A LINDDUN-Based Framework for Privacy Threat Analysis on Identification and Authentication Processes

Antonio Robles-González[1], Javier Parra-Arnau*[,2] and Jordi Forné[1]

[1] Department of Telematics Engineering, Universitat Politècnica de Catalunya, C. Jordi Girona 1–3, E-08034, Barcelona, Spain

[2] Department of Computer Science and Mathematics, Universitat Rovira i Virgili, CYBERCAT-Center for Cybersecurity Research of Catalonia, E-43007 Tarragona, Spain

Abstract — Identification and authentication (IA) are security procedures that are ubiquitous in our online life, and that constantly require disclosing personal, sensitive information to non-fully trusted service providers, or to fully trusted providers that unintentionally may fail to protect such information. Although user IA processes are extensively supported by heterogeneous software and hardware, the simultaneous protection of user privacy is an open problem.

From a legal point of view, the European Union legislation requires protecting the processing of personal data and evaluating its impact on privacy throughout the whole IA procedure. Privacy Threat Analysis (PTA) is one of the pillars for the required Privacy Impact Assessment (PIA). Among the few existing approaches for conducting a PTA, LINDDUN is a very promising framework, although generic, in the sense that it has not been specifically conceived for IA.

In this work, we investigate an extension of LINDDUN that allows performing a reliable and systematically-reproducible PTA of user IA processes, thereby contributing to one of the cornerstones of PIA. Specifically, we propose a high-level description of the IA verification process, which we illustrate with an UML use case. Then, we design an identification and authentication modelling framework, propose an extension of two critical steps of the LINDDUN scheme, and adapt and tailor the trust boundary concept applied in the original framework. Finally, we propose a systematic methodology aimed to help auditors apply the proposed improvements to the LINDDUN framework.

Keywords — Privacy Threat Analysis, Privacy Impact Assessment, LINDDUN, trust boundary, authenticable attribute, trust-based attribute.

1. Introduction

Internet services are becoming increasingly sophisticated and are ubiquitously reaching nearly every daily-life environment. Among these services, emerging smart communities and social networks demand more and more user interaction to perform identification (I) and authentication (A) procedures. Typically, these procedures are quite repetitive, interrupt the primary task of actually using and enjoying the service itself, and more importantly, might have an impact on user privacy. Usually, to carry out an IA process, users send personal, sensitive information to a service provider that might not be fully trusted, or being so, might want to share this information with other providers and third parties. Therefore, an IA process embracing different domains of responsibilities could result in unwanted information disclosure and/or linkability, and ultimately jeopardize user privacy. Although user IA processes are present in a large variety of procedures and are supported by heterogeneous software and hardware, the simultaneous protection of user privacy is an open problem and is the focus of this paper.

* Corresponding author. Tel.: +34 977 558 270

Email addresses: antonio.robles@entel.upc.edu (Antonio Robles-González), javier.parra@urv.cat (Javier Parra-Arnau), forne@entel.upc.edu (Jordi Forné)

Manuscript revised October 22, 2019

From a legal point of view, the European Union legislation requires protecting the processing of personal data throughout the whole IA procedures. Among others, privacy objectives are identified by performing a Privacy Impact Assessment (PIA) and several recommendations of how to conduct a PIA are given by governments, the European Union itself and scientists. All them demand to perform a Privacy Threat Analysis (PTA) as one pillar for a reliable PIA. The recommendations on how to conduct a PIA, however, focus predominantly on describing the procedure to follow but without neither guiding the auditor through the necessary PTA nor providing specialized systematic tools or methods for a reliable PTA.

To the best of our knowledge, LINDDUN [1] is the most promising systematic PTA framework, that uses an information-flow-oriented system representation and relies on a Data Flow Diagram (DFD) methodology. LINDDUN, nonetheless, is a generic framework in the sense that it has not been originally conceived for the IA procedures tackled in this work. The fact that IA procedures focus solely on authenticity and non-repudiation and do not aim to safeguard user privacy motivates the development and study more systematic PTA methodologies and frameworks that are applicable to user IA processes.

1.1. Contribution and Plan of this Paper

The purpose of this paper is to investigate an extension of LINDDUN that allows performing a reliable and systematically-reproducible PTA of user IA processes, and thus to contribute to one of the pillars of a reliable PIA. The realization of a high-level description of the whole verification (IA) process, the creation of a systematic modelling framework and the improvement of the LINDDUN PTA framework are crucial, further aspects investigated in this work. Also, from an instructional-guidance perspective, our works aims to provide step-by-step instructions for auditors to systematically apply the proposed methodology. The ultimate objective of this paper is to provide them with a comprehensive tool-set to analyze their environment.

We would like to stress, in the context of this work, the relevance of LINDDUN, whose usage is predominant when tackling threat modelling problems. We would like to emphasize, however, that LINDDUN largely addresses general security threat modelling and currently cannot be applied directly to identification and authentication processes.

More specifically, the main contributions of this work are described next:

I. We propose a high-level description of the IA verification process, which we illustrate with an UML use case. We describe the process of a user demanding access to a service, including the sequence *user demand – service login – user verification - service access*. The creation of the UML is accompanied by the categorization of the IA processes into centralized and decentralized, and the definition if they are realized as one or two components (unit/threat).
II. We develop an identification and authentication modelling framework and give a generic overview of possible combinations of IA methods. We extend the modelling of user verification introducing, among others, the DFD representation, a user data repository, DFD related trust boundaries, the concept of centralized and decentralized and local and external authentication.
III. We propose an extension of two critical steps of the LINDDUN scheme (specifically, steps 1 and 2) with the previously created DFD-based IA modelling framework, and further develop the trust boundary concept applied in the original LINDDUN framework.
IV. We propose a systematic methodology aimed to help auditors apply the proposed improvements to the LINDDUN steps 1 and 2, so that they can continue with step 3 of the original LINDDUN framework.

The remainder of the paper is organized as follows. §2 presents the background, state of the art of PIA as well as PTA, and the LINDDUN framework. The developed IA modelling framework, the extended LINDDUN methodology and one-page instructions list are presented in §3. The evaluation of a proof of concept with two variants is done in §4. Finally, conclusions are drawn in §5.

2. BACKGROUND AND STATE OF THE ART

We review the background and state of the art of related technologies. We start with PIA and PTA. Existing PTA approaches are derived from Security Threat Analysis (STA) solutions but do not tackle

PTA from a systematically enough perspective. LINDDUN is as far as we know one scientifically substantiated systematic methodology exclusively used for PTA.

2.1. Privacy Impact Assessment and Threat Analysis

2.1.1. Privacy Impact Assessment

A PIA [2] is performed for determining the privacy objectives of a system. In Europe, the PIA Framework recommendation was created in the project "A Privacy Impact Assessment Framework for data protection and privacy rights" (PIAF) [3]. Generally, it is recommended that a PIA initially should be done in a short and if necessary in an extended version. Going through handbooks, guides or other formal descriptions of how to perform a PIA, e.g., [4], [5], [6], the conclusion is the same as for the PIAF project. All present a widespread set of recommendations, procedure descriptions and/or check lists, etc., and all require a high degree of intuition by the person realizing the PIA. This person is not always the necessary expert for a substantiated PTA and the PIA procedure does not offer special PTA support to guide the person realizing the PIA. A PTA is the starting point to perform a PIA. According to the ENISA Privacy Report [2], existing privacy risk analysis methods use adopted security analysis methods, e.g., EBIOS [7] and STRIDE [8].

In the specific context of RFID, a couple of PIAs are as follows. One is proposed by the European Commission [9] and the second by the BSI [10]. The PIA guideline [10] for RFID created by the BSI considering the European Privacy and data protection Impact Assessment Framework for RFID applications [9] is usable for dedicated RFID-based scenarios and offer a solid guideline. The European PIA Guideline [9] is based on BSI and guides through three RFID-based scenarios from the retail, public transportation and automotive environment.

The "Conducting PIA" of the UK information Commissioner's office [6] describes the process to carry out a PIA and guide to identify the privacy related risk on a very high level without explicitly referring to privacy threats.

In the context of the present paper, we will focus on PTA, the indispensable fundament for every convincing PIA. A systematic approach for PTA is required to make it easier for the auditor to perform a reliable PIA.

2.1.2. Privacy Threat Analysis

A PTA is the starting point to perform a PIA. According to the ENISA Privacy Report 2014 [2], the only existing privacy risk analysis methods adopt security risk analysis methods, e.g., EBIOS [7] and STRIDE [8]. The former focuses more on the methodology for privacy risk management, considers threats in a high abstraction level and tackles security needs such as confidentiality, integrity and availability [11]. The STRIDE methodology, on the other hand, is the initial point to develop LINDDUN. As explained in the next subsection, LINDDUN is a specialized PTA framework that instructs the pertinent stuff performing the PTA on how to make a system model and provides for this purpose a list of threat types. Also, it instructs how to map them to elements on the system model. Next, we elaborate more on this framework.

2.2. LINDDUN Framework: A Systematic Approach for Privacy Threat Analysis

Throughout different scientific documents LINDDUN[1] is referenced as one applicable PTA methodology and/or is used to analyze concrete scenarios, e.g., in health systems [12, 13]. LINDDUN is to the best of our knowledge the only promising PTA framework that is systematically and scientifically proven.

The LINDDUN methodology offers a systematic procedure for eliciting and fulfilling privacy requirements and is based on STRIDE [8], an approach for security threat modelling. The LINDDUN framework was first presented in [1] and, according to their authors, the primary contribution is the systematic methodology to model privacy specific threats. A further important contribution is that it provides an extensive catalogue of privacy specific threat tree patterns [14] and defines a mapping of most commonly known privacy enhancing technologies (PET) to identified privacy threats.

[1] LINDDUN is an acronym of these privacy threat categories: Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness, Noncompliance.

One of the authors of LINDDUN evaluated the framework in [15] and provided some improvements; a contribution of [15] we want to stress at this point is the extension of the LINDDUN privacy threat catalog. Another contribution to be highlighted is the reduction of interaction between LINDDUN and STRIDE.

The improvement of the LINDDUN framework proposed in [15] leads to the improved methodology LIND(D)UN, that is described in the tutorial [16] and the corresponding updated "LIND(D)UN privacy threat tree catalog" [14]. We will use throughout the paper LINDDUN, since we will consider the information disclosure threat.

The LINDDUN framework is divided in two phases. The former is the "PROBLEM SPACE" and the latter the "SOLUTION SPACE", as shown in Figure 1 (original figure taken and identically redrawn by ourselves).

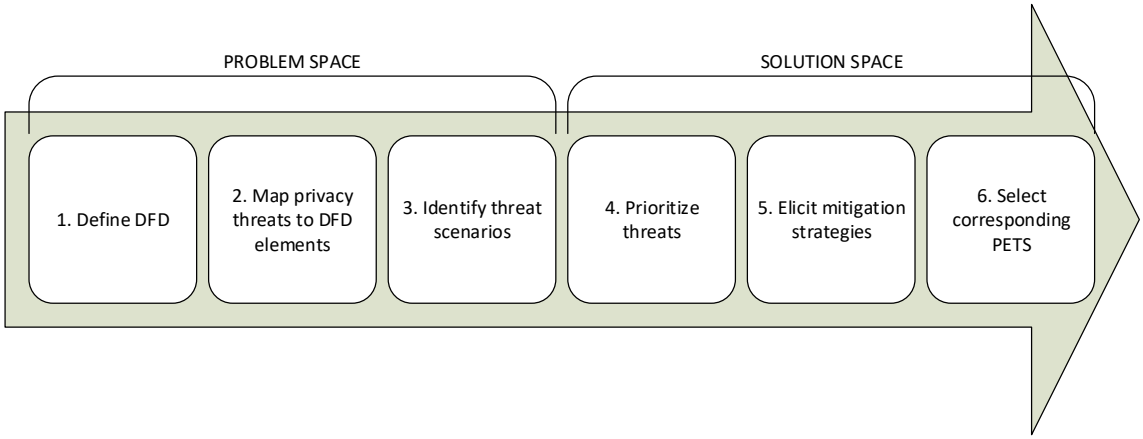

Figure 1. The formalized LINDDUN steps [1].

The emphasis throughout the present paper is the PTA, and for this reason the focus will be on the "PROBLEM SPACE" of the LINDDUN framework (see Figure 12 in the Appendix), and hence on steps 1 and 2. The problem-oriented steps of LINDDUN rely on [16], [14], [17] and [18].

3. IA MODELLING FRAMEWORK DEVELOPMENT AND APPLICATION TO THE ENHANCED LINDDUN FRAMEWORK

In this section, we propose an IA modelling framework suitable to extend the subsequent privacy-aware analysis and illustrate its application with a use case. We start in §3.1 with the presentation of a preliminary background for I and A. More specifically, in §3.2 the high-level description of the IA verification process for the use case of a user demanding service access is shown with UML notation. The IA modelling framework is developed in §3.3. During the development, common IA methods are gathered and presented in Table 1 and Table 2. The IA methods are modelled using the DFD, sub-phases are defined, and trust boundaries are considered. The extension of the LINDDUN framework, shown in §3.4, is contrived to be able to perform PTA on IA methods. The LINDDUN Privacy Threats are mapped to the DFD-based IA modelling framework and the trust boundary concept of the LINDDUN framework is tailored. A straightforward usable instruction list of how to use the previously worked out contributions is presented in §3.5.

3.1. Background on Identification and Authentication

In the present section a basic background for I and A processes is given to be used in the course of §3.3.

Throughout the paper, the definition used for identity is: "An identity is any subset of attribute of an individual person which sufficiently identifies this individual person within any set of persons. So usually there is no such thing as "the identity", but several of them." [19].

An identity required for the use of a certain service represents a "partial identity" [19], also "a subset of attribute values of a complete identity" of an individual person and "where a complete identity is the union of all attribute values of all identities of this person". Throughout the paper

we will use the term *Identity* representing a *partial identity* of all attributes related with one user (person).

The concept of identity mentioned in [19] comprises a subset or all identity attributes a service can require to be proved by the user passing an IA process.

In accordance to [20] we use the definitions for I and A: "Identification is the process of using claimed or observed attributes of an entity to deduce who the entity is." "Authentication is the corroboration of a claimed set of attributes or facts with a specified, or understood, level of confidence". In this context, we would like to point out that, for an auditor, identification is sometimes used as a synonym of authentication [21].

Authentication factors are used by the user to give evidence of their claim done by presenting the identity. The authentication-factors are grouped in three recognized categories that are as follows. The user can give evidence by demonstrating to know a knowledge (something you know), to have something in his possession (something you have) or to be him (something you are, biometric) [22, 23].

3.2. Use Case of User Demanding Service Access

The generic use case *service provision* for a user demanding service access is presented for modelling purpose using UML in Figure 2. The steps *service demand*, *service Login*, *user verification* and *service usage* represent at an overview of the steps the process to be passed by the user.

Depending on the user interaction throughout the *user verification,* we introduce the categorization into *centralized user verification* (user only communicates with Service) and *decentralized user verification* (user communicates with service and I / A components).

We assume that the I and A components can be realized together as one component (IA) or in two different components (I)-(A), so real circumstances can be considered. I and A components can be realized as hardware or software artefacts. The arrows interconnecting the categories and components below *user verification* indicate common combinations.

Depending if the service to be used and the components (IA), (I), (A) belong to the same or different domains, the user verification is determined as local or external authentication; further details will be given in the context of trust boundary consideration in §3.3.5.

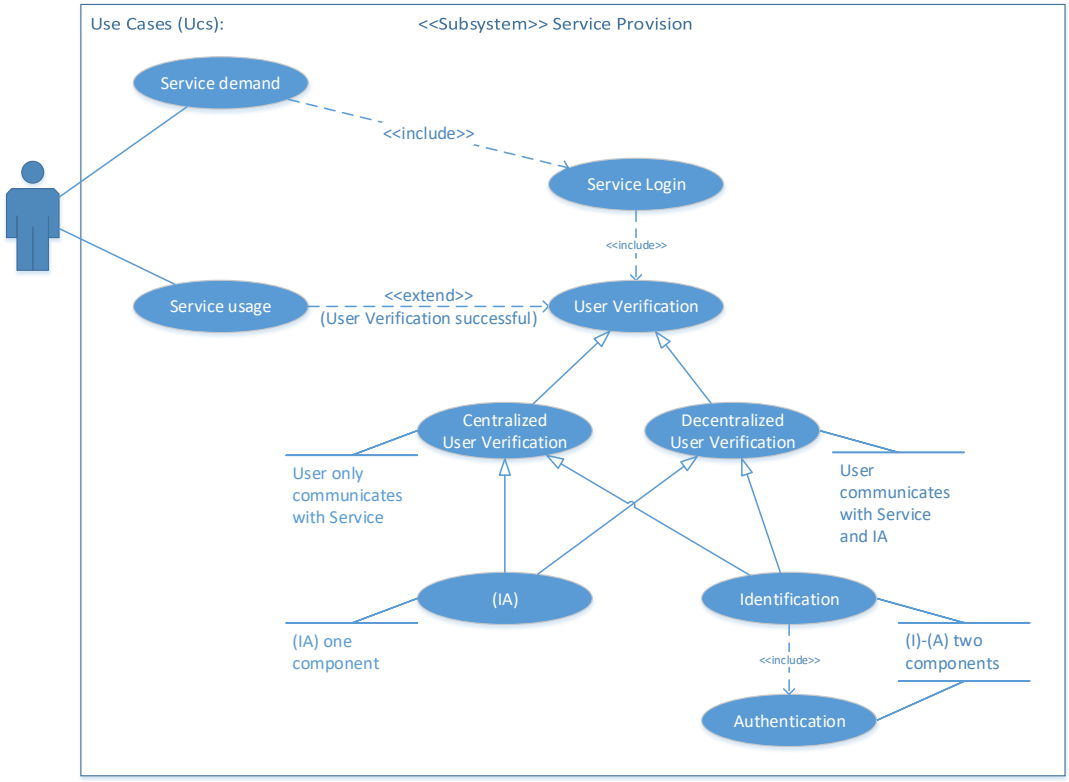

Figure 2. UML use case of user demanding service access.

3.3. Identification and Authentication Modelling Framework

In §3.3.1, the three step I and A process is defined. §3.3.2 presents tables with IA methods and possible combinations. §3.3.3 introduces the DFD for modelling purpose. The phases and subphases scope of the identification and authentication process is given in §3.3.4. §3.3.5 relates the concept of trust boundaries with the characteristics of identification and authentication methods.

3.3.1. Three-Step Identification and Authentication Process

Our starting point consider the definition for identification and authentication [20] that is a process in two steps. We parse the two steps I and A as follows into three steps: *Identity presentation*, *Identification* and *Authentication*. Now, before defining the three-step I and A processes we want to point out authenticable and not-authenticable attributes.

AUTHENTICABLE AND NOT-AUTHENTICABLE ATTRIBUTES

Theoretically, the provision of information by the user can be done during the whole IA process and will depend on the service requirements and used IA methods. We categorize the information a user can provide into *authenticable attributes* and *not-authenticable attributes.*

Authenticable attributes require that the user on his part can prove towards the service provider the correctness and/or legitimate usage of the presented attributes; these belong to one identity or partial identity of the user.

Not-authenticable attributes are passed to the service provider without any direct prove of correctness or if the user is legitimated to use it. These can be grouped into *free collected attributes* by the service provider or by the user additionally *voluntary given attributes*. Based on the applied *transitivity of trust* the service provider assumes the additionally voluntary given attributes are true, therefore, they are called *trust-based attributes*.

We want to point out the arising risk of privacy threats when the user, in addition to proving authenticable attributes, gives trust-based attributes. The consideration of all given user information is of major interest for an integral PTA that is beyond the scope of the present paper. The scope of the present paper is the PTA for authenticable attributes in the context of IA methods.

Accordingly, we define the three IA process steps *Identity presentation*, *Identification* and *Authentication* (IIA). In a two-step IA process, step 1 is usually included in step 2. We describe these three steps next:

- Step 1: Identity presentation is the consideration of how a subset of identity attributes are presented by the user. The user presents the required subset of identity attributes to a service, so that the user claims to be someone (or something), e.g., presenting a userID, username or other attributes. In step 1 we only consider attributes that are required to pass the (I)IA process, therefore, to be proved. The introduction of step 1 *Identity presentation* was done to cover, if necessary, all possibly existing technical realization of IA methods.
- Step 2: Identification in the present context is defined as the verification of the plausibility of the presented "subset of" (identity) "attributes" [19]. The plausibility verification can comprise the verification of the technical correctness (e.g., syntax, format, length, etc.), but can include the semantical verification of the plausibility (e.g., age in realistic range, age minimum is given, etc.) before proceeding with the proof of the presented attributes.
- Step 3: Authentication is the prove of the claim done by the user with the presentation of the subset of identity attributes in step 1 and/or 2, therefore, to confirm the legitimate usage and/or correctness of the presented identity attributes. This step in the best case is done self-determined by the user, e.g., introducing a password or personal identification number (PIN).

3.3.2. IA Methods: Creation of Tables for I-Methods and A-Methods

The three steps "Identity presentation", "I" and "A" (IIA) defined in §3.3.1 require a technical base. For this purpose, technical IA methods and authentication-factors and -Protocols are used to create IA methods tables.

IDENTITY PRESENTATION IN THE CONTEXT OF IDENTIFICATION AND AUTHENTICATION

Identification methods comprises the procedure and technical components that the user applies to present his identity to the service. The selected identification method facilitates the user to manually or electronically pass the required attribute to the service, therefore, the user manually types in the required details of the identifier or electronically passes the information with technology based, e.g., on barcode, magnetic strip, NFC and/or a smartcard.

Furthermore, recall that the acronym IA imply that "I" includes the Identity presentation and identification (II) and "A" is the abbreviation of authentication. In the remaining part of the subsection the compilation of the I-method Table 1 including the most common methods for realizing Step 1: Identity presentation, Step 2: Identification and to gather user provided attributes is done. We show in Table 2 the compilation of A-methods including the most common methods for Step 3: Authentication. Table 2 also depicts part of the possible combination of IA-methods.

Next, we describe the manual and electronic identity presentation methods

The categorization of identification methods is conducted depending on the provision method applied to pass the required user attributes (e.g., loginID, username, name, etc.) to the service and are the categories manually and electronically.

- *Manually*: The user types in the required attributes, e.g., his loginID he knows or is printed on a smartcard, magnetic card or similar plastic card.

  Access (protection) to the attributes is

  "free": the access to the attribute, e.g., printed on the card is without any restriction.
- *Electronically*: The user presents a smartcard, magnetic card or another similar card that is electronically readable using at least one of the following methods: optically (barcode, machine readable zone), magnetic strip card, smartcard with contact or by proximity using NFC (e.g., NFC smartcard or RFID tag).

  Access (protection) to the attributes is

  "free": the attribute is accessible without any restriction (barcode, RFID, smartcard),

  "restricted": the identity/attribute can only be read by (authorized) terminals (RFID, smartcard readable only with, e.g., cryptographic key) or

  "auth": the identity/attribute can only be read or verified by (authorized) terminals after additional user authorization with, e.g., password/pin and are called authenticable attributes.
- We introduce the *user information storage/user* data repositoryin the context of identity presentation methods for the user environment towards a more reliable and systematic user centric analysis; this implies the presence of a storage/database usable by the user and could be his brain for accessing the username or another identifier or medium, e.g., smartcard, smartphone or capability he possess to access the cloud[2].

Table 1 shows Identity presentation methods including one group of rows for authenticable attributes, therefore, to be proved by the user and a group of rows for trust-based attributes provided voluntary by the user without additional prove. For more details see not-authenticable attributes in §3.3.1. The *User-ID* is one possible attribute of the identity of the user and for that an authentication proof ("authenticable attribute") could be required and of course it could be demanded the proof of more than one attribute.

The input row in Table 1 describes how the trust-based attributes will be passed to the system, therefore, typed in, by a barcode, MRZ, contact reader or proximity (NFC) Reader. The row storage describes where the attributes are stored, e.g., on optical readable barcode, smart card and NFC Tag. We add to these storages the user memory and named it known to user. In Table 1 the identity presentation method properties of the presented authenticable attributes can be gathered and which trust-based attributes (see §3.3.1) are provided additionally by the user.

[2] There are still ideas and first realization of IA solutions based on attributes stored in the cloud

Identity Presentation/ Identification Method (ID-M)	ID-M properties			Authenticable (A) Attributes				Input method can vary from that used for A-Attributes	Trust Based (TB) Attributes given by user during IA process or afterwards		
ID-Method-Name	Storage	Input	Access (protection)	A-Attr1 e.g. User ID	Au-Attr 2 e.g. address	Au-Attr 3 e.g. adult	...	Input	TB-Attri1 e.g. hobby	TB-Attr2 e.g. name	...
Manually											
M-user	Known to user	Typed in	Free								
M-card	Printed on a card	Typed in	Free								
Electronically											
E-barcode	Optical readable	barcode	free								
E-MRZ	Optical readable	machine readable zone (MRZ)	free								
RFID-Tag	NFC-Tag	proximity	free								
RFID-Tag	NFC-Tag	proximity	restricted								
RFID-Tag	NFC-Tag	proximity	auth								
E-magnetic	Magnetic card	Reader	free								
E-contact-SC	Contact smart card (SC)	Reader	free								
E-contact-SC	Contact smart card (SC)	Reader	restricted								
E-contact-SC	Contact smart card (SC)	Reader	auth								
E-NFC-SC	NFC smart card	proximity	free								
E-NFC-SC	NFC smart card	proximity	restricted								
E-NFC-SC	NFC smart card	proximity	auth								

Table 1 Identity Presentation methods.

COMPILATION OF I- AND A-METHODS COMBINATION

We assemble in Table 2 A-methods, authentication factors, general recognized procedures (protocols) and requirements for securing user authentication:

Multi-Factor-Authentication: usage of two or more authentication factors. Verification process using more than one authentication factor is called multi-factor authentication [22].

Challenge Response (CR) Based Authentication Procedure: An entity (claimant) proves his identity to another entity (verifier) by demonstrating knowledge of a secret, without revealing the secret itself to the verifier during the protocol [21]. Known variants of CR-based authentication [21] could rely on techniques like “One-time password”, “symmetric-keys” or “public-key”. A special CR-based procedure is the Zero knowledge procedure [22].

Challenge Response Procedure (each authentication with a new password/credential) summarized: One-time password-based (One-time password, e.g., S/Key (Lesli Lampert), OTP RFC2289), Symmetric cryptosystem, Asymmetric cryptosystem. Zero-knowledge procedure (is special CR procedure [22]): Ask randomly a subset of available credentials.

Strong Authentication: the definition is ambiguous and could mean that multiple answers have been requested (CR Zero-Knowledge), it must be based on a challenge response protocol or that the verification may not be accomplished by sending the secret. In the following consideration we will use the definition of strong authentication, see [22], therefore, the methods based on Challenge Response (CR) and without sending the secret.

In Table 2 A-factors can be used in combination with different authentication procedures (protocols) that are ordered from weak to strong and, e.g., that *secret not revealed* is marked with (X) indicates that in the meantime it is an accepted and recognized practice and indispensable. The authenticable attributes are either provided during the identity presentation step (see in Table 2 in column *attributes* the cell with the text “Table 1”) or implicitly with the authentication method (see in Table 2 in column *attributes* the cell with the text *A-method*). When considering Table 2 for a PTA in §3.4 with LINDDUN framework the (X) will indicate that it is (quasi) mandatory to fulfill this requirement. Table 2 is a template for gathering information of the system to be analyzed. Systems using whatever IA-methods could require (and is recommendable) to apply in their realization the procedure of “mutual authentication and secure communication channel” (secure channel).

As explained in §3.3.5, the concept of trust boundary and trusted third party (TTP) related authentication, local authentication (inside the same domain) and external authentication (cross domain), is used throughout diverse IA-methods, too. Both concepts are used to expand Table 2The that has at the end two more categories. These are "mutual authentication and secure communication channel" and "trust boundary and trusted third party (TTP) related authentication".

The Table 2 for identification-methods (I-methods) and authentication-methods (A-methods) show a few of the possible and usually used combinations of I-methods and A-methods. Each combination is a generic IA-Type. Table 2 will serve as template to guide the auditor to elicit the analyzed IA environment for applying LINDDUN [1]. Table 2: has embedded in the center an *authentication method* table.

The output of the present section is a set of tables related with IA methods usable as part of a tool set by the auditor for gathering the actual status of the environment and model it afterwards. To our best knowledge we did not found similar tables for I- and A-methods.

User ID presentation/identification methods	Attribute properties / Authenticable (A) Attributes	Authentication Method	Authentication factors and attributes							Authentication Procedure (Protocols) and concepts									Mutual Authentication & Secure Communication channel		Trust Boundary & Trusted Third Party (TTP) Related Authentication				Combined IA-Method
	Table 1 Identity Presentation Methods	A-Method Name	Authentication factors			Attributes		A-Method	Weak	Weak toward strong auth	Challenge Response based (CR) (underlying protocol)			Strong Auth			Generic CR secret not revealed	Mutual Auth		Local (security domain) authentication		External (security domain) authentication		I-Auth-Method
			Knowledge	Possession	Biometric	Table 1: ID Presentation Methods Access protection	Table 1: ID Presentation Methods Authenticable	Authenticable	Static password	OTP	OTP	Symmetric key	Asymmetric key	Zero knowledge				Peer Authentication	Secure Channel e.g. https	Local	Local SSO	Federated	Fed. SSO	Name
M-user		Single Factor	Password			free			x								(X)		(X)					
M-user		Two-factor	Password	OTP by smartphone		free/auth					x						(X)		(X)					
E-NFC-SC		Two-factor	Pin	Smart card		auth							x				(X)		(X)					
E-NFC-SC		Single Factor		Smart card	fingerprint	free							x				(X)		(X)					
E-NFC-SC		Two-factor		Smart card		auth							x				(X)		(X)					
E-barcode		Single Factor		Smart card		free											(X)		(X)					
E-contact-C		Single Factor		Smart card		free											(X)		(X)					
E-contact-C		Two-factor	Pin	Smart card		auth											(X)		(X)					
E-MRZ						free											(X)		(X)					
Id-Service	Auth-Service																			Trustboundary				

Table 2: A-methods, Combinations of I-methods and A-methods and trust boundary.

3.3.3. Data Flow Diagram

The application of the LINDDUN framework [1, 15] is based on a DFD describing the environment to analyze. The core components required for identification and authentication are user, identification service, authentication service and (application) service provision and for each component one database/data store is assumed. To illustrate the application of the LINDDUN framework, we present a generic DFD for the identification and authentication environment (Figure 3):

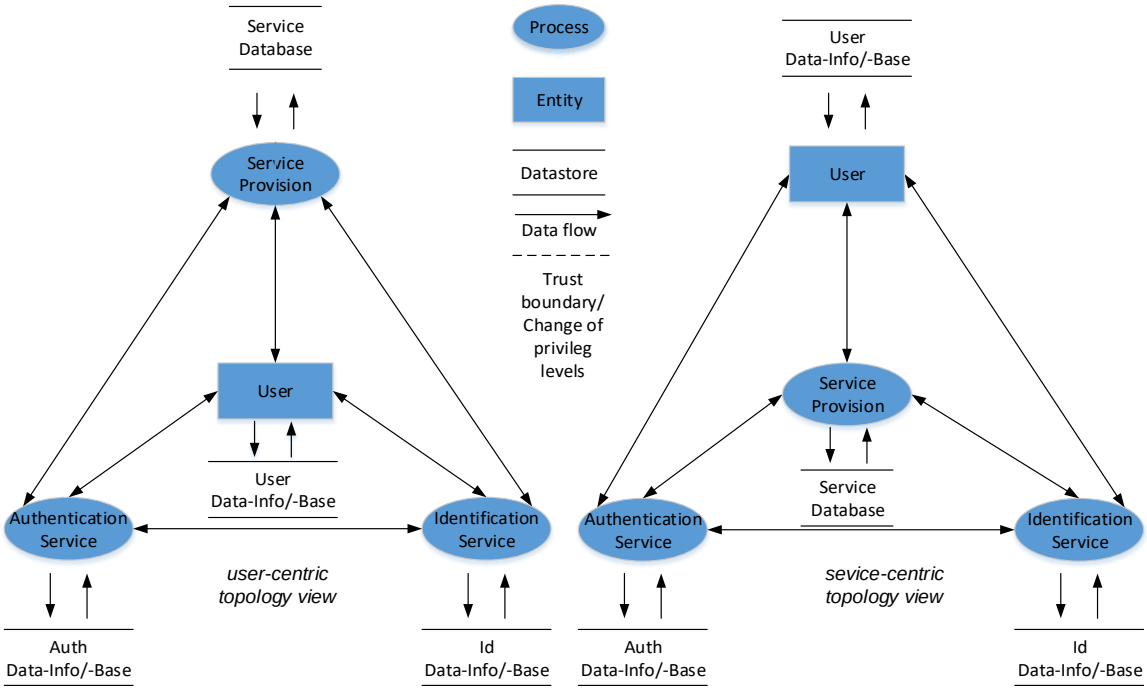

Figure 3. Generic DFD of the identification and authentication processes (user-centric or service-centric topology view).

In contrast to the presented DFD in the LINDDUN paper [1], we introduce a *user data-info/-base (repository)* that can be used for a more detailed analysis of IA methods. An example for the location of a user data-info/base could be a device brought along by the user to provide or confirm required attributes, therefore, for proving the claim as well as the attributes could be stored in the cloud. Further details will be given later in §3.3.2.

An arrow with two arrowheads between two components indicates that in principle a communication in both directions is possible and could be subdivided in two arrows with opposite head direction. The detailed communication to be considered will finally depend on the IA methods implemented.

The DFD elements of Figure 3 are:

Entity: User U; Processes: identification (I) ≙ (I)-P, authentication (A) ≙ (A)-P, service provision (S) ≙ S-P, identification-authentication (IA) ≙ (IA)-P; data Store: user data-/info-base ≙ U-DB, identification database ≙ (I)-DB, authentication database ≙ (A)-DB, identification-authentication database (IA)-DB, service provision database ≙ S-DB; data flow: "bidirectional arrows" ≙ "↔", "unidirectional arrows" ≙ "→" or "←".

User or service centric representation

In Figure 3 both views are given, the service-centric as well the user-centric view. Is it possible to gain different benefits for the LINDDUN analyses depending on which of both views have been used, therefore, the user- or the service-centric representation?

Applying the DFD-IA-Modell on the one hand, user-centric and, on the other hand, service-centric in our opinion gives only an advantage in the visualization that can be useful when the components depending on the real implementation belong to different domains and differ from the user domain and must be grouped together. Another conceivable visualization of the content could be a three-dimensional figure offering different perspectives. In this paper, we consider the service centric DFD element arrangement as depicted in Figure 3.

3.3.4. Process Phases P1 – P4 and Sub-Phases

In the present section, the IA phases and sub-phases are investigated. The derived extended generic DFD, including the (Sub-)Phases, is shown in Figure 4. The user access process to the service is divided in four phases P1 to P4, as explained below.

Figure 4 shows four phases in which the user, service provider and IA-service can be involved, and the details depend on the IA System to be analyzed; here it is assumed the user in P1 demands the usage of the service. The identification and authentication process are carried out in phases P2 and P3. Phase P4 represents the authorization to use the service after successful authentication.

The sub-phases P1 to P4 and the resulting phase diagram for a complete identification and authentication processes will depend on the system to be analyzed, so that only P1 and P4 are detailed and the rectangle for P2 and P3 will be replenished later by the auditor depending on the real system to be examined. The auditor can use for this purpose Figure 4 as template and gather for the place holders *Auditor verifies for P2 to P3 range of influence* which components and/or user of the analyzed system are participating in each of these sub-phases.

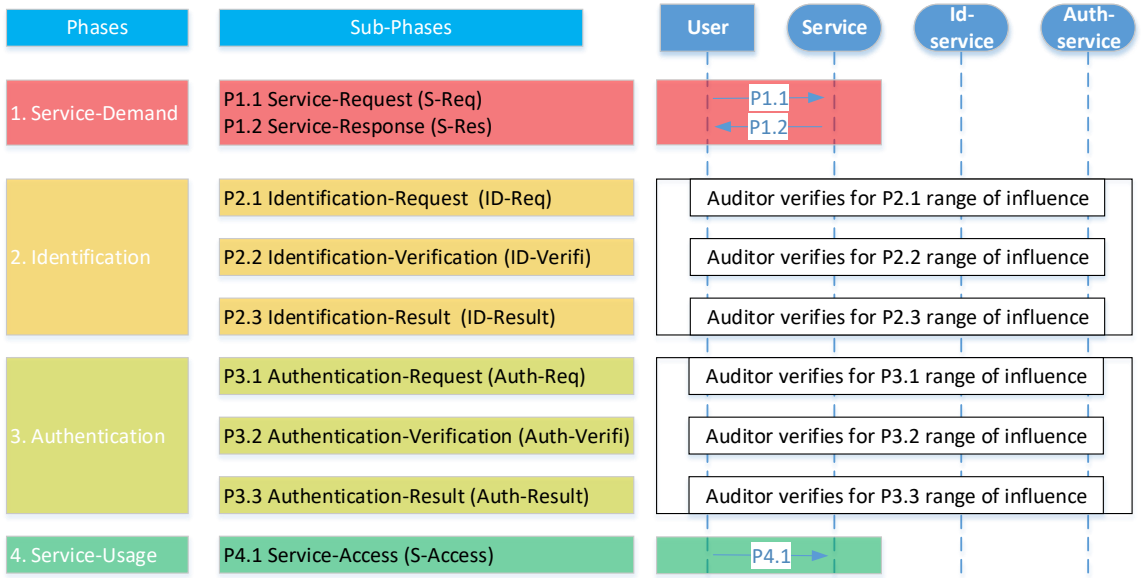

Figure 4. Extended DFD with (Sub-)Phases P1 to P4.

3.3.5. Trust Boundaries

In the present subsection we introduce the concept of Centralized and Decentralized User Verification, Local and External Authentication and Mutual authentication and secure channel.

CENTRALIZED AND DECENTRALIZED USER VERIFICATION

The categories *centralized user verification* (user only communicates with Service) and *decentralized user verification* (user communicates with service and I / A components) introduced in §3.2 are depicted in Figure 5 including the trust boundaries given by the domain borders and used for further explanation.

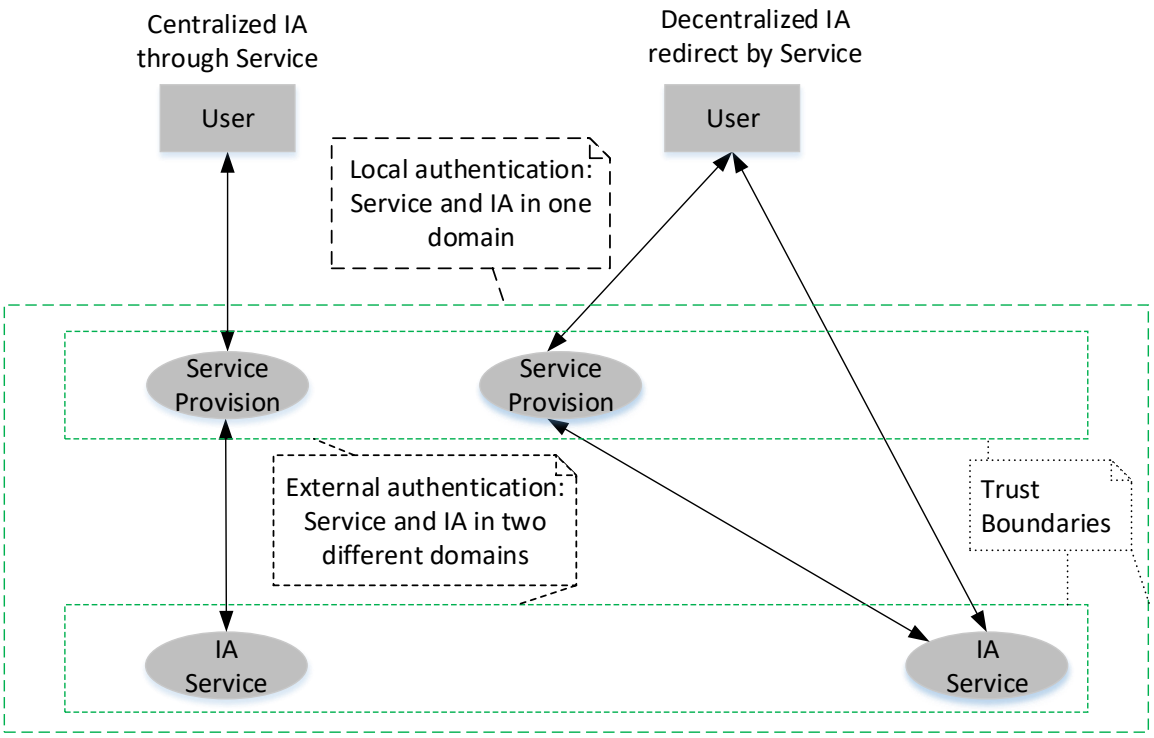

Figure 5. Trust boundaries centralized- and decentralized-user verification.

LOCAL AND EXTERNAL AUTHENTICATION

The categorization into *local authentication* and *external authentication*, see Table 3 and Figure 5, refers to the domain where the authentication is performed, therefore, if in the local domain (where the service reside) or at an external domain and we assume that the identification is done together in the same domain with the authentication, too. The presented model and concept could be applied for the case that the identification is performed locally and only the pure authentication is done through the external domain, too. The definition of what is to be considered local or external depends on the trust relation between the components, the environment and the user, therefore, on the course of trust boundaries.

Local Authentication (inside one domain):

The service (S) provider receives the service-request and will perform the identification and authentication processes in a centralized or decentralized manner, only communicating with the IA service (TTP) in the own local domain (see Figure 5).

As depicted in Figure 6, the service n in domain 1 to be accessed by a user of domain 1 (D1) will contact inside his own local domain 1 an instance, e.g., called IA service domain1, for performing the IA of the user. The user access to all other services of domain 1 will rely on the same IA service of domain1.

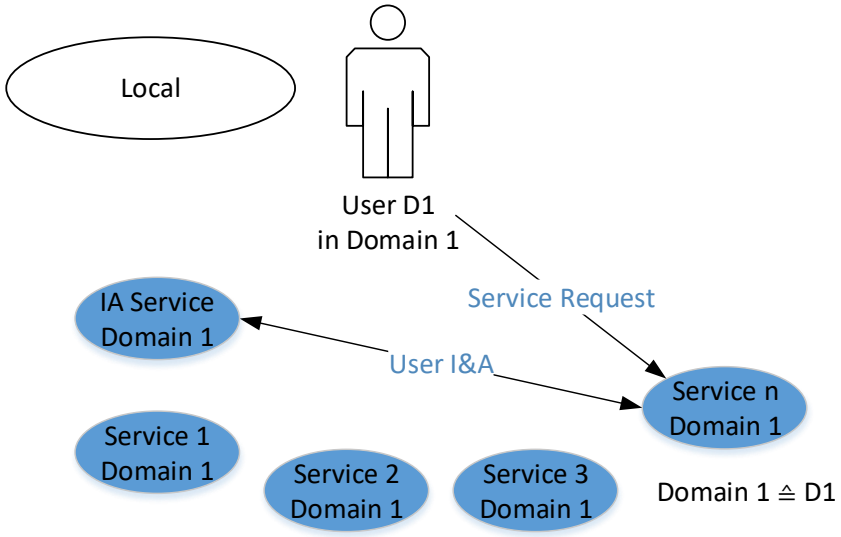

Figure 6. Local authentication (inside one domain).

External Authentication (cross domain):

The service (S) provider receives the service-request and will perform the identification and authentication processes in a centralized or decentralized manner, contacting an IA service (TTP) of an external domain (see Figure 5).

The concept of external authentication is often named delegated. As depicted in Figure 7, the services in domain n to be accessed by a user of domain 1 will contact an instance, e.g., called IA service domain 1, of the external domain 1 for performing the IA of the user of domain 1.

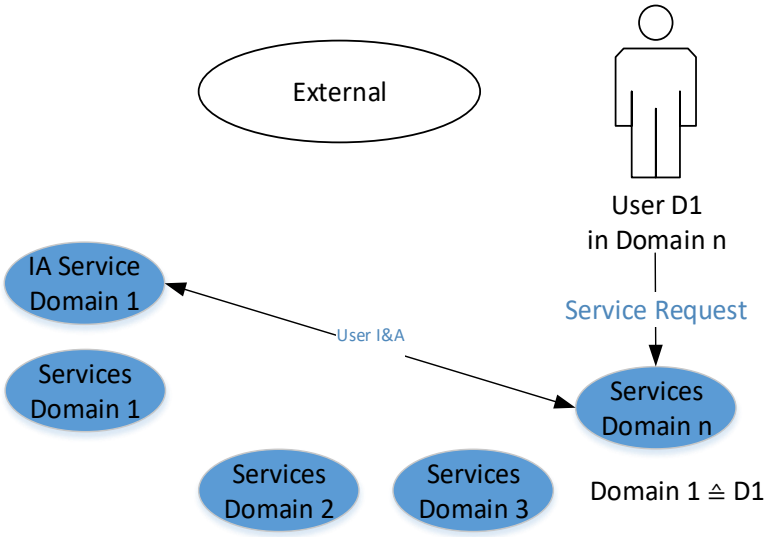

Figure 7. external authentication (cross domain)

Single Sign On (SSO) for local and external authentication:

Regardless if the user passes successfully the IA service in the local or external context SSO is determined as follows:

SSO is defined as the possibility of a user to access continuously after passing the IA service (successful authentication) for a period t one or more services in the domain(s) for that the initial authentication was performed. The validity period t and SSO domain together constitute the Auth-Result presented in Table 3 in §3.3.5. The Table 3 is a further instrument for the auditor to elicit the environment to be analyzed.

In Table 3, despite if the IA process is realized locally or externally or with centralized- and decentralized-IA, different combinations with possible realizations of IA process, therefore, as one server (unit/threat) (IA) service or two server (units/threats) (I)-(A) service, are presented.

								Authentication in		Auth-Result	
IA-Service		Centralized		decentralized				Security Domain		Validity	SSO
		A1	A2	B1	B2	C1	C2	Local Domain (LD)	External Domain (ED)	Time	Domain(s)
One component	(IA)	(IA)		(IA)	(IA)						
Two component	(I)-(A)		(I)-(A)			(I)-(A)	(I)-(A)				
Auth-Result-to-S	S	S	S	S			S				
Auth-Result-to-U	U				U	U					

Table 3 Authentication-Results in the context of centralized- and decentralized- IA process, trust boundaries and SSO.

In Table 3 the rows *authentication in security domain* including the options *Local Domain (LD)* and *External Domain (ED)* and *Auth-Result* including the options *Validity* and *SSO* specially expands the possibility to note more precise the details of the real system.

MUTUAL AUTHENTICATION AND SECURE CHANNEL

Mutual authentication in Table 2 is related with the underlying communication channel as for example *https* or *TLS layer* used between server and client independent of the possible mutual authentication at the level of user identification and authentication. Mutual authentication is according to [23] "When both the client and the server must be authenticated, the process is known as mutual authentication ". The server identify himself with a certificate towards the client and if required by the server the client can be requested to authenticate himself towards the server with an own client certificate. Therefore, mutual authentication on communication channel level is of interest for fulfilling security requirements but lower the possibility of the user to maintain his privacy, e.g., it can be possible to determine easier if a user is accessing independently from the IP address from the same client device. Client certificates could belong to the operating system or application, e.g., browser for surfing environment and, therefore, the corresponding store can vary and reveal more information as intended about the changing user environment.

Secure channel communication, e.g., https and TLS layer [22] are for granting the confidentiality on the communication channel, therefore, observers cannot access the encrypted content in the communication.

Mutual authentication and *secure channel communication* nowadays has become as pointed out in Table 2 The indispensable from the security point of view but can contribute to compromise user privacy.

3.4. Extension of LINDDUN Framework

In §3.4, we apply the PROBLEM SPACE of the LINDDUN framework (see §2.2) to the previously developed DFD-based IA modelling framework. The mapping of LINDDUN privacy Threats to the IA DFD model is specified in §3.4.1 and the extension of the LINDDUN trust boundary concept and application to IA DFD is shown in §3.4.2.

3.4.1. LINDDUN Privacy Threats Mapping to DFD IA Modelling Framework

The LINDDUN privacy threats [1] and related privacy properties are shown in Table 4, which is borrowed (but drawn by our self) from the LINDDUN framework to explain the terminology definition presented by their authors and used in this present work.

	Privacy properties	Privacy threats
HARD	Unlinkability	**L**inkability
	Anonymity & Pseudonymity	**I**dentifiability
	Plausible deniability	**N**on-repudiation
	Undetectability& Unobservability	**D**etectability
	Confidentiality	**D**isclosure of information
SOFT	Content awarness	content **U**nawarness
	Policy and sonsent compliance	policy and consent **N**oncompliance

Table 4. In the LINDDUN framework [1] privacy properties and the corresponding privacy threats are categorized as: hard privacy and soft privacy.

The LINDDUN framework differentiates (as shown in Table 4) between hard privacy as data minimization, and soft privacy where the data controller (entity getting user information) getting the information (should) honestly preserve the data privacy as agreed.

The service-centric topology view of Figure 3 is used to map the DFD Elements to LINDDUN privacy threats and, therefore, obtain Table 5 considering that the pure IA process could be implemented centralized as one component (IA) one server (unit/threat) or decentralized as two servers (as two units/threats) (I)-(A). Table 5 can be used as template to determine the susceptible LINDDUN Privacy Threats of the system during the analysis, e.g., as done in the proof of concept scenarios in §4.

DFD Elements of the Identification and Authentication model				Mapping LINDDUN privacy threats to DFD elements of the Identification and Authentication model						
		I-A on two server								
		IA on one server								
				L	**I**	**N**	**D**	**D**	**U**	**N**
Entity			L	X	X				X	
	User	U		X	X				X	
Process			L	X	X	X	X	X		X
	Identification (I)	I-P		X	X	X	X	X		X
	Authentication (A)	A-P		X	X	X	X	X		X
	Service Provision (S)	Service-P		X	X	X	X	X		X
	Identifi-Authent (IA)	IA-P		X	X	X	X	X		X
Data Store			L	X	X	X	X	X		X
	User Data-/Info-Base	U-DB		X	X	X	X	X		X
	Identification Database	I-DB		X	X	X	X	X		X
	Authentication Database	A-DB		X	X	X	X	X		X
	Identifi-Authent Database	IA-DB		X	X	X	X	X		X
	Service Provision Database	Service-DB		X	X	X	X	X		X
Data Flow			L	X	X	X	X	X		X
	User data stream	with {U-DB, I-P, A-P, Service-P}								
		U- I-P		X	X	X	X	X		X
		U- A-P		X	X	X	X	X		X
		U- IA-P		X	X	X	X	X		X
		U- Service-P		X	X	X	X	X		X
		U- U-DB		X	X	X	X	X		X
	Service data stream	with { Service-DB, U, I-P, A-P}								
		Service-P U		X	X	X	X	X		X
		Service-P I-P		X	X	X	X	X		X
		Service-P A-P		X	X	X	X	X		X
		Service-P IA-P		X	X	X	X	X		X
		Service-P Service-DB		X	X	X	X	X		X
	Identification data stream	with {I-DB, U, A-P, Service-P}								
		I-P U		X	X	X	X	X		X
		I-P A-P		X	X	X	X	X		X
		I-P Service-P		X	X	X	X	X		X
		I-P I-DB		X	X	X	X	X		X
	Authentication data stream	with {A-DB, U, I-P, Service-P}								
		A-P U		X	X	X	X	X		X
		A-P I-P		X	X	X	X	X		X
		A-P Service-P		X	X	X	X	X		X
		A-P A-DB		X	X	X	X	X		X
	Identifi-Authent data stream	with {IA-DB, U, IA-P, Service-P}								
		IA-P U		X	X	X	X	X		X
		IA-P Service-P		X	X	X	X	X		X
		IA-P IA-DB		X	X	X	X	X		X

Table 5: DFD elements of IA modelling framework mapping to LINDDUN privacy threats distinguishing (IA) and (I)-(A).

Table 5 will be the pattern (template) to be used when applying LINDDUN for IA process analysis despite if it is realized on one or on different (two or more) servers (units/threats), e.g., (IA)-P stands for identification and authentication on one server and (I)-P and (A)-P are identification and authentication on two (or more) servers. We highlighted in different shadows of grey IA components combinations that usually will be considered together or disregarded together depending on the realization, therefore, they are mutually exclusive.

3.4.2. Trust Boundary Concept Extension and Application to IA Data Flow Diagram

In the further development the term trust boundary (in LINDDUN [1] called trust boundaries/change of privileges) will be employed and will be extended for IA process. A description of how the trust boundary should be considered in the required interaction of the user and the components of the IA model environment will be given and is illustrated referring to Figure 3. Trust boundaries are illustrated by broken closed lines imbedding inside the components or entities trusting each other and will imply that the connecting data flow arrow between two components are not crossed by any trust boundary.

The smallest unit surrounded completely by a trust boundary comprises a component or entity and the accompanying database/information storage, so that the communication between these two parties is considered trustworthy. The database (information store) of the components and entity will be detailed in a latter step together with the considered IA methods.

The requirements of the possible realizations of IA systems result in the necessity to concretize the trust types to apply, since the trust boundaries delimit changes of competence and the possibility to take influence in the further handling of user and communication information.

We introduce three concepts of trust: The first is the *Exclusive-Trust*, the second is the *Non-Exclusive-Trust* and the third is *Enclosed-Exclusive-Trust*. The terminology is applied according to the DFD introduced in §3.3.3 and in addition brackets "(", ")" and "[", "]" are used to depict which components are inside one trust boundary and to distinguish different overlapping trust boundaries. Furthermore, the three concepts of trust are defined and applied to the IA DFD presented in §3.3.3. Additionally for *Exclusive-Trust Trust Boundary DFD-(U)(S)(I)(A)* Figure 8 and for *Non-Exclusive-Trust/Overlapping Trust Boundary [U ({I A}] S)* Figure 9 are exemplarily presented; of course for the other trust concepts analogous figures could be derived, too.

EXCLUSIVE-TRUST TRUST BOUNDARY

Definition of "Exclusive-Trust Trust Boundary"

Entity, components and data flows grouped together are only imbedded inside a single trust boundary, therefore, there are no overlapping trust boundaries.

3.4.2.1 DFD-(U) (S) (I) (A)

Each component only trusts his own database/information store imbedded by the broken line for trust boundary including the accompanying component and there is no further trust between the other components (See Figure 8).

One example can be a service delegating the identification to an I service and this one involves an A service to perform the authentication and afterwards the authorization confirmation could be provided by the I service or the A service.

3.4.2.2 DFD-(U) (S) (I A)

Each component trusts his own database/information store imbedded by the broken line for trust boundary including the component. The identification (I)- and authentication (A)-service are inside one broken trust boundary line, therefore, these are the only components trusting each other, thus additionally to Figure 8 one more broken trust boundary line including the I- and A-service would be added to the DFD.

One example can be a service delegating the IA Process to an external IA service (e.g., LDAP, RADIUS) located outside the own domain of responsibility, e.g., authentication in the environment of EDUROAM[3] access at universities. Another example could be that of a faculty service offered at a University and the service server contacts an identification and authentication service offered by the computation center inside the local University campus domain.

[3] user roaming in the education and research area, www.eduroam.org

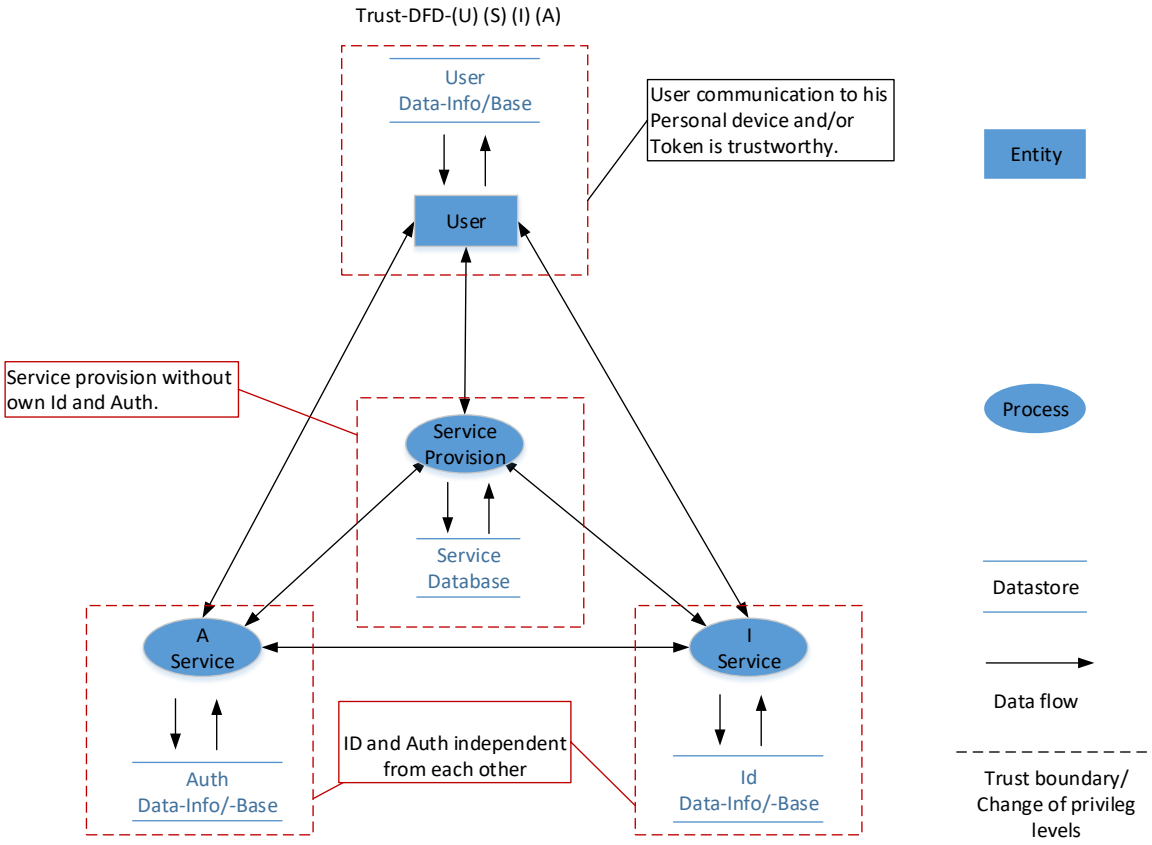

Figure 8. Exclusive Trust-DFD-(U) (S) (I) (A).

3.4.2.3 DFD-(U) (S I A)

Each component trusts his own database/information store imbedded by the broken line for trust boundary including the component. The service provision (S), the identification (I)- and authentication (A)-service are inside one broken trust boundary line, therefore, these are the components trusting each other, thus additionally to Figure 8 one more broken trust boundary line including the S-, I- and A-service would be added to the DFD.

One example can be a service having an own IA service, e.g., a company applying LDAP and authenticates the users using his own user DB.

3.4.2.4 DFD-(U S I A)

Each component trusts his own database/information store imbedded by the broken line for trust boundary including the component. The service provision (S), the identification (I)- and authentication (A)-service and user (U) are all inside one broken trust boundary line, therefore, these components and user trust each other.

This constellation could be an environment where the user uses all hardware provided by one operator, e.g., an employee using a computer (without any other physical access possibility, despite the keyboard and mouse) inside the company with a company account; the computer could be a fix PC (specially hardened and) only configurable by the company system administrator. This constellation would require an "hermetic" isolation towards the outer "world" of all the domain communication and is depreciated, because nowadays it is not a realistic constellation.

NON-EXCLUSIVE-TRUST/OVERLAPPING TRUST BOUNDARY [U ({I A}] S)

Definition "Non-Exclusive-Trust/Overlapping Trust Boundary"

Entity, components and data flows grouped together can be imbedded inside several overlapping trust boundaries.

One common example for the Non-Exclusive-Trust/Overlapping trust boundary concept is that of a user possessing a trusted third party issued IA method set who presents it to a service provider

that on his part is trusting the same trusted third party issuing the IA method set of the user. See Figure 9.

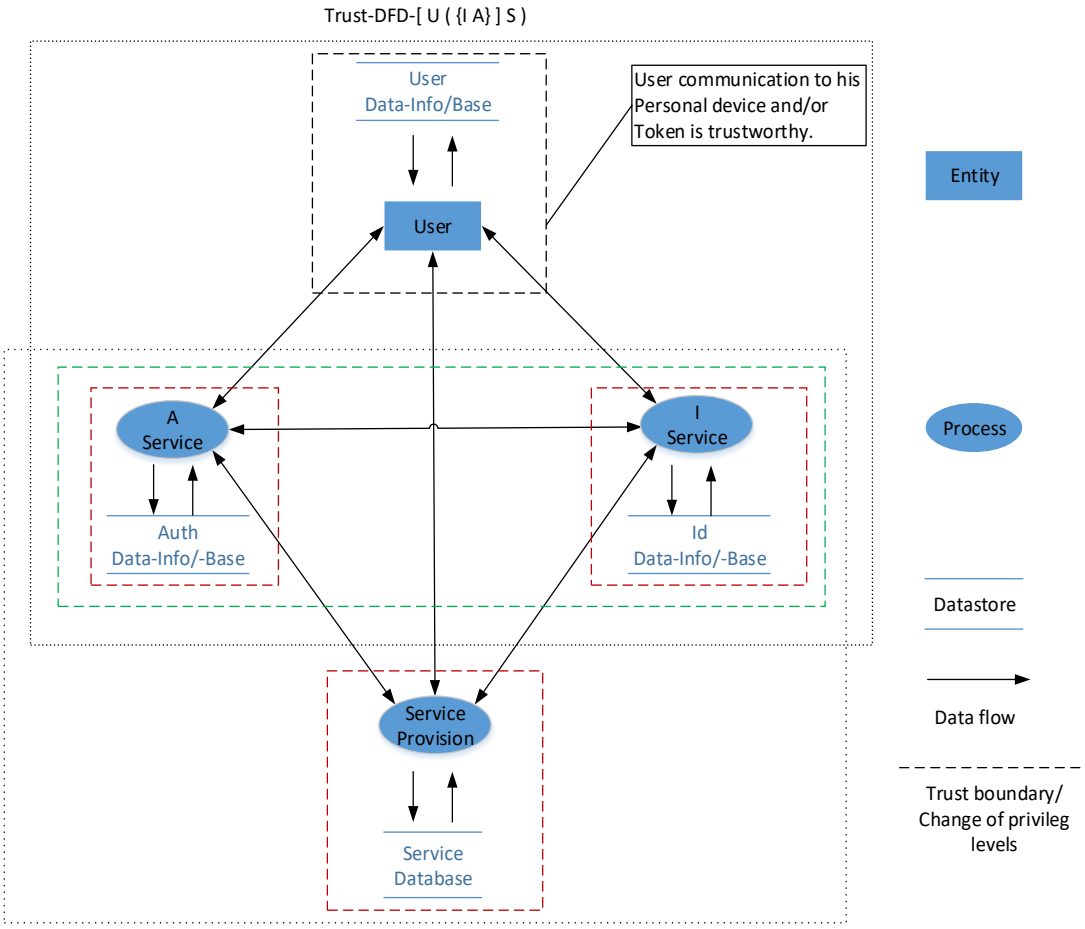

Figure 9. Non-Exclusive-Trust/Overlapping trust boundary [U ({I A}] S).

One concrete example can be a trusted third party issuing, e.g., an electronic ID (eID) (e.g., national identity (smart)card, etc.) and providing the necessary infrastructure for offering the identification and authentication service, too. One realization could be, e.g., authentication as a service based on an external TTP system that has the trust of the service provider company and the trust of the user possessing an eID issued by this TTP.

ENCLOSED-EXCLUSIVE-TRUST TRUST BOUNDARY: DFD- U [S (I A)]

Definition "Enclosed-Exclusive-Trust trust boundary"

Entity, components and data flows grouped together are imbedded inside a single trust boundary (Exclusive-Trust) and a further surrounding outer trust boundary (Enclosed-Trust) encloses such a group and further individual elements, without an overlap of the existing trust boundaries.

This constellation could be the trust concept 3.4.2.2 DFD-(U) (S) (I A) replenished with one additional broken line for trust that imbeds the S-, I- and A-Service.

3.5. Procedure (Instructions) to Apply Enhanced LINDDUN Step 1 and Step 2 for Analyzing IA Modelling Framework-Based Systems

The auditor before proceeding with the present section should first pick up from §3.2 the use case depicted in Figure 2.

LINDDUN STEP 1: DEFINE DFD

0. Replenish the tables (see §3.3.2):
   Table 1: Identity Presentation Methods
   Table 2: Authentication methods, combinations of I-methods and A-methods
1. DFD (introduction in §3.3.3)
2. Process Phases (see §3.3.4)
   Consider DFD in context of sub-phases, see Figure 4, and
   Categorize the IA process of your system.
   Note down in Table 1 and Table 2 the phases when the attributes are provided

3.1 a) Is your I A system (see §3.3.5):
   centralized U->S
   Or
   decentralized U->S and U-> I A
   b) verify if your system uses (IA) on one server or (I)-(A) on two servers
   c) determine if the S and I A are in one or two domains

} Use *Figure 5* to categorize

3.2 Using Figure 5 and Table 3 is for determining which of the constellation from A1 to C2 could be applicable to your system (see §3.3.5):
   which combinations {A1, A2, B1, B2, C1, C2} describes the system
   -> {A1, B1, B2} for (IA) on one server
   -> {A2, C1, C2} for (I)-(A) on two servers

   determining if the system is centralized or decentralized
   -> {A1} on one server for centralized or
   -> {B1, B2} on one server for decentralized

   -> {A2} on two servers for centralized or
   -> {C1, C2} on two servers for decentralized

3.3 -> draw the DFD for the analyzed system considering as guide §3.4.2 with the accompanying figures.

LINDDUN STEP 2: MAP PRIVACY THREATS TO DFD ELEMENTS

4. with the details of step 3.3 above in LINDDUN step 1 and Table 5 choose whether to consider the cells for (IA) on one server or the cells for (I) (A) on two servers.

5. Reduce the table you selected in the previous step by disregarding (removing) the lines not corresponding to your choice (real system).

6. Is your IA realized as Local authentication, see Figure 6
   External authentication, see Figure 7
   This step is to determine further trust boundary and apply it to the resulting table in step 5 above.

At this point the auditor finished step 2 of LINDDUN framework depicted in Figure 12 (see Appendix) applying the contributions of the present paper and must now continue with step 3 of LINDDUN framework [1].

4. EVALUATION

In this section, we conduct an evaluation in conjunction with a proof of concept. Recall that the central contribution of the present paper is the creation of a tool set and procedure description of how to model and analyze a system for identification and authentication of user identity attributes. The presented identification and authentication methods (see §3.3) make possible numerous combinations. For this reason, only a limited selection could be presented exemplarily for describing the application of the procedure summarized in §3.5. The proof of concept scenario, see §4.1, considers on the one hand, a user login (authentication) with user name and password and, on the other hand, the user authentication with a pin protected smartcard; in both cases towards the University Library Service. In §4.2 the application of the procedure summarized in §3.5 is presented. In §4.3 we discuss the application of the proposed framework to the proof of concept scenario of §4.1.

4.1. Proof of Concept Scenario

A state University with the accompanying information technology infrastructure (IT) including all services usually provided to members is chosen for the proof of concept of the developed IA modelling framework and enhancement of the steps 1 and 2 of the LINNDDUN framework. The scenario is based on a user, member of the state University, having access to diverse University IT infrastructure services; for the proof of concept a user accesses from outside to the library service of the University on the one hand, to reserve, e.g., a printed book, on the other hand, e.g., to pay the lending fee.

The state University issues smart cards including chip-based authentication using a personal identification number (PIN), chip-based cash, a barcode for the library with associated password, associated University user account (username/password) and printed on the smartcard are the user identity number, user name and surname, photo and validity of the smartcard.

In the context of the state University for a user there are plenty of constellations conceivable that require the user identification and authentication, e.g., VPN to the University campus, login at server of different faculties and usage of trust relationship through EDUROAM[4]. The user of the state University has for user verification (login) purpose at least two possibilities on the one hand, the username/password combination and, on the other hand, a PIN protected smartcard with an access protected public key infrastructure (PKI) private key. The selected proof of concept depicts on the one hand, a username and password-based login to the University Library Service and, on the other hand, a smart card-based authentication for electronic payment of lending fee, see Figure 10.

[4] user roaming in the education and research area, www.eduroam.org

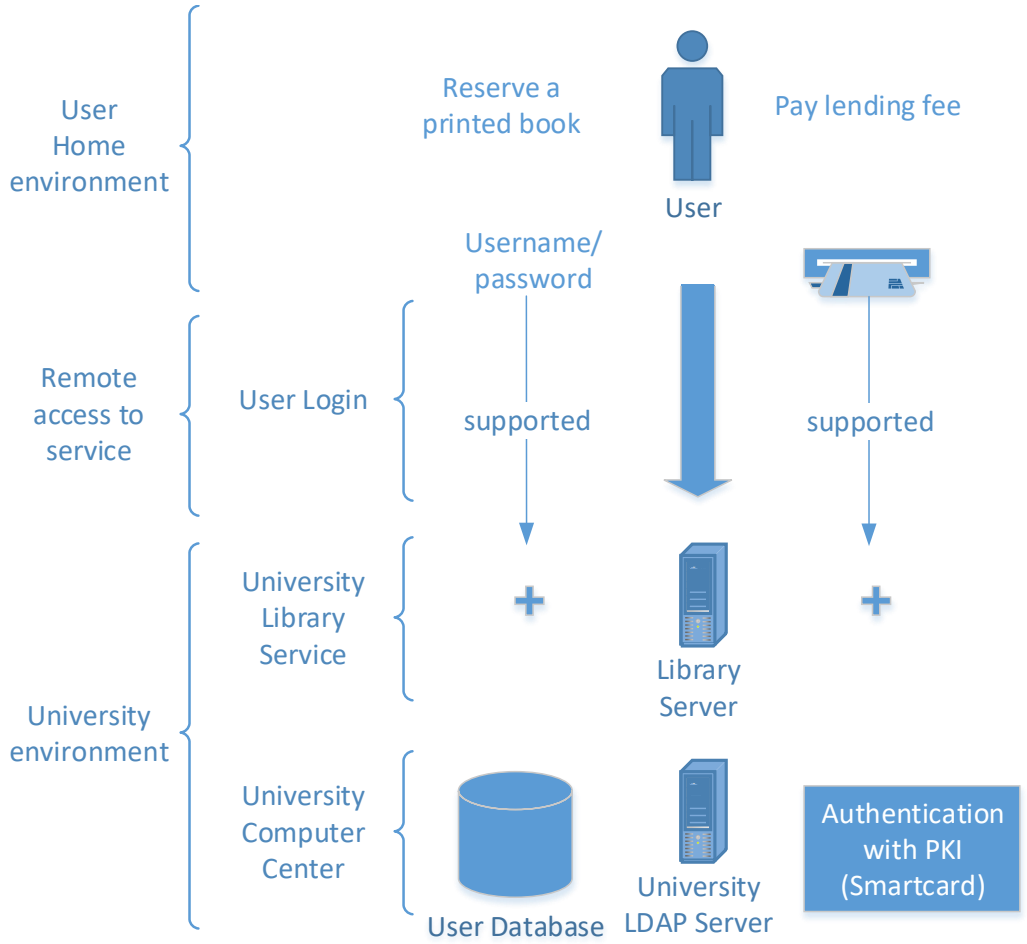

Figure 10. Proof of concept: user reserve a book or pay lending fee at University Library server.

4.2. Application of the Proposed Framework

In §4.2 we apply the contributions of §3 to the two variants of the proof of concept scenario depicted in §4.1. The first variant uses a username/password, the second variant a smartcard-based user identification and authentication.

USERNAME AND PASSWORD BASED LOGIN TO THE UNIVERSITY LIBRARY SERVICE

The scenario in Figure 10 is scrutinized based on §3. First consider the use case depiction in Figure 2 from §3.2 for visualizing the service access process.

According to §3.3.1 and §3.3.2 the identity, the attribute username, presentation is done manually and is authenticable. The user has no other information storage than his memory. From §3.3.3 the service centric DFD representation from Figure 3 will be taken. Following §3.3.4 phase 2, the identification, is done through the library server, therefore, centralized by contacting the University LDAP server and phase 3, the authentication, is done centralized, too. It depends on the realization of IA, if it is on one or two servers, therefore, if the LDAP has an own user database or contacts an external one for performing the authentication. The present proof of concept assumes an LDAP with own user database, thus one server (IA). Considering Figure 5 and Figure 6 in 3.3.5 the verification of the legitimate usage of the username is determined as local authentication. The user is not giving further attributes.

Table 5 in §3.4.1 presents the global table of DFD elements IA mapped to LINDDUN privacy threats for (IA) and (I) (A). Based on §3.4.2 the scenario presents the Exclusive-Trust (U) (S) (I A) property and obey to the DFD example 3.4.2.2 in §3.4.2. As commented in §3.3.5, the recommended secure channel communication between user client and server is given accessing the University servers by using https. Mutual authentication between the user client and the server is not used and no SSO with the done authentication is offered.

The considerations done results in the left light grey components of the DFD shown in Figure 11 and to consider from Table 5 the cells for (IA) on one server. The auditor at this point of the selected proof of concept variation would have to continue with step 3 of LINDDUN framework [1].

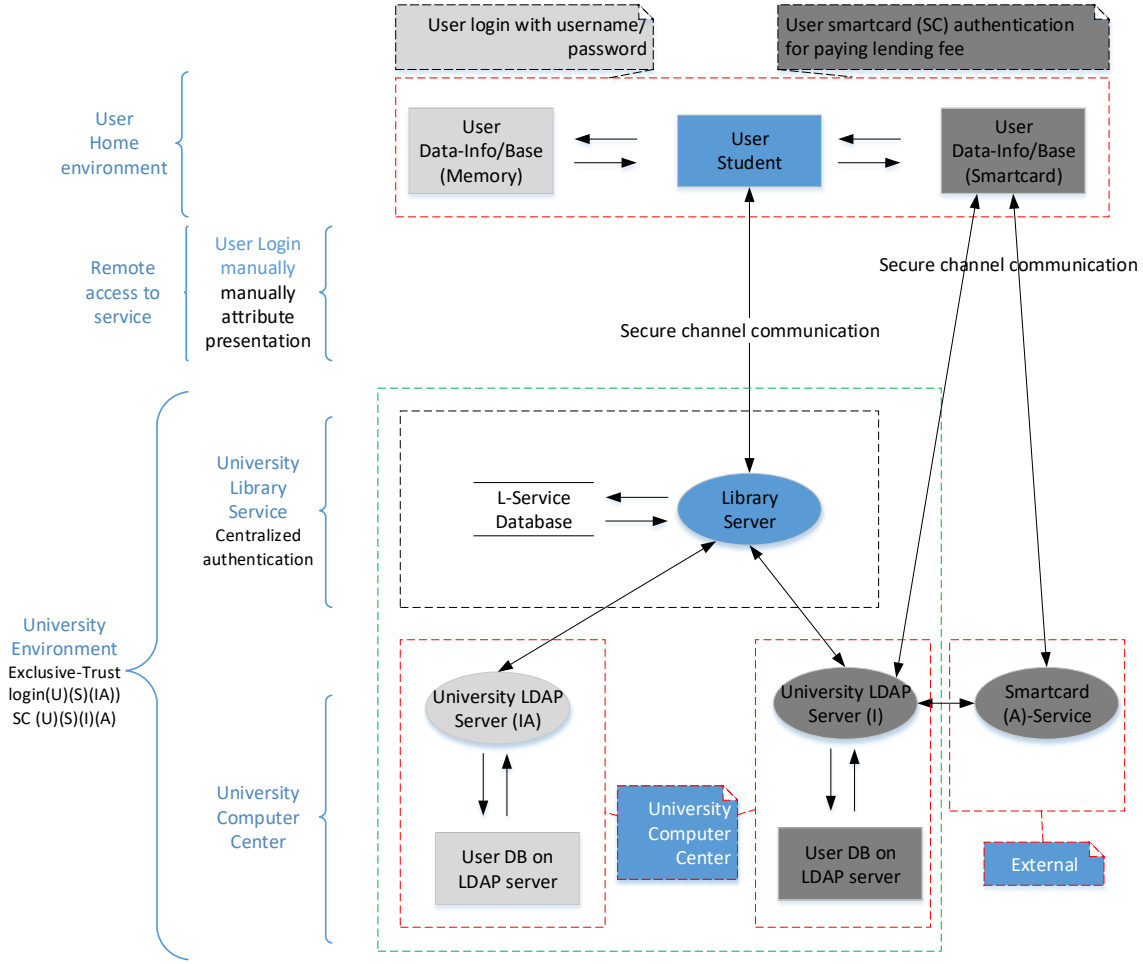

Figure 11. DFD for proof of concept: User/Password login and smartcard-based authentication.

SMART-CARD-BASED AUTHENTICATION FOR ELECTRONIC PAYMENT OF LENDING FEE AT UNIVERSITY LIBRARY SERVICE

Consider again as in the beginning of §4.2 the use case in Figure 2 from §3.2 for visualizing the whole service access process.

According to §3.3.1 and §3.3.2 the identity, the attribute username, presentation now is done electronically and is authenticable. The user brings along the information storage in the smartcard. From § 3.3.3 the service centric DFD representation from Figure 3 will be taken. Following §3.3.4 phase 2, the identification, is done through the library server contacting the University LDAP server and phase 3, the authentication, is done directly between the user device and the external smartcard-authentication server, therefore, decentralized. Assuming for the present proof of concept variation that the identification is done by the University LDAP server and the authentication is delegated to the external smartcard authentication server the present subsystem is based on two servers, (I) (A). Considering Figure 5 and Figure 7 in 3.3.5 the verification of the legitimate usage of the username is determined as external authentication. The user is not giving further attributes.

Table 5 in §3.4.1 presents the global table of DFD elements IA mapped to LINDDUN privacy threats for (IA) and (I) (A). Based on §3.4.2 the scenario presents the Exclusive-Trust (U) (S) (I) (A) property and obey to the DFD example 3.4.2.1 in §3.4.2. As commented in §3.3.5, the recommended secure channel communication between user client smartcard reader and smartcard authentication server is given by using TLS. Mutual authentication between the user client smartcard

reader and the smartcard authentication server is used and no SSO with the done authentication is offered.

The considerations done result in the right dark grey components of the DFD shown in Figure 11 and to consider from Table 5 the cells for (I) (A) on two servers. The auditor at this point of the selected proof of concept variation would have to continue with step 3 of LINDDUN framework [1].

4.3. Discussion

This section discusses several aspects of our contributions, particularly with regard to the application of the two use cases of user authentication described in the previous subsections.

In **§3.2**, Figure 2 offers a high-level entry point into the system analysis for the general use case of *user demanding service access*. The presented subdivision facilitates the auditor a first assignment of parts of their system to the general use case. At this stage, for both proof of concept variants (two uses cases of user authentication) in **§4.2**, we would like to stress the different specificity of (i) the user login with username and password, and (ii) user authentication with smartcard. These two variants could be regarded as *centralized user verification* or *decentralized user verification*.

On the other hand, **§3.3** provides the auditor with a tool set to break down the user verification process in their system. §3.3.1 and §3.2.2 facilitate the auditor to itemize their used identification and authentication methods with Tables 1 and 2. For both uses cases of user authentication, the core findings are:

- In the first use case with username and password login, no additional user data base (repository) is present. Phase 2, the identification, and phase 3, the authentication, are conducted on the same server including a user DB. The verification of the legitimate usage of the username is determined as local authentication and is a centralized verification based on one server (IA).
- Per contra, in the second use case with smartcard authentication, an additional user data base (repository) is present. Phase 2, the identification, is carried out through one server and phase 3, the authentication, is performed directly between the user device and a second external smartcard-authentication server. The verification of the legitimate usage of the username is given by an external authentication and is a decentralized verification based on two server (I) (A).

A comparison of both use cases shows that the results can vary largely depending on the assumptions made. Concerning the centralized and decentralized user verification on the one hand, and on the other the one (IA) or two (I)(A) server solution for identification and authentication, we notice that the results could be switched. What this means is that means that the first use case with username and password login could be conducted in a decentralized manner, and therefore on two (I)(A) servers, e.g., using a separate user database server. Consequently, the second use case with smartcard authentication could take place in a centralized environment and therefore carried out in one (IA) server. In this case, a smartcard authentication service would be integrated in the University LDAP server.

To our best knowledge, for the first time a set of tables of user identification and authentication methods are introduced. Likewise, our work is the first to introduce, in combination with the aforementioned tables, a user data base store (repository) to the DFD representation. Furthermore, we have extended the verification process representation and trust boundary concept. A remaining limitation is the lack of further adaptation of the LINDDUN framework for more environments.

The relevance of our work also lies in the practical applicability of the proposed solution. In particular, auditors can easily map the LINDDUN privacy threats to the DFD IA model created in §3.3. More specifically, Table 5 in §3.4.1 presents the IA DFD elements mapped to the LINDDUN privacy threats for one-server (IA) and two-server (I) (A) solutions. In this manner, the most suitable trust boundary concept can be selected. For both proof of concepts, the most important remarks are described next:

- In the first use case (with username and password login), from Table 5 the cells for (IA) on one server are considered and the scenario presents the Exclusive-Trust (U) (S)(I A) property.
- In the second use case (with smartcard authentication), from Table 5 the cells for (I)(A) on two servers are contemplated and the scenario shows the Exclusive-Trust (U) (S) (I) (A) property.

The combination of both remarks highlight that the auditor is supported in eliciting trust boundaries and that they should be aware of the fact that the cell groups for one server (IA) and server (I)(A) solution in Table 5 are mutually exclusive.

That being said, the most important aspect of our proposal is the adaptation of the LINDDUN framework to allow identification and authentication processes. The extension and application of the trust boundary concept to LINDDUN are undoubtedly a major advance in the systematic modeling of privacy threats in the context of those two processes. One of the limitations of such an adaptation, however, is that our solution is constrained to the assumptions made after step 2, and that the extension is obviously tailored for IA processes. We elaborate further in the concluding section §4.4 that one important challenge is to extend the application of LINDDUN to more environments.

Finally, we would like to emphasize that, with the user data repository (user data-info/-base), we proposed a more precise modelling of the location of attributes and authentication factors. This permits analysing more specific privacy threats.

4.4. Related Work

This section reviews the state of the art relevant to this work and emphasizes the value and novelty of our contributions. We proceed first by stressing the relevance of LINDDUN, the privacy threat analysis framework we build upon.

The usage of LINDDUN is predominant in the context of threat modelling methodologies. We would like to emphasize, however, that the focus is largely on security threat modelling, where LINDDUN is mentioned as one systematic modelling framework focusing on privacy threat analysis. This is specifically stated in [24], where a systematic literature review of threat modelling is conducted on the basis of more than one hundred works. In the cited paper, the authors contemplate that LINDDUN can address security threats in the environment of software application with focus on privacy.

The usage of LINDDUN is also suggested in [25] as central threat modelling methodology in the context of privacy by design, to directly achieve privacy guaranteeing systems. In that paper, the authors utilize LINDDUN as the core threat modelling methodology and propose the usage of LINDDUN in an iterative way.

A further recent paper [26] gives a summary of available methods for threat modelling coming along with 12 threat modelling methods that tackle most security services. Particularly only for LINDDUN, the authors emphasize its relevance on privacy. Most of the proposed threat modelling methods are based on a data flow diagram (DFD) to describe the system to be analyzed. In [27], the threat methodology STRIDE and LINDDUN are shown to be susceptible to certain threat explosion vulnerabilities, which the authors attempt to mitigate by first applying the threat methodology PASTA and afterwards LINDDUN. The authors claim that “PASTA also mitigates the threat explosion weaknesses of STRIDE and LINDDUN by utilizing risk and impact analysis”. In the context of threat explosion, [28] proposes a refinement of LINDDUN to mitigate its vulnerabilities.

We agree with the authors of [27] to use LINDDUN for threat modelling with a focus on privacy. However, we do not completely agree to previously apply PASTA to mitigate the threat explosion weaknesses of LINDDUN. We believe that, at that stage, possible relevant threats might be disregarded. In the cited paper, the authors evaluate LINDDUN based on the core categorizations “Strengths and weaknesses and Tailorability” and conclude that its level of maturity is high enough and that no consistent results could be achieved. As for tailorability, [5] states that “since none of these methods were designed with a specific type of system in mind, all may be applied to any kind of system.”

On the one hand, we agree with [5] that LINDDUN has achieved a high level of maturity, and on the other, we acknowledge the previously mentioned weaknesses and limitations as far as tailorability is concerned. In this present work, we aimed to achieve consistent results to increase the reproducibility of the application of LINDDUN, e.g., a more detailed and systematic approach to create the DFD of a system. One further contribution of the paper is a step-by-step guide to be used by analysts. This last step additionally guarantees a higher reproducibility, since the guided DFD creation depends less on the knowledge of the analyst.

Our focus on LINDDUN, and therefore the relevance of our contributions, are then justified by the extensive literature succinctly reviewed above. The adaptation of the LINDDUN framework for the specific services of identification and authentication may not need justification. Identification and authentication are essential and nearly ubiquitous security services nowadays.

Now we discuss different aspects related to privacy in the context of identification and authentication.

In privacy enhanced authentication systems (e.g., attribute based credentials [29]), we find systematic analyses of privacy threats based on system-related weaknesses. The authors of the cited work give an example: “Even though an attribute may be anonymous, the ‘leaking’ of information from another level in the infrastructure, such as an IP address, could make the attribute pseudonymous or even fully identifying…”. A further example of privacy threats in IA is given in [30], where a privacy vulnerability of OpenID was found.

The vulnerabilities mentioned in [29] and [30] can be analyzed systematically with our extended LINDDUN methodology, which we enhance to contemplate IA process modelling components. Our work supports the systematic development of privacy-by-default fulfilling systems that guarantee a higher reproducibility based on our LINDDUN methodology.

In the review of LINDDUN-related papers in [31], the authors propose a further improvement of LINDDUN consisting in the so-called Interaction-based Privacy Threat Elicitation which, as the authors acknowledge, comes along with threat explosion too. Similar to this approach, we have introduced independently the subdivision Process Phases P1 -P2 and Sub-Phases in the context of systematically describing more detailed identification and authentication processes. Our subdivision of IA processes is to perform a reproducible, reusable and detailed segregation of the subphases of identification and authentication.

Finally, to stress the novelty and relevance of our versatile contributions to the DFD-based modelling, and for the sake of completeness we would like to briefly comment on [32]. In this paper, the authors mention the Privacy Knowledge for Threat Elicitation, list six different knowledge bases including LINDDUN and assume for all of them a common underlying DFD modelling of the system. From this standpoint, our enhanced DFD modelling methodology could be used across all these so-called knowledge bases for Privacy Knowledge for Threat Elicitation.

5. CONCLUSION

Systematic approaches for PTA are a central pillar for a reliable PIA, but this task is in general not carried out systematically. The LINDDUN framework has become a promising approach for a systematic PTA framework, as we stated in related work.

- Our first main contribution is a novel modelling framework for identification (I) and authentication (A) process that is usable with LINDDUN framework [1, 14–18]. To our best knowledge, the proposed novel DFD based I and A modelling framework provides for the first time a compilation of tables including I and A methods linked with well-known procedures.
- Our second main contribution applies the privacy threat mapping of the LINDDUN framework to our Data Flow Diagram (DFD) based IA modelling framework. More specifically, we have extended the LINDDUN trust boundary concept to the developed DFD-based IA modelling framework. We have also adapted the step 1 and step 2 of the LINDDUN framework to be usable for PTA of I and A process. This contribution facilitates a generic mapping of the DFD elements of the IA modelling framework to LINDDUN privacy threats. It

distinguishes the realization of the process as one component (IA) and as two component (I)-(A) and is furthermore usable, too, as generic template by the auditor.

- The third contribution is the generic UML drawing in Figure 2, see §3.2, that serves as entry point for the auditor for a first categorization of the system.
- The fourth contribution is the compilation of straightforward instructions, see §3.5, that guides the auditor through the application of the contributions of the present paper.
- Finally, with the fifth contribution we have introduced a more detailed modelling of the user data repository called user data-info/-base to the DFD applied in LINDDUN. The user data-info/-base makes possible a more precise representation of the location where user attributes are stored, e.g., in the user memory, smartcard or in the cloud. This is a major further step towards the analysis and realization of user self-determination.

The specific objectives for the design of the PTA framework are described next:

- Rely on a mature and widely used privacy threat analysis framework.
- Satisfy upcoming demands stated in the literature such as privacy by default, adequate reduction of threat explosion weakness, reproducibility and adaptability.
- Capable of being extended to encompass identification and authentication, which are core processes to guarantee trustworthiness.
- Create a DFD-based system modelling method applicable with different privacy knowledge bases for threat elicitation.

As future research, we intend to extend our results to more environments (apart from that of I and A), develop more modelling procedures and hence systematic PTA methodologies focusing on specific user requirements.

ACKNOWLEDGMENT

The authors are thankful for the support through the research project “INRISCO”, ref. TEC2014-54335-C4-1-R, “MAGOS”, TEC2017-84197-C4-3-R, and the project “Sec-MCloud”, ref. TIN2016-80250-R. J. Parra-Arnau is the recipient of a Juan de la Cierva postdoctoral fellowship, IJCI-2016-28239, from the Spanish Ministry of Economy and Competitiveness. J. Parra-Arnau is with the UNESCO Chair in Data Privacy, but the views in this paper are his own and are not necessarily shared by UNESCO.

REFERENCES

[1] Deng, M., Wuyts, K., Scandariato, R., Preneel, B., and Joosen, W. 2010. LINDDUN: A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements.

[2] Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.-H., Le Métayer, D., and Schiffner, S. 2014. *Privacy and Data Protection by Design – from policy to engineering*. *ENISA*.

[3] Kloza, D. 2012. A Privacy Impact Assessment Framework for data protection and privacy rights. Recommendations for a privacy impact assessment framework for the European Union. Microsoft Word - PIAF D3 recommendations v4.2 pr clean.docx.

[4] Wright, D. and Hert, P. d. 2012. PRIVACY IMPACT ASSESSMENT. *Law, Governance and Technology Series,* VOLUME 6.

[5] Oetzel, M. C. and Spiekermann, S. 2013. A systematic methodology for privacy impact assessments: a design science approach. *Eur J Inf Syst* 23, 2, 126–150.

[6] Christopher, G. and Information Commissioners Office. 2014. Conducting privacy impact assessments. code of practice. pia-code-of-practice.

[7] CNIL - French Data protection Authority. É d i t i o n 2 0 1 2. Methodology for Privacy Risk Management - English version. How to implement the Data Protection Act (É d i t i o n 2 0 1 2).

[8] Michael N. Johnstone. 2010. Threat Modelling with Stride and UML. *Originally published in the Proceedings of the 8th Australian Information Security Mangement Conferencee, Edith Cowan University, Perth Western*.
[9] European Commission. 2011. *Privacy and Data Protection Impact Assessment Framework for RFID Applications*. Accessed 1 October 2015.
[10] (BSI) Bundesamt für Sicherheit in der Informationstechnik. 2011. Privacy Impact Assessment Guideline for RFID Applications.
[11] CNIL - Commission Nationale de l'informatique et des libertés. 2015. PIA, METHODOLOGY. PRIVACY IMPACT ASSESSMENT (PIA) Methodology (how to carry out a PIA) (Jun. 2015).
[12] Prasser, F., Kohlmayer, F., Spengler, H., and Kuhn, K. 2017. A scalable and pragmatic method for the safe sharing of high-quality health data. *IEEE journal of biomedical and health informatics*.
[13] Brandizi, M., Melnichuk, O., Bild, R., Kohlmayer, F., Rodriguez-Castro, B., Spengler, H., Kuhn, K. A., Kuchinke, W., Ohmann, C., Mustonen, T., Linden, M., Nyronen, T., Lappalainen, I., Brazma, A., and Sarkans, U. 2017. Orchestrating differential data access for translational research. A pilot implementation. *BMC medical informatics and decision making* 17, 1, 30.
[14] Wuyts, K., Joosen, W., and Scandariato, R. 2014. LIND(D)UN privacy threat tree catalog (Sep. 2014).
[15] Wuyts, K. 2015. Privacy Threats in Software Architectures. PhD (Jan. 2015).
[16] Wuyts, K. and Joosen, W. 2015. LINDDUN privacy threat modelling: a tuorial (Jul. 2015).
[17] LINDDUN - DistriNet Research Group. 2014. *LINDDUN in a nutshell*. https://distrinet.cs.kuleuven.be/software/linddun/linddun.php. Accessed 2 June 2016.
[18] Wuyts, K. 2015. LINDDUN 2.0. Privacy knowledge (tables) (Jul. 2015).
[19] Pfitzmann, A. and Hansen, M. 2010. *A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management*. *Anon_Terminology_v0.34.pdf*.
[20] Modinis IDM Study Team. 2005. Modinis Study on Identity Management in eGovernment. Modinis Workshop Discussion Paper (Nov. 2005).
[21] Menezes, A. J., Vanstone, S. A., and Van Oorschot, Paul C. 1997. *Handbook of applied cryptology*. CRC Press series on discrete mathematics and its applications. CRC, [S.l.].
[22] Eckert, C. 2013. *IT-Sicherheit*. *Konzepte - Verfahren - Protokolle*. 8. aktualisierte und korrigierte Auflage. Oldenbourg, München.
[23] Robert Havighurst. 2007. User Identification and Authentication Concepts. In *Mechanics of User Identification and Authentication: Fundamentals of Identity Management*, D. Todorov, Ed.
[24] Xiong, W. and Lagerström, R. 2019. Threat modeling – A systematic literature review. *Computers & Security* 84, 53–69.
[25] Veseli, F., Olvera, J. S., Pulls, T., and Rannenberg, K. Engineering privacy by design. In *Hung (Hg.) 2019 – The 34th Annual ACM Symposium*, 1475–1483. DOI=10.1145/3297280.3297429.
[26] Shevchenko, N., Chick, T. A., O'Riordan, P., Scanlon, T. P., and Woody, C. 2018. Threat Modeling: A Summary of Available Methods.
[27] Nataliya Shevchenko, Frye, B. R., and Woody, C. 2018. THREAT MODELING FOR CYBER-PHYSICAL SYSTEM-OF-SYSTEMS: METHODS EVALUATION.
[28] Wuyts, K., Van Landuyt, D., Hovsepyan, A., and Joosen, W. Effective and efficient privacy threat modeling through domain refinements. In *Haddad, Computing (Hg.) 2018 – The 33rd Annual ACM Symposium*, 1175–1178. DOI=10.1145/3167132.3167414.
[29] Koning, M., Korenhof, P., Alpár, G., and Hoepman, J.-H. The ABC of ABC. - An Analysis of Attribute-Based Credentials in the Light of Data Protection, Privacy and Identity - 2014.
[30] Urueña, M., Muñoz, A., and Larrabeiti, D. 2014. Analysis of privacy vulnerabilities in single sign-on mechanisms for multimedia websites. *Multimed Tools Appl* 68, 1, 159–176.
[31] Sion, L., Wuyts, K., Yskout, K., Van Landuyt, D., and Joosen, W. 2018. Interaction-Based Privacy Threat Elicitation. In *3rd IEEE European Symposium on Security and Privacy Workshops*. *Proceedings : 24-26 April 2018, London, United Kingdom*. Conference Publishing Services, IEEE Computer Society, Los Alamitos, California, 79–86. DOI=10.1109/EuroSPW.2018.00017.
[32] Wuyts, K., Sion, L., Van Landuyt, D., and Joosen, W. IEEE 2019. Knowledge is Power: Systematic Reuse of Privacy Knowledge for Threat Elicitation (IEEE 2019).

APPENDIX

A STEP-BY-STEP OVERVIEW OF THE LINDDUN FRAMEWORK EXAMPLE

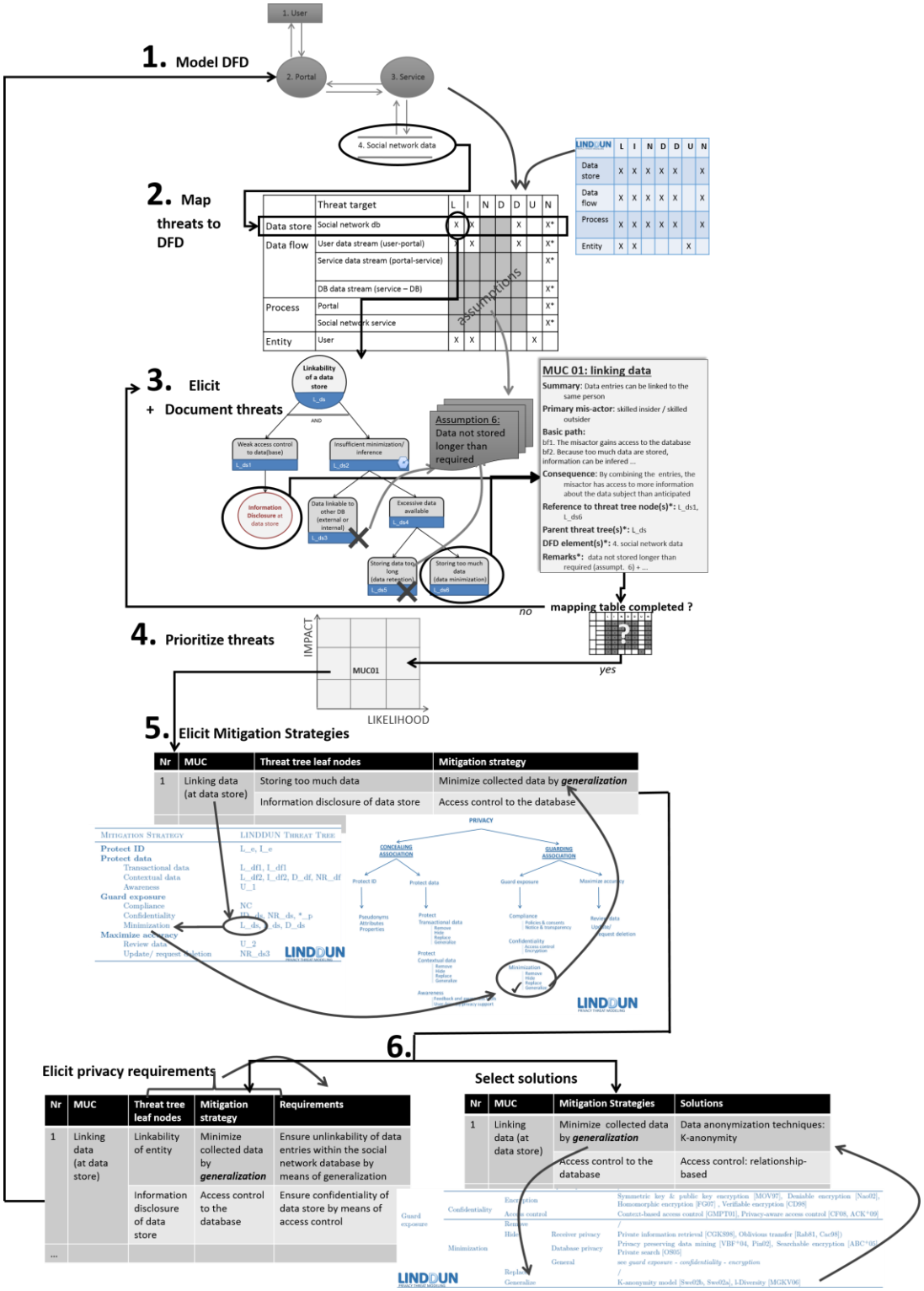

Figure 12. A step-by-step overview of the LINDDUN framework using a simple social network system as running example[5].

[5] https://distrinet.cs.kuleuven.be/software/linddun/linddun.php

ACRONYMS

A	Authentication
BSI	Bund für Sicherheit in der Informationstechnik (Federal Office for Information Security)
CR	Challenge response
DB	Data Base
DFD	Data Flow Diagram
ED	External Domain
EBIOS	Expression des Besoins et Identification des Objectifs de Sécurité – Expression of needs and identification of security objectives
ENISA	European Union Agency for Network and Information Security
I	Identification
IA	Identification and authentication
ID	Identifier
LD	Local Domain
MRZ	Machine Readable Zone
NFC	Near Field Communication
PIAF	A Privacy Impact Assessment Framework for data protection and privacy rights (project name)
PKI	Public Key Infrastructure
PET	Privacy Enhancing Technologies
PIA	Privacy Impact Assessment
PTA	Privacy Threat Analysis
RFID	Radio-Frequency Identification
SC	Smartcard
SSO	Single Sign On
S/SP	Service Provision
STRIDE	An acronym for Spoofing identity, Tampering, Repudiation, Information disclosure, Denial of service and Elevation of privilege
TTP	Trusted Third Party
UML	Universal Markup Language

Table 6: List of acronyms.