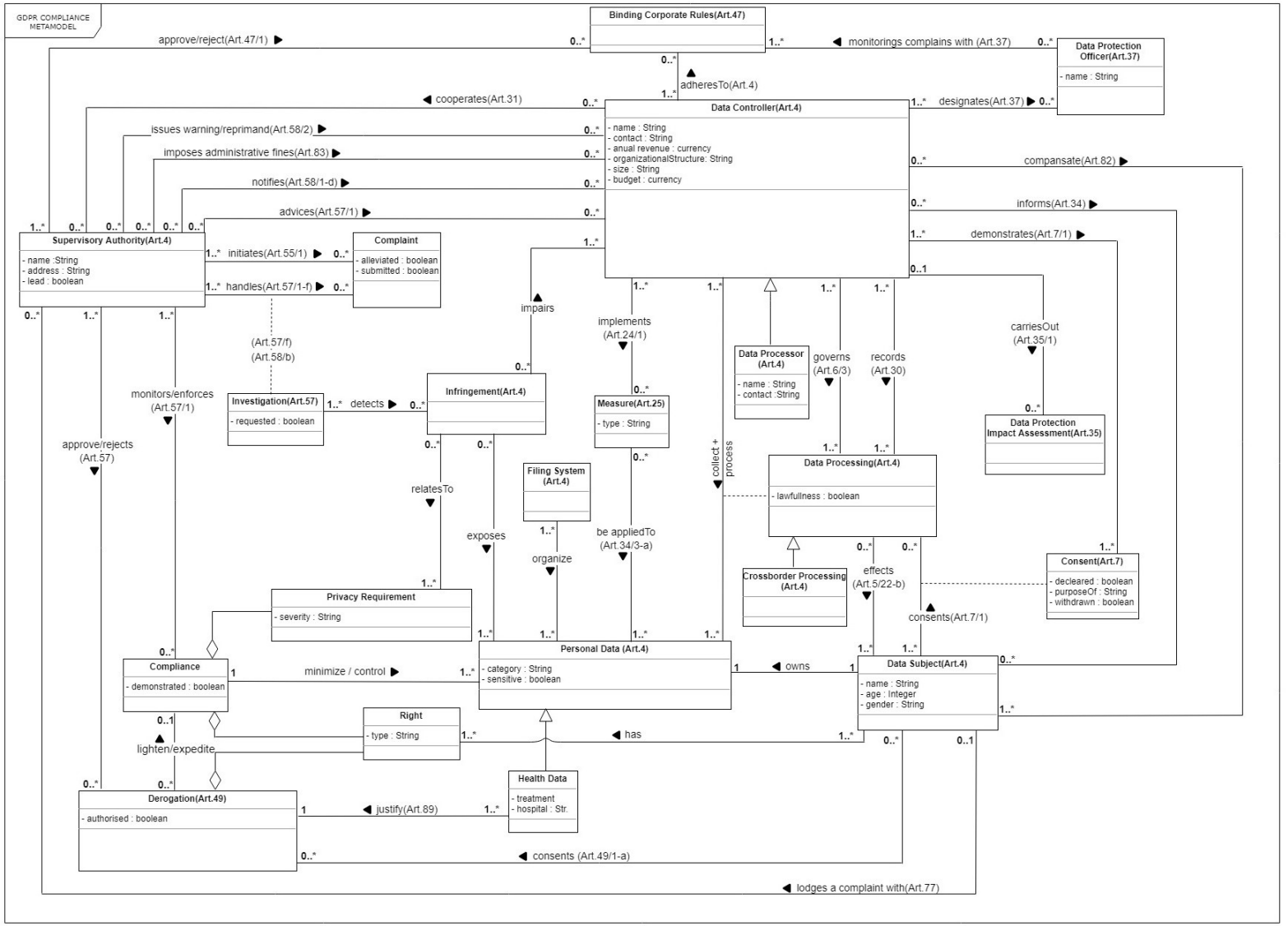

Figure 3. Meta-model of GDPR Compliance

Derogation: This legal term means an exemption from or relaxation of a rule or law. Here, in that paper, whenever we refer that term a reader should understand an exemption from some articles of GDPR under specific situations. For example, Article 66 of the GDPR emphasizes that in a case of urgency if there is a need to act in order to protect the rights and freedoms of data subjects, derogation could be take place based on an official approval given by the Supervisory Authority [GDP].

Data Protection Impact Assessment: In GDPR, Article 35 is all about Data Protection Impact Assessment. It is a preventive study to identify appropriate measures for mitigating the risks to data protection in case an intended processing activity, in particular using new technologies, is considered to result in a high risk to the rights and freedoms of the data subjects. If the results of the assessment do not enable data protection officer (or eventually controller) determine which safeguards could be applied, it will have to consult with the Supervisory Authorities. The latter might issue black- and whitelists in the future that clarify what processing activities will require a Data Protection Impact Assessment [VB17].

Infringement: Infringement is any kind of acts that might break GDPR rules. Personal data breach, organizational ineffectiveness and lack of technical measures are examples of infringements. According to Article 4 of the GDPR, personal data breach means the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed [GDP].

Binding Corporate Rules (BCR): Legally binding internal corporate privacy rules for transferring personal information within a corporate group. BCR are typically used by multinational corporations that operate in multiple jurisdictions, in order to compensate for a lack of data protection in a third country that has not been declared as safe under Article 45 of the GDPR. BCR must be approved by the EU data protection authorities of the member states in which the corporation operates [VB17].

Complaint: According to GDPR,a complaint can be initiated by two different ways:

1.Data subjects can initiate complaints with courts of the appropriate Member State and with the supervisory authority of the Member State where they reside, where they work, or where the infringement occurred. This leaves open the possibility that a controller or processor could face both judicial and administrative proceedings for infringing the Regulation [GDP]. (Article 77/1)

2.A supervisory authority is competent to initiate its own complaints within its Member State [GDP]. (Article 55/1)

Investigation: It means any investigations on the application of the GDPR, including on the basis of information received from another Supervisory Authority or other public authority. Investigations may be initialized based on a complaint, in that case Supervisory Authority, who conducts the investigation, should inform the complainant of the progress

and the outcome of the investigation within a reasonable period according to Article 57/f [GDP].

According to GDPR compliance meta-model that we created: Zero or more (0..*) Supervisory Authorities notifies zero or more data controllers/data processors for an alleged infringement of the GDPR. Conversely zero or more data controllers/data processors can be notified for an alleged infringement of the GDPR by zero or more Supervisory Authorities. Zero or more Supervisory Authorities advices zero or more data controllers/data processors on legislative and administrative measures relating to the protection of natural persons' rights and freedoms. Reciprocally zero or more data controllers/data processors can be advised by zero or more Supervisory Authorities on legislative and administrative measures relating to the protection of natural persons' rights and freedoms. Moreover, zero or more Supervisory Authorities can issue warning or reprimand against to zero or more data controllers/data processors due to infringements. Conversely zero or more data controllers/data processors can be faced with warnings or reprimands from zero or more Supervisory Authorities. Likewise, zero or more Supervisory Authorities can impose administrative fines against to zero or more data controllers/data processors. Reciprocally, zero or more data controllers/data processors can be charged by zero or more Supervisory Authorities. Furthermore, a Supervisory Authority can approve or reject zero or more Binding Corporate Rules. Conversely zero or more Binding Corporate Rules should be approved or rejected by one or more Supervisory Authorities. One or more Supervisory Authorities initiate and later also handle zero or more complaints (some complaints may lodged by data subject, some other may be initiated as Supervisory Authorities' initiatives) which lead to investigations made by Supervisory Authorities eventually. One or more investigations may detect zero or more personal data breaches or/and infringements which expose one or more personal data. A Supervisory Authority or more of them can approve or reject zero or more derogations. Similarly, one or more Supervisory Authorities can monitor and enforce zero or more compliances which are aggregated by law and privacy requirements.

Zero or more data controllers cooperate with zero or more Supervisory Authorities. One or more data controllers designates zero or more data protection officers. Reciprocally zero or more data protection officers can be designated by one or more data controllers. Zero or more data controllers inform zero or more data subjects depends on if there are any data breaches or changes related with data processing. Zero or more data controllers compensate zero or more data subjects if there are any damage occurs as a result of an infringement of the GDPR. One or more data controllers/data processors process one or more personal data. Similarly, one or more data controllers/data processors govern and record one or more data processing. One or more data controllers implement zero or more measures to one or more personal data. Zero or more personal data breaches impair one or more controllers/data processors. Zero or one data controller carries out zero or more data protection impact assessments. One or more data controllers should

demonstrate one or more consents that are given by one or more data subjects for zero or more processing of their personal data. Finally, one or more data controllers adhere to zero or more Binding Corporate Rules.

A data subject owns a set of personal data that can be organized by one or more filing systems. One or more data subjects have one or more rights, which aim to protect their personal data. One or more data subjects consent to zero or more data processing. Likewise, zero or more data processing effects one or more data subjects. Zero or one data subject can lodge a complaint with zero or more Supervisory Authorities. Zero or more data subjects consent to zero or more derogations. Derogations need to be justified by a reason, that can display an exceptional case. For example, data concerning health could be collected by devices without asking any consents, in case of emergency. One or more health data justify a derogation. Thus, zero or more derogations expedite zero or more compliance incidents. Eventually, a compliance controls one or more personal data.

3.4 Summary

The purpose of that chapter was to find answer for those sub-research questions:
SRQ1.4 - What is GDPR? : GDPR is a data protection regime that brought into force by European Union, in order to simplify lawful processing of personal data, while prohibiting unlawful data processing.
SRQ1.5 - What are GDPR actors? : GDPR actors represent natural persons and legal entities who have different rights and competences asserted by GDPR. Those actors are: Data Subject, Data Controller, Data Processor, Data Protection Officer and Supervisory Authority.
SRQ1.6 - How PETs and PbD strategies can guarantee GDPR compliance? : Article 25 in GDPR obliges to implement appropriate technical and organizational measures by considering the concept of data protection by design. PbD strategies elaborate the information system by assessing appropriate PET/PETs as technical measures that will satisfy GDPR requirements of the system by actualizing the concept of data protection by design.

4 Secure Tropos Framework

In this chapter, Secure Tropos and its underlying frameworks, which Secure Tropos is built on top of, are covered. This chapter aims to provide readers with the necessary background about i* Framework, Tropos methodology and Secure Tropos. Also, we discuss how Secure Tropos can be extended for privacy and GDPR related concepts. So, the chapter aims to answer several sub-research question of RQ2. Those questions are : SRQ2.1 - What is Tropos methodology? SRQ2.2 - Why Tropos methodology is suitable for acquiring GDPR-compliance?

4.1 Introduction to Secure Tropos

In order to understand Secure Tropos methodology, first we should mention about i* (iStar) Framework and social modeling. Thereafter, we introduce Tropos methodology and requirement analysis for composing a comprehensive background.

4.1.1 i* Framework and Social Modeling

i* is a modeling and reasoning framework which is adopted by Tropos [DFH16]. The framework specifies the types of objects and the relationships between these objects that can exist in a model which is employing i* [OME00]. Primary aim of i*star to capture early and late system requests based on the concept of strategic social actors.

i* brings social modeling concept with it and from our point of view that is one of the features which make PESTOS suitable for designing privacy-sensitive systems. Social modeling concept carries out a motivation that nowadays our software models should reflect the social characteristics of complex systems since software systems become ever more complex and densely intertwined with the human social environment [Yu09]. In that sense, characterizing and analyzing privacy within a model demands a high social complexity [LYM03].

The rationale of the i* model can provide answers not only for the what or the how, but also "the why a piece of software is developed ?" questions as well. As a matter of fact, it supports a more refined analysis of system dependencies and encourages a uniform treatment of the system's functional and non-functional (such as privacy or security requirements) requirements in that way [BPG$^+$04].

Actors: In i* framework, actors are the central conceptual modeling constructs. They have strategic goals, beliefs, abilities, and commitments. They are active and intentional entities that aim at achieving their goals by exercising independent actions based on their know-how, in collaboration with other actors [DFH16] [Yu09]. Actors are represented graphically as circles [DFH16].

In i*, we focus on intentional properties and relationships rather than actual behavior. The analyses requires to ask some questions like [Yu09]: "what does each actor want?",

"how do they achieve what they want?", "who do they depend on to achieve what they want?" and "what reconfigurations of those relationships can help actors advance their strategic interests?"

Intentionality of actors is made explicit through the actor boundary, which is a graphical container for their intentional elements such as goal, quality, task and resource together with their inter-relationships. In other saying, elements and relationships will appear inside this boundary [DFH16].

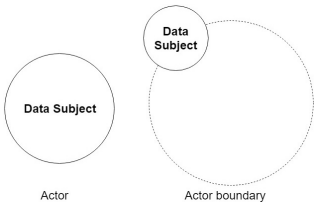

Figure 4. Actor and Actor Boundary

Intentional Elements: An intentional element which is emerging inside an actor boundary denotes something that is desired or wanted by that actor. An intentional element can also appear outside of actor boundaries, as part of a dependency relationship between two actors [DFH16]. Those intentional elements and their graphical representation in the model described on below.

- **Goal:** "a state of affairs that the actor wants to achieve and that has clear-cut criteria of achievement.", in other words, "an actor's strategic interests" It is graphically represented as oval [DFH16].
- **Soft Goal:** "goals without clear-cut criteria whether they are satisfied or not". Soft-goals can represent non-functional requirements (such as privacy requirements). It is graphically represented as cloud or bubbles [Ref11].
- **Task:** "represents actions that an actor wants to be executed, usually with the purpose of achieving some goal." It is graphically represented as hexagon [DFH16].
- **Resource:** "a physical or informational entity that the actor requires in order to perform a task." It is graphically represented as rectangle [DFH16].

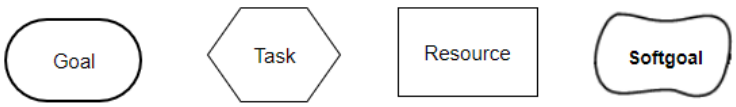

Figure 5. Internal Elements

Intentional Links: Intentional links identify different kinds of refinements and relationships between intentional elements (also called as entities). Moreover, these links can be connected with external dependencies when the reasoning of the analysis goes beyond the actor's boundary [Mou04] [FLM$^+$04].

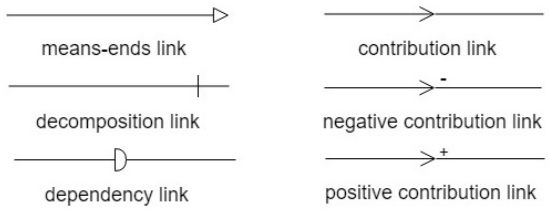

Figure 6. Intentional Links

- **Dependency Links:** They identify dependency relationships between different actors and also between actors and software system by representing those dependencies [Mou04] .
- **Means-Ends Links:** Means-end analysis is employed to identify goals, soft-goals, tasks, and/or resources that can provide means for reaching a goal (end) [Mou04]. Thus, each element connected to a goal by a means-ends link is an alternative way to achieve the goal [FLM$^+$04].

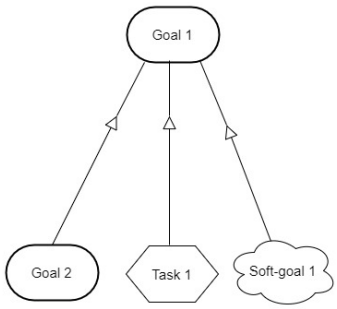

Figure 7. Means-end Analyses

- **Decomposition Links:** Decomposition links define a refinement for a task or a goal [FLM$^+$04]. Particularly, AND/OR decomposition provides an AND and OR decompositions of a root goal/task into sub-goals/tasks [BPG$^+$04]. AND/OR decomposition allows developers to consider alternatives when decomposing the goals/tasks of an actor into sub-goals/sub-tasks. Whereas AND decomposition means all the sub- goals/sub-tasks must be achieved for the root goal/task to be achieved, OR decomposition means that the achievement of one of the sub-goals/sub-tasks leads to the achievement of the root goal/task [Mou04].

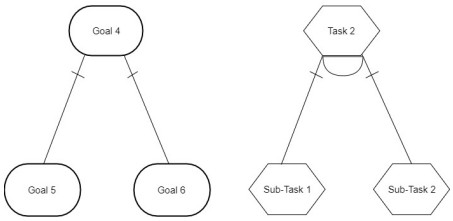

Figure 8. OR decomposition (left) - AND decomposition (right)

- **Contribution Links:** A contribution links describes the impact that an element has on another. This can be negative (-) or positive (+) [FLM$^+$04]. A positive contribution link associates two nodes when one node helps in the fulfillment of the other. A negative contribution link, on the other hand, indicates that a node contributes towards the denial of another node.

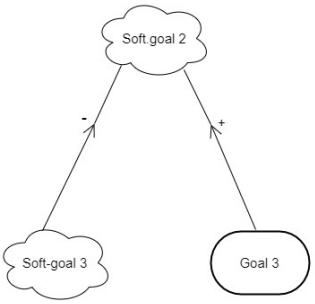

Figure 9. Contribution

Actor Dependencies: Actors depend on each other for goals to be achieved, tasks to be performed, and resources to be furnished [Yu09]. Those dependencies represent social relationships for defining the obligations of actors (dependees) to other actors (dependers) [DFH16] [MGM03]. In order to construct this kind of representation, there are five notions should be involved into it. They are described via [DFH16] on below :

- **Depender:** is the actor that depends for something (the dependum) to be provided.
- **dependerElmt:** is the intentional element within the depender's actor boundary where the dependency starts from, which explains why the dependency exists.
- **Dependum:** is an intentional element that is the object of the dependency.
- **Dependee:** is the actor that should provide the dependum.
- **dependeeElmt:** is the intentional element that explains how the dependee intends to provide the dependum.

Four types of dependencies are distinguished based on the type of dependum, such as resource dependency, task dependency, goal dependency and softgoal dependency [Yu09].

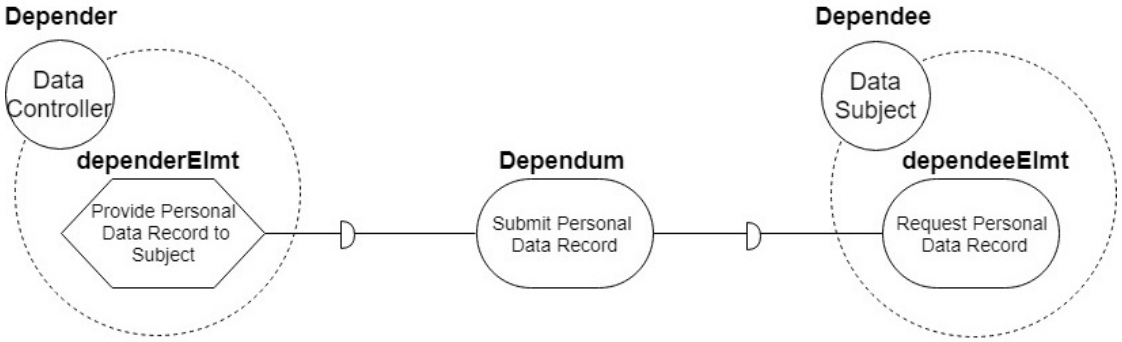

Figure 10. Dependency

4.1.2 Tropos Framework and Requirement Analysis

Tropos is an agent-oriented software development methodology, that embraces the whole software development process with concepts of the agent paradigm. i* offers a small set of well-worked out concepts for modeling social setting, however, i* was intended for early requirements modeling and analysis, so it needed to be revised and extended to cover the scope of the Tropos methodology [Gio]. Tropos is initiated by Mylopoulos as a project [BPG+04] and its main function is to adopt the social components of i* framework from the early requirements level down to the actual implementation. It is intended for becoming a UML-type language and methodology for agent-oriented software but better than the UML precedent [Myl]. "Tropos" means a turn, direction or a way, style, fashion such as "a way of life" in Greek [Hen].

Furthermore, Tropos covers also the very early phases of requirement analysis, thus allowing for a deeper understanding of the environment where the software must operate and congeneric interactions that should occur between software and human agents [BPG+01]. But on top of that, it continues to support and track those requirements until implementation phase with AOP (Agent Oriented Programming) in mind [BPG+01].

There are five main development phases of the Tropos methedology : Early Requirements, Late Requirements, Architectutal Design, Detailed Design and Implementation. In tropos methodology, notions of i* framework such as agent, goal, task and social dependency are used to model and analyze early and late software requirements, architectural and detailed design, and eventually to implement the final system. The requirement analysis in Tropos is divided into 2 different phases: early requirements and late requirements. Both share the same conceptual and methodological approach [BPG+04]. On the other hand, the Architectural Design and the Detailed Design phases focus on the system specification, according to the requirements resulting from the above phases [BPG+04].

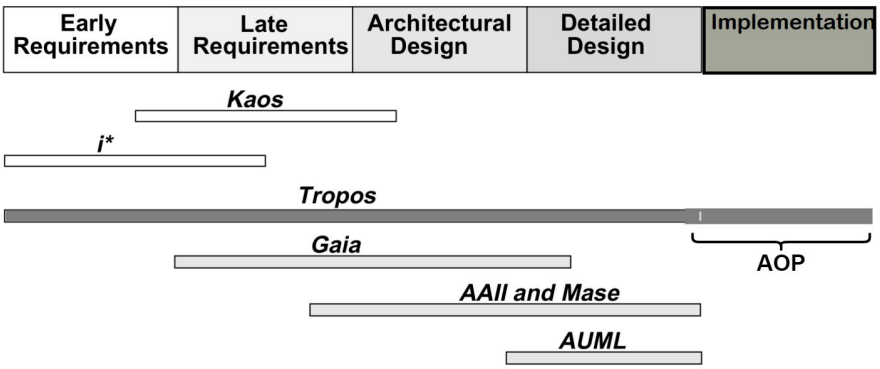

Figure 11. Comparison of Tropos with other software development methodologies [BPG+04] [Gio]

- **Early requirements**
  - Environment (the socio and organizational setting) is analyzed. It means identifying the domain stakeholders and modeling them as social actors [Gio].
    - Which are the main actors?
    - What are their goals?
    - How can they achieve them ?
    - Does an actor depend on another one to achieve its goals?
  - We are not interested in describing the system-to-be [Gio].
  - We are interested in modeling the analyzed environment in terms of relevant actors and their respective dependencies [Gio].
- **Late requirements**
  - The system-to-be is introduced as a new actor of the social domain which is analyzed in the previous phase [Gio].
  - This new actor (the system-to-be) is analyzed in terms of Tropos concepts [Gio].
    - goals decomposition
    - means-end analysis
    - contribution analysis
  - Identifying system's functional and non-functional(e.g: softgoals) requirements after analyze [Gio].

- **Architectural Design**
  - The actor system-to-be is designed [Gio].
  - More system actors (sub-actors) are introduced and are delegated sub-goals to system-assigned goals [BPG+04].
  - Software agents are identified [Gio].
  - Agent capabilities are identified [Gio].
- **Detailed Design**
  - Each architectural component is defined in further detail in terms of inputs, outputs, control, and the security aspects analysed in the previous stages [MGM03].
  - For this stage, Tropos is using elements of UML to complement the features of i* [MGM03].

Tropos extention of intentional elements : Tropos brings extension to the initial elements which are introduced in i* section by two additional concepts : capability and plan.

- **Plan:** "represents, at an abstract level, a way of doing something. The execution of plan can be a means for satisfying a goal or for satisfying a soft-goal" [BPG+04].
- **Capability:** " represents the ability of an actor of defining, choosing and executing a plan for the fulfillment of a goal, given certain world conditions and in presence of a specific event" [BPG+04].

Modeling Activities: In order to develop and to refine Tropos models, multifarious activities, such as actor, dependency, goal, task, and capability modeling, and various graphical diagrams, such as actor, goal, capability and plan diagrams, are used in the Tropos methodology [Mou04]. Please refer to Appendix VIII for further information about those modeling activities.

4.1.3 Secure Tropos

Secure Tropos [MG07] is an approach introduced by Mouratidis et al.and based on the Tropos methodology, with the scope to define security during system development [Mat17]. Secure Tropos considers the basic Tropos concepts such as dependency, goal, task, resource, and capability and adds security concepts such as security constraint, secure goal, secure plan, secure resource, and secure capability [PIM12]. The major aspects of the methodology are [Mat17]: (i) social issues of security are analyzed during the early requirements stage; (ii) security is considered simultaneously with the other

requirements of the system-to-be; (iii) security is addressed in depth during the system design phases.

According to Mouratidis, Tropos methodology was partially tackling modeling secure software systems by allowing developers to use soft-goal concept in order to capture security requirements together with other non-functional and functional requirements. However, utilization of the soft-goal concept was not sufficiently address those requirements may define constraints on the system that effect its stakeholders [Mou04]. The reason for this is soft-goals mostly represent qualities which are properties or characteristics of the system that its stakeholders care about, whereas constraints are restrictions, rules or conditions imposed to the system and unlike qualities are non negotiable [Mou04].

4.2 Secure Tropos for Security & Privacy Modeling

In previous section, we extracted the motivation behind the idea of extending the Tropos as growing into the Secure Tropos, by cited a several articles. If it has to be mentioned briefly one more time, we would like to address [IMJ10]: "In order to develop more secure software systems, security needs to be considered from the early stages of the development process and software systems developers should analyze not only the system but also its environment. This is important since all software systems operate within an environment and the various elements : stakeholders, users, relevant laws and regulations of the environment might influence the security aspects of the system." Here, a question should be raised: What about privacy needs and privacy-aware systems?

We believe that the need for developing privacy-aware systems constantly is growing due to legal regulations go on the stage for advocating human rights and also increasing privacy threats which become more harmful and frequent by means of misusing fast-enhancing technology. Thus, privacy needs to be considered also from the early stages of the development and should be planned according to special privacy design techniques offered by PbD based on the captured privacy requirements. However, Islam et al. assert in [IMJ10] that Secure Tropos language does not provide support for modeling legal dependencies which are needed for eliciting privacy requirements from legal contexts such as GDPR text.

Moreover, privacy-related concepts and privacy-oriented processes are overlooked by regarding them as a part of security matters within the Secure Tropos framework. There is neither any good definition of privacy nor any segregation on modeling entities and activities, as between privacy-related and secure, offered by the Secure Tropos. In this section, we exert our effort to cover those entities and activities in order to reveal the fact we mentioned above and we propose to extend the current Secure Tropos ontology (naturally it means also extension of the current Tropos ontology) by novel notations and well-structured methods in order to fill this gap, and thus enable the Tropos methodology to introduce needs of privacy engineering with social system engineering for focusing GDPR compliance.

In order to fulfill this proposal, firstly, we believe that several new concepts which are shaped by keeping GDPR on the mind; such as privacy constraint, privacy dependency, privacy mechanism (also known as PET), privacy feature (also known as PbD strategy), GDPR goal (as protection objectives cover both privacy and security objectives) and finally privacy reference diagram, must be on-boarded as separate concepts rather than being considered as a part of security modeling. Secondly, existing concepts of the Secure Tropos such as actor, goal, task, resource, capability must be utilized according to definitions and conceptions in the GDPR text.

In this regard, Secure Tropos gives us very convincing reasons for legitimacy of that prospective preference. Since Mouratidis propounds in [Mou04], the Tropos is a widely known and published agent oriented software engineering methodology. Therefore, it is well-supported by many researchers by way of various projects. Moreover, it is requirements-driven that specifies both the environment of the system and the system itself, by using the same concepts and notations through the whole development life-cycle. Finally, it is easily extensible and well integrated with other approaches and therefore existing work can be considered and incorporated within the proposed approach [Mou04].

4.2.1 Security-oriented Concepts and Overlooked Privacy in The Secure Tropos

In this section, we are going to present security-oriented concepts and their notations in the Secure Tropos methodology as described in the thesis of Mouratidis. Also this section has some examples about how privacy is overlooked under the security-oriented concepts.

Security Constraint: In [Mou04], Mouratidis takes constraints as a separate concept of Tropos ontology and propose a more tailored definition for them: "A restriction that can influence the analysis and design of the multi-agent system under development by restricting some alternative design solutions, by conflicting with some of the requirements of the system, or by refining some of the system's objectives".

A constraint restricts zero or more (0 ... *) dependencies, goals and/or tasks. Conversely zero or more (0 ... *) dependencies, goals and/or tasks are restricted by one or more (1 ... *) constraints. When a constraint is imposed to a goal (or task), two analysis processes are applied: Constraint decomposition, which aims to further decompose the constraint into sub-constraints; and goal introduction, which identifies prospective goals that the constraint might introduce to the system (known as goal introduction)[Mou04]. During the process of goal introduction, the purpose of these goals is to help towards the achievement of the constraint. In other words, the developer refines the goals of an actor to allow the satisfaction of a constraint [Mou04].

Mouratidis emphasizes that constraints can be valuable in modeling assorted non-functional requirements, from this point of view, Mouratidis defines a security constraint. According to his definition, secure constraint is "a restriction related to security issues,

such as privacy, integrity and availability, which can influence the analysis and design of a multi-agent system under development by restricting some alternative design solutions, by conflicting with some of the requirements of the system, or by refining some of the system's objectives" [Mou04].

It is very noticeable, in that definition above, how privacy overlooked by being counted as a security issue in Mouratidis paper. However, a highly secure information system can be non-competent and disregardful in the meaning of personal data privacy. Imagine a software developed as a customer loyalty program for a chain store which keep a lot of records of customers like shopping date, shopping time, payment way, name of brands and categories of goods that customers buy, brunch where shopping activity actualizes, etc. without taking consent of customers. Even this information system is equipped with very proactive and cutting edge technologies for security, can we call it privacy-sensitive?

In order to define GDPR-related restrictions, which may conflict with some system requirements, we need to have an independent concept of privacy constraint segregated from the security constraint. Section 5.2 presents the concept of privacy constraint, for more detailed information please refer to section 5.2.

Secure Dependency: Mouratidis introduces a special class of dependency as secure dependency. A secure dependency introduces security constraint/constraints that must be fulfilled for the dependency to be satisfied. Both the depender and the dependee must agree for the fulfillment of the security constraint in order for the secure dependency to be valid. That means the depender expects from the dependee to satisfy the security constraint/constraints and also that the dependee will make an effort to deliver the dependum by satisfying the security constraint/constraints [Mou04].

Mouratidis classifies secure dependencies under three classes based on who introduces security constraint. Those are distinguished as *Dependee Secure Dependency, Depender Secure Dependency, Double Secure Dependency* respectively. In *Double Secure Dependency*, security constraints are introduced on both sides [Mou04].

Previously, on the last paragraph of the security constraint headline, we emphasized that there is a need for introducing a concept of privacy constraint in order to build a privacy modeling language. From this point of view, we can also talk about a concept of privacy dependency between the actors. Section 4.2.2 presents the concept of privacy dependency, for more detailed information please refer to section 4.2.2.

Secure Entities: As we mentioned previously in section 4.1.3, Secure Tropos considers the Tropos entities and adds them security concepts. In this sense, as Tropos provides concepts of goal, task, resource, et cetera Secure Tropos provides concepts of secure goal, secure task, secure resource which are known as secure entities.

A **secure goal** represents the strategic interests of an actor with respect to security. Secure goals are mostly introduced in order to achieve potential security constraints, which are imposed to an actor or exist in the system. Nevertheless, a secure goal does

not particularly define how the security constraints can be fulfilled, since alternatives can be considered. On the other hand, a **secure task** is defined as a task that represents a particular way for satisfying a secure goal [Mou04]. Eventually, a **secure resource** can be defined as an informational entity that is security-critical for the system under development [Mou04] [Mat17].

With a similar motivation, PESTOS extends the Secure Tropos language by introducing *GDPR entities* such as GDPR goal, GDPR task and GDPR resource, while, on the other hand it involves the secure entities like Assignment, Contribution and adapt some of them for new privacy entities. Section 5.2 introduces the GDPR entities, for more detailed information please refer to section 5.2

Secure Capability: In [Mou04], secure capability is described as the ability of an actor/agent to achieve a secure goal, carry out a secure task and/or deliver a secure resource.

By PESTOS, instead of capability, GDPR Rights appear as a similar but GDPR specific concept represent the powers and the legal rights that actors are entitled to execute under favour of the GDPR.

Secure Reference Diagram: It is a diagram that constructed by actualization of a modeling activity, which involves the identification of security needs of the system-to-be and introduces problems, related to the security of the system (such as threats and vulnerabilities), possible solutions (usually these solutions are identified in terms of a security policy that the organization might have) to those security problems. Thus, the security reference diagram represents the relationships between security features, threats, protection objectives, and security mechanisms [Mou04].

Mouratidis explains the main purpose of the security reference diagram in [Mou04], according to him, it allows flexibility during the development stages of a multi agent system and also to save time and effort. He claims that many systems under development are similar to systems already in existence. Therefore the security reference diagram can be used as a reference point that can be modified or extended according to specific needs of particular systems.

Mouratidis prefers to use the same notation of Tropos for elements of the security reference diagram [Mou04]. Therefore, concepts from the Tropos methodology such as soft-goals, goals and tasks are used to model security features, protection objectives and security mechanisms respectively. His motivation behind this decision is allowing developers to work with well-known concepts and allowing them to use the same concepts throughout the development process.

With this work, we propose to have a concept of Privacy Reference Diagram, supported with a similar motivation which Security Reference Diagram has, since GDPR compliance requires to identify both security and privacy needs of the system-to-be. However, in contrast to preference of Mouratidis, we submit new notations for some elements of Privacy Reference Diagram, which are Privacy Features, Privacy Constraint

and Privacy Mechanism. In order to make software engineers to concentrate better on privacy-oriented problem analyzing and aware about independent concepts of privacy, that are different than the concepts of security, we offer to follow that approach since Tropos methodology is highly applicable to extend its notations together its application area.

- **Security Features:** (also protection properties) As Mouratidis identifies, "it represent features associated to security that the system-to-be must have". In Mouratidis' work the concept of a soft-goal is used to capture security features on the security reference diagram. He take this decision because the concept of soft-goal is used, in the Tropos methodology, to model quality attributes for which there are no a priori, clear criteria for satisfaction but are judged by actors as being sufficiently met. He claims that similarly, security features are not subject to any clear criteria for satisfaction. "Examples of security features are privacy, availability, and integrity" [Mou04].

  Here by that definition, we can clearly re-comprehend that privacy was counted as a security feature by Mouratidis. Within PESTOS system-to-be (a GDPR compliant system) has clear, well-distinguished privacy features like anonymity, pseudonymity, unlinkability, et cetera.

- **Protection Objectives:** They represent a set of principles or rules that contribute towards the achievement of the security features. These principles identify possible solutions to the security problems and usually they can be found in the form of the security policy of the organization [Mou04].

  In PESTOS point of view, if there is a need to define "privacy objectives", that need is supposed to be fulfilled by seven principles of PbD.

- **Security Mechanism:** It represent standard security methods for helping towards the satisfaction of the protection objectives. Mouratidis claims that some of these methods are able to prevent security attacks, whereas others are able merely to pinpoint security breaches. The concept of a task is used to model security mechanisms in Mouratidis' work. As he clarifies, this decision took place, because in Tropos, a task represents a particular way of doing something, such as the satisfaction of a goal. In the same sense, a security mechanism represents a particular way of satisfying a protection objective [Mou04].

- **Threats:** They represent circumstances that have the potential to cause loss; or problems that can put in danger the security features of the system. Mouratidis introduced a new notation within Secure Tropos since Tropos doesn't provide any related concept to model threats [Mou04].

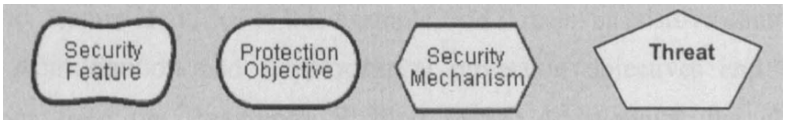

Figure 12. Notations of elements of Security Referance Diagram [Mou04]

4.2.2 Privacy-oriented Extension of Secure Tropos Framework

This thesis presents a language to extend Secure Tropos framework in order to support the consideration of GDPR during the development of privacy-sensitive, compliant software systems. In particular, the PESTOS framework enables software developers (i) to correctly elicit privacy requirements from the GDPR. (ii) to trace these requirements during the development stages in order to ensure design phase supports to build compliant systems that can demonstrate GDPR accountability. (iii) to consider privacy and security simultaneously with the other requirements of the system-to-be.

Existing notations of Secure Tropos does not completely support privacy components and it is not capable to represent GDPR-related concepts. PESTOS extends Tropos methodology in order to : 1)Introduce new privacy-related concepts and GDPR-related concepts 2)Introduce GDPR-oriented processes, Privacy-oriented processes and the integration of this process into the development stages of the Tropos methodology.

4.3 Summary

The purpose of that chapter was to find answer for those sub-research questions:
SRQ2.1 - What is Tropos methodology? : Tropos is an agent-oriented software development methodology which is built on top of i* framework.
SRQ2.2 - Why Tropos methodology is suitable for acquiring GDPR-compliance? : Tropos covers the very early phases of requirement analysis, thus allowing for a deeper understanding of the environment where the software must operate and congeneric interactions that should occur between software and human agents. But on top of that, it continues to support and track those requirements until implementation phase. Additionally, Tropos methodology is well supported by available materials and tools. It is also easily extensible for analyzing specific requirements of the system, which in our case GDPR requirements.

5 Privacy Enhanced Secure Tropos

This chapter will be focusing on explaining GDPR compliance approach of PESTOS through privacy modeling and also the language itself by covering its abstract syntax (meta-model), semantics and concrete syntax. It will also illustrate how PESTOS extends Tropos methodology for building the GDPR compliant software systems.

5.1 Meta-model

Meta-model is used here as a means to define syntax and semantics of PESTOS formally in order to analyze software models that are created by using the language.

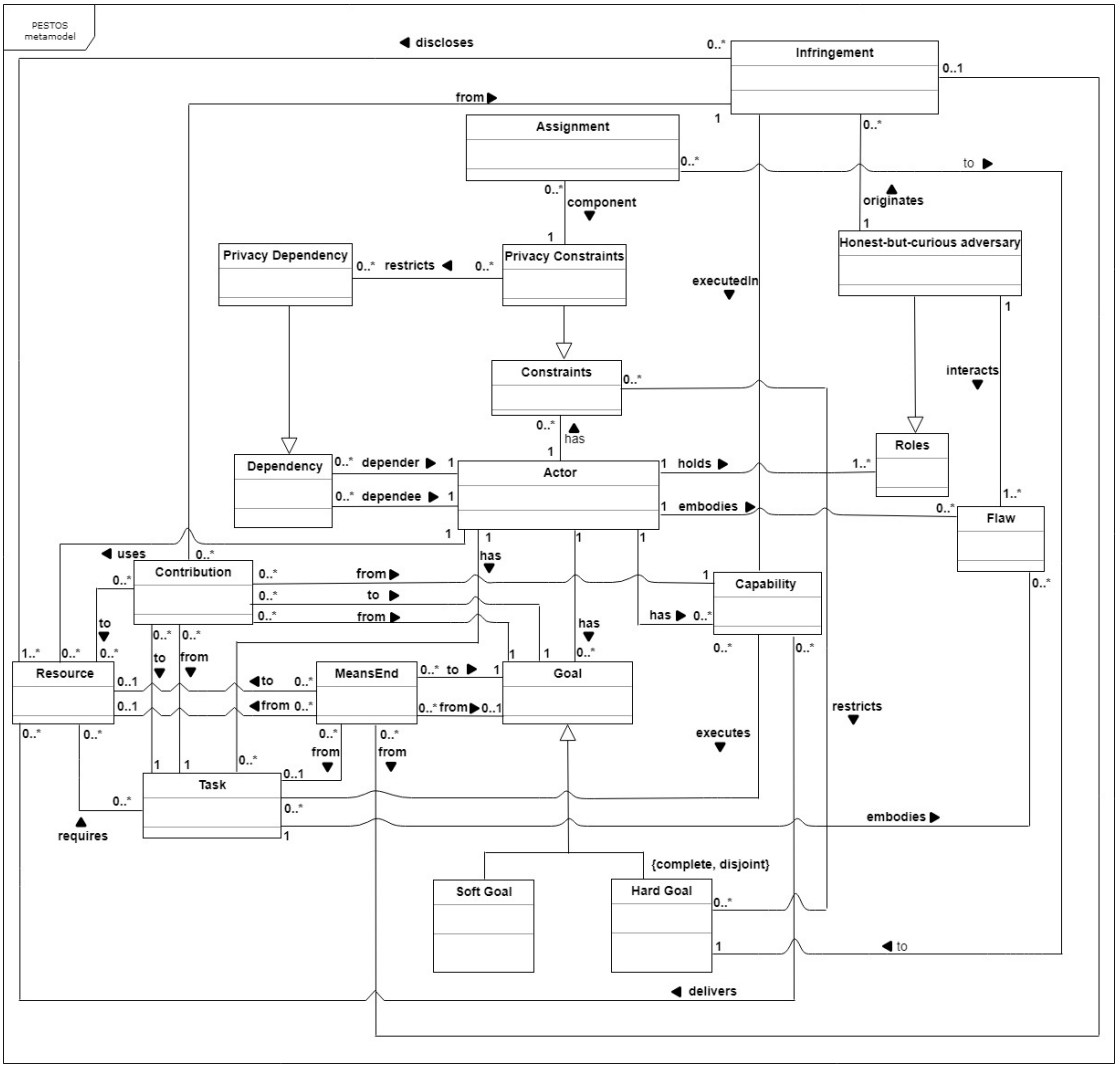

Figure 13. PESTOS meta-model

According to meta-model, one actor has zero or more constraints which some of them could be privacy constraints. Zero or more of those privacy constraints restricts zero or more privacy dependencies between different actors. One actor has zero or more strategic goals which can be restricted zero or more constraints. One goal may be accompanied by zero or more assignments.

An actor also has zero or more tasks which contributes different goals. One goal can be supported by zero or more contributions which might come from a task or another goal. One or more negative contributions might come from an infringement. One or more infringements can be originated by a honest-but-curious adversary. Honest-but-curious adversary (HBC) here represents an actor who doesn't intent to prepare an attack against to a filing system but can misuse to any flaws that an actor includes or/and a task includes in order to reveal personal data of individuals due to curiosity. One HBC initiates one or more infringements which disclose one or more system resources.

An actor has zero or more capabilities for executing zero or more tasks and delivering zero or more resources. A capability can contribute to zero or more resources and goals of both the actor who has that capability itself and the other actors. A capability also can contribute to a task. One actor can use zero or more resources. An actor embodies zero or more flaws that allows HBC to interact with it.

5.2 Semantics and Concrete Syntax

The entities of the Tropos methodology need to be extended with GDPR compliance and data privacy in mind. Therefore, the following figures (Fig. 14 and Fig. 18) show how concepts from the GDPR compliance meta-model can be combined by concepts from the PESTOS meta-model in order to introduce GDPR constructs.

Privacy Constraint: Privacy constraint is a restriction related to privacy issues that can interfere with design of the multi-agent system under development by restricting some alternative design solutions, by conflicting with some of the requirements of the system which are not eligible according to PbD concept and may refine some objectives of the system. As we mentioned before in section 4.2.1, a constraint restricts zero or more dependencies, goals and/or tasks. Hereby, it is important to recall Privacy by Design strategies of Hoepman that we introduced in section 2.2.2. In his paper [Hoe14] Hoepman states that design strategies do not necessarily impose a specific structure on the system but they certainly limit the possible structural realizations of it. It shows how to utilize PbD strategies and PbD patterns in order to describe *privacy features* that will clarify *privacy constraints*.

Privacy Dependency: A type of dependency that introduces privacy constraint/-constraints that must be fulfilled for the dependency to be satisfied. Both the depender and the dependee must agree for the fulfillment of the privacy constraint in order for the privacy dependency to be legit. Accordingly, the depender anticipates from the dependee to satisfy the privacy constraint/constraints and meanwhile the dependee will

make an effort to deliver the dependum by satisfying the privacy constraint/constraints. A dependum may get a contribution by a privacy feature in the case of satisfying the privacy constraint/constraints.

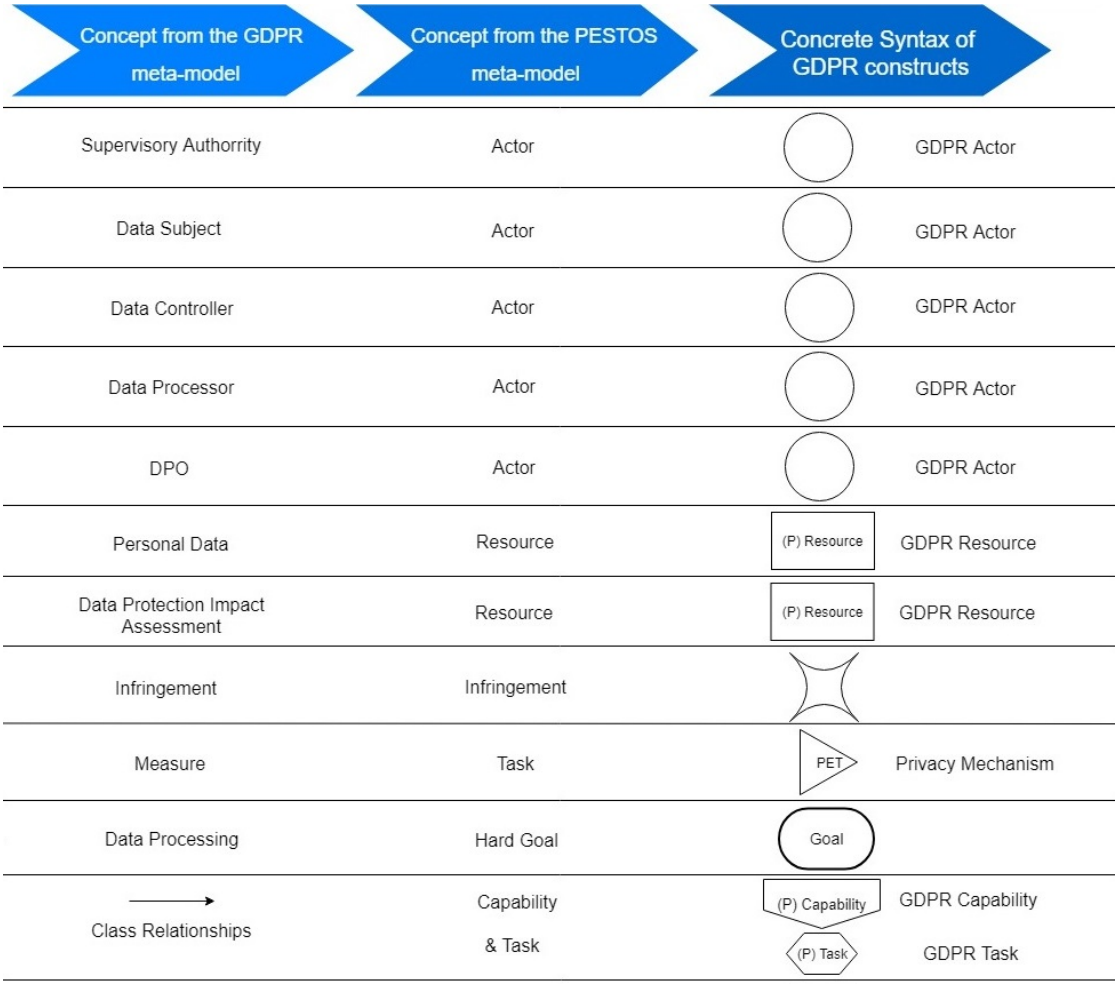

Figure 14. GDPR-related concepts I

We can set out three different privacy dependencies based on who introduces the privacy constraint. They will be listed as *Dependee Privacy Dependency, Depender Privacy Dependency, Double Privacy Dependency* and will be displayed respectively on below.

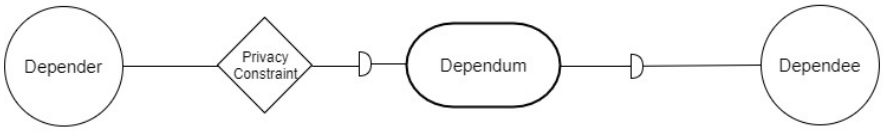

Figure 15. Dependee Privacy Dependency

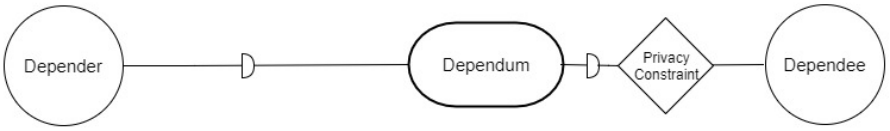

Figure 16. Depender Privacy Dependency

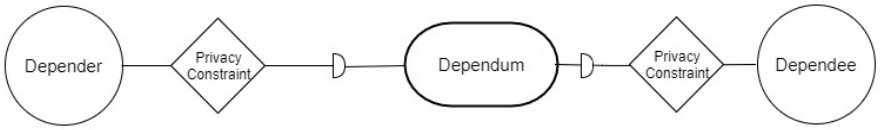

Figure 17. Double Privacy Dependency

GDPR Actor: It represents either a natural or legal person who is subject of the GDPR and that's why has certain goals and abilities in respect to GDPR. GDPR actors are represented in section 3.2, please refer it if you need more information.

GDPR Resource: An informational entity that is security-critical and privacy-critical for the system aiming to be GDPR compliant, for example : personal data. The GDPR resource is graphically represented as a *Resource*, but it is denoted with a (P) label additionally.

GDPR Goal: represents the strategic interests of a *GDPR actor* with respect to legal compliance and data protection. GDPR goals are derived from *Rights of the Data Subject* within the GDPR. In this sense, we can talk about a kind of bilateral relationship between the data subject and the other GDPR actors in order to succeed a mutual achievement based on the *Rights of the Data Subject*. Just to clarify, subject should be able to execute the rights in any time and the other GDPR actors should make that certain.

GDPR goals are well-defined and well-circumscribed since they are derived from the *Rights of the Data Subject* within the GDPR text. In this context, GDPR goals are introduced as hard goals according to Tropos methodology by PESTOS. They are mostly introduced in order to accomplish potential privacy constraints which are imposed to a *GDPR actor* or already exist in the system. Although *privacy constraints* restrict GDPR goals, a GDPR goal does not particularly define how the *privacy constraint* can be achieved.

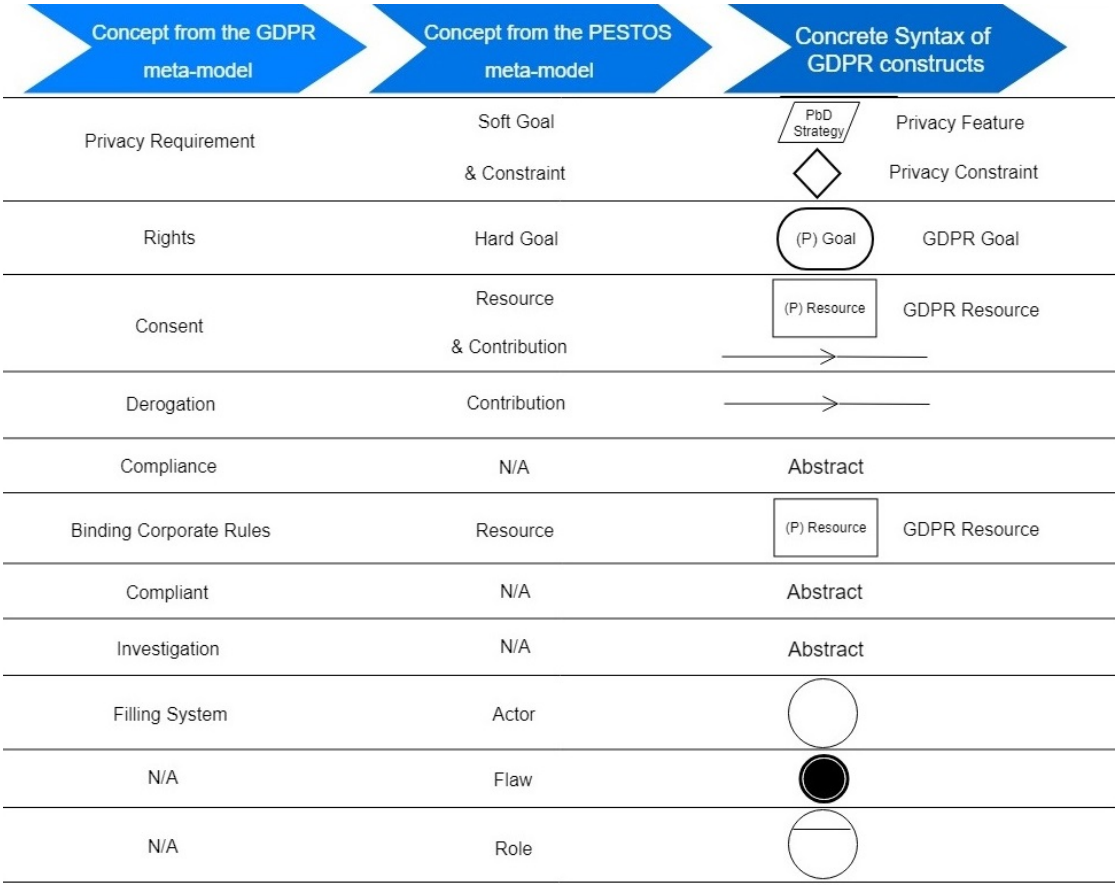

Figure 18. GDPR-related concepts II

The following table (Tabl. 1) displays GDPR goals with related GDPR articles that explains what they are all about and which GDPR tasks that users shall perform in order to reach those goals.

GDPR Task: GDPR task represents actions that a GDPR actor wants to perform in order to satisfy its GDPR goals. The GDPR task is adapted from task concept of Tropos methodology. Its graphical representation is similar with the graphical representation of *Task*.

GDPR Capability: GDPR capability represents legal powers, responsibilities and abilities of a GDPR actor to fulfill its GDPR goals. Those legal powers, responsibilities and abilities are defined under article 15, 16, 17, 18, 19, 20, 21, 24, 28, 35, 37, 38, 39, 57, 60, 77 in the GDPR. In PESTOS you can refer section 3.2 in order to have a listed version of them per GDPR actor.

For example, a supervisory authority has a legal power for enforcing the GDPR. It is represented as a GDPR capability of the GDPR actor, which is Supervisory Authority here, and its graphical representation is similar with the one for *Capability* from Tropos except it has a (P) label additionally.

Table 1. GDPR Goals Associated GDPR Articles

GDPR Goals	GDPR Articles
Security (Confidentiality, Availability, Integrity)	24, 25, 32
Organizational & Legal goals (Lawfulness, Privacy Risk Assessment, Privacy Policy, Privacy Management Trainings)	5, 6, 7, 8, 9, 10 24, 25, 27, 28, 29 31, 35, 36, 37, 38 39, 40, 41, 42, 43 44, 45, 46, 47, 48 49
Right of access	15, 24, 25
Right to rectification	16, 24, 25
Right to data portability	20, 24, 25
Right to erasure	17, 24, 25
Right to restriction of processing (including Right to object)	9, 10, 18, 21, 22 24, 25
Data Minimization	89
Transparency	12, 13, 14, 19, 24 25, 30, 33, 34, 49

Privacy Reference Diagram: It is a diagram that built by actualization of a modeling activity involves the identification of privacy needs of the system-to-be (GDPR compliant system) by introducing GDPR irregularities of the system (privacy related problems, regulatory problems) and possible technical privacy measures that are applicable to those irregularities. Thus, the privacy reference diagram represents the relationship between privacy features, flaws, infringements and privacy mechanisms.

The infringement meta-model (Fig. 19) shows how situating an event that may cause infringement could take place within privacy reference model of PESTOS. Infringement occurs whenever a HBC(Honest-but-curious adversary) finds out a flaw in the information system which processes personal data or in a task which is supposed to executed in order to reach a strategic goal. On the other hand, infringement may emerge also exclusively due to unlawful data processing.

Flaw: Flaw represents a condition, a behavior, a mechanism or a design decision which brings imperfection to an actor or a task. Such imperfections may potentiality cause infringements in respect of the GDPR.

Infringement: Infringement describes a situation where a potential data breach or an inconsistency in respect to GDPR may be occurred where an HBC as an agent takes action. An infringement can be modeled in two different way: 1. as a means-end relationship between a capability, which an HBC has, and GDPR resource 2. as an independent concrete syntax (which looks like a paraboloid figure), in cases where an

unlawful processing takes place. The both representations can be seen on the below in Figure 20.

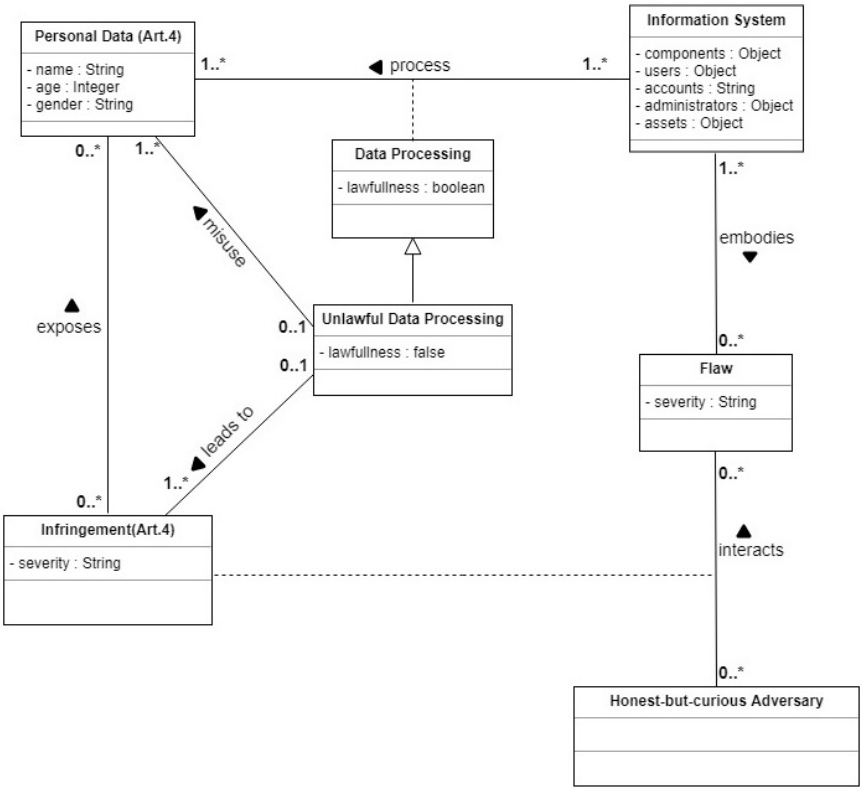

Figure 19. Infringement Meta-model

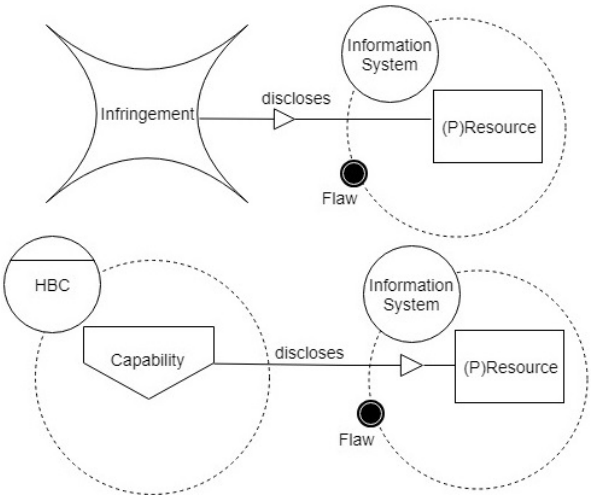

Figure 20. Graphical representations of infringement

Privacy Features: Privacy Features represent PbD strategies (e.g minimize, aggregate) and characteristics associated to privacy patterns(e.g anonymity, pseudonymity), such as the most appropriate ones between them (here, we mean picking appropriate ones from data oriented strategies, all process oriented strategies should be take place for each systems) for satisfying privacy requirements of the system should be considered through the development life cycle of the system-to-be from the beginning of the design phase. The following figure shows how those strategies can be imposed by different GDPR actors of the system .

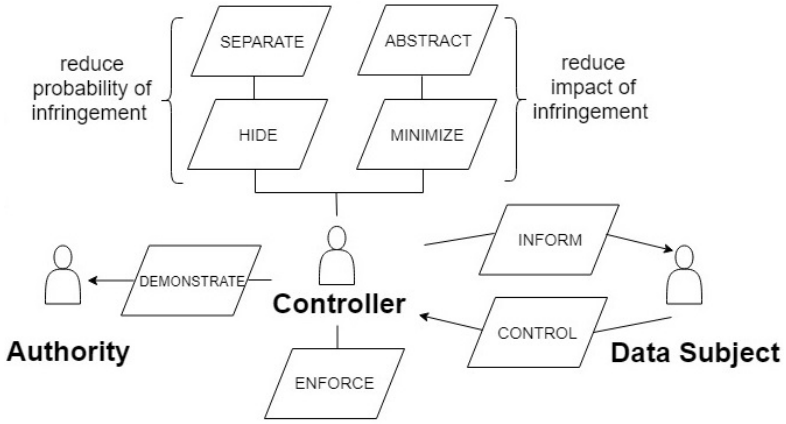

Figure 21. PbD strategies by GDPR actors, adapted from [CHH16]

Privacy Objectives: Privacy Objectives represent PbD principles that Privacy Features take inspiration from. It is an abstract concept so PESTOS doesn't introduce any notation for a privacy objective. PbD principles are mentioned in section 2.2

Privacy Mechanism: It represents *Privacy Enhancing Technologies(PETs)* to contribute actualizing GDPR tasks that are supposed to satisfy GDPR goals, eventually. Privacy Mechanism should enable the application of system's *Privacy Features* that satisfy privacy objectives.

5.3 Summary

The purpose of this chapter was to introduce extended Tropos methodology for building the GDPR-complaint software systems. PESTOS model aims to offer compliance solutions with the scope to address privacy and GDPR during system development. To fulfil this aim, this chapter introduced concrete syntax for GDPR concepts and explained the semantics of the language.

6 Example of PESTOS Model

This chapter introduces an application of PESTOS methodology based on a real-world case. It aims to introduce also application method of PESTOS in general. Finally, the chapter aims to to evaluate the proposed privacy-oriented & GDPR-oriented approach and better understand its advantages.

6.1 Case Presentation

In order to apply PESTOS and analyze the models, that are obtained after the implementation, an Identity and Access Management (IAM) platform of an organization is decided as a case study. Due to organizational secrecy, we prefer to not mention here the names of the organization and the platform. The vision of IAM is to enable effective administration of identities and access rights in the organization. In this public version of the thesis, there is no any further information about IAM platform.

6.2 Application Method

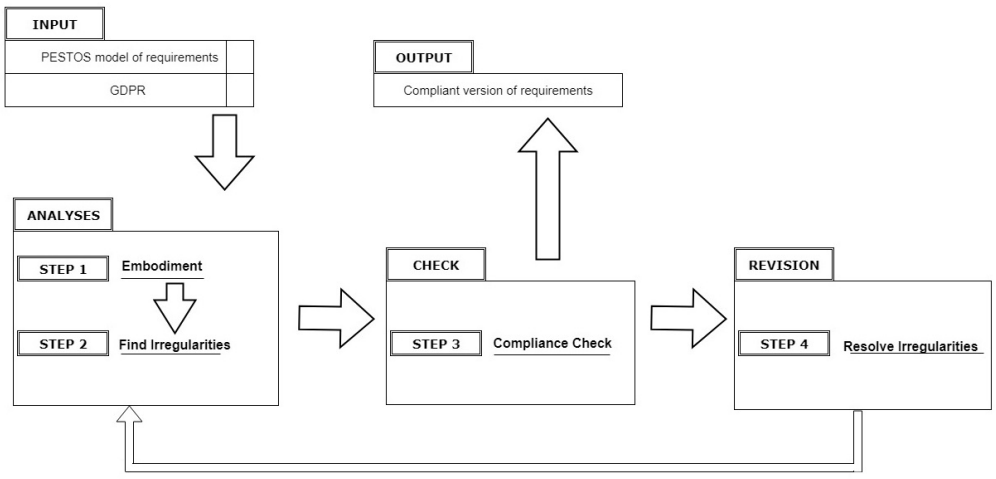

Figure 22. 3 Phases 4 Steps, adapted from [ISM+13]

The approach for modeling application of PESTOS within one iteration requires to follow 3 steps :

1. Embodiment
2. Finding Irregularities
3. Compliance Check

based on building two types of model :

1. Privacy Reference Model
2. GDPR Compliance Model

After compliance check took place the model is going to be revised by engineers, if they think there are still unsatisfied requirements with potential unsolved irregularities. Thus, revision phase may appear based on a decision. The revision phase brings one more step and an iteration of first two phases :

4. Resolving Irregularities

Revision phase can be repeated over and over in order to re-mediate the compliance model. In our implementation case, we won't go to that much details rather only first two phases, analyses and check, will be covered.

Embodiment: Modeling the GDPR requirements of the system and system actors by using the concrete syntax of PESTOS within GDPR compliance model. Modeling privacy threats and privacy requirements of that system within privacy reference model.

Finding Irregularities: Finding irregularities of the system as-is with respect to GDPR and PbD strategies based on *privacy reference model*.

Compliance Check: The important goal of this step is to evaluate whether the given GDPR compliance model satisfies the compliance properties. This step takes as input the *privacy reference model* of requirements where all the irregularities and possible measures have been highlighted.

Resolving Irregularities: Addressing all irregularities identified in the model in order to revise it by engineers proposing new elements, new operationalizations and changes. A discussion will evaluate whether the changes made are sufficient to consider the irregularity solved. The model is revised until the discussion rejects the validity of the irregularity. At the end of this step, each irregularity has been considered and then resolved through revisions of the model.

6.3 Privacy Reference Model

By privacy reference model, we constructed an infringement scenario which displays HBCs and their capabilities which might be used in order to disclose sensitive personal data.

In privacy reference model, we can also introduce privacy constraints, privacy features and privacy mechanisms in order to design a system which has preventive measures in order to prevent potential infringements, that may caused by different internal actors. Some privacy constraints should be considered whenever actors want to reach their essential goals. This public version of the thesis doesn't include any further information about Privacy Reference Model.

6.4 GDPR Compliance Model

GDPR compliance model is introduced with a goal modeling diagram. In the diagram, different GDPR actors realizes their goals and executes their tasks by as much as their capabilities allow them in order to perform a Data Subject's Right. So this diagram is a part of the GDPR compliance model. This public version of the thesis doesn't include any further information about GDPR Compliance Model.

6.5 Validation of PESTOS Model

As mentioned before in the *Secure Tropos Framework* chapter, this study aims to extend and customize an agent oriented software engineering methodology to enable it to model privacy and compliance issues within the scope of the GDPR, rather than building one from scratch.

By validation of this work, we mean validation of usability and correctness of the models introduced in this chapter. For this purpose we prepare a validation strategy that aims to collect unbiased feedbacks from identity and security experts in the organization. The feedbacks were collected after the presentation of the PESTOS framework, the case study and the PESTOS models. The validation strategy is defined and conducted of survey questionnaire and interview, based on open-ended questions. Due to the hardship in finding a large number of qualified people, who can evaluate such a specific work, our sample size is limited with 21 people.

Survey Question	Semantic Differential Scale	Mean	Median	Variance
1. How much do you think you know about GDPR?	(Not at all : 1) - (10 : Totally)	7,09	7	2,79
2. How much do you think you know about privacy?	(Not at all : 1) - (10 : Totally)	8,00	8	1,5
3. How easy to understand the syntax of the PESTOS?	(Not at all : 1) - (10 : Very easy)	6,42	8	6,65
4. How efficient to describe the GDPR actors and the relationships between them by using PESTOS ?	(Not at all : 1) - (10 : Very efficient)	7,28	8	3,91
5. Do you think how efficient would be to model a GDPR compliant system by using PESTOS?	(Not at all : 1) - (10 : Very efficient)	7,24	7	2,89
6. How easy to understand PETs and PbD strategies as technical measure for preventing infringements?	(Not at all : 1) - (10 : Very easy)	6,80	6	2,86
7. How suitable do you think to implement modeling techniques of PESTOS on any potential system in your organization ?	(Not at all : 1) - (10 : Very suitable)	6,86	8	6,32
8. How likely you would use this language to model a GDPR-compliant system?	(Very unlikely : 1) - (10 : Very likely)	6,90	8	4,19
9. How much do you think it is extensible and flexible as a privacy modeling language?	(Not at all : 1) - (10 : Very extensible)	7,09	8	2,69
10. How easy do you think to gain expertise on PESTOS?	(Not at all : 1) - (10 : Very easy)	7,19	8	4,06
11. How easy to understand application method of PESTOS ?	(Not at all : 1) - (10 : Very easy)	7,33	8	5,03
12. Is the modeling methodology restricted to a particular implementation choice?	(Very restricted : 1) - (10 : Very flexible)	6,76	6	3,99
13. How likely do you think that model could have some shortages?	(Very likely : 1) - (10 : Very unlikely)	7,43	8	1,85

Figure 23. Results of the PESTOS Evaluation Survey

Within survey, those questions shown in Figure 23 were asked to be marked anonymously between 1 to 10. We use semantic differential scale as the response scale of the

PESTOS evaluation survey. In Figure 23 mean(average score), median(middle score) and variance for each question are written on the right side of the table. According to this result, after a careful evaluation of PESTOS model by identity & security experts, who allege, that they have a solid knowledge about GDPR and privacy concept, it may safely be said that PESTOS is a modeling language which has a moderately easy syntax and flexibility, whereas it is fairly efficient to model GDPR compliant systems. It is fairly extensible as a privacy modeling language and considered as it may have neglectable shortages or some edge cases, due to being a new modeling language.

6.6 Summary

The main aim of this chapter was to illustrate how the proposed modeling language can be applied in the redesign of a real life information system. To fulfill this aim, this chapter described how the extended and customized Tropos methodology was employed for Identity and Access Management (IAM) platform of an financial institution in order to identify GDPR and privacy requirements of the system. In addition, this chapter described the validation approach for the proposed modeling language and evaluation feedbacks from identity & security experts.

7 Conclusion

This thesis gives an outline of privacy and privacy components by touching on privacy by design strategies and privacy enhancing technologies. The study commentates GDPR and approaches GDPR compliance as a modeling activity that could be realized by identifying system requirements related with privacy and GDPR in an early phase of design stage. In this direction, the study covers Secure Tropos methodology and enhance it by introducing GDPR and privacy concepts with new concrete syntax for each element. In order to explain Secure Tropos methodology, the thesis briefly touches on i* framework and Tropos methodology, beforehand. Moreover, the study introduces a new modeling language, PESTOS, with its meta-model, semantics and concrete syntax. Finally, it displays an application of the PESTOS modeling language in a real-world system and validation of the framework based on evaluation of the real-world case by identity & security experts.

7.1 Limitations

There are certain limitations about this dissertation that could be mentioned here. First of all, there is a paucity of literature about requirement engineering which addressing legal regulations so far and this case becomes more obvious if we talk especially about the limited literature on GDPR from requirement analysis and modeling perspective.

By its nature a model has a certain level of abstraction which reflects the understanding of the modeler about system requirements. By PESTOS, we enable modelers to introduce technical measures, by "Privacy Feature" and "Privacy Mechanism" concepts, against to potential infringements. However, compliance with GDPR doesn't stand on only technical measures but also organizational measures. Organizational measures aren't taken into the scope of this work due to there could be belong to various categories and the fact that how they should be implemented may differ by the organizational culture and the structure of an organization.

Additionally, another limitation may emerges by employing PETs as technical measures. As we mentioned before in the second chapter, PETs are not widely adopted yet due to three different obstacles; lack of availability of PETs, lack of user friendliness, subsistence of drawbacks in infrastructure deployment.

7.2 Answers to Research Questiones

In the introduction chapter we state our main research question *"How to extend Tropos methodology for building the GDPR compliant software systems ?"*. We disintegrate this question into three minor research questions.

RQ1 - How to make software systems GDPR-compliant? : GDPR obliges to implement appropriate technical and organizational measures via considering the concept of

data protection by design. Thus, PESTOS aims to help system designers and developers in order to plan, implement, maintain and demonstrate GDPR compliance through ensuring compliant processing of personal data by pointing proper PETs. As well as it proposes to assign PETs to GDPR resources according to PbD strategies which reflects the concept of data protection by design.
RQ2 - How can Tropos methodology implement GDPR-compliance principles? : Tropos methodology provides enough extensibility to represent GDPR-related concepts by introducing new modeling constructs such as *Privacy Mechanism*, *Privacy Feature*, *GDPR Resource*, *GDPR Actor*, *GDPR Goal* and *Infringement*. Through those entities Tropos methodology is used to build a number of models within goal and actor orientation.
RQ3 - How to apply enhanced Tropos methodology to the GDPR compliance? : Application method consists of three phases and four steps as illustrated in Section 6.2. Those steps are *Embodiment, Finding Irregularities, Compliance Check* and *Resolving Irregularities*. PESTOS model can be improved with additional iterations of the application method until the complaint version of requirements are considered as fully satisfied.

7.3 Conclusion

We proposed PESTOS as a main contribution, a privacy modeling language for GDPR compliance. To the best of our knowledge, PESTOS is the first framework, enhanced based on Secure Tropos methodology in order to elicit GDPR requirements by various methods.

Since all new methods require a form of validation, our second contribution consists of the through evaluation of the PESTOS methodology, in terms of correctness and usability, based on its applications.

7.4 Future Work

Privacy modeling is still considered as a novel area and similar studies on the GDPR haven't reached a meaningful amount of numbers yet. So this study is a pioneer in that field and it's our very firmly belief that it would create a basis for following works.

Once, the number of similar studies hit a sufficient level of existing works then different validations methods that may perform to measure performance of the PESTOS by comparing with the other studies who focus on the requirement engineering for the GDPR compliance. For now, it is left for the future due to lack of time and resource. Also a modeling tool could be developed for PESTOS framework which may increase the ease of application of PESTOS. Thus, it would facilitate the workload of system engineers caused by modeling with PESTOS.

References

[ABB+04] William Aiello, Steven M. Bellovin, Matt Blaze, Ran Canetti, John Ioannidis, Angelos D. Keromytis, and Omer Reingold. Just fast keying: Key agreement in a hostile internet. *ACM Transactions on Information and System Security (TISSEC)*, 7(2):242–273, Jan 2004.

[ABC+07] Michel Abdalla, Mihir Bellare, Dario Catalano, Eike Kiltz, Tadayoshi Kohno, Tanja Lange, John Malone-Lee, Gregory Neven, Pascal Paillier, Haixia Shi, and et al. Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous IBE, and Extensions. *Journal of Cryptology*, 21(3):350–391, 2007.

[ACK+10] Claudio A. Ardagna, Jan Camenisch, Markulf Kohlweiss, Ronald Leenes, Gregory Neven, Bart Priem, Pierangela Samarati, Dieter Sommer, and Mario Verdicchio. Exploiting cryptography for privacy-enhanced access control: A result of the PRIME Project. *Journal of Computer Security*, 18(1):123–160, Jan 2010.

[AF04] Martin Abadia and Cédric Fournet. Private authentication. *Theoretical Computer Science*, page 690–690, September 2004.

[And06] Anne H. Anderson. A Comparison of Two Privacy Policy Languages: EPAL and XACML. In *Proceedings of the 3rd ACM Workshop on Secure Web Services*, SWS '06, pages 53–60, New York, NY, USA, 2006. ACM.

[AP98] R.J. Anderson and F.A.P. Petitcolas. On the limits of steganography. *IEEE Journal on Selected Areas in Communications*, 16(4):474–481, 1998.

[Aus] Australian Human Rights Commission. What is the Universal Declaration of Human Rights ? "https://www.humanrights.gov.au/publications/what-universal-declaration-human-rights". Accessed: 2017-08-01.

[Bab17] Chris Babbel. Dark Reading, Endpoint, The High Costs of GDPR Compliance . "https://www.darkreading.com/endpoint/the-high-costs-of-gdpr-compliance/a/d-id/1329263", 2017. Accessed: 2018-04-28.

[Bar06] Susan B. Barnes. A privacy paradox: Social networking in the united states. *First Monday*, 11(9), 2006.

[BC] Stefan Brands and David Chaum. Distance-Bounding Protocols. *Advances in Cryptology — EUROCRYPT '93 Lecture Notes in Computer Science*, page 344–359.

[BeV95] Lillian R. BeVier. Information About Individuals in the Hands of Government: Some Reflections on Mechanisms for Privacy Protection. *William & Mary Bill of Rights Journal*, 4(2):455–506, 1995.

[BFK01] Oliver Berthold, Hannes Federrath, and Stefan Köpsell. *Web MIXes: A System for Anonymous and Unobservable Internet Access*, page 115–129. 2001.

[BGB04] Nikita Borisov, Ian Goldberg, and Eric Brewer. Off-the-record communication, or, why not to use PGP. *Proceedings of the 2004 ACM workshop on Privacy in the electronic society - WPES 04*, page 77–84, 2004.

[BGS01] Adam Back, Ian Goldberg, and Adam Shostack. Freedom systems 2.1 security issues and analysis. *White paper, Zero Knowledge Systems, Inc.*, 2001.

[Bor11] John J. Borking. *Why Adopting Privacy Enhancing Technologies (PETs) Takes so Much Time*, page 309–341. Springer, 2011.

[BPG+01] Paolo Bresciani, Anna Perini, Paolo Giorgini, Fausto Giunchiglia, and John Mylopoulos. A knowledge level software engineering methodology for agent oriented programming. *Proceedings of the fifth international conference on Autonomous agents - AGENTS 01*, 2001.

[BPG+04] Paolo Bresciani, Anna Perini, Paolo Giorgini, Fausto Giunchiglia, and John Mylopoulos. Tropos: An Agent-Oriented Software Development Methodology. *Autonomous Agents and Multi-Agent Systems*, 8(3):203–236, 2004.

[Cac98] Christian Cachin. On the foundations of oblivious transfer. In Kaisa Nyberg, editor, *Advances in Cryptology — EUROCRYPT'98*, pages 361–374, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg.

[Cav11] Ann Cavoukian. Privacy by Design: Strong Privacy Protection - Now, and Well into the Future. In *A Report on the State of PbD to the 33rd International Conference of Data Protection and Privacy Commissioners*, page 31, 2011.

[CD98] Jan Camenisch and Ivan B. Damgård. Verifiable Encryption and Applications to Group Signatures and Signature Sharing. *BRICS Report Series*, 5(32), Feb 1998.

[CF08] Barbara Carminati and Elena Ferrari. Privacy-Aware Collaborative Access Control in Web-Based Social Networks. *Lecture Notes in Computer Science Data and Applications Security XXII*, page 81–96, 2008.

[Cha81] David Chaum. Untraceable Electronic Mail, Return Addresses and Digital Pseudonyms. *Communications of the ACM*, 24(2):84–88, 1981.

[Cha85] David Chaum. Security without identification: Transaction systems to make big brother obsolete. *Communications of the ACM*, 28(10):1030–1044, 1985.

[Cha88] David Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1(1):65–75, 1988.

[CHH16] Michael Colesky, Jaap-Henk Hoepman, and Christiaan Hillen. A Critical Analysis of Privacy Design Strategies. *2016 IEEE Security and Privacy Workshops (SPW)*, 2016.

[CKGS98] Benny Chor, Eyal Kushilevitz, Oded Goldreich, and Madhu Sudan. Private information retrieval. *Journal of the ACM*, 45(6):965–981, Jan 1998.

[CL04] Jan Camenisch and Anna Lysyanskaya. Signature Schemes and Anonymous Credentials from Bilinear Maps. *Advances in Cryptology – CRYPTO 2004 Lecture Notes in Computer Science*, page 56–72, 2004.

[CP11] Ann Cavoukian and Claudia Popa. Privacy by ReDesign : A Practical Framework for Implementation. "`http://securityandprivacy.ca/download/new/Informatica-PbRD-framework.pdf`", November 2011. Accessed: 2017-09-26.

[CPHH02] Sebastian Clauß, Andreas Pfitzmann, Marit Hansen, and Els Van Herreweghen. Privacy-enhancing identity management. *The IPTS Report 67*, (2):8–16, Sep 2002.

[CW10] Shan Chen and Mary-Anne Williams. Privacy: An Ontological Problem. In *Pacific Asia Conference on Information Systems, PACIS 2010, Taipei, Taiwan, 9-12 July 2010*, page 134. AISeL, 2010.

[DCIO16] Angelo De Caro, Vincenzo Iovino, and Adam O'Neill. Deniable functional encryption. In Chen-Mou Cheng, Kai-Min Chung, Giuseppe Persiano, and Bo-Yin Yang, editors, *Public-Key Cryptography – PKC 2016*, pages 196–222, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg.

[DDH+15] George Danezis, Josep Domingo-Ferrer, Marit Hansen, Jaap-Henk Hoepman, Daniel Le Métayer, Rodica Tirtea, and Stefan Schiffner. Privacy and Data Protection by Design - from policy to engineering. *CoRR*, abs/1501.03726, 2015.

[DFH16] Fabiano Dalpiaz, Xavier Franch, and Jennifer Horkoff. iStar 2.0 Language Guide. *CoRR*, abs/1605.07767, 2016.

[DMS04] Roger Dingledine, Nick Mathewson, and Paul Syverson. Tor: The Second-Generation Onion Router. In *13th USENIX Security Symposium*, Jan 2004.

[EuG17] GDPR Portal. "http://www.eugdpr.org/", 2017. Accessed: 2017-11-19.

[EuP] Information society for all, European Parliament Debates (Thursday, 16 March 2000, Strasbourg). "http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+CRE+20000316+ITEM-002+DOC+XML+V0//EN". Accessed: 2017-06-28.

[FG07] Caroline Fontaine and Fabien Galand. A Survey of Homomorphic Encryption for Nonspecialists. *EURASIP J. Inf. Secur.*, 2007:15:1–15:15, January 2007.

[FLM+04] Ariel Fuxman, Lin Liu, John Mylopoulos, Marco Pistore, Marco Roveri, and Paolo Traverso. Specifying and analyzing early requirements in Tropos. *Requirements Engineering*, 9(2):132–150, Jan 2004.

[fNE16] The European Union Agency for Network and Information Security (ENISA). Privacy Enhancing Technologies: Evolution and State of the Art A Community Approach to PETs Maturity Assessment. "https://www.enisa.europa.eu/publications/pets-evolution-and-state-of-the-art/at_download/fullReport", December 2016. Accessed: 2017-09-28.

[GCW18] May Fen Gan, Hui Na Chua, and Siew Fan Wong. *Personal Data Protection Act Enforcement with PETs Adoption: An Exploratory Study on Employees' Working Process Change*, pages 193–202. Springer Singapore, Singapore, 2018.

[GDP] General Data Protection Regulation (GDPR). "https://gdpr-info.eu/". Accessed: 2017-10-02.

[GGM16] Mohamad Gharib, Paolo Giorgini, and John Mylopoulos. Ontologies for Privacy Requirements Engineering: A Systematic Literature Review. *CoRR*, abs/1611.10097, 2016.

[GH09] Editors Of Nolo Gerald Hill, Kathleen Hill. *Nolo's Plain-English Law Dictionary*. NOLO, 1 edition, 2009.

[Gio] Paolo Giorgini. Tropos: basics. "http://www.troposproject.eu/files/8-Tropos-Basics.pdf".

[GMPT01] Christos K. Georgiadis, Ioannis Mavridis, George Pangalos, and Roshan K. Thomas. Flexible team-based access control using contexts. *Proceedings of the Sixth ACM symposium on Access control models and technologies - SACMAT 01*, page 21–27, 2001.

[HBC+04] Marit Hansen, Peter Berlich, Jan Camenisch, Sebastian Clauß, Andreas Pfitzmann, and Michael Waidner. Privacy-enhancing identity management. *Information Security Technical Report*, 9(1):35–44, 2004.

[Hen] Henry George Liddell, Robert Scott, An Intermediate Greek-English Lexicon. "http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.04.0058%3Aalphabetic%2Bletter%3D%2At%3Aentry%2Bgroup%3D23%3Aentry%3Dtro%2Fpos".

[Hoe14] Jaap-Henk Hoepman. Privacy Design Strategies. *ICT Systems Security and Privacy Protection IFIP Advances in Information and Communication Technology*, page 446–459, 2014.

[HZNF15] Johannes Heurix, Peter Zimmermann, Thomas Neubauer, and Stefan Fenz. A taxonomy for privacy enhancing technologies. *Computers & Security*, 53:1–17, 2015.

[IMJ10] Shareeful Islam, Haralambos Mouratidis, and Jan Jürjens. A framework to support alignment of secure software engineering with legal regulations. *Software & Systems Modeling*, 10(3):369–394, Sep 2010.

[ISM+13] Silvia Ingolfo, Alberto Siena, John Mylopoulos, Angelo Susi, and Anna Perini. Arguing regulatory compliance of software requirements. *Data & Knowledge Engineering*, 87:279–296, 2013.

[KM01] Darko Kirovski and Henrique Malvar. Robust Covert Communication over a Public Audio Channel Using Spread Spectrum. *Information Hiding Lecture Notes in Computer Science*, 2137:354–368, 2001.

[KW14] Inga Kroener and David Wright. A Strategy for Operationalizing Privacy by Design. *The Information Society*, 30(5):355–365, October 2014.

[LBW08] Heather Richter Lipford, Andrew Besmer, and Jason Watson. Understanding Privacy Settings in Facebook with an Audience View. In *Proceedings of the 1st Conference on Usability, Psychology, and Security*, UPSEC'08, pages 2:1–2:8, Berkeley, CA, USA, 2008. USENIX Association.

[LHDL04] Scott Lederer, Jason I. Hong, Anind K. Dey, and James A. Landay. Personal privacy through understanding and action: five pitfalls for designers. *Personal and Ubiquitous Computing*, 8(6):440–454, 2004.

[LLV07] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. *2007 IEEE 23rd International Conference on Data Engineering*, 2007.

[LYM03] Lin Liu, Eric S. K. Yu, and John Mylopoulos. Security and Privacy Requirements Analysis within a Social Setting. In *11th IEEE International Conference on Requirements Engineering(2003), 8-12 September 2003, Monterey Bay, CA, USA.*, pages 151–161, 2003.

[Mat17] Raimundas Matulevičius. *Fundamentals of secure system modelling*. Springer Berlin Heidelberg, 2017.

[MCNM03] Ira S. Moskowitz, Daniel P. Crepeau, Richard E. Newman, and Allen R. Miller. Covert Channels and Anonymizing Networks. *In Workshop on Privacy in the Electronic Society*, page 79–88, 2003.

[MCV01] A. J. Menezes, Van Oorschot Paul C., and Scott A. Vanstone. *Handbook of applied cryptography*. CRC Press, 2001.

[MG07] Haralambos Mouratidis and Paolo Giorgini. Secure Tropos: A Security-Oriented Extension of the Tropos methodology. "`http://roar.uel.ac.uk/415/`", Apr 2007.

[MGKV06] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam. L-diversity: privacy beyond k-anonymity. *22nd International Conference on Data Engineering (ICDE06)*, page 24, 2006.

[MGM03] Haralambos Mouratidis, Paolo Giorgini, and Gordon A. Manson. Integrating Security and Systems Engineering: Towards the Modelling of Secure Information Systems. In *Advanced Information Systems Engineering, 15th International Conference, CAiSE 2003, Klagenfurt, Austria, June 16-18, 2003, Proceedings*, pages 63–78, 2003.

[Moo97] James H. Moor. Towards a Theory of Privacy in the Information Age. *SIGCAS Comput. Soc.*, 27(3):27–32, September 1997.

[Mou04] Haralambos Mouratidis. A security oriented approach in the development of multiagent systems : applied to the management of the health and social care needs of older people in England. "`http://etheses.whiterose.ac.uk/14864/`", Jan 2004. Accessed: 2018-01-27.

[Myl] John Mylopoulos. The Tropos Project: A Research Overview - università di trento. "`http://www.science.unitn.it/tropos/tropos-workshop/slides/mylopoulos.pdf`".

[Nao02] Moni Naor. Deniable Ring Authentication. *Advances in Cryptology — CRYPTO 2002 Lecture Notes in Computer Science*, 2442:481–498, 2002.

[NN01] Moni Naor and Kobbi Nissim. Communication preserving protocols for secure function evaluation. *Proceedings of the thirty-third annual ACM symposium on Theory of computing - STOC 01*, page 590–599, 2001.

[OAL12] Muyiwa Olurin, Carlisle Adams, and Luigi Logrippo. Platform for privacy preferences (P3P): Current status and future directions. *2012 Tenth Annual International Conference on Privacy, Security and Trust*, page 217–220, 2012.

[oas17] OASIS eXtensible Access Control Markup Language (XACML) Version 3.0. "http://docs.oasis-open.org/xacml/3.0/errata01/os/xacml-3.0-core-spec-errata01-os-complete.pdf", Jul 2017.

[oIa] UC Berkeley School of Information. Mixminion: A Type III Anonymous Remailer. "https://www.mixminion.net/". Accessed: 2018-02-24.

[oIb] UC Berkeley School of Information. Privacy Patterns. "https://privacypatterns.org/patterns/". Accessed: 2018-02-20.

[OME00] OME Power User Tutorial. http://www.cs.toronto.edu/km/ome/docs/poweruser/poweruser.html, 2000. Accessed: 2017-11-21.

[OS07] Rafail Ostrovsky and William E. Skeith. Private Searching on Streaming Data. *Journal of Cryptology*, 20(4):397–430, 2007.

[Pet11] CIS PET wiki. http://cyberlaw.stanford.edu/wiki/index.php/PET, November 2011. Accessed: 2017-11-19.

[Phi04] David J. Phillipes. Privacy policy and PETs: The influence of policy regimes on the development and social implications of privacy enhancing technologies. *new media & society*, 6:691–706, 2004.

[PIM12] Michalis Pavlidis, Shareeful Islam, and Haralambos Mouratidis. A CASE Tool to Support Automated Modelling and Analysis of Security Requirements, Based on Secure Tropos. *Lecture Notes in Business Information Processing IS Olympics: Information Systems in a Diverse World*, page 95–109, 2012.

[Pin02] Benny Pinkas. Cryptographic Techniques for Privacy-preserving Data Mining. *SIGKDD Explor. Newsl.*, 4(2):12–19, December 2002.

[PK09] Sameer Patil and Alfred Kobsa. Privacy Considerations in Awareness Systems: Designing with Privacy in Mind. *Human-Computer Interaction Series Awareness Systems*, page 187–206, 2009.

[PMB17] Pille Pullonen, Raimundas Matulevicius, and Dan Bogdanov. PE-BPMN: Privacy-Enhanced Business Process Model and Notation. In Josep Carmona, Gregor Engels, and Akhil Kumar, editors, *Business Process Management - 15th International Conference, BPM 2017, Barcelona, Spain, September 10-15, 2017, Proceedings*, volume 10445 of *Lecture Notes in Computer Science*, pages 40–56. Springer, 2017.

[PPW91] Andreas Pfitzmann, Birgit Pfitzmann, and Michael Waidner. ISDN-Mixes: Untraceable Communication with Very Small Bandwidth Overhead. *Kommunikation in verteilten Systemen Informatik-Fachberichte*, page 451–463, 1991.

[PW] Andreas Pfitzmann and Michael Waidner. Networks Without User Observability — Design Options. *Advances in Cryptology — EUROCRYPT' 85 Lecture Notes in Computer Science*, page 245–253.

[Rab81] Michael O. Rabin. How to Exchange Secrets with Oblivious Transfer. "https://eprint.iacr.org/2005/187.pdf", 1981. Accessed: 2018-02-25.

[RB11] Martin Rost and Kirsten Bock. Privacy By Design und die Neuen Schutzziele - Grundsätze, Ziele und Anforderungen. *Datenschutz und Datensicherheit*, 35(1):30–35, 2011.

[RDB+10] Alfredo Rial, Mina Deng, Tiziano Bianchi, Alessandro Piva, and Bart Preneel. A Provably Secure Anonymous Buyer–Seller Watermarking Protocol. *IEEE Transactions on Information Forensics and Security*, 5(4):920–931, 2010.

[Ref11] Information Science Reference. *Virtual Communities: Concepts, Methodologies, Tools and Applications*. IGI Global, 2011.

[RR98] Michael K. Reiter and Aviel D. Rubin. Crowds: anonymity for Web transactions. *ACM Transactions on Information and System Security*, 1(1):66–92, Jan 1998.

[SFEa] Intrinsic vs. Extrinsic Value. "https://plato.stanford.edu/entries/value-intrinsic-extrinsic/". Accessed: 2017-08-20.

[SFEb] Privacy and Information Technology. "https://https://plato.stanford.edu/entries/it-privacy/". Accessed: 2018-04-28.

[SGR] P.F. Syverson, D.M. Goldschlag, and M.G. Reed. Anonymous connections and onion routing. *Proceedings. 1997 IEEE Symposium on Security and Privacy (Cat. No.97CB36097)*.

[SJBK03] Vanja Seničar, Borka Jerman-Blažič, and Tomaž Klobučar. Privacy-enhancing Technologies: Approaches and Development. *Comput. Stand. Interfaces*, 25(2):147–158, May 2003.

[Sol06] Daniel J. Solove. A Taxonomy of Privacy. *University of Pennsylvania Law Review*, 154(3):477–560, January 2006.

[SP11] Y Shen and S Pearson. Privacy enhancing technologies: A review. *HP Laboratories Technical Report*, pages 1–30, 01 2011.

[STP09] Koen Simoens, Pim Tuyls, and Bart Preneel. Privacy Weaknesses in Biometric Sketches. *Proceedings of the 2009 30th IEEE Symposium on Security and Privacy*, page 188–203, 2009.

[Swe02a] Latanya Sweeney. Achieving k-Anonymity privacy protection using generalization and suppression. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):571–588, 2002.

[Swe02b] Latanya Sweeney. k-Anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570, 2002.

[TCF12] Juan Pablo Timpanaro, Isabelle Chrisment, and Olivier Festor. I2p's usage characterization. In Antonio Pescapè, Luca Salgarelli, and Xenofontas Dimitropoulos, editors, *Traffic Monitoring and Analysis*, pages 48–51, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[UDH48] Universal declaration of human rights. "http://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf", December 1948. Art. 12. Accessed: 2017-07-31.

[Uni14] United Nations, Office of the United Nations High Commissioner for Human Rights. The Right to Privacy in the Digital Age. "http://www.ohchr.org/EN/Issues/DigitalAge/Pages/DigitalAgeIndex.aspx", January 2014. Accessed: 2017-08-31.

[VB17] Paul Voigt and Axel von dem Bussche. *The EU General Data Protection Regulation (GDPR): a practical guide*. Springer, 2017.

[VBF+04] Vassilios S. Verykios, Elisa Bertino, Igor Nai Fovino, Loredana Parasiliti Provenza, Yucel Saygin, and Yannis Theodoridis. State-of-the-art in Privacy Preserving Data Mining. *SIGMOD Rec.*, 33(1):50–57, March 2004.

[Wik] English Wikipedia. General Data Protection Regulation. "https://en.wikipedia.org/wiki/General_Data_Protection_Regulation". Accessed: 2017-11-22.

[WP] Michael Waidner and Birgit Pfitzmann. The Dining Cryptographers in the Disco: Unconditional Sender and Recipient Untraceability with Computationally Secure Serviceability. *Lecture Notes in Computer Science Advances in Cryptology — EUROCRYPT '89*, page 690–690.

[Wuy15] Kim Wuyts. Privacy Threats in Software Architectures. *KU Leuven, ARENBERG DOCTORAL SCHOOL, Faculty of Engineering Science*, 2015.

[Yao82] Andrew C. Yao. Protocols for Secure Computations. In *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science*, SFCS '82, pages 160–164, Washington, DC, USA, 1982. IEEE Computer Society.

[Yu97] Eric S. K. Yu. Towards Modeling and Reasoning Support for Early-Phase Requirements Engineering. In *3rd IEEE International Symposium on Requirements Engineering (RE'97), January 5-8, 1997, Annapolis, MD, USA*, pages 226–235, 1997.

[Yu09] Eric S. K. Yu. Social Modeling and i*. In *Conceptual Modeling: Foundations and Applications - Essays in Honor of John Mylopoulos*, pages 99–121, 2009.

Appendix

I. List of Abbreviations

AD Active Directory

AOP Agent Oriented Programming

BCR Binding Corporate Rules

BEH Behörighet för Sparbankerna (Authority for Savings Banks)

CVF Core Value Framework

DPO Data Protection Officer

FIP Fair Information Practices

GDPR General Data Protection Regulation

GUID Globally Unique Identifier

HBC Honest-but-curious Adversary

HR Human Resources

IAM Identity and Access Management

ICT Information and Communication Technologies

ICO Information Commissioner's Office

PbD Privacy by Design

PbRD Privacy by ReDesign

PESTOS Privacy Enhanced Secure Tropos

PETs Privacy Enhancing Technologies

PIA Privacy Impact Analysis

pID Personal Identification

QM Queue Manager

SD Strategic Dependency

SR Strategic Rationale

UDHR Universal Declaration of Human Rights

UNGA United Nations General Assembly

II. Privacy as a Human Right

Accompanied by the Universal Declaration of Human Rights (UDHR), which is adopted by United Nations General Assembly (UNGA) on 10 December 1948 at the Palais de Chaillot in Paris, France, privacy gained important ground for itself in that international document that states basic rights and fundamental freedoms to which all human beings are entitled. Here, It has to be admitted that UDHR is not legally binding, however, it has a profound influence on enacting international laws for human rights by being widely accepted by almost all members of the international community [Aus].

In the Declaration, article 12 reflects that "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks." [UDH48] and privacy, in turn, it has been a plank of UDHR and has been enshrined as a fundamental right since the very beginning of general assembly herewith.

Furthermore, in the course of time UNGA became more concerned about privacy because of the new technologies which are very popular, however, at the same time also vulnerable to electronic surveillance and interception. Ultimately, in December 2013, UNGA adopted resolution 68/167, which expressed deep concern at the negative impact that surveillance and interception of communications may have on human rights. UNGA affirmed that the rights held by people off-line must also be preserved online, and it called upon all countries to respect and protect the right to privacy in digital communication [Uni14].

The General Assembly called on all States to respect and protect the right to privacy, including in the context of digital communication and to review their procedures, practices and legislation related to communications surveillance, interception and collection of personal data and emphasized the need for States to ensure the full and effective implementation of their obligations under international human rights law. UNGA also encourages them to to take measures to put an end to violations of those rights and to create the conditions to prevent such violations, including by ensuring that relevant national legislation complies with their obligations under international human rights law [Uni14].

Lastly, UDHR is not the only international document which affirms the privacy as a fundamental human right. Similarly, The Council of Europe (CoE) declares a right to respect for one's private and family life, his home and his correspondence by Article 8 in European Convention on Human Rights and the Charter of Fundamental Rights of the European Union defines the same values and very alike principles by Article 7 and Article 8 [DDH+15].

III. Seven Principles of Privacy by Design

1. **Proactive & Preventative :** "The Privacy by Design approach is characterized by proactive rather than reactive measures. It anticipates and prevents privacy invasive events before they happen. PbD does not wait for privacy risks to materialize, nor does it offer remedies for resolving privacy infractions once they have occurred − it aims to prevent them from occurring. In short, Privacy by Design comes before-the-fact, not after." [Cav11]

2. **Privacy as the Default Setting :** "We can all be certain of one thing − the default rules! Privacy by Design seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. If an individual does nothing, their privacy still remains intact. No action is required on the part of the individual to protect their privacy − it is built into the system, by default." [Cav11]

3. **Privacy Embedded into Design :** "Privacy by Design is embedded into the design and architecture of IT systems and business practices. It is not bolted on as an add-on, after the fact. The result is that privacy becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality." [Cav11]

4. **Full Functionality :** "Privacy by Design seeks to accommodate all legitimate interests and objectives in a positive-sum “win-win” manner, not through a dated, zero-sum approach, where unnecessary trade-offs are made. Privacy by Design avoids the pretence of false dichotomies, such as privacy vs. security, demonstrating that it is possible, and far more desirable, to have both." [Cav11]

5. **End-to-End Security :** "Privacy by Design, having been embedded into the system prior to the first element of information being collected, extends securely throughout the entire life-cycle of the data involved — strong security measures are essential to privacy, from start to finish. This ensures that all data are securely retained, and then securely destroyed at the end of the process, in a timely fashion. Thus, Privacy by Design ensures cradle to grave, secure life-cycle management of information, end-to-end." [Cav11]

6. **Visibility & Transparency :** "Privacy by Design seeks to assure all stakeholders that whatever the business practice or technology involved, it is in fact, operating according to the stated promises and objectives, subject to independent verification. Its component parts and operations remain visible and transparent, to both users and providers alike. Remember, trust but verify!" [Cav11]

7. **User-Centric :** "Above all, Privacy by Design requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options. Keep it user-centric!" [Cav11]

   The 7 foundational principles of PbD have been already translated into 23 different languages. In this way, by this potential to reach global consent, PbD unites the essential components for effective data protection across boarders and in world society [Cav11].

IV. Privacy by ReDesign

In May 2011, Privacy by Design was extended into Privacy by ReDesign (PbRD) to provide a framework for enhancing privacy safeguards in existing and legacy systems, where the chance of embedded design by default had already passed. PbRD is introduced by [CP11]. It is claimed as a transformative process which targets to legacy systems for privacy remediation within 3 phases. These three phases are also called as 3 R's of PbRD: Rethink, Redesign and Revive in respectively [CP11].

Phase	Rethink	Redesign	Revive
Objective	Identify business and privacy requirements associated with the target system	Design and develop new controls to meet business and privacy requirements	Rollout redesigned, privacy-enhanced system
Key Activities	Confirm/establish business requirements Evaluate existing system privacy controls against *PbD* Principles Identify deficiencies (gap analysis) Define strategic business objectives, control requirements and initial implementation strategy	Design and build controls that meet business objectives while supporting *PbD* principles Eliminate earlier existing non-compliant controls Implement new controls Test new controls	Revalidate the redesigned target system against *PbD* Principles Deploy Confirm successful integration of redesigned target system
Outcome	Clear project objectives developed	Redesigned target system with new privacy controls in place	Organizationally-integrated target system aligned with *PbD* Principles

Figure 24. Diagram : Implementing Privacy by Design [CP11]

Figure 2 on the above displays the phases with their objectives, key activities and anticipated outcomes. According to the diagram, it is aimed to identify the business and privacy requirements that are related with the target system in the rethink phase. Framework achieves this objective by evaluating existing privacy controls and identifying deficiencies based on PbD principles (gap analyses). Finally this will help to define the strategic business objectives and to develop an initial implementation strategy. By the end of this phase, enhancing clear project objectives is the anticipated outcome [CP11].

In the redesign phase, the main objective is to design, build and test new controls that will be compliance in/with both the business and privacy requirements which identified in the rethink phase. In that way, eliminating non compliant controls and implementing the new controls for testing become the key activities of redesign phase. Redesigned target system is the anticipated outcome [CP11].

Finally, in the revive phase, the main objective is the integration of the redesigned system into the organization. Revalidation of the redesigned system by considering PbD principles, deployment of the redesigned system and its integration with other systems in the organization are key activities through the phase. At the end, anticipated output is having a fully-functional integrated, privacy-enhanced system [CP11].

V. Classification of the PETs by Pullonen et all.

Goal	Target	Examples of technology
Communication protection	Secure	Client-Server encryption, TLS, IPSec, End-to-End encryption, PGP, OTR
	Anonymous	Proxies and VPN, onion routing, mix-networks, broadcast
Data protection	Integrity	Message authentication codes, signatures
	Confidentiality	Encryption, secret sharing
Entity authentication	Identity based	Username and password, single-sign-on
	Attribute based	Credential used only once, zero-knowledge proofs
Privacy-Aware computation	Confidential inputs	Homomorphic encryption, secure multiparty computation, private information retrieval
	Privacy adding	Differential privacy, k-anonymity, cell suppression, noise addition, aggregation, anonymisation
Human-Data interaction	Transparency of data usage	Information flow detection, logging, declarations about information usage
	Intervenability	Information granularity adjustment, access control

Figure 25. Table : Classification of privacy enhancing technologies [PMB17]

VI. Association of PETs with PbD Strategies

Table 2. Association of PETs (Privacy Mechanisms) with PbD Strategies (Privacy Features) (H.ul - Hide / unlinkability, H.ud - Hide / undetectability, H.e - Hide / encryption, M - Minimize, A - Abstract, P - Plausible deniability, C - Control, I - Inform), Adapted from [Wuy15]

Category	Privacy Mechanisms: PETs	H.ul	H.ud	H.e	M	A	P	C	I
Anonymity system	Mix-networks (1981) [Cha81]	X		X	X				
	DC-networks (1985) [Cha85, Cha88]	X		X	X				
	ISDN-mixes [PPW91]	X		X	X				
	Onion Routing (1996) [SGR]	X		X	X				
	Crowds (1998) [RR98]	X		X	X				
	Single proxy (90s) (Anonymizer, SafeWeb)	X		X	X				
	Anonymous Remailer(Mixminion Type 3 (2003) [oIa])	X		X	X				
	Low-latency communication (Freedom Network [BGS01], Java Anon Proxy (2000) [BFK01], Tor (2004) [DMS04])	X		X	X				
	Invisible Internet Project (I2P [TCF12])	X		X	X				
	DC-net & MIX-net + dummy traffic, ISDN-mixes [PPW91]	X	X	X	X				
	Broadcast systems [PW, WP] + dummy traffic	X	X		X				
Privacy preserving authentication	Private authentication [AF04, ABB+04] + dummy traffic	X			X				
	Anonymous credentials (single show [BC], multi show [CL04])	X			X				
	Deniable authentication [Nao02]	X			X		X		
	Off-the-record messaging [BGB04]	X		X	X		X		
Privacy preserving cryptographic protocols	Multi-party computation (Secure function evaluation) [Yao82, NN01]	X		X					
	Anonymous buyer-seller watermarking protocol [RDB+10]	X		X	X				
Information retrieval	Private information retrieval [CKGS98] + dummy traffic	X	X		X				
	Oblivious transfer [Rab81, Cac98]	X		X	X				
	Privacy preserving data mining [VBF+04, Pin02]	X		X	X				
	Searchable encryption [ABC+07], Private search [OS07]			X	X				
Data anonymization	K-anonymity model [Swe02b, Swe02a], L-diversity	X			X	X			
	[MGKV06], T-closeness [LLV07]	X			X	X			
Information hiding	Steganography [AP98]	X	X		X				
	Covert communication [MCNM03]	X	X		X				
	Spread spectrum [KM01]	X	X		X				
Pseudonymity systems	Privacy enhancing identity management system [HBC+04]	X			X				
	User-controlled identity management system [CPHH02]	X			X				
	Privacy preserving biometrics [STP09]	X			X				
Encryption techniques	Symmetric key & public key encryption [MCV01]			X					
	Deniable encryption [DCIO16]			X			X		
	Homomorphic encryption [FG07]			X					
	Verifiable encryption [CD98]			X					
Access control techniques	Context-based access control [GMPT01]			X					
	Privacy-aware access control [CF08, ACK+10]			X					
Policy and feedback tools	Policy communication (P3P [OAL12])								X
	Policy enforcement (XACML [oas17], EPAL [And06])								X
	Feedback tools for user privacy awareness [LHDL04, PK09, LBW08]							X	
	Data removal tools (spyware removal, browser cleaning tools, activity traces eraser, harddisk data eraser)							X	

VII. İ* Model Views

A model that is developed within i* framework can be visualized into multiple views. Three most common views that are originated from the original i* proposal and some extensions : the Strategic Rationale (SR) view, the Strategic Dependency (SD) view, and the Hybrid view [DFH16]. In i*, each operational configuration is typically expressed through an SD model. The alternatives that are explored in an SR model refer to the alternative SD configurations that have different implications for the various strategic interests held by each actor [Yu09].

- **Strategic Dependency (SD):** The SD view shows each actor in the model, the actor association links, and the dependency relationships among various actors in an organizational context. The SD model is a higher level abstraction than typical process models such as data-flow diagrams, activity diagrams...etc. [DFH16] [Yu09].
- **Strategic Rationale (SR):** Goals, tasks, resources and softgoals are attributed to each actor in SR view, this time as internal intentional elements that the actor wants to achieve. The link is used to connect a task to a goal is called as a means-end link, indicating a specific way to achieve the goal. Typically there is more than one way to achieve a goal, so a goal in an SR model prompts the question – how else can this goal be achieved ? [Yu09]
- **Hybrid SD/SR:** "It is often useful to combine SD/SR views where some of the actors are open, but not all, focusing on the strategic rationale of a particular set of actors, and the actor links are hidden." [Yu09]

VIII. Tropos Modeling Activities

- **Actor Modeling:** It consists of identifying and analyzing different stakeholders of the system as social actors together with their intentions on specific goals. In particularly, during the late requirement phase, actor modeling focuses on the definition of the actor system-to-be [BPG+04].

- **Dependency Modeling:** It involves the identification of the dependencies between the different actors which depend on one another for goals to be achieved, plans to be performed, and resources to be furnished. Dependency modeling takes place through the first three Tropos development phases. During the early requirements analysis stage, dependency modeling is focused on identifying dependencies between the actors of the organization setting in which the system will operate. In late requirements analysis stage, the dependencies between the system and the actors of its organization setting are identified and some of the actors dependencies identified in the previous stage are refined due to the system introduction. During the architectural design the data and control flows between the different actors of the system are modeled in terms of dependencies providing the basis for mapping the system's actors to software agents [BPG+04] [Mou04].

- **Goal Modeling:** The internal goals of each actor identified through actor modeling are furthered analyzed by goal modeling for providing a more precise definition of the actor. It involves three basic reasoning techniques: means-end analysis, contribution analysis, and AND/OR decomposition. During the early requirements analysis, goal modeling helps to refine the initially identified actors by further analyzing their goals and identify new dependencies, or refine existing ones, whereas during the late requirements analysis, goal modeling helps to further analyze the goals of the system. Finally, in the architectural design phase, goal modeling motivates the first-decomposition of the system actors into a set of sub-actors [BPG+04] [Mou04].

- **Plan Modeling:** It can be considered complimentary to the goal modeling activity and it employs similar reasoning techniques [BPG+04].

- **Capability Modeling:** It starts at the end of the architectural design and it involves the identification of capabilities for of the actors of the system according to the goals, tasks and dependencies of each actor. "Individual" capabilities are assigned to the actors of the system to make them to define, choose and execute tasks for achieving their goals together with "social" capabilities that let actors to manage dependencies with the other actors [BPG+04] [Mou04].

IX. Licence

Non-exclusive licence to reproduce thesis

I, **Ilhan Çelebi**,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to:

   reproduce, for the purpose of preservation, including for the purpose of preservation in the DSpace digital archives until expiry of the term of validity of the copyright

   of my thesis

   Privacy Enhanced Secure Tropos: A Privacy Modeling Language for GDPR Compliance

   supervised by Raimundas Matulevičius

2. Making the thesis available to the public is not allowed.

3. I am aware of the fact that the author retains the right referred to in point 1.

4. I certify that granting the non-exclusive licence does not infringe the intellectual property rights or rights arising from the Personal Data Protection Act.

Tartu, 10.05.2018