GLOBAL IDENTITY NETWORKING OF INDIVIDUALS

A Privacy Policy Framework for the INDI ecosystem

Version 1.0

Project Name	GLOBAL IDENTITY NETWORKING OF INDIVIDUALS
Work Package	WP4: The Privacy Dimension of the INDI Domain
Activity	Activity Title
Editors	Shuzhe Yang, Goethe University Frankfurt Ioannis Krontiris, Goethe University Frankfurt Kai Rannenberg, Goethe University Frankfurt
Date of Initial Creation	
Date of Last Change	23.04.2012
Status	[ ] Draft [ ] Internal Commenting [x] Release
CM Path	...

Further Document Information

Authors	Seda Gürses (KUL) Ioannis Krontiris (GUF) Ahmad Sabouri (GUF) Andreas Pashalidis (KUL) Shuzhe Yang (GUF) Berd Zwattendorfer (TUG)
Base Documents	1) GINI DoW 2) D1.1: The Individualised Digital Identity (INDI) Model: A User-centric Framework of identity management services

Change History

Modification			Affected Chapters	Motivation / Description	Author	State
No.	Date	Version				
1	01.05.2011	0.1	All	Initial creation		Draft
2	09.05.2011	0.2	4	Addition		Draft
3	20.05.2011	0.3	All	Modifications		Draft
4	15.06.2011	0.4	All	Modifications		Draft
5	20.06.2011	0.5	All	Internal Review		Draft
6	29.02.2012	0.6	All	Internal Review		Draft
7	21.03.2012	1.0	All	Final Draft		Draft

Audit and Quality Assurance History

No.	Date	Version	Remarks	Auditor	State
1	22.11.2011	0.6	The Privacy Policies are use case related and not covering all situations.		Draft

Table of Contents

Executive Summary

In order to give meaningful effect to the right to informational self-determination, it is clearly necessary for users to have the possibility of "information self-awareness". Even though its importance is being emphasized more and more by current research in online privacy, at the same time the limitations of currently available tools prohibit their wide adoption and applicability. INdividual Digital Identity (INDI) operators offer consumers the ability to designate "privacy agents" as proxies for exercise of their rights, abstracting the complexity and making online management of identity comprehensive, convenient and secure.

The objective of this deliverable is to formulate a Privacy Policy Framework for the new set of user-centric services within the INDI domain, based on a model of data provisioning where the individuals provide the data on their own and at their discretion. That means, a Privacy Policy Framework has to be developed that defines the core mode of operation for the GLOBAL IDENTITY NETWORKING OF INDIVIDUALS (GINI) services in an abstract manner. This allows for the users to express their preferences rather than having to define the way privacy protection is realized in detail. This requires that the privacy of the user is guaranteed at a lower level, through the functionality of the INDI operators, and the user is involved only to make decisions related to his digital identities and to define the informational balance on what information is made available for what purposes. So we first need to define privacy-related requirements for data handling that should be satisfied by the INDI operators. A necessary input to this analysis is the definition of the components of the architecture (GINI Deliverable D2.1), their functionality and the corresponding information flows within the INDI ecosystem.

Neither the current European legal framework nor the US approach towards private sector self-regulation has been effective so far in the protection of online privacy, particularly with regard to new business models, such as behavioural targeting, user profiling, social networking and location-based services. One reason is that there seems to be too much reliance on *ex post* securing of data rather than on *ex ante* avoidance of privacy risks through conformance to principles like data minimization [FHK+11]. In this deliverable, we focus on the latter approach and examine privacy by design solutions. We carry out a comprehensive and iterative privacy risk and impact assessment within the INDI ecosystem, examine how state-of-the-art privacy technologies can be adopted, and identify technological gaps. Legal requirements and their connection to the technological advances discussed here are presented separately in GINI Deliverable 3.1.

In this deliverable we concentrate on privacy requirement analysis and the development of a Privacy Policy Framework. The requirement analysis is broken down with respect to each INDI architectural entity. Our methodology consists of providing a structured table for each entity, all of them will be later unified to identify interconnections and/or missing elements. This is demonstrated in Section 4.1. Section 5 deals with the Privacy Policy Framework. In the first step, we extract use case related privacy policies based on use cases taken from GINI Deliverable D1.1. After that, we generalize them to high-level privacy policies, which fulfil the requirements from Section 2, and form them to the generic Privacy Policy Framework.

Readers Guide: Section 1 introduces the privacy threats that users face everyday and overviews the technology landscape of protection measures, as well as the reasons why these solutions have not been widely adopted today. Section 2 overviews the privacy requirements of identity management systems in general, as we know them from previous works and under the light of multilateral security. Section 3 focuses specifically on the INDI ecosystem and reviews the basic modules, architectural entities, and flows behind it. Then, Section 4 analyses the privacy requirements that each architectural entity of the INDI ecosystem should satisfy and the privacy threats that it should defend against. Finally, the Annex compiles an overview of privacy enhancing technologies that are available today. This has the goal on one hand to map the existing technologies to the privacy requirements of Section 4, and on the other hand to help identify existing gaps between these two. In Section 5 several use cases are analysed and based on the analysis use case related privacy policies are extracted. After that, generic privacy policies are formulated and formed to a framework based on the knowledge gained from the use case related privacy policies. In the last step, the generic privacy policies are mapped to the privacy requirements from Section 2.

1 Introduction

The principle of informational self-determination [FHK+11] is of particular importance for online privacy due to the infrastructural and interactive nature of modern online communication and to the options that modern computers offer, even though it is much older than the notion of "Online Privacy". Well before the advent of Web 2.0, the term informational self-determination originated in the context of a German constitutional ruling, related to the 1983 census, making Germany the first country to establish the principle of informational self-determination for its citizens. The German Federal Constitutional Court ruled that[1]: "*[...] in the context of modern data processing, the protection of the individual against unlimited collection, storage, use and disclosure of her personal data is encompassed by the general personal rights of the [German Constitution]. This basic right warrant in this respect the capacity of the individual to determine in principle the disclosure and use of her personal data. Limitations to this informational self-determination are allowed only in case of overriding public interest.*"

To put it simply, this provision gave individuals the right to determine what personal data is disclosed, to whom, and for what purposes it is used. Informational self-determination also reflects Westin's description of privacy as "*the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others*" [Wes70]. Despite this legal development, the path of the Internet did much to undermine the elements of informational self-determination and nowadays, individuals have effectively lost control over the collection, disclosure and use of their personal data.

With the evolution and commercialization of the Internet and the advent of Web 2.0, including its search engines and social networks, the environment in which we need to support online privacy and informational self-determination, became more complex. Certain new business models, like ad-financed "free" services for Internet users, rely on a wide-range collection of user data for various purposes, such as marketing of online shops or targeted advertising and include user profiling. It appears that many users of online services are unaware of this business model. In other contexts, it is believed that data collected for commercial uses has been later employed for government purposes; this has been possible by the fact that the rules of "free services" place little restriction on reuse of data.

It has been pointed out that identity management is instrumental to the implementation of online privacy management. Indeed, identity management can be used to manage handling of data relevant to satisfy privacy requirements. User-centric identity management in this context implies that personal data – even in cases that is created by a service – is handed back to the user on request. If the user desires consistency across service invocation, it is her decision to hand over the data again to the same or another service. This way, individuals can supervise and limit personal data disclosure and exercise rights of access to their data held by third parties.

1.1 Privacy Threats

Privacy is subjective, contextual and therefore hard to evaluate. In this regard, one of the main challenges that researchers are currently exploring is linked with the analysis of individual attitudes on privacy. For instance, research has shown that most users of websites with customizable privacy settings, such as Online Social Networks (OSNs), maintain the default permissive settings, which may lead to unwanted privacy outcomes [KW08]. The explanation to this behaviour

[1] *BVerfGE 65,1 – Volkszählung,* availabe in English at http://en.wikipedia.org/wiki/Informational_self-determination

is not necessarily that users do not care about their privacy. Instead, existing studies demonstrate an ambivalence of the users' attitudes towards privacy [THM+08, BAL10]. What makes it more difficult to interpret people's attitude against privacy is that the notion of privacy differs or changes, depending on the culture that individuals are coming from. So, there is still much need for experiments with individuals to allow a broader range of privacy related analysis to be tested and enable a better understanding of people's concerns and the actions they take to address these concerns. However, there are several surveys that introduce the following concerns as typical for people regarding their privacy in online environments:

- Collection and storage of extensive amounts of personal data.
- Unauthorized secondary use by the collecting organization.
- Unauthorized secondary use by an external organization with which personal data has been shared.
- Unauthorized access to personal data, e.g., identity theft or snooping into records.
- Errors in personal data, whether deliberately or accidentally created.
- Poor judgement through decisions made automatically based on incorrect or partial personal data.
- Combination of personal data from disparate databases to create a combined and thus more comprehensive profile for a person.
- The inference of additional profile information through analysis of user populations and inferences made thereupon.
- The use of profiles to discriminate or manipulate access to systems or to automate decision making in a non-transparent manner.

This analysis becomes particularly difficult, since frequently there is no immediate damage for individuals. Even though in some cases, an individual may directly experience an offense, if harassed, manipulated or embarrassed as a result of a prior privacy violation, more frequently the consequences may occur only later or not at all, as for example in third-party tracking of online behaviour for targeted advertisements. Nevertheless, in general we can say that there are two core informational privacy concerns:

Observability: The possibility that others (potential observers) will gain information. Observers might include the parties communicating (for example, two people emailing back and forth), the service providers facilitating the communication (for example, email or Internet service providers), and eavesdroppers (for example, attackers sniffing email content or Internet traffic).

Linkability: The potential to link between data and an individual as well as potential links between different data sets that can be tied together for further analysis. Controlling linkability involves both maintaining separate contexts that observers cannot accumulate sensitive data and being cautious when identity information is requested to keep track of information disclosure.

At present, only few users examine the privacy policies of online systems even though they contain important information, namely how the provider promises to treat the personal users' data, e.g., for which purpose data are stored or when they will be deleted. An identity management system could analyse them and show the user what is essential for her privacy rights by using a

language easier to understand or even visual symbols. The user could decide on the basis of this information whether to give consent for data processing, which data to disclose or whether to refrain from interacting with the site at all. Even more sophisticated requirements may be negotiated, e.g., how long the data may be stored, which third parties may get access to personal data for specific purposes, or that the data may only be used if the provider pays for that. The privacy policies would be stored together with the information on disclosed data, like keeping a copy of the general terms and conditions [FHK+11].

1.2 Current landscape of available solutions

After more than 20 years of research in the area of privacy and privacy-enhancing technologies (PET), there exists a wide variety of mechanisms [DG10]. Broadly speaking, we could distinguish between opacity tools and tools that enforce other legal privacy principles, such as transparency, security or purpose binding. Opacity tools can be seen as the "classical" PETs, which "hide information", i.e. strive for data minimization and unlinkability. They cover a wide variety of technologies, ranging from cryptographic algorithms and protocols (e.g., [homomorphic] encryption, blind and group signatures, anonymous credentials, oblivious transfer, zero-knowledge proofs etc.) to complex ecosystems like user-centric identity management. Opacity tools can be further characterized depending on whether they focus on data minimization at the network layer or at the application layer. Proposals for achieving sender or recipient anonymity at the network layer comprise protocols such as Chaumian Mixes, DC-Net etc. At the application layer, a much greater variety of technology proposals exists, such as private information retrieval (PIR), privacy preserving data mining (random data perturbation, secure multiparty computation), biometric template protection, location privacy, digital pseudonyms, anonymous digital cash, privacy-preserving value exchange, privacy policies, etc.

Transparency-enhancing tools (TETs) belong in the second category of PETs and focus on enforcing transparency, in cases where personal data need to be processed. TETs frequently consist of end-user transparency tools and services-side components enabling transparency [RRD09]. The end-user tools include, among other techniques, (1) tools that provide information about the intended collection, storage and/or data processing to the user when personal data are requested from her (via personalized apps or cookies) and (2) technologies that grant end-users online access to their personal data and/or to information on how their data have been processed and whether this was in line with privacy laws and/or negotiated policies[2].

Examples are Platform for Privacy Preferences (P3P) User Agents, Amazon's Recommendation System, or the Data Track developed in the PrimeLife EU project [WH10]. TETs should also encompass an obligation to design interfaces that give the users insight into the full spectrum of their privacy risk exposure. Innovations like Google's Dashboard[3], although not comprehensively fulfilling this purpose, are a step in this direction [FHK+11].

[2] A third type of TETs, which has so far only been discussed in the research community, include tools with "counter profiling" capabilities helping a user to "guess" how her data match relevant group profiles, which may affect her future opportunities or risks [Hil09].

[3] https://www.google.com/dashboard/

1.3 Adoptability of privacy respecting solutions today

There is a growing amount of research in the field of PETs, proposing technologies for solving various aspects of the privacy problem; yet PETs are not widely adopted in practice, including when designing identity management solutions. Generally speaking, there is a lack of clear incentives for enterprises to manage personal data in a privacy-respecting manner, to design privacy-preserving products, or to make the use of personal data transparent to the user and data subject respectively. We identify the following root causes for this situation [FHK+11]:

1.3.1 Lack of customer and market demand

There is a lack of customer (individuals, business partners) and market demand for privacy respecting information and communication technologies (ICT), systems, services and controls (beyond punishments for breaches and other excesses).

Usage models for PETs cannot currently be targeted to customer demand. One reason for this is the lack of user awareness with respect to privacy problems, which can be partly attributed to missing transparency of data acquisition and the related information processing. In the current state, once the data has been submitted to an online information system, individuals get no information about any further processing. But, even if we assume that the data processing of such complex systems like Facebook, Apple iTunes or Google Search could be transparent to the public, it would be hard or impossible for ordinary individuals to understand what happens with their data. Consequently, this limitation leads to the observation that it is more important for individuals to understand the outcome and implications of data flows in complex online information systems than understanding the full data movements. However, some first steps have already been taken place to achieve this kind of transparent outcome-based approach. One of them is the creation of ad-preferences by some third-party advertisers, where users are allowed to see the set of outcomes, based on which data has been forwarded to the third-party (e.g. Google Ad Categories[4] or the Deutsche Telekom Privacy Gateway for location-based services on the Web).

1.3.2 Current economic environment fosters personal data collection

Some industry segments' norms, practices and other competitive pressures currently favour exploiting personal data in ways contrary to privacy and the spirit of informational self-determination (resulting in erosion of transparency and accountability).

In the current identity ecosystem, doing nothing about privacy or even aggressively collecting data sometimes pays off, as some companies seem to acquire new clients with new features based on creative data use and serendipity. Furthermore, for some players implementing complex data minimization schemes is costly and time consuming and makes information filtering to the user's best interest much harder, if not impossible.

[4] http://www.google.com/ads/preferences

1.3.3 Poor awareness

There is poor awareness, desire, or authority within some industry segments on the operationalization of privacy, e.g., to integrate existing PETs, to design privacy-respecting technologies and systems, and to establish, measure and evaluate privacy requirements and claims. When building applications, far too often engineers do not even know that by employing PETs, they can indeed achieve the required functionalities and security properties while at the same time protecting privacy through data minimization.

On the other side, employing PETs to their full potential is far from trivial. Existing PETs still need to overcome several shortcomings to become easier for engineers to deploy, as real world solutions require properties like usability, scalability, efficiency, portability, robustness, preservation of system security, etc. Today, only a patchwork of mechanisms exists, far from a holistic approach to solve the privacy problems. The interaction between these mechanisms and their integration in large-scale infrastructures, like the Internet, is not well understood so far.

1.3.4 Weaknesses in the current legal framework

There is currently a lack of clarity, consistency, and international harmonization in legal requirements governing data privacy within and across jurisdictions (avoided, for example, by migrating data somewhere up in the cloud).

Neither the current European legal framework, nor the US approach towards private sector self-regulation has been effective for the protection of privacy online, particularly with regard to new business models, such as behavioural targeting, user profiling, social networking and location-based services. Key weaknesses in the EU framework include that:

- Services based predominantly in the US are effectively outside European jurisdiction.
- European users have little choice but to "consent" to companies' terms of use and privacy policies in the absence of alternatives of comparable functionality.
- The concept of "personal data" is currently the necessary trigger for the applicability of the Data Protection Directive (Directive 95/46/EG [EU95]).
- There seems to be too much reliance on *ex post* securing of data rather than on *ex ante* elimination of privacy risks through data minimization.

2 Requirements in the light of Multilateral Security

Securing an identity management system against online fraud and identity theft is the main concern of the digital service providers and individual users. It is clear that the security of the organization and the privacy of individuals will not be protected if the aforementioned condition does not hold. Transparency, consent, data minimization and security are the most highlighted requirements in the context of user-centric identity management systems. In order to achieve these, a secure infrastructure is required as well as strong identity and access control enhanced with innovative solutions that minimize the collection and linkability of data. Employing such mechanisms helps protecting personal information against malware and unauthorized access.

2.1 Requirements of the sites and services using the system

Protecting the services against fraudulent parties while ensuring the customers' access to their personalized services is a major interest of service providers. On the other hand, flexibility and adaptability of access controls for organizations without the need for major modifications in the organizational infrastructures is necessary, since business partnerships are likely to change over time. Furthermore, it is important to reduce the risks that might result in damages due to sensitive information leak or abuse. This can be done by either protecting effectively the sensitive information or storing "derived claims" – e.g. an assertion by a trusted party – instead of the source data. The latter would be helpful in case of "data loss" as well.

It is highly desirable that compliance with relevant statutes, standards and audit requirements is an automatic outcome of the Identity Metasystem (for further information about Identity Metasystem see [RRD09]) as instances are deployed.

Both INDI operators as well as service providers have an economic interest in collecting, processing and distributing as much data as possible as part of their business model. They also have an interest in obtaining the trust of their users for the management of their online identities in the INDI space. However, large datasets and the exchange of identity data across a highly interoperable ecosystem of INDI operators increase the security risks and hence the security costs. Especially delegation mechanisms and models that rely on the transitivity of trust decision increase the flexibility of the INDI space while also introducing new risks. Abuse or misuse of the information of the data anywhere in the INDI space infrastructure can lead to massive privacy breaches as well as identity theft, finally hurting the companies and threatening the existence of the INDI space.

2.2 Requirements of the people using the system

Users have different security and privacy requirements towards the INDI space. Their privacy concerns are mainly about activities that could infringe their rights and freedoms. This infringement is more likely if data is collected and processed in mass. Such data collections increase the risk of internal and external abuse, function creep, and surveillance. Approaches to addressing these concerns are to provide the users with mechanisms to minimize the data that is collected and to provide them with transparency and controls over the data that is then processed (which can be greater than the data that is directly collected from the user). Users' concerns hence include the following:

- Discrimination based on aggregation of data (categorization and social sorting) or personal profile

- Limitations with respect to access to services

- Being forced to create identities in order to have access to services (inability to access services without using the digital infrastructure. This is especially important for basic governmental services but also others).

- Being forced to use the same identity across service providers or, vice versa, being expected to create multiple identities.

- In case of leakages, misuse and abuse, not being informed, or not being protected against future harms. Lack of due process with respect to past harms.

Below we elaborate on the technical requirements that identity management systems should satisfy, in order to address these concerns.

2.2.1 Transparency

One of the major principals required for detecting and evaluating the privacy risks in the INDI space is transparency. It is necessary for all the parties involved in the INDI space to have a clear understanding of the different aspects of personal data collection and processing in the system. The system has to provide an informative representation of the legal and technical aspects of the purpose of data collection, how the personal data flows, where and how long it is stored, what type of controls the user will have after submitting the personal data, who will be able to access the information, etc. Having these questions answered, the user as well as different stakeholders in the INDI space will be able to evaluate the privacy risks and make informed decisions about their participation in the INDI space.

Furthermore, mechanisms that allow the user to verify whether the data has been processed according to her policies are needed. This implies that mechanisms to prove the contradictions between the negotiated policy and actual collection and processing should also be available to users or authorized third parties [LeSH08].

Transparency can increase the trust of the users in the INDI space. Social science researchers within the PRIME project state that trust in an application can be enhanced if procedures are clear, transparent and reversible, so that users feel that they are in control of their data as well as their interactions.

2.2.2 User Control

In the INDI space, the trusted parties are not supposed to disclose any part of the personal information on behalf of the end user before getting her consent. The end user should have the ultimate control over instances of her own identity information. This control should be expanded to the whole life cycle of the users' identity information and any profiling activities that these identities are linked to. Further, it should be possible for the user to revoke her consent at any time.

To further enhance users' control over their data, it is assumed that the information about collection, processing and further use of data is presented to the end user in an understandable, convenient and unambiguous manner. User control can also be enhanced if mechanisms to measure and evaluate the potential impacts of new data items being created on the other partial identities can be implemented in a user-friendly manner.

2.2.3 Minimum Disclosure

The entities in the INDI space are required to minimize the risk of personal data misuse by avoiding or minimizing the collection, processing, distribution or centralization of collected data. Minimum Disclosure applies to all the processes that deal with personal data. All of the following processes should work with the minimum amount of personal data and be designed in that way:

Collection: In the process of information collection, the end user should not be asked for information that is not necessary for the given service. The minimization of data collection should be ensured through advanced computational methods as well as through clear purpose specification and appropriate requirements analysis. These principles should hold for cases in which the user provides some information directly or indirectly to another party (in case of GINI, when the operator request is forwarded to a Data Source).

Aggregation: Aggregation of identifying information also bears security and privacy risks. Putting limits on the aggregation of data minimizes the related risk. The fact that the released information cannot be linked to an individual is not a protection against the aggregation of data sets about large populations, the collation of data to infer additional information through statistical analysis that can then be used to control and manage those populations. Hence, aggregation should only be done in light of specific needs and under strict control of authorized parties, users should be able to opt out of profiling practices and resulting mechanisms of control, as well as from the inclusion of the data they generate in such aggregation practices.

Storage: If the user's personal data needs to be stored by one or more of the actors in the INDI space, the stored data must be reduced to a minimum. In line with the requirement for transparency and user control, any information should only be stored with the user's consent.

Retention: The life time of the stored data must be known to the user and it should be minimized as much as possible to reduce the risks for privacy and security violations. All the sensitive data should be discarded after its usage for the given purpose at time of collection. To enforce this, legal, organizational and technical mechanisms should be used.

Replication: Collected information may be replicated on multiple systems in order to improve service quality or reliability. From a privacy perspective, the replication of personal information should be minimized in order to improve the protection level.

Distribution: Information must be handed over only to the parties who really need to know it. Therefore another aspect of data minimization concerns the recipients of the users' data. The distribution of the data for the legitimate purposes of fulfilling the service should also be minimized and limited to legitimate third parties. The user should be informed of the transaction with these third parties and provided with mechanisms of oversight with respect to these third parties.

Linkage: *"Unlinkability ensures that a user may make multiple uses of resources or services without observing others being able to link these uses together. [...] Unlinkability requires that users and/or subjects are unable to determine whether the same user caused certain specific operations in the system"* [ISO99]. Advanced data mining technology can allow data controllers to construct links between different partial identities of the same entity. Protection of the users' activities from being linked together is another require-

ment of the users. The trusted parties who work on behalf of the users must try to reduce the information that can result in associating different activities to the same individual.

Even though the requirement "collect as little information from individuals as possible" could seem intuitive in protecting the privacy of participants in the INDI space, it might not always apply.

Let us consider for example the privacy principle of minimizing data collection. From the service provider point of view, identity-risk analysis involves determining the probability that an individual engaged in a particular transaction is using a stolen or forged identity. To detect such fraudulent behaviour, the service provider may want to gather as much information as possible about the individual involved that would enable a comparison of the transaction to the individual's history or profile. For example, if the credit card involved in the transaction is being used to make purchases in countries where it's never been used before, someone might be using the individual's identity fraudulently. Therefore, designers must consider alternative usable security mechanisms to detect fraudulent behaviour over those mechanisms that are implemented at the price of users' privacy.

2.2.4 Contextual Separation

Appearances of personal data in different contexts enable context spanning linkage and thereby the development of increasingly detailed profiles. The need for Contextual Separation is a corollary of Data Minimization, since the introduction of links between activities in different contexts is a form of aggregation and collection. Identities are a key instrument to efficiently connect personal data between different contexts. Therefore, there is a need for mechanisms that make it easy for the users to actively use several identifiers related to the context they are involved in, and keep these identities separated to avoid merging or mixing of contexts.

2.2.5 Delegation

In the context of identity management throughout life, one focus lies on investigating the necessity of delegation for people who are not able to manage their needs of privacy for a limited time or permanently. In modern systems and service-oriented architectures, it is common that an application is spread over different hosts or even companies. However, this means delegation of access control decisions, as some personal data is not stored and protected by the users themselves, but by delegate identity service providers. The system should be able to provide mechanisms to delegate an identity, limit the privileges and the context of it, and revoke the delegation when it is required. The security and privacy risks associated with delegation mechanisms should be under strict observation, delegation and especially delegation based on the transitivity of trust between organizations should be avoided as far as possible. If implemented, such mechanisms should be subject to heavy oversight mechanisms.

2.2.6 Accountability

Accountability can be discussed in two different perspectives. From the users' point of view, the collection of personal information includes a duty of care for its protection. So the parties who are involved in the data collection and processing are in charge of keeping this information safe. There is need for the end users to ensure that the actors in the INDI space provide an account of

their data collection and processing activities, as well as all automated decision making based on this data. The users and authorized parties should be able to keep these organizations accountable to what they promise to deliver and to be able to reprimand them or ask for compensation if this is not the case. On the other hand, the relying parties should be able to reveal the identity of the end user in case of any abuse. An example for this case is where a user rents a car just by anonymously proving that she has a driving license, and then she does not return it. Here, the service provider needs to be able to find her identity after going through the legal procedures. In addition, such a mechanism for accountability should not be used to abuse the users' rights, or used as a backdoor or other functionality into a system that was built for other purposes.

2.2.7 Purpose Binding

Purpose binding is another requirement in the same line as Contextual Separation and Unlinkability. Personal data should be relevant to the purposes for which they are to be used and to extend necessary for those purposes, and should not be usable in other contexts. Binding to purpose might be done in two ways: limiting/prohibiting the use outside the given context, or making the context stick to the data. Purposes for which personal data are collected should be specified no later than the time of data collection, and the subsequent use must be limited to the fulfilment of those purposes. The purpose limitation has central importance for business, since it attempts to set the boundaries within which personal data may be processed, and those within which data collected for one purpose may be used for other purposes. Furthermore, data must not be used for further purposes incompatible with the original purposes once they have been properly collected. The end user has to give his consent for every change of purpose.

2.2.8 Proportionality

The user expects that the proportionality of the information system she is interacting with has been determined, i.e., the system is absolutely necessary, does not fringe upon basic privacy rights, does not collect information unnecessary for the purpose (or the collected data is proportional to the purpose). Sometimes the proportionality requirement is related to the data minimization requirement. However, the focus of proportionality requirement is data collection and the focus of the data minimization requirement is more on data processing (For further information about the proportionality requirement and the difference between proportionality and data minimization see GINI Deliverable D3.1 [Van11]).

3 The INDI ecosystem

3.1 Claims

A claim is an assertion made by one subject about another subject that is defined to be "in doubt" until passing "Claims Approval".

By doubt we mean:

- The integrity and origin of the claim needs to be verified (e.g. through cryptography and evaluation of a security presentation); and
- The meaningfulness of a given party making a given claim about a given subject needs to be determined.

Through cryptographic methods, "doubt" may be resolved without any need to "call home" to the subject's claims provider. The degree to which a Relying Party is willing to believe or act upon a claim from an originating party constitutes part of a Relying Party's technical policy. Elaboration of this technical policy is the responsibility of the Relying Party's administrative domain.

3.2 Actors

GINI envisions an operator-based trust model (i.e. 'brokered' trust relationship) enabling the establishment of trust between the INDI Users, Operators, Data Sources and Relying Parties.

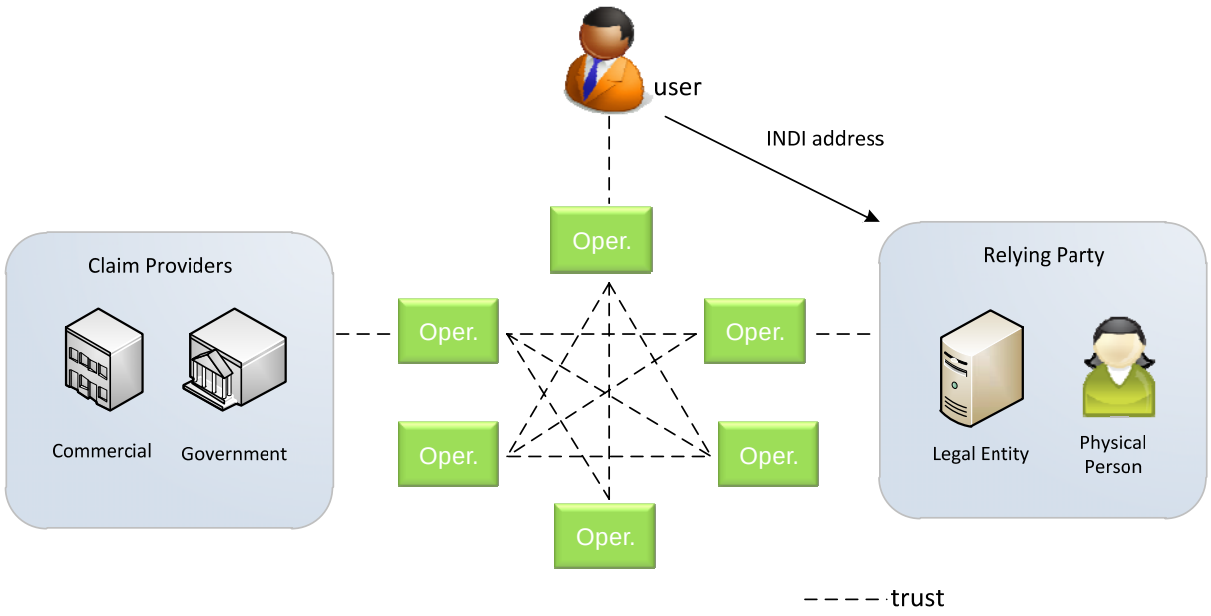

Figure 1: Relations between actors.

3.2.1 INDI user

In the INDI ecosystem, users can act in various roles, for instance citizen, employee, or customer. The user chooses which roles to act in and what information to reveal in the different roles. The user is able to manage its partial identities similarly with the physical world, by providing the relevant information to each situation.

3.2.2 Relying Party

In the INDI environment, a Relying Party could be either a legal entity or a physical person, with which users wish to perform transactions for personal, business or official purposes. One of the most prominent functionalities of an INDI environment (and the INDI infrastructure in general) is that it allows its users to present information about themselves in a verifiable fashion, i.e. in a manner which provides relying parties with appropriate assurance regarding the authenticity of the data that is presented (i.e. that the data originates from the identified source and has not been manipulated during the transmission).

3.2.3 INDI Operator

The underlying assumption in GINI is that users cannot manage multiple trust decisions in many different contexts, where they have to present their partial identities. If they need to understand and evaluate large amount of trusted third parties, users want entities, who they can trust and who can "represent" the whole infrastructure.

The INDI operator serves to represent the user in many different contexts, in such a way that users have sufficient technical assurances and legal safeguards, as represented by the INDI Operator, without having to question every entity of the underlying infrastructure. So, an operator is, together with other operators, responsible for enabling and managing the INDI environment. It acts as a gateway to the INDI environment on behalf of users, Relying Parties and Data Sources. The gateway can be standalone software or part of another software, e.g. a browser plugin.

The INDI users rely upon the Operator to deliver the basic INDI environment functionality, i.e. to facilitate the disclosure/presentation of information about them maintained in one or more Registers/Data Sources for the benefit of Relying Parties. The INDI user also relies upon the Operator to interact with these entities in a way, which will make the desired data exchange(s) possible.

In order to create and use an INDI, an individual must establish and maintain a relationship with at least one INDI Operator, or have multiple operators representing him in different interactions. This relationship may be contractual and should be sufficient for attaining access to the whole INDI environment (removing the need of additional one-to-one contracts). However, before the individual can use the INDI, and the operator fulfils its function, the identity must be created and enrolled with the operator service.

All INDI Operators can be a legal entity, but some of them could also be implemented as software running locally at the user's side. Whether some services are legal entities even if they are components that are under the control of the user is to be clarified in GINI Deliverable D3.1 [Van11] and D3.2 [Van12]. The following are the different types of INDI Operators as they are defined in GINI Deliverable D2.1 [Cau11]:

User Agent: A digital entity which can be used for selection of attributes to disclose, identity related service to connect to, allocation of required credentials, and the proof of attribute or assertions. In other words, the user can utilize the User Agent to: provision, propagate, maintain and de-provision digital identities, process service policies, control the flow of identity (related) information when the user makes these requests.

Identity Service: A service responsible for the verification and certification of the user's identity information.

Attribute Service: An entity that provides reliable attributes of a given user's "real world" identity.

Pseudonymization Service: Provides means to protect the user's (real-world) identity from e.g. identity attribute consuming business services. It substitutes identifiers and other attributes (that might potentially be used to discover the user's (real-world) identity) with persistent or transient pseudonyms.

Discovery Service: Provides the User Agent (and other services) with reliable and verifiable information on e.g. which identity related service to contact to acquire a certain attribute or pseudonym.

Business Service: A business service encapsulates any application logic and controls the access to its resources. The releasing of resources is based on identity related or attribute information of the user such as his age.

Policy Template and Protocol Directory (PTPD): manages information about policies and protocols. In particular it provides the User Agent with information on approved Service Sequencing Protocols. In addition the PTPD Service manages Policy Templates that may be used by INDI Services and Business Service for instantiating their particular INDI Policies.

Cross-Realm Service: Digital services are also very likely to cross national and regulatory borders with little or none indication of that aspect towards the user. However, due to specific regulatory requirements in the business service domain and concrete safeguarding responsibilities towards the user, as well as service deliver facilitation needs, an extended service is to be provided that mediates between the diverging dimension in order to make a service consumption regulatory compliant, fiscally sustainable, and technically feasible. The Cross-Realm Service decides according to a set of policies like government-issued lists of legitimate purposes of use, and GINI-specific harmonized mapping tables for substitution, correlation, availability, and statements of equality.

3.2.4 Claim Providers (Data Sources)

A claims provider is a digital service through which an individual or organization makes a claim about another individual, organization, device or service.

4 Privacy Requirements within the INDI ecosystem

Each consumption of digital identity data leaves traces (e.g. due to logging requirements). As with any digital data the ability for easily making copies quickly leads to a multiplication of identity data sets that are spread among multiple systems. Each copy makes it harder to control the proper use of this information and imposes new privacy risks through illegitimate linking and profiling of identity data. The goal of this section is to set the requirements that will guarantee privacy respecting handling of digital identities by the INDI ecosystem.

In requirements engineering it is useful and often necessary to structure and formalize the documentation of requirements. Requirements written in natural language are flexible and powerful in articulating the necessary functionality of the system, but as the number of requirements soar, then flowing text becomes difficult to handle, hampering the analysis of requirements. Hence, it is necessary to introduce different techniques that help in structuring requirements written in natural text and that enable elaborate analyses.

In analysing the requirements of the INDI ecosystem, we used basic guidelines to articulate the requirements in natural language and we used templates to further structure them. We expect these measures to assist us in achieving consistency and completeness with respect to the privacy requirements in the INDI ecosystem. In the following, we provide the reader with an overview of the techniques we utilized. As we do so, we also outline the scope of the privacy requirements analysis that we executed.

Basic guidelines for the documentation and analysis of requirements include the definition of a controlled vocabulary for formulating requirements (e.g., MUST is used for the specification of mandatory requirements, SHOULD for the specification of optional requirements). Team members cross-checked the requirements to make sure that requirements describe problems instead of solutions and amalgamated requirements are disjoined. Each of these requirements is also numbered and given a reference name (Name). The reference name is later used for the analysis of interactions between requirements.

Further, the focus of our analysis is on privacy requirements. Privacy requirements are comparable in many ways to security requirements: they are non-functional requirements that depend on the desired functionality of the system-to-be to be articulated. In our analysis, we depend on the functional description of the components of the INDI ecosystem, as these are defined in GINI Deliverable D2.1. Hence, the requirements below do not elaborate on the functionality of the INDI ecosystem, but focus solely on privacy requirements.

Since privacy and security requirements are comparable, we relied on existing literature on security requirements to create the requirements templates. Our objective is to complete a multilateral analysis of the privacy requirements: meaning requirements towards the system-to-be are analysed from the perspective of the different stakeholders and actors in the system [RPM99]. It is for this reason that we derived our templates from the tables proposed in [GJO+05] for executing multilateral confidentiality requirements analysis.

In order to capture the different stakeholder views, we included a field in the template to indicate the stakeholder or actor of the INDI ecosystem who has an interest in a given requirement statement (Subject). The component refers to the INDI entity – usually called asset – towards which the stakeholders' interest is articulated (Component). For each privacy requirement we also provide a more detailed account of the context of the requirement (Context) and the justification of the requirements (Justification).

Security requirements analysis can benefit from an explicit articulation of the trust assumptions, as well as the potential threats to the system posed by the various adversaries of the system. For these reasons, we adjusted our templates to articulate our assumptions explicitly and to capture threats posed to the INDI ecosystem by the different adversaries.

When we refer to trust assumptions, we refer to the decision on the side of the requirements analyst to include or exclude a domain with respect to the analysis of the security of the system-to-be, *assuming* that the security of that domain is taken care of otherwise [HLM+06]. These are the analysts' set of trust assumptions. They are implicit and explicit choices to trust some characteristic of the system environment. Assumptions can have a substantial impact on the security of the system. Each trust assumption also includes a detailed account of its context and its justification.

Privacy threats describe attacks using the vulnerabilities of the system that can lead to the breach of an individual user's privacy or that can violate some of the privacy related functionality of the INDI system. For each privacy threat, we include a definition of the threat, the vulnerability that it exploits, the impact of the execution of the threat, and the adversaries likely to execute the attack leading to the threat.

What is beyond the scope of this document is risk analysis. A security risk analysis demands the evaluation of the likelihood of occurrence and the negative impact of the combination of a threat with one or more vulnerabilities. Threat and vulnerabilities are part of the cause of the risk and impact is the consequence of the risk [MHM07]. Risk analysis requires knowledge of the context in which a system is going to be implemented, which is not in the scope of the GINI project.

However, a type of risk analysis can be extended to include an analysis of the trust assumptions, e.g., is the impact of the breach of the trust assumptions so great that it would be more reasonable to include them in the security analysis. We plan to execute such an analysis, once the first round of multilateral privacy requirements analysis is completed.

Template for requirements, threats and assumptions

We use the following template to document requirements and assumptions in our current high-level analysis.

Name	Subject	Comp	Assumption/Requirement
			Context
			Justification
			Related To

In addition to the requirement or assumption statement, the template for privacy requirements and trust assumptions include the following fields:

Name: refers to the number and reference of the requirement or assumption.

Subject: refers to the stakeholder(s) or actor(s) that have the requirement, make the assumption.

Component: refers to the components(s) in the architecture that are being analysed

Context: contextualizes the requirements, threats and assumptions in the given system

Justification: expresses why the requirement or assumption is reasonable to include in the analysis

Related To: allows a high level analysis of the dependencies and conflicts among and across requirements, threats and assumptions.

The template used for privacy threats has the following fields (notice the re-definition of the subject to template of requirements):

Name	Own	Comp	Threat
			Vulner.
			Advers
			Impact
			Related To

Name: refers to the number and reference of the threat.

Subject (here owner): refers to the primary actors and stakeholders that can be subject to the impact of the threat.

Component: refers to the components(s) in the architecture that are subject to the threat

Threat: Potential attack or incident which, in combination with one or more vulnerabilities, could have a negative effect on the privacy of the users.

Vulnerability: Characteristic of the information systems in the INDI ecosystem that can constitute a weakness or a flaw in terms of the privacy of the user or the privacy related functionality of the ecosystem. It could be accidentally or intentionally exploited in an attack.

Impact: The impact is the potential negative consequence of the realization of a threat that may harm the privacy of the user or the privacy related functionality of the ecosystem.

Adversary: The actors, stakeholders or other parties those have an interest in exploiting vulnerabilities of the INDI ecosystem.

All actors and stakeholders of the INDI ecosystem have an interest in some subset of the privacy requirements, i.e. have stakes in avoiding some subset of the privacy threats. An additional stakeholder is the entity hosting the User Agent (UA), if this is not on a device controlled by the user. One or more of any of the other actors/stakeholders of the INDI ecosystem may be an adversary of the User Agent. Further, third parties that attack the User Agent may also be expected. The adversaries may collude. If the User Agent is hosted on a third party, then that party is also a distinct adversary. The following table shows which INDI insider adversaries are envisioned; note that outsider adversaries, such as hackers and attackers are also considered.

Adversaries (Insiders)
Identity service (IS)
Business service (BS)
Attribute service (AS)
Pseudonimization service (PS)
Discovery Service (DS)
Cross-Realm Service (CRS)
Policy Template and Protocol Directory Service (PTPD)
Other Users (U)

Third Parties (TP)
UA Hosting Service (UAH)

This sub-section documents our trust assumptions, the identified privacy threats, and the resulting privacy requirements for the INDI Operator. This covers the entities that fall under INDI operator status, namely User Agent, Identity Service, Attribute Service, Pseudonymisation Service, Discovery Service, Business Service, Cross-Realm Service, and PTPD Service.

4.1 Trust Assumptions

In the following we state the trust assumptions with respect to the INDI Operator using our template. Notice that some of these assumptions are requirements towards the environment of the system, in this case, the environment of the User Agent [AnLe11].

From a privacy analysis perspective, the User Agent can be considered as the most critical type of INDI Operator and, in fact, the most critical entity in the GINI environment overall. This entity represents the real user in the GINI world and performs on her behalf desired transactions in order to deliver convincing proofs regarding the required attributes of the user. Therefore, it is essential for the user to have this User Agent under her control as much as possible. The user inevitably has to trust the User Agent to behave as expected. There are several ways in which the User Agent might be implemented. It could be tool running on the local machine, or software that resides somewhere outside of the user's sphere (e.g. in the cloud), or somewhere in between. It is worth mentioning that as the User Agent gets closer to the user, it becomes more privacy friendly but less portable.

In any case, the user must have full control over the User Agent. However, this control depends on where the User Agent resides and whether the environment of the User Agent is such that it provides full control to the user. Hence, we state this as an assumption on the environment but also later as a requirement towards the component.

Name	Own	Comp	Assumption	
TA.1 UAControl	User	User Agent	The user has control over her User Agent	
			Context	No technical backdoors or exceptions can be introduced to provide the information in the User Agent to other parties. Legal provisions to provide the user with such control must also be in place, but should not replace technical measures.
			Justif.	The centrality of the User Agent in providing the user with control over her data, assisting her in decisions, and informing her about any violations, makes the User Agent a target for attacks and manipulation.
			Related To	

Name	Own	Comp	Assumption	
TA.2 UAEnviron	User and all other actors of the INDI	User Agent	The environment in which the User Agent is located (be it on a device belonging to the user or on devices run by third parties), the security of the environment must be guaranteed. The security measures include the confidentiality of the User Agent content, but also the User Agent's traffic data.	
			Context	The device on which the User Agent is hosted should be secured such that this environment cannot be used to breach

		the security of the User Agent and hence the privacy of the user.
	Justification	It is not only enough to secure the User Agent, but also the devices/servers on which a User Agent is hosted. If the User Agent is hosted on remote servers than the User Agent is also susceptible to traffic analysis attacks.
	Related To	

Similar assumptions must hold for the other types of INDI operator.

Name	Own	Comp	Assumption	
TA.3 Infrastructure-Security	User and all other actors of the INDI	Identity, Attribute, Pseudonymisation, Directory, Business, Cross-Realm, PTPD Services	The service ensures that it is not vulnerable to outside attacks such as compromise of its infrastructure by hackers or attackers. Denial of service attacks are not included in this assumption.	
			Context	If the service is vulnerable to being taken over by outsider attackers, then the privacy of all users is at stake, and any other internal technical measures aimed to protect privacy will be circumvented.
			Justification	Denial of service attacks are outside the scope of this assumption since an interruption of service by itself does not lead to leakage of sensitive information.
			Related To	TA.2

4.2 Privacy Threats

The following are privacy threats pertain to the User Agent.

- Non-minimal disclosure: If the User Agent discloses more information than it is required to perform a transaction, or agrees to some data collection policies that are not necessary.

- Autonomous participation: If the User Agent gets access to the attributes of the user, then it might be capable of performing different transactions or disclosing some private information, without the user being aware.

- Traffic Analysis: Assuming the User Agent to be well behaved, it is always exposed to some traffic analysis attacks, which expose user to risk of profiling. This is a bigger concern when the Agent is deployed remotely. The cloud provider that is hosting the Agent should try to obtain some information about the performed transactions and build a profile of the user according to that.

- Stored Information: If the User Agent saves sensitive information such as attribute values, transaction logs, etc. it might be a source of information leaking and can threaten the user's privacy. It can happen with or without the knowledge of the user.

Name	Own	Comp	Threat	
PT.1 UADataMinimization	User	User Agent	The User Agent discloses more information than it is required to perform a transaction, or agrees to some data collection policies that are not necessary. The User Agent policy enforcement is manipulated, or the policy enforcement decision-making algorithm is manipulated to disclose more information than necessary.	
			Vulner.	The UA decision-making is not transparent to the user with respect to how it interprets policy rules and hence it discloses more information than necessary. The INDI space is designed such that the User Agent has to disclose unnecessary information by default (see case of P3P). The User Agent discloses the identity of the business service to the identity, directory or attribute service provider. Traffic analysis on the User Agent's communications reveals confidential information.
			Advers	The Business Service is the main adversary. However, all other adversaries may actively attack and manipulate the UA to disclose more information than necessary. The UA itself may also be the one that causes this threat as a result of its design.
			Impact	Data minimization is not practiced. Additional information is available about the user, increasing profiling activities by parties who collect the data. The user will likely be unaware of the disclosure, meaning the user is not in control of her data and the system's user centricity is questionable. All of these are likely to lead to a loss of reputation in the GINI system and loss in trust towards the infrastructure. The disclosed information may also directly lead to privacy breaches.
			Related To	

Name	Own	Comp	Threat	
PT.2 UAAutoParticipation	User	User Agent	The UA gets access to the attributes of the user, performs different transactions or discloses some private information, without the user being aware because the internal processes are not transparent.	
			Vulner.	The UA decision-making is not transparent to the user with respect to how it interprets policy rules and hence the user inadvertently authorizes the disclosure of personal information to the UA. The UA then starts disclosing information beyond the user's control and awareness. The user sets up policies without being aware of their consequences, leading the UA to disclose much data autonomously.
			Advers	The Business Service is the main adversary to manipulate the UA. Further, the design of the UA is likely to lead to this threat. All other adversaries may actively attack and manipulate the UA to disclose the attributes it collected.
			Impact	Same impact as PT1.
			Related To	

Name	Own	Comp	Threat	
PT.3 UATraffi-cAnalysis	User	User Agent	UA is subject to traffic analysis	
			Vulner.	UA on user controlled device: different service providers may collude to use traffic data to do timing attacks (linking identity of the user to services she uses). If the UA is hosted elsewhere, it may be subject to traffic analysis by the host who can create a profile of her activities. The host may also collude with the other service providers, increasing the strength of her observations.
			Advers	If the UA is hosted elsewhere, the hosting service (UAH) may perform the traffic analysis. All other parties may collude to increase the scope and strength of their observations.
			Impact	The analysis of the UA's traffic may make it possible to link the transactions of the UA with different services, making her transactions and further information vulnerable to being linked across services.
			Related To	TA.2 UAEnviron

Name	Own	Comp	Threat	
PT.4 UACon-tentConfi-dential	User, Business Service, Identity Service	User Agent	The confidentiality of the content stored in the UA is breached	
			Vulner.	The UA itself is not secured or the environment in which the UA is hosted is not secured.
			Advers	An adversary other than the UA attacks the UA to leak the content stored in it. An adversarial user may also attack the UA to gain access to the user's information.
			Impact	The loss of confidentiality may lead to privacy breaches and to identity theft. If the user is not aware, such leakage may also be used for profiling the user over time. If identity theft or profiling occurs, these will have a negative impact on the use of and trust in the GINI ecosystem.
			Related To	TA.2 UAEnviron

The following are privacy threats pertaining to the all other INDI operator types, namely Identity, Attribute, Pseudonymisation, Directory, Cross Realm, PTPD, and Business Services. Note that the Discovery Services may not be subject to the threats, depending on whether or not they handle personal data.

- Consent breach: The service uses personal data in ways that are not covered by user consent.

- Stored Information: If the service stores sensitive information such as attribute values, or transaction logs, these may be leak due to internal attacks and leakage. The Identity, Attribute, Pseudonymisation, and Business services are expected to handle personal information.

Name	Own	Comp	Threat	
PT.5 ConsentBreach	User	Identity/Attribute/Directory/Pseudonymisation/Business services/Cross-Realm/PTPD Agent	The service uses personal information for any purpose or in any context that is not covered by user consent (except mandated processing).	
			Vulner.	If personal data is used outside the limits of the law, makes the service provider vulnerable to legal action. However, from a privacy perspective, the damage is already done if data is used excessively and/or not covered by consent. Unwanted Dissemination: The service might violate the user's privacy by disseminating the collected information to other parties, e.g. advertising company. Profiling: In case there is some information in the provided attributes that could be used to identify the user, the service might utilize this information to track the user or make a profile of her without the user's consent. Traffic Analysis: If the service uses some traffic analysis techniques, such as Internet Protocol (IP) logging, it threatens the user's privacy by increasing the chance of identifying the user and profiling her. Tracing: The service is able to find out the identity of another service that the user wants to use. For that, it could collude with the other services or even the Pseudonymization Service, perform traffic analysis, or use some hidden triggering mechanisms that notify the source upon consumption of the credential, in order to get more information about the service, which the user is interested in. This eventually leads to profiling of the user.
			Advers	External adversaries are likely to attempt to coerce the service to share and use personal data for purposes such as direct marketing and targeted advertising, as well as the constructions of profiles that can be used to discriminate types of users.
			Impact	The user is likely not to be aware of the authorised usages of her personal information, meaning the user is not in control of her data and the system's user centricity is questionable. All of these are likely to lead to a loss of reputation in the GINI system and loss in trust towards the infrastructure. The unauthorised usage of personal information may also lead to direct privacy breaches.
			Related To	

Name	Own	Comp	Threat
PT.6 Stored-DataAbuse	User	Identity/Attribute/Directory/Pseudonymisation/Business services/Cross-	The database and/or log files of the server are compromised, and personal data is leaked or otherwise abused.

	Realm/PTPD Agent		
		Vulner.	Internal adversaries may gain access to personal data of users/customers/employees/citizens, either in an unauthorized or an authorized manner, and then leak the data to others without the approval of the organisation, leak the data to the public at large, or otherwise abuse the data for their personal benefit. Security Abuse: Additional to the unwanted dissemination threat, the service can also be subject to security abuse or internal misuse, leading again to privacy breaches. Retention: Keeping the collected attributes longer than the time period declared to the user is known as a kind of data misuse, which is against the privacy rules. Even when these data have been anonymized before, in many cases they can still be linked back to specific users by de-anonymization techniques [NS08]. Unnecessary Inspection: Inspection mechanisms are designed to be used for identifying the user when she violates the terms of use of a service. The service can try to disclose the user's identity by breaking into the inspection tokens provided by the user. Linkability: In case of a standalone implementation of the Pseudonymisation Service, it is possible for the provider to link different credentials coming from a specific user and extract additional knowledge which threatens the user's privacy.
		Advers	Internal adversaries, such as rogue employees.
		Impact	Same or more severe impact as PT5.
		Related To	PT.5

The following threat applies only to Pseudonymisation Services.

Name	Own	Comp	Threat	
PT.7 PSTrafficAnalysis	User	Pseudonymisation	The Pseudonymisation Service is subject to traffic analysis. For example, a weak implementation of the Pseudonymisation Service, where there is a correlation between the pseudonyms, will lead to an opportunity for the other GINI operators (e.g. Business Service) or even external attackers to link different transactions of a user, build a profile and eventually identify the user.	
			Vulner.	Service providers and external adversaries may collude to use traffic data in order to link data items that the Pseudonymisation Service was supposed to render unlinkable. The data that may be used for conducting the traffic analysis can be timing data, but also pseudonym values, and any other metadata that the Pseudonymisation Service transmits in its anonymized or pseudonymized messages. Since service providers may collude in this traffic analysis, decryption is within the abilities of the adversary.
			Advers	Other services in the INDI ecosystem as well as outsiders.
			Impact	This type of traffic analysis undermines the value of the pseudonymisation service. The pseudonymisation service has an interest to counter such attacks as they undermine its business case.
			Related To	PT.3

4.3 Privacy Requirements

Since the User Agent works on behalf of the user, it is responsible for providing technological and legal guarantees, showing that it is utilizing all the necessary mechanisms to protect the user's privacy.

Accountability. In order for the user to trust the User Agent and ensure about the correct behaviour of this entity, there is a need for a legal framework to guarantee the accountability. There must be adequate oversight tools to ensure that providers of User Agent tools (local or remote) can be held responsible for any malicious behaviour or lagging in the User Agent.

Data Minimization. User Agent is where the major Data Minimization requirements must be taken care of. It should guarantee that no more information than it is required would be disclosed.

Transparency and User Control. The User Agent is not allowed to reveal any attribute or other sensitive information before the user gives her consent. In order for the user to agree on revealing certain attributes in a transaction, the User Agent must provide a descriptive set of information about the steps of the process and the details regarding the parties who will participate in the transaction, the information which will be stored on the other sides, or the location and duration of the storage. In order for this to be comprehensive, the other entities of the INDI operator, i.e. Identity, Attribute, Directory, Business, Cross-Realm, PTPD, and Pseudonymisation Services, must also support transparency technologies and processes. In particular, they must indicate which data items are used for which purposes and what are applicable data retention periods.

Name	Own	Comp	Requirement	
R.1 UASecurity	User and all other actors	User Agent	The integrity, confidentiality and the availability of the User Agent MUST be guaranteed. The data collected and processing executed by the User Agent MUST be confidential towards all parties other than the user. The communication channels of the UA with all services MUST be secured.	
			Context	The User Agent has great powers to collect, aggregate and manage the user's data. Hence, this device must be secured such that it only does what it is supposed to do and nothing else. This includes protecting the confidentiality of collection, processing and use of the user's data from all other parties. If the data and processing of the User Agent is not confidential, then any party can observe the user's data breaching the control requirement. Further, if confidentiality is not guaranteed it makes no sense to use any Pseudonymisation or Anonymization Service.
			Justif.	Any vulnerability of the User Agent is likely to cause breaches of privacy, increase the risk of identity theft, and cause damages to the business services relying on the INDI framework. Hence, all actors and stakeholders share this requirement.
			Related To	PT.1 UANonMinDisclosure PT.3 UATrafficAnalysis PT.4 UAContentConfidential TA.2 UAEnviron

Name	Own	Comp	Requirement	
R.2UAAccountability	User, Business Service, Identity Service	User Agent	The accountability of the UA towards the user and third parties MUST be guaranteed	
			Context	The User Agent must include technical and legal measures to

		guarantee that it works as expected. Technically, this can be guaranteed through transparency mechanisms like internal audit functionality that logs the activities of the UA.
	Justif.	There should be mechanisms to back-track the correct functioning of the User Agent in order to provide the user with oversight and also with a mechanism to prove her actions towards third parties, if this were deemed necessary.
	Related To	PT.1 UANonMinDisclosure PT.2 UAAutoParticipation TA.1 UAControl

Name	Own	Comp	Requirement	
R.3UAData Min	User	User Agent	The User Agent MUST guarantee that no more than the information required for any transaction is disclosed.	
			Context	The User Agent must be designed not to disclose unnecessary information. Mechanisms should be in place to detect or retract the disclosure of unnecessary information.
			Justif.	Data minimization is one of guiding principles of privacy by design and ensures that the user remains in control of her data.
			Related To	PT.1 UANonMinDisclosure PT.2 UAAutoParticipation TA.1 UAControl

Name	Own	Comp	Requirement	
R.4 UAControl	User	User Agent	The user has full control over her User Agent regardless of where it is hosted	
			Context	No technical backdoors or exceptions can be introduced to provide the information in the User Agent to other parties. Legal provisions to provide the user with such control must also be in place, but should not replace technical measures.
			Justif.	The centrality of the User Agent in providing the user with control over her data, assisting her in decisions, and informing her about any violations, makes the User Agent a target for attacks and manipulation.
			Related To	PT.1 UANonMinDisclosure PT.2 UAAutoParticipation PT.3 UATrafficAnalysis PT.4 UAContentConfidential

Name	Own	Comp	Requirement	
R.5UATransparency	User	User Agent	The User Agent MUST provide a descriptive set of information about the steps of the process and the details regarding the parties who will participate in the transaction, the information, which will be stored with different parties, the location of the data, the duration of the storage and the intended use (purposes). All this data must be made available to the User Agent in order for it to make informed decisions when supporting the user in selecting personal data that will be disclosed in various situations.	
			Context	The User Agent is not allowed to reveal any attribute or other sensitive information before the user gives her consent. In order for the user to agree on revealing certain attributes in a transaction, the Agent must provide a descriptive set of information about the steps of the process. In order for this to be comprehensive, all GINI operators must support according protocols and processes.
			Justif.	Data minimization is one of guiding principles of privacy by design, can be inferred from data protection, and contributes to the concept of user control and centricity.
			Related To	PT.1 UANonMinDisclosure PT.2 UAAutoParticipation

TA.1 UAControl

The following are privacy requirements for all other INDI operator entities, namely Identity, Attribute, Pseudonymisation, Directory, Cross-Realm, Policy Template and Protocol Direcotry, and Business Services. Note that the Discovery Services may not be subject to the requirements, depending on whether or not they handle personal data.

- Data Minimization: The processing of personal data is not excessive in its collection and usage of personal data (data minimization). The service is required to minimize the information included in the personal information data fields so that no more information than it is necessary will be revealed than neccessary. Data Minimization also implies that the service must be obliged to dispose all the information about the user when the service contract expires (if no archiving is necessary).
- Security of stored Information: If the service stores sensitive information such as attribute values, or transaction logs, then this information must be sufficiently protected against leakage; against internal and external attackers.
- User-informed Inspection: The inspection mechanism must be implemented in such a way that when the inspection token is unpacked, the user will get informed, unless legal framework allows otherwise. Therefore, if there is a misuse by the service, the user gets notified and reacts appropriately.
- Accountability: Preservation of evidence is necessary to protect the user's rights in case of any unauthorized action by the service.
- Transparency: The service is obliged to offer (1) tools that provide information about the intended collection, storage and/or data processing to the user when personal data are requested from her (via personalized apps or cookies) and (2) technologies that grant end-users online access to their personal data and/or to information on how their data have been processed and whether this was in line with privacy laws and/or negotiated policies. It is also important to offer tools with "counter profiling" capabilities helping a user to "guess" how her data match relevant group profiles, which may affect her future opportunities or risks [Hil09].
- Unlinkability: The presentation of different credentials from the pseudonymisation service should be unlinkable with each other.

Name	Own	Comp	Requirement
R.6DataMin	User	Identity/Attribute/Pseudonymisation/Directory (if applicable)/Business/Cross-Realm/Policy Template and Protocol Direcotry Service	The service MUST minimize the information included in the personal information data fields so that no more information than it is necessary will be revealed. Data Minimization also implies that the service must be obliged to dispose all the information about the user when the service contract expires (if no archiving is necessary).

	Context	If data minimization is not practised throughout the GINI ecosystem, the risk of privacy breaches rises both on an individual and on the collective level.
	Justif.	Data minimization is one of guiding principles of privacy by design, can be inferred from data protection, and contributes to the concept of user control and centricity.
	Related To	PT.5

Name	Own	Comp	Requirement
R.7StoredInfoSecurity	User	Identity/Attribute/Pseudonymisation/Directory (if applicable)/Business/Cross-Realm/Policy Template and Protocol Direcotry Service	The service MUST implement technical measures to ensure that the likelihood of insiders being able to abuse personal data of consumers/customers/citizens is eliminated or at least minimized. Internal audits and other appropriate measures should complement the technical measures.
			Context Encrypting data at rest, distributing decryption keys to multiple persons, exercising separation of duty, strict internal access control policies that honour the principle of least privilege and "need to know", are examples of technical measures that can be utilised to satisfy this requirement.
			Justif. If data at rest is vulnerable to leakage, the privacy of the affected customers/consumers/citizens is endangered.
			Related To PT.5,6

Name	Own	Comp	Requirement
R.8UserInfoOnInspection	User	Identity/Attribute/Pseudonymisation/Directory (if applicable)/Business/Cross-Realm/Policy Template and Protocol Direcotry Service	Any service provider that may release personal data to a third party outside the context of "business as usual", or any party that may trigger the revocation of the anonymity of the user, either by itself or by means of a third party, MUST inform the user whenever the data disclosure or anonymity revocation process is triggered, along with a description of the reasons behind this. In this case of user-informed inspection, the communication MUST be such that the user is informed either before or shortly after the process was triggered.

	Context	This requirement applies only in situations where data can be disclosed or anonymity can be revoked within the context of a well-defined process, e.g. an on-going criminal investigation. For example, the service provider may obtain an "inspection token" from an authority, which enables anonymity revocation or inspection of a particular set of data fields.
	Justif.	If users are not informed in good time, then services might abuse their ability for revoke anonymity. This undermines the entire purpose of the GINI ecosystem.
	Related To	PT5,6

Name	Own	Comp	Requirement	
R.9Accountability	User	Identity/Attribute/Pseudonymisation/Directory (if applicable)/Business/Cross-Realm/Policy Template and Protocol Direcotry Service	All INDI operators must have clearly defined and compliant terms of service. Moreover, it should be clear and easy for users to complain, trigger appropriate dispute resolution processes, and even to be able to gather all information required in order to initiate legal action if necessary.	
			Context	In complex ecosystems, it is often unclear which particular entity holds one's data, or where a particular privacy breach occurred. This requirement addresses some of the negative side-effects of this situation.
			Justif.	While some jurisdictions require this requirement to be addressed, it is important to stress its importance and to implement consistently across the GINI ecosystem even in countries where laws are less stringent.
			Related To	PT5,6

Name	Own	Comp	Requirement	
R.10Transparency	User	Identity/Attribute/Pseudonymisation/Directory (if applicable)/Business service	The services MUST provide descriptions about who will participate transactions, the information, which will be stored with themselves and forwarded to different parties, the location of the data, the duration of the storage and the intended use (purposes). All this data MUST be made available to the User Agent in order for it to make informed decisions when supporting the user in selecting personal data that will be disclosed in various situations.	
			Context	The User Agent MUST be in a position to support the user in selecting attributes and personal data for usage in transactions. In order for this to be effective, all GINI operators must support protocols and processes in accordance with this requirement. The overarching purpose of this is to achieve better data minimization.
			Justif.	Data minimization is one of guiding principles of privacy by design, can be inferred from data protection, and contributes to the concept of user control and centricity.

	Related To	R.5UATransparency PT5,6

Name	Own	Comp	Requirement	
R.11Unlinkability	User	Pseudonymisation service	The presentation of different credentials from the PS should be unlinkable with each other.	
			Context	The goal of the PS is to help the user access a service without disclosing her identity to the Business Service. There are two approaches to implement such a service. One way is to have this service as a standalone entity in the system (i.e. a legal entity). This requires a trust relationship between the PS and the Business Service so that the user can provide the Business Service with a credential that is pseudonymized by the PS and have it accepted. The second approach is to use state of the art cryptographic mechanisms to integrate the PS into the User Agent and pseudonymize the received attribute claims while they are still certified under the signature of the Attribute Service Provider. Using these technologies, the User Agent is able to selectively disclose the attributes and hide the identity information, and the Business Service can ensure the authenticity of the attribute values without knowing the user's identity. The important point is that, the closer the PS is to the user, the more privacy friendly it becomes.
			Justif.	Unlinkability of credentials is a very important feature that enables data to be significantly minimised and privacy to be protected to a much higher degree than without unlinkability. This is because, without unlinkability, all services will be able to unambiguously link all transactions of a given user while, with unlinkability, this linking is significantly hampered.
			Related To	PT.7

5 Privacy Policy Framework for INDI ecosystem

In this chapter we derive a generic Privacy Policy Framework. At the first step we stated assumptions, which are essential within the INDI ecosystem. Then, the general structure of the upcoming privacy policies is explained based on available literatures. After that we took four use cases from GINI Deliverable D2.1 [Cau11] to extract use case related privacy policies. These four use cases were chosen because they cover many aspects of privacy, which allow us to generalize them to high-level privacy policies. They form the generic Privacy Policy Framework. In the last step we map the derived high-level privacy policies to the requirements stated in the Chapter 2.

5.1 Assumptions

Before we derive and set up a Privacy Policy Framework for the INDI ecosystem some assumptions about the term "trust" have to be made. In the INDI ecosystem trust is a fundamental element. Since the GINI conceptual model is based on a network of operators, an individual must establish and maintain a relationship with at least one IDNI operator.

GINI Deliverable D4.1 builds upon the trust assumptions laid out in GINI Deliverable D1.1 [AnLe11] and D3.2 [Van12]. First of all, trust is a very versatile term with a lot of different definitions, views and interpretations. In sociology it deals with influence and impact of trust in social systems [Luh79]. From a psychological view, *"*trust is a bet about the future contingent actions of others*"* [Gol05]. The INDI model is an operator-based model; therefore the operative view is important. So it is essential that the different operators, regardless if they know each other or not, build up a basic trust relationship for further interaction.

For basic trust relationships we assume that the trust relationship to the corresponding INDI operator includes:

- The INDI operator can run locally, which assumes direct trust relationship with the user, or remotely which assumes, that the user trusts the remote party,
- The INDI operators act within the policy boundaries,
- The INDI operator delivers the promised functions and security mechanisms.

The trust to the own INDI operators is very important because they act as trust anchors. They establish trust relationships without knowing the other parties, which means that they contact and interact with trusted and untrusted domains. In the last case, the INDI operators have to build trust to untrusted domains, which implies a basic or weak trust relationship (i.e. the user won't provide as much data as to a known and trusted domain). To enable this basic trust relationship, we have to assume that the INDI operators are 'clean', which means that the systems are not compromised or infected with malware. This assumption is essential and necessary because the interacting INDI operators do not have the possibility to check the integrity of other INDI operators and nobody wants to exchange data and information with a compromised and/or infected system.

For the same reasons we also assume that only authorized persons have access permissions to the INDI operators and that the INDI operators use advanced technologies, which enable a higher security level. This ensures the authenticity of information and the security of certificates.

The last assumption is about the accuracy and trustworthiness of the provided information. We assume and it also can be seen as a requirement to the organisations, that information, which is confirmed or vouched by a third party, is reliable. This leads to more accuracy and trustworthi-

ness of the information. This is a very abstract assumption, because in reality a confirmation or vouch does not in and of itself render the information trustworthy.

5.2 Structure of Privacy-Policy

According to Karat et al. [KKB+09] a privacy policy can have up to six elements:

- Data user,
- Action,
- Data,
- Purpose,
- Conditions,
- Obligations.

This format is specified in security access control standards like Enterprise Privacy Authorization Language (EPAL) or eXtensible Access Control Markup Language (XACML). Not every element needs to be in a privacy policy. For example the element "obligation": Obligation is defined as *"actions to be performed after an action has been executed on data objects"* [KKB+09]. For example, a nurse is allowed to forward corresponding medical patient data to the head physician for treatment if the patient is assigned to this head physician. In this case most of the elements can be mapped:

- Data user = nurse, head physician
- Action = forward
- Data = medical patient data
- Purpose = treatment
- Condition = if the patient is assigned to this head physician.

This policy is valid and does not need an obligation, because there is no need to perform an action after the forwarding action. Of course, the medical patient data must be archived when the treatment is done and the access permissions must be revisited when the patient leaves the hospital and these actions are performed after the forwarding action. But they are not related to the forwarding action. In the following chapters and sections, different use case related and generalized privacy policies will be created, but most of them do not contain every element because the purpose of the privacy policy can be reached without using all of them.

In addition to the upper elements, different vocabularies are used for the privacy policies. According to the TAS3 Project [AAV+11] three verbs are reasonable for distinguishing different privacy policy levels. We adopt this vocabulary for this document. The term "MUST" is used to clarify that the privacy policy is formulated to comply direct legal obligation (e.g. EU Data Protection Directive 95/46/EC [EU95]). The term "SHOULD" is used to show that the privacy policy does not reflect clear and direct legal obligations, but rather is best practice. The term "SHALL" indicates that the formulated privacy policy is not a clear and direct legal obligation, but it is needed to achieve the GINI objectives.

5.3 Use Cases

The development starting point of the Privacy Framework is the use cases. Due to the generality of the GINI model and INDI ecosystem, it is hardly possible to cover all situations and use cases. So this deliverable will only analyse some use cases which contain the most important and common privacy threats. Privacy policies can be derived from this information and combined to a Privacy Policy Framework. The use cases, privacy policies and Privacy Policy Framework are transferable to other use cases and situations.

In the following parts of the deliverable, four use cases will be explained at first. These use cases are presented and described in GINI Deliverable D1.1. The user-centric use cases were chosen because the GINI conceptual model is a user-centric approach. From this perspective, the privacy threats to the user can be revealed and the Privacy Policy Framework maximizes its user protection.

The next step is to apply the INDI ecosystem on it, which forms the base of the analysis to derive privacy policies. According to GINI Deliverable D3.1 [Van11] the INDI ecosystem has an operator-based trust model. On this account, the focus in the analysis lies on the interaction between the different INDI operators considering the trust assumptions described in Section 5.1. The goal of this analysis is to generalize the extracted privacy policies and form them to a Privacy Policy Framework. The high-level privacy policies enable the transferability to other use cases and situations.

In the best case, the Privacy Policy Framework can be applied to every case. However, with the information of four use cases and although there are a lot of more use cases, situations and conditions in the future we cannot imagine the demand of modification now and deployment is very likely.

5.3.1 Use Case: Person-to-Person Transactions

This chapter analyses the use case 6 of the GINI Deliverable D1.1 [AnLe11]. It is about a consumer-to-consumer (c2c) online auction and shopping website. The original description of the use case reads as follows:

"GINIbay.com is a website which enables private actors to sell goods to one and other. Like other websites, trust is increased due to the fact that customers are able to rate their experience with a particular seller. However, GINIbay.com wants to go one step further and ensure that the personal attributes asserted by the sellers (e.g., professional qualifications of the seller, their credit history, …) are in fact reliable. On the other hand, GINIbay.com does not have the resources available to verify all the attributes asserted by their individual users."

One abuse case can be "outing". If a seller sells used books about discovering homosexuality and is outed by a co-worker in the company or friends/family get informed about it from a stranger. Such problems arise because the identity of the seller might be revealed.

The service provided by GINIbay.com is comparable to other existing c2c online auction and shopping websites. The value adding part of GINIbay.com is the integration of reliable personal attributes, which increases the trust relationship between buyer and seller. The main problem of the use case is the conflict of interests and sharply varying properties/requirements from different perspectives. From the seller's perspective, receiving the money from the buyer is the most important goal, independently from the transaction value. For that purpose, the sellers want to collect as much information as possible, regardless if the information is needed for the transaction or not. The buyers normally desire the diametrically opposed way. The provided information

should be as less as possible. The best case for buyers would be not to disclose or publish any information and still be able and trustworthy enough to buy every product. Even though this case is obviously not realistic, the GINI conceptual model is a user-centric model, which provides the user (the user is this case is the buyer) full control of her data to increase privacy. From GINIbay's view, apart from providing the service and charging the money from one or both sides, the service provider is a mediation party between buyers and sellers. They have to find a balance between the different interests of buyers and sellers and use policies to enforce the balance (E.g. Not to reveal the identity of the seller but still to guarantee the trustworthiness of the seller).

The seller has always to present attributes, which are directly related to the transaction (e.g. professionalism etc.). It means that the buyer can always request transaction-related attributes from the seller. These transaction-related attributes, which are proofed by third parties, increase the reliability of the seller. Considering the abuse case, the identity of the seller must not be revealed.

The buyer has to provide some attributes under special circumstances. One case is when the commitment to buy is particularly important and the value might be high, the buyer has to provide her identity. From this point of view, GINIbay.com added most value to transactions with high transaction volume (like buying a car or valuable jewellery) since the buyer can be sure about the attributes from the seller and the seller can be sure to receive the money, if the value of the transaction object is high.

The business model of the GINIbay service provider can be charging both parties for different transactions. One option can be to charge both sides with a basic fee based on the auction price. But normally the seller would be most willing to pay for an auction or product on the platform because it provides the seller a distribution channel. GINIbay can also charge both parties for requesting additional attributes such as credibility. In this case, the buyer would probably be most willing to pay for it. Considering that the buyer first pays and then the seller delivers the product or service, the buyer takes the risk. But there are also cases where the seller needs to check the buyer (E.g. transaction object is a car and the seller needs to know if the buyer is already of legal age).

5.3.1.1 Information flows within the INDI ecosystem

The following chapter puts the generally described use case of the previous chapter into the context of the INDI ecosystem. The template of the information flow model goes back to the GINI Deliverable D2.1 [Cau11]. This use case comprises two users. The first user represents the buyer with an independent User Agent A and the second user represents the seller whereas GINIbay.com acts as her User Agent (User Agent B).

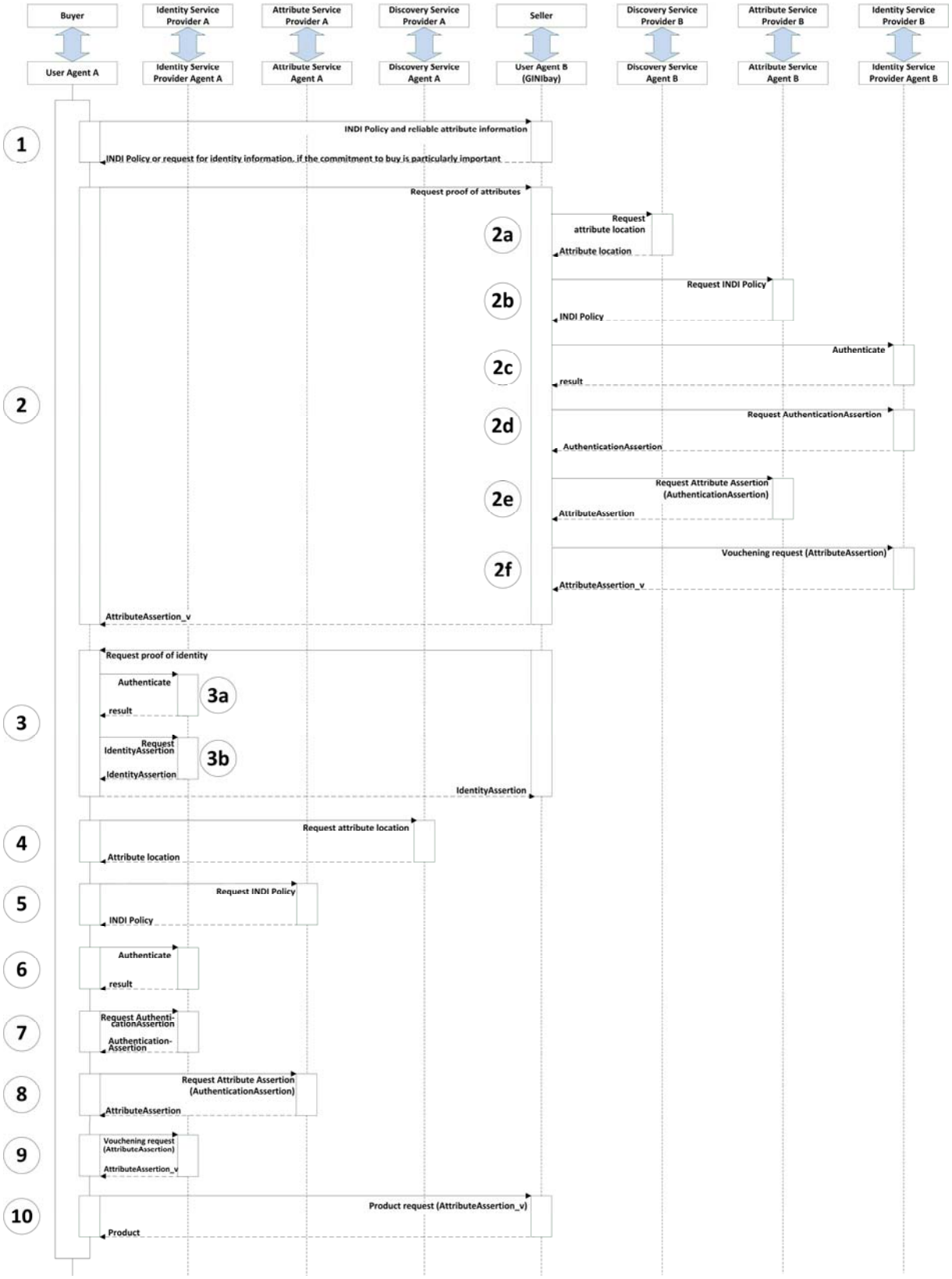

Figure 2: Person-to-Person transactions information flow.

The following process steps are carried out:

1. The User Agent A requests the INDI Policies of the User Agent B of the seller. The policies state that a certain set of attributes is needed to access the service or product. The buyer also requests reliable attributes from the seller.
2. The following steps 2a-f) are directly or indirectly controlled and/or supervised by the seller.
    a. The User Agent B contacts the Discovery Service B and retrieves information on where (at which Attribute Service B) the set of required attributes can be obtained.
    b. The User Agent B requests the Service Policy of the respective Attribute Service B. The policy states that an Authentication Assertion is required to access the service.
    c. The seller authenticates herself against the Identity Service Provider B. The authentication is mediated by the User Agent B.
    d. User Agent B requests an Authentication Assertion that fulfils the service policy requirements of the Attribute Service B contacted earlier.
    e. User Agent B provides the requested assertion to the Attribute Service B and requests the attributes. An Attribute Assertion containing the respective attributes is returned by the Attribute Service B.
    f. If no direct trust relation between the Attribute Service B and the User Agent A exists, the User Agent B sends the Attribute Assertion to the Identity Service Provider B. The Identity Service Provider B vouches for the authenticity of the Attribute Assertion and returns an extended version of the assertion.
    g. User Agent B sends the extended assertion to the User Agent A. Now the User Agent A has reliable attributes and the INDI Policies of the User Agent B.
3. The User Agent B is allowed to request for identity information, if the commitment to buy is particularly important (high value product and service). Depending on the commitment/transaction value this step is optional and can be left out.
    a. The buyer authenticates herself against the Identity Service Provider A. The User Agent A mediates the authentication.
    b. The User Agent A requests an Identity Assertion that fulfils the service policy requirements of the seller. The User Agent A sends the Identity Assertion to the User Agent B.
4. The User Agent A contacts the Discovery Service A and retrieves information on where (at which Attribute Service A) the set of required attributes can be obtained.
5. The User Agent A requests the service policies of the respective Attribute Service A. The policy states that an Authentication Assertion is required to access the service.
6. The buyer authenticates herself against the Identity Service Provider A. The User Agent A mediates the authentication.
7. The User Agent A requests an Authentication Assertion that fulfils the service policy requirements of the Attribute Service A contacted earlier.
8. The User Agent A provides the requested assertion to the Attribute Service A and requests the attributes needed to access the service or needed to buy the product. An At-

tribute Assertion containing the respective attributes is returned by the Attribute Service A.

9. If no direct trust relationship between the Attribute Service A and the User Agent B exists the User Agent A sends the Attribute Assertion to the Identity Service Provider. The Identity Service Provider vouches for the authenticity of the Attribute Assertion and returns an extended version of the assertion.
10. The User Agent A sends the extended assertion to the User Agent B.

All steps except step 2a-f) are directly or indirectly controlled and/or supervised by the user.

5.3.1.2 Extraction of Privacy Policies

The basis for this section is the information flow diagram of the previous section. The aim is to analyse each step in the diagram in order to identify possible privacy gaps. Privacy policies are used to close these gaps, thereby increasing users' privacy protection.

In the first step of this use case, User Agent A requests a product and the INDI Policy of the GINIbay agent. According to the minimum disclosure requirement, User Agent A is only allowed to send the request for the INDI Policy of GINIbay in the process:

- In the first process of the person-to-person transaction User Agent A SHALL only send a request, which only contains the request for INDI Policy of GINIbay and nothing more.

A policy must be set for requesting the reliable attribute information. As described in the use case, User Agent A (buyer) is always allowed to ask for reliable attributes, but it should not be allowed to ask for every attribute. For preventing the mentioned "outing" abuse case, the identity of the seller should not be revealed to the buyer as long as it is not required since GINIbay has the identity information of the seller and the transaction is mediated by GINIbay. The only case for revealing identity information is if the product should be returned to the seller.

- In private person-to-person transactions the identity of the seller MUST not be revealed to the buyer for the transaction in any case except when the buyer wants to return the product.

This policy is in line with the requirements of data minimization of Section 2.2.3. The elements of the privacy policy of Section 5.2 can be mapped onto it. In this policy the buyer is the data user, the action is revealing, the data is identity information about the seller, the purpose is the transaction process and the condition is "in any case except the buyer wants to return the product".

Given that the identity information of the user is not provided, other information is needed to increase the trustworthiness of the seller. In this case background information about the seller vouched by a third party can be used to reach the goal. According to Resnick and Zeckhauser [RZ01] trust or reputation can be created through context. For example, if a professional angler wants to sell his old fishing rod the trust on her description and quality of the product would be higher than if an inexperienced angler would do the same. A similar way to increase trust is to provide information about education history. If a motorcar mechanic sells her car, then the possibility is high that her description is precise and the car is in a good status. Context information could be also the credibility of the seller derived from previous transactions. The credibility can be shown in different ways. One possibility is to request a credibility assertion from an independent evaluation system where buyers can evaluate the seller. If the seller in this use case has a lot of good valuations, this will be reflected in a credibility assertion. Another possibility to get a credi-

bility assertion is to ask a third party (e.g. a company or an authority) to vouch for the credibility. But in this case, the reliability of the assertion depends on the entity that vouches for it. If a professional angler uses GINIbay for the first time, then the angler club, where she is a member of, can vouch for her. Based on this we can derive following privacy policy:

- The buyer SHALL only allow to request attributes about the seller's professionalism, education and/or credibility for increasing trustworthiness, if they are directly related to the object of the transaction.

In the same step the User Agent gets the INDI Policy of the User Agent B, which states a certain set of attributes which are needed to buy the product. Each product needs different attributes and so which attributes are requested in the INDI Policy is a case-to-case decision. If the transaction object is a car, the buyer has to proof the legal age, which is not required when the transaction object is a fishing rod. But in all cases we can find some common ground on which to base.

After the User Agent A has received the INDI Policy it will be displayed to the buyer. If the buyer accepts the INDI Policies she gives her consent and the process can be continued. If the buyer does not give the consent, then the process stops in this step. To ensure that the consent is obtained before requesting/processing personal data, a privacy policy must be set:

- The buyer MUST obtain her consent to the INDI Policy of GINIbay, before the transaction process continues.

For decision-making, the buyer must know what happens to her data. For that purpose, the user must be informed about the reasons why which attribute is needed to perform the transaction. On the basis of this information the user can decide either to buy the product or not.

- The INDI Policy MUST justify the request for each attribute and specify the purpose in a way that the user can understand it.

In some cases the purpose of data use may change. For example: Usually, the professionalism of the buyer is used for assessment of the reliability of the information. But if GINIbay wants to create statistics about average professionalism of the buyer on the platform, then purpose changes. In such cases a new consent must be obtained from the buyer to the new or changed usage.

- GINIbay MUST obtain a new consent from the buyer for data processing, if the original specified purpose of the data use changes, unless law obliges it.

Regarding the justify policy, the data minimization requirement becomes important. The seller can ask for every attribute in an INDI Policy as long as it can be reasoned. If the justify conditions are not limited, the seller can always state that the attributes are needed for (anonym) static analysis and evaluations. If a lot of sellers state this reason in their INDI Policies, the user might be substantially restricted in her choice when she does not want to provide information for static analysis and evaluations. To prevent this situation the requested attributes must be minimized to those attributes, which are only directly relevant for providing the requested product. One attribute could be the legal age if the transaction object is a vehicle or game/movie with an age limit. For some goods (like weapons) the buyer needs a licence (at least in some countries like Germany). In such case, the seller is allowed to request a reliable assertion, i.e. a certificate confirmed by an official authority. Another example is buying some books on the ban list (e.g. "Mein Kampf"/"My Battle"/"My Struggle" from Adolf Hitler). It is prohibited to buy such books, except they are used for research purposes. The buyer needs a reliable permission assertion, vouched by the university or research department to buy these books. The privacy policy fulfils the accountability requirement of the previous sections.

- User Agent B SHALL only allow to request for legal age, licences/certificates and/or permission assertion if it is required for the requested product.

In some cases the seller might request an official authority as Identity Service Provider. Like the weapon example, the seller must be sure that the licence or certificate is valid in the current country. An official authority has to act as Identity Service Provider, but this requirement does not affect the following processes.

For this use case, the seller is also allowed to request identity information of the buyer if the commitment to buy is particularly important. This is rather the case if the transaction value is high. The problem is to define the term "high". Since this use case is about private person-to-person auctions, the amount varies from case to case. The privacy policy does not set a static amount, but it should be adjusted individually.

- User Agent B SHALL allow to request for identity information of the user if the transaction value is higher than xxxxx €.

The two last policies fit with the data minimization requirement of Section 2.2.3.

Normally, information flow is not a high priority issue in a traditional private person-to-person transaction. A private person normally has no third party, which handles the accounting or billing for the transaction. For one transaction the seller probably would not set up a website or a shop on GINIbay. But in some cases one or more third parties must be informed due to legal requirements. One case would be the weapon deal again. According to the legal framework, the change of the owner of a weapon must be reported to authorities, which means that some information will flow to a third party. If an Internet domain changes the ownership, it must be also reported to the registry of the domains (e.g. Deutsches Network Information Center - DENIC[5]). The INDI Policy of the seller has to show the buyer which information will be transferred to which third party and for what reason. The INDI Policy of the third party must also be forwarded to the user and obtain her consent. If the buyer does not accept the INDI Policy of the third party, GINIbay is not allowed to forward any data to the third party and the transaction may stop here. These policies must be set to fulfil the transparency requirement.

- The INDI Policy of the seller MUST contain information about information flow to third parties, it has to declare which information will be transferred to which third party and for what reason.
- GINIbay MUST forward the INDI Policy of the respective third party to the buyer and obtain her consent, if data is forwarded. GINIbay MUST not forward any data to a third party without buyer's consent.

Another point must be also included in the information of the INDI Policies. Since the User Agent of the seller is GINIbay, GINIbay has to set up policies about storage time of personal data in person-to-person transactions. According to FIDIS Deliverable D14.8 [MG09], storage of personal data is a privacy threat if data storage temporally or quantitatively exceeds the user's consent. The data collection must be archived as soon as the purpose of the data is fulfilled. The purpose is fulfilled, if the seller has delivered the product because the user's consent for data processing ends with the end of transaction. In special cases, like legal issues, GINIbay has to store the data after the transaction has ended. For example the European Retention Directive 2006/24/EG [EU06] explicitly allows the provider the storage of communication details for tracing purposes in case of illegal activities or suspicion. After archiving the access to the data must be restricted to the responsible person and every action performed on the data must be logged.

[5] http://www.denic.de

- GINIbay MUST store the provided personal data of the buyer as long as the transaction has not ended. Collected data MUST be archived immediately when the transaction has ended, unless the buyer gives the consent for further usage and/or the legal framework obliges GINIbay not to archive the data.
- GINIbay MUST restrict access to the archived data to the responsible person only and log the access to the archived data.

The difference between this privacy policy and the policies before is the additional element obligation. This obligation enforces GINIbay to not delete the data collection if legal authority obliges to store it for a longer period. According to the trust assumption, we trust that the legal orders will be respected and the storage of the data does not need to be monitored.

In the whole second step the buyer requests attributes from the seller to prove the requested attributes. The User Agent B represented by GINIbay performs the sub-steps (a-f).

In sub-step 2a GINIBay contacts the Discovery Service Agent B to request the location where the required attributes can be obtained. For this request GINIbay has to forward some information about the requested attribute because the Discovery Service has to find the right location based on this information. The Discovery Service does not need to know for which purpose or transaction the attributes are needed, and it also does not need to know the requestor, in this case the buyer, of the attributes. A privacy policy has to ensure that the request only contains the requested attributes and nothing more.

- GINIbay SHOULD only forward the request of the attributes and nothing more.

Since the Discovery Services have their own lists of available Attribute Service Providers and the associated attributes, there is no need to forward information to any third party.

- The Discovery Service Agent B SHOULD not forward any information to any third parties in any case.

Another privacy policy needs to be set on the storage time of the request. This is comparable to the storage policy before. All data, provided by GINIbay, must be archived, when the transaction has ended. The transaction is over, when GINIbay received the Attribute Service Provider's location from Discovery Service Agent B. Of course, the Discovery Service provider has to follow legal orders. If the data is archived, the access to the data must be restricted to the responsible person and every action must be logged.

- Discovery Service Agent B MUST store the provided request data of GINIbay as long as the transaction has not ended. The request data MUST be archived immediately when the transaction has ended, unless the legal framework obliges GINIbay not to archive the data.
- Discovery Service Agent B MUST restrict access to the archived data to the responsible person and log the access to the archived data.

In the next sub-step 2b GINIbay contacts the Attribute Service Provider B and requests for the INDI Policy, which states that an Authentication Assertion is needed for access. In this step GINIbay is not permitted to provide any information. The reason is to match the minimum disclosure and purpose binding requirements. The purpose in this sub-step is to request the INDI Policy for further information about the required attributes for access. There is no need for disclosure of other data now.

- GINIbay SHOULD not forward any details to the Attribute Service Provider B as far as it has all required credentials for access.

In return, the INDI Policy of the Attribute Service Provider B is solely allowed to request for an Authentication Assertion because it is the only credential needed to verify the access permissions.

- The Attribute Service Provider B SHALL only state an INDI Policy for access purpose, which requests for an Authentication Assertion. The Justification of data request and specification of purpose must be articulated to the user in an understandable way.

Comparable to step one, the seller must give her consent for further processing.

- The seller MUST obtain her consent to the INDI Policy of Attribute Service Provider B, before the transaction process continues.

There is also the possibility that GINIbay has to contact more than one Attribute Service Provider. This is the case, if the required attributes are not available at one provider. In this case the Discovery Service Agent returns the location of more than one Attribute Service Provider and the sub-steps 2b to 2f have to be run through several times.

To get the Authentication Assertion the seller has to contact the Identity Service Provider. Therefore the sub-steps 2c and 2d are required and from a privacy policy perspective they can be seen as one step because it is the same interaction partner and both processes have the same purpose. First, the seller has to authenticate towards the Identity Service Provider Agent B itself (sub-step 2c). The technical site of the authentication process is not the focus of this deliverable. As we state in the trust assumptions that the process is secure and the seller can be uniquely identified. The result can be a Boolean variable. True is returned to the seller if the authentication was successful and false otherwise. In sub-step 2d GINIbay requests for the Authentication Assertion and the prerequisite for this process is a positive result of sub-step 2c. The Authentication Service Provider returns an Authentication Assertion to GINIbay.

Since these procedures are standard processes including the assumption of secure interaction, there are not many privacy policies to make, except storage time and information flow to third parties. Like the circumstances for Discovery Service Agent B, there is no need to forward information to third parties because the entire authentication and issuing process are performed internally. The data should also be archived after the end of the transaction except legal orders oblige not to archive them. The transaction regularly ends with sub-step 2d, when GINIbay gets the Authentication Assertion. In one exceptional case the transaction has ended with returning an extended Attribute Assertion. In this case, User Agent A and the Attribute Service Provider Agent B do not have a direct trust relationship and User Agent A demands a confirmation about the authenticity of the Attribute Assertion.

The privacy policies from sub-step 2a can be adopted almost unmodified.

- The Identity Service Provider Agent B SHOULD not forward any information to any third parties in any case.
- The Identity Service Provider Agent B MUST store the provided authentication information of GINIbay as long as the Authentication Assertion or the extended Attribute Assertion has not been returned. The information SHOULD be archived immediately when the Authentication Assertion or extended Attribute Assertion has been returned, unless legal order obliges Identity Service Provider Agent B not to archive the data.
- Identity Service Provider Agent B MUST restrict access to the archived data to the responsible person and log the access to the archived data.

In sub-step 2e GINIbay provides the Authentication Assertion obtained by Identity Service Provider B to the Attribute Service agent B together with a request for the required attributes. The Authentication Assertion confirms that GINIbay is permitted to request the attributes. The attribute request states which attributes GINIbay needs. The Attribute Service Provider collects the

requested attributes and packs them in an Attribute Assertion, which is returned to GINIbay.com. The Attribute Service Provider is not permitted to question the Authentication Assertion because according to the trust assumptions the seller trusts his User Agent (in this case GINIBay), controls every activity and the interaction processes between the INDI operators are secure.

- The Attribute Service Provider Agent B SHALL not question the access permissions of GINIbay, if the request contains an Authentication Assertion and all steps before are directly or indirectly controlled/supervised by the seller.

Privacy policies about information flow and storage are comparable to those for Discovery Service and Identity Service Provider Agents since the request can be performed completely internally and storage beyond the end of transaction is not needed due to the purpose binding requirement. The transaction has ended in this step, when identity attribute agent B returns the Attribute Assertion to GINIbay.

- The Attribute Service Provider Agent B SHOULD NOT forward any information to any third party in any case.
- The Attribute Service Provider Agent B MUST store the attribute request of GINIbay as long as the transaction has not ended. All requested information SHOULD be archived immediately when the transaction has ended, unless legal order obliges Attribute Service Provider Agent B not to archive the data.
- Attribute Service Provider B MUST restrict access to the archived data to the responsible person and log the access to the archived data.

The sub-step 2f is an optional step for vouching the Attribute Assertion. This step can be left out if a direct trust relationship (e.g. contractual trust or Attribute Service Provider Agent B also possesses attributes about the buyer and there were interaction between them before) between User Agent A and Attribute Service Provider Agent B exists. If there is no direct trust relationship, GINIbay has to send the received Attribute Assertion to the Identity Service Provider B and request a vouch for the authenticity. The Identity Service Provider Agent B returns an extended Attribute Assertion. This process does not need a privacy policy due to the privacy policies of the sub-steps 2c and 2d. The information flow privacy policy is still valid since the entire process can be performed within the Identity Service Provider B. The storage time of the data is certainly longer than without the vouching process, but this case is already be considered in the privacy policy. The data must be deleted if the extended Attribute Assertion is returned to GINIbay.

The next step is also an optional step. As we mentioned in the use case description, GINIbay is only allowed to request for identity information when the commitment to buy is particularly high. And for this, we derive a privacy policy. If this is not the case, these steps can be left out. In the optional step 3 GINIbay requests a proof of the Identity Assertion. But even the commitment to buy is particularly high and hence the value of the object is also high, the seller should not be permitted to request all identity information. The goal of this process is to show that the buyer is credible and trustworthy and with this revelation to increase the trust relationship between them. It is enough to disclose the name of the buyer and the address. For the requesting process, the following policy should be hold:

- Given that GINIbay is permitted to request identity information from User Agent A, it SHALL only request for name and address of the buyer.

The next sub-steps (3a and 3b) and final step of 3 (provider sends Identity Assertion to GINIBay) is similar to the processes we analysed in the sub-steps of 2. The same privacy policies must hold. The Identity Service Provider Agent A can perform the requested information within the

organisation, so the information does not need to leave the boundaries of the organisation. When the Identity Assertion is returned to the User Agent A the transaction has ended and the request information and the created Identity Assertion should be archived. The corresponding privacy policies are:

- The Identity Service Provider Agent A SHOULD not forward any information to any third party in any case.
- The Identity Service Provider Agent A MUST store the identity information request of User Agent A and the created Identity Assertion as long as the transaction has not ended. All referring information SHOULD be archived immediately when the transaction has ended, unless legal order obliges Identity Service Provider Agent A not to archive the data.
- The Identity Service Provider Agent A MUST restrict access to the archived data to the responsible person and log the access to the archived data.

Since GINIbay also gets the Identity Assertion we need a privacy policy to ensure the privacy of the buyer. The provided assertion is used for increasing trust between buyer and seller and proof of the identity, so the information must not leave the boundaries of GINIbay:

- GINIbay SHOULD not forward any information to any third parties in any case.

The storage time is here also a relevant point. Considering that the product has to be dispatched to the buyer when the seller receives the money and that identity information is needed for that, then policy for storage time is slightly different than for the Discovery Service Agent or Attribute Service agent. The time of archiving is still the end of the transaction. So the relative storage time stays the same but the absolute storage period is longer. The privacy policy for the storage time is equally worded, but the meaning of "end of transaction" is different.

- GINIbay MUST store the identity information of User Agent A as long as the transaction has not ended. The identity information SHOULD be archived immediately when the transaction has ended, unless legal order obliges GINIbay not to archive the data.
- GINIbay MUST restrict access to the archived data to the responsible person and log the access to the archived data.

Since this use case based on a person-to-person transaction, the steps 4 to 9 are comparable to the steps 2a to 2f because the position of the buyer is similar to the position of the seller. Both are private persons requesting reliable assertions for a transaction on GINIbay. The steps 2a to 2f describe the process of GINIbay getting the required Attribute Assertion and its interaction process with Discovery Service Agent B, Attribute Service agent B, and Identity Service Provider B. User Agent A has to do the same to get the required Attribute Assertion, but it interacts with Discovery Service Agent A, Attribute Service agent A, and Identity Service Provider A. Because of that, the privacy policies also stay the same. The request for specific attributes was already sent in the first step within the INDI Policy. The steps 4 to 10 will be performed if the User Agent A receives the Attribute Assertion and when the seller gets the optional proof of identity. The following table contains the adjusted privacy policies for steps 4 to 9. Besides that, the entities changed in many cases, the privacy policies by themselves do not change.

Step	Privacy Policies
4	• User Agent A SHOULD only forward the request of the attributes and noth-

	ing more. • The Discovery Service Agent A SHOULD not forward any information to any third party in any case. • Discovery Service Agent A MUST store the provided request data of User Agent A as long as the transaction has not ended. The request data SHOULD be archived immediately when the transaction has ended, unless the legal framework obliges User Agent A not to archive the data. • Discovery Service Agent A MUST restrict access to the archived data to the responsible person only and log the access to the archived data.
5	• User Agent A SHOULD not forward any details to the Attribute Service Provider A if it requests the INDI Policy. • The Attribute Service Provider A SHALL only state an INDI Policy for access purpose, which requests for an Authentication Assertion. The Justification of data request and specification of purpose must be articulated to the user in an understandable way. • The buyer MUST obtain her consent to the INDI Policy of Attribute Service Provider A before the transaction process continues.
6	• The Identity Service Provider Agent A SHOULD not forward any information to any third party in any case. • The Identity Service Provider Agent A MUST store the provided authentication information of User Agent A as long as the Authentication Assertion or the extended Attribute Assertion has not been returned. The information SHOULD be archived immediately when the Authentication Assertion or extended Attribute Assertion has been returned, unless legal order obliges Identity Service Provider Agent A not to archive the data. • Identity Service Provider A MUST restrict access to the archived data to the responsible person only and log the access to the archived data.
7	
8	• The Attribute Service Provider Agent A SHALL not question the access permissions of User Agent A, if the request contains an Authentication Assertion and all steps before are directly or indirectly controlled/supervised by the buyer. • The Attribute Service Provider Agent A SHOULD not forward any information to any third parties in any case. • The Attribute Service Provider Agent A MUST store the attribute request of User Agent A as long as the transaction has not ended. All request information SHOULD be archived immediately when the transaction has ended,

	unless legal order obliges Attribute Service Provider Agent A not to archive the data. • Attribute Service Provider A MUST restrict access to the archived data to the responsible person only and log the access to the archived data.
9	• The Identity Service Provider Agent A SHOULD not forward any information to any third party in any case. • The Identity Service Provider Agent A MUST store the provided authentication information of User Agent A as long as the Authentication Assertion or the extended Attribute Assertion has not been returned. The information SHOULD be archived immediately when the Authentication Assertion or extended Attribute Assertion has been returned, unless legal order obliges Identity Service Provider Agent A not to archive the data. • The Identity Service Provider Agent A MUST restrict access to the archived data to the responsible person only and log the access to the archived data.

Table 1: Person-to-Person use case related privacy policies for the steps 4 to 9.

In the last step 10, User Agent A requests the product. This request also includes the (extended) attribution assertion. In this case, the User Agent A should be only allowed to send the request with the (extended) attribution assertion (it fulfils the data minimization requirement) because with this information GINIbay has enough information to deliver the product.

- The User Agent A SHALL only send the product request and the (extended) attribution assertion and nothing more.

With the product request GINIbay respectively the seller can deliver the product. In this step, the privacy policy for GINIbay is very likely to the Identity Service Provider Agent B. The process can be completed internally and without leaving the boundaries of GINIbay and the seller itself. So any information flow to third parties is prohibited.

- The GINIbay SHOULD not forward any information to any third party in any case.

The storage time must also be considered. The deletion of the collected data should be carried out at the end of the transaction and the end is reached when the product is dispatched to the buyer. The privacy policy sounds again very similar to the policy for Identity Service Provider Agent B.

- GINIbay MUST store the provided request information of User Agent A as long as the product has not been dispatched to the buyer. The information SHOULD be archived immediately when the product has been dispatched, unless legal order obliges GINIbay not to archive the data.
- GINIbay MUST restrict access to the archived data to the responsible person and log the access to the archived data.

5.3.2 Use Case: Job-related attestations

This use case is also taken from GINI Deliverable D1.1. It is about getting reliable certificates of previous employers and visited education institutions with decreased paper work and high usability compared to the current situation. The original description of the use case reads as follows:

"Roberto is a Spanish citizen temporarily working in Belgium. During his stay in Belgium, he sees an opportunity to apply for a position in Finland. However, in order to apply he has to submit a certified copy of his grades and his degree. He also needs to provide certain attestations relating to prior work experience. In the current state of affairs Roberto must contact each institution (university, previous employer) separately and request them to provide him with a certified copy of this information. This is very time-consuming and requires much planning, orchestration and follow-up.

Using his INDI environment however, Roberto could simply request the authoritative sources which have to issue the necessary Attribute Assertions to the prospective employer."

The goal of the INDI ecosystem is to increase reliability and usability. In fact, the certificates are sent directly to the prospective employer, Roberto needs more effort to modify the documents (the case that the institutions sends the certified copies to Roberto's User Agent and it forwards them to the prospective employer will be discussed later). With the trust assumption that the processes are secure it should not even be possible to fake details on the certificates. From this perspective, the INDI model adds value because the trust in the evaluation of the user can be increased.

Roberto profits from the INDI ecosystem since he does not need to send a letter to every institution and request certified copies. The whole process can be accomplished faster because the certificates will be delivered digitally and the prospective employer receives the data within some minutes, while the "old fashioned way" needs significantly more time as the certified copies needs to be sent via normal mail. The value added here is the higher usability of own information and an optimized process.

If the certified copies of the records are sent directly to the employer, there might be a lack of privacy because Roberto's previous employer knows his new or desired employer, which Roberto might want to avoid. In the INDI ecosystem the new employer does not need to be revealed. Roberto has also privacy benefits when only necessary information is revealed. This case will be preferred because the User Agent just forwards the information to the new employer and Roberto does not need to step in. The usability from Roberto's perspective is not decreasing, but the privacy protection increases.

The business model can be charging both sides of the markets. Charging the user can be justified because he profits from usability and better privacy protection. The prospective employer wants to use the INDI ecosystem since the provided information about previous records is more reliable and on that base he can make better decisions. Both of them have an incentive to use the INDI ecosystem as long as the costs are proportional to the benefits (privacy, usability, trust) they gain and so both can be charged for using it.

5.3.2.1 Information flows within the INDI ecosystem

This job related attestation use case is a special case. Roberto wants attributes from university and his former employers, which means that these entities are likewise Identity and Attribute Service Providers because Roberto worked or studied there. They can act as Identity Service Provider

because in most cases they provide identity data such as e-mail addresses or other credentials for computer login or application to exams.

Figure 2 shows the information flow of the use case.

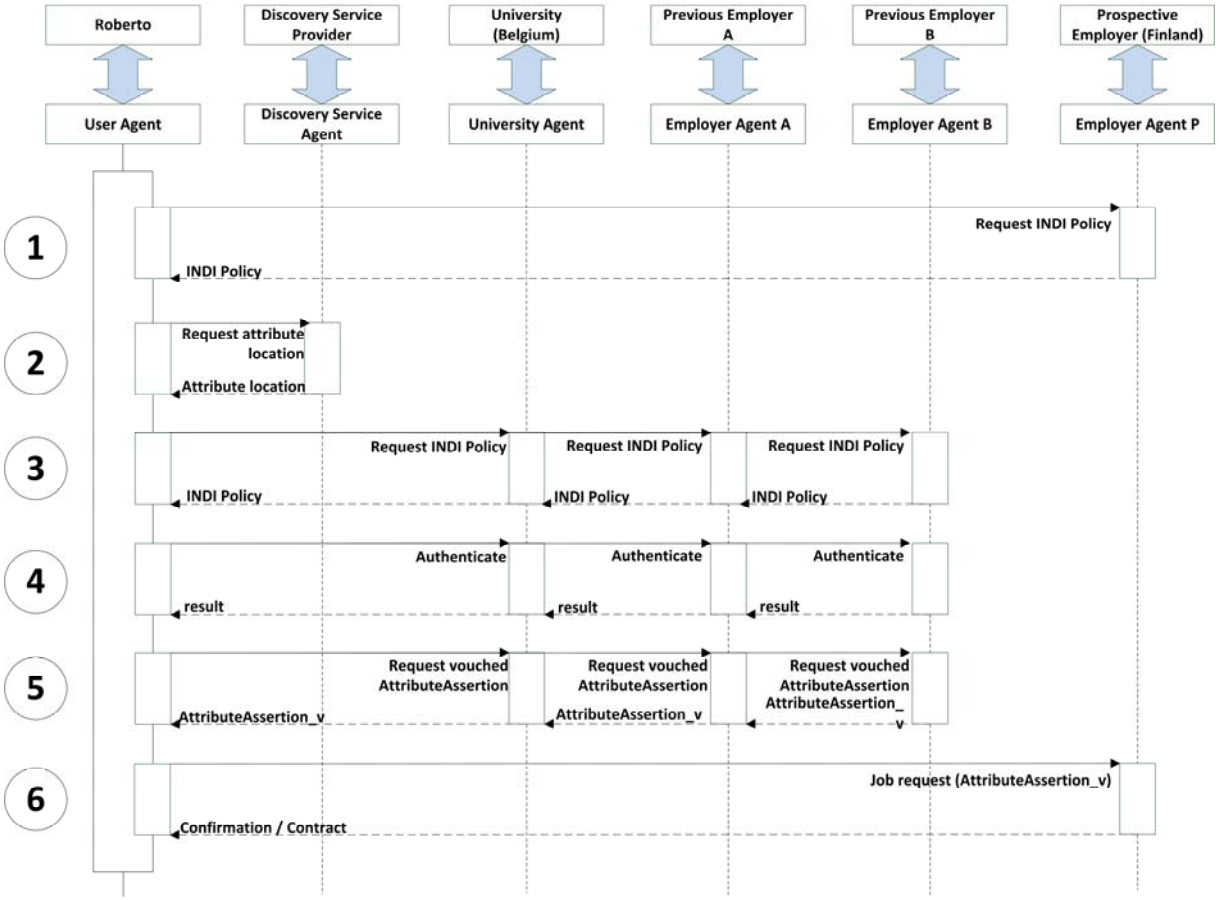

Figure 3: Job-related attestation information flow.

Compared to the first use case, the number of steps decreased. The main reason is that the steps three to five can be performed simultaneously. Roberto's User Agent can authenticate, request the INDI Policies and Attribute Assertions from different entities in parallel. Another reason is that only the prospective employer requests for attributes and Roberto does not. In this use case, there is no need for Roberto to request something from his employer.

The steps can be described as follows:

1. The User Agent requests the INDI Policies of the Employer Agent P of the prospective employer. The policies state that a certain set of attributes is needed to get a job interview or contract.
2. The User Agent contacts the Discovery Service and retrieves information on where (at which Attribute Services) the set of required attributes can be obtained.
3. The User Agent requests the service policies of the university and previous employers. The policy states that an authentication is required to access the attributes.
4. Roberto authenticates himself against the university and previous employers. The User Agent mediates the authentication.

5. After successful authentication the User Agent requests vouched attributes from the university's and the previous employers' agents, which are needed for the job interview or contract at the prospective employer. The contacted agents return multiple vouched Attribute Assertions containing the respective attributes.
6. The User Agent sends the vouched Attribute Assertions with a job request to the Employer Agent P.

All steps are directly or indirectly controlled and/or supervised by Roberto.

5.3.2.2 Extraction of Privacy Policies

The extraction of privacy policies is similar to the actions performed in Section 5.3.1.2. We take the use case as the basis for extraction of use case related privacy policies, which are fundamental for the generalized policies in subsequent sections.

In the first step Robert's User Agent requests the INDI Policy from his prospective employer. According to the assumptions about the trust relationship between user and the User Agent (e.g. the INDI operators act within the policy boundaries or the INDI operator delivers the promised functions and security mechanisms), there is only need for one policy: what is allowed to send. In this step Roberto and his User Agent want to know what is going to be needed in future steps if he is invited for a job interview. According to data minimization and purpose binding requirements it is not needed to reveal any information because Roberto has not applied for the position yet and maybe he decides not to apply for this job. The respective use case related privacy policy would be:

- The User Agent SHALL only send a request for INDI Policy in the first contact to the prospective employer and no other information in any case.

The storage time of the request should also be considered. Since the request does not contain a lot of information the information can be deleted or archived according to the legal requirements of the respective country. The information can be deleted if the requestor does not answer within a defined period of time (e.g. three month). If the requestor responds to an application with the required attributes the information should be archived with the application after the application process is completed. The application is completed, when Roberto signs the job contract.

- The Employer Agent P SHOULD delete the requested information if the requestor does not apply for the position within three month and if the requestor applies with the requested attributes, the request information MUST be archived when Roberto signs the job contract.
- The Employer Agent P MUST restrict access to the archived data to the responsible person and log the access to the archived data.

The Employer Agent P returns an INDI Policy containing a set of attributes required for the job interview or contract. According to the use case description, the attributes are limited to records and employment references. This means that the employer is only permitted to request for these attributes. Additionally, Roberto has to give his consent to the INDI Policy.

- Roberto MUST obtain his consent to the INDI Policy of the Employer's Agent P before the transaction process continues.

- The Employer Agent P SHALL only request certified copies of the university records and former employment references and it MUST specify the purpose of data use in a way that Roberto can understand.

The purpose of data use may change. Normally the certified copies of university records and former employment references are used to assess the abilities and skills of Roberto. But this information also enables the prospective employer background check. If this purpose is not specified in the INDI Policy, the prospective employer must obtain a new consent from Roberto.

- The Employer Agent P MUST obtain a new consent from Roberto for data processing if the original specified purpose of the data use changes, unless law obliges it.

In consequence of the user-centric GINI model, the user should also be informed about the information flow. This requirement is also stated in Section 2.2.2 on user control. The human resource department of the prospective Finnish employer might be outsourced to a third party. In that case, some data will be transferred to the third party. The employer also might have outsourced parts of their infrastructure to a cloud service (e.g. Amazon S3). The transaction process, or part of the process or parts of the process will be performed on the cloud server, which might be hosted in other countries with a different legal framework or data privacy act. Partly outsourced company structure is often used and it is not the focus of this deliverable to discuss about the privacy gaps in this company structure, these have to be accepted as given. We also cannot force the company to insource the outsourced parts, but for decision-making, Roberto must be informed about the information flow process before he sends his attributes. The information should contain which data flows to which third party, for what reason and what does it mean to the user's privacy.

- The Employer Agent P MUST inform Roberto about information flow to third parties and it must contain:
    - Which information is affected,
    - To which third party does the information flow (name and location of the third party),
    - What are the reasons for the transfer,
    - And how is the privacy of the user affected.

Another point, which must be considered if data flows to a third party, is consent. Before the respective employer can forward data to third parties, the INDI Policy of the third party must be forwarded to Roberto and his consent must be obtained.

- The Employer Agent P MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent, if data is forwarded. The Employer Agent P MUST not forward any data to a third party without Roberto's consent.

In the second step, Roberto's User Agent contacts a Discovery Service Agent. This step is a standard step and if the process is treated isolated from the other steps, the privacy gaps are the same like in the person-to-person use case. The User Agent contacts the Discovery Service Agent for the location of the required attributes. For that reason, the privacy policies stay the same, too.

- The User Agent SHOULD only forward the request of the attributes and nothing more.
- The Discovery Service Agent SHOULD not forward any information to any third party in any case.
- The Discovery Service Agent MUST store the provided request for data of the User Agent as long as the transaction has not ended. The transaction ends when the User

Agent receives the location information. The request data SHOULD be archived immediately when the transaction has ended, unless the legal framework obliges the Discovery Service Agent not to archive the data.

- The Discovery Service Agent MUST restrict access to the archived data to the responsible person and log the access to the archived data.

In the next step, the User Agent requests the INDI Policies of the University Agent, Employer Agent A and Employer Agent B. This step is comparable to, maybe even not the same, like the first step because they have the same goal. The User Agent requests the information about what is required for the access to the attributes. Because of that, the privacy policies stay the same.

- The User Agent SHALL send a request for the INDI Policy to the University Agent, Employer Agent A and B, but no other information in any case.

The privacy policy for the storage time has to be modified slightly. In this state of the application process we can assume that Roberto will continue because he already accepted the INDI Policies of the prospective employer and requests the required attributes for the application.

- The University Agent, Employer Agent A and B MUST store the request information as long as the contract has not been signed or Roberto has not been rejected. If the contract is signed or Roberto is rejected the request SHOULD be archived immediately.
- The University Agent, Employer Agent A and B MUST restrict access to the archived data to the responsible person and log the access to the archived data.

The University Agent, Employer Agent A and B return an INDI Policy containing a set of attributes, which is required for access. Because the university and the former employers act as Identity and Attribute Service Provider, they only need an authentication to verify the identity of the user. The INDI Policy is only allowed to request for an authentication.

- The INDI Policy of the University Agent, the Employer Agent A and B SHALL only request for authentication and no other information in any cases. The data use purpose MUST be specified before authentication.

Information of the authentication process should be also considered. Like e.g. the human resource department, the identity management infrastructure might be outsourced to a third party, too. If authentication is performed in cloud services or external companies the legal framework and the related privacy regulation may vary. Roberto should be informed about this fact and the privacy policy to ensure is the same policy as in the first step.

- The University Agent, Employer Agent A and B MUST inform Roberto about information flows to third parties and it must contain:
    - Which information is affected,
    - To which third party does it flow (name and location of the third party),
    - What are the reasons for the transfer,
    - And how is the privacy of the user affected.
- The University Agent, Employer Agent A and B MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent, if data is forwarded. The Employer Agent P MUST not forward any data to third party without Roberto's consent.

In the fourth step Roberto authenticates himself against the university and the previous employers and the User Agent mediates this process. In this process Roberto respectively his User Agent

has to present a predefined kind of credential and it will be proven by the corresponding entity. The result of the authentication will be returned to Roberto. To match the data minimization and purpose-binding requirement, the User Agent is only permitted to send authentication information like username and password since no more information is needed in this step.

- The User Agent SHALL only send authentication information (e.g. username and password) to the university and Employer Agents, but no other information in any case.

One other possible privacy gap in this step is the storage of the provided authentication details. Depending on the implemented authentication technology, the (encrypted) username and password might be transferred to the Relying Party. To increase privacy, the transferred data must be deleted when the validation is finished.

- The university and Employer Agents MUST store the provided authentication information of the User Agent as long as the authentication has not been validated. The information SHOULD be archived immediately when the authentication has been validated, unless legal order obliges not to archive the data.
- The university and Employer Agents MUST restrict access to the archived data to the responsible person and log the access to the archived data.

In the case that the authentication process is performed on outsourced servers, Roberto is already informed about it in step 3. We assume that Roberto accepts the information flow to third parties and he trusts that the university and/or former employers informed him about all facts.

In the fifth step Roberto requests for the vouched Attribute Assertions from the University Agent and the two former Employer Agents. The request contains the required certified copies of university records and employment references from employer A and B. The contacted agents return the requested assertions. According to the contextual separation, minimum disclosure and purpose-binding principle, the request should only contain information about what is requested. In this case, the only information allowed is the request for university records and employment references.

- The requests for Attribute Assertions sent by the User Agent SHALL only contain requests for university records and employment references and no other information in any cases.

Like in steps 3 and 4, the university and former employers A and B are not allowed to forward information to a third party because the process can be performed completely internally.

- The University Agent, Employer Agent A and B SHOULD not forward any information to any third party in any case.

For the creation of the certified copy of the required records and employment references, some data have to be collected and temporally stored. The created certified copies must be deleted after returning them to Roberto since the task is finished and in the current process the copies are not needed anymore. The only exception is when legal order obliges them to store it for a longer time.

- The University Agent, Employer Agent A and B MUST store the certified copies of records, employment references and the data needed for creation as long as the certified copies have not been returned to Roberto, which marks the end of the transaction. These SHOULD be archived immediately when the transaction has ended, unless legal order obliges not to archive the data.
- The University Agent, Employer Agent A and B MUST restrict access to the archived data to the responsible person and log the access to the archived data.

In the last step Roberto sends a job request with the certified copies of records, employment references and a job application to the prospective Finnish employer. Normally, if the prospective Finnish employer receives the certified copies, these need to be verified at the issuer. But we mentioned that this procedure decreases the privacy of Roberto because his former employers and university would know his new employer. Since the implementation of technologies is not in the focus and concern of this deliverable, we assume that Roberto cannot change the certified copies or the prospective employer will always detect modification.

The prospective Finnish employer receives the documents and they will start an evaluation process to decide if Roberto gets a job interview. The job request should only contain information required for the evaluation process. It is sufficient when the request includes the certified copies, cover letter and curriculum vitae. With this information, the prospective employer can assess the skills of Roberto and on that basis decide whether to invite him for an interview or not.

- The User Agent SHALL only send a job request containing the certified copies of records, employment references, cover letter and curriculum vitae. Any other information is not permitted in the job request.

The prospective Finnish employer receives the request and if the evaluation process can be done completely internally, forwarding information to third parties is prohibited. As mentioned before, the human resource department could be outsourced. In that case, forwarding information to a third party is inevitable and Roberto is informed in the very first step about it. In this process, a privacy policy must ensure that the provided information stays within the prospective employer's organisation and human resource department.

- If the applicant evaluation process is performed internally, the Finnish prospective employer SHOULD not forward any information to any third party in any case. If the human resource management is outsourced to a third party, then forwarding the provided information is allowed but the third party SHOULD not forward this information to other third parties in any case.

Like in the steps before, Roberto must be informed about information flow to third parties and his consent must be obtained before forwarding data.

- The Employer Agent P MUST inform Roberto about information flows to third parties and it must contain:
    - Which information is affected,
    - To which third party does it flow (name and location of the third party),
    - What are the reasons for the transfer,
    - And how is the privacy of the user affected.
- The University Agent, Employer Agent A and B MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent, if data is forwarded. The Employer Agent P MUST not forward any data to third party without Roberto's consent.

The storage time of the data in this process must be treated differently because there are three possible results of the evaluation. The first possible result is that the prospective employer invites Robert for a job interview and after that he gets the job and becomes an employee of the company. In that case, the data has to be stored and archived with strict access permission because an employer is legally obliged to have this information about its employees. Depending on the national legal framework, the employer has to store data for a defined period of time, even if Roberto has already left the company. It is also in line with the accountability requirement.

- In case that Roberto becomes an employee of the prospective Finnish employer, the employer MUST store the provided data as long as Roberto is an employee of the company. The permission to archive the data is granted after Roberto left the company, if the national legal framework obliges the company to store and archive it.

As mentioned before, data access must be restricted to fit the minimum disclosure privacy requirement and most of the people in the company do not need to access to this information. The person who needs access is the person in the human resource department, who is responsible for Roberto and his supervisor because he must be able to access Roberto's ability and skills. The access permission must be restricted to these two people/roles – the human resources employer and the supervisor.

- The access permission to Roberto's archived data MUST be restricted to his supervisor and the responsible persons in the human resource department.

The second possible result is that Roberto has been interviewed but the prospective employer rejected him after that. In that case, the data provided by Roberto is not needed after the rejection and the data must be deleted after the rejection.

- In case that the prospective Finnish employer rejects Roberto after the interview, the employer SHOULD delete the provided data as soon as Roberto receives the rejection, unless legal order obliges the employer to store the data.

The third and last possible result is that Roberto gets rejected immediately, i.e. he is not invited for a job interview. In that case the data must be deleted after Roberto receives the rejection. The privacy policy is quite similar to the second possible result; the difference is the time of deletion.

- For the case that the prospective Finnish employer rejects Roberto immediately, the employer SHOULD delete the provided data as soon as Roberto receives the rejection, unless legal order obliges the employer to store the data.

5.3.3 Use Case: Online Petition

The use case of "Online Petition" has been identified as important by the GINI-SA consortium for the INDI privacy considerations described in this deliverable. Again, this use case is taken from GINI Deliverable D1.1 [AnLe11]. In GINI Deliverable D1.1, the use case on online petition has been described the following:

"A group of concerned parents wishes to ensure that their children are taught creationism in their local schools. To this end, they have launched a petition in the hope of convincing the local-school board. However, in order to keep things fair they want to ensure that only residents can vote, and that every resident is able to sign the petition once."

Basically, for fulfilling this use case in a privacy respecting manner certain requirements must be met. In detail, users' or parents' full identity must not be revealed for fulfilment. This means, that e.g. the parents' names are completely irrelevant for signing the petition. For instance, the school or the city must only verify whether the persons who sign the petition are residents in the school's city or not. Additionally, a certain age level must also be assured. Furthermore, the parents should be able to vote for launching a lecture on creationism anonymously or pseudonymously respectively. Pseudonymously because parents are allowed to sign the online petition only once and this must be transparent for the online petition platform.

Applying the INDI environment for this use case has several advantages. Citizens can sign the petition in a privacy preserving way whereas the city or the online petition platform profits from

reliable attributes received from the citizen. Additionally, online petitions are cost-effective means for carrying out such fundamental rights of democracy. Hence, cities' costs can be decreased without compromising citizens' privacy.

Citizens profit from the INDI ecosystem because they do not need to leave their houses for signing the petition or need to take care on office hours of respective public authorities. Additionally, due to the privacy protecting INDI environment, citizens need not to disclose their complete identity compared to paper-based processes when showing a driving license, passport, or any other ID which usually incorporates a bigger set of identity attributes. Hence, citizens need not to be afraid of having to reveal any other identity information which is not important for the context of petition signing but could harm the citizen's personality or allow profiling.

The city or school, which organizes the online petition, benefits from INDI adoption as just an online platform needs to be set up compared to traditional petitions where people wait in offices to manually verify citizens' identities. This allows saving costs and offers citizens great flexibility and comfort. Due to the design of the INDI architectural model, the city's online platform can rely on and trust the attributes sent by the citizen for identification. To protect the citizens' privacy, only a proof of residence, age, and the information on having the petition not signed yet are required for identity verification. As discussed, this offers citizens enough data protection when using online petitions.

Concerning the business model, both parties - the city (or school) and the citizens – could be willing to pay for the services offered by the INDI environment. However, it is more likely that the organizer of the online petition will be more willing to pay because the benefits (quantitatively expressed by costs savings) are higher than for the citizen. The organizer of the petition (e.g. city or school) saves costs because just an online petition framework based on INDI needs to be set up by achieving the same or even better trustworthiness and reliability of citizens' attributes in comparison to traditional paper-based petitions. Nevertheless, also citizens could pay for the INDI services as they save time for signing the petition by not having the necessity showing up personally in an office. Additionally, their privacy is protected by disclosing only required attributes and leaving fewer possibilities for profiling.

5.3.3.1 Information flows within the INDI ecosystem

In this sub-section, the INDI model is applied to the online petition use case. The complete information flow is shown in Figure 4. In general, this information flow describes the interactions between a citizen and the online petition platform (city or school) and required steps in between (the communication with different INDI services or INDI operators).

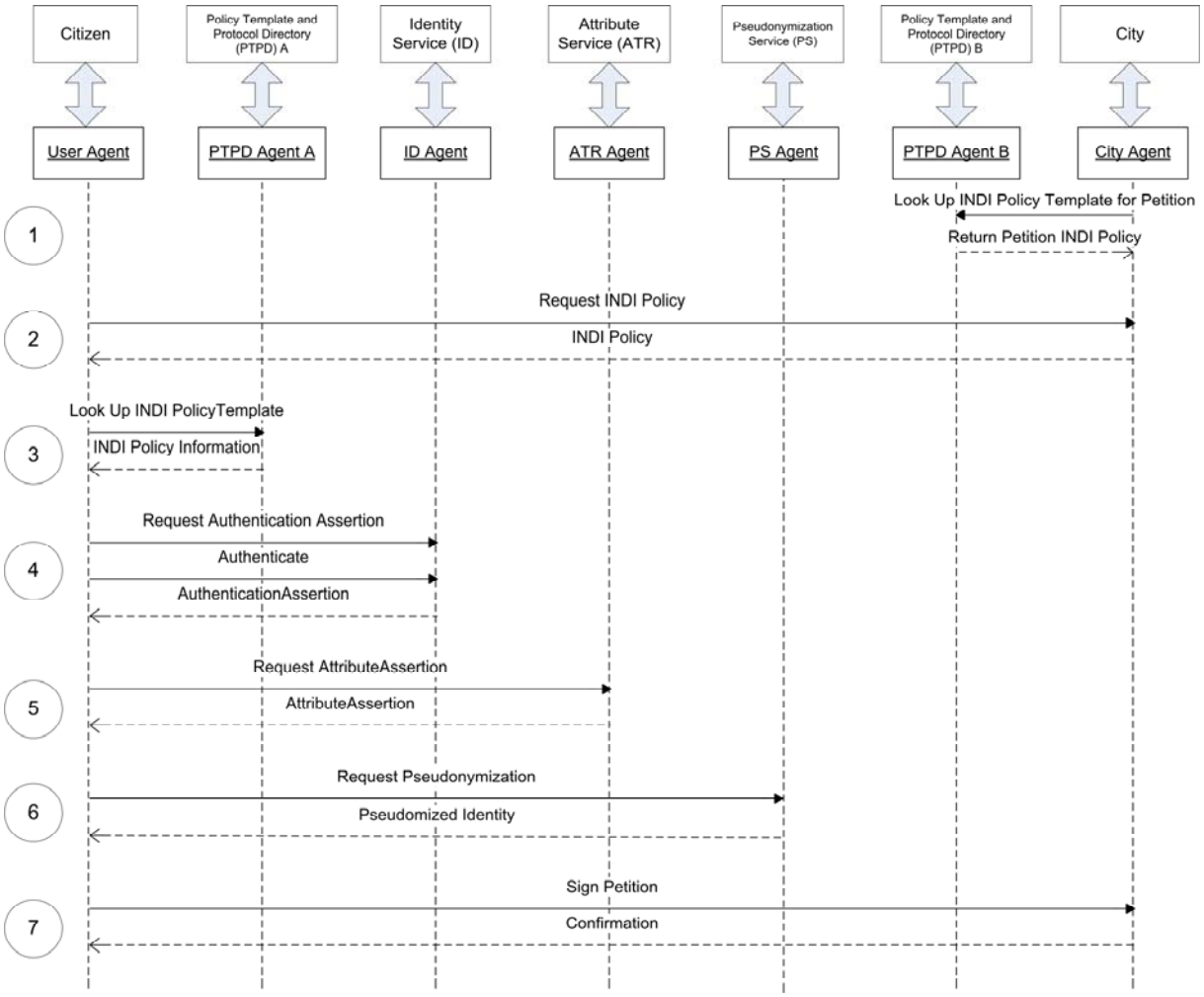

Figure 4: Online petition information flow

The information flow for online petition in detail:

1. The PTPD B has storeed various INDI Policies. Since online petitions are frequently used by cities, it also has stored INDI Policies for the setup of online petitions. In this use case, the city triggers its INDI operator to look up for an appropriate INDI Policy at the PTPD B for launching an online petition for creationism at schools. The result of this query constitutes an INDI Policy, which contains requirements and a set of attributes (residence, age, signing only once) required for signing the petition.
2. A parent or citizen wants to sign the online petition for a creationism lecture at schools. Therefore, she triggers her User Agent for requesting the according INDI Policy from the petition arranging city. The returned INDI Policy contains information on which reliable and approved attributes are required.
3. Using her User Agent, the citizen triggers her affiliated PTPD A to look up which services or operators must be contacted to retrieve all required attributes for fulfilling the INDI online petition policy from the city. By returning the desired information, the User Agent knows which operators or services must be communicated with.
4. In a first step, the received INDI Policy information states that for contacting the Attribute Service for attribute retrieval proper identification is required. Therefore, the citizen's User Agent requests an Authentication Assertion from the designated Identity Service or operator respectively. For successful identification, the citizen has to authenticate at the

Identity Service. The authentication mechanism to be used or the authentication process in general is out of scope in this information flow.

5. After successful authentication and identification, the Authentication Assertion together with an Attribute Assertion request is transmitted to the corresponding Attribute Service. The Attribute Service now fetches the attributes required for the INDI online petition policy (residence, age) corresponding to the claimed identity. After that, an Attribute Assertion containing those attributes is returned to the User Agent.
6. According to the INDI Policy of the city, citizens are allowed to sign the petition only once. To guarantee this requirement, the Pseudonymization Service is contacted to pseudonymize the citizen's identity. By the help of this service, the city only knows that the user has already signed the petition but not exactly who the user actually is.
7. In this step, the citizen presents her identity and attributes to the city's operator to sign the petition. Due to the trust relationship between the INDI operators, the user is allowed to sign the petition. If the action is successful some kind of confirmation information will be returned to the User Agent and the citizen respectively.

5.3.3.2 Extraction of Privacy Policy

Similar to the use cases discussed before, this sub-section extracts privacy policies according to the information flow for online petitions. Those extracted privacy policies build a solid basis for the design and the development of the generalized INDI Privacy Policy Framework in Section 5.4.

Referring to Figure 4, in the first step the city that wants to arrange an online petition queries its affiliated PTPD B for an appropriate INDI Policy. This operator hosts various INDI Policies for different use cases for the city. These INDI Policies constitute predefined policies, which e.g. contain information on what kind of data or what attributes are required from citizens in certain processes or how and under what circumstances transmitted data are processed by the city. For the online petition use case, the city queries its operator for an appropriate INDI online petition policy to electronically model the online petition for creationism lectures at school. According to the purpose binding and minimal disclosure principle, the INDI Policy for the arrangement of online petitions must only contain information requests really required for being able to successfully sign the petition.

- The INDI Policy for online petitions MUST only request the citizen's age, residence, and verifiable information for signing only once as personal attributes. No other, non-contextual information MUST be requested.
- The INDI Policy MUST contain information how the data will be processed and if third parties are involved.
- The user MUST give her consent to the INDI Policy of the online petition, before the transaction process continues.

The INDI Policy for signing the online petition is offered on the city's web site or the according online petition platform. If a citizen wants to sign the petition, the citizen's User Agent requests the corresponding INDI Policy from the city. For retrieving the INDI Policy, no personal information from the citizen is required.

- The User Agent SHALL only send a request for the desired INDI Policy of the city but no further information in any case.

In our example, the citizen and the User Agent, respectively, have a trust relationship with another PTPD Agent A. For the city, the PTPD Agent stores INDI Policy templates. In case of the user, it is assumed that the citizen has already signed a couple of online petitions before. Therefore, the according INDI Policies are stored at the affiliated PTPD Agent. If referring to the use case on job-related attestations of Section 5.3.2, after having received the INDI Policy in that example the User Agent actually queries a Discovery Service to get information on which Attribute Services must be contacted. However, in this example on online petitioning we assume that this information is already known and has been previously stored at the user's PTPD Agent. Hence, the User Agent just presents the city's INDI Policy for online petitions to the PTPD Agent and in turn receives information how this INDI Policy can be fulfilled. This information may contain details on which services must be contacted, in which order, or any other side constraints.

- The User Agent SHOULD only forward the INDI Policy to the PTPD Agent and no other information.
- The PTPD Agent SHALL not forward the received information to any other third party.
- The PTPD Agent MUST inform the user which attributes are required for fulfilling the INDI Policy and which actions must be taken.
- The PTPD Agent MUST store the policy look-up request only as long as the transaction has not ended.

Since the actions for fulfilling the policies have already been stored at the PTPD Agent, the User Agent knows that in the next step user identification is required. Therefore, the User Agent contacts an appropriate INDI Identity Service. At this service, authentication is required and the user is requested to present appropriate claims such as username and password.

- The Identity Service SHALL only request for authentication information and nothing more.
- The User Agent SHALL only send authentication information (e.g. username and password) and no other information to the Identity Service.
- The Identity Service SHOULD store the authentication information only during the authentication process. After credential verification, the presented credentials by the User Agent MUST be deleted immediately.

After having the user successfully authenticated at the Identity Service, the Identity Service transmits an Authentication Assertion back to the User Agent. This Authentication Assertion may contain information on when the user has successfully authenticated or which authentication mechanisms had been used. Additionally, the assertion must contain some kind of identifier to re-identify the user at the INDI Attribute Service contacted in the next step. The Attribute Service requires the identification information to retrieve the requested attributes for fulfilling the INDI online petition policy. Which kind of authentication and identity information is required for successfully querying the Attribute Service is usually stated in the INDI Policy of the Attribute Service. In the previous use cases, the INDI Policy of the Attribute Service has been requested before communication with the Identity Service. However, in this use case it is assumed that this policy has been already required at least in one transaction before and hence was stored at the PTPD Agent.

When triggering the Attribute Service, the Attribute Assertion request shall only contain the Authentication Assertion and the request for the personal attributes age and residence. Hence, this request follows the contextual separation, minimum disclosure, and purpose binding principles.

- The User Agent SHOULD only transmit the authentication information to the Attribute Service according to its policy.
- The Attribute Service SHOULD not forward any identity or authentication information to any other service or third party.
- The attribute request, sent to the Attribute Service, SHALL only contain requests for age and residence attributes and no other information.
- The Attribute Service MUST store the attribute request of the User Agent as long as the transaction has not ended. All request information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data.

One of the main requirements of the city for signing an online petition is that citizens are allowed to do so only once. Therefore, it must be possible to check if a citizen has already signed the petition. Hence, for fulfilling this requirement the simple presentation of age and residence attributes to the city's online petition platform is not sufficient. However, to achieve this requirement the User Agent contacts a so-called Pseudonymization Service after having received the authentication and Attribute Assertions. By the help of this service, a pseudonym for the citizen is generated or derived from an identifier. This pseudonym can be used at the online petition platform to check whether the citizen has already signed the petition or not. At this point it is important to mention that by presenting the pseudonym the citizen's identity will not be disclosed. The pseudonym is only useful for verifying the sign-only-once requirement but does not help when trying to find out user's real identity.

- The User Agent SHOULD not transmit more information than necessary to the Pseudonymization Service (e.g. identifier).
- The Pseudonymization Service SHOULD not forward the user's identity information to any other service or third party.
- The pseudonym calculated by the Pseudonymization Service MUST not contain any information on the user's real identity.
- The Pseudonymization Service MUST store the pseudonymization request of the User Agent and the identity data as long as the transaction has not ended. All request information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Pseudonymization Service to store the data.

Having successfully processed the previous steps, the User Agent now holds the citizen's pseudonym, age, and residence as user and identity information according to the city's INDI online petition policy. In the last process step, the citizen sends this information to the city's online petition platform to sign the petition. If the signing process was successful, the citizen receives some kind of confirmation message. The city stores these data for counting the number of petition signers. Since petitions usually can only be signed during a certain period, the stored data should be deleted after the petition signing phase ends. If the city wants to use the citizen's data for some kind of analysis or evaluation (e.g. how many citizens of a specific age range have signed the petition) this desire must have been stated in the INDI Policy before.

- The User Agent SHALL only send the pseudonym, age, and residence to the online petition platform of the city.
- The city SHOULD not forward the user's information to any other service or third party if not explicitly stated in the INDI Policy.

- The city MUST store the signing request of the User Agent as long as the signing period has not ended. All request and user information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data. Otherwise, if data wants to be stored for a longer period this MUST be stated in the INDI Policy.

5.3.4 Use Case: Renewal of Authoritative Documents

The use case of "Renewal of authoritative documents" defines the last INDI use case to be taken as a basis for general privacy policies extraction. According to GINI Deliverable D1.1 [AnLe11], the use case has been described as follows:

"Citizens usually are assigned many authoritative documents (passport, ID-card, driving license, student attestation, social welfare attestation, etc.) and special-purpose smart cards (library card, fitness centre membership card, etc.) with limited time of validity. Usually one document/card is used as an attestation of certain citizen attributes which are needed to renew another document."

In current traditional and paper-based processes, citizens are usually forced to visit public authorities' offices personally for authoritative document renewal. However, in the information technology ages these processes are not up-to-date anymore. To fill this gap, applying the INDI environment can help both citizens and public authorities to save time and costs.

Citizens benefit because they do not have to care about public authorities' offices opening hours and are able to renew their documents around the clock. Additionally, because of the privacy friendly design of INDI and depending on the document to renew the citizens must not provide or disclose all their personal information. For example, the citizen's religious affiliation is totally irrelevant for renewing a driving license. In this sample use case, citizens also need not to keep in mind any document validity expiration date. This and which information is required for document renewal is completely managed by the INDI environment.

In addition, by adopting the INDI environment also public authorities can profit when documents can easily be renewed online. Costs can be saved because personnel can be pared down due to less necessity for office opening hours. Furthermore, the document renewal process can be implemented more transparent for the citizen. Out of the INDI Policy the citizen can see which data are required for document renewal and how they are processed. Alternatively, more service comfort can be offered to citizens if e.g. a reminder e-mail is sent out when the document is going to expire soon.

In this use case it is obvious that both sides (citizens as well as public authorities) could be charged for using the INDI environment because both sides could benefit from higher cost savings rates. Citizens could be willing to pay because they save time and money but additionally can rely on trustworthy and privacy preserving online processes. Public authorities can offer more citizen friendly online processes and at the same time can decrease personnel costs because less staff must be present during opening hours.

5.3.4.1 Information flows within the INDI ecosystem

Equally to the previously described use cases the INDI ecosystem is applied to the authoritative document renewal process in this sub-section. It shows the information flow between a citizen, a public authority, and of all additionally involved parties and services in more detail. Figure 5 illustrates this information flow in detail.

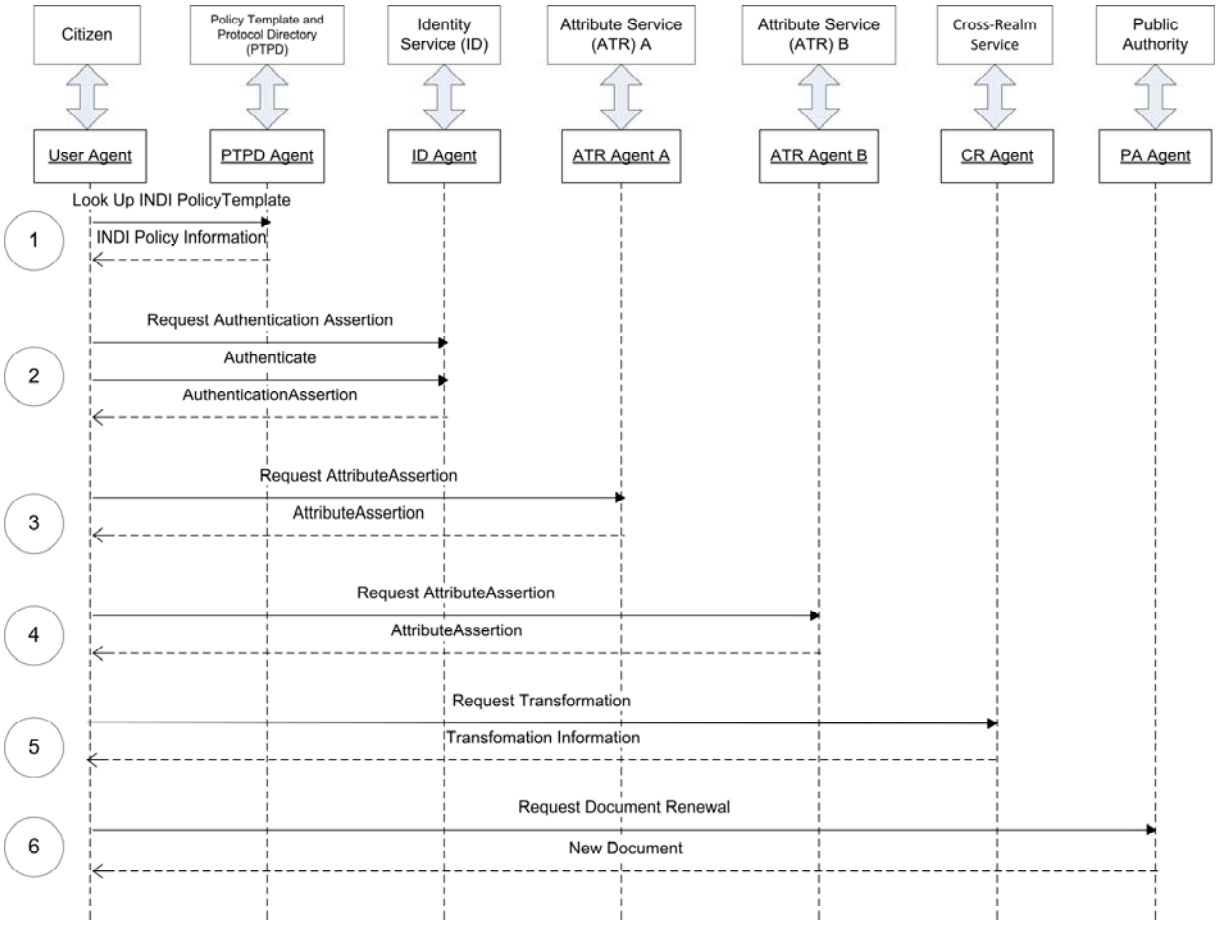

Figure 5: Renewal of authoritative documents information flow

1. Basically, in this use case it is assumed that the INDI environment takes care on document validity expiration dates. This means that the citizen is either reminded by the public authority applying INDI or the User Agent. In this example, we assume that an authoritative document, e.g. a driving license, needs to be renewed soon and the citizen receives a reminder message from her User Agent. To launch the renewal process, the citizen does not need to fetch necessary information from the public authority but instead just authorizes the User Agent to launch the process. The User Agent can retrieve all information required from the affiliated PTPD. This directory has stored various INDI Policy templates including the specific one required for authoritative document renewal. This INDI Policy contains all information on e.g. which personal attributes or other documents are required for fulfilling this process and where this information can be retrieved.
2. Due to the received INDI Policy the User Agent knows which process steps must be triggered for fulfilling the document renewal request. As a first step of the INDI Policy, secure identification and authentication of the citizen is required. Hence, a request for an Authentication Assertion is sent to a trustworthy Identity Service. In our example, we do not specify details on the authentication process to be carried out but since for document renewal sensitive data are processed, the authentication mechanism must follow high security and quality standards (e.g. two-factor authentication mechanisms). The Authentication Assertion returned to the User Agent contains the citizen's unique identity as well as additional authentication information.
3. This unique identity information is further used for attribute retrieval at an Attribute Provider. For instance, in our example we assume that for renewing a driving licence a proof

of residence is required. Thus in this step, a copy of such proof is requested from Attribute Service A. This proof can be issued because of identity and authentication information is transmitted by the User Agent. The returned Attribute Assertion includes the proof of residence information.

4. In this process step, a second Attribute Service is contacted. We assume that the new driving license should contain the new family name of the citizen who got married a couple of weeks ago. Therefore, the request to the second Attribute Service B contains a request for a marriage certificate. We further assume that the marriage certificate was issued by a foreign government as the wedding took place in another country. To achieve equality between foreign and domestically issued marriage certificates a Cross-Realm Service is contacted in the next step.
5. The User Agent requests a proof of equality for the marriage certificate by communicating with the Cross-Realm Service. The Cross-Realm Service confirms equality and returns any transformed information of the marriage certificate.
6. Having received the transformed information, the citizen contacts the public authority for requesting the authoritative document (driving license) renewal. The request contains the citizen's identity information, the proof of residence, as well as the marriage certificate.

As can be seen by the sequence diagram in Figure 5, the user is always in control which information and data are transferred because every process step is routed through the citizen's User Agent.

5.3.4.2 Extraction of Privacy Policy

The aim of this sub-section is to extract privacy policies from the designated use case of authoritative document renewal. The extracted policies help in developing a GINI Privacy Policy Framework, which will be shown in the next section.

The general idea behind this use case is document renewal without any media breaks. This means for example that a new issued driving license can be fully requested electronically by providing all necessary documents in a digital format. No personal visits in public authorities' offices or the provision of paper-based documents are required. To achieve this goal, both the citizen and the public authority that issues new driving licenses rely on the INDI ecosystem. The information flow for document renewal has been illustrated in Figure 5. We take this information flow as a basis for the extraction of privacy policies.

The processing of the use case starts by contacting a PTPD. This directory has stored the INDI Policy for authoritative document renewal. The INDI Policy contains information on which attributes are required for document renewal and where those attributes can be retrieved. For our INDI use case we assume that the citizen wants to renew her driving licence (having been reminded by the User Agent that the validity would expire) and to additionally change her family name in it because the citizen got married a couple of weeks ago. To get the desired information for renewal, the User Agent requests the appropriate INDI Policy from the PTPD Service.

- The User Agent SHALL only send a request for the desired INDI Policy but no further information in any case.
- The PTPD Agent SHALL not forward the received information to any other third party.
- The PTPD Agent MUST inform the user which attributes are required for fulfilling the INDI Policy and which actions must be taken.

- The PTPD Agent MUST store the policy look-up request only as long as the transaction has ended.
- The user MUST give her consent to the INDI Policy before the transaction process continues.

The INDI Policy contains all information on which operations must be carried out for successfully fulfilling the request. Hence, the next step foreseen by the policy is secure citizen identification and authentication. Which authentication mechanisms should be used and which Identity Service should be contacted for that is also stated in the INDI Policy. However, certain requirements should be considered in this process step.

- The Identity Service SHALL only request for authentication information and nothing more.
- The User Agent SHALL only send authentication information (e.g. username and password) and no other information to the Identity Service.
- The Identity Service SHOULD store the authentication information only during the authentication process. After credential verification, the presented credentials by the User Agent MUST be deleted immediately.

As response from the Identity Service the User Agent receives an Authentication Assertion including the citizen's identity as well as additional more detailed information on the conducted authentication process. This assertion is further used for communication with different Attribute Services. Regarding to our information flow diagram in Figure 5, two separate Attribute Services are contacted in the next steps. For simplicity we have assumed that the issued Authentication Assertion is sufficient for both services for attribute or document retrieval. The first Attribute Service issues or transmits a proof of residence certificate, which is required for driving license renewal. The second Attribute Service is contacted for providing a marriage certificate for the family name change in the new issued driving license. For both requests the following privacy policies apply:

- The User Agent SHOULD only transmit the authentication information to the Attribute Service according to its policy.
- The Attribute Service SHOULD not forward any identity or authentication information to any other service or third party.
- The attribute request sent to the Attribute Service SHALL only contain requests for a proof of residence or a marriage certificate and no other information.
- The Attribute Service MUST store the attribute request of the User Agent as long as the transaction has not ended. All request information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data.

For the duration of the transaction the User Agent stores both certificates. As described in Section 5.3.4.1, the marriage certificate was issued by a foreign government since the wedding had been taken place in a foreign country. To guarantee equality between the foreign marriage certificate and domestic-issued certificates, the Cross-Realm Service is contacted for transformation. This service proves equality and returns any required transformation information to the User Agent.

- The User Agent SHOULD only transmit the marriage certificate to the Cross-Realm Service and no other information.

- The Cross-Realm Service SHOULD not forward the user's marriage certificate or any included information to any other service or third party.
- The transformed information SHOULD only contain information related to the marriage certificate and no other information.
- The Cross-Realm Service MUST store the marriage certificate as long as the transaction has not ended. All information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data.

If all the previous steps have been processed successfully the User Agent sends a request for driving license renewal to the according public authority. The request contains the citizen's identity information, the proof of residence, and the marriage certificate. By the help of this information the public authority is able to renew or issue a new driving license to the citizen. The issuance of the driving license will be processed in the back-office facilities of the public authority. Finally, the citizen receives her new driving license (including the new family name) with a new expiration date. The driving license is wrapped in the return message to the User Agent.

- The User Agent SHALL only send the user's identity information, proof of residence, and marriage certificate to the public authority.
- The public authority SHOULD not forward any user information to any other service or third party if not explicitly stated in the INDI Policy.
- The public authority MUST store the document renewal request of the User Agent as long as the renewal process has not ended. All request and user information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data. Otherwise, if data wants to be stored for a longer period this MUST be stated in the INDI Policy.

5.4 Generalization of extracted Privacy Policies

For now, we have analysed four use cases and we have collected sufficient information of respective use case-related privacy policies to derive generalised privacy policies and form them to a framework.

Since the INDI ecosystem is a very common and generic construct, the Privacy Policy Framework must be as common and generic as the INDI ecosystem itself. From this perspective, the goal of the generic Privacy Policy Framework is to articulate a set of privacy policies, which can be applied on the most common use cases to ensure the privacy of the user (e.g. avoiding profiling or collecting information). For that reason we chose these four use cases, which include the most common elements. Based on that it is possible to transfer and apply the gained knowledge and constructed privacy policies to other use cases. To make this framework applicable and adoptable to as much use cases and situations as possible, it is necessary to formulate the high-level privacy policies as generic as possible. At best, the Privacy Policy Framework can be applied to all current use cases. But considering the quantity of variables and factors affecting the nearly infinite amount of use cases, it is unlikely that the presented generic Privacy Policy Framework can be applied to all (future) use cases without modification. This framework should be seen as a basis and starting point for further development, which already can be applied to most of the current use cases. In the future, new generic privacy policies should be added. The goal is to find a universally valid Privacy Policy Framework for all conceivable and possible use cases through an iterative finding and development process.

Considering all use cases, some recurring and similar procedures as well as information flows can be observed. For example: every User Agent has to contact the Discovery Service since only they know where the required attribute can be found. Or every User Agent has to contact the transaction partner for requesting the INDI Policy. Due to the similarity of procedures and information flows, similar privacy policies have been defined for them.

5.4.1 Reveal Information/Attributes

Use Case	Reveal Information/Attributes
Person-to-Person Transaction	In private person-to-person transactions the identity of the seller MUST not be revealed to the buyer for the transaction in any case except the buyer wants to return the product.
	Given that GINIbay is permitted to request identity information from User Agent A, it SHALL only request for name and address of the buyer.
	User Agent B SHALL only allow to request for legal age, licences/certificates and/or permission assertion if it is required for the requested product.
	The buyer SHALL only allow to request attributes about the seller's professionalism, education and/or credibility for increasing trustworthiness, if they are directly related to the object of the transaction.
	User Agent B SHALL allow to request for identity information of the user if the transaction value is higher than xxxxx €.
	The Attribute Service Provider B SHALL only state an INDI Policy for access purpose, which requests for an Authentication Assertion. The justification of data request and specification of purpose must be articulated to the user in an understandable way.
	Given that GINIbay is permitted to request identity information from User Agent A, it SHALL only request for name and address of the buyer.
	The User Agent A SHALL only send the product request and the (extended) attribution assertion and nothing more.
Job-related Attestation	The Employer Agent P SHALL only request certified copies of the university records and former employment references and it MUST specify the purpose of data use in a way that Roberto can understand.
	The requests for Attribute Assertions sent by the User Agent SHALL only contain requests for university records and employment references and no other information in any cases.
	The requests for Attribute Assertions sent by the User Agent SHALL only contain requests for university records and employment references and no other information in any cases.
	The User Agent SHALL send a request for INDI Policy to the University Agent, Employer Agent A and B, but no other information in any case.
	The INDI Policy of the University Agent, the Employer Agent A and B SHALL only request for authentication and no other information in any cases. The data use purpose MUST be specified before authentication.
	The INDI Policy for online petitions MUST only request the citizen's age, residence, and verifiable information for signing only once as personal attributes. No other, non-contextual information MUST be requested.
nli ne Pe ti- tio	The PTPD Agent MUST inform the user which attributes are required for fulfilling the INDI Policy and which actions must be taken.

	The User Agent SHALL only send authentication information (e.g. username and password) and no other information to the Identity Service.
	The attribute request sent to the Attribute Service SHALL only contain requests for age and residence attributes and no other information.
	The User Agent SHOULD not transmit more information than necessary to the Pseudonymization Service (e.g. identifier).
	The pseudonym calculated by Pseudonymization Service MUST not contain any information on the user's real identity.
	The User Agent SHALL only send the pseudonym, age, and residence to the online petition platform of the city.
	The User Agent SHALL only send a request for the desired INDI Policy of the city but no further information in any case.
	The Identity Service SHALL only request for authentication information and nothing more.
	The User Agent SHALL only send authentication information (e.g. username and password) and no other information to the Identity Service.
Renewal of Authoritative Documents	The User Agent SHOULD only transmit the authentication information to the Attribute Service according to its policy.
	The attribute request sent to the Attribute Service SHALL only contain requests for a proof of residence or a marriage certificate and no other information.
	The User Agent SHOULD only transmit the marriage certificate to the Cross-Realm Service and no other information.
	The transformed information SHOULD only contain information related to the marriage certificate and no other information.
	The User Agent SHALL only send a request for the desired INDI Policy but no further information in any case.
	The Identity Service SHALL only request for authentication information and nothing more.
	The User Agent SHALL only send the user's identity information, proof of residence, and marriage certificate to the public authority.

Table 2: Reveal Information/Attribute related Use Case Privacy Policies.

Table 2 lists all use case privacy policies related to revelation of information or attributes. Limited attribute revelation is one of the most important factors in the GINI vision and one of the most critical factors in privacy research. In 2000 Sweeney [Swe00] worked out that 87% (216 million of 248 million) of the U.S. population could be unambiguously identified by their five-digit zip code, gender and date of birth. Over half of the population (132 million of 248 million or 53%) could be uniquely identified by place, gender and date of birth, where place is the main residence of the people. 18% of the U.S. population still could be uniquely identified by state, gender and date of birth. This study illustrates that every disclosed data is or can be a potential threat to individual's privacy. Even if the data is not directly identifiable, the combination of indirect identifiable data can result in a direct identifiable profile of an individual. The result and impact of the study get stronger nowadays since social networks such as Facebook are taking a big part in social life. Due to that, information used by Sweeney to identify people is easily available and provided by the user on their social network profiles.

Therefore, it is essential to minimize the revealed attributes to a minimum. In most of the transactions, attributes must be revealed or otherwise the transaction cannot continue. For example, in the "Job-related Attestation" use case Roberto has to provide his records and/or certificates

from his university and/or previous employees to get the job. In the "online petition" use case the citizens have to proof their age and residence to be authorized to participate in the petition. From this point of view, for most of the cases the request for attributes cannot be forbidden, but it has to be reduced to a minimum set of attributes, which are directly relevant and required for further process of the transaction. These attributes must be predefined and articulated to the user. The problem is that either the user or the User Agent has to decide at every request, if the requested attributes are the minimum set of required attributes. Another possibility would be that some independent institution verifies it. In the former case the usability will be decreases significantly and maybe the user has insufficient knowledge to evaluate whether the Relying Party is really requesting the minimum set of required attributes necessary. In the latter case the user must trust the independent third party. Both cases are possible but user experiences must show, which way is the better one. So for simplicity, we assumed, that the users decide if the requested set of attributes is the minimum set. The generic privacy policy is formulated as follows:

- The Relying Party SHALL only request attributes which are directly relevant and required for the current transaction, and no other information in any case.

In return, the User Agent must only send the set of attributes, which were requested by the Relying Party.

- The User Agent SHALL only send the minimum set of attributes, which were requested by the Relying Party and no other attributes/information.

A comparable privacy policy should be applied to the Identity Service and Attribute Provider. The Identity Service Provider should issue an Identity Assertion, which confirms the identity of the user. This assertion should only contain the identity confirmation and no other information about the user. This also applies for Attribute Service Providers. If the user requests an Attribute Assertion, the assertion should only contain the requested attribute and no other information.

- The Identity Service Provider SHOULD only provide an Identity Assertion to the User Agent, which fulfils the requirements of the INDI Policy of the respective Relying Party, and no other information.

- The Attribute Service Provider SHOULD only provide an Attribute Assertion to the User Agent with the minimum set of Attributes, which were requested by the Relying Party, and no other attributes/information.

Attributes with direct and unique identifiable characteristics have a special position. Examples of such attributes are the mobile phone number (since the mobile phone number and the user have normally a 1:1 relationship [SCV08]), social security number or bank account information. There might be some cases where this information is absolutely necessary for the transaction, for example if the customer wants to extend her mobile phone contract. Also in the "Person-to-Person Transaction" use case, the request of identity information from the seller's side is allowed, if the commitment to buy (respectively the value of the goods) is very high. This means, that the request of identity information/direct identifiable information needs predefined policies with clear stated circumstances, where it is allowed to request for identity information/direct identifiable information.

- The respective Agent SHALL only request for or send identity information/direct identifiable information in situations, where this information is required and indispensable. In any other case identity information/direct identifiable information MUST not revealed.

In all four use cases it can be seen that the first request is widely used to get the INDI Policy of the Relying Party. This step is needed to inform the user, which information and attributes will be revealed and required to complete the transaction. At that time and based on the received INDI Policy, the user can still decide if she wants to continue or cancel the transaction. Since the user has not decided on further interaction, no more information should be sent, when the Agent contacts another party to get the INDI Policies, even if it is possible to expedite the process if more information is sent in this request.

- The respective Agent SHALL only send a request for the INDI Policy of the Relying Party and no other information in any case, if this is the sole purpose of the request and the user has not decided to continue the transaction.

In all use cases, an Identity Service Provider must be contacted for an Identity Assertion. In this authentication process, the only information transmitted to the Identity Service Provider is authentication information required to ensure the identity of the user:

- The User Agent SHOULD only send the authentication information, which is required to ensure the identity of the user and no further information or attributes.

5.4.2 Storage

Use Case	Storage
Person-to-Person Transactions	GINIbay MUST store the provided personal data of the buyer as long as the transaction has not ended. Collected data MUST be archived immediately when the transaction has ended, unless the buyer gives the consent for further usage and/or the legal framework obliges GINIbay not to archive the data.
	Discovery Service Agent B MUST store the provided request data of GINIbay as long as the transaction has not ended. The request data MUST be archived immediately when the transaction has ended, unless the legal framework obliges GINIbay not to archive the data.
	The Identity Service Provider Agent B MUST store the provided authentication information of GINIbay as long as the Authentication Assertion or the extended Attribute Assertion has not been returned. The information SHOULD be archived immediately when the Authentication Assertion or extended Attribute Assertion has been returned, unless legal order obliges Identity Service Provider Agent B not to archive the data.
	The Attribute Service Provider Agent B MUST store the attribute request of GINIbay as long as the transaction has not ended. All requested information SHOULD be archived immediately when the transaction has ended, unless legal order obliges Attribute Service Provider Agent B not to archive the data.
	The Identity Service Provider Agent A MUST store the identity information request of User Agent A and the created Identity Assertion as long as the transaction has not ended. All referring information SHOULD be archived immediately when the transaction has ended, unless legal order obliges Identity Service Provider Agent A not to archive the data.
	GINIbay MUST store the identity information of User Agent A as long as the transaction has not ended. The identity information SHOULD be archived immediately when the transaction has ended, unless legal order obliges GINIbay not to archive the data.

	GINIbay MUST store the provided request information of User Agent A as long as the product has not been dispatched to the buyer. The information SHOULD be archived immediately when the product has been dispatched, unless legal order obliges GINIbay not to archive the data.
Job-related Attestations	The Employers Agent P SHOULD delete the requested information, if the requestor does not apply for the position within three month. If the requestor applies with the requested attributes, the request information MUST be archived when Roberto signs the job contract.
	The Discovery Service Agent MUST store the provided request for data of the User Agent as long as the transaction has not ended. The transaction ends when the User Agent receives the location information. The request data SHOULD be archived immediately when the transaction has ended, unless the legal framework obliges the Discovery Service Agent not to archive the data.
	The University Agent, Employer Agent A and B MUST store the request information as long as the contract has not been signed or Roberto has not been rejected. If the contract is signed or Roberto is rejected the request SHOULD be archived immediately.
	The university and Employer Agents MUST store the provided authentication information of the User Agent as long as the authentication has not been validated. The information SHOULD be archived immediately when the authentication has been validated, unless legal order obliges not to archive the data.
	The University Agent, Employer Agent A and B MUST store the certified copies of records, employment references and the data needed for creation as long as the certified copies have not been returned to Roberto, which marks the end of the transaction. These SHOULD be archived immediately when the transaction has ended, unless legal order obliges not to archive the data.
	In case that Roberto becomes an employee of the prospective Finnish employer, the employer MUST store the provided data as long as Roberto is an employee of the company. The permission to archive the data is granted after Roberto left the company if the national legal framework obliges the company to store and archive it.
	In case that the prospective Finnish employer rejects Roberto after the interview, the employer SHOULD delete the provided data as soon as Roberto receives the rejection, unless legal order obliges the employer to store the data.
Online Petition	The PTPD Agent MUST store the policy look-up request only as long as the transaction has not ended.
	The Identity Service SHOULD store the authentication information only during the authentication process. After credential verification, the presented credentials by the User Agent MUST be deleted immediately.
	The Attribute Service MUST store the attribute request of the User Agent as long as the transaction has not ended. All request information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data.
	The Pseudonymization Service MUST store the pseudonymization request of the User Agent and the identity data as long as the transaction has not ended. All request information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Pseudonymization Service to store the data.
	The city MUST store the signing request of the User Agent as long as the signing period has not ended. All request and user information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data. Otherwise, if data wants to be stored for a longer period this

	MUST be stated in the INDI Policy.
Renewal of Authoritative Documents	The PTPD Agent MUST store the policy look-up request only as long as the transaction has not ended.
	The Identity Service SHOULD store the authentication information only during the authentication process. After credential verification, the presented credentials by the User Agent MUST be deleted immediately.
	The Attribute Service MUST store the attribute request of the User Agent as long as the transaction has not ended. All request information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data.
	The Cross-Realm Service MUST store the marriage certificate as long as the transaction has not ended. All information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data.
	The public authority MUST store the document renewal request of the User Agent as long as the renewal process has not ended. All request and user information SHOULD be deleted immediately when the transaction has ended, unless legal order obliges the Attribute Service to store the data. Otherwise, if data wants to be stored for a longer period this MUST be stated in the INDI Policy.

Table 3: Storage related Use Case Privacy Policies.

All privacy policies related to data storage are listed in Table 3. Looking at these storage privacy policies of the use cases, we can see a big overlap. They are mostly constructed in similar ways. The responsible Agent or party must store the data as long as the respective process has not ended. If the process has ended, the data must be deleted or archived. The data must not be deleted, but archived, if the legal framework mandates the corresponding party to do so or if the data is needed for billing purposes (e.g. GINIbay.com charges the seller monthly, if she sold something on the platform. To document the activities, the data must be archived for at least one month). Mandated storage of data can occur when the user is under (uninformed) investigation. For example the European Retention Directive 2006/24/EG [EU06] explicitly allows the provider the storage of communication details for tracing purpose in case of illegal activities or suspicion.

- The responsible party MUST store the provided data as long as they are required for the transaction and the transaction has not ended. This collected data MUST be deleted immediately when the transaction has ended, unless the user gives the consent for further storage and/or the legal framework obliges the Relying Party to store the data.

- The responsible party SHALL archive the provided data instead of deleting if the user gives the consent for archiving and/or the legal framework mandates the Relying Party to archive and not to delete the data.

5.4.3 Access

Use Case	Access
Person-to-Person Transaction	GINIbay MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	Discovery Service Agent B MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	Identity Service Provider Agent B MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	Attribute Service Provider B MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	The Identity Service Provider Agent A MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	GINIbay MUST restrict access to the archived data to the responsible person and log the access to the archived data.
Job-related Attestation	The Employers Agent P MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	The Discovery Service Agent MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	The University Agent, Employer Agent A and B MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	The university and Employer Agents MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	The University Agent, Employer Agent A and B MUST restrict access to the archived data to the responsible person and log the access to the archived data.
	The access permission to Roberto's archived data MUST be restricted to his supervisor and the responsible persons in the human resource department.
	For the case that the prospective Finnish employer rejects Roberto immediately, the employer SHOULD delete the provided data as soon as Roberto receives the rejection, unless legal order obliges the employer to store the data.

Table 4: Access related Use Case Privacy Policies.

In the use cases we set up access privacy policies for data, which is archived instead of deleted. Table 4 lists all access restricting related privacy policies from the use cases. The numbers of the privacy policies are significantly fewer, because only in the "Person-to-Person Transaction" and "Job-related Attestation" use cases the provided data archived instead of deleted.

The reason to restrict the access to the archived data is to hold the purpose binding, minimum disclosure and accountability requirements. For example, in the "Job-related Attestation" use case: Roberto sends his record certificates to the prospective employer. The purpose of the data is to enable an evaluation of Roberto's abilities and skills. This evaluation process properly involves several people (e.g. some officials from the human resource department, the manager of the department searching for new employees and some employees, who can verify some of the stated skills).

- The Relying Party MUST restrict the access to the provided data to a minimum set of people, who are predefined according to their tasks and the purpose of the data.

After evaluation and if Roberto is employed, not all people should have access to the archived records, since the purpose of the data should be fulfilled. The access to archived records must be restricted to a minimum amount of people (e.g. only the responsible official in the human resource department has access to the archived records). If the employer needs information from the records, the respective official can only access the archived records and deliver the required information. This case can occur if the employer wants to verify information (e.g. Roberto claims to know nothing about mobile app development. The employer can check with the provided and now archived records if his stated information is true). But the people with access to the archived data must be also controlled and audited. To ensure this and satisfy the accountability requirement, every access to the data must be logged.

- The Relying Party MUST restrict access to the archived data of the user to the responsible person and log the access to the archived data.

5.4.4 Consent

Use Case	Consents
Person-to-Person Transactions	The buyer MUST give her consent to the INDI Policy of GINIbay before the transaction process continues.
	GINIbay MUST obtain a new consent from the buyer for data processing if the original specified purpose of the data use changes, unless law obliges it.
	GINIbay MUST forward the INDI Policy of the respective third party to the buyer and obtain her consent, if data is forwarded. GINIbay MUST not forward any data to third parties without the buyer's consent.
	The seller MUST obtain her consents to the INDI Policy of Attribute Service Provider B, before the transaction process continues.
Job-related Attestations	Roberto MUST give his consent to the INDI Policy of the Employers Agent P before the transaction process continues
	The Employer Agent P MUST obtain a new consent from Roberto for data processing if the original specified purpose of the data use changes, unless law obliges it.
	The Employer Agent P MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent if data is forwarded. The Employer Agent P MUST not forward any data to third parties without Roberto's consent.
	The University Agent, Employer Agent A and B MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent if data is forwarded. The Employer Agent P MUST not forward any data to third parties without Roberto's consent.
	The University Agent, Employer Agent A and B MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent if data is forwarded. The Employer Agent P MUST not forward any data to third parties without Roberto's consent.
Online Petition	The user MUST give her consent to the INDI Policy of the online petition before the transaction process continues.

Renewal of Authoritative Documents	The user MUST give her consent to the INDI Policy before the transaction process continues.

Table 5: Consent related Use Case Privacy Policies.

Table 5 lists all consent related privacy policies of the four use cases. The amount of consent related policies is less than revealing attributes related or storage related privacy policies. The reason for that are the characteristics of the INDI ecosystem. In the first contact the INDI Policy is requested and in that step all information about further data processing, required attributes, information flow, etc. is provided. The user has to explicitly accept the conditions of the INDI Policy. If the user accepts the INDI Policy, the consent for further data processing according to the INDI Policy is given. It means that the user received a request for consent with the INDI Policy, since she gets all required information in the first step of the transaction. Table 5 also shows that the first two use cases have more consent related privacy policies than the last two use cases. The reason for this situation is that the first two use cases consider data forwarding to third parties and changes the data usage purpose. These two potential possibilities require additional privacy policies.

As mentioned before, the user must be informed through the INDI Policy of desired data processes. The Relying Party is not allowed to proceed until the user has accepted its INDI Policy. In other word, without consent data processing is forbidden.

- The Relying Party MUST obtain consent through the acceptance of the INDI Policy from the user for data processing before the transaction process continues.

Like discussed in the first two use cases there might be cases where the specified purpose of the data is not enough and/or it changes because the company wants to perform a data analysis. In such case, the Relying Party has to obtain new consent from the user.

- The Relying Party MUST obtain a new consent from the user for data processing if the original specified purposes of the data usage changes, unless law obliges it.

In case of data forward to third parties, the INDI Policy of the third party must be sent to the user. In that case, the user has to give new consent. It means that the party, which wants to forward the provided data to a third party, has to forward the INDI Policy of the third party to the user and obtain a new consent. The responsibility is on the side of the forwarding party since it requests the forward. If the user accepts the INDI Policy of the third party, consent can be seen as given.

- The Relying Party MUST obtain a new consent from the user for data processing if data is forwarded to a third party. The consent is given if the user accepts the INDI Policy of the third party.

5.4.5 Inform

Use Case	Inform
Person-to-Person Transaction	The INDI Policy MUST justify the request for each attribute and specify the purpose in a way that the user can understand it.
	The Attribute Service Provider B SHALL only state an INDI Policy for access purpose, which requests for an Authentication Assertion. The justification of data request and specification of purpose must be articulated to the user in an understandable way.
	The INDI Policy of the seller MUST contain information about the information flow to third parties, it has to declare which information will be transferred to which third party and for what reason.
	GINIbay MUST forward the INDI Policy of the respective third party to the buyer and obtain her consent if data is forwarded. GINIbay MUST not forward any data to third parties without the buyer's consent.
Job-related Attestation	The User Agent SHALL only send a request for an INDI Policy in the first contact to the prospective employer and no other information in any case.
	The Employer Agent P MUST inform Roberto about information flow to third parties and it must contain: • Which information is affected, • To which third party does the information flow (name and location of the third party), • What are the reasons for the transfer, • And how is the privacy of the user affected.
	The Employer Agent P MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent if data is forwarded. The Employer Agent P MUST not forward any data to third parties without Roberto's consent.
	The University Agent, Employer Agent A and B MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent if data is forwarded. The Employer Agent P MUST not forward any data to third parties without Roberto's consent.
	The University Agent, Employer Agent A and B MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent if data is forwarded. The Employer Agent P MUST not forward any data to third parties without Roberto's consent.
	The University Agent, Employer Agent A and B MUST inform Roberto about information flows to third parties and it must contain: • Which information is affected, • To which third party does it flow (name and location of the third party), • What are the reasons for the transfer,

	• And how is the privacy of the user affected.
	The Employer Agent P MUST inform Roberto about information flows to third parties and it must contain: • Which information is affected, • To which third party does it flow (name and location of the third party), • What are the reasons for the transfer, • And how is the privacy of the user affected.
	The University Agent, Employer Agent A and B MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent if data is forwarded. The Employer Agent P MUST not forward any data to third parties without Roberto's consent.
Online Petition	The PTPD Agent MUST inform the user which attributes are required for fulfilling the INDI Policy and which actions must be taken.
	The INDI Policy MUST contain information how the data will be processed and if third parties are involved.
Renewal of Authoritative Documents	The PTPD Agent MUST inform the user which attributes are required for fulfilling the INDI Policy and which actions must be taken.

Table 6: Inform related Use Case Privacy Policies.

The user-centric approach of the GINI model and INDI ecosystem gives the user the full control of her data. One central point of the user-centric approach is to inform the user about the collected data and why it is absolutely necessary to request and collect it. Table 6 shows all inform related privacy policies from all four analysed use cases.

The first step of every transaction between two parties starts with the request for the INDI Policy. The INDI Policy must inform the user which attributes are requested and for what reason. With this information the user can decide if she is willing to provide the requested attributes and/or information. Another point is the way how the information is articulated to the user. This means that the information should be formulated in natural language with grammar on a level, which can be easily understood by people without technical knowledge.

- The INDI Policy MUST inform the user about the required attributes and justify the request for each attribute by specifying the purpose in a user-understandable form.

The upper generic privacy policy is important to make the user stay informed about her data, especially about local processing. In the use cases we mentioned some circumstances where data might be forwarded to a third party. This situation can be applied to all companies concentrating on their core competency and have outsourced other departments. In such cases, the Relying Party must contact the user to inform her about the information flow to third parties. This in-

cludes which information is affected, to which third party does the information flow (especially the location of the third party is an important point because if the third party is in a foreign country, there might be completely different legal frameworks, which can strongly influence the privacy of the user), the reason of the information flow to third parties and how the user's privacy is affected. All conditions can be presented in a new INDI Policy, since the INDI Policy must be forwarded to the user on which the user can decide to continue or stop the transaction. The Relying Party can only forward the data if the user gives her explicit consent to do so. The following two generic privacy policies synthesize the requirements above into privacy policies.

- The Relying Party MUST forward the INDI Policy of the respective third party to the user and obtain her consent if data is forwarded. The Relying Party MUST not forward any data to third parties without user's consent.
- The Relying Party MUST inform the user about information flows to third parties and this information must contain:
  - Which information is affected,
  - To which third party does it flow (name and location of the third party),
  - What are the reasons for the transfer,
  - And how is the privacy of the user affected.

5.4.6 Information Flow

Use Case	Information Flow
Person-to-Person Transaction	The INDI Policy of the seller MUST contain information about information flow to third parties, it has to declare which information will be transferred to which third party and for what reason.
	GINIbay SHOULD only forward the request of the attributes and nothing more.
	The Discovery Service Agent B SHOULD not forward any information to any third parties in any case.
	GINIbay SHOULD not forward any details to the Attribute Service Provider B as far as it has all required credentials for access.
	The Identity Service Provider Agent B SHOULD not forward any information to any third parties in any case.
	The Attribute Service Provider Agent B SHOULD not forward any information to any third party in any case.
	The Identity Service Provider Agent A SHOULD not forward any information to any third party in any case.
	GINIbay SHOULD not forward any information to any third parties in any case.
	The GINIbay SHOULD not forward any information to any third party in any case.

Job-related Attestation	The User Agent SHOULD only forward the request of the attributes and nothing more.
	The Discovery Service Agent SHOULD not forward any information to any third party in any case.
	The University Agent, Employer Agent A and B MUST inform Roberto about information flows to third parties and it must contain: • Which information is affected, • To which third party does it flow (name and location of the third party), • What are the reasons for the transfer, • And how is the privacy of the user affected.
	The University Agent, Employer Agent A and B MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent if data is forwarded. The Employer Agent P MUST not forward any data to third parties without Roberto's consent.
	The User Agent SHALL only send authentication information (e.g. username and password) to the university and Employer Agents but no other information in any case.
	The University Agent, Employer Agent A and B SHOULD not forward any information to any third party in any case.
	The User Agent SHALL only send a job request containing the certified copies of records, employment references, cover letter and curriculum vitae. Any other information is not permitted in the job request.
	If the applicant evaluation process is performed internally, the Finnish prospective employer SHOULD not forward any information to any third party in any case. If the human resource management is outsourced to a third party, then forwarding the provided information is allowed but the third party SHOULD not forward this information to other third parties in any case.
	The Employer Agent P MUST inform Roberto about information flows to third parties and this information must contain: • Which information is affected, • To which third party does it flow (name and location of the third party), • What are the reasons for the transfer, • And how is the privacy of the user affected.
	The University Agent, Employer Agent A and B MUST forward the INDI Policy of the respective third party to Roberto and obtain his consent if data is forwarded. The Employer Agent P MUST not forward any data to third parties without Roberto's consent.
Online Petition	The INDI Policy MUST contain information how the data will be processed and if third parties are involved.
	The User Agent SHOULD only forward the INDI Policy to the PTPD Agent and

	no other information.
	The PTPD Agent SHALL not forward the received information to any other third party.
	The User Agent SHALL only send authentication information (e.g. username and password) and no other information to the Identity Service.
	The Attribute Service SHOULD not forward any identity or authentication information to any other service or third party.
	The Pseudonymization Service SHOULD not forward the user's identity information to any other service or third party.
	The city SHOULD not forward the user's information to any other service or third party if not explicitly stated in the INDI Policy.
Renewal of Authoritative Documents	The PTPD Agent SHALL not forward the received information to any other third party.
	The Attribute Service SHOULD not forward any identity or authentication information to any other service or third party.
	The Cross-Realm Service SHOULD not forward the user's marriage certificate or any included information to any other service or third party.
	The public authority SHOULD not forward any user information to any other service or third party if not explicitly stated in the INDI Policy.

Table 7: Information Flow related Use Case Privacy Policies.

This sub-section includes many privacy policies related to "inform", which are analysed in Section 5.4.5. But these privacy policies also target the information flow. This sub-section analyse them according to the aspect of information flow.

After a party has received the requested data, it will be processed according to the defined purpose of the data. There are two possible cases to process data: locally or remotely.

In the former case, all data processing actions will be performed within the company's (digital) boundaries. For smaller companies with place of business in only one country, data can flow within these boundaries without re-informing the user since the company is subject to the same legal framework and we assume that all departments have the same privacy and security policies. In large companies with many places of business in different countries, the situation is much more complicated. Although the invisible digital company boundaries still exist with higher range, the companies are subject to different legal frameworks, which may put entirely various (privacy) requirements on them, especially if the company operates globally and not only in the European Union. Those different legal frameworks also might be incompatible, which means that some privacy policies are not allowed to be set in some countries. Since the analysis of different privacy legal frameworks is not the focus of this deliverable, we assume that worldwide operating companies, which underlie different (privacy) legal frameworks, use the most restrictive framework for all worldwide departments. This means that the user receives the highest protection level regarding to their privacy. With this assumption, larger and worldwide operating companies need not to re-inform the user if the provided data stay within the company boundaries.

A different situation is in the latter case where the provided data is processed remotely. As discussed many times in this deliverable, such situation is often given when the company concentrates on the core business respective competence and outsources the rest. In that case, the commissioned third party might have different internal privacy policies and/or underlies other legal frameworks. Depending on the power and/or contract, the outsourcing party might be able to force the third party to adopt other and more restrictive privacy policies. But not every com-

pany has the power to force a third party to adopt their own privacy policies. And it is also not the interest of the third parties since they would have to adopt many sets of privacy policies from all companies outsourcing and delegating their jobs to them. Considering these circumstances, the responsibilities for the data must be shifted to the outsourcing party and the user. As mentioned in the generic privacy policies about informing the user, the Relying Party must provide the INDI Policy of the third party and different explanations to the user. Based on this information, the user can decide if the process should continue. So these generic privacy policies shift some part of the responsibility to the user. The party, which wants to forward information to a third party, must verify if information forward is really needed. If it is unavoidable (but it does not mean that a company should change its business model or philosophy), then the Relying Party must choose a third party, which has a similar privacy understanding and policies. It should not have any privacy policies, which are completely in contradiction with the privacy policy of the delegating party.

- The Relying Party MUST verify if information forward is unavoidable and if it is, then it must choose a third party, which does not have contradictory privacy policies but similar understanding of privacy.

Legally, the user has a contract with the Relying Party and not with the third party. From this perspective the Relying Party has the responsibility to enforce the rights of the user at the third party according to the INDI Policy provided by the third party. For example the Relying Party has to verify if the third party deletes the provided data after the transaction has ended.

- The Relying Party MUST verify if the third party acts correctly according to its privacy policy.

5.5 Mapping Generic Privacy Policies to the Requirements

The last step of creating the generic Privacy Policy Framework is to map the formulated generic privacy policies to the eight privacy requirements, which are listed in Section 2.2. The goal of this section is to show that the generic Privacy Policy Framework does not only cover use cases but also fits to general privacy requirements. This shows that this Privacy Policy Framework, as it is now, can be applied on use cases, which have the same or comparable privacy requirements. But it also means that modification of the framework might be necessary if a use case or situation has different privacy requirements.

The following Tables (Table 8 to Table 15) show, which privacy policy fulfils which requirement. The mapping of the privacy policies to the requirements is not one-by-one, instead most of the privacy policies fulfil several requirements. As already mentioned before, this framework should be considered as a starting point or development basis. Additional privacy policies should be added if it is required for a use case. With additional privacy policies, the mapping might be changed. So this mapping should also be seen as a basis, which can change and should be modified in further development.

Transparency
The Relying Party SHALL only request for attributes, which are directly relevant and required for the current transaction and no other information in any case.
The INDI Policy MUST inform the user about the required attributes and justify the request for each attribute by specifying the purpose in a user-understandable form.
The Relying Party MUST forward the INDI Policy of the respective third party to the user and obtain her consent if data is forwarded. The Relying Party MUST not forward any data to third parties without user's consent.
The Relying Party MUST inform the user about information flows to third parties and this information must contain: • Which information is affected, • To which third party does it flow (name and location of the third party), • What are the reasons for the transfer, • And how is the privacy of the user affected.
The Relying Party MUST verify if information forward is unavoidable and if it is, then it must choose a third party, which does not have contradictory privacy policies but similar understanding of privacy.
The Relying Party MUST verify if the third party acts correctly according to its privacy policy.

Table 8: Generic Privacy Policies for Transparency

User Control
The Relying Party MUST obtain consent through the acceptance of the INDI Policy from the user for data processing before the transaction process continues.
The Relying Party MUST obtain a new consent from the user for data processing if the original specified purposes of the data usage changes, unless law obliges it.
The Relying Party MUST obtain a new consent from the user for data processing if data is forwarded to a third party. The consent is given if the user accepts the INDI Policy of the third party.
The responsible party SHALL archive the provided data instead of deleting, if the user gives the consent for archiving and/or the legal framework mandates the Relying Party to archive and not to delete the data.
The INDI Policy MUST inform the user about the required attributes and justify the request for each attribute by specifying the purpose in a user-understandable form.
The Relying Party MUST forward the INDI Policy of the respective third party to the user and obtain her consent if data is forwarded. The Relying Party MUST not forward any data to third parties without user's consent.
The Relying Party MUST inform the user about information flows to third parties and this information must contain: • Which information is affected, • To which third party does it flow (name and location of the third party), • What are the reasons for the transfer, • And how is the privacy of the user affected.

Table 9: Generic Privacy Policies for User Control.

Minimum Disclosure
The Relying Party SHALL only request for attributes, which are directly relevant and required for the current transaction and no other information in any case.
The User Agent SHALL only send the minimum set of attributes, which were requested by the Relying Party and no other attributes/information.
The Identity Service Provider SHOULD only provide an Identity Assertion to the User Agent, which fulfils the requirements of the INDI Policy of the respective Relying Party and no other information.
The Attribute Service Provider SHOULD only provide an Attribute Assertion to the User Agent with the minimum set of Attributes, which were requested by the Relying Party and no other attributes/information.
The respective Agent SHALL only request for or send identity information/direct identifiable information in situations only, where this information is required and indispensable. In any other case identity information/direct identifiable information MUST not be revealed.
The respective Agent SHALL only send a request for INDI Policy of the Relying Party and no other information in any case, if this is the sole purpose of the request and the user has not decided to continue the transaction.
The User Agent SHOULD only send the authentication information, which is required to ensure the identity of the user, and no further information or attributes.
The Relying Party MUST restrict the access to the provided data to a minimum set of people, who are predefined according to their tasks and the purpose of the data.
The Relying Party MUST restrict access to the archived data of the user to the responsible person and log the access to the archived data.
The Relying Party MUST obtain a new consent from the user for data processing, if data is forwarded to a third party. The consent is given if the user accepts the INDI Policy of the third party.
The Relying Party MUST verify if information forward is unavoidable and if it is, then it must choose a third party, which does not have contradictory privacy policies but a similar understanding of privacy.
The Relying Party MUST forward the INDI Policy of the respective third party to the user and obtain her consent if data is forwarded. The Relying Party MUST not forward any data to third parties without user's consent.

Table 10: Generic Privacy Policies for Minimum Disclosure.

Contextual Separation
The Relying Party MUST obtain a new consent from the user for data processing, if the original specified purposes of the data usage changes, unless law obliges it.
The Relying Party MUST obtain a new consent from the user for data processing if data is forwarded to a third party. The consent is given if the user accepts the INDI Policy of the third party.
The INDI Policy MUST inform the user about the required attributes and justify the request for each attribute by specifying the purpose in a user-understandable form.
The Relying Party SHALL only request for attributes, which are directly relevant and required for the current transaction, and no other information in any case.
The Identity Service Provider SHOULD only provide an Identity Assertion to the User Agent, which fulfils the requirements of the INDI Policy of the respective Relying Party, and no more information.
The respective Agent SHALL only send a request for an INDI Policy of the Relying Party and no other information in any case, if this is the sole purpose of the request and the user has not decided to continue the transaction.
The User Agent SHOULD only send the authentication information, which is required to ensure the identity of the user and no further information or attributes.

Table 11: Generic Privacy Policies for Contextual Separation.

Delegation
The Relying Party MUST obtain a new consent from the user for data processing, if data is forwarded to a third party. The consent is given if the user accepts the INDI Policy of the third party.
The Relying Party MUST forward the INDI Policy of the respective third party to the user and obtain her consent if data is forwarded. The Relying Party MUST not forward any data to third parties without user's consent.
The Relying Party MUST inform the user about information flows to third parties and this information must contain: • Which information is affected, • To which third party does it flow (name and location of the third party), • What are the reasons for the transfer,

• And how is the privacy of the user affected.
The Relying Party MUST verify if information forward is unavoidable and if it is, then it must choose a third party, which does not have contradictory privacy policies but a similar understanding of privacy.
The Relying Party MUST verify if the third party acts correctly according to its privacy policy.

Table 12: Generic Privacy Policies for Delegation.

Accountability
The Relying Party MUST obtain consent through the acceptance of the INDI Policy from the user for data processing before the transaction process continues.
The Relying Party MUST obtain a new consent from the user for data processing if the original specified purposes of the data usage changes, unless law obliges it.
The Relying Party MUST obtain a new consent from the user for data processing if data is forwarded to a third party. The consent is given if the user accepts the INDI Policy of the third party.
The Relying Party MUST verify if information forward is unavoidable and if it is, then it must choose a third party, which does not have contradictory privacy policies but a similar understanding of privacy.
The Relying Party MUST verify if the third party acts correctly according to its privacy policy.
The INDI Policy MUST inform the user about the required attributes and justify the request for each attribute by specifying the purpose in a user-understandable form.
The Relying Party MUST restrict the access to the provided data to a minimum set of people, who are predefined according to their tasks and the purpose of the data.
The Relying Party MUST restrict access to the archived data of the user to the responsible person and log the access to the archived data.

Table 13: Generic Privacy Policies for Accountability.

Purpose Binding
The Relying Party SHALL only request for attributes, which are directly relevant and required for the current transaction, and no other information in any case
The Relying Party MUST obtain consent through the acceptance of the INDI Policy from the user for data processing before the transaction process continues.
The Relying Party MUST obtain a new consent from the user for data processing if the original specified purposes of the data usage changes, unless law obliges it.
The Relying Party MUST obtain a new consent from the user for data processing if data is forwarded to a third party. The consent is given if the user accepts the INDI Policy of the third party.
The INDI Policy MUST inform the user about the required attributes and justify the request for each attribute by specifying the purpose in a user-understandable form.
The User Agent SHALL only send the minimum set of attributes, which were requested by the Relying Party, and no other attributes/information.
The respective Agent SHALL only request for or send identity information/direct identifiable information in situations only, where this information is required and indispensable. In any other case identity information/direct identifiable information MUST not be revealed.
The responsible party MUST store the provided data as long as they are required for the transaction and the transaction has not ended. This collected data MUST be deleted immediately when the transaction has ended, unless the user gives the consent for further storage and/or the legal framework obliges the Relying Party to store the data.
The Relying Party MUST restrict the access to the provided data to a minimum set of people, who are predefined according to their tasks and the purpose of the data.

Table 14: Generic Privacy Policies for Purpose Binding.

Proportionality
The Relying Party SHALL only request for attributes, which are directly relevant and required for the current transaction, and no other information in any case.
The User Agent SHALL only send the minimum set of attributes, which were requested by the Relying Party, and no other attributes/information.
The Identity Service Provider SHOULD only provide an Identity Assertion to the User Agent, which fulfils the requirements of the INDI Policy of the respective Relying Party, and no more information.
The Attribute Service Provider SHOULD only provide an Attribute Assertion to the User Agent with the minimum set of Attributes, which were requested by the Relying Party, and other attributes/information.
The respective Agent SHALL only request for or send identity information/direct identifiable information in situations only, where this information is required and indispensable. In any other case identity information/direct identifiable information MUST not be revealed.
The respective Agent SHALL only send a request for an INDI Policy of the Relying Party and no other information in any case, if this is the sole purpose of the request and the user has not decided to continue the transaction.
The User Agent SHOULD only send the authentication information, which is required to ensure the identity of the user, and no further information or attributes.
The Relying Party MUST restrict access to the archived data of the user to the responsible person and log the access to the archived data
The INDI Policy MUST inform the user about the required attributes and justify the request for each attribute by specifying the purpose in a user-understandable form.

Table 15: Generic Privacy Policies for Proportionality.

Appendix A: Enabling Technologies

A.1.1 Anonymous Communication Systems

People increasingly use the Internet for a wider range of activities that are also relevant for applications in the INDI environment: reading the newspaper, shopping, staying in contact with family and friends, finding a partner, booking holidays, expressing their opinion, keeping an on-line diary, etc. While performing an online activity, even if the confidentiality of the information being transmitted is protected through encryption, the source and destination of the communication are easily traceable. The information on who communicates with whom may reveal critical information that could be used against the Internet user. For example, someone accessing a web site with information on a life-threatening disease may be excluded by health insurance companies, or lose his job, if her employer observes the interaction.

The analysis of all traffic information generated by an Internet user (e.g., through the IP address, national ID number or social security number) allows for sophisticated profiling of each user. Some of the data that could be gathered and stored directly or indirectly, just by monitoring the users communication are: email address, age, gender, location, religious preferences, sexual orientation, bank, job, type of products bought on the Internet, period of holidays, political orientation, lifestyle, or social network. In the current communication infrastructure, traffic data is available at moderate cost to anyone willing to harvest it, without the data subject being aware of it. There is already an emerging market of personal data that criminals use to impersonate their victims.

Anonymous communication networks serve as tools for the protection of privacy in electronic applications, and they are a key component of PETs. They protect the privacy of Internet users towards the other end of the communication and towards observers in the network. This is achieved by hiding the link between the initiator of the communication and the responder. If the communication layer is not anonymized, then privacy-enhancing techniques applied at the application layer can be rendered ineffective by observations at the communication layer, as users would be identified by their IP addresses. This is true for applications such as electronic voting and electronic payments.

Anonymous communication systems are systems that provide probabilistic unlinkability between inputs and outputs (this definition includes, for example, anonymous remailers [MCPS03, DDM03a]). Anonymous communication infrastructure is a more restrictive label that applies to systems that provide an application independent, real-time, bidirectional anonymous communication layer. An anonymous communications infrastructure or system to be implemented in the INDI environment needs to fulfil a number of requirements as described in the next section.

A.1.2 Requirements towards anonymous communications systems

Enabling anonymous communications is not a trivial task and requires constant developments in technology and a rigorous understanding of new attacks. This holds for both research on the topic as well as implementations of anonymous communication infrastructures and systems, e.g., The Union Router (TOR)[6]. For an anonymous communication to provide certain guarantees, certain basic requirements should hold.

[6] http://www.torproject.org/

First, the system should provide anonymous access to the Internet at the Transmission Control Protocol (TCP) layer. This means that any application with connection-oriented communication is anonymized.

Second, the system should be designed in such a way that only the initiator needs to know about the infrastructure and is required to install the necessary (software) interface to use it. The responder may of course be aware that it receives (some of it's) communication through this infrastructure, but does not need to install any special software or hardware.

Third, the anonymous communication infrastructure should provide a layer that supports anonymous and privacy-enhanced applications. On the quantitative side, the effective size of the anonymity set can be measured using the methods proposed in [DSCP02, SD02]. In this respect, the anonymous communication layer should aim at the largest possible anonymity set size. On the qualitative side, the anonymity provided must be secured, that is, robust against passive anonymity attacks (traffic analysis), mix corruption and active attackers. Here we present a list of requirements related to the quality (security) of the anonymous communications system:

Unlinkability: Different connections established by the same user should be unlinkable to each other for any other entity.

Load balancing: The system should provide a good level of anonymity even in low traffic conditions. The system should also remain usable. Techniques such as generation of dummy traffic may help keeping a reasonable performance/anonymity balance.

Implementation issues: It would be desirable that the nodes are geographically distributed and run by independent institutions or individuals, in order to minimize the probability of collusion of different entities in the system and to increase the surface of attack. Care must be taken in the choice of sources of randomness, key lengths, cryptographic algorithms, crypto service providers, etc.

User experience: Nodes run and maintained by non-expert users or running on insecure platforms are likely to be more vulnerable to attacks. This aspect should also be taken into account as a potential security issue. On the other hand, the system should have as many users as possible, that is, it should be usable (and provide an acceptable level of security) for average users. In INDI, the usage of mix cascades should be considered, where each mix is operated by an infrastructural entity that can guarantee a certain quality of service.

Attack model: The system should be designed to resist powerful attackers, as the profile of a realistic attacker in some scenarios (e.g., governments) has control over a large number of resources. Ideally, the attacker that the system should resist is the Global Active Attacker (GAA). This attacker controls all communication links, some of the users and some of the nodes of the network. The possibility of attacks on the availability (next point) should also be considered.

Availability of anonymity services includes three steps: first, the possibility to reach an access point, second, the correct operation of the mixes and their communication, and third, the accessibility of requested information.

Access points: The first step concerns possible blocking efforts of authoritarian states or organisations. Depending on the details of the filter method, a powerful blocker can always restrict access to anonymity services. However, it is obvious that publicly known or even central access points are more vulnerable than distributed networks.

Operation of the network: The network should be robust against availability attacks. It should be able to resists floods as well as drops in traffic levels. It should also be resistant against malicious or faulty nodes. Note that distributed systems present a larger surface of attack than centralized systems.

Exit points: A large number of exit points would make the deployment of a (passive or active) global attack difficult. The nodes may comply with the legislation of the country (e.g., they may be required to blacklist certain banned sites). The wide geographical distribution of exit points will ensure that, as long as a few countries respect freedom of expression and access to information rights, the access to information will be granted for all users. Wise exit policies may also help node operators modulate the risks in difficult contexts.

Incentives to cooperate and Usability: The quantitative dimension of the anonymity (degree of anonymity obtained) depends on two parameters: the number of users of the system and the probability distribution of the correlation of inputs/outputs. The second term depends on the statistical behaviour of the anonymous network. The limit to this anonymity will be selected as a trade-off with the acceptable delay of the network. A good level of anonymity can therefore be achieved if a large numbers of users join and use the network. The design of an anonymous communication system must provide the users with incentives to cooperate, in order to provide an acceptable level of anonymity. The performance of the communications must be acceptable for the users. This may lead to limits on the minimization of statistical correlation of incoming and outgoing streams. Finally, the software run by the users should be easy to install, configure and use.

Distributed systems: The population of users should be as large as possible; therefore the system should scale well with the number of users, and also with the number of nodes. In this sense, it seems that distributed systems scale better than centralised systems. For example, mix networks scale better (with the increase of number of nodes and users) than mix cascades.

Performance: Users are not willing to pay for anonymity with a sensitive loss of performance. The system should perform well in all traffic conditions, implementing mechanisms for load balancing. Also, the degree of anonymity should be guaranteed to stay above a certain level. This may be a difficult trade-off in extreme scenarios.

Unobservable access: In certain scenarios of extreme danger for the users of anonymous systems (where the simple fact of accessing an anonymous service may put the user in danger), it may be worth considering the use of covert channels to provide unobservable access to the anonymous service.

A.2 Anonymous Credentials

A.2.1 Anonymous credential systems and the use of pseudonyms

Since David Chaum [Cha82] first defined the concept of digital credentials and pseudonyms, a lot of thought has been invested into protecting the privacy of individual users while still providing companies with the required security to do business.

The anonymous credential system proposed by Chaum is sometimes also referred to as a pseudonym system [LRSW00]. This stems from the fact that the credentials of such a system are obtained from and shown to organizations using different pseudonyms that cannot be linked. In certain extraordinary situations trusted organizations might be authorized to link two pseudonyms or even to reveal the identity of the user. This procedure is called anonymity revocation. In some cases, this is an ethically questionable practice, as it provides a false sense of security and privacy to the users and could be abused [DG10]. In other cases, anonymity revocation can be hard-coded into the protocol, ensuring that the revocation rules are clear to all parties from the beginning of the development and use of systems. While the ethical problems are not mitigated, such revocable protocols are more difficult to abuse by the so called trusted parties.

The introduction of pseudonyms [Cha82] is a useful anonymity tool. Pseudonyms allow users to choose a different name with each organization. Generally, these pseudonyms cannot be linked without the help of the user. Nevertheless, certain statements about the relationship of a user with one organization, under a pseudonym, can be shown to another organization that knows the user only under a different pseudonym [LRSW00]. While pseudonyms allow organizations to create accounts for individual users, organizations cannot determine the real identities of their customers.

The extent of influence that law enforcement will have on credential systems is still a matter of debate. Key escrow has been discussed in the nineties as a possibility for allowing law enforcement authorities to eavesdrop on encrypted connections [AABB97]. Anonymity revocation serves a similar purpose by revoking the anonymity of communication and credentials and may face similar resistance. While the cryptographic techniques for identity revocation already exist [CL01], implementing the required trust infrastructure and getting the necessary support from all parties involved is not an easy task. However, if the only alternative to anonymity revocation and trustee based tracing is to have no anonymity at all, then a discussion on clearly stated rules for revocation with legal consequences need to be put into place.

A.2.2 Requirements towards anonymous credential systems

In anonymous credential based systems, one can generally distinguish three types of players: a certification authority (CA), a user, and a verifier. In some cases, the CA and the verifier are controlled by the same entity. The CA issues a credential to a user who fulfils certain conditions. In exchange for goods and services, the user may be required to prove to a verifier (service provider) possession of a valid credential from the CA. The user may also be required to prove a predicate on the attributes encoded in his credential. The service provider may later decide to deposit a transcript of the interaction it had with the user, to the CA.

Credential systems are very heterogeneous and may have very special requirements specific to the application at stake. Nevertheless, it can be conjectured that the following set of requirements are commonly desired in most credential systems, and can therefore be used as a basis for research and development of generic solutions:

Non-forgeability: It should not be possible to forge a credential on behalf of a CA, or to alter the attributes already encoded by the CA in a previously issued credential.

Non-transferability: There should be mechanisms in place to discourage credential holders from sharing their credentials with third parties.

Non-modifiability: Any modification to a credential showing transcript should be detectable with overwhelming probability. This property is desirable to preserve evidence and prevent framing.

Privacy with respect to the CA: The CA should not be able to link the showing transcripts of a credential to the issuing protocol instance that generated it.

Privacy with respect to the verifier: A verifier should not be able to learn any information about the attributes embedded in the credential being shown, beyond what the credential holder wilfully reveals and the a priori knowledge.

Selective disclosure: a credential holder should be able to selectively disclose any partial information or property about the identity attributes embedded in his or her credential, without necessarily revealing their exact values.

Selective depositing: It is desirable sometimes for verifiers to be able to deposit showing transcripts that reveal only partial information about the transaction that took place between the cre-

dential holder and verifier. The deposited information should be consistent with the initial showing transcript.

Suitability for smartcard implementations: Special attention should be given to efficiency when designing credential systems for smartcards, because of their limited computation and storage resources.

Revocability: In exceptional cases, as discussed above, it should be possible to revoke credentials in case they are used for abusive behaviour. In some cases, this also means the unveiling of the identity of the credential holder.

Unlinkability: It should not be possible for a verifier to link different credential showings by the same credential holder.

Anonymous credential systems continue to evolve rapidly, as can be witnessed by concrete implementations. Notable is especially the appearance of two technologies, IBM's Identity Mixer [Ide] and Microsoft's U-Prove [UPr], as well as extended work done in past EU projects. In particular, the EU-funded projects PRIME[7] and PrimeLife[8] have actually shown that the state-of-the art research prototypes of anonymous credential systems can indeed confront the privacy challenges of identity management systems.

Despite all of this, the effort of understanding anonymous credential technologies so far was rather theoretical and limited to individual research prototypes. Indeed, PRIME and PrimeLife showed that these technologies provide the desirable level of privacy protection, but so far this has been demonstrated in a very limited number of actual production environments with real users. Furthermore, there are no commonly agreed set of functions, features, formats, protocols, and metrics to gauge and compare these technologies, and it is hard to judge the pros and cons of the different technologies to understand which ones are best suited to which scenarios.

Recently, the initiation of the EU project ABC4Trust[9] came to address these problems. It produces an architectural framework for Privacy-ABC[10] technologies that allows different realizations of these technologies to coexist, be interchanged, and federated. This enables users to obtain credentials following different Privacy-ABC technologies and use them indifferently on the same hardware and software platforms, as well as service providers to adopt whatever Privacy-ABC technology best suits their needs. In particular, the ABC4Trust architecture [Kro11] has been designed to decompose future (reference) implementations of Privacy-ABC technologies into sets of modules and specify the abstract functionality of these components in such a way that they are independent from algorithms or cryptographic components used underneath. The functional decomposition foresees possible architectural extensions to additional functional modules that may be desirable and feasible using future Privacy-ABC technologies or extensions of existing ones.

As we said above, there is still a gap between the technical cryptography and protocol sides of these technologies and the reality of deploying them in production environments. ABC4Trust makes for the first time considerable progress to this direction, by deploying Privacy-ABCs in two large-scale pilots. The experiences gained by these pilots will show us for the first time how these technologies can be used in real production environments and what problems emerge in practice.

[7] www.prime-project.eu

[8] www.primelife.eu

[9] www.abc4trust.eu

[10] Privacy-ABCs (or Privacy Attribute Based Credentials) is a more accurate term for anonymous credentials suggested by ABC4Trust.

Another line of research looks at interoperability issues between conventional identity management infrastructures and anonymous credentials. The identity management paradigm that is currently hyped by the industry uses only conventional cryptographic techniques has clear basic principles, and an already large products and standards portfolio. Still, the interoperability issues between different vendors and different domains define it as a moving target. The ABC4Trust architecture takes a big step ahead in helping the integration of anonymous credentials, due to the unified format and specification of the corresponding artifacts. Deliverable D2.1 [Kro11] provides an analysis showing that the applicability of the ABC4Trust architecture to the popular existing identity protocols and frameworks such as WS-*, SAML, OpenID, OAuth and X.509 is not only possible but can also help to alleviate some of the security, privacy and scalability issues of the latter.

A.3 Electronic cash

A.3.1 Electronic cash

Electronic cash systems are, in many ways, similar to anonymous credential systems. The main difference is, in many instances, that the attributes in credentials, which function as 'electronic coins', encode the right to 'spend' the coin, i.e. to obtain a good or service of a particular value when showing the coin. An extension of electronic cash are e-tokens that one can spend up to n times [CHK+06]. Electronic coins are such tokens where n=1.

A.3.2 Requirements for privacy-preserving electronic cash

The requirements for anonymous credential systems, discussed earlier, also apply to electronic cash systems. However, certain differences exist, most notably the need for double spending protection. The differences are summarized in the following.

- **Protection against Double-Spending:** An e-cash user should not be able to spend a given electronic coin more than once. Achieving this requirement is non-trivial since, it is always easy to copy information in the digital world and use a copy instead of the original. However, systems with double-spending protection have been developed that detect double spending instances. In some cases, this automatically reveals the double spender's identity, thus providing strong incentives for users to refrain from double-spending.

- **Transferability:** While it is generally desirable for anonymous credentials to not be transferable between individuals, in some situations it may be desirable that e-cash is transferable between users. This, after all, simulates the physical world where people can give cash to each other. Enabling transferable electronic cash while not violating the other desirable properties is challenging, but proposals exist in the literature.

- **Divisibility:** At the time of e-cash withdrawal, it is typically unknown how the obtained coins will be spent. It is, therefore, desirable for an electronic coin to be divisible, i.e. to enable the user to spend only a part of its value in a given showing protocol, and retain the remainder of its value for later.

A.4 Private Information Retrieval

A.4.1 Private Information Retrieval and Oblivious Transfer

With over 15 years of development, the field of PIR contains multiple subareas and overlaps with several other fields. At its core, lies the problem of a client that needs to access a particular entry in a database that is held in a server. A PIR protocol enables the client to access the desired database record without the server learning which record it is. Of course, there exists a trivial solution to this problem, which is for the client to simply download the entire database. In this way, it certainly sees the record it wants to see, and the server has no idea about which record it actually is.
Many PIR protocols aim to improve this trivial solution in multiple respects, most notably in reducing the communication overhead. Can the same or a similar degree of privacy be achieved without the need to transfer the entire database? Another aspect that some PIR protocols support is the notion that the client may not be allowed to see more than a predetermined proportion of the database. Other protocols ensure that the client pays a certain amount before obtaining the database entry; still the server doesn't learn which entry is revealed to the client. Also computational aspects of PIR protocols must be optimized, since it is not practical if a protocol requires only small amounts of data transfer at the cost of prohibitively costly computations at either the server or the client. More information on the PIR field can be obtained from [OS07, Tre04, GAS]. Oblivious transfer protocols [CDN10, RP10] can be seen as a particular type of PIR protocols. Some of these protocols, e.g. [RP10], support the notion of protecting individual database records with access control policies. This type of protection is important in multiple scenarios, for example in the area of e-health where access to medical data is strictly regulated. Oblivious transfer protocols with support for access control enable users to obtain access to records to which they are explicitly authorized to obtain access to, but without revealing to the database server (a) they identity, (b) the very attributes that guarantee their access, and (c) which records are accessed. Yet, the database server is assured that only authorized clients obtain access.

A.4.2 Requirements for Oblivious Transfer Protocols

Requirements for oblivious transfer protocols are as follows.

- **Privacy with respect to the accessed record:** The database server should not learn which record the client accesses. Depending on the context, this could mean different things. In the most privacy-preserving case, the client may have accessed any one from all records in the database, and this privacy guarantee holds unconditionally (i.e. without the need to make cryptographic assumptions). In the least privacy-preserving case, the database server knows that the client accessed one from only two records, but doesn't know which one, and this guarantee holds only computationally, i.e. under some cryptographic assumption.

- **Privacy with respect to the client's identity:** In some contexts the client must possess certain certified attributes in order to obtain access to a database record. In such cases, it is not always necessary for the database server to learn which attributes a client possesses in order to enforce an access control policy. Protocols that ensure that attributes are not disclosed have a clear advantage over protocols that do not.

- **Security:** In the context of database records that either require an access control policy to be satisfied by clients, or that require payment, it must be guaranteed that dishonest clients cannot obtain access to more records that they are entitled to.

- **Practicality:** A PIR or oblivious transfer protocol must not lead to prohibitive communication or computation overheads either at the client or the server.

A.5 Reputation systems

A.5.1 Reputation systems

Reputation systems are used to establish trust in the online world. Reputation systems typically work as follows. Users first interact with a reputation subject. This subject is some type of entity, for example another user, an online shop, a particular product, or an online article. After interacting with this subject, the users deposit ratings to a typically centralised entity called the "Reputation Provider". These ratings contain information about the user's opinion about the subject, which she formed during or as a result of her interaction with it. The reputation provider collects everybody's ratings and, based on these ratings (and, perhaps also other parameters such as their identities and context data), computes a "reputation" for each reputation subject. This reputation value is typically meant to represent, in a compact form, the aggregate "general public" opinion that exists about the subject.
In an electronic world that is vibrant and diverse, reputation systems are a proven tool to establish trust between initially strangers. This is mostly true in marketplaces where many competitors "fight" for their market share, but also holds in less commercial environments such as online chat rooms, dating sites, online forums, blogs and wikis.
The passive or active usage of a reputation system bears privacy risks, especially in context where one's reputation becomes "sticky". That is, in some situations users switch context but keep their reputation. Consider, for example, an anonymous author who writes online articles in a variety of blogs and who has been rated based on his writings. If her resulting reputation, which is published along with her articles, is unique, then it may become trivial to determine that all articles originate from the same author.
Since a positive reputation accelerates trust establishment, a sticky reputation is desirable from the subject's point of view. However, a negative reputation that is sticky is undesirable. Nevertheless, the usefulness of the entire reputation system is undermined if it is too easy to get rid of negative reputations.

A.5.2 Requirements for privacy-preserving reputation systems

Privacy-preserving reputation systems should fulfil a larger requirement set than reputation systems that do not aim to protect privacy. We first list desirable properties for reputation systems. Note that [ENISA07] elaborates on some of these requirements further.

- **Utility:** a reputation system should produce reputation values that help users to distinguish reputation subjects based on the ratings they received in the past. That is, if a reputation subject A has received significantly more positive ratings than subject B, then it should be possible to extract this information from A and B's reputation values.

- **Protection against whitewashing:** an attacker may reset a negative reputation by simply creating a new identity and then registering with the system again. A reputation system should therefore ensure that re-joining the system implies a cost.

- **Protection against sybil attacks:** A sybil attack is very difficult to prevent, as it applies to many types of system, including reputation systems. A Sybil attacker simply creates many identities for herself in the reputation system and abuses these identities in order to boost the reputation of friends or, equivalently, harm the reputation of others. A reputation system can counteract this threat by ensuring that a certain cost exists when joining the system as well as when submitting ratings.

- **Protection against bad mouthing and ballot stuffing:** An attacker may feed superfluous ratings to a victim, for example in order to exercise revenge or in order to boost a friend's reputation. In order to prevent such situations, a reputation system should enforce strict limits on the number of ratings that can be submitted, as well as the condition under which ratings can be submitted.

- **Liveliness:** This requirement calls for a particular form of utility. Namely, a lively reputation system ensures that a subject's reputation can change depending on future ratings. Some reputation systems are not lively, for example reputation systems where only positive feedback is necessary and where a "maximum reputation" (e.g. five out of five stars) can be reached. Someone who reaches the maximum reputation cannot be downgraded again; this limits the usefulness of that subject's reputation, and undermines the utility of the entire system.

- **Protection against discrimination:** while sometimes it is desirable that the identity of raters may play a role when calculating a subject's reputation (e.g. the ratings of raters who themselves have a high reputation may weigh more heavy than the ratings of other raters), the identity of the subject should not make a difference on reputation values.

Privacy-related requirements for reputation systems are as follows. A more formal model for privacy in reputation systems is described in [SPT11].

- **Bootstrap issues:** the initial reputation value given to a newcomer may become a privacy threat if only few newcomers exist. This is because this "default" reputation value then reveals the newcomer status of participants which, in some cases, can lead to the leakage of the newcomer's identity.

- **Privacy for ratees:** Ratees may appear in the system under different pseudonyms. For example, a blogger may publish articles in different blogs under different pseudonyms. A reputation system should not enable outsiders to link these pseudonyms together by exploiting reputation values.

- **Privacy for raters:** A reputation system should not enable ratees or outsiders to deduce how a particular rater rated a particular ratee. If, for example, it is possible for an observer to see that a ratee's reputation decreased after the conclusion of a transaction with a particular rater, then the observer may deduce that the rater submitted a negative rating. This presents a privacy breach. Reputation systems that guard against this type of abuse typically update the reputation of ratees after multiple transactions have taken place.

- **Privacy for reputation queriers:** Also the privacy of users that query the reputations of ratees should be protected. Ideally, reputation values are proactively published, or are available without the need for identification.

- **Distributed vs. centralised systems:** Distributed reputation systems may or may not be more privacy-friendly than centralised ones. This depends on the replication mechanism used; if a distributed system replicated the entire rating database then privacy suffers since ratings are exposed to a much larger circle of people. If, however, participating servers only see subsets of the rating database, then privacy may be protected to a better extent because trust is more distributed.

A.6 Accountability Systems

A.6.1 Accounting for Integrity

Accounting for Integrity: A number of accountability mechanisms that cover integrity have been developed or proposed in recent years. Many provide accountability for specific applications and systems [AMI+07, YuCh07, MiSG07, HARD09]. Other systems provide accountability for general distributed systems [HaKD07, HARD10].

Early work focused on making the case for accountability, later work contributed mechanisms for specific applications like p2p content distribution and cooperative storage, while the most recent work provides integrity accountability for general distributed systems and even for arbitrary binary software images that execute inside an *accountable virtual machine* (AVM).

One element of practical accountability mechanisms is a *tamper-evident log* of a distributed execution [MaBa02, HaKD07]. Briefly, each participating node in a distributed system maintains a local log of all the messages it sends and receives, along with certain other events. The log entries are connected by a hash chain, i.e. each log entry includes a secure hash of the previous entry. Each message includes a cryptographic signature, covering the message contents and headers, and the hash of the most recent entry in the log (which describes the message being sent). In this way, whenever a node sends a message, it commits to the entire sequence of events recorded in the log up to and including the message transmission.

6 Abbreviations

CA	Certification Authority
EPAL	Enterprise Privacy Authorization Language
GINI	GLOBAL IDENTITY NETWORKING OF INDIVIDUALS
ICT	Information and Communication Technologies
INDI	INdividual Digital Identity
IP	Internet Protocol
P3P	Platform for Privacy Preferences
PET	Privacy-Enhancing Technology
PIR	Private Information Retrieval
PTPD	Policy Template and Protocol Directory
TCP	Transmission Control Protocol
TET	Transparency-Enhancing Tools
TOR	The Onion Routing
XACML	eXtensible Access Control Markup Language

7 List of Figures

8 List of Tables

9 References

[AABB97] H. Abelson, R. Anderson, S. M. Bellovin, J. Benaloh, M. Blaze, W. Diffie, J. Gilmore, P. G. Neumann, R. L. Rivest, J. I. Schiller and B. Schneier, "The risks of key recovery, key escrow, and trusted third-party encryption", in: World Wide Web Journal, vol. 2, no. 3, pp. 241–257, O'Reilly & Associates, 1997.

[AAV+11] J. Alhadeff, B. Van Alsenoy, G. Verhenneman, L. Vervenne, L. Polman, D. Pruis, K. Legal, Q. Reul, L. Schilders and K. Böhm, "Legal and Policy handbook for TAS3 implementations", TAS3 Deliverable D6.1-2, Version 1.0, Draft, November 2011.

[AMI+07] K. Argyraki, P. Maniatis, O. Irzak, S. Ashish and S. Shenker, "Loss and Delay Accountability for the Internet", in: IEEE International Conference on Network Protocols (ICNP), October 2007.

[AnLe11] T. Andersson and L. Leontaridis (Eds.), "The Individualised Digital Identity (INDI) Model: A User-centric Framework of identity management services", GINI Deliverable D1.1, 2011.

[BAL10] L. Brandimarte, A. Acquisti and G. Loewenstein, "Privacy Concerns and Information Disclosure: An Illusion of Control Hypothesis," in: Proceeding of the ninth Workshop on the Economics of Information Security (WEIS 2010), June 2010.

[Cau11] J. Caumanns (Ed.), "Technology Gaps for Longer-Term Research", GINI Deliverable D2.1, Version 1.1, Draft, 2011.

[CDN10] J. Camenisch, Maria Dubovitskaya and Gregory Neven, "Unlinkable Priced Oblivious Transfer with Rechargeable Wallets", in: Financial Cryptography and Data Security, Lecture Notes in Computer Science, vol. 6052, ISBN 978-3-642-14576-6, IFCA/Springer-Verlag, Berlin Heidelberg, pp. 66, 2010.

[Cha82] D. Chaum "Blind signatures for untraceable payments", in: Proceedings of CRYPTO'82, pp. 199-203, Plenum Press, 1982.

[CHK+06] J. Camenisch, S. Hohenberger, M. Kohlweiss, A. Lysyanskaya and M. Meyerovich, "How to win the clonewars: efficient periodic n-times anonymous authentication", in: Proceedings of the 13th ACM Conference on Computer and Communications Security, CCS 2006, Alexandria, VA, USA, October 30 – November 3, 2006.

[CL01] J. Camenisch and A. Lysyanskaya, "An efficient system for non-transferable anonymous credentials with optional anonymity revocation", in: B. Pfitzmann (Ed), "EUROCRYPT", vol. 2045 of Lecture Notes in Computer Science, pp. 93–118, Springer, 2001.

[DDM03a] G. Danezis, R. Dingledine and N. Mathewson, "Mixminion: Design of a Type III Anonymous Remailer Protocol", in: Proceedings of the 2003 IEEE Symposium on Security and Privacy, May 2003.

[DG10] G. Danezis and S. Gürses, "A critical review of 10 years of Privacy Technology," in: Proceedings of Surveillance Cultures: A Global Surveillance Society?, London, UK, April 2010.

[DSCP02] C. Diaz, S. Seys, J. Claessens and B. Preneel, "Towards measuring anonymity", in: Designing Privacy Enhancing Technologies, Proceedings of PET'02, pp. 54-68. Springer-Verlag, LNCS 2482, 2002.

[ENISA07] E. Carrara and G. Hogben (Eds), "Reputation-based Systems: a security analysis", in: ENISA Position Paper No. 2, October 2007, http://www.enisa.europa.eu/act/it/privacy-and-trust/reputation-systems/reputation-based-systems-a-security-analysis. Link functional as of Feb 2012.

[EU06] European Union, "Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC", Luxembourg, April 2006.

[EU95] European Union, "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data", Luxembourg, November 1995.

[FHK+11] S. Fischer-Hübner, C. Hoofnagle, I. Krontiris, K. Rannenberg and M. Waidner (Eds.), "Online Privacy: Towards Informational Self-Determination on the Internet (Dagstuhl Perspectives Workshop 11061)", in: Dagstuhl Manifestos, vol. 1, no. 1001, pp. 1-20, 2011, http://drops.dagstuhl.de/opus/volltexte/2011/3205. Link function as of Feb 2012.

[GAS] W. Gasarch, "Web page on private information retrieval", http://www.cs.umd.edu/~gasarch/pir/pir.html. Link functional as of Feb 2012.

[GJO+05] S. Gürses, J. H. Jahnke, C. Obry A. Onabajo, T. Santen and M. Price, "Eliciting confidentiality requirements in practice", in: Proceedings of the 2005 Conference of the Centre for Advanced Studies on Collaborative research, pp. 101-116, 2005.

[Gol05] J.A. Golbeck, "Computing and applying trust in web-based social networks", College Park, MD, USA, 2005.

[HaKD07] A. Haeberlen, P. Kuznetsov and P. Druschel, "PeerReview: Practical Accountability for Distributed Systems", in: ACM Symposium on Operating Systems Principles (SOSP), 2007.

[HARD09] A. Haeberlen, I. Avramopoulos, J. Rexford and P. Druschel, "NetReview: Detecting when interdomain routing goes wrong", in: 6th Symposium on Networked Systems Design and Implementation (NSDI), 2009.

[HARD10] A. Haeberlen, P. Aditya, R. Rodrigues and P. Druschel, "Accountable Virtual Machines", in: USENIX Symposium on Operating Systems Design and Implementation (OSDI), 2010.

[Hil09] M. Hildebrandt, "Behavioural Biometric Profiling and Transparency Enhancing Tools", FIDIS Deliverable D7.12, March 2009.

[HLM+06] B. Haley, C. Laney, D. Moffett and B. Nuseibeh, "Using trust assumptions with security requirements", in: Requirements Engineering 11, 2, pp. 138-151, February 2006.

[ide] The Identity Mixer. http://www.zurich.ibm.com/security/idemix/. Link functional as of Feb 2007.

[ISO99] ISO/IEC 15408-1:1999, "Information technology - Evaluation criteria for IT security, Part 1: Introduction and general model", First Edition, 1999.

[KKB+09] J. Karat, C.-M. Karat, E. Bertino, N. Li, Q. Ni, C. Brodie, J. Lobo, S. B. Calo, L. F. Cranor, P. Kumaraguru, and R. W. Reeder, Policy framework for security and privacy management. IBM J. Res. Dev. 53, pp. 242-255, March 2009,.

[Kro11] I. Krontiris (Ed.), "D2.1 Architecture for Attribute-based Credential Technologies - Version 1", ABC4Trust Deliverable D2.1, 2011.

[KW08] B. Krishnamurthy and C. E. Wills, "Characterizing privacy in online social networks," in: Proceedings of the first workshop on Online social networks (WOSP '08), Seattle, WA, USA, pp. 37–42, 2008.

[LeSH08] R. Leenes, J. Schallaböck and M. Hansen, "PRIME White Paper", third and final version, 15. May 2008.

[LRSW00] A. Lysyanskaya, R. L. Rivest, A. Sahai, and S. Wolf, "Pseudonym systems", in: Howard M. Heys and Carlisle M. Adams (Eds), "Selected Areas in Cryptography", vol. 1758 of Lecture Notes in Computer Science, pp. 184-199, Springer, 2000.

[Luh79] N. Luhmann, "Trust and Power", John Wiley & Sons, 1979.

[LYM03] L. Liu, E. Yu and J. Mylopoulos, "Security and Privacy Requirements Analysis within a So-cial Setting", in: Proceedings of the 11th IEEE International Conference on Requirements Engineering, p. 151, September 2003

[MaBa02] P. Maniatis and M. Baker, "Secure History Preservation Through Timeline Entanglement", in: USENIX Security Symposium, 2002.

[MCPS03] U. Möller, L. Cottrell, P. Palfrader and L. Sassaman, "Mixmaster Protocol", Version 2. Draft, July 2003

[MG09] G. Müller and M. Gilliot (Eds.), "Privacy in Business Processes", FIDIS Deliverable D14.8, Version 0.9, Final, May 2009.

[MHM07] N. Mayer, P. Heymans and R. Matulevičius, "Design of a Modelling Language for Information System Security Risk Management", in: Proceedings of the 1st International Conference on Research Challenges in Information Science (RCIS 2007).

[MiSG07] N. Michalakis, R. Soulè and R. Grimm, "Ensuring Content Integrity for Untrusted Peer-to-Peer Content Distribution Networks", USENIX Symposium on Networked Systems Design and Implementation (NSDI), 2007.

[NS08] A. Narayanan and V. Shmatikov, "Robust De-anonymization of Large Sparse Datasets," in: IEEE Symposium on Security and Privacy, pp. 111-125, 2008.

[OS07] R. Ostrovsky and W.E. Sketch, "A survey of single database private information retrieval: Techniquesand applications", in: Proceedings of the 10th International Conference on Practice and Theory in Public-Key Cryptography LNCS 4450 (Beijing, Apr.16–20). Springer, pp. 393-411, 2007.

[RP10] A. Rial and B. Preneel, "Blind Attribute-Based Encryption and Oblivious Transfer with Fine-Grained Access Control", WISSec 2010.

[RPM99] K. Rannenberg, A. Pfitzmann and G. Müller, „IT security and multilateral security", in: G. Müller and K. Rannenberg (Eds), "Multilateral Security in Communications - Technology, Infrastructure, Economy", pp. 21-29, Addison-Wesley, 1999.

[RRD09] K. Rannenberg, D. Royer and A. Deuker (Eds.), "The Future of Identity in the Information Society – challenges and Opportunities", Springer, 2009, p. 507.

[RZ01] P. Resnick and R. Zeckhauser, "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System", 2001.

[SCV08] G.B. Shelly, T.J. Cashman and M.E. Vermaat, "Discovering Computers", Introductory, Course Technology, Boston, Mass, 2008.

[SD02] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity", in: Designing Privacy Enhancing Technologies, Proceedings of PET'02, pp. 41–53. Springer-Verlag, LNCS 2482, 2002.

[SPT11] S. Schiffner, A. Pashalidis and E. Tischhauser, "On the limits of privacy in reputation systems", in: Proceedings of the 11th ACM workshop on Privacy in the electronic society (WPES 2011), ACM, 11 pages, 2011.

[Swe00] L. Sweeney, "Uniqueness of Simple Demographics in the U.S. Population", LIDAP-WP4 Carnegie Mellon University, Laboratory for International Data Privacy, Pittsburgh, 2000.

[THM+08] J. Turow, C. J. Hoofnagle, D. K. Mulligan, N. Good and J. Grossklags, "The Federal Trade Commission and Consumer Privacy in the Coming Decade", in: A Journal of Law & Policy for the Information Society, 723, 2007-08.

[Tre04] L. Trevisan, "Some applications of coding theory in computational complexity", in: Quaderni di Matematica 13, pp. 347–424, 2004.

[UPr] The U-Prove SDK. http://www.credentica.com/uprove_sdk.html. Link functional as of Feb 2007.

[Van11] B. Van Alsenoy (Ed.), "Legal Provisions for Deploying INDI Services", GINI Deliverable D3.1, Version 1.0, 2011.

[Van12] B. Van Alsenoy (Ed.), "A Regulatory Framework for INDI Operators", GINI Deliverable D3.2, Version 1.0, internal draft, 2012.

[Wes70] A. Westin, "Privacy and freedom", New York: Atheneum, 1970.

[WH10] J. E. Wästlund and S. Fischer-Hübner (Eds.), "End User Transparency Tools: UI Prototypes", PrimeLife Deliverable D4.2.2, June 2010.

[YuCh07] A.R. Yumerefendi and J.S. Chase, "Strong Accountability for Network Storage", in: USENIX Conference of File and Storage Technologies (FAST), 2007.

[ZJ97] P. Zave and M. Jackson, "Four dark corners of requirements engineering", in: ACM Transactions on Software Engineering and Methodology (TOSEM), vol. 6 no. 1, pp. 1-30, Jan. 1997.