A Critical Analysis of Privacy Design Strategies

Michael Colesky, Jaap-Henk Hoepman
Digital Security
Radboud University Nijmegen
Nijmegen, The Netherlands
{mrc, jhh} @cs.ru.nl

Christiaan Hillen
Valori Security
Valori
Nieuwegein, The Netherlands
christiaanhillen@valori.nl

***Abstract*—The upcoming General Data Protection Regulation is quickly becoming of great concern to organizations which process personal data of European citizens. It is however nontrivial to translate these legal requirements into privacy friendly designs. One recently proposed approach to make 'privacy by design' more practical is privacy design strategies. This paper improves the strategy definitions and suggests an additional level of abstraction between strategies and privacy patterns: 'tactics'. We have identified a collection of such tactics based on an extensive literature review, in particular a catalogue of surveyed privacy patterns. We explore the relationships between the concepts we introduce and similar concepts used in software engineering. This paper helps bridge the gap between data protection requirements set out in law, and system development practice.**

Keywords—data processing; privacy; software engineering; legal factors; data protection; privacy by design; design patterns

I. Introduction

The Council of the European Union is close to finalizing the General Data Protection Regulation (GDPR) [1], a binding legislation to replace the Data Protection Directive. This has important repercussions for organizations dealing with personal data. In addition to enforcing one implementation for all EU member states and any organizations operating in them, it introduces substantial fines for violations. Many IT systems and services, even in the EU, are unprepared for this [2]. Therefore, organizations need a way to address this problem.

Methodologies which translate legal requirements into software requirements do exist. One example is test casing legal requirements as though they were functional ones [3]. One other widely accepted approach to address data protection during software development is 'privacy by design' (PbD) [4]. This design philosophy 'bakes-in' privacy throughout the system development lifecycle [5]. However, PbD in itself lacks concrete tools to help software developers design and implement privacy friendly systems. It also lacks clear guidelines on how to map specific legal data protection requirements into system requirements. Software design patterns (i.e. privacy patterns) address the former shortcoming of PbD. They provide guidelines for solving recurring software development problems [6]. To address the latter shortcoming, 'privacy design strategies' have recently been proposed [5].

These strategies are intended as an accessible model in which system engineers can consider privacy protection during the analysis and requirements engineering phase. They thus provide a potential bridge between the legal and engineering domain, especially if they were to be mapped and correlated with design patterns. Unfortunately, their original definitions are broad and vague, needing refinement if they are to be used in practice. We redefine them more concretely as engineering approaches to PbD which correspond to architectural goals.

In addition to making the definitions more concrete, we correlate and map the strategies against privacy patterns. We use both an extensive literature review and argumentation (similar to e.g. [7]) to achieve this. The greater area of privacy engineering was examined, with a focus on privacy patterns. From this a pattern catalogue was formed to link with and improve the strategies [8].

During this endeavor we discovered the potential for a useful additional layer of abstraction: approaches to privacy by design which contribute to an overarching strategy. Referring to these as 'tactics', we created a hierarchy that allows comprehension and classification opportunities. Our approach to this compares to the thought organization tools of Wuyts et al. [7] and Urquhart et al. [9]. However, our approach uses a greater assortment of patterns, applies to both legal and engineering contexts, and still remains usable, straightforward, and consistent. We subsequently found that 'architectural tactics' already exist in the software architecture domain [10].

This turns out to provide an interesting connection between the work on strategies and the software architecture domain. Architectural tactics serve to achieve system quality attributes (like performance, usability or security), much like our tactics facilitate privacy protection goals expressed as strategies. In fact, we believe privacy protection is another quality attribute. This paper is a first step towards exploring this relationship.

Tactics are the primary contribution of this paper, as detailed in Section III. This is accompanied by definitions for both the strategies and tactics, and the association of tactics to patterns. We describe this in Section IV. Section V provides a critical analysis of our approach, where we associate strategies with GDPR entities and personal data processing examples. We reveal the inner structure of our strategy definitions and provide concise alternatives to them. We discuss the conclusions and limitations of our work in Section VI.

II. Terms and Concepts

We begin with describing a couple of relevant concepts and define a few necessary terms.

This research is conducted within the Privacy and Identity Lab (PI.lab).
This research is supported by the Netherlands Organization for Scientific Research (NWO) as project 'Patterns for Privacy' (CYBSEC.14.030).

DOI 10.1109/SPW.2016.23

A. Software Architecture

In software development, *architecture* is considered the highest level of abstraction, consisting of structures which include elements, their properties, and the relationships among them [10] [11]. It is the first stage in which requirements are addressed. These requirements do not only encompass functionality, but also system quality attributes like security.

System quality attributes capture important non-functional properties of a system. Architectural tactics are used to achieve these quality attributes. For example, in the case of security as a quality attribute, it is characterized by the confidentiality, integrity, and availability (CIA) triad, and includes tactic categories like 'detect' and 'resist' attacks [10]. Tactics belong to these overarching categories.

B. Privacy Protection

Both the engineering domain and US legal framework use the term 'privacy', while the EU legislative variation is termed 'data protection'. These are not however interchangeable. We therefore choose to combine the ideas present in each and, as opposed to using them interchangeably, refer to the combined concept as 'privacy protection' [12]. We believe that privacy protection, like security, is a quality attribute. For systems that process personal data, their design has a strong influence on how well they protect privacy. Functionality provided by any organization, especially processing European data, should give quality assurances of privacy protection [13].

In the US context, privacy protection is sector specific and less comprehensive, requiring additional concerns with the replacements for Safe Harbor. Notably, the GDPR focuses on the limited ways in which organizations may lawfully facilitate the 'free flow of information', with a strong focus on purpose specification, limitation, proportionality, and consent [1].

C. Processing of Personal Data

A concept pivotal to EU data protection legislation, 'processing of personal data' is used to account for various activities. Article 2(b) of the European Data Protection Directive [14], and Article 4(3) of the proposed GDPR [1], give specific examples to which the term applies, like collecting and storing data. This broad usage introduces confusion when an engineer wants to refer to actual data processing – that is, performing operations on data. So instead, we propose to use the word 'operate' in this case – TABLE I. presents mutually exclusive actions that together represent the examples of processing of personal data and aid us in making our revised definitions more precise.

In our revised strategy definitions, which we present below, we furthermore use the term 'agreed upon purposes' to denote: *specified purposes, for which the data subject has freely given specific informed consent, or where required by indicated legitimate grounds*. This usage is intended as either a comprehensive reflection of GDPR 'purpose' and 'consent', or depending on specific jurisdiction, in as much as it is legal to process personal, private, or sensitive information.

TABLE I. ACTIONS COMPARED TO 'PROCESSING OF PERSONAL DATA' IN DATA PROTECTION LEGISLATION

Action	Relevant GDPR Personal Data Processing Examples
Operate	Adaptation; Alteration; Retrieval; Consultation; Use; Alignment; Combination
Store	Organization; Structuring; Storage
Retain	opposite to (Erasure; Destruction)
Collect	Collection; Recording
Share	Transmission; Dissemination; Making Available; opposite to (Restriction; Blocking)
Change	unauthorized third party (Adaptation; Alteration; Use; Alignment; Combination)
Breach	unauthorized third party (Retrieval; Consultation)

III. STRATEGIES AND TACTICS

This section provides a consistent and comprehensive framework for our strategy and tactic definitions. It includes a separation of data/policy oriented strategies, goal oriented groupings, and information processing examples per strategy.

With the preceding terms set out in the previous section, we now provide our new definition for privacy design strategies.

DEFINITION

Privacy Design Strategy: *specifies a distinct architectural goal in privacy by design to achieve a certain level of privacy protection*

We note that this is different from what is understood to be an architectural strategy within the software engineering domain [10]. Instead our strategies can be seen as the goals of the privacy protection quality attribute. We explore these strategies by introducing tactics for each of them. Our tactics contribute to the strategies, and therefore privacy protection as a quality attribute. Hence our usage of the term tactic is very close to its use within software engineering. In the context of our paper, this means facilitating the strategies which achieve privacy protection. Tactics are defined as follows.

DEFINITION

Tactic: *an approach to privacy by design which contributes to the goal of an overarching privacy design strategy.*

These tactics are summarized in TABLE II. They are based on findings from an extensive privacy pattern literature review, which catalogued each of around 100 privacy patterns against a corresponding strategy [8].

TABLE II. STRATEGIES BY TACTICS

MINIMISE	HIDE	SEPARATE	ABSTRACT
EXCLUDE SELECT STRIP DESTROY	RESTRICT MIX OBFUSCATE DISSOCIATE	DISTRIBUTE ISOLATE	SUMMARIZE GROUP
INFORM	**CONTROL**	**ENFORCE**	**DEMONSTRATE**
SUPPLY NOTIFY EXPLAIN	CONSENT CHOOSE UPDATE RETRACT	CREATE MAINTAIN UPHOLD	AUDIT LOG REPORT

These patterns were further categorized by the tactics – that is, less general descriptors of their approaches to accomplish the strategy goals, and therefore privacy protection quality. The results are the tactics that will be presented in more detail alongside the strategies further on in this paper.

In addition to those tactics found through the method described above, HASH, ENCRYPT and TEST were considered for the HIDE and DEMONSTRATE strategies in addition to those identified through the patterns. It was however more pragmatic to feature both ENCRYPT and HASH in OBFUSCATE. For TEST, absence from the privacy patterns can be explained by the more conventional view of testing as a quality assurance process. It is instead featured in software design patterns [15]. Also, AUDIT can accommodate testing in a privacy context. TABLE III. shows a summarized sample of the collection of privacy patterns, demonstrating some of the patterns which fell under the ENFORCE strategy and its tactics.

TABLE III. ENFORCE TACTICS FOR PRIVACY PATTERNS

	Tactics & Patterns	Description
CREATE	Creating Privacy Policy [16]	A legal document which conveys the risks an organization's activities may pose to a person's privacy and how it endeavors to reduce them.
	Fair Information Practices [17]	The FTC's proposed principles concerning informational privacy in the US online market - less comprehensive than the EU or OECD ones.
	Respecting Social Organizations [17]	Disparity between the intimacy and trust of systems and users may cause invasions of privacy. This pattern suggests Involving users in the privacy policy creation process.
MAINTAIN	Appropriate Privacy Feedback [18]	"Appropriate feedback loops are needed to help ensure people understand what [information] is being collected and who can see [it]"
	Maintaining Privacy Policy [16]	"As services evolve so does the amount of personal information they require, [this] pattern tackles [the evolution] of privacy policies"
	Privacy Management System [5]	Personalized systems may cater to privacy preferences on a user by user basis. These preferences should be adhered to.
UPHOLD	Usage Control Infrastructure [19]	A system which supports protocol and application independent data flow tracking, sticky policies, and external policy enforcement.
	Distributed Usage Control [20]	Once access to data has been granted, control over that access may be lost. This pattern maintains rules through distributed systems.
	Sticky Policies [21]	When personal information is processed through multiple entities, they act under obligatory previous disclosed policies to prevent violations.

IV. IMPROVED DEFINITIONS FOR THE PRIVACY STRATEGIES

This section uses the discussion from previous sections, coupled with a consistent framework, to analyze and redefine each of the strategies. They are examined against descriptions by Hoepman [5] and similar concepts in the field.

The framework for defining the aforementioned privacy design strategies focuses on the main goal of each strategy. This is either a form of limitation, prevention, provision, or ensurance. Each strategy's goal is realized through various methods, and may be approached in multiple listed ways. The strategies are also examined for their significance in privacy protection. An explicit definition is provided for each strategy towards the end of the subsection describing it, followed by a layer of abstraction comprising of numerous tactics. These tactics are highlighted in the initial discussions around each strategy in italics and are formally defined after each strategy. The definition notes the types of actions performed on the data which are affected by the strategy, and an emphasis is made on exceeding legal requirements. This excessive mind-set falls within the constraints suggested by 'agreed upon purposes'.

A. *MINIMIZE*

The first of these strategies, MINIMIZE, specifically 'data minimization', advocates minimal collection and operation on personal data. The two main ways as per [5] to approach the strategy include: an all or nothing refusal of processing (*exclusive*), or granular privacy settings (*selective*). Data should be open to selection prior to collection, during operations, and also while stored. This can be done through data *stripping* – removing fields entirely.

Unlike ISO 29100 [22] and Cavoukian [23], who separate data minimization from 'collection limitation', the strategy includes it. This decision was likely in favor of avoiding a distinct 'limit' strategy. Fortunately, there is little to discourage combining collection limitation into the strategy. Cavoukian's [23] description of data minimization begins "…the collection of [personal data] should be kept to a strict minimum…" as opposed to collection limitation, in which, she stresses fairness, lawfulness, and 'specified purposes' – another distinct principle. Despite this, it is presented separately. We conclude that it is more practical for MINIMIZE to encourage the non-collection of purposeless data.

In ISO 29100 [22], minimization 'strictly minimizes the [operating]' on personal data, yet it includes a 'need-to-know' principle describing access control. The strategy does not provide for this. This includes sharing with third parties. This is because MINIMIZE is more concerned with the data itself than access to it. Instead, the HIDE strategy (see Section B) accounts for access. The ISO 29100 variation also features unlinkability – another aspect which, for the same reason HIDE should cover. This variation also features data retention. However, due to the fact that permanently *destroying* data is decidedly more minimization oriented than merely unlinking or removing access, retention is more suitable as an element of MINIMIZE. We formally define our revision as follows.

DEFINITIONS

MINIMIZE: *limiting usage as much as possible by excluding, selecting, stripping, or destroying any storage, collection, retention or operation on personal data, within the constraints of the agreed upon purposes*.

This strategy features the following proposed tactics:

EXCLUDE: *refraining from processing a data subject's personal data, partly or entirely, akin to blacklisting or opt-out.* ('Don't Disturb' [24])

SELECT: *decide on a case by case basis on the full or partial usage of personal data, akin to whitelisting or opt-in.* ('Partial Identification' [17])

STRIP: *removing unnecessary personal data fields from the system's representation of each user.* ('Strip Metadata' [25])

DESTROY: *completely removing a data subject's personal data.* ('Limited Data Retention' [25])

B. *HIDE*

Access control and sharing is covered by the second strategy, HIDE [5]. While it may be inferred that HIDE also provides for data retention limitation, this is not explicitly stated. On the other hand, data which is deleted but recoverable, or *dissociated*

through removal of links is HIDE orientated. The strategy supports confidentiality, unlinkability, and unobservability, assumedly working towards anonymity, undetectability, and pseudonymity, based on the way in which these terms are defined [26]. This includes measures to *dissociate*, and encrypt or *obfuscate*. It supports data security, as per the Information Security principle of ISO 29100 [22] and Safeguards in OECD's basic principles [27]. It also implicitly adheres to the principle of collection and use or purpose limitation. As an abstracted term, HIDE supports data quality and subject rights in the context of dissociation.

DEFINITIONS

HIDE: *preventing exposure as much as possible by mixing, obfuscating, dissociating, or restricting access to any storage, sharing or operation on personal data, within the constraints of the agreed upon purposes*.

This strategy includes:

RESTRICT: *preventing unauthorized access to personal data.* ('Access Control' [25])

MIX: *processing personal data randomly within a large enough group to reduce correlation.* ('Mix Networks' [25])

OBFUSCATE: *preventing understandability of personal data to those without the ability to decipher it.* ('Encryption' [5])

DISSOCIATE: *removing the correlation between different pieces of personal data.* ('Delayed Routing' [19])

C. SEPARATE

Data or procedure separation pertains to the *distribution* or *isolation* of personal data, in storage or operation, in order to make correlation for misuse more difficult. In essence, the strategy prevents putting together enough information about a data subject to endanger their privacy, which presents a close relation to the HIDE strategy. The separation intended is not only in system operations and tables in a database, but also in even physically distributed systems within reason. It does not cover collection, though data may be separated prior to storage.

The strategy can be related to purpose specification and limitation in the context of separating data according to its purpose, but this relation is not explicit. Furthermore, SEPARATE does not cover security safeguards, instead it can apply abstractly through access-controlled distributed storage.

DEFINITIONS

SEPARATE: *preventing correlation as much as possible by distributing or isolating any storage, collection or operation on personal data, within the constraints of the agreed upon purposes*.

Separation features the following variations:

DISTRIBUTE: *partitioning personal data so that more access is required to process it.* ('Privacy-Sensitive Architectures' [17])

ISOLATE: *processing parts of personal data independently, without access or correlation to related parts.* ('Physical Privacy Zones' [17])

D. ABSTRACT

The former AGGREGATE strategy accounts for data at the point of collection, storage and operation. It *summarizes* or *groups* data to the coarsest granularity still useful for operating. However, without sufficient k-anonymity [28] (size and diversity of the group over which it is aggregated), aggregation does not provide privacy protection. It should be used responsibly. AGGREGATE has two explicit variations: summarizing and grouping data. The relation between this strategy and collection, use, and data minimization is not explicit, but it can be derived from the strategy's intent. The strategy is redefined as 'ABSTRACT' (see Section V.D) below.

DEFINITIONS

ABSTRACT: *limiting detail as much as possible by summarizing or grouping any storage, collection or operation on personal data, within the constraints of the agreed upon purposes.*

Abstraction comprises of:

SUMMARIZE: *extracting commonalities in personal data by finding and processing correlations instead of the data itself.* ('Data Abstraction' [25])

GROUP: *inducing less detail from personal data prior to processing, by allocating into common categories.* ('Dynamic Location Granularity' [5])

These concealment and limitation orientated strategies are utilized between the data controller and or processor. While the implementations of these may rely on other strategies, or other actors, the strategies themselves are data oriented. The following strategies are instead process or policy oriented.

E. INFORM

Derived from transparency, the need for timely notification and informed decisions regarding personal data is embodied in the INFORM strategy. This includes the *supply* of verbose information related to policies and currently held data, the *explanation* of necessary detail in a concise and understandable form, and *notification* of any changes. Of which, only relevant information need be given to avoid notification fatigue [29]. The information presented to the data subject must include what is stored, operated or disseminated, why and how it is done, who has access to it, and when it will be destroyed. It also features a need for informing users about how their information is secured. Any changes, including breaches, should be promptly communicated. It is defined as follows.

DEFINITIONS

INFORM: *providing as abundant clarity as possible for supplying, explaining, and notifying on storage, collection, retention, sharing, changes, breaches or operation on personal data, in a timely manner, within the constraints of the agreed upon purposes.*

The INFORM strategy comprises of these tactics:

SUPPLY: *making available extensive resources on the processing of personal data, including policies, processes, and potential risks.* ('Privacy Policy Display' [30])

NOTIFY: *alerting data subjects to any new information about processing of their personal data in a timely manner.* ('Data Breach Notification' [5])

EXPLAIN: *detailing information on personal data processing in a concise and understandable form.* ('Privacy Icons' [30])

This strategy can be interpreted to apply to legal basis, legitimate grounds, and data subject rights through the need to inform data subjects about choices – including information on limitation of use and retention. It can also implicitly relate to security safeguards and data quality, as notifying users of the state of their information allows them to better CONTROL it.

F. CONTROL

The next strategy is CONTROL. Specifically, this relates to the data subject's control over their information's collection, storage, operation, and dissemination. They should be capable of informed consent, retrieval, modification, and retraction in an intuitive and timely fashion. While INFORM covers retrieval of stored personal data, and contributes to informed consent, the ability of the data subject to choose to consent, or not, is part of the CONTROL strategy. This closely relates to exercising a data subject's right to self-determination [31], and to keeping information up to date and accurate [31]. It also allows data subjects to explicitly have a say in collection, use and retention limitation. The strategy is defined as follows.

DEFINITIONS

CONTROL: *providing as abundant means as possible for consenting to, choosing, updating, and retracting from storage, collection, retention, sharing or operation on personal data, in a timely manner, within the constraints of the agreed upon purposes.*

The CONTROL strategy comprises:

CONSENT: *only processing the personal data for which explicit, freely-given, and informed consent is received.* ('Obtaining Explicit Consent' [16])

CHOOSE: *allowing for the selection or exclusion of personal data, partly or wholly, from any processing.* ('Discouraging Blanket Strategies' [32])

UPDATE: *providing data subjects with the means to keep their personal data accurate and up to date.* ('Reasonable Level of Control' [17])

RETRACT: *honoring the data subject's right to the complete removal of any personal data in a timely fashion.* ('Invisible Mode' [17])

What sets CONTROL and INFORM aside from the remaining strategies is that they focus on the data subject and controller. The following strategies focus on the controller and authority.

G. ENFORCE

This strategy advocates creating, ensuring, and complying with contractual and legal policy obligations. It accounts for technical controls and organizational controls through policies. It is however more in line with adhering to these than outright data security. It stipulates creation and maintenance of these prior to, during, and after development. It is defined as follows.

DEFINITIONS

ENFORCE: *ensuring as abundant commitment as possible for creating, maintaining, and upholding policies and technical controls regarding storage, collection, retention, sharing, changes, breaches or operation on personal data, in a timely manner, within the constraints of the agreed upon purposes.*

The ENFORCE strategy includes the following tactics:

CREATE: *acknowledging the value of privacy and deciding upon policies which enable it, and processes which respect personal data.* ('Fair Information Practices' [17])

MAINTAIN: *considering privacy when designing or modifying features, and updating policies and processes to better protect personal data.* ('Appropriate Privacy Feedback' [18])

UPHOLD: *ensuring that policies are adhered to by treating personal data as an asset, and privacy as a goal to incentivize as a critical feature.* ('Distributed Usage Control' [20])

This strategy covers purpose limitation and data quality [5]. It can be abstracted to include purpose specification, legal basis and legitimate ground (through creation), as well as security safeguards in the context of access control.

H. DEMONSTRATE

This strategy stems directly from compliance [5]. It specifies that a controller should be readily capable of showing that it adheres to legal requirements. Focus is placed on legal grounds and the controller's presented compliance with them. This can be shown through *auditing*, *logging*, and *reporting*.

DEFINITIONS

DEMONSTRATE: *ensuring as abundant evidence as possible for testing, auditing, logging, and reporting on policies and technical controls regarding storage, collection, retention, sharing, changes, breaches or operation on personal data, in a timely manner, within the constraints of the agreed upon purposes.*

The DEMONSTRATE strategy features these tactics:

LOG: *tracking all processing of data, without revealing personal data, securing and reviewing the information gathered for any risks.* ('Non-repudiation' [33])

AUDIT: *examining all day to day activities for any risks to personal data, and responding to any discrepancies seriously.* ('Privacy Audit Trails' [16])

REPORT: *analyzing collected information on tests, audits, and logs periodically to review improvements to the protection of personal data.* ('Building Trust and Credibility' [6])

Relations between DEMONSTRATE and legal concerns could arguably consist of all requirements, as adherence to all obligations should be demonstrated. However, this exists

primarily for what is enforced, and as such data subject rights, data minimization, and data quality are merely extensions on purpose specification, legal basis, and collection and use limitation. This correlation, and that of the previous strategies, is summarized along with definitions in Section V.

V. Privacy Design Strategies Analysis

We redefined the strategies and introduced tactics in Section IV. In this section, we analyze the ways privacy protection may be breached, we provide an overview of the definition framework, present concise variations, study the correlation with the various entities in data protection legislation, and discuss the validity of ABSTRACT and SEPARATE as strategies.

A. Strategies by GDPR Data Processing Examples

One way to examine the strategies is to determine their impact on actions that affect personal data, comparable to Solove's [34] taxonomy of privacy. The difference is the taxonomy has a wide subset of malicious activities, not explicitly including storage and retention. We group these with operate under Solove's term, 'processing'. The 'intrusion' and 'decisional interference' activities under the 'invasion' group are related to unauthorized changes and breaches. These are shown together in relation to the strategies in TABLE IV.

In this, it can be seen that 'retain' as opposed to 'operate' or 'store' is useful for relating between strategies and privacy affecting actions. Some strategies effect retention while others do not. However, the distinction between 'operate' and 'store' as well as 'change' and 'breach', is not as beneficial. While operations or changes are distinct when considering integrity, the use of both terms in definitions mainly serves for clarity. Inferring terms for brevity is not worth the extra complexity.

TABLE IV. Privacy Affecting Actions Impacted by Strategies

Actions	Taxonomy Group	Applies to	ENFORCE	DEMONSTRATE	INFORM	CONTROL	MINIMIZE	ABSTRACT	SEPARATE	HIDE
Operate	Processing									
Store										
Retain										
Collect		Collection								
Share		Dissemination								
Change	Invasion									
Breach										

B. Strategies by Definition

The strategy definitions are summarized within TABLE V. which illustrates the internal consistency of the definitions. Within the first four (ENFORCE, DEMONSTRATE, CONTROL, and INFORM) 'process oriented' [5] strategies, each goal is ensured or provided in abundance. The second, 'data oriented' group (MINIMIZE, ABSTRACT, HIDE, and SEPARATE), advocate a near-excess approach of limitation and prevention. Excess in either group with negative impact is accounted for through 'agreed upon purposes' as introduced in Section II.C.

Limitation and prevention refer to the reduction of impact and probability of a privacy protection failure. In this sense the data oriented strategies accomplish a kind of privacy protection risk mitigation. The difference between these orientations can be compared to Porekar, Jerman-Blažič, and Klobučar's [16] suggested 'privacy agreements' and 'dataflow' distinction. Some strategies have a unique focus on the data subject or authority. Others associate with dataflow around the controller.

It is also useful to provide a more concise alternative. When a number of factors are first given, the strategy definitions can safely be summarized without losing much detail. By sacrificing mention of associated actions, definitions for each privacy design strategy can be aptly shortened. This however requires that these definitions include the following considerations.

- *Concern for personally identifiable information (PII) or personal data, where personal information in these definitions is processed personal data* [35] [36]*;*
- *Association with one or more kinds of* ***personal information*** *(processing of personal data), where 'processing' may include all examples in Article 4 (3) of the proposed GDPR* [1]*;*
- *An 'as much as possible' perspective; and*
- *Constraint on agreed upon purposes as previously defined.*

Summarized Definitions

HIDE: *preventing exposure of access, association, visibility, and understandability of personal information to reduce the likelihood of privacy violations.*

MINIMIZE: *limiting usage of personal information to reduce the impact of privacy violations.*

TABLE V. Privacy Design Strategy Definition Framework

Strategy	Underlying Goals				Effects on Actions Regarding Personal Data										
ENFORCE	ensuring	as *abundant*	**commitment**	as *possible* for	**creating**, **maintaining** and **upholding**	on policies and technical controls regarding	storage,	collection,	retention,	sharing,	changes, breaches	or operation on	personal data,	in a timely manner,	within the constraints of the agreed upon purposes.
DEMONSTRATE	ensuring	as *abundant*	**evidence**	as *possible* for	**testing**, **auditing**, **logging**, and **reporting**	on policies and technical controls regarding	storage,	collection,	retention,	sharing,	changes, breaches	or operation on	personal data,	in a timely manner,	within the constraints of the agreed upon purposes.
CONTROL	providing	as *abundant*	**means**	as *possible* for	**consenting** to, **choosing**, **updating**, and **retracting**	From	storage,	collection,	retention,	sharing,	changes, breaches	or operation on	personal data,	in a timely manner,	within the constraints of the agreed upon purposes.
INFORM	providing	as *abundant*	**clarity**	as *possible* for	**supplying**, **explaining**, and **notifying**	On	storage,	collection,	retention,	sharing		or operation on	personal data,	in a timely manner,	within the constraints of the agreed upon purposes.
MINIMIZE	limiting		**usage**	as *much as possible* by	**excluding**, **selecting**, **stripping**, or **destroying**	Any	storage,	collection,	retention			or operation on	personal data,	in a timely manner,	within the constraints of the agreed upon purposes.
ABSTRACT	limiting		**detail**	as *much as possible* by	**summarizing** or **grouping**	Any	storage,	collection				or operation on	personal data,	in a timely manner,	within the constraints of the agreed upon purposes.
SEPARATE	preventing		**correlation**	as *much as possible* by	**distributing** or **isolating**	Any	storage,	collection				or operation on	personal data,	in a timely manner,	within the constraints of the agreed upon purposes.
HIDE	preventing		**exposure**	as *much as possible* by	**mixing**, **obfuscating**, **dissociating**, or **restricting** access to	Any	storage,	sharing				or operation on	personal data,	in a timely manner,	within the constraints of the agreed upon purposes.

SEPARATE: *preventing the correlation of personal information to reduce the likelihood of privacy violations.*

ABSTRACT: *limiting the detail of personal information to reduce the impact of privacy violations.*

CONTROL: *providing data subjects with means to consent to, choose, update, and retract from personal information in a timely manner.*

INFORM: *providing data subjects with clear explanation and timely notification on personal information.*

ENFORCE: *ensuring commitment to continually create, maintain, and uphold policies and technical controls regarding personal information.*

DEMONSTRATE: *ensuring available evidence to test, audit, log, and report on policies and technical controls regarding personal information.*

Since the definitions abandon the granularity of associated actions and involve a number of prerequisites, their usage is only suggested where contextually sufficient.

C. Strategies by Data Subject, Controller, and Authority

As mentioned during Section IV, strategies are associated with specific entities in the data protection legislation context. This relationship is shown in Fig. 1.

A subject controls their personal data through the controller, who informs them about that data. The controller is involved in all strategies, including enforcing policy on any processors, also represented as the controller. It demonstrates compliance to the authority, and applies data oriented privacy risk mitigation strategies. These interactions are somewhat cyclic, as only the controller and or processor use them. The following section investigates whether ABSTRACT and SEPARATE are justifiable as strategies as opposed to mere variations of MINIMIZE and HIDE.

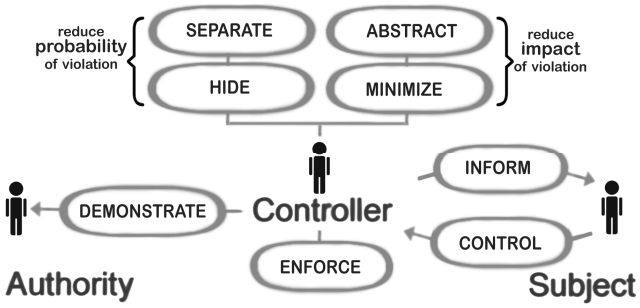

Fig. 1. Strategies by data protection legislation actors

D. ABSTRACT and SEPARATE

The ABSTRACT (AGGREGATE) and SEPARATE strategies border on the notion of a distinct architectural goal. Between limiting 'detail' as opposed to 'usage' (processing), or preventing 'correlation' as opposed to 'exposure' (access), these strategies may, like tactics, contribute to the goals of other strategies. However, this alone does not adequately justify their conversion to tactics.

In the case of SEPARATE, one may argue that its focus on decentralization, an important privacy aspect, grants it good reason to retain its position. Though with AGGREGATE, coupled with an unfortunate negative connotation (aggregation is often misused or otherwise detrimental to privacy, for example through inference [12] [18] [28] [32]), the limitation of detail goal is perhaps better encapsulated in the term 'abstract'.

Abstraction can be defined in multiple ways [37], including "considering something as a general quality or characteristic". While merely renaming a strategy does not relieve it of misuse, we believe ABSTRACT better represents the goal, and therefore, the purpose of the AGGREGATE strategy.CONCLUSIONS

This paper introduces architectural tactics in-between strategies and patterns, and performs a critical analysis on the privacy protection characterizations known as privacy design strategies. Hoepman's original strategies have been examined in detail, resulting in refined definitions which are more precise with less overlap, and conform to a consistent framework. We have exposed the relationship between the strategy approach and the software architecture domain, providing evidence for considering privacy protection a quality attribute.

This definition framework includes clear and explicit relations between the strategies and the GDPR concepts of 'purpose' and 'consent', as well as examples of processing on personal data. Adding to this are the mapping to privacy patterns through underlying architectural tactics, which are given their own formal definitions. The ABSTRACT and SEPARATE strategies are challenged and confirmed as legitimate strategies, and better justified for their retention, while AGGREGATE has been renamed to ABSTRACT to better cover its intent.

A. Limitations of Our Approach

This work was limited in scope for the purpose of clarity and succinctness. We have focused on the use of strategies and privacy patterns to realize PbD as an alternative approach to requirements translation methodologies like [3]. PbD is a popular design philosophy, therefore it is important to make it more concrete. Our work is a first step towards this goal. Even though we have focused on the European data protection framework (GDPR) to steer the definitions, we believe our results are also relevant and valuable in other jurisdictions.

Our selection of tactics is open to interpretation and is driven by the fact that we take a legal point of departure. We believe we have examined a sufficient number of privacy patterns, in sufficient depth, to achieve precision in our results. However, the suitability of the tactics requires further evaluation.

B. Suggestions for Future Work

Considering the aforementioned limitations, there are multiple directions in which our results can be improved, adapted or otherwise built upon. In particular, we suggest to investigate the effectiveness of strategies and tactics in practice including assessments of usability, and measurable quality improvements. We suggest that this be achieved through various case studies and based on argumentation for adequate

measurements of usefulness. We also note the opportunity for introducing further tactics, though this was not our goal. These could perhaps include tactics inspired by social and ethical instead of legal theories of privacy, or perhaps anti-tactics.

We would like to thank the anonymous referees and Daniel Smullen for valuable feedback and discussions.

REFERENCES

[1] European Commission, "Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)," *COM(2012) 11 final including SEC (2012) 72 final and SEC (2012) 73 final*, vol. 2015, no. June, pp. 1–201, 2015.

[2] Ipswitch, "European IT Teams Woeful Lack of Preparation for General Data Protection Regulation (GDPR) May Mean Painful Compliance Audits Ahead," *Ipswitch.com*, 2014. [Online]. Available: http://www.ipswitchft.com/about-us/news/press-releases/2014/11/gdpr-may-mean-painful-compliance-audits-ahead. [Accessed: 29-Jun-2015].

[3] T. D. Breaux, A. I. Antón, and E. H. Spafford, "A distributed requirements management framework for legal compliance and accountability," *Computers and Security*, vol. 28, no. 1–2, pp. 8–17, 2009.

[4] A. Cavoukian, "Operationalizing Privacy by Design : A Guide to Implementing Strong Privacy Practices," pp. 1–72, 2012.

[5] J.-H. Hoepman, "Privacy Design Strategies," *IFIP SEC 2014*, pp. 446–459, 2014.

[6] E. S. Chung et al., "Development and Evaluation of Emerging Design Patterns for Ubiquitous Computing," *DIS '04 Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, pp. 233–242, 2004.

[7] K. Wuyts, R. Scandariato, B. De Decker, and W. Joosen, "Linking privacy solutions to developer goals," in *Proceedings - International Conference on Availability, Reliability and Security, ARES 2009*, 2009, pp. 847–852.

[8] "privacypatterns.eu - collecting patterns for better privacy." [Online]. Available: https://privacypatterns.eu/. [Accessed: 20-Oct-2015].

[9] L. Urquhart, T. Rodden, and M. Golembewski, "Playing the Legal Card : Using Ideation Cards to Raise Data Protection Issues within the Design Process," *Proc. CHI'15*, pp. 457–466, 2015.

[10] L. Bass, P. Clements, and R. Kazman, *Software Architecture in Practice*, 3rd ed. Addison-Wesley Professional, 2012.

[11] M. Hamilton, *Software Development: Building Reliable Systems*. 1999.

[12] ISO/IEC, "ISO/IEC 15944-8:2012 Information technology -- Business Operational View -- Part 8: Identification of privacy protection requirements as external constraints on business transactions," 2012.

[13] The European Commission, *EU Commission and United States agree on new framework for transatlantic data flows: EU-US Privacy Shield*, no. February. Strasbourg, 2016.

[14] European Parliament and Council of European Union, "Directive 95/46/EC of the European Parliament and of the Council," *Official Journal of the European Communities*, vol. 281, no. 31, pp. 31–50, 1995.

[15] A. V. Feudjio, I. Schieferdecker, and A. Vouffo, "Test Automation Design Patterns for Reactive Software Systems," *EuroPLoP*, 2009.

[16] J. Porekar, A. Jerman-Blažič, and T. Klobučar, "Towards organizational privacy patterns," *Proceedings - The 2nd International Conference on the Digital Society, ICDS 2008*, 2008.

[17] H. Baraki et al., *Towards Interdisciplinary Design Patterns for Ubiquitous Computing Applications*. Kassel, Germany: Kassel University Press GmbH, 2014.

[18] G. Iachello and J. Hong, "End-User Privacy in Human-Computer Interaction," *Foundations and Trends® in Human-Computer Interaction*, vol. 1, no. 1, pp. 1–137, 2007.

[19] M. Hafiz, "A Pattern Language for Developing Privacy Enhancing Technologies," *Software - Practice and Experience*, vol. 43, pp. 769–787, 2013.

[20] F. Kelbert and A. Pretschner, "Towards a policy enforcement infrastructure for distributed usage control," *Proceedings of the 17th ACM SACMAT (Symposium on Access Control Models And Technologies)*, p. 119, 2012.

[21] S. Pearson and Y. Shen, "Context-aware privacy design pattern selection," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6264 LNCS, pp. 69–80, 2010.

[22] ISO/IEC, "ISO/IEC 29100:2011 Information technology -- Security techniques -- Privacy Framework," 2011.

[23] A. Cavoukian, "Privacy by Design The 7 Foundational Principles Implementation and Mapping of Fair Information Practices," *Information and Privacy Commissioner of Ontario, Canada*, 2009.

[24] T. Shümmer, "The Public Privacy – Patterns for Filtering Personal Information in Collaborative Systems," in *Proceedings of CHI workshop on Human-Computer-Human-Interaction Patterns*, 2004.

[25] C. Bier and E. Krempel, "Common Privacy Patterns in Video Surveillance and Smart Energy," in *ICCCT-2012*, 2012, pp. 610–615.

[26] A. Pfitzmann and M. Hansen, "A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management," *Identity*, pp. 1–98, 2010.

[27] Organisation of Economic Co-Operation and Development, "OECD Guidelines governing the Protection of Privacy and Transborder Flows of Personal Data," 2013.

[28] L. Sweeney, "k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[29] Committee on Civil Liberties Justice and Home Affairs, "Draft Report on the proposal for a regulation of the European Parliament and of the Council on the protection of individual with regard to the processing of personal data and on the free movement of such data," 2014.

[30] C. Graf, P. Wolkerstorfer, A. Geven, and M. Tscheligi, "A Pattern Collection for Privacy Enhancing Technology," *The Second International Conferences of Pervasive Patterns and Applications (Patterns 2010)*, vol. 2, no. 1, pp. 72–77, 2010.

[31] N. N. G. de Andrade, "Oblivion : The Right to Be Different … from Oneself Reproposing the Right to Be Forgotten," *IDP. Revista de Internet, Derecho y Política*, vol. 13, no. February, pp. 122–137, 2012.

[32] S. Ahern et al., "Over-Exposed ? Privacy Patterns and Considerations in Online and Mobile Photo Sharing," *CHI '07*, pp. 357–366, 2007.

[33] L. Compagna, P. El Khoury, F. Massacci, R. Thomas, and N. Zannone, "How to capture, model, and verify the knowledge of legal, security, and privacy experts : a pattern-based approach," *ICAIL '07*, pp. 149–153, 2007.

[34] D. J. Solove, "Understanding Privacy," *Harvard University Press*, 2008.

[35] A. Sølvberg and D. C. Kung, *Information Systems Engineering: An Introduction*, Illustrate. Springer Science & Business Media, 2012.

[36] A. Narayanan and V. Shmatikov, "Myths and fallacies of 'personally identifiable information,'" *Communications of the ACM*, vol. 53, no. 6, p. 24, 2010.

[37] Dictionary.com Unabridged, "abstraction," *Random House, Inc.* [Online]. Available: http://dictionary.reference.com/browse/abstraction. [Accessed: 02-Dec-2015].